



Les emplois adnominaux des démonstratifs tchèque et français dans la langue orale informelle : étude de sémantique référentielle contrastive

Jan Dvorak

► To cite this version:

Jan Dvorak. Les emplois adnominaux des démonstratifs tchèque et français dans la langue orale informelle : étude de sémantique référentielle contrastive. Linguistique. Université de Lyon, 2021. Français. NNT : 2021LYSEN040 . tel-03461656

HAL Id: tel-03461656

<https://theses.hal.science/tel-03461656>

Submitted on 1 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Numéro National de Thèse : 2021LYSEN040

THESE de DOCTORAT DE L'UNIVERSITE DE LYON

opérée par

l'Ecole Normale Supérieure de Lyon

Ecole Doctorale N° 484

Lettres, Langues, Linguistique et Arts

Discipline : Sciences du langage. Linguistique

Soutenue publiquement le 24/09/2021, par :

Jan DVOŘÁK

Les emplois adnominaux des démonstratifs tchèque et français dans la langue orale informelle

Étude de sémantique référentielle contrastive

Devant le jury composé de :

CARLIER, Anne, Professeur des universités, Sorbonne Université, Rapporteur

PEŠEK, Ondřej, Professeur, Université de Bohême du Sud, Rapporteur

CREISSELS, Denis, Professeur émérite, Université Lumière-Lyon II, Examineur

STOSIC, Dejan, Maître de conférences – HDR, Université Toulouse Jean Jaurès, Examineur

GUILLOT-BARBANCE, Céline, Maître de conférences – HDR, École normale supérieure de Lyon,
Directrice

NÁDVORNÍKOVÁ, Olga, Professeure, Université Charles de Prague, Co-directrice

« Ce n'est point chose vicieuse, mais grandement louable, emprunter d'une langue étrangère les sentences et les mots, et les approprier à la sienne. »

Joachim Du Bellay

„V jazyku je naše národnost.“

Josef Jungmann

Remerciements

Mes plus grands remerciements s'adressent, sans conteste aucun, à **Mme Céline Guillot** pour son encadrement hors du commun tout au long du travail sur cette thèse, dès la conception et la définition du sujet de recherche jusqu'aux dernières touches apportées au texte avant son impression. Je tiens à lui manifester ici ma plus sincère reconnaissance pour toute l'aide précieuse qu'elle m'a apportée ainsi que pour la capacité d'écoute dont elle a fait preuve à de nombreuses reprises et en particulier aux moments difficiles. Je la remercie d'avoir su se montrer à la fois empathique et patiente face à mes crises de défaitisme. Si le travail sur une thèse de doctorat est, sans aucun doute, un processus de mûrissement personnel et intellectuel, c'est également un moment d'apprentissage de l'humilité. Avec l'expérience et le recul que j'espère avoir acquis au terme des quatre années de la durée de ce projet, je dois donner raison, a posteriori, à une série de conseils prévoyants de Mme Guillot.

Je tiens à remercier chaleureusement **Mme Olga Nádvorníková**, tout d'abord d'avoir accepté d'encadrer mon projet malgré la distance physique qui nous séparait et son agenda bien rempli, mais encore davantage pour toutes ses remarques, suggestions et recommandations extrêmement pertinentes et précieuses, y compris en ce qui concerne la structuration du travail et la présentation des faits. Je lui suis tout particulièrement reconnaissant d'avoir su me convaincre de l'utilité des outils statistiques dans les recherches en sciences du langage et de m'avoir guidé dans mon initiation à la linguistique de corpus. Ses nombreuses remarques au sujet du démonstratif tchèque ont alimenté ma réflexion et m'ont permis d'aller au-delà de ma zone de confort intellectuelle.

Ce projet n'aurait jamais vu le jour sans l'assistance technique de plusieurs personnes, assistance qui a pallié mon manque de compétences en matière d'outils informatiques et statistiques. Ces personnes, que je remercie humblement pour leur temps et leur patience (il en a fallu beaucoup), sont : **Alexei Lavrentev**, **Bénédicte Pincemin** et **Serge Heiden** du groupe CACTUS (groupe de recherche en textométrie) à l'ENS de Lyon et **Matthieu Quignard** du laboratoire ICAR de Lyon. Un grand merci revient également à **Michal Křen** du Corpus national tchèque et à **Clément Plancq** du projet Orféo pour l'amabilité qu'ils ont eue de me fournir, respectivement, les textes des transcriptions des conversations du corpus ORTOFON v1 et du Corpus d'Études pour le Français Contemporain. M. Křen a également eu l'amabilité de répondre à mes questions relatives aux dictionnaires des fréquences tchèques dont il est co-auteur. Je remercie aussi **Achille Falaise** pour ses informations concernant le corpus ScienQuest.

La réalisation de ce projet s'est nourrie d'échanges fructueux et de dialogues avec nombre de chercheurs confirmés. Ceux-ci ont accepté, en dépit du clivage qui nous séparait sur le plan de l'expérience et du statut institutionnel et symbolique, de me traiter en égal et ils ont eu la grande

gentillesse de répondre à mes sollicitations. Outre mes deux directrices, je tiens, à cet égard, à remercier notamment :

Adrian Czardybon, de n'avoir jamais manqué de répondre avec une extrême bienveillance à mes nombreuses questions et sollicitations au sujet de la définitude dans les langues slaves occidentales et d'avoir toujours été prêt à discuter avec moi de l'interprétation d'exemples puisés dans mes données de corpus et résistant à une analyse claire et univoque.

Georges Kleiber, d'avoir bien voulu partager avec moi son expertise inégalée dans le domaine des démonstratifs.

Sebastian Löbner, d'avoir eu l'amabilité de m'éclairer sur des aspects de sa fascinante approche, qui m'a fourni le principal outil d'analyse sémantique pour l'appréhension de la grammaticalisation de TEN en tchèque.

Josef Jodas, d'avoir bien voulu répondre à mes questions au sujet des différences dialectales du système du démonstratif tchèque.

Anne Carlier, d'avoir bien voulu me recevoir à Paris pour discuter avec moi de la grammaticalisation de TEN. Sa participation à mon jury de soutenance me réjouit particulièrement.

Radek Šimík, que je remercie à la fois de m'avoir permis de participer à un atelier consacré à la problématique de la définitude et dont il fut co-organisateur dans le cadre de la 41^{ème} conférence de la Société allemande de linguistique, tenue en mars 2019 à Brême, pour les échanges très stimulants au sujet du démonstratif tchèque auxquels il a consenti et pour sa volonté de partager avec moi une partie de ses constats et de ses perspectives de recherche.

Tilman Berger, d'avoir eu la gentillesse de partager avec moi un de ses articles, et à qui je rends hommage pour sa contribution majeure aux études sur le démonstratif tchèque et, plus généralement, pour sa contribution, via ses activités de recherche foisonnantes, à la compréhension des nombreux liens linguistiques entre le tchèque et l'allemand.

Nicolas Laurent, d'avoir accueilli et encadré mon projet de recherche de master 1, événement qui fut à l'origine de ma venue en France en 2015 et qui me permit de commencer cette belle aventure. Je le remercie, de manière générale, d'avoir fait naître une bonne partie de mon intérêt pour la linguistique.

Le présent projet n'aurait jamais pu aboutir sans l'aimable aide de mes informateurs français, tchèques, slovaques, polonais, russes et ukrainien, qui ont trouvé le temps, souvent malgré un agenda fort chargé, de remplir mes questionnaires et de me faire part de leurs très précieux retours. Ces personnes sont : **Sarah Andriamahakajy, Petra Baštanová, Lucie Berg, Geoffroy Blocx, Emma Bratková, Boumediene Briksi Sid, Jan Dvořák, Jorika Dvořáková, Hasnaa El Awad, Jacques-Éloi Genot, Patrik Felčer, Petr Felčer, Onřej Hrách, Elżbieta Kaczmarska, Marcela Kočandrlová, Tomáš Komm, Kristýna Knoppová, Kristína Málková, Boris Masar, Benoit**

Meunier, Piotr Plocharz, Václav Rameš, Kristína Ramešová, Gilles Ramponi, Thomas Scott, Eva Ulrichová, Marika Zecconi, Alla Zhuk et Ivona Zorn. Je remercie également **Waldemar Blech** pour ses retours fort éclairants sur l'allemand.

Un grand merci à **Beltamiro Patrício** pour ses encouragements, ses conseils amicaux et son aide avec le traitement formel du document, à **Boumediene Briksi Sid** pour son indispensable soutien informatique ainsi qu'à **Éric Bordas** et à **Thomas Zacher** pour leur soutien administratif.

Enfin, je prends ici le temps de remercier, du fond de mon cœur, **Zeïna**, pour son amour, son soutien moral, psychologique et intellectuel intarissable tout au long des quatre années de ce travail. Je la remercie d'avoir partagé avec moi cette passion aux deux sens du terme. Je remercie **ma famille** et tout particulièrement **ma mère** pour l'amour inconditionnel qu'ils m'ont toujours témoigné.

Thèse préparée sous la double direction de Mme Céline Guillot et de Mme Olga Nádvorníková à l'École normale supérieure de Lyon et au sein du laboratoire Institut d'Histoire des Représentations et des Idées dans les Modernités (UMR 5317)

**École normale supérieure de Lyon
15 parvis René Descartes
BP 7000
69342 Lyon Cedex 07
FRANCE**

Les emplois adnominaux des démonstratifs tchèque et français dans la langue orale informelle. Étude de sémantique référentielle contrastive

Résumé

Cette thèse a pour objectif de comparer les usages du démonstratif défini en emploi adnominal dans les langues tchèque et française, plus précisément dans leur registre parlé informel (le plus sujet au changement et à l'innovation). Pour ce faire, nous nous appuyons sur des données de corpus composées essentiellement de conversations. Ces données sont annotées et exploitées dans le logiciel textométrique TXM, développé par le groupe de recherche CACTUS à l'École normale supérieure de Lyon. Outre une comparaison statistique de la fréquence du démonstratif adnominal dans les deux langues, le travail opte, avant tout, pour une approche sémantico-référentielle articulée autour de quatre axes : (i) le fonctionnement du démonstratif adnominal des deux langues au sein des différents régimes d'indexicalité (emplois exophoriques et endophoriques), (ii) les singularités des emplois « émotionnels » du démonstratif – ceux où le démonstratif ne participe pas directement à l'identification d'un référent sémantiquement unique (Löbner, 1985 et 2011), mais véhicule un effet sémantico-pragmatique ressortissant à la subjectivité du locuteur (cf. Adamec, 1983 ; Kleiber, 1991 ; Jonasson 1994) –, (iii) la grammaticalisation en cours du démonstratif tchèque TEN, neutre eu égard à l'expression de la distance du référent et en perte de sémantique déictique (cf. Adamec 1983 ; Berger 1993 ; Mathesius 1926 ; Meyerstein 1972 ; Orlandini 1981 ; Vey 1946 ; Zíková 2017), ainsi que, dans une moindre mesure, (iv) le cumul, en tchèque, de plusieurs mots fonctionnels (Abney, 1987), dont le démonstratif, au sein d'un même complexe nominal. C'est à la question de la grammaticalisation de TEN en article défini qu'une attention toute particulière est accordée. Pour examiner cette question, nous nous en référons, avant tout, au cadre de la théorie appelée « Concept Types and Determination Theory » et développée par Löbner (1985 et 2011). L'ensemble des analyses s'inscrivent dans une approche contrastive, sous-tendue par des interrogations qui relèvent de la typologie linguistique.

Mots-clés : démonstratif adnominal défini, tchèque, français, grammaticalisation, linguistique de corpus, linguistique contrastive, sémantique référentielle, indexicalité, définitude, article défini

The Adnominal Uses of the Czech and the French Demonstrative in Informal Spoken Language. A Study in Contrastive Referential Semantics

Abstract

The objective of this dissertation is to compare the usages of the definite demonstrative in adnominal uses in Czech and in French, more particularly in their informal spoken register (the one most prone to change and innovation). In order to do so, we rely on corpus data consisting mostly of informal spoken exchanges (conversations). The data is annotated and analysed in TXM, a textometric software developed by the CACTUS research group at the École normale supérieure de Lyon. In addition to a statistical account of the frequency of the adnominal demonstrative in both languages, the dissertation opts, first and foremost, for a semantic-referential approach, centred around four axes : (i) the functioning of the adnominal demonstrative of both languages within the domains of deixis (exophoric and endophoric), (ii) the particularities of “emotional” uses – those where the demonstrative does not take part directly in the identification of a semantically unique referent (Löbner, 1985 and 2011), but produces a semantic-pragmatic effect falling within the speaker’s subjectivity (cf. Adamec, 1983; Kleiber, 1991; Jonasson 1994) –, (iii) the ongoing grammaticalization of the Czech demonstrative *TEN*, neutral as regards the expression of distance of the referent and undergoing a process of weakening of its deictic semantics (cf. Adamec 1983; Berger 1993; Mathesius 1926; Meyerstein 1972; Orlandini 1981; Vey 1946; Ziková 2017), and, to a more limited extent, (iv) the accumulation of several functional words (Abney, 1987), including the demonstrative, within the same noun complex in Czech. It is the issue of the grammaticalization of *TEN* into a definite article which is paid particular attention in this work. In addressing this issue, we majorly base our analyses upon the theoretical framework of the “Concept Types and Determination Theory”, developed by Löbner (1985 and 2011). In general, the analyses presented in this work are couched within a comparative approach and underlined by questionings related to linguistic typology.

Keywords: definite adnominal demonstrative, Czech, French, grammaticalization, corpus linguistics, contrastive linguistics, referential semantics, deixis, definiteness, definite article

Sommaire

Remerciements	5-7
Résumé	9-10
Sommaire	11
 Introduction générale	 13-37
 Partie I : Fondements théorico-méthodologiques et état de l’art	 39-264
Chapitre 1. Le démonstratif : présentation générale	41-80
Chapitre 2. Présentation des corpus et de la méthodologie du travail avec les données	81-135
Chapitre 3. Le démonstratif en tchèque	137-210
Chapitre 4. Le démonstratif en français	211-263
 Partie II : Analyses quantitatives	 265-377
Chapitre 5. Le démonstratif adnominal : bilan quantitatif	267-377
 Partie III : Analyses qualitatives	 379-554
Chapitre 6. La grammaticalisation du démonstratif TEN en article défini	381-494
Chapitre 7. Les emplois « émotionnels » du démonstratif TEN	495-554
 Conclusion générale	 555-587
 Bibliographie, index et annexes	 589-652
Bibliographie	591-617
Index des auteurs	618-620
Annexes	621-646
 Table des matières	 647-652

Introduction générale

Introduction générale

Présentation du projet

Dans cette thèse de doctorat, préparée sous la double direction de Mme Céline Guillot de l'École normale supérieure de Lyon et de Mme Olga Nádvorníková de l'Université Charles de Prague, nous nous donnons pour objectif de comparer les usages du démonstratif défini en emploi adnominal¹ dans les langues tchèque et française, plus précisément dans leur registre parlé informel. Il s'agira ainsi de réaliser une étude contrastive des deux systèmes. Pour ce faire, nous nous appuierons essentiellement sur des corpus composés de productions orales informelles. Après une comparaison générale, notre étude se focalisera sur les spécificités du démonstratif adnominal dans le tchèque parlé informel, à savoir notamment la grammaticalisation présumée de la forme TEN en article défini et les emplois dits « émotionnels ».

Problématique, hypothèses et principaux axes du projet

Notre projet de recherche part d'une fascination. La toute première idée qui en constitue l'origine remonte au moment où nous nous sommes aperçu d'une surabondance d'emplois du démonstratif adnominal qui semble opposer notre langue maternelle à la langue française. Ce phénomène est notamment caractéristique de la langue parlée informelle de tous les jours, dont on sait qu'elle est, en général, la plus ouverte aux innovations (voir à cet égard par exemple Blanche-Benveniste, 2007 ; Cvrček *et al.*, 2015 [2010] et Zíková, 2018 pour la langue parlée en général et Short, 1933 et Zíková, 2018 pour le tchèque en particulier).

C'est ainsi que, dans son étude consacrée à la grammaticalisation du démonstratif en article défini en tchèque, en haut-sorabe et en slovène, Trovesi justifie son choix d'un matériel composé exclusivement de productions orales informelles :

L'analyse de l'emploi des pronoms démonstratifs en fonction d'article défini a été menée sur un matériel de langue parlée en situations informelles, car ce phénomène fait partie des traits du registre substandard, et les locuteurs cherchent consciemment à l'éviter dans les textes écrits ou supervisés². (2002 : 131)

¹ Les paradigmes des démonstratifs indéfinis TEL et TAKOVÝ et ceux marquant l'identité (TÝŽ, TENTÝŽ (LE MEME)) ne seront pas au centre de la présente étude.

² Nous traduisons : „Analýza používání ukazovacích zájmen ve funkci určitého členu byla prováděna na materiálu mluveného jazyka v neformálních situacích, protože tento jev patří k rysům substandardním a mluvčí se mu v psaných nebo monitorovaných textech záměrně vyhýbají.“

Souček, qui se penche sur la grammaticalisation du démonstratif tchèque et finnois tout en les comparant au fonctionnement de l'article défini *the* anglais, écrit ceci :

Comme le démontrent les études de la langue parlée (cf. Cvrček *et al.*, 2010, Šonková, 2008, Hakulinen *et al.*, 2004), dans les productions orales, les pronoms démonstratifs imitent l'article défini dans une bien plus grande mesure, et leur fréquence en fonction d'article grammatical y est aussi beaucoup plus élevée, ce qui peut constituer le signe d'une tendance évolutive vers l'article défini (et, dans le cas de l'expression *jeden (un)*, aussi vers l'article indéfini) dans les deux langues comparées³.

Enfin, citons encore Cvrček *et al.*, qui, proposant la première grammaire du tchèque non normative et entièrement basée sur des données du Corpus national tchèque, écrivent ceci :

Dans le tchèque parlé, le déterminant *ten* possède une fréquence et des fonctions similaires, en position adnominale, à l'article défini *the* en anglais, même si dans le tchèque écrit, il n'est que rarement utilisé dans la détermination nominale. Les deux formes du tchèque s'influencent mutuellement, d'où probablement le fait que la détermination au moyen du pronom *ten* n'est pas figée, dans le tchèque parlé, de la même façon qu'elle l'est dans les langues à articles définis [...] ⁴ (2015 [2010] : 176)

Le pronom *ten* est souvent utilisé, dans la langue parlée, comme un déterminant des substantifs, exprimant ainsi la définitude, similairement à l'article *the* en anglais [...] ⁵. (2015 [2010] : 262)

Au vu de ce qui vient d'être dit, le choix de corpus de langue orale informelle comme principale source de données s'est imposé à nous, dans le cadre du présent travail, comme une nécessité.

Nous tenons aussitôt à préciser que notre projet ne se veut pas une étude synchronique exhaustive du démonstratif en tchèque, en français ou dans les deux langues à la fois, mais une comparaison qui permette de faire ressortir les principales similitudes et différences entre les deux systèmes. Si la linguistique contrastive franco-tchèque orientée sur la problématique du démonstratif n'en est pour l'instant qu'à ses débuts⁶, il existe, pour chaque domaine séparément, des travaux d'une grande qualité. Pour le domaine du tchèque, nous pensons notamment, à côté d'une série d'études éclairantes, quoique plutôt exploratoires et fragmentaires (cf. Adamec, 1983 ; Kresin, 2008 ; Křížková, 1971 ; Machač et Zíková, 2014 ; Mathesius, 1926 ; Meyerstein, 1972 ; Nédoluzko, 2004 ; Schneiderová, 1993 ; Šimík, 2015 ; Souček, 2011 ; Štícha, 1999 et 2001 ; Uhlířová, 1992 et 2006 ;

³ Nous traduisons : „Jak ukazují studie mluveného jazyka (srov. Cvrček *et al.* 2010, Šonková 2008, Hakulinen *et al.* 2004) ukazovací zájmena napodobují v mluvených projevech člen v mnohem větší míře, jejich frekvence v pozici funkce gramatického členu je také mnohem vyšší, což může být znakem tendence vývoje určitého členu (a v případě výrazu *jeden* i neurčitého členu) v obou porovnávaných jazycích.“

⁴ Nous traduisons : „V mluvené češtině má determinátor *ten* v pozici před substantivem podobnou frekvenci i podobné funkce jako určitý člen *the* v angličtině, v psané češtině se ale k určení substantiv užívá jen zřídka. Obě formy češtiny se ovlivňují, a patrně proto není determinace zájmenem *ten* v mluvené češtině ustálena jako v jazycích s určitým členem [...]“

⁵ Nous traduisons : „Zájmena *ten* se užívá v mluveném jazyce často jako determinátoru substantiv, vyjadřuje tedy určitost podobně jako v angličtině člen *the* [...]“

⁶ À l'heure actuelle, nous ne connaissons pas d'autres travaux que les études exploratoires suivantes de Pešek : « Nominální anafora a determinace – kontrastivní analýza francouzských a českých systémových možností » (2014) et « Progression thématique et anaphorisation. L'apport informationnel des procédés de reprise » (2012).

Trovesi, 2002 ; Voráč, 1946) et des sections consacrées au démonstratif par les grammaires d'importance – telle *Akademická gramatika spisovné češtiny* de Štícha –, à la thèse d'habilitation non-publiée de Berger *Das System der tschechischen Demonstrativpronomina. Textgrammatische und stilspezifische Gebrauchsbedingungen* (soutenue en 1993). Un autre travail d'importance – s'intéressant, plus particulièrement, à la grammaticalisation du démonstratif TEN en article défini dans un corpus de productions orales narratives – est la thèse de doctorat de Zíková *Gramatikalizační potenciál anaforické funkce lexému ten v mluvených narrativech* (soutenue en 2017).

Certaines études de bohémistes et de slavistes consacrées (entre autres) au sujet du démonstratif tchèque s'inscrivent dans une approche contrastive, comparant ainsi le tchèque à d'autres langues : cf. Pešek (2012 et 2014) pour le tchèque et le français, Souček (2011) pour le tchèque, l'anglais et le finnois, Erbenová (2016) pour le tchèque et l'espagnol, Nědolužko (2004) pour le tchèque et le haut-sorabe, Czardybon (2017) pour le tchèque et bon nombre d'autres langues slaves et Trovesi (2002) pour le tchèque, le haut-sorabe et le slovène.

Pour le domaine du français, il existe, à notre connaissance, un nombre plus élevé de travaux sur le démonstratif en synchronie, même s'il serait naïf de prétendre que ceux-ci parviennent à rendre compte du sujet dans toute sa complexité. Que soient ici mentionnés – sans aucun souci d'exhaustivité ou de hiérarchisation – au moins les ouvrages suivants : *Indéfini, défini et démonstratif : constructions linguistiques de la référence* (1987) de Corblin, *La référence et les expressions référentielles en français* (2002) de Charolles, *Le démonstratif en français : essai d'interprétation psychomécanique* (1995) de Guénette et *La détermination nominale : quantification et caractérisation* (1982) de Wilmet, sans oublier, bien sûr, une pléiade d'articles publiés par une série d'auteurs (Bonnard, 1972 ; Corblin, 1983 et 1995 ; De Mulder, 1998 ; Gary-Prieur, 1998 et 2001 ; Gary-Prieur et Noailly, 1996 ; Guillot, 2004 et 2006a ; Jonasson, 1998b ; Kleiber, 1983b, 1984, 1986a, 1986b, 1986c, 1986d, 1990, 2003, 2004, 2006 et 2007, etc.).

S'agissant des travaux francophones sur le démonstratif, les descriptions démonstratives y sont parfois étudiées avec les autres types d'expressions référentielles, telles les descriptions définies, indéfinies et les noms propres, cette approche ayant pour incontestable avantage de faire ressortir à la fois les similitudes et les divergences qui existent entre le démonstratif et ces autres expressions. Certains travaux du domaine du français, dont une quantité importante d'articles de Kleiber, appréhendent le démonstratif dans le cadre plus général de la problématique des expressions indexicales (cf. Kleiber, 1986 ; Jungbluth et da Milano, 2015). Ensuite, un nombre considérable de travaux de ce domaine (d'expression française ou pas) s'intéressent également au démonstratif français en diachronie (y compris à ses origines latines) et à la relation de parenté entre celui-ci et l'article défini (cf. Caboli, 1990 ; Carlier et De Mulder, 2007 et 2010 ; Carlier et Guillot, 2015 ;

Epstein, 1993 et 1994 ; Guillot, 2005a, 2005b, 2010a, 2010b, 2013 et 2017 ; Marchello-Nizia, 2003, 2004, 2006a et 2006b ; Orlandini, 1981, etc.). Ces références nous serviront d'appui à la fois lorsque nous retracerons l'histoire du démonstratif français et lorsque nous étudierons les emplois où le démonstratif adnominal tchèque TEN s'approche, de par sa sémantique et sa fonction, d'un article défini.

En dehors de la linguistique tchèque et française, il existe également, bien évidemment, des études qui se penchent sur le démonstratif et la problématique de la deixis dans une perspective soit générale (cf. Bühler, 2008 [1934] ; Diessel, 2006 ; Fillmore, 1982 et 1997 ; Hanks, 1992 ; Himmelmann, 1996 ; J. Lyons, 1977), soit typologique et translinguistique (cf. Dixon, 2003 ; Diessel, 1999 ; Himmelmann, 1997 ; Anderson et Keenan, 1996). Par souci d'exhaustivité, rappelons aussi une série de travaux d'orientation plus générale (nombre de ces travaux se penchent sur les problématiques de la définitude et de l'émergence des articles définis), qui, à un moment donné, sont amenés à traiter de l'indexicalité⁷ et du démonstratif (cf. Christophersen, 1939 ; Czardybon, 2017 ; Greenberg, 1978 ; Halliday et Hasan, 1976 ; Hawkins, 1978 ; Himmelmann, 2001 ; Hlavsa, 1975 ; Löbner, 1985 et 2011 ; C. Lyons, 1999 ; J. Lyons, 1990 [1978], Schwarz, 2009 ; Zimová, 1994).

Notre propre projet se cantonne à une étude des démonstratifs en position dite « adnominale », c'est-à-dire là où le démonstratif introduit un nom (ou un élément substantivé), dans un corpus contenant essentiellement des productions orales informelles. Ce corpus est composé de deux parties de tailles et de natures à peu près comparables pour chaque langue. Outre ce corpus principal, nous recourrons également et de manière complémentaire à un sous-corpus composé de pièces de théâtre du dramaturge tchèque V. Havel et de leurs traductions françaises. En travaillant avec ces données, nous espérons éclairer certaines des spécificités des deux langues et mettre en évidence les usages dominants ou émergents des démonstratifs dans la langue orale.

Le présent travail sera structuré en trois parties. Une première partie visera à poser un cadre théorique général concernant le démonstratif et sa place au sein de la problématique de la deixis, à présenter les sources des données empiriques analysées et le cadre méthodologique adopté (prélèvement des échantillons et annotation) et, enfin, à proposer un état de l'art de la recherche sur le démonstratif avec ses paradigmes dans les deux langues. La seconde partie du travail sera avant tout consacrée aux analyses quantitatives du matériel empirique : nous y présenterons des relevés quantitatifs comparant les fréquences des différentes formes adnominales du démonstratif ainsi que

⁷ Les termes de « deixis » et « indexicalité » seront employés comme synonymes dans le présent travail.

de leurs types d'emploi dans les deux langues. La troisième partie sera, quant à elle, consacrée aux singularités du système tchèque. Voici une présentation plus détaillée du plan adopté.

Structure du travail

Partie I

L'enjeu de cette première partie sera le suivant : poser les fondements théorico-méthodologiques de notre projet tout en présentant 1) un état de l'art de la problématique du démonstratif, à la fois dans une perspective générale et dans les perspectives des langues tchèque et française, et 2) le matériel analysé avec la démarche exacte de l'analyse (notre question de recherche ainsi que les principales hypothèses seront présentées ci-après, dans la Section 0.3).

Cette partie s'ouvrira sur le Chapitre 1, où nous ferons le point sur la problématique du démonstratif et de sa place au sein de l'ensemble des expressions indexicales. Cette démarche sera à la fois sémasiologique (le fonctionnement sémantico-référentiel du démonstratif) et onomasiologique (les différents types d'emploi du démonstratif). Nous y mobiliserons à la fois des références théoriques ressortissant au cadre franco-belge (notamment pour ce qui est des spécificités sémantico-référentielles du démonstratif) et des références ayant une vocation plus universelle, appréhendant le démonstratif sur les plans morphologique, syntaxique et sémantique dans une approche typologique. Une introduction à la diachronie du démonstratif figurera également à la fin de ce chapitre.

Le Chapitre 2 sera consacré à la description de nos corpus et de la méthodologie d'analyse. C'est là que nous décrirons nos catégories et nos valeurs d'annotation, en partant d'une typologie avant tout onomasiologique de l'indexicalité et en étayant les explications avec des exemples concrets. Notre parti pris méthodologique pour la comparaison des langues sera exposé : outre la description des corpus, il s'agira de présenter notre propre point de vue sur le problème de l'exploitation de corpus « comparables » et « multilingues » dans une démarche de linguistique contrastive comparant deux langues typologiquement différentes, avec tout ce que cette pratique comporte en termes de difficultés et de pièges.

Dans les deux chapitres qui suivront (Chapitres 3 et 4), nous traiterons du démonstratif dans les systèmes des deux langues qui nous intéressent (en tchèque, d'abord, puis en français). Ce bilan sera essentiellement de nature morphosyntaxique, sémantico-référentielle et pragmatique, tout en s'accompagnant d'un rappel diachronique pour chaque langue. Nous passerons également en revue, dans ces deux chapitres, les différents modes d'emploi du démonstratif adnominal en français et en tchèque, compte-tenu à la fois des régimes d'indexicalité (endophoriques et exophoriques) et de la forme du démonstratif. Force est cependant de constater que ces deux chapitres ne seront pas uniquement de nature théorique, car nous y travaillerons déjà avec les résultats de nos annotations.

Ainsi, de nombreux exemples puisés dans notre matériel empirique seront cités et commentés dans le but d'étayer nos explications. Afin d'avoir une idée de la fréquence du démonstratif adnominal dans l'oral informel, nous effectuerons, pour les deux langues, une série de comparaisons de sa fréquence dans les données d'oral informel avec sa fréquence dans les données d'autres types de corpus, composé d'autres types de productions langagières. Nous verrons les points communs quant au fonctionnement du démonstratif dans nos deux langues (son indexation forte du référent sur le contexte, sa mise en valeur dans la (re)classification du référent, le paradoxe dit « de la reprise immédiate », etc.), mais de cette analyse ressortiront déjà des divergences saillantes, sur lesquelles nous nous centrerons au premier chef dans le reste du travail. Ici, nous aborderons pour la première fois les cas qui nous intéresseront avant tout dans la troisième partie, à savoir la grammaticalisation du démonstratif tchèque TEN en article défini, les emplois « émotionnels » du démonstratif tchèque et la possibilité de cumuler, en tchèque, le démonstratif avec plusieurs autres éléments ressortissant au domaine fonctionnel (Abney, 1978). Nous sommes convaincu que ces trois phénomènes soulèvent des pistes de recherche qui méritent qu'on y prête attention, et que leur étude peut approfondir notre compréhension actuelle des démonstratifs en tant que catégorie universelle.

Partie II

La seconde partie sera, quant à elle, consacrée à un bilan contrastif plus poussé, et qui s'appuiera sur une analyse essentiellement quantitative du matériel empirique (Chapitre 5). Nous y comparerons les deux langues eu égard à la fréquence du démonstratif adnominal tout court, compte-tenu de ses différentes formes, de ses différents types d'emploi, compte-tenu, également, de la position de l'expression démonstrative par rapport au verbe de la proposition, eu égard à la présence de disfluences, à la grammaticalisation présumée touchant certaines occurrences de la forme tchèque TEN, etc. La fréquence exceptionnelle du démonstratif tchèque mise en évidence dans notre corpus d'oral informel constitue pour nous un indice clair que celui-ci remplit des fonctions en partie différentes dans les deux langues, en grande partie parce qu'en tchèque, il empiète sur le domaine qui est réservé en français à l'article défini. Afin de mieux examiner cette hypothèse, nous procéderons à une comparaison de la fréquence de TEN en emploi adnominal dans notre corpus de conversations tchèques avec celle de l'article défini français LE dans différents types de productions langagières (types de textes).

Partie III

Cette partie consistera en une étude approfondie de deux phénomènes concernant davantage ou exclusivement le tchèque. Le premier phénomène, traité dans le Chapitre 6, est le cas où le tchèque – notamment dans son registre oral informel – recourt à la forme du démonstratif adnominal TEN –

neutre à l'égard de l'expression de la distance du référent par rapport au locuteur – alors que le français (et les langues à article défini en général) recourt automatiquement ou plus volontiers à l'article défini dans le même contexte⁸. Ce phénomène touche les contextes exophoriques situationnels, anaphoriques (y compris certains contextes d'anaphore associative), mémoriels, mais aussi les contextes pour lesquels certains auteurs emploient le terme de « cataphore structurale » ou « intra-syntagmatique » (cf. Fraser et Joly, 1979 : 48 ; Haliday et Hassan, 1976 : 72 ; Kesik, 1989 : 18 et 22), à savoir les cas où le nom est rendu défini par la présence des compléments tels que les adjectifs superlatifs, les adjectifs numéraux ordinaux et les propositions relatives dites « établissantes » (cf. Hawkins, 1978 : 131)⁹. Que soient ici cités juste deux exemples à titre indicatif, un exemple de TEN dans le régime de la deixis situationnelle et un exemple de TEN introduisant un nom dont la nature définie (la définitude) est instaurée par un numéral ordinal :

- 1) Dans une situation où une seule pochette bleue est présente :

Pod-ej	mi	prosim	ty	modr-é	desk-y
passer-IMP.2SG	moi.DAT	s'il te plaît	ten.F.ACC.PL	bleu-ACC.PL	pochette-ACC.PL tantum

Passe-moi la pochette bleue, s'il te plaît.

- 2) a **ten** **druh-ej** **film** ten dáva-l-i
et ten.M.ACC.SG second-M.ACC.SG film.ACC.SG celui-ci.M.ACC.SG donner-PST-3PL.M

včera na dvoj-ce ..
hier sur deux-LOC.SG

le second film, lui, il est passé hier sur la chaîne deux ORTOFON v1 (Corpus national tchèque)

Notre hypothèse explicative partira de l'intuition selon laquelle le démonstratif tchèque TEN subit un processus de grammaticalisation pour devenir un article défini. Une section liminaire du chapitre sera ainsi consacrée à une brève présentation de la notion de grammaticalisation et à la problématique de l'émergence des articles définis à partir de démonstratifs adnominaux et, plus généralement, à la question de la définitude. Des sources de référence de la littérature existante seront mobilisées dans cette présentation. Un parallèle avec l'évolution de l'article défini français à partir de la forme démonstrative ILLE en latin tardif sera également proposé. Par ailleurs, nous sommes d'avis que les données de notre sous-corpus composé de pièces de V. Havel pourraient nous aider à étayer les constats faits sur les données de l'oral informel : nous partirons de l'hypothèse que les

⁸ Jusqu'à présent, ce statut « articulaire » de TEN a été reconnu plus ou moins explicitement par toute une série d'auteurs, tels Adamec (1983), Berger (1993), Dvořák (2020), Hlavsa (1975), Mathesius (1926), Meyerstein (1972), Orlandini (1981), Štícha (1999), Souček (2011), Trovesi (2002), Vey (1946), Ziková (2017 et 2018), etc.

⁹ Dans l'approche de Löbner (1985 et 2011), ces différents contextes relèvent soit de la définitude « pragmatique » soit d'une zone de transition entre la définitude « pragmatique » et la définitude « sémantique ». Voir *infra*.

emplois de la forme *TEN* affichant des traits de la grammaticalisation seront davantage rendus dans les traductions françaises par l'article défini.

Le second phénomène, traité dans le Chapitre 7, concerne les emplois du démonstratif adnominal tchèque que nous avons fait le choix d'appeler « émotionnels », en empruntant ce terme à Mathesius (1926). En fonction des auteurs des deux traditions, ces emplois sont aussi appelés « secondaires » (Adamec, 1983 ; Štícha, 1999), « expressifs » (Štícha, 1999), « affectifs » (Kleiber, 1991 ; Trávníček, 1951 ; Voráč, 1946), « de familiarité/distanciation » (Kleiber, 1991), etc. Dans tous les cas, ils sont censés occuper la « périphérie du système du démonstratif »¹⁰. Ces usages du démonstratif sont intrigants pour deux raisons. D'abord, ils semblent défier le fonctionnement sémantico-référentiel du démonstratif tel qu'il est habituellement défini et en vertu duquel cette unité linguistique est dotée d'un sens instructionnel déictique permettant de diriger l'attention de l'interlocuteur vers le référent et de l'identifier. Dans les emplois en question, le démonstratif peut introduire dans le discours des noms dont le référent est parfaitement identifiable pour les personnes de l'interlocution. C'est typiquement le cas des noms à référent unique (*unica*), des noms à référent générique et des noms propres¹¹. Ensuite, ces emplois « périphériques » du démonstratif semblent intriguer les linguistes des deux langues par la diversité des valeurs pragmatiques qu'ils peuvent véhiculer : tantôt ils marquent l'inclusion du référent dans la sphère du locuteur, tantôt son rejet de celle-ci, la familiarité ou, au contraire, une prise de distance avec le référent, etc. Citons ici deux exemples pour chaque langue :

3) Dupont vient encore de faire des siennes. Ah ! ***Ce Dupont*** ! qu'est-ce qu'il nous en a fait voir !
(Kleiber, 1991a : 93)

4) Ils sont fous, ***ces Romains*** ! (Kleiber, 1991a : 95)

5) Jen	kdyby	mne	<i>ta</i>	<i>hlav-a</i>
seulement si.CONJ+AUX.COND.3SG	moi.DAT	ten.F.NOM.SG	tête-NOM.SG	
přesta-l-a	bolet!			
cesser-PST-3SG.F	faire mal.INF			

Si seulement je n'avais plus mal à la tête ! (Mathesius, 1926 : 40)

6) Une employée de banque parle de sa journée au travail :

¹⁰ « L'on peut, à mon avis, considérer comme généralement accepté (ou acceptable) le fait qu'il est possible d'attribuer au lexème *ten* trois ensembles de fonctions : 1) les fonctions déictiques (monstratives) dénotatives (d'identification), 2) les fonctions de renvoi contextuel et discursif (anaphoriques et cataphoriques) et 3), les fonctions expressives, dites "additionnelles". Les deux premiers ensembles se trouvent au centre du système, le troisième, à sa périphérie. » (Štícha, 1999 : 124) Nous traduisons : „Lze myslím pokládat za obecně akceptovaný (resp. akceptovatelný) fakt, že lexému *ten* lze přisoudit trojí okruh funkcí: 1. deiktické (ukazovací) funkce denotační (identifikační); 2. kontextově-diskurzivní (anaforické a kataforické) funkce odkazovací; 3. expresivní funkce „přídátečné“. První dva okruhy funkcí jsou systémově centrální, třetí je systémově periferní.“

¹¹ Sauf dans le cas des noms génériques, il s'agit de représentants de la définitude dite « sémantique » par Löbner (1985 et 2011).

jako	dneska	jsem	mě-l-a	devět	lid-í	vid' ..
MOD	aujourd'hui	AUX.PST.1SG	avoir-PTCP-1SG.F	neuf.NOM	gens-GEN	MOD
prostě	ty	lid-i	ne-chod-ěj	vůbec ..		
MOD	ten.M.NOM.PL	gens-NOM	NEG-vont-PRS.3PL	du tout		

par exemple aujourd'hui, j'ai reçu neuf clients [en tout et pour tout], tu vois... genre les gens, ils viennent pas chez nous du tout...

ORTOFON v1

Si de tels emplois sont nettement plus répandus en tchèque (telle est, en tout cas, notre hypothèse de départ), cela ne veut pas dire pour autant que l'on n'en trouve pas en français. Il existe un certain nombre d'études d'auteurs francophones sur ce sujet (cf. Gary-Prieur, 1994 et 2001 ; Jonasson, 1994 et 1998a ; Kleiber, 1991a), études qui nous serviront d'appui avec les sources disponibles pour le tchèque (notamment Adamec, 1983 ; Mathesius, 1926 et Kresin, 2008). Cependant, par comparaison avec le tchèque, ces emplois du démonstratif seraient régis en français par des contraintes beaucoup plus fortes, comme le suggère justement la littérature existante. Étant donné ces divergences entre les deux systèmes, il faudra voir, entre autres, si les ébauches de solutions formulées pour l'une et l'autre langue peuvent également être retenues pour l'autre et, si oui, dans quelle mesure : les approches fondées sur les spécificités de chaque système peuvent-elles se combiner entre elles pour s'éclairer mutuellement ou faut-il abandonner toute tentative de rapprochement ? Parmi d'autres questions, se pose ici la suivante : qu'en est-il, dans ce cas, de la sémantique indexicale (déictique) du démonstratif ? Nous partirons du principe que la présentation – mais non plus l'identification – du référent continue à s'effectuer via un recours explicite au contexte, opération dont le démonstratif est toujours responsable¹². Plus généralement, ce cas de figure nous amènera à dissocier dans le démonstratif, d'une part, ce qui relève de son sens indexical tel qu'il est codé par le système de la langue et, d'autre part, ses fonctions. Si la plus importante de ses fonctions, découlant directement de son sens indexical, est celle qui consiste en l'identification d'un référent (fonction « désignative » ou « instructionnelle »), d'autres fonctions sont également susceptibles d'apparaître (par inférence pragmatique) si certaines conjonctures particulières sont réunies.

Sera également abordée, dans les Chapitres 6 et 7 et en rapport avec les deux phénomènes susmentionnés, la possibilité de cumuler en tchèque devant un nom plusieurs éléments adnominaux appartenant au même domaine fonctionnel¹³, dont un démonstratif. On abordera donc des cas comme ceux-ci :

¹²Ainsi, Kleiber écrit dans le résumé d'un article consacré au démonstratif et à la question du point de vue : « Chemin faisant, nous serons amené à distinguer entre la valeur unitaire, basiquement déictique, du démonstratif et ses emplois discursifs qui exploitent de façon fort complexe sa déicticité première. » (2003 : 33)

¹³ Dans l'ensemble de ce travail, nous partons de la position adoptée par Himmelmann (1997), selon laquelle il est illégitime de parler de déterminants pour une langue sans article défini ; en réalité, il semble que la catégorie morphosyntaxique des déterminants tout comme la position syntactico-distributive déterminative et même l'unité

- 7) Kdy po-jed-em na **tu** **tvoj-i** **chat-u?**
 quand FUT-aller-1PL à ten.F.ACC.SG tien-F.ACC.SG chalet-ACC.SG
Quand est-ce qu'on partira à ton chalet ?

Sauf cas exceptionnels qui seront notamment décrits dans le Chapitre 4, et où le démonstratif se combine avec un pré-déterminant et/ou un post-déterminant, une telle cumulation est proscrite dans le français d'aujourd'hui par des contraintes syntaxiques et distributionnelles¹⁴.

En revanche, dans le tchèque d'aujourd'hui – et, de nouveau, notamment dans son registre parlé informel –, ce type de cumul est très courant, le démonstratif s'y associant non seulement à un possessif, mais aussi à un indéfini (NĚJAKÝ) ou à un démonstratif indéfini (TAKOVÝ). Dans certains cas, l'ordre de ces deux déterminants peut changer alors qu'il ne le peut pas dans d'autres, sachant que ce changement d'ordre semble toujours entraîner un (léger) changement de sens. Nous étudierons ces combinaisons plus en détail, en nous appuyant sur notre matériel empirique.

Conclusion

À la toute fin de ce travail, nous tâcherons d'apporter une conclusion synthétique à nos analyses et observations, tout en proposant une ouverture vers d'autres horizons de recherche à destination des auteurs qui auront l'intention de s'inscrire dans une veine proche de la nôtre.

Question de recherche

La question de recherche centrale qui sous-tend notre projet peut être formulée comme suit :

Y a-t-il des différences substantielles – aussi bien sur le plan quantitatif que qualitatif – en matière d'emploi des démonstratifs adnominaux au sein du registre parlé informel du français et du tchèque ? Si oui, alors comment expliquer ces différences ?

L'hypothèse dont nous partons – et que le projet cherche ainsi à confirmer – est que l'oral informel tchèque emploierait le démonstratif adnominal beaucoup plus et dans davantage de contextes que l'oral informel français, et ce notamment pour deux raisons.

syntactique appelée « syntagme/groupe nominal/déterminatif » apparaissent à l'issue du processus de grammaticalisation qui a également produit l'article défini. Ainsi, le terme de déterminant ne convient pas pour les emplois adnominaux du démonstratif tchèque. De même, n'étant pas du tout certain que le terme de « syntagme nominal/déterminatif » (SN/SD) convienne pour le tchèque, nous l'emploierons uniquement pour le français ; pour le tchèque, nous parlerons de « complexe nominal » (= le nom + tous ses compléments et actualisateurs). Ces choix seront argumentés dans la Section 3.1.3 du Chapitre 3.

¹⁴ Des contraintes menant à une structuration plus rigide au niveau de l'unité syntaxique que constitue le nom avec ses compléments sont typiques des langues à articles définis en général (voir Himmelmann, 1997).

La première en est la grammaticalisation en cours de TEN en article défini. Afin d'évaluer cette dernière, nous nous appuyons sur la distinction élaborée par Löbner (1985 et 2011) entre deux types de définitude, à savoir la définitude « pragmatique » et « sémantique ». Cette distinction sera présentée de façon détaillée dans le Chapitre 6 (Section 6.1.3.4). Contrairement à la définitude « pragmatique », établie moyennant une indexation du référent sur le contexte immédiat, la définitude « sémantique » « [...] suppose que le référent peut être identifié à partir du seul contenu descriptif du SN, compte tenu des associations stéréotypées qu'il entretient avec un cadre ou scénario à restituer. » (Carlier et De Mulder, 2006 : 102) L'usage habituel du démonstratif se limite au seul domaine de la définitude « pragmatique », domaine que le démonstratif partage avec certains emplois du défini ou, dans les langues sans articles, la détermination zéro. En revanche, sa présence est proscrite dans le domaine de la définitude « sémantique ». Par conséquent, lorsque l'emploi du démonstratif tend à devenir systématique en contexte de définitude « pragmatique » et qu'en plus de cela, il commence à pénétrer dans certains contextes situés à la charnière entre les définitudes « pragmatique » et « sémantique », il est légitime de conjecturer sa grammaticalisation en article défini. Par ailleurs, il convient de mentionner qu'un lien de causalité a été mis au jour par des chercheurs étudiant la grammaticalisation entre la grammatisation d'un démonstratif et une forte hausse de sa fréquence d'emploi (cf. Kroch, 1989 ; Laury, 1997 ; Zíková, 2017).

Les emplois dits « émotionnels » du démonstratif adnominal tchèque constitueraient, d'après nous, la seconde principale raison de sa haute fréquence tout comme du nombre plus élevé des types d'emploi auxquels il se prête. Les emplois « émotionnels » font en sorte que la présence du démonstratif s'étend à des contextes ressortissant à la pure définitude « sémantique ». Or ici, comme nous le montrerons notamment dans le Chapitre 7, le démonstratif n'adopte pas son mode de fonctionnement habituel dans la mesure où il ne contribue pas à l'identification du référent¹⁵, mais traduit une attitude pragmatique émotionnelle du locuteur vis-à-vis de ce dernier (voir *supra*).

Pour le premier phénomène (la grammaticalisation de TEN), nous renvoyons le lecteur aux exemples 1 et 2 *supra*, mais également aux trois exemples suivants. Le premier est un cas ressortissant à la reprise anaphorique (répétée) d'un référent activé par le discours :

- 8) L1 : ti soused-i jak jsem ti vo nich vyprávě-l-a
 ten.M.NOM.PL voisin-NOM.PL REL AUX.PST.1SG toi.DAT de eux.LOC raconter-PTCP-SG.F
- jak tam bydle-l-i u mě . ty NP jak . voni
 REL là-bas habiter-PST-3PL.M chez moi.ACC ten.M.NOM.PL NP.M.NOM.SG REL eux.NOM

¹⁵ Ou, en tout cas, il ne le fait pas de manière directe. Nous développerons ce point dans le Chapitre 7.

mě-l-i	[tapety] _i	tam	padesát	let
avoir-PST-3PL.M	papier peint-NOM.PL	là-bas	cinquante	an.GEN.PL
L2 : no tak to	já	je	ne*	
MOD	moi.NOM	eux.ACC	NEG	
L1 : a ona	ještě	byl-a	kuřák .	a vona rád-a
et elle.NOM.SG	en plus	être.PST-3SG.F	fumeur.NOM.SG	et elle.NOM.SG content-F.SG
děla-l-a	tydlet-y	panenk-y	víš	
fabriquer-PST-3SG.F	tenhleten-F.ACC.PL	poupée-ACC.PL	MOD	
L2 : já	v-í-m			
moi.NOM	savoir-PRS-1SG			
L1 : ona	skoro	byl-a	už	na smrt nemocn-á .
elle.NOM.SG	presque	être.PST-3SG.F	déjà	à mort.ACC.SG malade-F.NOM.SG
L1 : já	jsem	tam	přiš-l-a	ona mi
moi.NOM	AUX.PST.1SG	là-bas	arriver-PTCP-SG.F	elle.NOM.SG moi.DAT
ukazova-l-a	panenk-y .			
montrer-PST-3SG.F	poupée-ACC.PL			
já	řík-á-m	no	Dáš-o	ale je to na úkor .
moi.NOM	dire-PRS-1SG	MOD NP-F.SG.VOC	mais	être.PRS.3SG cela.NOM à détriment.ACC.SG
se	urazi-l-a .	no	poněvač . a	ted' tam přiše-l .
REFL vexer-PST-3SG.F	MOD	parce que	et maintenant	là-bas arriver-PST.3SG.M
nov-ej	majitel .			
nouveau-M.NOM.SG	propriétaire.NOM.SG			
a ty	představ	si	[ty tapet-y] _i	muse-l
et ten.F.ACC.PL	imaginer.IMP.2SG	REFL	ten.F.ACC.PL papier peint-ACC.PL	devoir-PST.3SG.M
strhat .	[to] _i	byl-y	ještě	takov-ý ze západn-fho
arracher.INF	ce	être.PST-3PL.F	encore	tel-F.NOM.PL de occidental-N.GEN.SG
Německ-a	rozumíš .			
Allemagne-GEN.SG	MOD			
proto	mi	tam	děla-l-y	sbíječ-k-y
c'est pour ça que	moi.DAT	là-bas	faire-PST-3PL.F	marteau-piqueur-NOM.PL
[to vše-cko] _i	šl-o	dolů	to	byl-o příšern-ý
cela.NOM	tout-N.NOM.SG	aller.PST-3SG.N	en bas	ce être.PST-3SG.N horrible-N.NOM.SG
tak dlouho	[ty	tapet-y] _i	tam	byl-y že jo
si longtemps	ten.F.NOM.PL	papier peint-NOM.PL	là-bas	être.PST-3PL.F MOD

L1 : ces voisins que je t'ai parlé d'eux, ceux qui vivaient chez moi, ces NP, ceux qui avaient chez eux [des papiers peints]_i depuis cinquante ans

L2 : bon, moi, je les connais pas

L1 : et en plus de ça, elle fumait et elle aimait faire ces espèces de poupées, tu vois

L2 : je vois

*L1 : elle était presque mortellement malade, j'y suis arrivée, elle m'a montré des poupées, je lui dis mais Dáša, ça, c'est au détriment de... elle s'est vexée parce que bon... et maintenant, il y a un nouveau propriétaire qui est arrivé, et imagine qu'il a été obligé d'arracher [ces papiers peints]_i, en plus [c']_iétait des papiers peints d'Allemagne de l'Ouest, tu vois, c'est pour ça que j'avais des marteaux-piqueurs à côté de chez moi, il fallait bien enlever [tout ça]_i, c'était horrible, ça faisait si longtemps que [**les papiers peints**]_i étaient là, tu vois*

ORTOFON v1

Le référent de *tapety* (*papiers peints*) est d'abord introduit dans le discours par un nom nu. Douze propositions plus tard, sa continuité est assurée par TEN. Cette seconde mention établit le nouveau référent dans le discours – nous considérons ainsi TEN comme un vrai démonstratif. Mais seulement cinq propositions plus tard et après que le référent a été repris deux fois par un TEN pronominal (*to* et *to všechno*), le même nom est de nouveau introduit par TEN adnominal. À ce moment de sa reprise, non seulement le référent est identifiable pour celui qui reçoit le message, mais il est aussi bien activé dans son esprit. Ainsi, nous sommes d'avis que TEN est ici plus proche d'un article défini, d'où **les papiers peints** dans notre traduction française.

Le second exemple, emprunté à Czardybon, relève d'un type particulier d'anaphore associative. L'anaphore associative se situe à la limite des deux types de définitude présentés *supra* :

- 9) Po kin-ě se ještě bavi-l-i o (**tom**) **film-u**.
 après cinéma-LOC.SG REFL encore discuter-PST-3PL.M de ten.M.LOC.SG film-LOC.SG
*Après être allés au cinéma, ils ont encore discuté **du film**.* (Czardybon, 2017 : 90)

En 9, le nom *film* est employé comme défini en première mention grâce à la situation générale qui est celle d'une sortie au cinéma, qui conduit habituellement au visionnage d'un film. Alors que le français (langue à article défini) peut uniquement employer ici LE, en tchèque (parlé informel), le nom peut apparaître nu ou précédé de TEN.

Le dernier exemple met en scène un emploi mémoriel¹⁶ de TEN. Comme nous le verrons plus tard dans ce travail, ce cas de figure n'est pas sans équivoque, dans la mesure où le locuteur d'une langue à article défini comme le français pourra tantôt préférer LE et tantôt CE, en fonction du coût d'activation du référent (à la fois pour lui-même et pour son interlocuteur). En 10, les locuteurs recourent à leur mémoire partagée pour identifier un référent qui se trouve en même temps représenté sur une photo qu'ils regardent :

- 10) L1 : tady jsou **ty** **promenáď-y** ne ? kolem
 ici être.PRS.3PL ten.F.NOM.PL promenade-NOM.PL MOD autour

¹⁶ La littérature d'expression anglaise emploie souvent le terme de « recognitional uses », proposé par Himmelmann (1996).

ty *chodník-y*
 ten.M.NOM.PL trottoir-NOM.PL

L2 : jo tady-hlenc je *ta* *promenád-a*
 MOD ici ici-SUFFIXE DEICTIQUE être.PRS.3SG ten.F.NOM.SG promenade-NOM.SG

L1 : ici, il y a les/ces promenades et les/ces trottoirs tout atour, non ?

L2 : ouais, là, il y a la/cette promenade ORTOFON v1

Pour les emplois « émotionnels », nous renvoyons le lecteur aux exemples 3, 4, 5, et 6 cités *supra*. Le comportement sémantico-pragmatique exact du démonstratif apparaissant dans ce type de contextes sera abordé de façon plus conséquente dans le Chapitre 3, avant de faire l'objet de l'ensemble du Chapitre 7. Dans ce dernier chapitre, il sera suggéré, entre autres, que la valeur « émotionnelle » du démonstratif pourrait s'analyser comme le fruit d'une implicature conversationnelle au sens de Grice (1975).

Méthodologie : une approche sémantico-référentielle contrastive basée sur un corpus exploité à l'aide du logiciel textométrique TXM

Étant de langue maternelle tchèque et maîtrisant la langue française comme langue seconde, nous nous sommes vite aperçu de l'intérêt d'une voie contrastive pour nos recherches linguistiques. C'est cette même voie qui, dans son stade encore très embryonnaire et intuitif, nous a permis de prendre conscience des divergences existant entre les deux langues et, plus précisément, des particularités propres au démonstratif tchèque. Nous avons fait le choix de combiner l'approche contrastive avec une approche sémantico-référentielle en synchronie, qui, grâce à son articulation de la théorie et de l'analyse de données, nous semble dotée d'un pouvoir explicatif considérable. Nous optons dans notre travail pour une orientation avant tout sémasiologique, inductive, en nous inspirant de la notion certes informelle et cocasse mais très parlante de « linguistique teckel » de Kleiber¹⁷. Ainsi, au tout début d'une entreprise de recherche se trouve une sensation d'étonnement, voire de sidération, déclenchée par un phénomène bien particulier, un emploi « dérangent » d'une forme. Vient ensuite un travail d'analyse et d'élucidation qui part de ladite occurrence pour à la fois en rendre compte et l'inscrire dans un cadre plus global.

Afin de donner à ce travail d'analyse une assise concrète, il nous semble nécessaire de recourir à un matériel empirique suffisamment riche, composé de sources authentiques et adaptées à l'orientation globale du projet de recherche. C'est la raison pour laquelle nous allons notamment travailler avec deux sources de langue parlée informelle, à savoir le Corpus national tchèque (CNT)

¹⁷ G. Kleiber a employé cette métaphore pour décrire son approche de chercheur en linguistique lors du séminaire consacré aux appellatifs intitulé « Mon nom est PERSONNE : qui est JE ? Qui est TU ? », qui s'est tenu le 26 février 2016 à l'EHESS.

et le Corpus d'Étude pour le Français Contemporain (CEFC). Le CNT est une plateforme développée par l'Université Charles de Prague (l'Institut du Corpus national tchèque) et très diversifiée, contenant aujourd'hui plusieurs corpus de langue tchèque parlée informelle, dont ORTOFON v1 (*Korpus neformální mluvené češtiny s víceúrovňovým přepisem*¹⁸), rendu public le 2 juin 2017. Ce corpus est composé de transcriptions de dialogues privés et informels entre locuteurs tchèques qui se connaissent et contient 1 236 508 positions (avec ponctuation, interjections et marques de disfluences ; voir le Chapitre 2). Ce corpus est composé de façon à être parfaitement équilibré eu égard aux différences diatopiques ainsi qu'à l'âge, au sexe et au niveau de formation des locuteurs.

Pour le français, nous avons opté pour le CEFC¹⁹, rendu pleinement public et accessible courant septembre 2018 sur le site de la plateforme Orféo et regroupant une série de plateformes déjà existantes, dont des corpus de langue parlée informelle. Un filtrage s'appuyant sur plusieurs critères (qui seront, eux aussi, présentés de manière détaillée dans le Chapitre 2) permet d'accéder uniquement à ce type de productions. La taille de l'ensemble des productions choisies grâce à ces critères est de 934 102 positions (avec interjections et marques de disfluences²⁰). Malheureusement, ce corpus, qui résulte d'un assemblage de données empruntées à diverses plateformes préexistantes, ne présente pas la même situation d'équilibrage que le corpus ORTOFON v1.

Les textes d'ORTOFON v1 datent tous de la période qui va de 2012 à 2017. L'écrasante majorité des textes retenus pour le français ne remontent pas avant l'an 2000. Nous tenons ces deux sources de données pour suffisamment comparables pour réaliser notre étude, même si les limites de cette comparabilité ne sont pas négligeables. Elles seront présentées dans le Chapitre 2.

L'une des hypothèses de départ – selon laquelle les emplois du démonstratif qui nous intéressent sont plus représentés dans ce registre éminemment progressiste qu'est l'oral informel – exige de travailler, en premier lieu, avec des sources authentiques de langue parlée informelle. Mais, si le tchèque emploie davantage de démonstratifs adnominaux, il n'est pas sans intérêt de se pencher sur les autres moyens que le français utilise à la place. Pour quelles autres « stratégies » optera-t-il si, précisément, il n'utilise pas le déterminant démonstratif là où le tchèque l'utilise ? Pour en avoir une idée, nous avons également décidé de travailler, quoique dans une optique davantage complémentaire, avec le corpus parallèle InterCorp²¹, lui aussi développé par l'Institut du CNT. Il s'agit d'une base composée de textes et de leurs traductions (le nombre de langues incluses s'élève aujourd'hui à 40) avec toujours le tchèque comme langue pivot. L'interface d'InterCorp rend possible pour chaque

¹⁸ *Corpus de tchèque parlé informel disposant d'une transcription à plusieurs niveaux* ; pour plus d'informations, voir le Chapitre 2 ou le site Internet du corpus : <http://wiki.korpus.cz/doku.php/en:cnk:ortofon>.

¹⁹ Pour plus d'informations, voir le Chapitre 2 ou le site Internet du corpus : <https://www.projet-orfeo.fr>.

²⁰ Ce corpus ne contient pas de signes de ponctuation.

²¹ Pour plus d'informations, voir le Chapitre 2 ou le site Internet du corpus : <http://wiki.korpus.cz/doku.php/en:cnk:intercorp>.

usager la définition de sous-corpus personnalisés, et ce via un choix de textes effectué à partir des textes disponibles dans la base. On peut ainsi décider d'inclure uniquement les textes pour lesquels il existe une traduction dans la langue cible qui nous intéresse, d'inclure uniquement les textes qui sont des originaux de la langue source, de choisir la datation des textes, etc.

Au sein de cette base, nous avons décidé de travailler avec un micro-corpus parallèle composé de quatre pièces du dramaturge tchèque Václav Havel et de leurs traductions françaises (48 568 positions au total en tchèque, tous signes compris). Ce choix se justifie avant tout par le fait que ces textes, destinés à être mis en scène, emploient globalement une langue proche de la langue orale informelle, même si des éléments de langue soutenue sont aussi présents et que la datation et l'influence de l'idiolecte de l'auteur/traducteur peuvent paraître problématiques (voir le Chapitre 2). Ce micro-corpus pourra notamment nous fournir un aperçu sur les moyens dont use le français pour rendre le démonstratif adnominal tchèque. Deux phénomènes nous intéresseront ici avant tout : quelles stratégies le français met-il en place pour rendre les effets émotionnels du démonstratif tchèque ? Recourra-t-il systématiquement à l'article défini là où le tchèque emploie le démonstratif adnominal TEN avec une valeur articulaire ou s'approchant de cette dernière ?

Nous disposons de l'ensemble des textes retenus pour l'oral informel sous forme de fichiers téléchargés, ce qui nous permet de les interroger et de les annoter en régime hors ligne. Ainsi, toutes les requêtes portant sur ces textes ont été formulées dans le logiciel de textométrie TXM, développé par le groupe de recherche CACTUS à l'École normale supérieure de Lyon. L'analyse des occurrences a procédé d'une annotation préalable, effectuée dans Excel, où les occurrences obtenues grâce à des requêtes pertinentes ont été exportées. Une fois ce travail d'annotation fait, nous avons réinjecté les occurrences dans TXM. Le logiciel a permis par la suite d'exploiter les annotations sous forme de requêtes portant sur les catégories d'annotation préalablement définies, dans le but de faire des statistiques de fréquence, de trier les occurrences en fonction de différents paramètres et pour les analyser qualitativement. Quant aux textes des quatre pièces de Václav Havel, ils ont été interrogés directement depuis l'interface du site du CNT, nommée KonText, qui permet de formuler des requêtes via le langage CQL (Corpus Query Language).

D'autres sources complémentaires de données ont été utilisées de façon à étayer certaines thèses avancées et afin d'acquérir une vue d'ensemble quant à l'emploi du démonstratif en français et en tchèque, y compris deux séries d'informateurs natifs dans les deux langues. Ces sources seront présentées dans les Sections 2.1.4 et 2.1.5 du Chapitre 2. Ressentant parfois la nécessité de compléter notre gamme d'exemples, nous nous sommes réservé le droit de mobiliser des exemples d'emplois du démonstratif adnominal collectés dans la vie de tous les jours, auprès de locuteurs natifs des deux

communautés linguistiques. Nous nous sommes également autorisé à nous servir d'exemples déjà utilisés par d'autres auteurs et, le cas échéant, à en forger nous-même, en nous fiant à notre intuition de locuteur. Mais si nous ne trouvons cette pratique nullement condamnable en soi (certains grands noms de la linguistique contemporaine ne jurent que par elle), vu la nature de notre projet, il est clairement préférable de minimiser sa place au profit de corpus en tant que sources authentiques.

En même temps, notre projet ne se veut absolument pas ce que l'on appelle parfois en linguistique de corpus « corpus-driven » (cf. McEnery et Hardie, 2012), c'est-à-dire un projet qui s'en remet entièrement au corpus pour formuler jusqu'à sa question de recherche et sa méthodologie. Il se veut plutôt « corpus-based », dans la mesure où il recourt au corpus pour tester sa question de recherche et ses hypothèses²². En ce sens, le projet prétendra à une bonne articulation des approches inductive et déductive, à un va-et-vient entre formulation d'hypothèses théoriques et analyse du matériel empirique. Comme nous l'avons déjà écrit *supra*, le tout sera étayé par une série d'analyses quantitatives où d'autres sources de données complémentaires seront mobilisées.

Modes de citation et de présentation des exemples

Nous tenons ici à dire un mot sur la façon dont ce travail présentera ses exemples et ses citations. Toutes les citations dans les langues française, anglaise et allemande seront présentées dans leur version originale. Toutes les citations en langue tchèque seront traduites par nous en français, et la version originale sera donnée en note de bas de page. Le même principe sera appliqué aux exemples cités, qu'il s'agisse des exemples puisés dans les corpus ou ailleurs. Nous donnerons systématiquement une traduction des exemples tchèques (et de ceux d'autres langues que le français, l'anglais et l'allemand). Cette traduction, comme il est de coutume dans la littérature linguistique, se fera en deux temps : nous donnerons d'abord une traduction littérale (sauf pour les exemples tirés d'InterCorp, où une traduction alignée avec le texte de l'original est déjà présente²³), puis une traduction davantage idiomatique. Pour ce faire, nous suivrons globalement les règles de glosage de Leipzig²⁴, à la seule exception près que la langue des gloses sera le français. Les gloses ont pour but de permettre au lecteur ne maîtrisant pas le tchèque de repérer et d'identifier les formes et les structures employées. Reprenons ici l'exemple 1 cité *supra* :

²² En même temps, les corpus, loin d'être de simples réservoirs à exemples, occupent une place tout à fait cruciale dans notre projet, car ce sont eux qui nous fournissent un échantillon représentatif d'occurrences sur lequel nous cherchons à confirmer ou à infirmer nos hypothèses. La question des limites de leur représentativité est elle aussi posée de manière consciente (voir le Chapitre 2., Section 2.1.1).

²³ Nous recourons uniquement à une traduction littérale lorsque la traduction existante, trop libre, ne contiendra pas les phénomènes qui nous intéresseront.

²⁴ <https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf>.

- 11) Pod-ej mi prosim *ty* *modr-é* *desk-y*.
passer-IMP.2SG moi.DAT s'il te plaît ten.F.ACC.PL bleu-F.ACC.PL pochette-ACC.PL tantum
Passe-moi la pochette bleue, s'il te plaît.

Les formes verbales sont rendues, en français, par la combinaison de l'infinitif auquel s'ajoutent des informations portant sur les différentes catégories verbales. Ces catégories sont :

INF = infinitif
PRS = présent
IMP = impératif
PST = passé
FUT = futur
PTCP = participe passé
COND = conditionnel
1-3= personnes
SG = singulier
PL = pluriel
NEG = négation
AUX = verbe auxiliaire
M = masculin
F = féminin
N = neutre²⁵

Voici un exemple :

- 12) Já ne-v-í-m.
je NEG-savoir-PRS-1SG
Je ne sais pas.

Le futur des verbes imperfectifs se construit, en tchèque, par la combinaison du verbe auxiliaire BÝT (ETRE) conjugué au futur auquel s'ajoute l'infinitif du verbe concerné. C'est pour cela que les différentes formes du futur du verbe BÝT seront glosées par AUX.FUT + les catégories de personne et de nombre chaque fois qu'elles apparaîtront dans ce rôle d'auxiliaire :

- 13) Bud-u mluvit.²⁶
AUX.FUT-1SG parler.INF
Je parlerai.

²⁵ L'indication du genre ne concerne que les formes du participe passé et du participe (passé) passif („přičestí minulé“, „přičestí trpné“).

²⁶ Le tchèque est une langue « pro-drop ».

Quant aux perfectifs, ces derniers expriment le futur en empruntant la forme morphologique du présent (leur sémantique perfective les empêche, dans ce cas, d'exprimer le présent comme tel, d'où ce décalage entre morphologie et sémantique²⁷). Comme c'est la sémantique qui importe davantage pour la compréhension de nos exemples, c'est FUT qui sera majoritairement employé dans ce type de contextes :

- 14) Pošl-e-te.
envoyer-FUT-2PL
Vous enverrez.

Le passé, lui, est formé par la présence de l'auxiliaire BÝT (ETRE) conjugué au présent suivi du participe passé pour les personnes 1 et 2 et par le seul participe passé pour la personne 3. Le participe passé s'accorde en genre et en nombre. Nous avons ainsi pris la décision de gloser les formes des personnes 1 et 2 par AUX.PST + les catégories de personne et de nombre suivi de PTCP + les catégories de genre et de nombre (exemple 15), contrairement à la personne 3, qui sera glosée par PST + les catégories de personne, de genre et de nombre (exemple 16) :

15) Mluvi-l	jsem.	Mluvi-l-a	jsem.
parler-PTCP.M.SG	AUX.PST.1SG	parler-PTCP-F.SG	AUX.PST.1SG
<i>J'ai parlé.</i>	(sujet masculin)	<i>J'ai parlé.</i>	(sujet féminin)
Mluvi-l-i	jsme.	Mluvi-l-y	jsme.
parler-PTCP-M.PL	AUX.PST.1PL	parler-PTCP-F.PL	AUX.PST.1PL
<i>Nous avons parlé.</i>	(sujet masculin)	<i>Nous avons parlé.</i>	(sujet féminin)
-
16) Mluvi-l.	Mluvi-l-a.
parler-PST.3SG.M	parler-PST-3SG.F
<i>Il a parlé.</i>	<i>Elle a parlé.</i>
(sujet masculin)	(sujet féminin)

Lorsque le participe passé se combine avec l'auxiliaire BY (1SG = *bych*, 2SG = *bys*, 3SG et 3PL = *by*, 1PL = *bychom/bysme*, 2PL = *byste*), le tout forme le conditionnel présent. Dans ces cas, le participe continuera à être glosé comme tel, et BY sera glosé par AUX.COND + les catégories de personne et de nombre :

17) Jed-l-a	by.
manger-PTCP-F.SG	AUX.COND.3SG
<i>Elle mangerait.</i>	

Deux formes amalgamées existent en tchèque, qui font apparaître BY : ABY et KDYBY. ABY combine le sens du conditionnel avec un sens conjonctif (souvent de but ; en français, le tout est

²⁷ Il est cependant des contextes où les verbes perfectifs au présent morphologique expriment bel et bien le présent, comme c'est le cas des présents historique et de narration ou bien le cas de l'expression d'un procès itératif au présent.

traduit soit par *que* soit par *pour que*). KDYBY combine le sens du conditionnel avec un sens conjonctif conditionnel (condition potentielle ou irréaliste). Les deux seront glosés comme suit :

(pour) que.CONJ+AUX.COND + les catégories de personne et de nombre

si.CONJ+AUX.COND + les catégories de personne et de nombre

18) Koléba-l ho, aby usnu-l.
bercer-PST-3SG.M lui.ACC pour que.CONJ+AUX.COND.3SG s'endormir-PTCT.SG.M
Il le berçait pour qu'il s'endorme.

Kdybyste spa-l-i.
si.CONJ+AUX.COND.2PL dormir-PTCP-PL.M
Si vous dormiez.

Quant aux formes nominales, adjectivales et pronominales, nous indiquerons systématiquement leur cas. S'ajoute à cela une information sur le nombre et le genre dans le cas des adjectifs et de certains pronoms :

SG = singulier

PL = pluriel

PLtantum = noms étant employés exclusivement au pluriel

M = masculin

F = féminin

N = neutre

NOM = nominatif

GEN = génitif

DAT = datif

ACC = accusatif

VOC = vocatif

LOC = locatif

INS = instrumental

Quelques autres signes apparaissent dans les gloses :

MOD = terme modal (adverbe, particule)

REFL = pronom réfléchi

NP = nom propre de personne ; ce signe a été introduit afin de faciliter l'orientation du lecteur francophone, ignorant de l'onomastique tchèque ; la marque F ou M spécifiera s'il s'agit d'un nom propre féminin ou masculin

SUP = sert à gloser le préfixe *nej-* du superlatif tchèque

SUFFIXE DEICTIQUE = *-hle/-dle* dans certains pronoms et adverbes déictiques tchèques

Le tchèque parlé informel fait souvent usage des mots *jak* (*comment*) et *co* (*que*) pour introduire une subordonnée relative avec un antécédent nominal. Ces formes, provenant d'un adverbe et d'un pronom relatifs, ne se déclinent pas (contrairement aux pronoms relatifs JAKÝ, KTERÝ et JENŽ)²⁸. Elles seront ainsi glosées par REL (= relatif) :

- 19) Ta žensk-á **jak / co** jsem ti o ní
ten.F.NOM.SG femme-NOM.SG **REL** AUX.PST.1SG toi.DAT de elle.LOC.SG
- vyprávě-l.
raconter-PTCP.SG.M
- La femme que je t'ai parlé (d'elle).*

Les démonstratifs tchèques adnominaux seront toujours cités dans les gloses au nominatif singulier, suivi d'abréviations renvoyant aux marques de genre, de cas et de nombre. Les démonstratifs en emplois pronominaux (ne faisant pas a priori l'objet de l'analyse) seront rendus par leurs équivalents français les plus directs, à savoir : *celui-ci*, *celui*, *ce* et *cela* :

- 20) a **ten** **druh-ej** **film** **ten** dáva-l-i
et ten.M.ACC.SG second-M.ACC.SG film **celui-ci.M.ACC.SG** donner-PST-3PL.M
- včera na dvoj-ce..
hier sur deux-LOC.SG
- le second film, lui, il est passé hier sur la deuxième chaîne...* ORTOFON v1 (Corpus national tchèque)
- 21) **To** je Petr.
ce²⁹ être.PRS.3SG NP.M.NOM.SG
C'est Pierre.
- 22) **To** se mi líb-í.
cela.NOM REFL moi.DAT plaître-PRS.3SG
Cela me plaît bien.

Les possessifs tchèques seront glosés au moyen du paradigme français des adjectifs possessifs, du fait de leur plus grande proximité syntaxique avec la catégorie des pronoms qu'avec celle des déterminants (voir le Chapitre 3 pour la justification de cette thèse) :

- 23) můj počítač
mien.M.NOM.SG ordinateur.NON.SG

²⁸ L'on peut parler de « particule relative » (cf. Haspelmath, 2001).

²⁹ Pour les démonstratifs en emploi « identificatoire » (Diessel, 1999 ; voir le Chapitre 1), le cas ne sera pas indiqué, car il s'agit toujours du nominatif.

Les exemples puisés dans les corpus oraux peuvent être cités avec les signes codant pauses et disfluences si cela sert le propos. Ils peuvent aussi être abrégés pour les besoins de la présentation. Si l'exemple se présente sous la forme d'un dialogue, nous indiquerons les prises de parole : L1 (locuteur 1), L2 (locuteur 2), etc. Les exemples du corpus des pièces de V. Havel seront cités dans leur forme originale et leur traduction française. Ils seront suivis d'informations indiquant le/les nom(s) du/des traducteur(s) ainsi que les dates de parution de l'original et de la traduction. Exceptionnellement, nous pouvons proposer notre propre traduction si la traduction existante s'éloigne trop de l'original, et n'affiche pas le phénomène étudié.

Pour citer les différents lexèmes (démonstratifs, autres mots fonctionnels, noms, pronoms, etc.), nous indiquerons ces derniers en petites capitales dès lors qu'il sera question du lemme, mais en lettres minuscules et en italiques dès lors qu'il s'agira uniquement d'une forme appartenant à ce lemme³⁰ :

TEN vs *to*

CELUI vs *celle*

TOUT vs *toute*

Les langues sources de nos exemples seront également glosées par des abréviations si nécessaire :

FR = français

CZ = tchèque

ENG = anglais

DE = allemand

IT = italien

ES = espagnol

Si nous le jugeons opportun, nous pourrions marquer certains exemples comme appartenant au registre informel (INF) ou formel (FOR).

Dans les exemples contenant des expressions coréférentes (notamment celles ressortissant au régime de l'anaphore), cette coréférence sera marquée moyennant un système d'index dans les cas où il importera de la mettre en relief :

- 24) L1 : donc du coup ils vous ont sorti [Bienvenue Chez Les Ch'tis]_i ils vous [l']_i ont diffusé là-bas
L2 : j'étais j'étais pas en France à ce moment-là

³⁰ Cette règle ne s'appliquera pas aux citations d'autres auteurs, dont nous respecterons la graphie originelle.

En 23, les expressions *Bienvenue Chez Les Ch'tis*, *le* et *ce film* étant coréférentes, elles sont marquées par le même index ([]_i). Lorsque la coréférence n'est que partielle, c'est l'indexe []_{i+?} que nous utiliserons :

25) L1 : ben ça reste [Titanic]_i quoi c'est

L2 : je sais pas moi j'aime ça me plaît pas trop [*ce genre de films*]_{i+?} si je peux éviter j'évite

CEFC

Afin de mettre en relief un élément, notamment une occurrence du démonstratif adnominal, nous le marquerons en gras. Dans les exemples, le gras sera combiné aux italiques (voir l'exemple 25 *supra*). Le soulignement sera lui aussi utilisé dans les exemples, comme outil de mise en relief d'éléments et/ou de passages de moindre importance que les passages en gras et en italiques. Il sera notamment utilisé dans le cas de l'anaphore dite « résomptive » pour indiquer l'extrait faisant l'objet de l'opération résomptive :

26) et puis si tu dis ça à partir de *ce moment-là* tu enfin les gens se

...

les gens font encore plus attention à leurs expressions CEFC

Il en ira de même de l'anaphore associative pour identifier l'argument contextuel permettant l'introduction du nom comme défini en première mention :

27) Nous sommes arrivés dans une ville inconnue. *La mairie* était un somptueux édifice baroque.

Dans le corps du texte, nous nous servirons également d'une série de signes typographiques, à commencer par les guillemets afin d'introduire des citations courtes et la terminologie que nous emprunterons à d'autres auteurs. Quant aux citations longues, elles seront présentées en retrait par rapport au reste du corps du texte. Le cas échéant, c'est la version traduite qui apparaîtra dans le corps du texte, l'original étant cité en note de bas de page³¹. Le gras sera utilisé notamment pour une meilleure organisation du texte et la mise en exergue de la structuration de notre propos. Ainsi, une notion sera susceptible d'être marquée en gras dès lors qu'elle constituera un thème important qu'il s'agira de souligner dans le texte. L'italique sera employé pour la paraphrase des exemples ainsi que pour la citation de termes et de passages ressortissant aux exemples. Parfois, nous recourrons également au surlignage jaune.

³¹ Il sera introduit par la formule « Nous traduisons ».

Partie I

Fondements théorico-méthodologiques et état de l'art

1 Le démonstratif : présentation générale

L'objectif de ce premier chapitre est de proposer une introduction générale au sujet du démonstratif. Moyennant le rappel d'une série de constats cruciaux formulés par les chercheurs à l'égard de cette unité linguistique singulière, il s'agira ainsi de poser les jalons pour le reste de notre propre travail. Précision importante : il ne saurait en aucun cas être question d'un compte-rendu exhaustif du sujet, et ce non seulement parce que cela dépasserait le cadre du présent travail, mais aussi parce que la question du démonstratif est loin d'être « résolue » par la recherche. Dans ce chapitre liminaire, nous nous référerons notamment au travail de Diessel (1999) et aux nombreuses études de Kleiber (1983b, 1984a, 1986a, 1986b, 1986c, 1986d, 1991b, 2007). Nous recourrons également aux travaux de Himmelmann (1996 et 1997), Dixon (2003), Anderson et Keenan (1985), Corblin (1987), Vuillaume (1980), Charolles (2002), Gary-Prieur (2011) et d'autres. Le plan du chapitre est le suivant : dans un premier temps, nous ferons un rappel sur la notion de deixis et ses approches théoriques et nous nous pencherons sur la place qu'occupe le démonstratif au sein des unités déictiques. Ce rappel nous semble indispensable non seulement pour la rigueur théorique du travail, mais aussi en vue des analyses de corpus à venir, où les occurrences du démonstratif seront systématiquement examinées eu égard au régime de deixis (d'indexicalité) dans lequel elles apparaissent. Ensuite, nous traiterons des singularités sémantico-référentielles du démonstratif. Dans un troisième temps, nous nous pencherons sur le démonstratif dans une perspective typologique et contrastive ; plusieurs plans seront ici pris en compte, notamment celui de la variabilité du système du démonstratif dans les langues du monde, ainsi que la sémantique et la morphosyntaxe du démonstratif. Nous finirons en abordant brièvement le démonstratif dans une optique diachronique.

1.1 La deixis : un grand débat linguistique

Tout travail traitant du démonstratif doit inévitablement se situer par rapport au débat sur la deixis. Aussi essaierons-nous ici de faire un état de la question. Compte-tenu, cependant, de la complexité que l'on reconnaît au sujet, cet état de l'art ne pourra se faire qu'au prix d'une brièveté et d'une simplification considérables. Ainsi, plutôt que d'aspirer à une exhaustivité feinte – car impossible par définition –, les lignes qui suivront auront davantage pour but d'afficher et de formuler, au fil d'un passage en revue critique des moments du débat qui nous semblent importants, nos partis pris théoriques et terminologiques. Parmi les différentes propositions, les questions suivantes peuvent être posées : quel(s) modèle(s) de la deixis trouvons-nous le(s) plus approprié(s) ? Comment définir

le rapport entre deixis et anaphore ? À laquelle des acceptions du terme de deixis – assez problématique par sa polysémie – souscrivons-nous ? C’est à ce type de questions que nous tâcherons d’apporter des réponses engageant le présent travail.

1.1.1 La double « découverte » de la deixis et de l’énonciation

Comme l’écrit Kleiber (1986d : 4) :

Les déictiques ont été, on le sait, à l’origine de deux importantes évolutions en linguistique : l’abrogation du dogme saussurien langue-discours, avec le cap mis sur les terres de l’énonciation, et l’avènement de la pragmatique par l’élargissement de la sémantique vériconditionnelle aux phrases hébergeant les déictiques.

Bühler (2009 [1934]), référence incontournable de la recherche moderne sur la deixis, accentue ce que l’on pourrait appeler, dans une légère hyperbole, le fonctionnement binaire du langage humain, phénomène qui se manifeste, en son sein, par l’existence de deux champs : « symbolique »¹ et « déictique » (« monstratif »). Cette binarité n’est cependant qu’apparente, car il s’agit, en réalité, d’une complémentarité, le langage humain étant ainsi fait qu’il mélange ces deux champs de la façon la plus naturelle. Par conséquent, c’est plutôt le terme d’« hybride », et non pas « binaire », qui convient pour sa description. Dans la deuxième partie de sa *Théorie du langage : la fonction représentationnelle*, Bühler – et c’est sans doute l’un des aspects révolutionnaires de son travail – nous semble insister sur cette hybridité. Selon lui, une partie de la logique (et même de la psychologie) se serait condamnée à une impasse dans son étude des langues naturelles précisément en ne tenant pas (suffisamment) compte du « champ déictique » et du mode référentiel très spécifique qui lui est associé.

Bühler n’est pas, loin de là, le premier à théoriser la notion de deixis. Des auteurs de l’Antiquité comme Priscien, certains auteurs médiévaux et de la Renaissance, des grammairiens classiques ainsi que des grammairiens allemands du 19^{ème} siècle comptent au nombre de ses prédécesseurs². Le terme de deixis lui-même provient de l’adjectif grec *deiktikos*, lui-même issu du verbe *deiknumi*, qui veut dire « montrer par geste », « indiquer par ostension »³. Avant même l’intérêt moderne pour le phénomène – que l’on pourrait, non sans une partie d’arbitraire, dater à partir de la publication par Brugmann de son texte *Die Demonstrativpronomina der indogermanischen Sprachen* en 1904⁴ –, plusieurs tournants se produisent qui contribuent à l’émergence de cette modernité.

¹ Ce terme est entendu, chez Bühler, au sens saussurien et peircien du signe linguistique, dont les deux composantes (signifiant et signifié) sont liées par un lien conventionnel.

² Pour un aperçu historique beaucoup plus conséquent, nous renvoyons le lecteur à la thèse de doctorat de Vuillaume (1980).

³ Rappelons ici que le terme de *démonstratif* (*demonstrativus*) n’est qu’une traduction latine du terme grec *deiktikos*.

⁴ C’est dans ce texte que Brugmann propose la typologie des déictiques à quatre classes : la « Dér-Deixis », la « Ich-Deixis », la « Du-Deixis » et la « Jener-Deixis ».

Il convient d'abord de citer Beuzée et son étude des pronoms, une catégorie à la source d'un grand nombre de difficultés pour la grammaire, qu'elle lui soit antérieure ou contemporaine. Beuzée décrit la spécificité des pronoms personnels en plaçant au centre de sa description ce qu'il appelle l'« acte de la parole » et en insistant sur les rôles que les trois personnes représentent au sein de cet acte. Contrairement aux noms, qui « [...] expriment des êtres déterminés, en les désignant par l'idée de leur nature [...] », les pronoms peuvent « [...] désigner des êtres de diverses natures [...] » (Beuzée, 1974 [1767] : 269), le seul critère étant le rôle que l'entité assume ou qu'elle se voit attribuer dans l'« acte de la parole ». Lorsque Beuzée remarque que la nature du référent est d'une importance bien moindre dans la référence pronominale que dans la référence nominale, il touche au problème de la complexité sémantico-référentielle de l'ensemble de cette classe assez hétérogène d'unités linguistiques « [...] qui fonctionnent comme indices de l'inscription dans l'énoncé du sujet d'énonciation [...] » (Kerbrat-Orecchioni, 2014 [1999] : 36). En effet, ce que les déictiques partageraient avec d'autres unités « subjectives » du langage, c'est une certaine difficulté que l'on éprouve à vouloir définir à la fois leur sens descriptif (dénotatif) et leur extension :

À la différence des termes objectifs, dont la classe dénotative a des contours relativement stables, celle des termes subjectifs est un ensemble flou : l'appartenance d'un x à la classe des professeurs, des célibataires, des anciens combattants, ou même des objets jaunes, est admise ou rejetée plus unanimement, et peut se vérifier plus facilement, que son appartenance à la classe des imbéciles, ou des beaux objets. (Kerbrat-Orecchioni, 2014 [1999] : 80)

Pour ce qui est ensuite du terme d'« acte de la parole » lui-même, il n'est pas sans évoquer celui d'« énonciation », théorisé plus tard par Benveniste dans son article « Appareil formel de l'énonciation » (1974)⁵. Ainsi, selon Vuillaume (1980 : 29) :

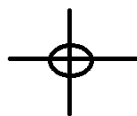
La grande découverte de Beuzée, c'est d'avoir vu que la parole est un acte et un événement singulier. Dans ce cas précis, on peut vraiment dire sans anachronisme que ce qu'il nomme « acte de la parole » coïncide très exactement avec ce que nous appelons « énonciation ».

Vuillaume explique en outre que c'est probablement grâce à l'enrichissement de la théorie de l'« acte de la parole » de Beuzée par la notion de subjectivité telle que l'a « découverte » la grammaire comparée allemande du 19^{ème} siècle que pourra naître la conception de la deixis centrée sur la personne du locuteur, si chère à tant d'auteurs à commencer par Bühler.

Ce sont donc ces deux notions – *acte de la parole* et *subjectivité* – qui ont favorisé le développement du fameux concept bühlerien d'« Origo », également dit « centre déictique », composé de trois éléments qui ressortissent à l'énonciation :

⁵ Benveniste décrit l'énonciation comme la « [...] mise en fonctionnement de la langue par un acte individuel d'utilisation. » (1974 : 80)

Soit un système de coordonnées composé de deux traits se croisant perpendiculairement avec O comme origine, la source des coordonnées.



J'affirme que les trois termes déictiques doivent être posés en O, si ce schéma est censé représenter le champ déictique du langage humain, en l'occurrence *ici, maintenant, je*. (Bühler, 2009 [1934] : 204)

La théorie bühliérienne d'*Ego-Hic-Nunc-Origo* reflète une vision égocentrée du langage. Elle fait, par ailleurs, directement écho au terme de « nynégocentrisme » que Damourette et Pichon (1936 : 12) emploieront à peine quelques années plus tard pour exprimer l'idée de la primauté des trois repères en question dans le fonctionnement du langage :

Le langage est naturellement centré sur le moi-ici-maintenant, c'est-à-dire sur la personne qui parle s'envisageant au moment même où elle parle ; c'est ce qu'on peut appeler le nynégocentrisme naturel du langage.

Le moment bühliérien a suscité le développement de tout un angle d'approche de la deixis, articulé autour des notions d'énonciation et d'énonciateur (éventuellement celle de co-énonciateur). Mais loin d'être unique, cet angle d'approche a été fortement critiqué pour ses points faibles, dont notamment sa trop grande insistance, justement, sur la présence d'un locuteur/énonciateur⁶. Or la « traque de l'énonciateur » ne saurait être pertinente dans le cas d'un énoncé tel que :

- 1) Assurance annulation. Cochez **ici** (formulaire d'inscription avec petit rectangle à droite)
(exemple employé par Kleiber lors de la conférence « Discours programmeurs et deixis », présentée le 20 octobre 2017 à l'ENS de Lyon)

De même, une théorie de la deixis qui se limite à cette vision des choses pourra difficilement rendre compte des emplois textuels (*endophoriques*) et mémoriels des déictiques (voir la Section 1.1.2 suivante). Aussi ce type d'approche est-il concurrencé par d'autres, dont notamment les approches « localisantes », « cognitives » et celles souscrivant à la théorie de la « *token-réflexivité* ».

1.1.2 Les approches « localisantes » de la deixis

Un autre grand angle d'approche de la deixis s'articule autour du site de présence du référent. Ses avantages résident notamment dans la possibilité d'opérer une séparation entre ce que nous appellerons dans ce travail les différents régimes de deixis/indexicalité. C'est sans doute chez Fraser et Joly (1979 et 1980) que l'on trouve, à cet égard, le modèle le plus exhaustivement élaboré. Pour

⁶ Ici, nous utiliserons ces deux termes comme synonymes, conscient que nous sommes par ailleurs de leur différence dans différents modèles de linguistique énonciative.

faire la généalogie de ce modèle, il est toutefois nécessaire de remonter à l'ouvrage *Cohesion in English* (1976) de Halliday et Hasan. Ces derniers, de même que, dans une certaine mesure, Fraser et Joly par la suite, restreignent le terme de « référence » à ce que Bühler désigne par le terme ici mentionné de « champ déictique ». Au sein de la référence entendue en ce sens, Halliday et Hasan font la distinction aujourd'hui très bien connue entre l'*endophore* et l'*exophore* (1976 : 33). Il y a endophore lorsque « [...] le référent se trouve dans l'espace textuel [...] » (Kleiber, 1991b : 3) et exophore lorsque le référent est fourni par la situation extralinguistique⁷. L'endophore se scindant ensuite en *anaphore* et *cataphore*⁸, il y a anaphore dans le cas d'une référence en amont et cataphore dans le cas d'une référence en aval. L'idée est bien saisissable intuitivement, mais demande une précision cruciale. Comme le remarque Kleiber, ce n'est évidemment pas le référent qui est physiquement logé dans l'espace textuel ; c'est l'identification de ce dernier qui s'appuie sur les informations apportées par le texte : « [...] une expression est anaphorique si son référent est (nécessairement) identifié par l'intermédiaire du contexte linguistique [...] » (Kleiber, 1991b : 9)⁹.

Cette définition « minimaliste » de l'anaphore¹⁰ parvient à rendre compte de toutes les expressions endophoriques, y compris de la cataphore intra-syntagmatique (à condition de remplacer *contexte* par *cotexte*), et de celles ressortissant à l'endophore non-coréférentielle :

- 2) Nous sommes arrivés dans une ville inconnue. **La mairie** était un somptueux édifice baroque.
- 3) Tu as aimé l'exposé de Paul ? J'ai préféré **celui de Patrick**.

L'énoncé 2 est un exemple d'anaphore associative. En 3, c'est la reprise du sens du terme exposé qui permet la bifurcation référentielle entre les deux expressions.

Avec Fraser et Joly, la typologie de Halliday et Hasan atteint un stade de symétrie plus avancé, lorsque la référence exophorique se scinde également en deux, entre référence « a-mémorielle » et « mémorielle » :

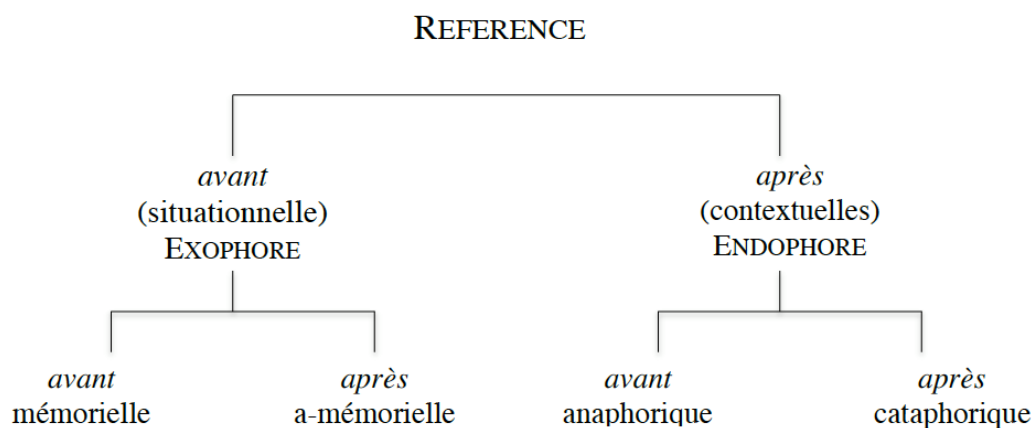
Nous nommerons *exophore* la référence situationnelle et *endophore* la référence contextuelle. Ces deux termes empruntés à Halliday et Hasan (1976, p.33) désignent de façon très économique l'opération par laquelle on renvoie (-*phore*) un signe à un référent extra-discursif (*exo-*) ou intra-discursif (*endo-*). Les deux auteurs cités, qui font une analyse intéressante, offrent sous forme de schéma une typologie générale de la référence qui nous paraît incomplète. Nous proposons donc la suivante :

⁷ « [...] linguistiquement, un objet peut avoir deux lieux d'existence : hors discours ou en discours [...] » (Fraser et Joly, 1980 : 24).

⁸ Rappelons que le terme a été forgé par Bühler.

⁹ Lorsque le référent est littéralement logé dans l'espace textuel, il ne s'agit pas d'endophore, mais de *deixis textuelle* (cf. J. Lyons, 1990 [1978] : 289-290).

¹⁰ L'*anaphore* est ici entendue, comme c'est souvent le cas, au sens générique de l'*endophore*. D'autres termes ont été proposés qui regroupent anaphore et cataphore, dont notamment « diaphore » (Maillard, 1974).



(Fraser et Joly, 1979 : 107)

L'on voit d'emblée en quoi réside la symétrie : l'opposition des deux types d'exophore reprend celle des deux types d'endophore en ce que l'exophore mémorielle réfère en amont et l'exophore a-mémorielle en aval. Ce parallèle est également corroboré par le fait que l'exophore mémorielle est souvent envisagée comme un type d'anaphore, où un laps de temps plus long sépare les expressions antécédente et anaphorique¹¹. Après avoir introduit leur schéma, Fraser et Joly se posent la question de la préséance entre les deux types de « référence » et en concluent à une double antériorité de l'exophore :

Halliday et Hasan (*op. cit.*, p. 32) se posent la question de savoir s'il faut placer l'exophore avant l'endophore et, si oui, quel serait le sens de cette antériorité. Antériorité historique ou antériorité logique ? À notre sens, les deux à la fois. Dans l'histoire particulière des langues que nous avons pu examiner, il semble que la deixis exophorique précède la deixis endophorique, celle-ci n'étant qu'un cas particulier de celle-là, dans la mesure où le « contexte » linguistique fait partie intégrante de la « situation », au sens large [...]. L'ostension contextuelle paraît donc devoir être interprétée comme une transformée de l'ostension situationnelle. (1979 : 107)

Le fait de poser l'antériorité de l'exophore par rapport à l'endophore a chez de nombreux auteurs pour conséquence de subsumer la seconde sous la première. Un argument justifiant cette décision est, comme nous pouvons le voir dans la citation, le fait que le discours fait lui aussi partie de la situation d'énonciation au sens large. Bühler parle déjà de l'anaphore comme d'un type de deixis et utilise parfois l'expression « deixis anaphorique » (cf. 2009 [1934] : 229). Il en va de même pour Fraser et Joly, qui, comme nous pouvons le constater, parlent tantôt de « deixis exophorique » et tantôt de « deixis endophorique ». Cette conclusion est également partagée par d'autres auteurs (cf. Ducrot et Todorov, 1972 ; Fillmore, 1982 ; J. Lyons, 1990 [1978]), et elle trouve globalement une justification en termes de transfert ou d'extrapolation du sens des déictiques considéré comme premier :

¹¹ Voir par exemple J. Lyons (1990 [1978] : 293-296) ou Berger (1993), parlant, lui, de « pseudo-anaphore ».

One extremely common extension of demonstrative categories is from spatial to textual identifications. Most commonly, perhaps, the Proximal category is used for text-referring cataphora, a non-Proximal category for text-referring anaphora. (Fillmore, 1982 : 53)

Une autre conclusion, extrême, qu'il est possible de tirer de l'opposition entre les régimes endo- et exophorique consiste à poser l'existence de deux paradigmes composés de formes – dont le démonstratif – pour la plupart identiques mais polysémiques, voire homonymiques. Une telle conclusion, si elle n'est pas indéfendable, défie à la fois l'étymologie, la sémantique et l'intuition :

Si l'on s'en tient à l'opposition déictique/anaphorique, il y aurait en fait deux adjectifs démonstratifs *ce*, un démonstratif qui fait partie de la classe des expressions déictiques comme *je*, *tu*, *ici*, etc., et un démonstratif qui appartiendrait à la catégorie des anaphoriques. Une telle position n'est théoriquement pas indéfendable, surtout du côté des philosophes-logiciens (D. Kaplan, 1977), qui n'ont aucun goût à loger les emplois anaphoriques à la même enseigne que les emplois situationnels, mais elle s'avère peu satisfaisante linguistiquement en ce qui concerne les démonstratifs. Il est en effet plus que difficile d'adhérer à une analyse éclatée du sens de l'adjectif démonstratif. Une expression comme *Cet homme* ne semble pas présenter, intuitivement, deux sens différents. (Kleiber, 1986d : 6)

Pour ce qui est de l'« exophore mémorielle », l'on sait que ce terme fait déjà son apparition chez Bühler, lequel parle de « Deixis am Phantasma » (« deixis à l'imaginaire »), par opposition à la « demonstratio od oculos » (2009 : 226-248). Elle a ceci de particulier qu'elle sollicite chez l'interlocuteur une expérience et des connaissances partagées avec le locuteur. Sont exigées chez lui une attitude participative et la mobilisation d'un effort actif de compréhension :

[...] le référent visé n'est pas « présent » (directement ou indirectement) dans la situation d'énonciation : il ne s'agit donc pas d'un emploi situationnel ; il n'est pas non plus « présent » dans le contexte antérieur immédiat : il ne s'agit donc pas d'un emploi anaphorique. [...] si le référent n'est présent ni dans la situation immédiate d'énonciation ni dans le contexte linguistique, il est présent dans la conscience, dans l'imagination.

(Kleiber, 2007 : 373 et 378)

Les approches « localisantes » possèdent avant tout l'avantage de présenter le phénomène de la deixis sous la forme d'un système symétrique. Toutefois, elles comptent aussi des faiblesses considérables, dont le risque, mentionné ci-dessus, d'un clivage des paradigmes. On serait tenté de conclure, du moins à première vue, que la majeure partie des limites de cet angle d'approche résident dans sa nature onomasiologique : la présence du référent étant le critère absolu, le sémantisme inhérent des expressions n'est pas suffisamment pris en considération. Ainsi, à suivre ce modèle au pied de la lettre, nous serons amené à dire que le pronom *je* dans l'énoncé 4 *infra* est anaphorique, car son interprétation référentielle s'effectue moyennant le contexte amont. D'où l'importance, immédiatement perçue par Kleiber, de ne jamais perdre de vue ce sémantisme :

4) Alors [Paul]_i s'est exclamé : « [*Je*]_i ne suis pas une anaphore ! » (Kleiber, 1991b : 5)

Il est tout aussi clair que si l'on veut éviter d'en arriver là, il faut faire intervenir autre chose que le seul critère onomasiologique de localisation que nous avons utilisé jusqu'à présent ; il faut se tourner vers le sens des expressions elles-mêmes pour y trouver l'élément susceptible de retirer un déictique comme *je* du giron des anaphoriques [...]. (Kleiber, 1991b : 5)

Une autre erreur à laquelle peut aboutir l'orientation onomasiologique des théories « localisantes » concerne le traitement de l'article défini et des descriptions définies de manière générale. Il est un fait incontestable que l'article défini connaît d'abondants emplois exophoriques et endophoriques¹². D'autre part, il est aberrant d'attribuer à cette unité linguistique un sémantisme déictique a priori, son rôle premier étant d'exprimer l'unicité dénotative¹³ du concept représenté par le nom qu'il introduit, et ce même en dehors des emplois indexicaux. Halliday et Hasan écrivent à ce propos : « [...] its function is to signal definiteness, without itself contributing to the definition [...] (1976 : 74) ». Cela explique d'ailleurs pourquoi l'existence de l'article défini est facultative dans les langues du monde. En ce qui concerne la cataphore intra-syntagmatique¹⁴, type d'emploi endophorique auquel le défini se prête souvent, Halliday et Hasan (1976 : 72) écrivent ceci :

Cataphoric or forward reference with *the* is limited to the structural type. Unlike the selective demonstratives [...] *the* can never refer forward cohesively. It can only refer to a modifying element within the same nominal group as itself.

The ascent of Mount Everest

The longest street

The best way to achieve stability

Mais dans ces emplois tout comme dans les autres emplois indexicaux, le défini n'est évidemment pas doté d'un *sens* indexical.

1.1.3 Les approches fondées sur l'accessibilité référentielle

C'est également pour échapper au formalisme onomasiologique des approches « localisantes » qu'une tradition désormais bien établie de la deixis prend pour point de départ la notion d'accessibilité/saillance référentielle. Il s'agit d'un ensemble d'approches – aussi appelées, entre autres, « mémorielles » et « cognitives » – orientées sur la personne de l'interlocuteur. C'est le degré de familiarité que celui-ci possède avec le référent qui en constitue la pierre angulaire :

¹² C'est dans ces emplois – ressortissant à la définitude dite « pragmatique » – qu'il est apparu d'abord, au début de son évolution, avant de s'étendre progressivement aux contextes de définitude « sémantique » (cf. Löbner, 1985 et 2011 ; Himmelmann, 1997 et 2001 ; Carlier et De Mulder, 2006, 2010 et 2011 ; Ortmann, 2014 ; voir le Chapitre 6 pour la situation en tchèque oral informel contemporain).

¹³ Le terme de dénotation est ici employé dans son acception extensionnelle (acception qui a cours en logique) et non pas dans son acception intensionnelle (qui a davantage cours en linguistique).

¹⁴ « Cataphore intra-syntagmatique » (Kesik, 1989 : 18 et 22 ; Fraser et Joly, 1979 : 48), aussi appelée « structural cataphora » (Halliday et Hasan, 1976 : 72) ou « authophoric uniqueness » (Ortmann, 2014 : 302).

Le choix des expressions référentielles se trouve ainsi crucialement lié aux présuppositions du locuteur sur la récupérabilité par l'interlocuteur du référent visé [...]. (Kleiber, 1991b : 10)

Le mode de donation du référent – situationnel, textuel ou mémoriel – fait ainsi place à la distinction entre le nouveau et le déjà connu. Dans cette perspective, le démonstratif serait toujours déictique, même lorsqu'il est employé textuellement ou mémoriellement, car dans tous ces emplois, son rôle consisterait à (re)placer le référent dans le foyer d'attention de l'interlocuteur, à le rendre accessible et saillant :

La deixis sert prototypiquement à orienter le foyer d'attention de l'allocutaire vers un nouvel objet de discours (ou vers un nouvel aspect d'un objet de discours déjà existant au moment de la prise de parole) qui est à construire mentalement, par défaut à partir de la situation d'énonciation [...]. L'emploi de la procédure référentielle déictique implique toujours une rupture, fût-elle brève, dans la continuité du discours jusqu'au moment de son occurrence [...]. (Cornish, 2010 : 114)

La notion de rupture est d'une grande importance pour le démonstratif. En effet, comme ce dernier identifie son référent moyennant un recours au contexte le plus immédiat, il introduit une coupure avec le contexte plus large, minant ainsi, paradoxalement, la continuité discursive tout en l'assurant. Dans le pôle opposé au démonstratif, nous retrouvons l'article défini et le pronom personnel de la 3^{ème} personne, expressions par définition anaphoriques car reprenant des référents déjà accessibles et qui doivent le rester :

Quant à l'anaphore, l'occurrence d'un anaphorique conjointement avec la proposition dans laquelle il apparaît dans son ensemble, constitue un signal ou une instruction de maintenir le foyer d'attention déjà établi (ou présumé déjà établi) au moment de la parole – ainsi, l'état du modèle de discours tel qu'il existe au moment de l'emploi de l'anaphorique : l'anaphore signale qu'il ne doit pas y avoir de changement à cet égard, donc que ce serait « Business as usual », en somme – au sens : « pas d'introduction d'un nouveau référent, ou d'un nouveau foyer d'attention » [...]. (Cornish, 2010 : 115)

En plus d'échapper au formalisme purement linguistique dont souffrent les approches « localisantes » et aux risques qui en découlent, les approches mémorielles ont pour vertu de tenir compte de réalités psychologiques et psycho-linguistiques. Elles présentent aussi l'avantage d'une analyse beaucoup plus unifiée des expressions : comme le démonstratif est toujours tenu pour déictique, par exemple, on n'a pas à se confronter au problème de ses emplois textuels et à leur rapport aux emplois situationnels. S'agissant plus particulièrement du démonstratif, la formulation de Cornish citée *supra* et selon laquelle ce dernier peut également mettre en relief « [...] un nouvel aspect d'un objet de discours déjà existant au moment de la prise de parole [...] » s'avère un moyen explicatif particulièrement puissant pour les cas où le démonstratif est employé dans un but de reclassification du référent, y compris dans ce que les auteurs appellent l'« anaphore résomptive » (cf. Kara et Wiederspiel, 2011) ou la « deixis discursive » (cf. Himmelmann, 1996 ; Guillot, 2006a) :

- 5) Hier, je me suis fait voler ma voiture. *Cette sale histoire* m'a empêché de dormir.

Toutefois, outre ces avantages apparents, Kleiber pointe également une série d'inconvénients non-négligeables. Car si l'approche mémorielle de la deixis s'avère, dans un premier temps, unificatrice pour les paradigmes des expressions, elle y entraîne aussi des divisions. Ainsi, le pronom IL, par exemple, n'est pas porteur du même degré d'accessibilité référentielle selon qu'il est employé ana- ou cataphoriquement. Dans une phrase comme

- 6) Quand [*il*]_i est entré, [Paul]_i enleva son chapeau. (Kleiber, 1991b : 3)

IL serait – paradoxalement – susceptible de passer du côté des déictiques, car il annonce un référent à venir et qui, par conséquent, n'est pas encore accessible. Un autre problème surgit par exemple avec le nom propre, expression qui ne se présente ni comme tout à fait accessible ni comme tout à fait nouvelle (inaccessible) sur le plan référentiel, et dont le mode de donation du référent ne dépend pas directement de la situation d'énonciation¹⁵.

En réalité, l'accessibilité référentielle est une affaire de progression, que l'on peut représenter sur une échelle. Nous donnons ici pour exemple celle de Cornish (2010 : 121) :

Deixis pure > **Deixis de discours** > **Deixis de « reconnaissance »** > **Anadeixis 'stricte'** > Anaphore pure
 <-----**anadeixis**----->

Nous pouvons donc voir qu'il existe une zone de transition entre deixis et anaphore définies en termes d'accessibilité référentielle, zone de transition pour laquelle Cornish emploie le terme d'« anadeixis »¹⁶:

L'« anadeixis » correspond au type de référence indexicale qui combine les procédures anaphorique et déictique à des degrés différents. Autrement dit, les expressions indexicales qui la réalisent sont anaphoriques dans la mesure où leur référent est déjà (potentiellement du moins) présent dans la représentation discursive supposée partagée par locuteur et allocutaire au moment de son emploi ; et ce référent est récupéré par ce biais. Cependant, ce référent est moins que hautement saillant au moment du renvoi, à la différence de la situation qui prévaut avec l'anaphore canonique (il est relativement peu saillant par rapport à d'autres référents qui

¹⁵ « En effet, au titre de désignateur rigide, il [le nom propre] constitue une procédure d'étiquetage univoque et constante au sens où il n'émane pas, contrairement à d'autres unités désignatives, de situations affectant, plus ou moins décisivement, le référent. Par ailleurs, au plan instructionnel, le Np a la capacité tout à fait remarquable de débrayer et de remettre les compteurs référentiels à zéro. » (Schnedecker, 1997 : 209)

¹⁶ La définition de la « deixis de discours » (« deixis discursive ») a été rappelée *supra*. La deixis « de reconnaissance » (Himmelman (1996 : 47) parle de « recognitional uses ») correspond approximativement à l'« exophore mémorielle » de Fraser et Joly et l'« anadeixis 'stricte' », elle, « [...] consiste en la reprise par des démonstratifs (pronoms et SN), et parfois même par des SN définis dans certaines circonstances, d'un référent déjà évoqué dans le discours environnant, mais qui n'est que peu ou pas topical au moment où la récupération du référent est à faire ; ou bien là où il y a 'compétition' entre plus d'un référent d'un même type donné [...] ; ou enfin où un référent de discours macro-topical vient d'être introduit, et doit être pleinement installé dans l'empan de mémoire à court terme de l'allocutaire. Dans aucun de ces cas de figure un pronom de 3^{ème} personne inaccentué ne serait à même d'assurer avec naturel le maintien pur et simple du référent visé – sa vocation indexicale étant canoniquement anaphorique [...] » (Cornish, 2010 : 123).

viennent d'être évoqués). En même temps, une référence anadéictique est déictique dans la mesure où le locuteur a recours au contexte d'énonciation afin de réorienter l'attention de son allocutaire vers un référent qui, bien que potentiellement disponible au sein du contexte discursif au moment de l'énonciation, n'est pas celui auquel on s'attendrait pour une référence ultérieure à ce moment-là du discours. (Cornish, 2010 : 117)

Dans cette optique, la deixis de discours est moins déictique que la deixis pure, tout en restant nettement plus proche de la deixis que de l'anaphore, et ainsi de suite.

Un sous-ensemble des approches de ce type, représentées par exemple par Yule (1981), Ariel (1988) et Givón (1983), se consacrent avant tout au domaine de la référence anaphorique. Ainsi, leur principale préoccupation est la construction d'un modèle théorique susceptible de rendre compte des conditions d'emploi de différents types d'expressions anaphoriques. Le modèle d'accessibilité référentielle élaboré par Ariel (1988) tient compte de quatre facteurs : a) la distance entre la reprise anaphorique du référent et sa dernière mention dans le discours antérieur (« distance anaphorique »), b) le nombre de candidats potentiels à la référence anaphorique (le nombre d'expressions éligibles pour la fonction d'antécédent), c) la topicalité (la saillance) de l'anaphorique (son statut thématique/rhématique) et d) les « cadres » ("frames") d'identification du référent. Pour ce dernier facteur, Ariel distingue entre les connaissances générales ("general knowledge"), la situation physique ("physical surroundings") et le matériel linguistique mentionné auparavant ("previous linguistic material"). En partant de ces facteurs, elle établit trois degrés d'accessibilité référentielle, faisant usage de trois types d'expressions référentielles : l'accessibilité basse, moyenne et haute :

[...] the degree of Accessibility of the marker used should correspond to the Accessibility of the referent it is used to refer to. The less accessible the representation of a given entity, the lower the Accessibility marker used. (1988 : 75-76)

Ces marqueurs sont, respectivement, le pronom personnel de la 3^{ème} personne pour l'accessibilité haute, le démonstratif pour l'accessibilité moyenne et l'article défini et le nom propre pour l'accessibilité basse. Ariel (1988 : 80) précise que

[l]ow Accessibility marked entities are often those stored in long-term memory, while High Accessibility marked entities are normally those held in short-term memory.

Concrètement, cela veut dire que globalement, les référents dotés d'une accessibilité haute sont identifiés via un contexte (discursif ou situationnel) immédiat, tandis que ceux dotés d'une accessibilité basse le sont via une connaissance à long terme, indépendante du contexte immédiat. Le principal point faible de cette approche nous semble cependant résider dans le fait qu'Ariel assume automatiquement que le défini marque, dans toutes les circonstances, une accessibilité basse. Cette conclusion est par exemple en nette contradiction avec le modèle de Cornish (2010) présenté *supra*,

où le défini est toujours plus « anaphorique » (donc introduisant un référent plus accessible) que le démonstratif. En l'occurrence, ce paradoxe est dû à la fois à une conception en partie différente que les deux auteurs ont de l'accessibilité référentielle et au fait que ni l'un ni l'autre n'analysent de manière assez conséquente les différents emplois auxquels le défini se prête. En réalité, une différence majeure existe entre une description définie référant à une entité « sémantiquement » définie comme *le Président* et une description définie reprenant un référent défini « pragmatique », car établi comme accessible par le discours antérieur (typiquement, via son introduction par une expression indéfinie en première mention et sa reprise, en deuxième mention, par une expression démonstrative). Dans le premier cas, nous pouvons admettre avec Ariel que le référent introduit par le défini ressortit à une accessibilité (plutôt) basse, son identification reposant sur la mémoire à long terme (“long-term memory” ; Ariel, 1988 : 80) des membres d'une communauté linguistique, alors que dans le second cas, le défini marque davantage une accessibilité haute, car résultant d'une apparition répétée du référent dans le discours. Comme le résume Diessel (1999 : 98) tout en se rapportant à des travaux existants :

Himmelmann (1996: 229) points out that the use of anaphoric demonstratives after the first mention of a new discourse participant is especially common in languages that do not have a definite article; however, even in languages that employ a definite article, demonstratives are often preferred in this position (cf. Christophersen 1939: 29). Once a new discourse participant has been established as topic, it is usually tracked by third person pronouns, zero anaphors, **definite articles**¹⁷, or pronominal affixes on the verb; but when a referent is mentioned for the second time, demonstratives are often the most common tracking device.

Ces deux partis-pris opposés au sujet du statut des descriptions définies constituent une autre preuve de la faiblesse des approches reposant sur la notion d'accessibilité référentielle. En effet, malgré leur apport incontestable¹⁸, la tentative de proposer un modèle explicatif puissant, permettant de rendre compte d'une problématique aussi complexe que la distribution complémentaire des expressions référentielles indexicales, n'est pas à l'abri de simplifications, de raccourcis, voire de contradictions.

Givón (1983) est lui aussi l'auteur d'une échelle d'accessibilité référentielle¹⁹. Au sommet de celle-ci, l'on trouve l'anaphore zéro (l'élément anaphorique est sous-entendu ; Ø), suivie, dans l'ordre d'une accessibilité décroissante : des pronoms (personnels) atones (clitiques) et de l'accord grammatical, des pronoms accentués, des descriptions définies en dislocation à droite, des descriptions définies (hors dislocations), des descriptions définies en dislocation à gauche, des

¹⁷ Nous soulignons.

¹⁸ Nous nous y référerons lorsque nous traiterons du démonstratif adnominal en emploi anaphorique, notamment dans le Chapitre 6.

¹⁹ Chose curieuse : à aucun moment, cette échelle ne fait mention du démonstratif adnominal.

constructions clivées et des SN indéfinis²⁰. Pour Givón, l’accessibilité du référent est le résultat de trois facteurs, à savoir : a) la distance anaphorique, b) l’ambiguïté potentielle entre le référent en question et d’autres référents et c) la « persistance du référent » (“decay”), à savoir le nombre de fois où celui-ci apparaîtra dans le discours subséquent. Deux différences principales apparaissent ainsi par rapport au modèle d’Ariel présenté ci-dessus : si Givón ne s’intéresse pas explicitement au cadre permettant l’identification du référent, il tient néanmoins compte de sa place dans le discours ultérieur.

Yule (1981) opère la différence entre référents « donnés » (“given”) et « nouveaux » (“new”)²¹. Au sein des premiers, il distingue ce qu’il appelle les référents « actuels » (“current”) et les référents « déplacés » (“displaced”). Le premier terme renvoie aux référents rendus récemment accessibles par le discours ou à ceux dont l’accessibilité est maintenue grâce à des reprises se succédant à peu de temps d’intervalle ; le second, lui, renvoie aux référents rendus accessibles moins récemment²².

De manière générale, les modèles proposés par ces trois auteurs semblent avoir pour point commun le constat que les formes de reprise de référents tendent à s’atténuer dans leur matérialité linguistique : un référent introduit en première mention le sera moyennant une expression plus étoffée (*a red horizontal line of about two inches* ; Yule, 1981 : 49), contrairement à sa reprise, recourant à une expression moins étoffée (*the line/it/Ø* ; Yule, 1981 : 49). Ceci dit, il faut immédiatement préciser que la tendance est en partie réversible, car les référents établis dont il n’a pas été question depuis un certain temps (« displaced » chez Yule, 1981) tendent vers une reprise moyennant des expressions linguistiquement plus étoffées que ceux qui viennent juste d’être rendus accessibles ou ceux dont l’accessibilité est constamment maintenue par le discours²³. Ainsi, le référent de *a red horizontal line of about two inches* déjà repris par *it* et *Ø* mais sorti du discours y sera de nouveau réintroduit par *the (red) line*, etc.

²⁰ Il faut cependant préciser qu’il parle le plus souvent de « continuité thématique » (“thematic continuity”), et que son modèle se base avant tout sur les contraintes du seul discours, tenant ainsi moins compte de facteurs de l’ordre cognitif et psychologique.

²¹ Contrairement à Ariel et à Givón, Yule procède à la mise en place d’un cadre expérimental n’admettant que la référence anaphorique et en excluant tout renvoi au contexte d’énonciation ou à des connaissances partagées des participants.

²² Nous retrouvons une typologie en partie similaire chez Laury (1997). Cette dernière emploie les termes de « donné » (“given”) et « accessible » pour, respectivement, les référents « actuels » et « déplacés » de Yule. Mais contrairement à Yule, elle emploie le terme de « nouveau » pour renvoyer à tous les référents en première mention (= nouveaux du point de vue du seul discours), y compris donc les référents identifiables indépendamment du discours (le plus souvent sur la base d’une connaissance mémorielle).

²³ “Thus the degree of attenuation increases after the first mention of the new entity, but it decreases after a while.” (Ariel, 1988 : 72)

1.1.4 Les approches fondées sur la « *token-réflexivité* »

Une autre conception de la deixis suit la direction indiquée par l'ouvrage *Elements of symbolic logic* (1966) de Reichenbach. L'idée au cœur de ces approches est celle que l'identification du référent d'une expression déictique s'effectue par le biais de l'occurrence (= le *token*) de cette expression :

To understand the function of these words we have to make use of the distinction between token and symbol, 'token' meaning the individual sign, and 'symbol' meaning the class of similar tokens [...]. Words and sentences are symbols. The words under consideration are words which refer to the corresponding token used in an individual act of speech, or writing; they may therefore be called token-reflexive words.

(Reichenbach, 1966 : 284)

Comme il s'agit d'expressions référentielles procédant à l'indexation du référent sur le contexte de leur propre occurrence, certains auteurs parlent également d'« indexicaux » ou de « symboles/expressions indexicaux/les » (cf. Burks : 1948 ; Bar-Hilley : 1954 ; Kleiber : 1984a, 1986a, 1986d). Dans l'espace francophone, cette approche est reprise par Vuillaume et Kleiber. Les deux auteurs se posent la question de savoir quelle est la nature exacte du mécanisme qui relie l'occurrence de l'expression en question à son référent, et en viennent à conclure que ce mécanisme est de deux sortes : ou le chemin entre l'occurrence et le référent va de soi, ou une opération d'inférence est nécessaire. Vuillaume (1980) appelle le premier cas de figure « deixis directe » et le second, « deixis indirecte ». Kleiber (1986d), lui, parle de « token-réflexivité transparente » et « opaque ». Les pronoms personnels JE et TU, tout comme les adverbes *aujourd'hui*, *hier* ou *maintenant*, sont des « token-réflexifs transparents ». Les autres « token-réflexifs » – dont notamment les expressions démonstratives – seraient par définition indirects/opaque.

Tout en décrivant le lien occurrence-référent comme un lien de contiguïté spatio-temporelle, Kleiber dit à propos des « token-réflexifs transparents » et « opaques » :

Les déictiques sont des expressions qui renvoient à un référent dont l'identification est à opérer nécessairement au moyen de l'entourage spatio-temporel de leur occurrence. La spécificité du sens indexical est de « donner » le référent par le truchement de ce contexte. (1986d : 19)

Par opposition aux symboles indexicaux comme *Je*, *Tu*, *maintenant*, *aujourd'hui*, etc., qu'on peut appeler symboles indexicaux **transparents** ou **complets**, parce que leur sens est tel qu'il détermine a priori le type de référent dénoté, les démonstratifs se présentent comme des symboles indexicaux **opaques** ou **incomplets**, en ce qu'ils sont susceptibles de plusieurs types d'emplois référentiels possibles. Les symboles indexicaux transparents ne connaissent qu'un type d'emploi situationnel possible. *Je*, *Tu*, *aujourd'hui* ne peuvent s'employer respectivement que pour la personne qui dit *je*, la personne à laquelle le locuteur dit *tu* et le jour où est prononcé *aujourd'hui*. Les symboles indexicaux opaques, comme les démonstratifs par exemple, nécessitent des informations supplémentaires (soit geste, soit contexte, etc.) : en dehors de toute autre indication, l'énonciation de *ce chien/cela est beau* ne nous apprend pas de quel référent il s'agit, ni ne nous renseigne sur l'endroit précis où il se trouve, la description démonstrative et le pronom démonstratif étant susceptibles d'emplois gestuel, symbolique et anaphorique. (1984a : 67)

Ainsi, les indexicaux connaissant des emplois uniquement déictiques sont transparents, alors que les indexicaux connaissant des emplois aussi bien déictiques qu'anaphoriques sont opaques. Simultanément à cette première division, Kleiber opère une autre distinction, plus poussée, selon que l'appariement occurrence-référent est « immédiat » ou « différé » (1986d : 19-20). Pour les indexicaux transparents, l'affaire est simple, car ladite opération est toujours immédiate. En revanche, au sein des indexicaux opaques (donc, typiquement, des démonstratifs), les deux cas de figure peuvent se présenter. La ligne de séparation est tracée entre, d'une part, les emplois situationnels dits « symboliques »²⁴ et, d'autre part, les autres emplois situationnels, les emplois endophoriques et les emplois mémoriels. Dans le premier cas, l'appariement référentiel entre l'occurrence de l'expression indexicale et son référent se fait sans intermédiaire. Dans le second cas, ce même processus passe obligatoirement par l'intermédiaire d'une autre instance, que celle-ci soit un geste d'ostension, une autre expression référentielle présente dans le discours environnant ou une connaissance mémorielle.

Sans rentrer dans le détail des débats entre les partisans de cette approche, nous trouvons que c'est de loin la plus solide, car débarrassée à la fois de l'écueil « énonciativiste », exigeant à tout prix la présence d'un locuteur-centre déictique, des écueils « localisants », dont notamment l'excès d'orientation onomasiologique avec tous les inconvénients qui en résultent, mais aussi des impasses des approches basées sur l'accessibilité référentielle que nous venons de voir. Sous-tendue par le critère sémasiologique où la prime importance est donnée au sens des expressions elles-mêmes, la « *token-réflexivité* » ne débouche pas sur un clivage paradigmatique artificiel en termes de polysémie/homonymie et elle réconcilie plus qu'elle n'oppose les emplois premiers (en l'occurrence situationnels) et « dérivés » des déictiques :

[...] la filière définitoire indexicale des déictiques, où la notion de deixis ne s'oppose plus directement à celle d'anaphore, mais s'établit, à partir des cas paradigmatiques comme *je*, *maintenant* et *ici*, comme un mode de référence original, l'objectif étant de définir la spécificité du fonctionnement référentiel des marqueurs indexicaux [...]. (Kleiber, 1991b : 12)

Ainsi, pour conclure à propos de cette première section, nous ferons usage des termes *deixis*, *indexicalité*, éventuellement aussi *token-réflexivité*, que nous emploierons comme des termes génériques, dans une acception qui recouvre celle de « référence » chez Halliday et Hasan (1978)²⁵. Si nous laissons de côté la question de savoir s'il faut voir dans l'endophore et les emplois mémoriels des sous-catégories de la deixis situationnelle (de l'« exophore a-mémorielle » de Fraser et Joly, 1979 et 1980) issues d'un transfert en diachronie plutôt que des catégories à part et situées sur un pied

²⁴ Un exemple de ces emplois : *Il fait chaud dans cette voiture !* prononcé par quelqu'un se trouvant à l'intérieur de la voiture en question.

²⁵ Nous sommes conscient que cette acception ne correspond pas tout à fait à celle de Kleiber (cf. 1986d, 1991b).

d'égalité avec celle-ci²⁶, nous veillerons à faire la distinction entre sens et emplois des unités. Le démonstratif, quant à lui, est l'un des représentants emblématiques de la classe des déictiques en ce que son sens est intrinsèquement indexical. Ensuite, nous parlerons de « régimes de deixis/indexicalité » pour les différents emplois – situationnels (éventuellement aussi temporels), discursifs/textuels et mémoriels – du démonstratif.

1.2 Le démonstratif : singularités sémantico-référentielles

Sera ici rappelé le comportement sémantico-référentiel spécifique au démonstratif et, plus précisément encore, au démonstratif nominal (Dixon, 2003) en emploi adnominal. Nous commencerons par traiter de la place du démonstratif parmi les expressions ressortissant à la définitude. À dessein de mieux illustrer la spécificité du démonstratif, nous opposerons son fonctionnement à celui de l'article défini. Nous présenterons la thèse de Kleiber selon laquelle un syntagme nominal (SN) démonstratif en français constitue une prédication classificatoire partitive en structure profonde. Après cela, nous reprendrons pour l'approfondir la thèse du démonstratif « indexical opaque » et nous verrons, à ce moment-là, plus en détail les différents emplois possibles du démonstratif ainsi que la relation entre ce dernier et le geste qui l'accompagne dans certains de ses emplois.

1.2.1 Le démonstratif : un *désignateur* contextuel

Quelle est la nature du démonstratif et comment réfère-t-il ? Le démonstratif est couramment rangé dans la classe des expressions linguistiques définies, expressions qui ont bénéficié d'un intérêt abondant non seulement chez les linguistes mais aussi chez les logiciens et les philosophes du langage. Kleiber (1984a : 63-64) résume à propos de cette classe :

Comme on sait, cette classe d'expressions, qui regroupe également descriptions définies, descriptions possessives, noms propres, pronoms personnels et pronoms démonstratifs, se laisse déterminer par des propriétés logiques – ces expressions correspondent à l'argument, dans la mesure où elles prennent la place de la variable individuelle x de la forme logique $F(x)$ –, syntaxiques – elles sont, par exemple, incompatibles avec l'impersonnel –, sémantiques – elles sont compatibles, dans une interprétation spécifique, avec des prédicats non spécifiques –, pragmatiques enfin – elles permettent toutes d'effectuer un acte de référence défini unique.

Quant aux expressions possédant le trait *définitude* (+ Défini), les logiciens utilisent pour elles l'appellation de « termes singuliers » (Kleiber, 1984a : 63). Certains linguistes, tel Corblin (1987),

²⁶ Outre que cette question n'a pas trop de pertinence pour notre travail, elle est, comme nous semble bien le montrer Kleiber (1991b), mal posée du fait de son orientation onomasiologique : en réalité, la réponse dépendra des unités appréhendées (cf. le démonstratif vs le pronom de la 3^{ème} personne).

emploient, à l'instar de Geach (1962), celle de « désignateur »²⁷. Pourtant, l'attribution inconditionnée du trait (+ Défini) au démonstratif suscite des réactions critiques d'une partie des chercheurs, le démonstratif partageant clairement des propriétés avec l'indéfini :

Bien que les grammairiens le [= le démonstratif] rangent [...] dans la catégorie des déterminants définis, avec lesquels il partage en effets certaines propriétés, d'autres propriétés le rattachent par contre plutôt à l'article indéfini. (Gary-Prieur, 2011 : 64)

Les exemples suivants font ressortir cette proximité partielle avec l'indéfini :

7) a) Je voudrais *le gâteau*.

b) Je voudrais *ce gâteau*²⁸. (Gary-Prieur, 2011 : 66 ; modifié)

Alors que *le gâteau* « [...] réfère, dans la situation, à l'unique objet correspondant au sens du nom *gâteau* [...] », *ce gâteau* « [...] réfère, parmi tous les objets de la situation, à celui que le locuteur distingue des autres. » (Gary-Prieur, 2011 : 67). C'est la raison pour laquelle 8a est parfaitement acceptable, contrairement à 8b :

8) a) Je voudrais *ce gâteau*, et aussi *ce gâteau*.

b) *Je voudrais *le gâteau*, et aussi *le gâteau*. (Gary-Prieur, 2011 : 66)

Au vu de ce constat, il semble évident que le présupposé (éventuellement, chez Russell (1905), l'implication²⁹) d'unicité³⁰ du référent, traditionnellement associé aux descriptions définies³¹, doit être relativisé ou entendu dans un sens différent pour les expressions démonstratives. En effet, les énoncés de 8a et 8b montrent que contrairement à l'article défini, le démonstratif adnominal « [...] n'implique pas que l'entité désignée est seule de son espèce dans une situation donnée [...] ». En

²⁷ Geach (1962) répartit les SN dans les deux catégories suivantes : « expressions indéfinies » et « désignateurs » ; dans les premières, il range les SN à article indéfini et, dans les secondes, les descriptions définies, démonstratives, possessives et les noms propres. D'autres taxinomies sont également possibles, comme celle de Kripke (1980), distinguant entre le nom propre d'une part et les descriptions définies d'autre part, catégorie où il range aussi le démonstratif.

²⁸ Vu qu'il est pour l'instant question d'une introduction théorique au problème, et que ce travail s'adresse *a priori* à un lecteur francophone, la plupart des exemples introduits dans ce chapitre seront des exemples du français. Ce n'est que dans les chapitres suivants que nous ferons à proprement parler entrer le tchèque.

²⁹ Entendue au sens fort, de l'assertion.

³⁰ Pour les référents pluriels et indéénombrables, il convient davantage d'utiliser d'autres termes, tels que « inclusivité », « totalité » ou « maximalité » (cf. Hawkins, 1978 ; Kleiber, 1981a et C. Lyons, 1999).

³¹ Voir notamment Russell (1905), Strawson (1950), Donnellan (1966), ainsi que Kleiber (1981a, 1984) ou Prince (1981 : 244-245) : « As is well known, a presuppositional definite NP often conveys an understanding that there exists just one entity describable by the NP. » Nous laissons ici de côté la nature présupposée ou impliquée (assertée) de cette unicité induite par le défini. Une précision importante s'impose par ailleurs : alors que la détermination et la définitude semblent représenter des phénomènes sémantiques universels (voir par exemple Krámský, 1972), la situation est plus délicate dans les langues comme le tchèque, où la définitude n'est pas marquée systématiquement, moyennant un article défini. Ainsi, en l'absence de marqueurs explicites de définitude (démonstratifs) ou d'indéfinitude (quantificateurs), le statut défini ou indéfini d'un complexe nominal s'avère être le fruit d'un processus d'inférence complexe, où d'autres facteurs (tels que, typiquement, la structure informationnelle) jouent un rôle important (cf. Krámský, 1972 ou Šimík et Demian, 2020).

revanche, il « [...] permet d'établir un contraste entre le référent du GN et d'autres membres de la catégorie N : ce N, c'est un N que le locuteur distingue des autres N. » (Gay-Prieur, 2011 : 66-67) Or cette possibilité de référer, dans une situation donnée, à plusieurs membres d'une classe référentielle dénotée par le sens descriptif d'un nom, le démonstratif la partage justement avec l'indéfini :

9) Ça, c'est *un gâteau* et ça, c'est aussi *un gâteau*.

Himmelman, s'appuyant sur la position prise par Hawkins dans ses travaux de 1978 et de 1991, remet lui aussi en question ce présupposé en rapport avec le démonstratif :

Für Demonstrativa gilt keine Einzigkeitspräsupposition (so Hawkins 1978:156) oder aber sie ist von anderer Art als die Einzigkeitspräsupposition für Definitartikel (so Hawkins 1991:415). Im letzten Falle ist dieses Konzept so auszudehnen, dass damit die für den Gebrauch aller Arten von definiten Ausdrücken charakteristische Präsupposition, dass der Referent für H relativ eindeutig zu identifizieren ist, erfaßt wird³².

Ainsi, selon Himmelmann, deux positions peuvent être adoptées : a) abandonner le présupposé d'unicité pour les expressions démonstratives tout court ou b) reconnaître que ce dernier est, dans ce cas, différent de l'article défini. Ce que l'auteur suggère par ailleurs, cependant, c'est que cette dernière position doit nécessairement aboutir à une fusion du présupposé d'unicité et du présupposé d'identification du référent par l'interlocuteur. Nous ne savons pas si une telle position serait acceptable. Chez Hawkins, elle a d'ailleurs évolué depuis son travail publié en 1978. Ainsi, la nouvelle position d'Hawkins (contrairement à ce que laisse paraître la citation de Himmelmann *supra*), explicitée dans son article de 1991, postule ouvertement pour le démonstratif l'existence d'un présupposé d'unicité distinct dudit présupposé d'identification :

Hawkins (1978) took the position that demonstrative expressions such as *that bucket* were not necessarily uniquely referring and were subject only to a pragmatic condition of referent identifiability. I now believe that it is desirable to extend the uniqueness generalization to all definite NPs, thereby achieving a general definition of definiteness, and to view the differences between subtypes of definite NPs as resulting from different realizations of this common generalization in language performance. (1991 : 416)

Löbner (1985 et 2011) va dans le sens de cette thèse tout en montrant qu'il semble justifiable de traiter les démonstratifs avec les descriptions définies dans la mesure où ils se plient – à condition d'une précision importante – au test de la négation. Nous reproduisons ici la définition de ce test ayant pour but de dissocier les expressions qui portent le trait (+ Défini) de celles qui ne le portent pas³³:

³² La même position est prise par C. Lyons, qui remplace le terme d'« unicité » par celui, plus universel, d'« inclusivité » : "Demonstratives are generally considered to be definite, but it is clear that their definiteness is not a matter of inclusiveness." (1999 : 17).

³³ Parmi celles-ci, l'on compte, entre autres, les expressions introduites par un quantificateur universel du type TOUT et *chaque*.

- a. **negation condition**
For the NP n and arbitrary predicate term p : ' $n + p$ ' is false iff ' $n + \neg p$ ' is true
- b. consistency
If ' $n + p$ ' is true, necessarily ' $n + (\neg p)$ ' is false.
Or: ' $n + p$ ' and ' $n + (\neg p)$ ' cannot both be true.
- c. completeness
If ' $n + p$ ' is false, necessarily ' $n + (\neg p)$ ' is true.
Or: ' $n + p$ ' and ' $n + (\neg p)$ ' cannot both be false. (Löbner, 2011 : 293)

Une expression définie telle que le nom propre valide l'intégralité du test :

10) Lucie est partie. $\vee$ Lucie est restée.

Lucie n'est pas partie. $\vee$ Lucie n'est pas restée.

Il en va de même pour une expression démonstrative, à condition cependant de préciser que la référence en est fixée comme étant la même pour les deux occurrences :

Note that – if the reference of two or more occurrences of one demonstrative NP is fixed for the whole sentence – the criterion holds, along with the other definiteness criteria. (Löbner, 2011 : 297)

Le malaise qui surgit au sujet du démonstratif et du présupposé d'unicité est directement lié à la manière qu'a cette unité linguistique d'instaurer la sémantique définie, reposant sur un autre mécanisme que celui de l'article défini et passant obligatoirement (semble-t-il) par l'acte de référence. Selon des auteurs comme Ducrot (1980a) et Kleiber (1984a), une autre différence importante existerait ainsi entre descriptions définies et démonstratifs, ayant trait, cette fois-ci, à la relation entre le présupposé existentiel et la référence. Alors que dans le cas du défini, présupposé existentiel et référence pourraient se trouver dissociés, cela ne saurait être le cas pour le démonstratif :

Et si donc l'emploi normal du démonstratif exige bien l'existence d'un [...] objet, cette existence est la condition préalable de l'acte de démonstration. L'exigence d'existence (et d'unicité) a un tout autre statut dans le cas du défini. Aussi bien lorsqu'elle désigne que lorsqu'elle ne désigne pas, la description définie constitue un objet – dont elle fait l'univers du discours : c'est ce que nous exprimons en disant qu'elle le présuppose. Et si elle sert parfois à désigner, c'est-à-dire à fournir les indications permettant de le repérer dans un monde indépendant de la parole, c'est par l'intermédiaire seulement de sa fonction constitutive ou présuppositionnelle – que nous tenons pour primordiale. (Ducrot, 1980a : 244-245)

Kleiber, qui partage cette position, la résume en écrivant qu'« [...] une description démonstrative, comme tout symbole indexical, est vouée à la désignation³⁴ [...] » (1984a : 67). L'avis selon lequel le démonstratif, du fait de sa « vocation référentielle », entraîne automatiquement le présupposé existentiel pour son référent est également partagé par J. Lyons, qui écrit : « L'acte même

³⁴ Nous employons les termes « désignation » et « référence » comme synonymes.

de montrer quelque chose engage la personne qui réalise cet acte à croire en l'existence de ce qu'elle montre. » (1990 [1978] : 278)

Même si nous comprenons très bien la motivation de cette thèse, qui nous semble avoir sa source à la fois dans l'intuition et dans une tendance généralisée, nous nous permettrons néanmoins ici d'émettre quelques réserves. Il suffit d'envisager un référent qui, bien que repris anaphoriquement par le démonstratif, n'en possède pas moins une existence purement virtuelle :

- 11) Si j'avais de l'argent je m'achèterais [une grande maison]_i. [*Cette maison*]_i serait toute peinte en vert.

Pourtant, de tels énoncés sont non seulement parfaitement envisageables, mais aussi couramment produits par les locuteurs. Il semble d'ailleurs acquis en linguistique et philosophie du langage que le statut spécifique³⁵ d'un référent n'est pas en lui-même suffisant pour activer pour celui-ci le présupposé existentiel. Ainsi Prince écrit-elle que « [...] specificity is a necessary (though not sufficient) condition for an NP to carry an existential presupposition [...] » et cite des exemples tels que 12³⁶ :

- 12) If Mary finds [a pink truck]_i, she'll buy [*it*]_i, but I don't think she'll find one. (1981 : 236-237)

Par ailleurs, cette réfutation d'un présupposé existentiel inhérent tient également pour le défini :

- 13) Si j'avais de l'argent je m'achèterais [une grande maison]_i et [une voiture]_{ii}. [*La maison*]_i serait toute peinte en vert, et [*la voiture*]_{ii}, serait rouge.

- 14) Un enseignant dit à sa classe : « En fin d'année, *le meilleur élève de la classe* gagnera un prix³⁷. »

Résumons : si nous sommes d'accord pour dire que la référence est une propriété primordiale et, semble-t-il, ineffaçable du démonstratif³⁸ (contrairement au défini, qui, lui, connaît également des

³⁵ La notion logico-sémantique de *spécificité* étant complexe, nous renvoyons le lecteur aux articles de Fodor et Sag (1982), Galmiche (1983), von Heusinger (2002), Karttunen (1968a, 1976) et Prince (1981). À titre d'exemple, l'énoncé suivant (Prince, 1981 : 237) peut avoir deux lectures, une spécifique et une non-spécifique : *John wants to marry a Norwegian*. Lecture spécifique : la Norvégienne existe, il s'agit d'un individu concret et identifiable par le locuteur. Lecture non-spécifique : n'importe quelle Norvégienne pourvu qu'elle soit Norvégienne ; l'expression ne renvoie pas à individu identifiable pour le locuteur, et donc ne réfère pas.

³⁶ Pourtant, cette règle n'est peut-être pas absolue : dans un énoncé comme *Donnez-moi un croissant, s'il vous plaît*, prononcé par un client dans une boulangerie, le présupposé d'existence semble présent (le client suppose que des croissants soient présents), mais pas la spécificité (*un croissant, peu importe lequel*).

³⁷ La description définie *le meilleur élève de la classe* est employée de manière attributive (*celui qui sera le premier de la classe, peu importe qui*), faisant écho à la lecture attributive de la description définie *l'assassin de Smith* de Donnellan (1966). Fodor et Sag (1982 : 364) font un rapprochement entre les SN indéfinis non-spécifiques et l'emploi attributif des SN définis : "[...] it appears that Donnellan's distinction is closely related to our distinction between a referential and a quantificational use of an indefinite [...]."

³⁸ Et encore : il existe au moins un cas de figure où le démonstratif peut figurer dans un emploi purement attributif, à savoir justement les tours périphrastiques du type *celui qui sera le premier de la classe* ; mais sans doute faut-il tenir cet

emplois non-référentiels³⁹), elle n'entraîne pas forcément le présupposé existentiel pour le référent. Si ce dernier est à sauver à tout prix, il faudrait alors insister sur une définition plus extensible de la notion d'existence, applicable aussi aux mondes possibles et hypothétiques au sens de Martin (1983). Ensuite, quant à l'article défini et ce même présupposé d'existence, le même constat que pour le démonstratif doit être formulé, raison pour laquelle nous tenons pour problématique toute tentative d'appréhender la notion de définitude en termes de présupposé existentiel.

1.2.2 Le démonstratif : référence par contraste interne et prédication (re)classifiant le référent

Nombre d'auteurs (cf. Blanche-Benveniste et Chervel, 1966 ; Corblin, 1987 et 1995 ; Reichler-Béguelin, 1998) emploient pour les modes référentiels respectifs de l'article défini et du démonstratif adnominal les termes de « contraste externe » et « contraste interne ». Ainsi, alors que le démonstratif établit une différence au sein d'une même classe de référents, le défini saisit son référent sur la base de la différence entre plusieurs classes référentielles :

[...] au défini, la référence virtuelle est l'élément dont le calcul autorise à fixer un designatum, alors qu'au démonstratif, elle reclassifie un designatum isolé sur des critères externes [...]. (Corblin, 1987 : 217)

Pour paraphraser Corblin, lequel se sert ici du terme milnérien de « référence virtuelle » (Milner, 1976), une description définie gagne son référent grâce au sens dénotatif du nom introduit par l'article défini, et c'est seulement après, lorsque cette référence effectuée sur une base dénotative s'est avérée insuffisante, qu'il recourt au contexte. En revanche, pour le démonstratif, le contexte est primordial, la « référence virtuelle » n'entrant, à proprement parler, que de manière secondaire dans ce processus. Plusieurs phénomènes corroborent ce constat, à commencer par le fait que le démonstratif sera utilisé, comme nous venons de le voir, chaque fois que la situation présente un paradigme de référents potentiels dont tous ressortissent à la même « référence virtuelle », et entre lesquels il faut obligatoirement opérer une distinction contextuelle. Soit un autre exemple, emprunté à Corblin (1987 : 221) :

15) [Une voiture]_i était rangée devant la porte. C'est [*cette voiture*]_i que j'ai prise et non la tienne.

emploi du démonstratif comme un cas à part, car ce dernier assume une simple « Platzhalterfunktion » (Himmelfmann, 1997 : 78) au sein de la construction.

³⁹ Rappelons que le présupposé d'identification du référent par l'interlocuteur, lui, est conditionné par l'emploi référentiel de la description définie : "[...] when a definite description is used referentially, not only is there in some sense a presupposition or implication that someone or something fits the description, as there is also in the attributive use, but there is a quite different presupposition; the speaker presupposes of some particular someone or something that he or it fits the description. [...] No such presupposition is present in the attributive use of definite descriptions." (Donnellan, 1966 : 288-289)

16) [Une voiture]_i était rangée devant la porte. C'est [**la voiture**]_i que j'ai prise et non la moto.

C'est de nouveau Corblin (1987 : 196) qui remarque que du fait de la dépendance absolue du démonstratif au contexte, une expression démonstrative peut davantage faire l'économie de sens descriptif, comme on le voit dans le cas des pronoms démonstratifs *cela*, *ça* et *ceci*, sémantiquement assez ténus et pouvant à eux seuls constituer des syntagmes. C'est en traitant des approches cognitives de l'indexicalité au sein de ce même chapitre (Section 1.1.3) que nous avons insisté sur la nature directe et immédiate du lien entre le démonstratif et son contexte, ainsi que sur les conséquences qui en découlent pour la présentation du référent. Ce constat est régulièrement rappelé par la littérature :

La saisie du référent se faisant par investigation de la situation d'énonciation, il y a détachement avec tout ce qui peut toucher à la façon dont les participants conçoivent les événements auxquels ils sont mêlés. De là il découle que **le SN démonstratif installe le référent dans le focus d'attention des destinataires comme une entité nouvelle ou, en tout cas, en faisant abstraction de la représentation que les destinataires peuvent ou pourraient en avoir dans les circonstances où intervient l'échange**⁴⁰. (Charolles, 2002 : 120)

Cette saisie à chaque fois nouvelle du référent fait en sorte que celui-ci peut être présenté sous un jour nouveau : contrairement au défini, le démonstratif peut attribuer à l'objet désigné une propriété, permet sa (re)classification, sa (re-)catégorisation. En dehors des contextes d'anaphore résomptive (voir la Section 1.1.3 *supra*), il en va ainsi, par exemple, toutes les fois où le locuteur⁴¹ vise à reprendre le référent à l'aide d'un terme hyperonyme :

17) [Un chien]_i n'en finissait pas d'aboyer. [**Cet animal**]_i l'exaspérait.

Mais également lorsqu'il vise à assortir l'acte de référence d'un point de vue personnel portant sur le référent, l'expression de ce point de vue pouvant aller jusqu'à l'emprunt d'une métaphore :

18) [Un chien]_i n'en finissait pas d'aboyer. [**Ce sale clebs**]_i l'exaspérait.

19) Ils considérèrent [le plancher du wagon]_i, et décidèrent que pour une nuit, on pouvait se satisfaire de [**ce lit**]_i. (Corblin, 1987 : 212)⁴²

Le fait que le démonstratif soit doté de cette fonction de (re)classification du référent constitue l'une des raisons pour lesquelles Kleiber (1984a : 65) a proposé la thèse en vertu de laquelle toute description démonstrative « [...] doit être analysée comme étant l'abréviation d'une structure

⁴⁰ Nous soulignons.

⁴¹ Ou bien celui à qui le jugement en question est attribué, comme c'est typiquement le cas du discours indirect libre. Ce travail ne s'intéressera pas à la place du démonstratif dans les contextes de polyphonie, même si notre bibliographie contient quelques références en la matière.

⁴² Dans la littérature francophone, les reprises anaphoriques dans les énoncés 18 et 19 sont couramment désignées par le terme d'« anaphore infidèle » : « Les anaphores infidèles sont des syntagmes nominaux dont la tête est différente de celle de l'antécédent. » (Le Pesant, 2002 : 39-40) Dans le milieu tchèque, l'on parle plutôt d'« anaphore coréférentielle indirecte » („nepřímá koreferenční anafora“ ; Pešek, 2014).

attributive classificatoire » [...]. Kleiber précise dans le même texte que le nom faisant l'objet de cette structure renvoie à une représentation partitive du référent, et qu'il est, par conséquent, introduit par l'article indéfini ou partitif. Cette extensivité partitive⁴³ du nom s'explique aisément par la sémantique classificatoire de la structure, laquelle implique la présence d'autres référents (ou parties de référents) appartenant à la même classe référentielle. Ainsi, les SN *ce gâteau* et *cette bière* renverraient, dans la structure profonde, à *c'est un gâteau* et *c'est de la bière/une bière*⁴⁴. Cette analyse est d'autant plus convaincante qu'elle rapproche les SN démonstratifs de la forme pronominale *ce*, cette dernière étant elle-même souvent engagée dans une opération de (re)classification :

20) [Paul]_i, [c']_j est un monstre. (Kleiber, 1984a : 71)

Kleiber indique que dans les énoncés comme 20, le démonstratif apparaît parce qu'il est question d'un prédicat classifiant : Paul est (re)classifié par le locuteur comme étant un monstre. Cela expliquerait en bonne partie pourquoi le démonstratif *ce* et ses équivalents dans d'autres langues (les « identificational demonstratives » de Diessel, 1999 ; voir les Sections 1.3.2 et 1.3.3 *infra*) ont plus tendance à introduire les référents non animés⁴⁵ que les référents animés (et encore moins les animés humains) – la nature ontologique des premiers appelant davantage la classification opérée par le démonstratif –, mais aussi pourquoi le démonstratif reste en même temps possible avec un animé (humain) lorsque celui-ci fait l'objet d'un acte d'identification ou de (re)classification comme en 20.

1.2.3 Le démonstratif : un indexical opaque

Un autre constat qui permet à Gary-Prieur de justifier le rapprochement entre le démonstratif et l'indéfini, toujours en ce qui concerne l'identification du référent, c'est le fait que le démonstratif, « [...] contrairement à l'article défini, [...] ne suffit pas à donner au destinataire de l'énoncé les moyens d'opérer cette identification. » (2011 : 70) De cela s'ensuit la nécessité de disposer d'autres indices pour que l'acte référentiel utilisant le démonstratif réussisse. Selon Gary-Prieur (2011 : 71-72), ces indices peuvent être au nombre de quatre, à savoir :

- 1) la situation d'énonciation
- 2) un geste
- 3) une forme linguistique présente dans le contexte

⁴³ L'extensivité renvoie au « [r]apport de l'extensité à l'extension. L'extensivité est *extensive* en cas d'égalité déclarée de l'extensité à l'extension ; elle est *partitive* en cas d'inégalité déclarée de l'extensité à l'extension. » (Wilmet, 1986 : 194).

⁴⁴ Kleiber de préciser que « La prédication classificatoire véhiculée relève de la présupposition et non de l'assertion, de même que l'indication de l'unicité existentielle marquée par l'article défini est à porter au crédit de la présupposition et non de l'implication ». (1984 : 65)

⁴⁵ Dont les entités dites par J. Lyons (1990 [1978] : 77-78) « du second et du troisième ordre ».

4) une expansion du nom tête

Pour Gary-Prieur, cette insuffisance référentielle du démonstratif aurait pour cause une asymétrie dans l'accès des actants de l'énonciation au référent d'une expression démonstrative, car le démonstratif, par définition centré sur le locuteur, ne présupposerait, contrairement au défini, aucune connaissance préalable du référent chez l'interlocuteur :

C'est cette asymétrie des interlocuteurs face à l'identification du référent du GN démonstratif qui permet de comprendre une propriété relevée par la plupart des linguistes : l'incomplétude du démonstratif. (2011 : 69)

Cependant, l'existence de trois phénomènes nous permettrait de relativiser l'asymétrie dont parle Gary-Prieur, sans pour autant remettre en cause le fait que l'emploi du démonstratif est intrinsèquement lié à la perspective du locuteur. D'abord, très souvent, le référent d'une expression démonstrative est parfaitement identifiable par le destinataire de l'énoncé. C'est ce même fait qui, d'ailleurs, permet l'effet de connivence produit par les emplois mémoriels, ainsi que par une partie des emplois « émotionnels » du démonstratif (nous verrons cela plus concrètement dans le Chapitre 3, Section 3.5, et dans les Chapitres 6 et 7). Ensuite, parce que le défini semble lui aussi connaître des emplois où le référent est identifiable pour le seul émetteur de l'énoncé (il serait ainsi question d'un référent indéfini spécifique), comme le montre cet exemple cité par Epstein (2002 : 368), présentant le synopsis d'un film :

- 21) The film's setting and the story both have a mythic simplicity. In the aftermath of a drought that leaves most people surviving by selling themselves into lifelong servitude, a farmer and a weaver escape and set up residence in a desert ghost town. Their only contact with the outside world is a trader who keeps them in debt to him while also keeping them supplied with essentials. Then *the woman* arrives, like a fleeing animal. Her family has been killed in a flood.

Et, enfin, parce que des notions telles que « connaissance du référent » et « familiarité avec le référent » s'avèrent, de manière générale, plus que problématiques dans l'étude des emplois référentiels des descriptions définies⁴⁶.

⁴⁶ Contrairement à certains auteurs travaillant sur la définitude (cf. Christophersen, 1939 ; Heim, 2011 [1982] ; Roberts, 2003), nous sommes d'avis que ces termes n'ont aucune pertinence explicative. D'abord, parce que les approches qui en font leur clef de voûte ne permettent de rendre compte que d'une part des emplois référentiels du défini – notamment des emplois anaphoriques –, mais s'avèrent peu opérantes pour expliquer toute une série d'autres emplois ; comme le résume Czardybon (2017 : 14) au sujet de l'approche de Heim, pour qui la familiarité s'établit obligatoirement via une mention antérieure du référent dans le discours : "Other contexts, however, seem to be problematic; for example, SNs with complements establishing uniqueness, associative anaphors, NPs with superlatives and ordinal numbers as well as uniques. Here, the referents of the definite NPs need not be mentioned previously in order to occur with the definite article." Pour les deux autres auteurs susmentionnés, le bilan est seulement un peu plus satisfaisant. Ensuite, parce que les termes sont vagues par nature : que veut dire exactement « être familier avec le référent » ? Pour les trois auteurs, cela semble impliquer tout simplement le fait que ce dernier peut être identifié par l'interlocuteur. Or comme le pointe à juste titre C. Lyons (1999), « identifiable » ne veut pas forcément dire « familier ». Typiquement, un référent anaphorique est identifiable grâce à sa reprise par le discours, mais cela ne veut pas dire que l'interlocuteur en possède une connaissance intime. Les choses deviennent encore plus problématiques au sujet de SN comme *The bloke Ann went out with last night*

Après avoir introduit, dans la précédente section, le terme de « *token* réflexif opaque », nous nous concentrerons à présent sur les différents emplois auxquels le démonstratif peut se prêter. Dans son article « Les démonstratifs (démon)trant-ils ? Sur le sens référentiel des adjectifs et pronoms démonstratifs » (1983), Kleiber montre que l'étymologie du terme *démonstratif*, renvoyant à un acte de monstration, fausse en partie l'interprétation qui en est souvent proposée en termes sémantico-référentiels. Car si le démonstratif connaît des emplois où sa combinaison avec un geste, quelque discret qu'il puisse être, s'avère nécessaire, ces emplois ne constituent qu'un type possible (le type 2) dans la typologie de Gary-Prieur reproduite *supra*. Dans les emplois « symboliques », c'est la situation d'énonciation dans son ensemble qui fournit la partie manquante des indices, permettant ainsi l'aboutissement de l'acte référentiel. Dans ce type d'emplois, la référence démonstrative peut évidemment viser l'endroit dont l'occurrence du démonstratif fait directement partie (*Il fait chaud dans cette voiture !* ; l'occurrence *ce chocolat* figurant sur l'emballage d'un chocolat). Mais il s'agit, de manière générale, de tous les cas où l'appariement entre l'occurrence du démonstratif et le référent s'établit sur la base d'une contiguïté spatio-temporelle non textuelle, non mémorielle et non ostensive, y compris lorsqu'il est question « [...] d'une référence à des entités non concrètes comme celles de *ce soir, cette odeur, cette imagination* [...] » (Kleiber, 1983b : 104) ou des cas comme la notice suivante, accrochée au-dessus d'un urinoir :

22) Prière de garder *cet urinoir* propre.

Les emplois endophoriques relèvent du type 3) ci-dessus, l'appariement entre l'occurrence du démonstratif et le référent s'établissant grâce à la contiguïté avec une autre expression référentielle présente dans le discours. Une lecture plus extensive de ce même type pourrait aussi, à notre avis, rendre partiellement compte des emplois mémoriels (« pseudo-anaphoriques » ; Berger, 1993) du démonstratif.

Mais arrêtons-nous à présent davantage sur le type 2) : quel est le rapport entre le démonstratif et le geste qui l'accompagne⁴⁷ ? Kleiber examine la thèse d'un démonstratif qui, tel un doigt tendu ou une flèche, pointerait le référent. Il réfute aussitôt cette thèse en écrivant que

phoned a minute ago (C. Lyons, 1999 : 5), se prêtant volontiers à des emplois où l'interlocuteur (voire même le locuteur) ne saurait être familier avec le référent, car c'est la relative qui vient tout juste d'en établir l'existence dans son univers du discours : avant que le locuteur n'ait prononcé le SN contenant la relative « établissante », il se peut que l'interlocuteur n'ait même pas su qu'Ann était sortie avec quelqu'un la veille. L'on voit ainsi que le référent n'est pas familier pour lui, puisqu'il n'est même pas identifiable à proprement parler.

⁴⁷ Ce « geste » peut être de plusieurs natures, allant d'un pointage du doigt explicite jusqu'à la simple position adoptée par le regard du locuteur en passant pour un mouvement de la tête, des yeux, etc. (cf. Kleiber, 1983b : 104 ; Mc Ginn, 1981 : 183 ; Charolles, 2002 : 106). Charolles (2002 : 109) distingue l'« ostension forte » (le pointage du doigt) et « faible » (la direction du regard).

[s]i les démonstratifs étaient des « pointeurs », les gestes de désignation qui les accompagnent dans de tels emplois seraient en somme redondants. Or, ces gestes, bien loin de faire double emploi, sont nécessaires. Ce sont eux qui jouent le rôle de la flèche qui dirige l'attention de l'interlocuteur vers l'endroit précis où se trouve le référent et non les démonstratifs. (1983b : 105)

Une version plus modérée de cette thèse consisterait à inclure « [...] dans le contenu sémantique des démonstratifs non pas la direction d'application précise, qui est réservée au geste d'ostension, mais uniquement un trait de monstration ou de localisation. » (Kleiber, 1983b : 107) Ainsi, plutôt que d'agir comme un geste de pointage, une occurrence du démonstratif renverrait à la présence de ce geste, réalisé concomitamment à elle. Telle serait la position de Ducrot, pour qui le démonstratif « [...] ne présente pas lui-même son objet : il attire seulement l'attention sur un geste de démonstration concomitant. » (1980a : 244) Mais ce sont justement tous les autres types d'emplois – dépourvus, eux, de geste – qui donnent tort à la volonté d'inscrire une telle instruction dans la sémantique du démonstratif. Par conséquent, la seule instruction dont on puisse affirmer qu'elle s'y trouve vraiment codée est celle qui invite l'interlocuteur à traquer le bon référent par le biais d'un recours au contexte immédiat⁴⁸.

Il reste enfin à dire un mot sur les emplois du type 4), c'est-à-dire ceux où l'indice nécessaire est fourni par une expansion du nom tête. Dans de tels cas, le démonstratif est en l'occurrence doté d'un fonctionnement (du moins partiellement) cataphorique, rentrant ainsi dans le phénomène de la cataphore intra-syntagmatique⁴⁹:

- 23) Toutes choses, doucement, tendrement, se laissaient aller à l'existence, comme *ces femmes lasses* qui s'abandonnent au rire et disent « C'est bon de rire » d'une voix mouillée.
(J.-P. Sartre, *La Nausée*, emprunté à Gary-Prieur, 1998 : 45)

Ici, la relative semble essentielle pour le fonctionnement de la référence, car elle restreint l'extension du nom tête, construisant une sous-classe générique composée d'un certain type de femmes lasses. Sans elle, la référence ne serait pas complète, de même que dans l'énoncé suivant, où il n'est plus question de référence générique :

- 24) Tu te rappelles *cette vieille dame* qui roulait toujours à vélo en jupe ?

⁴⁸ Kleiber emploie la comparaison suivante : « Soit par exemple une lumière qui s'allume sur le tableau de bord d'une machine ou une sonnerie qui retentit pour signaler une panne, sans en préciser toutefois la nature exacte. L'utilisateur est obligé de chercher la panne qui a déclenché l'occurrence de la lumière ou de la sonnerie. » (1983b : 114-115)

⁴⁹ La position de Gary-Prieur, plus radicale, postule même un sens cataphorique par défaut pour presque tous les emplois du démonstratif : « On a beaucoup discuté du statut du démonstratif (anaphorique ? déictique ? les deux ?), mais sans accorder me semble-t-il la place qu'elle mérite à une propriété que le démonstratif est pourtant le seul élément de la langue (ou presque) à mettre en œuvre, et que j'appellerai la dimension cataphorique. J'entends par là que le démonstratif renvoie presque toujours, d'une manière ou d'une autre, à la suite du discours [...] » Ainsi, une phrase comme *Regarde cet avion* anticiperait une suite : « [...] si le locuteur utilise un démonstratif pour désigner tel objet, c'est qu'il a quelque chose de particulier à dire à son propos, et il serait assez étrange que le discours s'arrête là. » (1998 : 44)

Sock et Kleiber (2006) résument que de tels emplois du démonstratif exigent « [...] la présence du contexte aval, puisque, comme la plupart des commentateurs l'ont noté, la suppression de l'expansion leur est fatale, la perte, tout particulièrement celle de la relative, suscitant une interrogation sur l'identité du référent visé [...] » (2006 : 253). En même temps, de nombreux linguistes et grammairiens (cf. Charolles, 2002 ; Corblin, 1987 ; Fuchs, 1987 ; Gary-Prieur, 1998 ; Kleiber et Sock, 2006) ont relevé que la nature de l'expansion d'un SN démonstratif – notamment lorsque celle-ci est une relative – se trouve partiellement brouillée pour ce qui est de l'opposition nécessaire/accessoire. Gary-Prieur affirme à cet égard :

On sait que ces relatives ne sont ni « restrictives » ni « descriptives », et que la corrélation établie avec le déterminant n'est pas la même que dans les GN introduits par un article défini. (1998 : 45)

Charolles écrit qu'

[...] avec le défini l'adjectif participe à l'identification du référent, il fournit le trait qui permet de distinguer celui-ci comme unique de sorte que, quand cet adjectif se présente sous la forme d'une relative [...] celle-ci satisfait à l'appellation de déterminative. Avec le démonstratif, l'adjectif paraît par contre rajouté après coup (sans être pour autant apposé ni explicatif), au vu par exemple des hésitations de l'interlocuteur. (2002 : 120)

Une autre position sur ce sujet est celle de Wilmet, qui affirme que même lorsqu'elle apparaît avec le nom d'un SN démonstratif, une relative peut être dite « déterminative » ou « prédicative » (2010 [1997] : 497). Personnellement, nous penchons vers un compromis : si ce brouillage de la nature de la relative constitue, selon nous, une réalité des SN et complexes nominaux démonstratifs⁵⁰, il ne serait pas juste d'aller jusqu'à rejeter catégoriquement l'opposition en question. Cela nous paraît évident lorsque nous comparons 23 et 24 à 25 :

25) *Ces programmes de télé qui exaspèrent les français.* (emprunté à Kleiber et Sock, 2006 : 259)

Comme l'expliquent les auteurs, dans cet énoncé, qui constitue le titre d'un article au sujet de l'ensemble des programmes de télé en France, la relative ne restreint pas l'extension du nom⁵¹.

Le brouillage en question est en tout cas dû à un double fonctionnement référentiel du démonstratif, lequel renvoie ici en aval – à l'expansion du nom –, mais aussi en amont, mobilisant

⁵⁰ Force est toutefois de préciser que ce type de brouillage est également attesté pour les relatives accompagnant les noms introduits par le défini et l'indéfini. Quant au critère de la virgule pour le français, celui-ci n'est pas très fiable : même si nous nous limitons exclusivement aux SN définis – où la différence apparaît le plus nettement –, la présence de la virgule semble automatiquement annoncer une relative appositive, mais son absence ne veut pas forcément dire que la relative soit déterminative (voir notamment Fuchs, 1987), raison pour laquelle certains auteurs comme Corblin (1987) parlent de « relatives internes », dont le critère définitoire est purement syntaxique, à savoir l'absence de virgule. En tchèque, toute distinction graphique est abolie par les règles orthographiques, les relatives, indépendamment de leur statut, étant toujours précédées d'une virgule.

⁵¹ Même si, comme ils le soulignent par ailleurs, elle est nécessaire pour la construction de la référence, et pourrait difficilement être enlevée.

des connaissances partagées par le locuteur et l'interlocuteur. Le premier mouvement référentiel est cataphorique, le second, mémoriel (« pseudo-anaphorique ») :

Il s'agit de l'introduction d'un nouveau référent dans la mémoire immédiate [...], mémoire commune au locuteur et à l'interlocuteur au moment de l'énonciation du SN démonstratif. Ce référent nouveau ne l'est pas totalement, mais est supposé (ou supposé pouvoir être) déjà disponible dans la mémoire longue. Il se trouve réactivé ou réintroduit dans la mémoire immédiate discursive via l'expression d'une proposition portant sur ce référent, proposition supposée également faire partie du stock des propositions déjà acquises dans l'univers de croyance [...] de l'interlocuteur. (Kleiber et Sock, 2006 : 261)

Ainsi, pour simplifier un peu la comparaison des relatives dans les SN (complexes nominaux) démonstratifs à celles dans les SN (complexes) définis, nous dirons que pour les premières, la construction de la référence est, de manière générale, moins dépendante de l'expansion que dans le cas des relatives déterminatives (restrictives) tout en étant plus dépendante d'elle que dans le cas des relatives descriptives (appositives, prédicatives). Ensuite, au sein même des relatives modifiant les SN (complexes) démonstratifs, certaines s'approchent davantage du pôle « restriction », et d'autres du pôle « apposition ».

1.3 Le démonstratif : un « primitif » du langage humain

D'une part, le démonstratif semble être une catégorie universelle, que l'on retrouve dans tous les systèmes linguistiques. D'après les chercheurs qui ont étudié quantité de langues du monde (cf. Anderson et Keenan, 1985 ; Himmelmann, 1997 ; Diessel, 1999 ; Dixon, 2003), il n'y aurait pas une seule langue sans démonstratif, celui-ci représentant, de fait, une « catégorie primitive » du langage humain (voir aussi Clark, 1978 et Diessel, 2006 pour la place primordiale du démonstratif dans le processus d'acquisition du langage). Cette conclusion est également corroborée par le constat que contrairement aux autres mots grammaticaux (« catégories mineures » ; Marchello-Nizia, 2009), les démonstratifs, en l'occurrence, ne sont pas issus d'unités lexicales (« catégories majeures » ; Marchello-Nizia, 2009) par grammaticalisation :

Demonstratives are used to orient the hearer in the speech situation, focusing his or her attention on objects of interest. This is one of the most basic functions of human communication for which there might be a particular class of linguistic expressions that emerged very early in the evolution of language. [...] All this suggests that demonstratives are not ordinary grammatical markers, and hence it would make sense if they do not derive from lexical items as all other grammatical markers. [...] If demonstratives are not derived from lexical items, there would be two different sources from which grammatical markers may emerge: lexical expressions and demonstratives. The grammaticalization of items from both domains would be unidirectional. That is, grammatical items develop from lexical expressions and demonstratives but never vice versa. Furthermore, there would be no transitions between the two source domains — demonstratives do not develop into lexical items, nor is there evidence that lexical items have ever been reanalyzed as demonstratives. (Diessel, 1999 : 152)

D'autre part, une grande diversité existe entre les langues eu égard au système du démonstratif et à l'organisation de l'espace déictique. Nous nous contenterons ici d'en fournir une comparaison très cavalière tout en nous focalisant, avant tout, sur ce qui nous semble utile pour le futur traitement du français et du tchèque. Nous nous pencherons ensuite sur la sémantique du démonstratif ; nous nous interrogerons, entre autres, sur le rapport entre le démonstratif et l'expression de la distance tout en nous posant la question – capitale – de savoir si celle-ci constitue une condition nécessaire pour la définition de la catégorie du démonstratif. Puis nous verrons brièvement la morphosyntaxe du démonstratif et, enfin, nous aborderons le thème du démonstratif en diachronie.

1.3.1 Le démonstratif dans la variabilité typologique

Diessel (1999 : 26) écrit que :

All languages have at least two demonstratives locating the referent at two different points on a distance scale: a proximal demonstrative referring to an entity near the deictic center, and a distal demonstrative indicating a referent that is located at some distance to the deictic center.

Dixon (2003 : 86), quant à lui, arrive à la conclusion que

The majority of languages have two demonstratives, relating to “near speaker” and “not near speaker”. There are also a fair number of languages with a three-term system.

Le système à deux formes est appelé « minimal » (“minimal system”) par Anderson et Keenan (1985 : 81). De nombreuses langues possèdent cependant des systèmes à trois, quatre, voire plus de formes encore⁵².

S'agissant de l'orientation du système du démonstratif et de l'organisation de l'espace déictique des langues, vu l'emprise des approches « énonciativistes » et de la notion bühlérienne d'*Origo* sur les études de la deixis, ce sont la place occupée par le locuteur (ou, éventuellement, l'interlocuteur) dans l'espace et la distance du référent par rapport à ce dernier qui constituent un angle d'approche courant. Cette approche, dite souvent « distance-oriented » (Anderson et Keenan, 1985 : 282-286) dans la littérature d'expression anglaise, est dotée d'un grand pouvoir explicatif. Toutefois, on s'est vite aperçu qu'elle ne suffit pas pour rendre compte du fonctionnement du démonstratif dans une approche typologique. C'est pour cela qu'une autre approche est apparue aussitôt, essayant de rendre compte de l'usage des différentes formes du démonstratif au sein d'un système en prenant pour point de départ les rôles assumés par les actants de l'énonciation. Une telle approche, de même que les

⁵² Anderson et Keenan (1985 : 295) citent l'exemple de la langue youpik, parlée en Alaska, qui contiendrait jusqu'à 30 formes de démonstratif, permettant d'exprimer un éventail d'informations très riche sur le référent, y compris si le référent est en haut ou en bas par rapport au locuteur, s'il se trouve dans la direction du courant d'une rivière, sur une île, etc. À l'opposé, certains systèmes peuvent, sous certaines conditions, être envisagés comme ne comprenant qu'une seule forme, dont (à la limite) le français moderne (voir la Section 1.3.2 *infra* et le Chapitre 4).

systèmes qui affichent ce fonctionnement, portent le nom de « person-oriented » (Anderson et Keenan, 1985 : 282-286). Au sujet des systèmes du français et du tchèque, nous pouvons constater qu'ils sont orientés par rapport à la distance, sachant que tous les deux distinguent entre trois degrés de celle-ci. Voici une synthèse sous forme de tableaux de ces deux systèmes, qui feront l'objet d'analyses détaillées dans les Chapitres 3 et 4 :

Distance par rapport au locuteur	Formes	
	Déterminants	Pronoms
Neutre	<i>ce</i>	<i>celui</i>
Proche	<i>ce X-ci</i>	<i>celui-ci</i>
Lointaine	<i>ce X-là</i>	<i>celui-là</i>

Tableau 1.1 : Système du démonstratif nominal en français (une simplification)

Distance par rapport au locuteur	Formes (en emplois adnominaux ou pronominaux)
Neutre	TEN
Proche	TENTO, TENHLE, TENHLETEN, TUHLETEN, TADYTEN, TADYHLETEN...
Lointaine	TAMTEN, TAMHLETEN, ONEN...

Tableau 1.2 : Système du démonstratif nominal en tchèque (une simplification)

Le latin classique, typiquement, avec ces trois formes – HIC, ISTE et ILLE – assez étroitement associées aux trois personnes de l'énonciation, compte parmi les langues orientées par rapport à la personne. C'est également le cas du japonais, où la symétrie est encore plus poussée, les trois formes du démonstratif nominal faisant écho à trois autres formes d'un démonstratif adverbial, dont elles sont morphologiquement proches :

Démonstratif nominal	Démonstratif adverbial
<i>kore</i> = proche du locuteur	<i>koko</i> = ici
<i>sore</i> = proche de l'interlocuteur	<i>soko</i> = là
<i>are</i> = distant des deux	<i>asoko</i> = là-bas

Tableau 1.3 : Système du démonstratif en japonais

Mais une telle répartition stricte en systèmes orientés par rapport à la distance et par rapport à la personne est en partie contestable. D'abord, parce que l'on peut aussi considérer qu'entre les deux, il s'agit davantage, comme l'explique Meira (2003 : 3), d'une affaire de continuum que d'une opposition nette :

An important distinction is often made in the literature on demonstratives between so-called *distance-* and *person-oriented* systems [...]. These terms are somewhat misleading, a distance-based system is still person-oriented in that it is speaker-centred, i. e. its 'proximal', 'medial' and 'distal' terms refer to the position of the speaker.

De même pourrait-on aussi soutenir que dans les systèmes dits « orientés par rapport à la personne », les différentes distances du référent vis-à-vis du locuteur finissent par être transférées aux acteurs de l'échange linguistique. Comme cette place (notamment celle de l'interlocuteur) est porteuse d'une importance pragmatique, elle finit par être codée par le système. Comme l'écrit de nouveau Meira (2003 : 3) :

It may be that, for distance-based systems, the position of the addressee is pragmatically important (it is an "unmarked salient place" for reference), and that this importance or salience has simply grammaticalized into a semantic feature in personal-based systems.

Ensuite, la notion de distance s'avère elle-même problématique. Comme l'ont démontré Jungbluth (2003) pour l'espagnol, Marchello-Nizia (2003, 2004, 2006a, 2006b) pour l'ancien français et Guillot et Carlier (2015) pour le latin, elle doit être entendue à la fois au sens littéral de distance spatiale et au sens figuré de distance *pragmatique*⁵³. Celle-ci dépendrait de toute une série de configurations que les personnes de l'interlocution peuvent adopter dans l'espace, mais aussi du choix fait par le locuteur d'inclure ou non l'interlocuteur dans ce que Marchello-Nizia (2003) appelle sa

⁵³ Comme le remarque Hanks (1992 : 52) : "The standard assumption that the space is always foundational in deixis is an inconvenient fiction not borne out comparatively." De même que Price (1971 : 121), traitant des démonstratifs anglais : "The distinction between *this* = 'nearness' and *that* = 'distance' is a subjective, not an objective one: it is fundamentally one of nearness or remoteness of interest." Ou encore Guénette (1995 : 12-13), écrivant au sujet des démonstratifs latins : « D'une opposition fondée sur la personne grammaticale, on en vient à une opposition de "distance textuelle" et finalement, à une opposition de "distance d'intérêt ou psychologique" ».

« sphère personnelle »⁵⁴. Cette double lecture de la notion de distance nous paraît incontournable non seulement pour la description des usages situationnels du démonstratif, mais aussi – ou surtout – de ses usages endophoriques, exophoriques mémoriels et « émotionnels »⁵⁵. À titre indicatif, nous donnons l'exemple suivant (construit par nous), où le locuteur compare deux livres d'un auteur :

- 26) J'aime beaucoup [son dernier livre]_i. Je me rappelle en avoir lu [un autre]_{ii} de lui, mais [*ce livre-là*]_{iii}, je ne l'ai pas du tout aimé. Par contre, [*ce livre-ci*]_i, je l'aime vraiment beaucoup.

Le référent d'*un autre* est repris par la forme distale CE X-*là*, bien que la reprise suive immédiatement son antécédent, alors que le référent de *son dernier livre* est repris par la forme proximale CE X-*ci*, même si la distance textuelle qui sépare cette reprise de son antécédent est plus longue que dans le premier cas. Selon nous, l'explication peut être cherchée, ici aussi, dans le domaine de la distance pragmatique : c'est le livre le plus récent qui se trouve au centre (foyer) de l'attention du locuteur et non l'ancien. Soit le même exemple en tchèque⁵⁶ :

27) Mně	se	hodně	lib-í	[jeho	posledn-í	knih-a] _i
moi.DAT	REFL	beaucoup	plaire-PRS.1SG	sien.F.NOM.SG	dernier-F.NOM.SG	livre-NOM.SG
Vzpomín-á-m	si,	že jsem		od něho	čet-l	i
rappeller-PRS-1SG	REFL	que AUX.PRS.1SG		de lui.GEN	lire-PTCP.SG.M	aussi
[nějak-ou		jín-ou] _{ii} ,				
quelque-F.ACC.SG		autre-F.ACC.SG				
ale [ta] _{ii}		se	mi	vůbec	ne-líbi-l-a.	
mais celui-ci.F.NOM.SG		REFL	moi.DAT	du tout	NEG-plaire-PST-3SG.F	
[<i>Tahle</i>		<i>knih-a</i>] _i	se	mi	ale	opravdu lib-í.
tenhle.F.NOM.SG		livre-NOM.SG	REFL	moi.DAT	MOD	vraiment plaire-PRS.3SG

Ici, le démonstratif proximal TENHLE apparaît au même endroit que CE X-*ci* en français. La forme distale française est traduite par la forme neutre TEN.

⁵⁴ L'auteure a forgé cette notion pour rendre compte des emplois des démonstratif CIST (forme proximale) et CIL (forme distale) de l'ancien français. En voici un résumé donné par Guillot : « La sphère personnelle du locuteur se définit comme une sorte d'espace très abstrait construit par le locuteur dans son propre discours et incluant tout ce qui le concerne de près ou de loin. Cette sphère comprend naturellement les objets qui lui appartiennent, les parties de son corps, les membres de sa famille, ses proches, ceux qu'il aime et chérit, parfois même les paroles qu'il a prononcées. C'est donc en fonction de la visée communicative de son énoncé que le locuteur choisira d'inclure ou d'exclure un référent de sa sphère et non en fonction de la position de cet objet dans l'espace. Pour inclure le référent à l'intérieur de sa sphère, le locuteur choisira l'une des formes du paradigme de CIST. Pour l'en exclure, il utilisera au contraire l'une des formes de CIL. » (2010b : 237-238)

⁵⁵ Comme on le verra, la forme se prêtant de loin le plus souvent aux emplois « émotionnels » en tchèque est la forme neutre pour la distance TEN. En français, il semble toujours être question de CE tout seul, sans particule localisatrice.

⁵⁶ Dans les deux cas, il ne s'agit, bien évidemment, que d'une formulation possible parmi d'autres, le but étant uniquement de montrer que les locuteurs des deux langues ont la possibilité de marquer la distance pragmatique.

1.3.2 La sémantique du démonstratif

La question qui se pose de façon incontournable pour le démonstratif est la suivante : comment définir son sens ? La conclusion en vertu de laquelle ce sens reposerait (essentiellement) sur l'expression de la distance qui sépare le référent du centre déictique semble séduisante. Comme l'écrit Diessel (1999 : 37-38) tout en se référant à une thèse radicale de Harris, portant sur le démonstratif français :

Harris (1978, 1980) argues that *ce*, *cette* and *ces* can be viewed as definite articles rather than demonstratives when they occur without a distance marker. This raises the interesting question whether distance is a necessary feature of the category demonstrative. Are demonstratives generally marked for distance or are there reasons to consider an item a demonstrative even if it does not indicate the relative distance of its referent to the deictic center? Anderson and Keenan (1985: 280) argue that a deictic expression unmarked for distance “would be little different from a definite article” or third person pronoun (cf. Frei 1944: 119). In their view, demonstratives are generally distance-marked.

D'une part, il semble généralement vrai que les démonstratifs neutres sur le plan de l'expression de la distance ont tendance à se rapprocher d'un article défini. Pour le tchèque (et les autres langues slaves occidentales, cf. Ortmann, 2014 et Czardybon, 2017), ce constat apparaît avec une évidence flagrante. Pour le français, certains auteurs (cf. Harris, 1977, 1978 et 1980 ; Carlier et De Mulder, 2006) penchent également vers l'hypothèse d'un CE déterminant déictiquement affaibli et se rapprochant de l'article défini (voir la Section 4.3 pour une meilleure analyse de cette thèse). L'on sait qu'en diachronie, ce sont les formes neutres ou marquées pour l'éloignement (éventuellement, dans les systèmes orientés par rapport à la personne, les formes associées à la 3^{ème} personne) qui constituent de loin le plus souvent une source de futurs articles définis, contrairement aux formes marquées pour la proximité (cf. Greenberg, 1978a ; Himmelmann, 1997 et 2001 ; Carlier et De Mulder, 2006, 2010 et 2011). De plus, dans la majorité des systèmes linguistiques, l'article défini – s'il est présent – n'exprime pas non plus d'oppositions de distance.

Mais faut-il en conclure pour autant que l'expression d'une distance constitue une condition définitoire nécessaire (voire principale) pour la classe des démonstratifs ? Il nous semble évident que non. Certes, les formes neutres du point de vue de la distance sont susceptibles de s'engager dans un processus aboutissant à une forme d'article défini (nous verrons dans ce travail que c'est précisément le cas du tchèque), mais ce fait n'enlève pas automatiquement à ces formes le statut de démonstratif. Ainsi, dans le français contemporain, CE reste bel et bien un démonstratif, car possédant le fonctionnement d'un déictique opaque et s'opposant clairement, par ailleurs, à la forme LE (voir la Section 1.2 *supra*). De même en tchèque, si le démonstratif TEN en emploi adnominal est souvent « javélisé » au point de rappeler fortement un article défini (voir le Chapitre 6), cette « javélisation » reste en grande partie une affaire de contexte d'emploi, et doit faire l'objet d'une évaluation au cas

par cas. D'autres contre-exemples peuvent être fournis, comme celui des formes allemandes DER (*das* au neutre) et DIES(ER)⁵⁷, agissant comme des démonstratifs, malgré leur incapacité à coder la distance :

Dies is almost always used adnominally, but *das* can be both an independent pronoun and a modifier of a cooccurring noun. *Dies* and *das* do not contrast deictically: both forms may occur with proximal and distal meaning. In order to indicate that *dies* or *das* are used contrastively, they are commonly accompanied by a demonstrative adverb (e.g. *das da* 'this/that there', *das Haus da* 'this/that house there'). (Diessel, 1999 : 38)

Aussi Diessel (1999 : 38), se référant en partie à Himmelmann (1997), conclut-il comme suit :

I assume therefore, with Himmelmann, that demonstratives are not generally distance-marked. Some languages have demonstratives that do not indicate a deictic contrast. The occurrence of distance-neutral demonstratives is, however, crosslinguistically infrequent. Apart from German and Supyire there are only five other languages in my sample in which some demonstratives are distance-neutral: Alambak, **French**, **Czech**⁵⁸, Koyra Chiini, and Tok Pisin.

En vérité, il s'avère que l'expression de distance ne constitue ni une condition nécessaire ni une condition suffisante pour définir la classe des démonstratifs. Nous venons de voir pourquoi elle n'est pas nécessaire. La raison pour laquelle elle n'est pas suffisante est la suivante : même si dans la plupart des langues à article défini, celui-ci est neutre eu égard à la distance, des exceptions existent également, comme le macédonien, langue où trois formes de démonstratif s'opposant quant à l'expression de la distance ont produit trois formes d'article défini où cette opposition se voit conservée⁵⁹ :

Articles neutres : *-ot*, *-ta*, *-to*, *-te*,

Article marquant la proximité : *-ov*, *-va*, *-vo*, *-vé*

Articles marquant la distance : *-on*, *-na*, *-no*, *-né*

Tableau 1.4 : Système de l'article défini en macédonien

En même temps, force est de constater avec Diessel (1999, 38-39) que

All languages employ at least some demonstratives that are distance-marked [...]. Distance is thus after all a feature that occurs in the demonstrative system of all languages even though individual elements of the system may lack a distance feature.

⁵⁷ DER doit être accentué. Dans les cas où il ne l'est pas, il fonctionne comme article défini.

⁵⁸ Nous soulignons.

⁵⁹ Sur le plan morphosyntaxique, il semble y avoir une hésitation au sujet du statut de suffixes ou de clitiques de ces articles (cf. C. Lyons, 1999). Nous transcrivons les formes en alphabet latin ; les oppositions en genre et en nombre, codées par les formes des trois paradigmes, ne sont pas détaillées ici.

L'auteur de préciser ensuite que si certaines formes du démonstratif nominal dans les langues du monde ne codent pas d'oppositions de distance, celles du démonstratif adverbial en codent toujours.

Pour Diessel (1999, 35), le sens du démonstratif comporte deux ensembles de traits, à savoir les traits « déictiques » et « qualitatifs » :

The meaning of demonstratives comprises two kinds of features: (i) deictic features, which indicate the location of the referent relative to the deictic center, and (ii) qualitative features, which characterize the referent [...].

Le premier ensemble inclut, selon l'auteur, les traits suivants : a) la distance, b) la visibilité, c) l'altitude ("elevation"), d) la position géographique et e) le mouvement du référent. Quant au second, il est composé des traits suivants : a) le statut ontologique (endroit, objet ou personne), b) l'opposition animé/inanimé, c) le trait « humain », d) le sexe, e) le nombre et f) la nature délimitée ("boundedness") ou non du référent.

Ainsi, l'expression de la distance relèverait des traits sémantiques déictiques du démonstratif, sans pour autant qu'elle doive être présente, constat qui tient également pour les autres traits de cet ensemble. Il en va de même pour les traits sémantiques qualitatifs. Ainsi, les formes françaises du déterminant *ce/cet/cette/ces* tout comme les formes pronominales *celui/celle/ceux/celles* sont-elles uniquement marquées pour le sexe (le genre) et le nombre. Les formes pronominales *ceci, cela, ça* et *ce* sont, quant à elles, marquées pour le genre – renvoyant toujours à des entités neutres ou (re)catégorisées comme telles par le locuteur, le nombre (toujours singulier), ainsi que pour la nature ontologique du référent, renvoyant (a priori et dans leurs usages prototypiques⁶⁰) à des référents inanimés du type « objet ». En tchèque, les formes *TEN, TENTO, TENHLE, TAMTEN*, etc. sont marquées pour le genre, le nombre et le cas en emploi adnominal et pronominal, cependant qu'elles possèdent toutes une forme pronominale neutre qui, sur le plan morphosyntaxique, agit comme la forme pronominale neutre française *ce* (voir la Section 1.3.3 *infra* et le début du Chapitre 3).

1.3.3 La morphosyntaxe du démonstratif

Sur le plan morphosyntaxique, Dixon (2003) opère une distinction entre démonstratifs nominaux, adverbiaux et verbaux. Tandis que les deux premiers sont bien attestés dans les langues du monde, les troisièmes seraient plutôt rares (leur sens peut être rendu par des paraphrases du type *être là*). Au sujet des démonstratifs nominaux, Dixon (2003 : 63) constate que :

⁶⁰ Evidemment, rien n'empêche qu'elles soient également employées pour des entités animées (humaines) avec, typiquement, un effet péjoratif pour résultat. La forme pronominale neutre *ce* doit être traitée complètement à part, car elle apparaît dans la construction présentative permettant de (re)catégoriser le référent, quelle que soit sa nature (voir la Section 1.2.2 *supra* et la Section 1.3.3 *infra*).

In most languages nominal demonstratives can (1) make up a complete NP; and (2) occur in an NP with a noun or a personal pronoun. Sometimes different – but related – forms are used for (1) and (2) but **generally the same forms are employed**⁶¹.

Alors que le tchèque est conforme à ce constat, le français s’y oppose. Dans les Chapitres 3 et 4, nous étudierons de façon détaillée les deux systèmes, avec, chaque fois, une section consacrée exclusivement à la syntaxe du démonstratif. Pour l’instant, contentons-nous de dire que le français connaît un double paradigme de démonstratif nominal, dont une série de formes réservée aux emplois de déterminant et une autre série réservée aux emplois pronominaux, par opposition au tchèque, où les formes peuvent être employées avec ou sans nom⁶².

La richesse du paradigme morphologique du système diffère, bien évidemment, d’une langue à l’autre. En tchèque, langue qui se caractérise par une flexion nominale extrêmement développée (trois genres grammaticaux et sept cas), le démonstratif nominal présente une variabilité morphologique plus grande qu’en français. Tandis que dans les deux langues, ce dernier est morphologiquement marqué pour le genre et le nombre, en tchèque, il est aussi marqué pour le cas.

Aux trois types de démonstratifs reconnus par Dixon, Diessel en ajoute un quatrième, à savoir les démonstratifs identificatoires (“the identificational demonstrative”). Ces démonstratifs apparaissent dans les constructions présentatives servant à identifier ou à (re)categoriser un référent. Diessel (1999 : 27) écrit que :

Like adnominal demonstratives, identificational demonstratives are often morphologically invariable in languages in which pronominal demonstratives are marked for gender, number and/or case.

Ce constat vaut pour nos deux langues⁶³, où les formes identificatoires *ce* et *to* (éventuellement *toto/tohle* et *tamto*), ne varient jamais, restant toujours au nominatif singulier du genre neutre.

Disons à présent un mot sur le comportement syntaxique des démonstratifs nominaux. Comme le résume la citation de Dixon (2003) ci-dessus, les démonstratifs nominaux peuvent soit introduire un nom (éventuellement un pronom) dans le discours soit fonctionner comme des pronoms, à savoir de manière autonome, ayant la même distribution qu’un complexe nominal. Néanmoins, ce qui apparaît comme controversé dans cette citation, c’est le fait de considérer que de tels complexes nominaux (que le démonstratif y introduise un autre élément ou qu’il apparaisse seul) soient automatiquement analysables comme des SN. Comme nous avons déjà pu le suggérer, nous ne sommes pas d’accord avec cette analyse pour la raison suivante : les SN tout comme la catégorie des

⁶¹ C’est nous qui soulignons.

⁶² Cependant que, comme nous le verrons également dans le Chapitre 3 (Section 3.1.4), celles parmi les formes appartenant au genre neutre sont parfois traitées à part et encouragent des analyses en termes de polysémie.

⁶³ À ceci près que les formes employées de manière adnominale (“adnominal demonstratives”) n’y sont pas, comme le suggère Diessel, morphologiquement invariables.

déterminants existent uniquement dans les systèmes contenant un article défini (Himmelman, 1997). Diessel (1999) souscrit à ce parti-pris (que nous nous appliquerons à détailler dans la Section 3.1.3 du Chapitre 3 à propos du statut syntaxique des démonstratifs tchèques en emplois adnominaux), le résumant ainsi :

[...] there are languages in which adnominal demonstratives are independent pronouns that are adjoined to a coreferential noun in apposition, and there are other languages in which adnominal demonstratives are determiners, which require a cooccurring noun. (Diessel, 1999 : 65)

Pour l'instant, nous nous contenterons de constater que cette opposition entre, d'une part, des démonstratifs adnominaux fonctionnant comme des pronoms apposés au nom qu'ils introduisent et, d'autre part, des démonstratifs adnominaux ressortissant à la classe des déterminants semble parfaitement reflétée par l'opposition entre le tchèque et le français.

1.3.4 Le démonstratif dans une perspective diachronique

Au début de ce chapitre, nous avons vu les principales différences qui opposent, sur le plan du fonctionnement sémantico-référentiel, les démonstratifs aux descriptions définies. Toutefois, étant donné que les démonstratifs représentent la source la plus commune des articles définis dans les langues du monde (cf. Christophersen, 1939 ; Krámský, 1972 ; Greenberg, 1978a ; Himmelman, 1997 et 2001 ; Diessel, 1999 ; C. Lyons, 1999 ; Carlier et De Mulder, 2011), l'on ne saurait éviter la question de la nature de la relation entre ces deux unités linguistiques. Autrement dit : faut-il plutôt parler d'une différence de degré ou d'une différence de nature ? En réalité, la réponse serait déjà contenue dans la question elle-même et dépendrait de la distinction entre les plans synchronique et diachronique.

Pour ce qui est du premier, il semble aujourd'hui plus courant d'admettre que les deux unités sont séparées par une différence de nature. Dans l'espace francophone, Kleiber (cf. 1983b, 1984a, 1986a et 1986d) compte parmi les partisans de cette ligne, qui insiste sur la nature logico-sémantique irréductible de l'ensemble de la classe des déictiques. Toujours dans l'espace francophone, l'hypothèse opposée, dite « réductionniste », est représentée par Bonnard (1972) et Wilmet (1986, 2010). Dans le système de la détermination nominale proposé par ce dernier, les déterminants démonstratifs français ressortissent à la catégorie des « quantifiants-caractérisants », à savoir des termes dont le rôle est à la fois d'indiquer l'extensité et de restreindre l'extension d'un nom (1986 : 477). Les types d'arguments que l'auteur avance pour étayer sa thèse sont au nombre de quatre, à savoir : historiques – l'origine de l'article défini ; formels – la symétrie morphologique *ce/le//ces/les* ; distributionnels – les deux formes s'excluent mutuellement au sein d'un SN tout en étant

partiellement interchangeables ⁶⁴; et, enfin, sémantiques – la possibilité de traduire une expression démonstrative par une expression définie assortie d'un complément déictique :

[...] les « adjectifs démonstratifs » *ce* (*cet* devant voyelle), *cette*, *ces* seuls ou combinés aux particules *-ci/-là*, cumulent la fonction quantifiante de *le*, *la* ou *les* avec un qualifiant Δ disponible (Δ 'delta' est le symbole du vide en attente de complétude), sur la valeur duquel on s'interrogera. (1986 : 477)

Un angle d'attaque contre cette thèse repose justement sur la nature et la récupérabilité de cet élément déictique restricteur d'extension. En effet, celles-ci ne vont pas toujours de soi. Nous pourrions à la limite soutenir, comme le fait Wilmet, que dans une reprise anaphorique, l'élément Δ s'apparente à un commentaire du type *que je viens de mentionner/en question*, etc. Cela expliquerait pourquoi, si le français éprouve souvent des réticences face à la présence du défini dans la reprise immédiate d'un référent en anaphore fidèle (cf. Corblin, 1983 ; Kleiber, 1986b), ce problème serait résolu par l'association du défini avec un commentaire en expansion caractérisante :

L'interdiction ne concerne que le nom privé de caractérisation : *UN homme entra dans la pièce. J'avais déjà vu L'homme EN QUESTION/ PRECITE*, etc. redevient légitime. (Wilmet, 1986 : 491)

La question reste ouverte pour nous de savoir dans quelle mesure une telle « équation » est naturelle et défendable, étant donné que contrairement au reste de l'énoncé, le commentaire se situe davantage au niveau du dire (de l'acte de l'énonciation lui-même) qu'au niveau du dit. Un autre argument présenté par Kleiber (1983 : 108) pour invalider cette position est le suivant : une expression comme *cet arbre* ne saurait être ramenée à *l'arbre que je te montre*, car si la proposition *Cet arbre est l'arbre que je te montre* « [...] est certainement vraie, elle n'est pas pour autant nécessairement vraie, puisque j'aurais fort bien pu désigner un autre arbre. »

En diachronie, en revanche, les deux unités doivent être rapprochées dans la mesure, précisément, où les démonstratifs représentent la source principale des articles définis dans les langues du monde. Le processus de grammaticalisation à l'origine de cette catégorie a fait l'objet de nombreuses recherches, et, force est de constater que si ses grandes lignes sont aujourd'hui relativement bien décrites, il continue toutefois à receler plus d'un mystère (nous verrons cela dans le Chapitre 6). Dans la majorité des cas, le défini serait le produit d'un démonstratif dont le sémantisme déictique est allé vers un affaiblissement progressif :

⁶⁴ Cette interchangeabilité est, bien entendu, limitée à certains contextes. Et même lorsqu'elle est possible, elle semble toujours s'accompagner d'une différence sur le plan sémantico-référentiel, contrairement à ce que dit Bonnard, pour qui « [...] l'adjectif *ce* est souvent substituable à *le* sans différence de sens » (1972 : 1207).

Even though the grammatical category of articles as such is far from being universal, the grammaticalization process that leads to its development exhibits cross-linguistic regularities: in the majority of cases, the definite article originates from a weakened demonstrative [...]. (Carlier et De Mulder, 2011 : 523)

Pour les langues romanes, dont la majorité ont développé un paradigme d'articles définis à partir de la forme du démonstratif latin de la troisième personne ILLE⁶⁵, le processus a fait l'objet de nombreuses études (cf. Aebischer, 1948 ; Buzzoni, 2012 ; Caboli, 1990 ; Carlier et De Mulder, 2010 ; Epstein, 1994 ; Fruyt, 2003 ; Marchello-Nizia, 2009 ; Orlandini, 1981 ; Selig, 1992). Les langues (et dialectes) slaves occidentales, où il est aujourd'hui possible d'observer la naissance d'un article défini, utilisent la forme neutre du démonstratif (TEN en polonais, en tchèque et en slovaque et TYN en haut-sorabe ; cf. Czardybon, 2017). Le cadre théorique que nous trouvons le plus adapté pour évaluer le degré de grammaticalisation d'un démonstratif est celui proposé par Löbner (1985 et 2011) et employé, depuis, par toute une série de chercheurs (cf. Carlier et De Mulder, 2006 et 2011 ; Czardybon, 2017 ; Gerland et Horn, 2010 ; Himmelmann, 1997 et 2001 ; Hofherr et Zribi-Hertz, 2014 ; Ortmann, 2014). Nous le présenterons en détail dans le Chapitre 6, lorsque nous traiterons de la grammaticalisation de TEN dans le registre parlé informel du tchèque contemporain.

Au passage, il convient également de souligner que l'article défini est loin d'être l'unique produit du démonstratif en diachronie. Diessel (1999 : 115) emploie à l'égard du démonstratif le terme de « polygrammaticalisation », terme désignant le développement de plusieurs unités grammaticales à partir d'une seule source. En effet, c'est une chose bien connue que le démonstratif représente, dans les langues du monde, une source courante d'une série de « catégories mineures » (Marchello-Nizia, 2009) : « [...] demonstratives provide a common historical source for a wide variety of grammatical items. » (Diessel, 1999 : 153-154) Concrètement :

Pronominal demonstratives are frequently reanalyzed as third person pronouns, relative pronouns, complementizers, sentence connectives, pronominal determinatives, possessive pronouns, and verbal number markers. Adnominal demonstratives may develop into definite articles, linkers, nominal number markers, adnominal determinatives, specific indefinite articles, and boundary markers of postnominal attributes. Adverbial demonstratives provide a common source for temporal adverbs, directional preverbs, and sentence connectives. Finally, identificational demonstratives may evolve into copulas, focus markers, and expletives.

(Diessel, 1999 : 7)

1.4 Résumé

Ce chapitre se voulait une introduction au sujet des démonstratifs. Dans la Section 1.1, nous avons fait le point sur la notion de deixis et les grandes approches existant en la matière. C'est là que nous avons exprimé notre préférence pour le modèle articulé autour de la notion de « token-

⁶⁵ À l'exception du sarde et de certaines variétés du catalan.

réflexivité », tout en relevant cependant les aspects utiles des autres approches. Dans la Section 1.2, nous avons présenté les singularités sémantico-référentielles du démonstratif. Nous l'avons fait en grande partie en nous appuyant sur la littérature francophone, où le sujet est appréhendé d'une manière assez conséquente, entre autres parce que le français possède un paradigme d'articles définis, auquel les démonstratifs sont souvent comparés. Nous étions partis du constat que le démonstratif est un *désignateur* contextuel comportant le trait + Défini, sans oublier de rappeler une série d'autres traits qui, pour leur part, le rapprochent davantage de l'indéfini. Dans cette optique, nous avons soutenu que l'unicité du référent véhiculée par une expression démonstrative doit être relativisée. En effet, bien qu'en employant le démonstratif, le locuteur entende rendre possible, pour l'interlocuteur, l'identification d'un seul référent, celui-ci n'est pas, contrairement à la référence moyennant le défini, le seul représentant de sa classe dans la situation. Après avoir reconnu, à l'instar de Ducrot et de Kleiber, la nature intrinsèquement référentielle du démonstratif, nous avons interrogé son rapport au présupposé existentiel, tout comme cette notion en elle-même. Nous avons souligné le rôle primordial du contexte dans le fonctionnement du démonstratif. Celui-ci non seulement permet d'expliquer pourquoi le démonstratif saisit son référent moyennant un contraste interne, mais il rend également compte de la possibilité pour lui de (re)categoriser le référent, en particulier en anaphore infidèle ou résomptive. L'incomplétude référentielle du démonstratif, due à son sémantisme indexical opaque, est comblée par quatre types d'indices, que nous avons examinés à tour de rôle.

Dans la dernière section (Section 1.3), nous avons vu que le démonstratif représente une catégorie universelle, quoique très variée dans ses formes et ses sens. L'importance de la distance physique, spatiale, souvent perçue comme le critère le plus pertinent pour la description des systèmes du démonstratif dans les langues du monde, doit être relativisée⁶⁶, de même que, d'ailleurs, la capacité de rendre des oppositions de distance tout court, qui ne saurait être tenue pour une condition nécessaire dans la définition de la catégorie. Des constats portant sur la morphosyntaxe du démonstratif ont été rappelés, dont le statut syntaxique différent des formes dans les langues selon qu'elles possèdent ou non la catégorie des déterminants. En nous intéressant plus avant à la relation entre le démonstratif et le défini, nous avons souligné l'importance de la dissociation, sur ce point, entre la synchronie et la diachronie, rappelant également que le démonstratif constitue, à l'échelle universelle, une source privilégiée d'un grand nombre de termes grammaticaux.

⁶⁶ C'est la raison pour laquelle nous avons insisté sur la possibilité d'une lecture davantage abstraite, pragmatique du terme.

2 Présentation des corpus et de la méthodologie du travail avec les données

Après avoir introduit le lecteur au sujet du démonstratif (Chapitre 1) et avant de faire le bilan de la situation des démonstratifs adnominaux en tchèque et en français (Chapitres 3 et 4), nous consacrerons ce chapitre à la présentation de nos corpus ainsi qu'à la description détaillée de notre traitement et de notre exploitation du matériel empirique. Dans un premier temps, nous présenterons les corpus pour ensuite exposer notre méthodologie d'analyse. Dans ce cadre, le système de nos annotations sera présenté. Pour finir, nous décrirons notre procédure d'échantillonnage des résultats obtenus à partir des requêtes appliquées aux corpus de langue orale informelle.

2.1 Présentation des sources de données utilisées et des informateurs sollicités

Ce premier volet sera consacré aux sources auxquelles nous avons recouru pour réaliser le présent travail et qu'il est possible de répartir en deux catégories : 1) les sources primaires et 2) les sources secondaires. Se trouvent, dans la première catégorie, les corpus de langue parlée informelle, à savoir ORTOFON v1 de la plateforme du Corpus national tchèque (CNT) et le Corpus d'Études pour le Français Contemporain (CEFT) de la plateforme Orféo, mais aussi le corpus parallèle multilingue InterCorp de la plateforme du CNT. Dans les Sections 2.1.2 et 2.1.3, il s'agira à la fois de donner une description à grands traits de ces trois sources et d'exposer la façon particulière dont nous les avons ajustées pour les objectifs de notre travail. Sous la seconde bannière se trouvent regroupés tous les autres corpus – notamment écrits – et sources de données que nous avons consultés de manière complémentaire, mais, en plus de cela, également une série d'informateurs natifs en tchèque et en français, autant de « ressources humaines », auprès de qui nous avons pu vérifier certains constats et intuitions au sujet de l'emploi du démonstratif dans les deux langues. Ces sources et ces informateurs seront présentés, respectivement, dans la Section 2.1.4 et la Section 2.1.5. Enfin, dans la Section 2.1.6, nous exposerons les méthodes statistiques de traitement des données auxquelles nous avons recouru dans le but d'étayer nos analyses quantitatives. Mais avant même de faire tout cela, nous tenons, dans un tout premier temps, à élucider certains choix terminologiques ayant trait à la typologie textuelle.

2.1.1 Types de textes, registres, genres... quels choix terminologiques ?

Il existe une pluralité d'approches et de parti-pris dans le domaine de l'étude et du classement des textes. Cette pluralité apparaît comme particulièrement frappante, voire irréconciliable, lorsqu'il est question de plusieurs traditions dont chacune est issue d'un milieu linguistique et savant différent. Dans notre cas, il est évidemment question des milieux intrinsèquement liés aux deux langues faisant l'objet de la présente étude, ce à quoi s'ajoute une certaine prégnance de la tradition anglosaxonne. Dans une telle conjoncture, que faire pour ne pas sombrer dans un chaos terminologique ? S'il est inévitable de faire des choix, ceux-ci devraient, à notre avis, entrer autant que faire se peut en résonance avec la nature du matériel empirique exploité. C'est donc à la fois au profil des corpus utilisés et aux choix terminologiques opérés par leurs créateurs que devraient s'adapter nos propres choix. Parallèlement à cela, il nous semble impossible d'éviter un certain biais terminologique, qui, dans notre cas personnel, procède de ce que notre formation en la matière a eu lieu dans le milieu tchèque.

En effet, celui-ci peut être fier d'une longue et fructueuse tradition en matière de stylistique et typologie textuelle, développée notamment au sein du paradigme structuraliste du Cercle linguistique de Prague. Au sein de cette tradition se trouve une approche fonctionnelle („funkční stylistika“), articulée autour de la notion de « style fonctionnel » :

Le concept de départ de la tradition stylistique tchèque est la fonction du texte/discours en tant que facteur objectif de genèse de style (le choix des moyens linguistiques et leur organisation dans un agrégat plus complexe obéissent à l'effort de remplir cette fonction de manière adéquate), donnant lieu à la définition du style fonctionnel¹. (Cvrček *et al.*, 2020 : 18)

Les styles fonctionnels fondamentaux traditionnellement reconnus sont : le style de la communication ordinaire („prostěsdělovací“), le style professionnel („odborný“), le style journalistique („publicistický“) et le style artistique („umělecký“ ; cf. Hausenblas, 1955 et 1991 ; Hoffmannová et Homoláč, 2016). Nous sommes là face à une taxinomie reposant essentiellement sur une perspective interne, à savoir les propriétés affichées par les textes, lesquels textes, répétons-le, obéissent à une (éventuellement à plusieurs) fonction de l'acte de communication. Afin de réaliser leur but communicationnel, les styles fonctionnels recourent à un ensemble de procédés appelés « procédés de style » („stylové postupy“), tels que la description, la narration, l'exposition, l'argumentation, etc.

¹ Nous traduisons : „Výchozím konceptem české stylistické tradice je funkce textu/projevu jako objektivního stylového faktoru (výběr jazykových prostředků a způsob jejich spojování ve vyšší celek jsou vedeny snahou tuto funkci adekvátně naplnit), na jehož základě vymezuje funkční styl [...]“

Ces procédés ne sont pas sans rappeler la typologie textuelle développée par Adam (cf. 1997, 2005 [1992] et 1999), qui reconnaît l'existence de cinq types textuels appréhendés en des unités appelées « séquences ». Le linguiste de Lausanne distingue ainsi cinq types de séquences : narrative, descriptive, argumentative, explicative et dialogale (2005 [1992] : 30). Chaque séquence est caractérisée par un certain nombre de traits prototypiques et elle est composée d'unités atomiques appelées « propositions énoncés », elles-mêmes regroupées en « macro propositions » (2005 [1992] : 35-43). L'auteur insiste par ailleurs sur le fait que s'il est possible de parler de « types textuels », il n'en va pas de même pour les « textes » tout court, car « [l']unité “texte” est trop complexe et trop hétérogène pour présenter des régularités linguistiquement cernables. » (1997 : 665)². Dans tous les cas, l'approche d'Adam nous semble avoir pour point commun avec la tradition tchèque l'orientation interne des analyses.

La tradition anglosaxonne, quant à elle, a coutume de travailler avec le concept de « registre » (cf. Biber, 2014). Ce concept, relevant également d'une approche orientée vers l'intérieur, c'est-à-dire vers les propriétés spécifiques des textes, correspond à peu près à celui de « procédé de style » (voire « style fonctionnel » ; Cvrček *et al.*, 2020 : 19) dans la tradition tchèque et à celui de « type textuel » dans le système de la linguistique textuelle adamienne.

Vus à travers le prisme externe, les textes se répartissent en « genres », terme employé de façon systématique et doté d'une acception très proche dans les deux (voire trois) traditions. Hoffmannová et Homoláč (2016 : 156) définissent les genres comme des unités textuelles « [...] figées, obéissant à des normes plus ou moins strictes, voire codifiées »³. Biber et Conrad (2009 : 7) précisent que si les genres sont eux aussi caractérisés par la présence d'une série de traits, ces traits sont rarement omniprésents, ayant davantage tendance à apparaître de manière ponctuelle :

Genre features are not pervasive; rather, they might occur only one time in a complete text, often at the beginning or ending boundary. They are also often conventional rather than functional.

Les créateurs du CNT appliquent un système de classement des textes qui opère la distinction entre plusieurs niveaux hiérarchiques, dont le genre représente le niveau le plus bas. Ce système est bien plus complexe dans le cas des corpus écrits que dans celui des corpus parlés. Dans les corpus écrits de la série SYN, les auteurs distinguent d'abord le niveau appelé « groupe de types de textes » (“text-type group”), suivi du niveau « type de textes » (“text-type”), sachant que les premiers sont au nombre de trois : la fiction, la non-fiction/littérature spécialisée („oborová literatura”⁴) et la presse.

² La même réserve est émise par Cvrček *et al.* (2020 : 128), qui attirent tout particulièrement l'attention sur l'hétérogénéité propre aux textes de fiction (ainsi, dans un roman, la narration cède tantôt à la description, tantôt au dialogue, etc.).

³ Nous traduisons : „[...] ustálené, různě striktně normované či přímo kodifikované.”

⁴ Les termes de « non-fiction » / « littérature spécialisée » ont été choisis par nous à défaut d'un meilleur équivalent de traduction. Cette catégorie englobe des textes spécialisés dans un domaine particulier, de nature philosophique,

Le deuxième niveau (« types de textes ») introduit un classement plus fin. Ainsi, le groupe de types de textes « fiction » contient des types de textes tels que « roman », « poésie », « pièces de théâtre », etc. Le groupe de types de textes « littérature spécialisée » contient les types de textes « littérature scientifique », « littérature professionnelle » et « littérature populaire »⁵. Il apparaît donc que la notion de « type de textes » relève, elle aussi, de la perspective externe du classement des textes. En cela même, elle entre en dissonance avec la perspective adamienne, où l'unité appelée « type textuel » correspond à un classement interne (voir *supra*). Restant plus proche ici de la tradition tchèque et des choix des créateurs du CNT, nous utiliserons dans l'ensemble de ce travail le terme de « type de textes » pour référer aux différents textes avec lesquels nous travaillerons (conversations, entretiens, fiction, textes de presse, textes de littérature spécialisée, etc.)⁶. Ce terme nous paraît également le plus à même d'éviter d'éventuelles confusions terminologiques.

À présent, focalisons-nous davantage sur le terme de « conversations », tout à fait crucial pour notre travail. Pour mieux circonscrire ce terme, nous nous rapporterons notamment à la publication *Registry v češtině (Registres du tchèque)* de Cvrček *et al.* (2020). Appliquant l'« analyse multidimensionnelle » de Biber (cf. 2014) à un corpus représentatif de plusieurs types de communication, les auteurs élaborent une taxinomie de registres en tchèque. Cette taxinomie repose sur une série de dimensions articulées en oppositions binaires, dont les plus importantes sont celles entre le dynamique vs le statique, ainsi que le spontané vs le préparé. Dans cette optique, la « conversation » („konverzace“) est un registre dynamique et spontané, caractérisé par un degré élevé d'interaction informelle, riche en mots phatiques, en mots de remplissage et en pronoms démonstratifs. Sur le plan syntaxique, l'on constate un degré de cohésion textuelle moindre qu'à l'écrit (ou dans les productions orales préparées), se manifestant notamment par une présence faible de conjonctions de subordination (c'est la juxtaposition qui sembler primer ; Cvrček *et al.*, 2020 : 126). Du point de vue des « styles fonctionnels », le registre des conversations recourt le plus souvent au style de la « communication ordinaire » (voir *supra*). Pourtant, contrairement à l'usage de la publication de Cvrček *et al.* (2020), le terme de « registre » sera employé, dans le présent travail, dans une acception beaucoup plus vague, à savoir pour référer à l'opposition entre la langue formelle et la langue informelle. Cette opposition fondamentale sera susceptible d'un classement plus raffiné,

scientifique (y compris académique), éducative, des textes de vulgarisation, ainsi que des manuels, des textes administratifs, des textes de nature professionnelle et des mémoires et autobiographies de nature non littéraire. Sur le plan des « styles fonctionnels », il s'agit de textes appartenant au style « professionnel » (voir *supra*). La tradition allemande emploie souvent le terme de « Fachstil » (cf. Berger, 1993).

⁵ Voir [https://wiki.korpus.cz/doku.php/en:cnk:klasifikace_textu_syn2015?s\[\]=text&s\[\]=types](https://wiki.korpus.cz/doku.php/en:cnk:klasifikace_textu_syn2015?s[]=text&s[]=types).

⁶ Tout en prenant cependant certaines libertés avec le terme et en l'employant avec une extension plus large, ce dernier pouvant ainsi s'appliquer à la fois à ce que les créateurs du CNT appellent « groupes de types de textes », « types de textes », voire plus encore. En revanche, ce terme sera toujours entendu dans l'acception qu'il reçoit au sein de la tradition tchèque, c'est-à-dire comme un terme correspondant au classement externe des textes.

et c'est à ce moment-là que nous parlerons de « variétés » des deux langues, telles que, dans le cas du tchèque, les variétés écrite et parlée du tchèque soutenu, le « tchèque commun », les différents dialectes de Moravie et de Silésie, etc. (voir à cet égard le Chapitre 3).

2.1.2 Les corpus de langue orale informelle : le CEFT et ORTOFON v1

Comme il a déjà été annoncé dans l'Introduction générale, deux sources majeures de données empiriques ont été mobilisées dans le cadre de ce projet. Il s'agit, pour la langue française, d'une partie du Corpus d'Études pour le Français Contemporain (CEFC) de la plateforme Orféo et, pour la langue tchèque, de la plateforme du CNT. De cette dernière, nous avons avant tout retenu le corpus de langue orale informelle ORTOFON v1. Le corpus multilingue InterCorp, également présent sur le site du CNT, a été utilisé comme une source additionnelle, et ce notamment pour nous donner un aperçu la traduction en français d'une partie des démonstratifs adnominaux tchèques.

Il sied à présent de faire une description plus détaillée des deux corpus principaux⁷, utilisés pour l'étude de la langue orale informelle. La toute première question qui se pose tout naturellement est la suivante : ces deux corpus sont-ils comparables ? Nous sommes convaincu que, globalement, ils le sont, et ce même si une série de différences importantes les opposent. De celles-ci, les plus « graves » sont sans doute les suivantes : 1) une différence de taille non négligeable et 2) des critères différents relatifs à la récolte et la nature des données, ainsi qu'à l'accessibilité des métadonnées.

Une différence de taille d'abord, car la taille de l'ensemble retenu pour le français ne fait qu'à peu près 75 % de la taille de la totalité des conversations d'ORTOFON v1. Mais ce n'est pas dans cette première disproportion que nous voyons un éventuel point faible de notre projet, et cela pour deux raisons. D'abord, ce n'est pas tant la fréquence absolue des phénomènes étudiés qui compte que leur fréquence relative (« items per million » ; ipm), à savoir le nombre d'occurrences du phénomène relativement à un million de positions (tokens)⁸ du corpus interrogé. Ensuite, notre travail se veut – malgré une série de comparaisons quantitatives nécessaires, notamment présentées dans le cadre du Chapitre 5 – avant tout qualitatif. Compte-tenu de tout cela, il n'est pas impératif de disposer de deux corpus de même taille ; tout ce qui compte, à ce moment-là, c'est que les deux corpus soient d'une taille satisfaisante eu égard au phénomène étudié, de telle sorte que les observations puissent être validées sur une masse de données suffisamment grande. À cet égard, la taille de nos deux corpus

⁷ Afin d'éviter toute confusion terminologique, nous distinguerons dans ce travail entre quatre termes relatifs à quatre niveaux hiérarchiques : « plateforme/base », renvoyant à l'ensemble du projet en question (cf. Orféo et le CNT) ; « corpus », renvoyant aux macro-corpus que sont ORTOFON v1, le CEFC (à la fois dans son intégralité et la partie retenue par nous), InterCorp et d'autres corpus encore ; « sous-corpus », renvoyant aux différentes sources dont est composé le corpus CEFC, ainsi qu'aux sous-corpus définis par nous au sein d'ORTOFON v1, d'InterCorp, etc., et, enfin, « conversation/texte », référant, respectivement, aux plus petites unités dont sont composés les corpus oraux et écrits (cf. les pièces d'un corpus de théâtre, les différentes conversations d'un corpus de conversations).

⁸ Aussi indiquée, pour certains corpus, relativement à un million de mots.

s'avère suffisante, les démonstratifs adnominaux comptant, dans les deux langues, parmi les mots apparaissant avec une fréquence élevée dans les discours parlés informels⁹. Enfin, le recours à des tests statistiques, notamment à la fonction de vraisemblance, permet d'étayer la validité des éléments quantitatifs obtenus (voir la Section 2.1.5 *infra*).

C'est certainement la seconde série de différences qui pose davantage problème, remettant partiellement en cause la nature comparable de nos corpus oraux et, sans doute, de l'ensemble de notre projet. La création d'ORTOFON v1, marquée par la volonté de parvenir à un équilibre aussi parfait que possible, a obéi à une série de contraintes strictes qui ont porté sur les quatre paramètres suivants : l'origine géographique des locuteurs, leur sexe, leur âge et leur niveau de formation. Comme l'expliquent les auteurs du corpus :

From the very beginning of data collection, special care was taken to achieve the maximum possible speaker variability with regard to dialectal regions. Over the course of the collection process, the material was adjusted in order to achieve a balanced corpus within the four basic sociolinguistic categories: gender, age, level of education and the dialectal region in which the speaker spent the majority of the first 15 years of his life. The first three categories, i.e. gender, age, education, were assigned binary values [...], while the fourth category was divided into ten groups i.e. ten dialectal regions. Each region should, therefore, contain the same number of words from men and women, from speakers of ages 18-34 years and those over 35 years, and from speakers with a high school education and those with a university education. [...] The basic concept was the idea of the same proportional representation of the sociolinguistic categories listed above, applied to the collection of material for all of the ČNK spoken corpora. Taking into account the target corpus size (1 000 000 words), the target for every category presented by the combination of four variables – gender (2) × age (2) × education (2) × dialectal region of residence up to the age of 15 years (10) – was set at 12 500 words. In the effort to achieve the highest possible speaker variability within the scope of each category, a minimum of five different speakers was set. The aim of this provision [was] to limit the influence of idiolect¹⁰.

C'est pour son caractère représentatif, mais également pour sa taille, que ce corpus a été choisi, et non pas un autre corpus de tchèque parlé informel de la plateforme du CNT. Outre cette composition très équilibrée, le corpus ORTOFON v1 ne contient – à deux exceptions négligeables près¹¹ – que des conversations informelles entre personnes qui se connaissent, le plus souvent dans le cadre d'une relation familiale, amicale, de couple ou professionnelle. Ces conversations se déroulent dans le régime face-à-face (à la maison, dans le jardin, sur le lieu de travail, à l'école pendant la récréation, etc.) ou par téléphone/sur Skype. Les enregistrements ont été réalisés dans la période 2012-2017¹². Le nombre de locuteurs par conversation est le plus souvent de deux, mais il peut aussi

⁹ Ce constat vaut évidemment davantage pour le tchèque ; néanmoins, même dans le corpus français, le démonstratif adnominal est employé avec une fréquence assez élevée pour nous permettre de comparer la situation dans les deux langues.

¹⁰ Pour plus d'informations, voir <http://wiki.korpus.cz/doku.php/en:cnk:ortofon>.

¹¹ Les corpus 13A087N et 13B028N, qui, pris ensemble, font la taille de 2 221 positions. Le premier (339 positions) est une conversation entre deux hommes inconnus qui se rencontrent dans un arrêt de bus. Le second est une conversation ayant lieu à la maison entre cinq personnes, dont seulement certaines d'entre elles ne se connaissent pas.

¹² Une version élargie contenant également des enregistrements de la période 2012-2019 a été rendue disponible en 2020 sous le nom d'ORTOFON v2. Toutefois, cette version n'est pas équilibrée eu égard aux paramètres susmentionnés et n'était pas disponible au moment où nous avons commencé le présent travail.

s'élever à trois et, dans des cas plutôt rares, à quatre, cinq, voire à six (un seul cas). Le corpus est annoté semi-automatiquement en lemmes et catégories grammaticales par l'outil d'étiquetage morphosyntaxique MorphoDiTa¹³. Cette pratique d'annotation, tout comme l'ensemble du corpus, s'inscrit dans la continuité de la série de corpus de langue parlée informelle ORAL, dont le corpus fait partie. Sa taille est indiquée à la fois en nombre de mots et en nombre de positions (tokens). L'interface du corpus présente les textes des conversations découpés en segments d'une longueur maximale de 25 mots (ou en segments plus courts, si ceux-ci correspondent aux prises de parole). Ces segments sont chacun appariés avec leur piste d'enregistrement, ce qui permet de suivre la voix du locuteur en même temps que la transcription de ses paroles. Le corpus contient trois niveaux de transcription : orthographique, phonétique et métalinguistique. Le premier est constitué par les textes. Le second est composé d'un système de signes de codage diversifié afin de transposer 1) les pauses observées par les locuteurs – dont trois longueurs différentes sont distinguées : celles dont la durée est inférieure à 120 ms, celles dont la longueur se situe entre 120 ms et 2 s et celles dont la longueur dépasse 2 s ; 2) la modalité interrogative, les interjections, les marques de disfluences, les interruptions, les énoncés et les mots inachevés, les passages incompréhensibles, etc¹⁴. Le niveau de transcription dit « métalinguistique » contient toutes les manifestations sonores extérieures à la conversation, qu'elles proviennent des locuteurs eux-mêmes (rire, éternuement) ou de leur entourage (un chien qui aboie, une voiture qui passe). Tous ces signes (y compris ceux du niveau métalinguistique) sont compris dans la taille du corpus exprimée en nombre de positions (tokens). Les chevauchements entre des prises de parole sont également signalés. Sur le site du CNT, il est possible d'accéder aux métadonnées relatives à toutes les conversations¹⁵. Celles-ci mentionnent notamment : la durée et la taille de la conversation (en nombre de positions), le nombre de locuteurs, le contexte de l'interaction (cf. visite privée, récréation à l'école), la relation entre les locuteurs (cf. amis, membres d'une même famille, compagnons, connaissances), l'année, le mois et la ville de l'enregistrement, le sexe, la tranche d'âge, le niveau de formation et le métier des locuteurs, ainsi que la ville et la région où ces derniers ont résidé jusqu'à l'âge de quinze ans et leur lieu de résidence au moment de l'enregistrement (voir la Figure 2.2). Voici d'abord un résumé des principales propriétés statistiques du corpus ORTOFON v1 avec une image de son interface sur le site du CNT :

¹³ L'outil a été développé par M. Straka et J. Straková à l'Institut de linguistique formelle et appliquée de la Faculté de mathématiques et physique de l'Université Charles : <http://ufal.mff.cuni.cz/morphodita>.

¹⁴ Voir <http://wiki.korpus.cz/doku.php/cnk:ortofon:pravidla> (site uniquement en tchèque) pour l'explication de l'ensemble du système de ponctuation.

¹⁵ Voir https://wiki.korpus.cz/doku.php/cnk:struktura_ortofon (site uniquement en tchèque) pour le détail de ces métadonnées, ainsi que les Annexes, où elles figurent traduites en français.

Nom	ORTOFON v1
Nombre de positions (tokens)	1 236 508
Nombre de positions (tokens) sans ponctuation, sons d'hésitation, interjections et sans commentaires (soit le nombre de mots)	1 014 786
Période des enregistrements	2012-2017
Nombre de conversations enregistrées	332
Nombre d'énoncés	172 736
Nombre de locuteurs uniques (différents)	624
Durée des enregistrements [hh:mm:ss.ms]	102:41:14.247

Tableau 2.1 : Propriétés du corpus ORTOFON v1 (CNT)



Figure 2.1 : Interface d'ORTOFON v1 (CNT)

doc.id	130003N	doc.year	2012
doc.month	12	doc.location	Velešovice
doc.situation	návštěva	doc.speakers	3
doc.relationship	partnerský, přátelský	doc.genders	smíšené
doc.generations	1	doc.length	25:49
doc.tokens	3198	sp.id	bab0
sp.nickname	Bronislav_182	sp.gender	M: muž
sp.confederate	ne	sp.age_binary	V: nad 35 let
sp.age	37	sp.edu_binary	A: vysokoškolské
sp.edu_level	VŠ	sp.edu_field	Právo, právní a veřejnosprávní činnost
sp.occupation	kriminalista	sp.occupation_category	42 úředník ve službách a obchodě
sp.reg_childhood	středomoravská	sp.loc_childhood	Rousínov
sp.locsize_childhood	město nad 100 tisíc	sp.reg_longest	středomoravská

Figure 2.2 : Métadonnées du corpus ORTOFON v1 (CNT)

Le corpus CEFC, quant à lui, est né de la fusion de quinze corpus déjà existants et obéissant, tels des projets indépendants, aux règles de leurs créateurs. Guidé par la nature de notre projet, nous avons établi un certain nombre de contraintes (voir *infra*) qui nous ont permis de filtrer cet ensemble, contenant, entre autres, de nombreux textes littéraires et de presse, pour n'en retenir que des productions appartenant à seulement neuf de ces corpus (voir *infra*). Les concepteurs du CEFC ont visé la mise en place d'un corpus reflétant les pratiques de la langue française contemporaine dans une diversité de formes et de types d'emploi :

La plate-forme ORFEO (Outils et Ressources sur le Français Écrit et Oral) propose un ensemble unique de textes écrits et de transcriptions d'enregistrements en français contemporain ainsi que des outils pour les exploiter, que ce soit dans le cadre de la recherche en linguistique, en traitement automatique des langues, pour l'enseignement ou pour simplement satisfaire sa curiosité à propos du fonctionnement du français contemporain : emploi d'un mot, d'une expression ou d'une forme grammaticale (fréquence, contexte, exemples à l'écrit ou à l'oral [...]). L'orientation principale a été néanmoins de se placer du point de vue du linguiste qui veut étudier les formes de l'énoncé et de ses composantes à travers les divers usages de la langue française¹⁶.

De par sa nature très ambitieuse, l'initiative qu'est le CEFC tâche sans doute de répondre à un besoin à nos yeux assez pressant du monde francophone contemporain : de disposer finalement d'un

¹⁶ Pour plus d'informations, voir <https://repository.ortolang.fr/api/content/cefc-orfeo/4/documentation/site-orfeo/home/index.html>.

corpus de langue de référence digne de ce nom. À cet égard, les outils à disposition des chercheurs des deux sphères linguistiques ne sont malheureusement pas comparables. La plateforme du CNT, disposant d'une variété considérable de corpus à la fois écrits et oraux et sans cesse élargie de nouvelles données, ne trouve aucune véritable contrepartie dans le monde linguistique francophone. Nous voyons dans le CEFC une première tentative de combler ce manque. Cependant, on est encore loin du niveau de perfection atteint par l'Institut du CNT. Si nous avons donc décidé de travailler avec le CEFC, c'est que ce corpus s'est imposé à nous comme la seule option dans la conjoncture actuelle¹⁷.

Fort heureusement, l'interface du CEFC permet de trier l'ensemble des sous-corpus et des conversations le constituant en fonction de plusieurs critères (voir la Figure 2.3). Les critères de tri que nous avons appliqués afin d'obtenir un corpus conforme à notre objectif de recherche ont été les suivants : nous n'avons retenu que des « conversations » (« type ») du « secteur privé ». Cela nous a permis de retenir uniquement des conversations du « milieu » « amical » et « familial », soit 178 conversations au total, se déroulant au régime face-à-face ou par téléphone. Ainsi, comme le lecteur peut le constater, les données pour les deux langues sont comparables de par leur nature (informelle, spontanée). Outre les propriétés mentionnées *supra*, la masse des métadonnées renseignées diffère en fonction de la plateforme d'origine des conversations. Si pour certaines, nous connaissons la situation de l'enregistrement (face-à-face, au téléphone), la durée (mais pas la taille), le nombre de locuteurs, le contexte de l'interaction, la relation entre les locuteurs, la date de l'enregistrement (le lieu étant renseigné pour l'ensemble de la plateforme d'origine), le sexe, la tranche d'âge (très approximative), le niveau de formation, la ville et la région de naissance des locuteurs (voir la Figure 2.4), pour d'autres conversations, ces informations ne sont que parcellaires ou font complètement défaut.

¹⁷ Il existe, bien entendu, d'autres corpus de conversations informelles, mais ceux-ci sont, pour la plupart, très limités en taille ou seulement en cours de développement à l'heure actuelle. Une liste de corpus de langue parlée, incluant aussi certains des sous-corpus du CEFC, se trouve par exemple sur le site <https://perso.atilf.fr/apotheloz/corpus/>. Il s'avère que le seul « concurrent » possible du CEFC en matière de taille aurait pu être le Corpus de français parlé au Québec (cf. Dostie, 2016). Cependant, ce corpus, en plus d'être plus petit (712 300 mots), est aussi – comme son nom le suggère – moins représentatif eu égard aux différentes variétés diatopiques du français. Et – surtout –, il n'est annoté ni en lemmes ni en catégories morphosyntaxiques.

Délimiter votre recherche

Corpus >

Statut de l'annotation >

Type ▾

conversation ✕ 176

Secteur ▾

privé ✕ 176

Milieu ▾

amical 157

familial 14

Vous avez demandé : Secteur > privé ✕ Type > conversation ✕

« Précédente | 1 - 10 sur 176 échantillons | Suivante »

1. CLAPI (O) > [jeux_video_foot_Lyon](#)
2. CORALROM (O) > [ffamcv01](#)
3. CORALROM (O) > [ffamcv02](#)
4. CORALROM (O) > [ffamcv03](#)
5. CORALROM (O) > [ffamcv04](#)

Figure 2.3 : Interface du CEFC (Orféo) : choix des textes

Métadonnées : general	
Corpus	TUFS
Nom du fichier	fr05-1_2005_07_04
Responsable(s)	Kawaguchi
Résumé	retour à la maison après une journée de travail
Date de l'enregistrement	04/07/2005
Durée de l'enregistrement	00:11:58
Nature du signal	audio
Qualité du son	environnement bruité
Anonymisation du signal	script Daniel Hirst
Niveaux d'annotation	Annotation automatique
Annotation	automatique
Type	conversation
Secteur	privé
Interaction en milieu ...	familial
Modalité	oral
Nombre de locuteurs	2
Situation de l'enregistrement	face_à_face
Adresse d'échantillon	/annis-sample/tufs/fr05-1_2005_07_04.html

Métadonnées : locuteur L1	
Identifiant du locuteur	L1
Âge du locuteur	21-60
Sexe du locuteur	M
Profession du locuteur	étudiant
Niveau d'études du locuteur	études supérieures
Lieu de naissance du locuteur	France, PACA, Marseille

Figure 2.4 : Métadonnées des conversations du CEFC (Orféo)

Comme déjà écrit *supra*, ces conversations proviennent de neuf sous-corpus différents (voir le Tableau 2.2). De cet ensemble, nous avons par la suite exclu deux conversations, dont les participants ne sont pas des locuteurs natifs francophones. Deux autres conversations de l'ensemble contiennent également une locutrice étrangère chacune ; cependant, cela ne nous a pas empêché de les garder, car la part de ces locutrices à l'ensemble de la conversation est négligeable (il s'agit des

corpus Kathy et Gaël du corpus TUF¹⁸). Voici les sous-corpus retenus pour notre travail avec, pour chacun, le nombre de conversations et la taille de l'ensemble en positions :

Sous-corpus	Nombre de conversations	Taille en positions (tokens)
TUFS	50	668 421
TCOF	62	132 098
C-ORAL-ROM	37	48 357
VALIBEL	8	38 947
OFROM	11	30 364
CLAPI	1	7 770
CRFP	2	4 760
Réunions de travail	3	2 528
Fleuron	2	857
Nombre total	176	934 102

Tableau 2.2 : Sous-corpus du CEFC (Orféo) retenus pour notre corpus de français oral informel avec, pour chacun, le nombre de conversations et le nombre de positions de l'ensemble

L'écrasante majorité des conversations ont été enregistrées dans la période 2000-2012, sauf deux conversations, datant de 1989 et 1999, et cinq corpus pour lesquels la date d'enregistrement est inconnue. Le nombre de locuteurs varie entre un et sept, hormis un cas exceptionnel, où dix-huit locuteurs (sic) apparaissent sur scène (PRO-NAR-1). Contrairement à ORTOFON v1, on ne peut ici parler d'un ensemble équilibré pour les paramètres susmentionnés relatifs aux locuteurs, car une bonne partie des sous-corpus privilégie clairement les locuteurs d'un certain type. Nous ne dirons ici qu'un mot sur les cinq plateformes les plus représentées (pour plus de détails, nous renvoyons le lecteur au site Internet du projet).

Une source majeure de notre matériel empirique pour l'oral informel français sont les conversations françaises privées de la base TUF¹⁸ du Center of Corpus-based Linguistics and Language Education de l'Université de Tokyo, enregistrées dans les universités d'Aix-Marseille, de Paris et de Bordeaux. Il s'agit, pour la plupart, de dialogues entre étudiants universitaires de la même formation (voire de la même promotion), éventuellement aussi entre collègues enseignants, membres du personnel universitaire, etc. Le cadre conversationnel est en partie imposé : très souvent, les enregistrements ont lieu dans une bibliothèque universitaire, des sujets de conversation sont suggérés aux participants en amont. Deux conséquences découlent de cela. D'abord, les

¹⁸ Les prises de parole des deux locutrices étrangères ne contiennent, respectivement, que quatre et une seule occurrence(s) de démonstratifs adnominaux. De plus, lors du travail d'échantillonnage, ces occurrences n'ont pas été incluses dans l'échantillon retenu pour les analyses qualitatives (voir la Section 4.2.4 *infra*).

conversations tendent à se ressembler : vacances d'été, études universitaires (notamment de japonais), voyages, culture japonaise, etc. Ensuite, elles manquent parfois de spontanéité et de fluidité : de temps en temps, les participants éprouvent des difficultés à trouver des sujets, à poursuivre la conversation ; des références au cadre de l'enregistrement apparaissent parfois (les locaux, la présence du dictaphone).

Les conversations empruntées à la base TCOF (Traitement de Corpus Oraux en Français) sont des conversations entre adultes enregistrées dans plusieurs régions françaises. La base C-ORAL-ROM est un projet de corpus comparables de plusieurs langues romanes, financé par l'Union européenne et élaboré sous la direction de l'Université de Florence. La plupart des locuteurs des conversations retenues sont originaires de la région PACA. La plateforme Valibel (Discours et Variation) est composée, elle, uniquement de conversations de locuteurs francophones belges (des régions Wallonie et Bruxelles-Capitale). Et, finalement, la base OFROM (Le Corpus Oral de français de Suisse Romande) contient exclusivement des productions de locuteurs francophones suisses¹⁹.

Mis à part les points faibles déjà mentionnés, dont l'absence de métadonnées sur une bonne partie des locuteurs et sur une partie des années de l'enregistrement tout comme l'absence de chiffres sur la taille des conversations – que nous avons donc dû établir nous-même, séparément pour chaque texte (aussi étonnant que cela puisse paraître²⁰) –, voici d'autres limites de ces sources. 1) Certaines conversations suggèrent une préparation partielle du sujet en amont et/ou comportent des éléments de langage formel. Ainsi, le sujet (« l'amour ») du corpus ffamcv01 de la base ORFROM semble partiellement préparé à l'avance et l'ensemble de la conversation est organisé en débat ; pourtant, ce débat – à bâtons rompus – reste en grande partie improvisé et inscrit dans le registre conversationnel informel. Le corpus styHC1r de la base Valibel représente une visite de musée, contenant des éléments de discours formel institutionnel avec de la terminologie spécialisée. Les enregistrements unine08a18m et unine11a04m du corpus OFROM sont des monologues (seul un locuteur est présent) ; pourtant, si nous avons décidé de les garder, c'est que ces monologues sont fortement dialogisés : il est évident que la personne s'adresse à un interlocuteur tacite, présent dans la situation. 2) Certaines conversations ont une taille très, voire extrêmement petite (dont la plus petite ne compte que 147 positions). 3) Certaines conversations sont liées entre elles, d'où la difficulté d'établir le nombre exact de locuteurs différents (uniques).

La plateforme Orféo permet d'interroger les données de deux façons : via les interfaces appelées, respectivement, « recherche simple » et « recherche avancée ». Nous avons recouru

¹⁹ Pour plus de détails concernant les plateformes d'origine, voir <https://repository.ortolang.fr/api/content/cefc-orfeo/4/documentation/site-orfeo/home/index.html>.

²⁰ Par conséquent, les chiffres établis pour les tailles des corpus peuvent comporter des erreurs.

essentiellement à la première d'entre elles, et ce non seulement pour travailler avec les textes et les enregistrements dans le mode en ligne, mais aussi pour accéder aux métadonnées. En ce qui concerne le mode de transcription, celui-ci a été unifié pour tous les sous-corpus. Les enregistrements sont simultanément consultables sous la double forme sonore et graphique ; les tours de parole sont marqués, mais l'ensemble de l'enregistrement n'est pas segmenté : il constitue un bloc. Les annotateurs utilisent beaucoup moins de signes de ponctuation que ceux d'ORTOFON v1 : seules les interjections et les disfluences sont marquées (y compris les mots inachevés), mais pas les pauses, la modalité interrogative, les répliques inachevées, les bruits ne faisant pas partie de l'interaction verbale, etc. Les chevauchements entre locuteurs ne sont pas non plus signalés. Les corpus sont annotés semi-automatiquement en lemmes, catégories grammaticales et même en fonctions syntaxiques (un avantage que le CEFC détient par rapport à ORTOFON v1, qui ne dispose pas de ce type d'annotation) dans l'outil MACAON²¹.

Les transcrip-teurs des deux corpus ont procédé à l'anonymisation des personnes faisant l'objet des conversations. Dans ORTOFON v1, seuls les prénoms sont gardés dans les transcriptions, alors que les noms de famille sont codés par le lemme « NP » ; dans le CEFC, la plupart des noms et prénoms sont codés par le lemme « NNAAMMEE », éventuellement par de faux prénoms, choisis au hasard (le cas du sous-corpus Valibel). Dans les enregistrements, les noms anonymisés sont rendus non identifiables²².

Voici un résumé des principales propriétés statistiques de notre corpus personnel créé sur le portail Orféo ainsi qu'un exemple de son interface :

Nom	Le CEFC (conversations privées)
Nombre de positions (tokens)	934 102
Nombre de conversations enregistrées	176
Nombre de locuteurs (non uniques)	487
Période des enregistrements	2000-2012 ; sauf les conversations Prov pin 89 et PRI-BAY-2, datant de 1989 et de 1999, et 5 conversations avec une date d'enregistrement inconnue

Tableau 2.3 : Propriétés de notre corpus de français oral informel constitué à partir des conversations privées du CEFC (Orféo)

²¹ Voir Nasr, Béchet et Rey (2009).

²² Le procédé d'anonymisation n'est pas appliqué, dans les deux cas, aux personnes de notoriété publique (politiques, acteurs, etc.).

Corpus TUFS – Center of Usage-based Linguistic Informatics

Métadonnées : general

Métadonnées : locuteur GJ

Métadonnées : locuteur IA

Métadonnées : locuteur PE

Texte et audio

0:00 / 55:35

Vitesse de lecture: 1x

☒ Surligner mot courant

GJ: voilà donc est-ce que tu as fait pendant les vacances

IA: euh pendant les vacances euh j' ai d' abord commencé par travailler pendant un mois

GJ: tu as travaillé

IA: où euh j' étais au Conseil Régional de la ville de NNAAMMEE

GJ: et qu' est-ce que tu faisais comme travail

IA: euh je devais trier les archives euh c' était assez intéressant mais c' était un travail difficile

Figure 2.5 : Interface du CEFC (Orféo)

2.1.3 Le corpus multilingue InterCorp

Le projet InterCorp de la base du CNT regroupe des textes avec leurs traductions vers d'autres langues avec le tchèque comme langue pivot :

InterCorp is a large parallel synchronic corpus covering a number of languages. The corpus is compiled mostly by teachers and students of the Faculty of Arts, Charles University in Prague, and by other collaborators of the Institute of the Czech National Corpus. It serves as a source of data for theoretical studies, lexicography, student research, (foreign) language learning, computer applications, translators and also for the general public²³.

Régulièrement mise à jour, le corpus en est aujourd'hui (mars 2021) à sa treizième version, incluant 40 langues différentes plus le tchèque comme langue pivot. Nous avons travaillé avec la dixième version, car les trois dernières versions (rendues publiques en 2011, 2019 et 2020) contiennent une pièce (*Petice*) dont la traduction en français n'est pas disponible dans la base. S'agissant de la nature des textes d'InterCorp, on y trouve, avant tout, des textes littéraires, mais aussi les textes des acquis communautaires de l'Union européenne, des articles de journaux des bases PressEurope et Syndicate, des sous-titres de films et autres. InterCorp permet la création de sous-

²³ Pour plus d'informations, voir <http://wiki.korpus.cz/doku.php/en:cnk:intercorp>.

corpus personnalisés. Aussi avons-nous défini, pour les objectifs de notre travail, un petit sous-corpus composé de quatre pièces de V. Havel traduites en français. Le Tableau 2.4 représente une synthèse des propriétés de ce sous-corpus :

Pièce	Traducteur(s)	Éditions tchèques	Éditions françaises ²⁴	Nombre de mots (tchèque/français)
<i>La fête en plein air</i> (<i>Zahradní slavnost</i>)	François Kérel	1963 / 1992	1964 / 1990	10 676 / 16 490
<i>Audience</i> (<i>Audience</i>)	Marcel Aymonin, Stéphan Meldegg	1975 / 1999	1979	5 175 / 7 547
<i>Vernissage</i> (<i>Vernisáž</i>)	Stéphan Meldegg	1975 / 1999	1979	5 171 / 7 426
<i>La grande roue</i> (<i>Žebrácká opera</i>)	Ivan Palec	1975	1987	17 098 / 21 163
Total				38 120 / 52 626²⁵

Tableau 2.4 : Propriétés du sous-corpus des quatre pièces de V. Havel constitué à partir d'InterCorp v10 (CNT)

Le tout contient, respectivement, 38 120 mots et 48 568 positions en tchèque et 52 626 mots et 64 540 positions en français. Un avantage que nous voyons dans cette composition est que les influences possibles de l'idiolecte d'un seul auteur tchèque se voient au moins partiellement neutralisées par quatre traducteurs différents en français. Toutefois, les textes ne se conforment pas tout à fait à la définition que donnent les auteurs du CNT de la notion de synchronie pour les textes de fiction. Cette définition veut que la première édition du texte ne remonte pas à plus de 75 ans et que l'édition présente dans le corpus ne remonte pas à plus de 25 ans²⁶. Seule la première condition est ici satisfaite par les quatre pièces et leur traduction. Cependant, nous ne voyons pas en cela un handicap trop grave, la langue des pièces ne s'éloignant pas ou guère de la langue contemporaine. D'autres paramètres des textes plaident, à nos yeux, pour ce choix, à commencer par leur nature de pièces, chose qui va de pair avec un usage exclusif du discours direct (les didascalies n'ont pas été

²⁴ Pour les deux langues, nous indiquons d'abord l'année de la toute première édition, suivie de l'année de l'édition enregistrée dans le corpus (si ces deux années diffèrent). Les traducteurs sont, bien entendu, ceux de l'édition enregistrée dans InterCorp v10.

²⁵ Pour chacune des pièces séparément, la taille est uniquement indiquée en mots, contrairement à celle de l'ensemble, qui est indiquée à la fois en mots et en positions.

²⁶ Pour plus d'informations, voir <http://wiki.korpus.cz/doku.php/en:cnk:syn2015>.

incluses). Ensuite, ces pièces de théâtre mettent en place une représentation de l'oral familial, même si des éléments de langage formel y sont présents (notamment dans *La grande roue* ; voir la Section 5.2.4 du Chapitre 5) ; la segmentation en dialogues rend l'orientation dans les textes plus aisée, y compris pour ce qui est de la correspondance entre l'extrait de l'original et celui de la traduction, particulièrement importante pour notre annotation. Le sous-corpus ainsi créé est annoté en lemmes et en catégories grammaticales²⁷. Les textes de l'original et de la traduction sont alignés au niveau des phrases :

Hits: 1,064 | i.p.m.: 12,353.81 (related to the subcorpus "havel do fr") | ARF: - | Result is shuffled

Line selection: simple (34 selected)

InterCorp v10 - Czech		InterCorp v10 - French	
☒ Havel-Zebracka_opera	Máte tu návštěvu , kapitáne -	Havel-Zebracka_opera	Une visite pour vous ...
☐ Havel-Audience	Ten Tonda Mašku je fakt slušnej kluk a fakt to potřebuje - tak ho nemůžem přece nechat ve štychu !	Havel-Audience	Ce Anton Mashek, c' est un type très bien, il en a besoin, je te jure - on ne peut pas le laisser dans la merde .
☐ Havel-Audience	Ale tu Bohdalku přitáhneš , víd že jo ?	Havel-Audience	Tu vas trouver la Bodalova et tu l' amèneras ici, dis, tu le feras, n' est - ce pas ?
☐ Havel-ZahradníSlavnost	A von i ten likvidační úředník má právo na ten kus toho vopravdu nějak jako plnýho - rozumíte, plnýho - života !	Havel-ZahradníSlavnost	Et un fonctionnaire de la liquidation a droit lui aussi à une vie, voyons, vous comprenez, à une vie pleine, vraiment pleine !
☐ Havel-Vemisaz	Nechápu, jak ses mohl zaplést s těmi komunisty -	Havel-Vemisaz	Qu' es -ce qui a pu te pousser à t' associer avec ces communistes ? !
☐ Havel-Zebracka_opera	Hned zítra vám to děvče pošlu na ukázk u !	Havel-Zebracka_opera	Je vous envoie l' échantillon dès demain .
☐ Havel-ZahradníSlavnost	A to jste tu poštu třídili celou noc ?	Havel-ZahradníSlavnost	Et vous avez passé toute la nuit à trier le courrier ?
☐ Havel-Zebracka_opera	To jmění už skoro mám .	Havel-Zebracka_opera	Le petit magot, je l' ai presque !
☒ Havel-Audience	Ale stejně : jsou to paradoxy, co ?	Havel-Audience	N' empêche, c' est bien ce que je disais, c' est un paradoxe, hein ?
☐ Havel-Zebracka_opera	Mám ten dojem - HAROLD Šroubek -	Havel-Zebracka_opera	J' en ai l' impression ...

Figure 2.6 : Interface du sous-corpus des quatre pièces de V. Havel (InterCorp v10, le CNT) et de leur traduction française

2.1.4 Les autres corpus et bases de données

Outre ces trois sources principales, nous en avons également mobilisé d'autres, bien que dans une moindre mesure. Nous l'avons notamment fait pour étayer la validité de certains constats formulés grâce aux données obtenues dans nos sources de base, de taille limitée, en regardant leur présence et leur fréquence au sein d'ensembles de données plus larges et pouvant, de ce fait, davantage aspirer au statut de corpus de référence. Pour le tchèque, il s'agit d'autres corpus faisant eux aussi partie du CNT, à savoir : les corpus de la série ORAL— notamment la version ORAL v1, où des conversations informelles constituent l'écrasante majorité des données —, le corpus PMK (Pražský mluvený korpus / Corpus parlé de Prague), qui est un corpus de tchèque parlé composé à la fois de conversations et d'entrevues enregistrés exclusivement à Prague, le corpus de référence pour le

²⁷ Les textes tchèques sont annotés avec le logiciel Morče (<http://ufal.mff.cuni.cz/morce/index.php>) et les textes français, avec TreeTagger (<https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>).

tchèque écrit synchronique SYN2015 et, enfin, le corpus Diakorp v6, corpus de référence pour le tchèque diachronique.

En ce qui concerne ORAL v1, le principal avantage de ce corpus est sa taille : 6 361 707 de positions (soit 5 368 392 de mots) contre seulement 1 236 508 de positions (soit 1 014 786 de mots) pour ORTOFON v1²⁸. De par cette dimension, ORAL v1 peut aspirer au statut de corpus de référence pour le tchèque parlé. En revanche, contrairement à ORTOFON v1, il ne s'agit pas d'un corpus représentatif, car la majorité des données provient de Bohême²⁹. Quant au corpus SYN2015, il contient 120 748 715 positions au total (soit 100 838 568 de mots), réparties de manière équilibrée entre trois principaux types de textes, à savoir : la fiction, la littérature spécialisée (la non-fiction) et les textes de presse. Aucun des textes n'a été publié pour la première fois avant 1961, sachant que les textes de littérature spécialisée ont tous été publiés pour la première fois entre 1990 et 2014, et les textes de presse entre 2010 et 2014. Le corpus inclut aussi bien des originaux tchèques que des traductions³⁰.

Au sein de la partie « fiction », nous avons défini un sous-corpus composé uniquement de pièces de théâtre publiées entre 2007 et 2009 et contenant 1 301 447 de positions. Ce sous-corpus, contenant à la fois des originaux tchèques et des textes traduits³¹, nous servira de référence pour certains phénomènes analysés dans notre sous-corpus des quatre pièces de V. Havel.

Les corpus ORAL v1 et SYN2015 sont lemmatisés³². Le corpus PMK contient 819 267 positions (soit 674 992 de mots). Il est composé à la fois de conversations et d'entretiens et il n'est pas lemmatisé. Dans le cadre des entretiens, la personne chargée de l'enregistrement posait des questions, majoritairement sociétales, ayant pour but de faire réagir le locuteur, l'inciter à s'exprimer de manière spontanée et informelle. À cette fin, les questions étaient elles aussi formulées en « tchèque commun » (voir la Section 3.1.1 pour l'explication de ce terme). Par ailleurs, chaque occurrence d'un lemme a été évaluée en termes de l'opposition « formel » vs « informel », déterminée à la fois par la forme elle-même et par le profil de l'ensemble du discours dans lequel elle s'inscrivait. La part des entretiens représente à peu près 38 % des données. Si le corpus n'est pas équilibré sur le plan diatopique, il l'est au regard de l'âge, du sexe et du niveau de formation des locuteurs. Le corpus

²⁸ La taille de tous les corpus du CNK est indiquée à la fois en mots et en positions (tokens). Rappelons que ce dernier calcul inclut également les signes de ponctuation et, pour les corpus oraux, les signes marquant les hésitations, les commentaires métalinguistiques, les interjections, etc.

²⁹ Voir <https://wiki.korpus.cz/doku.php/en:cnk:oral>.

³⁰ Pour plus d'informations, voir <https://wiki.korpus.cz/doku.php/en:cnk:syn2015>.

³¹ Soit à peu près un quart d'originaux et trois quarts de traductions.

³² La lemmatisation diffère à certains égards entre les corpus de langue écrite et ceux de langue parlée (voir la Section 3.1.1 du Chapitre 3).

Diakorp v6 contient 4 128 874 de positions (soit 3 450 142 de mots). Les textes, non lemmatisés, sont de nature écrite ou parlée formelle (prières, discours solennels) et datent de la période 1375-1939³³.

Pour le français, nous avons recouru à Frantext, base à dominante littéraire et contenant notamment des textes de la période 17^{ème}-21^{ème} siècles³⁴, ainsi qu'à la base ScienQuest, mettant notamment à la disposition des internautes des corpus de textes scientifiques³⁵. Enfin, des données statistiques ont aussi été empruntées au corpus sur lequel se basent le *Dictionnaire des fréquences : vocabulaire littéraire des XIXe et XXe siècles* d'Imbs et de Martin (1971) et le *Trésor de la langue française* d'Imbs et de Quémada (1971-1994)³⁶. ScienQuest recèle bien des avantages, dont la possibilité de choisir entre plusieurs corpus. Dans notre cas, nous en avons utilisé deux, à savoir : le corpus Écrits scientifiques en français (5 596 492 de mots)³⁷ et le corpus du journal *l'Est républicain*, composé de textes de presse parus dans le journal éponyme entre 1999, 2002 et 2003 (87 341 867 de mots)³⁸. Mais – surtout – la plateforme permet d'opérer la distinction entre pronoms et déterminants. Quant à la liste des concordances, elle ne contient quasiment pas d'occurrences bruitées. La lemmatisation des mots grammaticaux (démonstratif, article défini) est également opérée pour les deux corpus avec lesquels nous avons travaillé. Frantext propose des fonctionnalités diverses, permettant de rechercher des formes et des lemmes au sein d'un sous-corpus prédéfini par la plateforme ou créé par l'utilisateur lui-même. Il permet également d'afficher des statistiques portant sur la composition du (sous-)corpus et la répartition du phénomène recherché. Pour le démonstratif, le paradigme adnominal CE est distingué de façon minutieuse de la forme pronominale *ce*³⁹. Nous n'avons identifié aucune occurrence bruitée lors de notre examen des listes de concordances générées. Nous avons travaillé avec un sous-corpus personnalisé, composé uniquement de textes de fictions (« domaine » = « fiction » ; tous « genres textuels » confondus, voir *infra*) publiés pour la première fois entre 1961 et 2014, soit la tranche chronologique qui correspond aux textes de fiction de SYN2015. Ce sous-corpus contient 2 363 566 de positions (2 008 222 de mots)⁴⁰.

³³ Le corpus n'est cependant pas représentatif, car 48 % des textes datent du 19^{ème} siècle. Voir <https://wiki.korpus.cz/doku.php/en:cnk:diakorp>.

³⁴ 258 177 340 de mots et 299 308 409 de positions au total. Le plus grand nombre de textes remontent cependant au 20^{ème} siècle. Voir <https://www.frantext.fr/information>.

³⁵ Voir <https://corpora.aiakide.net/scientext20/?do=SQ.set&view=corpora>.

³⁶ Ce corpus est majoritairement composé de textes de fiction. Il est réparti en quinze tranches chronologiques d'une taille à peu près égale (entre 4 et 6 millions de mots). Les mots grammaticaux ne sont pas lemmatisés (voir la Section 4.1.4 du Chapitre 4).

³⁷ Les textes relèvent notamment du domaine des sciences humaines et, dans une moindre mesure, des sciences expérimentales et des sciences appliquées. Leur date de publication n'est pas renseignée. Pour plus d'informations, voir <https://corpora.aiakide.net/scientext20/?do=SQ.set&textsTab=list>.

³⁸ Ce corpus est également disponible sur l'interface du CNT. Il ne s'agit que d'une part de l'ensemble du Corpus Est Républicain disponible. Pour plus d'informations, voir <https://corpora.aiakide.net/scientext20/?do=SQ.set&textsTab=list>.

³⁹ Grâce à la requête ([lemma="ce"%cd & pos="DET"]).

⁴⁰ Nous avons décidé de nous contenter de cette taille, car Frantext ne permet d'afficher ni la totalité de la liste des concordances ni la fréquence absolue de l'élément recherché lorsque celle-ci dépasse le nombre de 100 000 occurrences.

Il faut préciser que de manière générale, le classement des textes dans Frantext est différent de celui de SYN2015. Les auteurs de ce dernier distinguent de manière rigoureuse entre quatre niveaux hiérarchiques : « groupe de types de textes » (“text-type group”), « type de textes » (“text-type”), « groupe de genres » (“genre-group”) et « genre » (voir la Section 2.1.1 *supra*). En ce qui concerne les textes de fiction, leur hiérarchisation n’est opérée qu’aux deux premiers niveaux, en raison de l’hétérogénéité intrinsèque à cette catégorie (ainsi, les romans, la poésie et les pièces de théâtre sont considérés comme autant de types de textes). À titre d’exemple, au sein de la littérature spécialisée (un groupe de types de textes), l’on trouve « littérature scientifique », « littérature professionnelle » et « littérature populaire » (des types de textes ; voir la Section 2.1.1 *supra*). Dans « littérature scientifique », l’on trouve « sciences humaines », « sciences sociales », etc. (soit des types de genres). Et, enfin, dans « sciences humaines », l’on trouve « philosophie et religion », « anthropologie et ethnographie », etc., qui sont donc des genres. Dans Frantext, en revanche, le classement n’opère qu’à deux niveaux : « domaine » et « genre textuel ». Le premier niveau étant assez large, il englobe des textes ressortissant à des champs aussi disparates que « littérature », « linguistique », « histoire », « arts », « sciences de la terre », « philosophie », « critique littéraire », « politique », « information », « industries », « ethnologie », « sémiologie », « biologie », « administration publique », etc. Le second niveau, lui, regroupe des genres tels que « essai », « roman », « autobiographie », « journal », « écrits personnels », etc. Dans SYN2015, chaque texte est associé à une seule catégorie au sein d’un niveau hiérarchique (d’où l’impossibilité pour un texte, par exemple, d’être à la fois étiqueté comme « autobiographie » et comme « roman »). Dans Frantext, les textes sont associés à plusieurs catégories à la fois, ce qui amoindrit la souplesse de la recherche et la définition de sous-corpus personnalisés : à titre d’exemple, si l’on coche, lors de la définition d’un sous-corpus personnalisé, à la fois les genres textuels « autobiographie » et « roman », le résultat sera filtré de telle sorte que seuls seront maintenus les textes associés à ces deux catégories à la fois. Autre différence importante : tandis que Frantext ne contient (à quelques rares exceptions près) que des originaux rédigés en français, les trois sections de SYN2015 mélangent originaux tchèque et traductions⁴¹.

Or les fréquences du déterminant démonstratif et de l’article défini (notamment celle de l’article défini) dépassent facilement ce seuil, au-delà duquel le calcul des fréquences devient une affaire compliquée. L’échantillon en question contient 34 textes. Il a été créé par la sélection, à partir d’une liste initiale de 349 textes de fiction publiés pour la première fois entre 1961 et 2014 et triés par l’ordre alphabétique du nom de l’auteur, de chaque dixième texte. Si quelques textes n’ont pas été réédités depuis leur première publication, cela ne constitue certainement pas un handicap, dans la mesure où, contrairement à celle du TEN adnominal tchèque, l’évolution de l’article défini français est achevée. Contrairement à la partie fiction de SYN2015, ce sous-corpus ne contient pas de textes de pièces de théâtre (toutefois, dans la partie fiction de SYN2015, ces dernières constituent à peu près 3 % de l’ensemble – leur influence sur cet ensemble est donc minimale).

⁴¹ Les auteurs de SYN2015 partent du présupposé que les traductions obéissant aux critères de synchronie (la première publication de l’original ne remonte pas à plus de 75 ans et sa traduction incluse dans le corpus ne remonte pas à plus de 25 ans) reflètent elles aussi l’usage de la langue (écrite) de l’époque concernée.

2.1.5 Présentation des informateurs

Comme autre source complémentaire d'informations, nous avons sollicité dix locuteurs francophones natifs et dix locuteurs tchécophones natifs. C'est au jugement et au sentiment de la langue de ces locuteurs-ci que nous avons soumis certaines de nos intuitions et c'est avec eux que nous avons consulté certaines thèses formulées au sujet du démonstratif dans les deux systèmes et présentées comme plus ou moins généralement acquises dans la littérature. Nous tenons ici, une fois de plus, à les en remercier vivement. Entre autres, nous avons recouru à l'usage de trois questionnaires simples, (deux pour le tchèque et un pour le français). Ces questionnaires figurent dans les Annexes⁴². Deux d'entre eux (Questionnaires n°1 et n°2) – le premier pour le tchèque et le second pour le français – avaient pour but d'enquêter sur le choix des formes au sein de l'emploi situationnel, permettant ainsi d'étudier la conceptualisation de l'espace en fonction de l'opposition proche vs distant et les préférences des locuteurs pour certaines formes au détriment d'autres (notamment en tchèque, où il existe une assez grande variété de formes pour exprimer le même degré de distance ; voir la Section 3.1.1 du Chapitre 3 ; quant au français, il s'agissait avant tout d'étudier la concurrence entre les formes en *-ci*, celles en *-là* et celles dépourvues de particule localisatrice). Le troisième questionnaire (Questionnaire n°3), s'adressant uniquement aux informateurs tchécophones, s'enquiert de l'emploi du démonstratif TEN en position adnominale dans les domaines déictique situationnel, anaphorique (y compris associatif), mémoriel et cataphorique intra-syntagmatique. Il vise notamment à étudier l'ampleur avec laquelle TEN est employé, dans une conversation entre amis et personnes proches, au sein des contextes qui semblent avoir plus ou moins de préférence, en français, pour l'article défini. Bon nombre d'exemples utilisés dans ce questionnaire sont également cités et commentés dans le texte du présent travail.

Nous avons tâché de choisir nos informateurs tout en ayant le souci d'une certaine diversité dans les profils, notamment pour ce qui était du tchèque, car certaines formes du démonstratif de cette langue sont marquées sur le plan diatopique (voir également la Section 3.1.1 du Chapitre 3). Pour le tchèque, nous avons ainsi sollicité dix personnes, dont cinq femmes et cinq hommes. Leurs années de naissance sont : 1972, 1990 (deux personnes), 1991 (quatre personnes) et 1992 (trois personnes). Toutes sauf une sont titulaires d'un diplôme universitaire (du niveau licence ou supérieur, dans les domaines des lettres et sciences humaines et du droit) ; celle qui fait exception est titulaire d'un bac + 2. Sur le plan diatopique, nous nous sommes intéressé à la fois à la ville et la région où les personnes ont résidé lors de la totalité ou la plupart de leur enfance (soit jusqu'à l'âge de dix-huit ans) et à la ville et la région où elles résident actuellement, y compris le nombre d'années (éventuellement de mois) passés à cet endroit. Voici un récapitulatif des profils de nos informateurs tchèques eu égard à

⁴² Les consignes et les exemples des questionnaires destinés aux locuteurs tchèques y sont traduits en français.

leur sexe, leur année de naissance, leur principal lieu de résidence jusqu'à l'âge de dix-huit ans et leur lieu de résidence actuel⁴³ :

Informateur	Sexe	Année de naissance	Principal lieu de résidence avant l'âge de 18 ans	Lieu de résidence actuel	Temps passé en ce lieu
Personne 1	F	1992	Ostrava, Moravie-Silésie	Černošice, Bohême centrale	5 ans
Personne 2	F	1992	Valašské Meziříčí, Région de Zlín	Prague	8 ans
Personne 3	F	1991	Nedabyle, Bohême du Sud	Prague	10 ans
Personne 4	F	1991	Prague	Prague	29 ans
Personne 5	F	1990	Olomouc, Région d'Olomouc	Prague	9 ans
Personne 6	M	1991	Jihlava, Vysočina	Prague	6 ans
Personne 7	M	1991	Lomnice nad Lužnicí, Bohême du Sud	Hradec Králové, Région de Hradec Králové	10 mois
Personne 8	M	1990	Prague	Prague	29 ans
Personne 9	M	1972	Rosice u Brna, Moravie du Sud	Prague	24 ans
Personne 10	M	1992	Plzeň, Région de Plzeň	Prague	8 ans

Tableau 2.5 : Informateurs tchécophones natifs avec leurs métadonnées

Nous avons ainsi six informateurs et informatrices originaires de la zone dialectale de Bohême (Prague, Vysočina, Bohême du Sud et Bohême de l'Ouest), une informatrice originaire de la zone dialectale de Moravie de l'Est (personne 2), une informatrice originaire de la zone dialectale

⁴³ Nous invitons le lecteur à consulter les Figures 3.1, 3.2 et 3.3 de la Section 3.1.1 du Chapitre 3 et reproduites dans les Annexes pour les éléments nécessaires à la compréhension de la situation territoriale et dialectologique de la République tchèque.

silésienne (personne 1), une informatrice originaire de la zone dialectale de Moravie centrale (personne 5) et un informateur originaire de la zone dialectale de Moravie du Sud⁴⁴ (personne 9). Comme le départ des jeunes personnes vers les métropoles devient de plus en plus fréquent, il devient également de plus en plus difficile de trouver des informateurs « purs » quant à la représentativité de leur zone dialectale. Ainsi, quasiment tous les informateurs ayant passé leur enfance hors de la capitale y sont maintenant installés depuis des années⁴⁵.

S'agissant de nos informateurs francophones natifs, nous avons également recherché une certaine diversité, à la fois sur le plan diatopique et sur le plan du sexe et de l'âge. En commençant par le premier paramètre, nous avons parmi nos informateurs cinq Français, un Belge, deux Marocaines, un Algérien et un Luxembourgeois. Il s'agit de six hommes et de quatre femmes. Tous sont titulaires d'un diplôme universitaire (du niveau licence ou supérieur, tous domaines confondus, mais avec une prévalence des lettres et sciences humaines) et sont nés entre 1977 et 1996. Le Tableau 2.6 synthétise les profils de nos informateurs francophones eu égard à leur sexe, leur année de naissance, leur principal lieu de résidence jusqu'à l'âge de dix-huit ans et leur lieu de résidence actuel :

Informateur	Sexe	Année de naissance	Principal lieu de résidence avant l'âge de 18 ans	Lieu de résidence actuel	Temps passé en ce lieu
Personne 1	M	1985	Luxembourg, Luxembourg	Luxembourg, Luxembourg	29 ans
Personne 2	M	1992	Sèvres, Île de France	Lyon, Rhône-Alpes	7 ans
Personne 3	M	1977	Lyon, Rhône-Alpes	Prague, République tchèque	16 ans
Personne 4	M	1982	Lobbes, Hainaut, Belgique	Liège, Belgique	21 ans
Personne 5	M	1970	Paris	Prague, République tchèque	22 ans
Personne 6	M	1994	Oran, Algérie	Lyon, Rhône-Alpes	5 ans

⁴⁴ Voir la Figure 3.3 présentée dans le Chapitre 3 et reproduite dans les Annexes.

⁴⁵ Le cas de la personne 7 fait légèrement exception à la règle, mais celle-ci a également passé dix ans de sa vie à Prague avant de déménager à Hradec Králové. De plus, une part de sa vie professionnelle continue à être localisée à Prague.

Personne 7	F	1996	Saint-Denis, La Réunion ; Nouvelle Aquitaine	Paris	2 mois
Personne 8	F	1994	Casablanca, Maroc	Lyon, Rhône-Alpes	5 ans
Personne 9	F	1994	Casablanca, Maroc	Lyon, Rhône-Alpes	4 ans
Personne 10	F	1981	Toulouse, Midi Pyrénées	Fonsorbes, Midi Pyrénées	2 ans

Tableau 2.6 : Informateurs francophones natifs avec leurs métadonnées

Les trois questionnaires furent introduits par une partie expliquant les enjeux⁴⁶ et les consignes. Exception faite de deux cas concrets du troisième questionnaire et qui ressortissent au registre formel, nous avons insisté sur le caractère parlé informel du contexte où les énoncés étaient employés en incitant littéralement les informateurs à choisir la forme qu'ils trouvaient la plus naturelle et qu'ils emploieraient au quotidien, dans une conversation informelle avec leurs amis. Ce caractère est d'ailleurs directement évoqué par la présence des traits morphologiques, syntaxiques et lexicaux typiques de la langue informelle⁴⁷. Dans les deux premiers questionnaires, les informateurs furent invités à choisir, dans chaque cas, entre une ou plusieurs formes du démonstratif. Dans le troisième questionnaire – se limitant au tchèque –, ils avaient pour consigne de choisir une seule variante sur deux ou trois variantes proposées, dont toujours une où le nom était présenté comme nu et une où il était introduit par TEN ; dans certains cas, une troisième possibilité était proposée, où le nom était introduit par TENHLE. À plusieurs reprises, les informateurs des deux communautés linguistiques nous ont fait part, sous forme de commentaires, de leurs impressions au sujet de certains exemples (dans la consigne, nous les avons explicitement invités à le faire), ce dont nous tenons ici à les remercier tout particulièrement.

Enfin, nous tenons ici à mentionner les noms de sept personnes natives d'autres langues slaves que le tchèque. C'est grâce à leurs aimables retours que nous avons pu nous faire une meilleure idée de la situation du démonstratif dans ces langues, notamment eu égard à sa grammaticalisation en article défini et à ses emplois « émotionnels » (voir les Chapitre 6 et 7). Ces personnes sont : E. Kaczmarska de l'Institut d'études slaves occidentales et méridionales de l'Université de Varsovie et P. Płocharz de l'École normale supérieure de Lyon pour le polonais, E. Bratková, B. Masar et I. Zorn

⁴⁶ Sans pour autant que fussent divulguées nos hypothèses, bien entendu.

⁴⁷ Il s'agit globalement du « tchèque commun ». Voir la Section 3.3.1 du Chapitre 3 pour la définition de ce terme.

de l'Institut de traductologie de l'Université Charles de Prague pour le slovaque et A. Zhuk et P. Felčer pour le russe et l'ukrainien.

2.1.6 Présentation des outils statistiques employés

Pas plus que n'importe quelle autre recherche linguistique s'appuyant sur des corpus de données, la nôtre propre ne saurait se passer du recours à quelques outils statistiques fondamentaux, permettant de tester la validité de certains constats d'ordre quantitatif. Dans le présent travail, nous avons, avant tout, employé à de multiples reprises la fonction de vraisemblance⁴⁸. L'application de cette fonction – également appelée « the likelihood function » ou « log likelihood » en anglais – dans le domaine de la linguistique permet de tester la significativité statistique d'une différence quantitative apparue au sein d'un ensemble de données (une population n). Pour tester la significativité statistique d'une différence de fréquence, nous avons donc opté pour cette fonction plutôt que pour d'autres fonctions plus ou moins analogues, tel le test du Khi 2 ou le test exact de Fisher (Fisher's exact test), car elle nous semble la plus appropriée au profil de nos données. En effet, si le Khi 2 s'avère peu fiable lorsqu'il est appliqué à des fréquences trop basses (en dessous de cinq), le test exact de Fisher permet, au contraire, de mettre en évidence les différences d'ordre de grandeur même entre des valeurs très petites. Mais c'est pour cela même qu'il se montre moins opérant lorsqu'il est appliqué à nos données, car, du moins sur l'interface du CNT avec laquelle nous travaillons, les chiffres sont livrés trop bruts, sans aller très loin dans la précision, ce qui fait qu'ils sont quasiment toujours nuls (l'interface ne distingue pas entre différentes valeurs très proches de zéro)⁴⁹.

En revanche, un type de test exact de Fisher, appelé « calcul de spécificités »⁵⁰, sera employé dans le Chapitre 5, afin de déterminer s'il existe des affinités entre plusieurs valeurs d'annotation, telles que, d'une part, le type/sous-type d'emploi et, d'autre part, la position occupée par l'expression démonstrative par rapport au verbe de la proposition, la présence des particules *-ci/-là*, la grammaticalisation de TEN en position adnominale, etc. (voir la Section 2.2.3 *infra* pour l'élucidation de ces valeurs). Le grand bénéfice du calcul de spécificités est son mode opératoire symétrique : appliqué à une série de données assemblées sous la forme d'une table de contingence importée dans le logiciel TXM, il explore le lien entre les deux variables A et B dans les deux directions. Ainsi, pour qu'il y ait affinité, il faut que B soit surreprésenté par rapport à A, mais aussi que A le soit par rapport à B.

⁴⁸ Notamment dans les Chapitres 3, 4 et 5.

⁴⁹ Pour plus d'informations sur ces différents tests, voir le site Wikipédia du CNT : <https://wiki.korpus.cz>.

⁵⁰ Ce calcul est couramment utilisé par le groupe de recherche en textométrie CACTUS, développeur du logiciel TXM. Pour plus d'informations sur le calcul, voir <http://txm.sourceforge.net/doc/manual/0.7.7/fr/manual36.xhtml#toc142>. Nous tenons ici à remercier tout particulièrement B. Pincemin, qui nous a été d'une aide inestimable pour la compréhension et l'application du calcul.

Parmi ses nombreux avantages, le site du CNT compte la possibilité d'employer des interfaces pour réaliser des tests statistiques de manière automatique, à travers une simple saisie des données observées. Ces interfaces sont regroupées sous l'application Corpus Calculator⁵¹. Les interfaces dont nous avons fait l'usage sont les suivantes : « 2 words in 2 corpora », « 2 words in 1 corpus », « 1 feature – many samples » et « many features – 1 sample ».

La première interface nous permet de comparer la fréquence d'un lemme – démonstratif, dans notre cas – dans deux corpus. L'enjeu consiste à tester la significativité statistique de l'écart entre les deux fréquences en employant, justement, la fonction de vraisemblance. Cette vérification est souhaitable au-delà de la simple comparaison des fréquences relatives (rapportées, dans notre cas, à un million de positions ou de mots) du lemme au sein des deux ensembles de données⁵². Il est recommandé de travailler avec les nombres absolus, tant en ce qui concerne les fréquences que les tailles des deux ensembles. Nous travaillons toujours avec un seuil de significativité (α) égal à 0,05, ce qui est d'habitude jugé suffisant pour le domaine des sciences humaines. La valeur critique citée par le tableau de contingence est, dans ce cas, égale à 3,841, le degré de liberté étant, lui, égal à 1. Concrètement, cela veut dire pour nous que lorsque le score de la fonction de vraisemblance (L) est égal ou supérieur à 3,841, il existe au moins une probabilité de 95 % que la différence entre les deux fréquences absolues soit statistiquement significative (le risque d'erreur étant ainsi égal ou inférieur à 5 %). Autrement dit, lorsque le score est égal ou supérieur à 3,841, il y a lieu de rejeter l'hypothèse nulle (H_0), selon laquelle il n'y a pas de différence de nature entre les deux ensembles de données, avec un risque de se tromper à hauteur de 5 % ou moins. Dans ce cas, la p -valeur est donc inférieure ou égale à 0,05⁵³. L'interface calcule également les fréquences relatives des lemmes, tout en donnant un intervalle de confiance, soit la marge d'erreur pour les résultats obtenus⁵⁴.

L'interface appelée « 2 words in 1 corpus » permet d'appliquer la même fonction pour établir la significativité statistique de l'écart entre deux fréquences, à cette différence près que les fréquences sont observées au sein d'une même population n (d'un même corpus). Nous y avons recouru afin de vérifier la significativité statistique de l'écart entre les fréquences de deux lemmes au sein d'un même

⁵¹ Pour consulter l'interface : <https://www.korpus.cz/calc/>.

⁵² Exemple : le lemme du démonstratif tchèque TEN atteint une fréquence relative de 73 159,3 occurrences pour un million de positions au sein du corpus ORTOFON v1 et une fréquence relative de 14 597,4 occurrences pour un million de positions au sein du corpus SYN2015. Cependant, cette différence est-elle significative sur le plan statistique ou uniquement due au hasard ?

⁵³ La p -valeur exprime la probabilité d'obtenir la même différence ou une différence supérieure à celle que l'on a constatée dans une situation où H_0 serait valide. Ainsi, lorsque la p -valeur $\leq 0,05$, la probabilité d'obtenir la même différence entre les fréquences dans le cadre de H_0 est égale ou inférieure à 0,05 (seuil de significativité). Il y a donc de fortes chances que H_0 ne soit pas valide, et qu'elle puisse être rejetée.

⁵⁴ Par exemple « 2 631,336 ipm ($\pm 210,9294$) », soit une fréquence relative de 2 631,336 occurrences pour un million de positions avec une fluctuation pouvant aller jusqu'à (2 631,336 + 210,9294) et (2 631,336 – 210,9294) occurrences.

corpus⁵⁵. Les Figures 2.7 et 2.8 présentent les interfaces « 2 words in 2 corpora » et « 2 words in 1 corpus » :

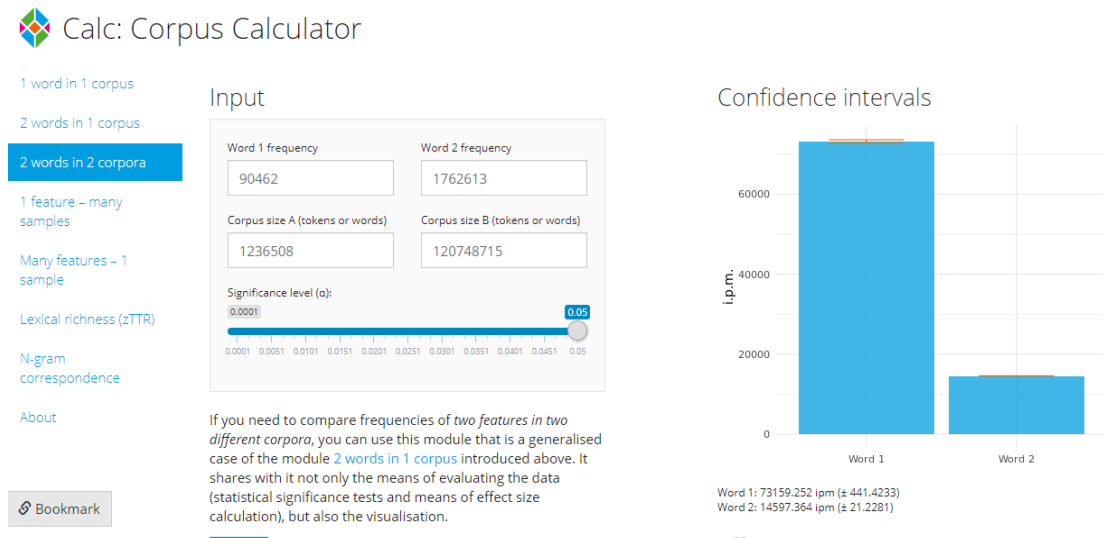


Figure 2.7 : Interface « 2 words in 2 corpora » (CNT) ; application de la fonction de vraisemblance aux fréquences absolues du lemme TEN dans les corpus ORTOFON v1 et SYN2015

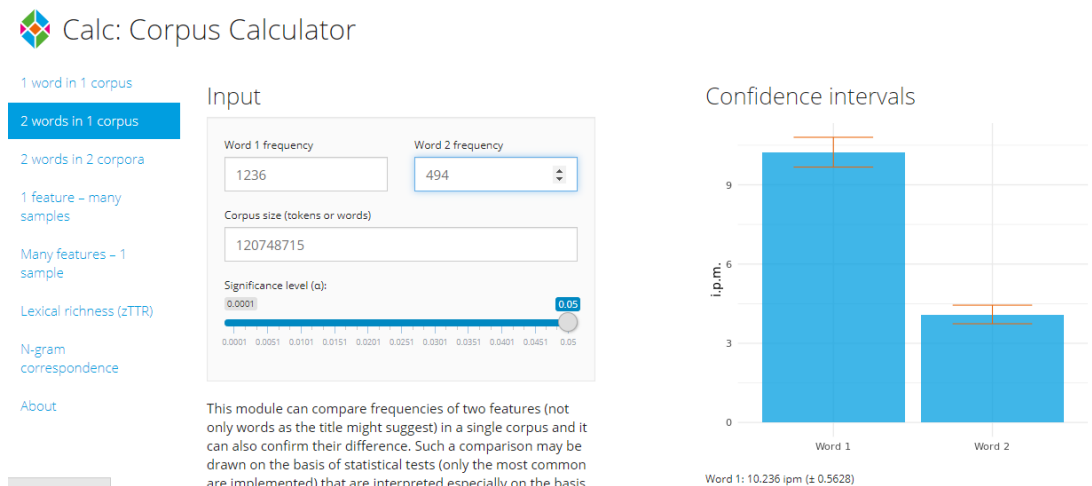


Figure 2.8 : Interface « 2 words in 1 corpus » (CNT) ; application de la fonction de vraisemblance aux fréquences absolues des lemmes TENHLE et TENTO dans le corpus ORTOFON v1

Quant à l'interface « 1 feature – many samples », elle sert à évaluer la représentativité d'un échantillon analysé manuellement, en utilisant la loi de Student ou la loi normale :

This module is intended to evaluate the reliability of results of manual analysis of a number of random samples (typically from an extensive concordance that is impossible to examine in detail). The usual problem of such analysis stems from our inability to determine the number and size of the individual samples necessary for a reliable estimate. There is no universal answer to this problem, because it depends on a number of factors (esp. on the frequency and dispersion of the feature), however, based on the confidence intervals, we can estimate

⁵⁵ Exemple : au sein du corpus ORTOFON v1, le lemme du démonstratif TENHLE atteint une fréquence relative de 999,6 occurrences pour un million de positions, alors que celui du démonstratif TENTO n'atteint qu'une fréquence relative de 399,5 occurrences pour un million de positions. Cependant, cette différence est-elle significative sur le plan statistique ou uniquement due au hasard ?

in what range of frequencies the feature in question appears in the whole concordance. This module shows that by adding samples, the frequency estimate precision increases (the confidence interval narrows). It is then possible to decide, based on this interval, if a given precision is satisfactory enough, or if more samples should be analysed to increase the precision of the estimate. The app provides two ways to calculate this: using the Student's and normal distribution. While the former usually returns a narrower confidence interval, the latter returns a more conservative estimate.

Dans notre cas, ce module nous a été utile lors du traitement des résultats obtenus suite à l'application de notre requête au corpus ORTOFON v1, plus précisément lorsqu'il a fallu calculer la part estimée des occurrences bruitées au sein de la liste des concordances générée par la requête (voir la Section 2.2.4 *infra* pour l'application concrète du module). Nous y avons également recouru pour calculer la part estimée, toujours dans le corpus ORTOFON v1, des occurrences appartenant aux lemmes du démonstratif autres que TEN (voir la Section 5.2.3.1).

Enfin, le module « many features – 1 sample » a notamment été utilisé pour établir la part approximative occupée, au sein de nos deux corpus de conversations, par les différents types et sous-types d'emploi du démonstratif adnominal et par les différents types de position par rapport au verbe de la proposition (voir la Section 2.2.3 *infra* pour la liste de ces valeurs). Ainsi, l'interface permet la saisie des fréquences observées pour les différents sous-types d'emploi à l'intérieur d'un échantillon représentatif. Cet échantillon est prélevé sur la liste initiale des concordances, elle-même obtenue par l'application de la requête. Dans notre cas, l'échantillon comptait 1 000 occurrences pour les deux corpus de conversations et la totalité des occurrences (soit 241) pour le sous-corpus des pièces de V. Havel (voir la Section 2.2.4 *infra* pour la démarche d'échantillonnage appliquée à nos sources de données). En utilisant la technique du bootstrapping, l'interface calcule ensuite les intervalles de confiance pour la fréquence de chaque valeur d'annotation, dans une situation où le seuil de significativité (α) est égal à 0,05. Nous détaillerons les résultats de ces calculs dans les Sections 5.2.2.1, 5.2.2.2, 5.2.3.1, 5.2.3.4 et 5.2.4.1 du Chapitre 5⁵⁶.

2.2 Travail avec les corpus et annotation des données

À présent, nous allons décrire plus avant notre façon de travailler avec nos sources de données. D'abord, nous parlerons de la nature des requêtes utilisées pour les corpus de langue orale informelle ainsi que du mode d'interrogation de ces corpus. Puis nous présenterons la requête et le mode d'interrogation employés pour notre sous-corpus des quatre pièces de V. Havel défini dans InterCorp. Nous passerons ensuite au sujet de l'annotation des données, commentant nos catégories et valeurs

⁵⁶ Nous y verrons que plusieurs intervalles de confiance peuvent se chevaucher, auquel cas il est impossible d'établir qu'un des phénomènes apparaît toujours plus que l'autre. Si un intervalle comprend la valeur zéro, il se peut que le phénomène en question ne soit pas du tout présent dans un échantillon différent de celui étudié.

d'annotation, et nous finirons par aborder la question de l'échantillonnage des résultats obtenus à partir de l'interrogation des corpus oraux.

2.2.1 Interroger les corpus : le CEFC et ORTOFON v1

Même si les conversations des deux langues peuvent être interrogées directement sur les sites des plateformes, nous avons également décidé de prendre contact avec les administrateurs afin de demander les transcriptions et de travailler avec elles dans le régime hors-ligne⁵⁷. La raison majeure justifiant cette démarche est que nous avons ainsi pu réimporter les occurrences annotées dans le logiciel TXM et, par la suite, exploiter les catégories d'annotation dans ce logiciel (procéder à des requêtes ciblées et afficher des statistiques). Nous avons procédé dans l'ordre suivant : une fois les textes des conversations reçus⁵⁸, nous avons sollicité M. Quignard, affilié au laboratoire ICAR de Lyon, qui a eu la gentillesse d'entreprendre le nécessaire pour leur import dans TXM ; une requête a ensuite été formulée pour relever le plus grand nombre possible de démonstratifs adnominaux dans les deux corpus. C'est A. Lavrentev, membre du groupe CACTUS et du laboratoire IHRIM de Lyon, qui nous a aidé à formuler ces requêtes. Leurs résultats ont été exportés en format Excel, où nous les avons annotés. Lors de ce travail d'annotation, nous avons très vite découvert la nécessité de travailler parallèlement avec les plateformes en ligne, et ce pour disposer à la fois de l'enregistrement apparié avec la transcription et d'un contexte plus large, le seul en mesure de trancher à propos de la nature d'un grand nombre d'occurrences⁵⁹. La dernière phase a été celle du ré-import des résultats des conversations annotés dans TXM, où les textes déjà stockés ont été enrichis de ces annotations.

L'élaboration des requêtes nous permettant d'accéder aux occurrences voulues s'est très tôt avérée délicate, notamment pour la partie d'ORTOFON v1. Les deux plateformes emploient une langue de requête assez semblable. Aux textes de langue tchèque, les créateurs du CNT appliquent leur propre version de l'interface No Sketch Engine appelée « KonText », ajustée aux propriétés de cette langue⁶⁰. Dans la plateforme du CEFC, c'est dans l'interface « recherche avancée » que l'utilisateur peut procéder à des requêtes plus sophistiquées, employant également un langage de requête ("corpus query language" – CQL) appelé « Annis Query Language »⁶¹. Le logiciel

⁵⁷ À ce titre, nous tenons tout particulièrement à remercier cinq personnes : M. Křen, directeur du CNT, d'avoir eu la gentillesse de nous envoyer l'intégralité des textes d'ORTOFON v1, C. Plancq d'avoir eu la gentillesse de nous envoyer les textes demandés du CEFC, S. Heiden d'avoir partiellement servi d'intermédiaire entre nous et C. Plancq et pour l'aide qu'il nous a apportée avec le traitement des données et des métadonnées, ainsi que M. Quignard et A. Lavrentev pour la conversion des textes du CEFC au format XML, l'import des textes et de nos annotations dans le logiciel TXM et l'aide avec la formulation des requêtes.

⁵⁸ Les textes des conversations d'ORTOFON v1 étaient en format XML ; ceux du CEFC étaient dans un format tabulaire convertible en format XML.

⁵⁹ L'interface du CNT ne permet pas d'afficher les textes dans leur intégralité. Néanmoins, nous disposons de ces textes dans le régime hors-ligne.

⁶⁰ Pour plus d'informations, voir http://wiki.korpus.cz/doku.php/en:pojmy:dotazovaci_jazyk.

⁶¹ L'interface « recherche simple » permet uniquement de rechercher des mots.

textométrique TXM travaille avec une CQL ne différant pas radicalement des versions utilisées par les deux bases. Concrètement, il s'agit de l'outil the IMS Open Corpus Workbench (CWB).

En revanche, ce sont les systèmes d'annotation en catégories grammaticales qui diffèrent davantage. Concernant le démonstratif, seul celui employé pour le français opère la différence très utile entre emplois adnominaux et pronominaux. Le système d'annotation tchèque suit, à cet égard, la tendance du discours grammatical, où les emplois adnominaux du démonstratif ne sont pas considérés comme une catégorie à part. Comme le tchèque ne possède pas de déterminants au sens propre du terme⁶², et que les deux types d'emploi (adnominal et pronominal) recourent au même paradigme de formes, les sources de référence faisant la distinction restent toujours rares (voir la Section 3.1.3). Ainsi, dans le CNT, les deux emplois sont subsumés sous la même catégorie de « pronom ». Sur le plan de l'annotation des corpus, l'absence de cette distinction est fatale. Il a ainsi fallu, sur le plan de la construction de la requête, essayer de trouver le compromis optimal. Plusieurs phénomènes ont dû être pris en compte en même temps. Avant tout, il a fallu écarter autant que faire se pouvait toutes les occurrences où le démonstratif était en emploi autonome (pronominal). Mais il a aussi fallu tenir compte de l'ordre des mots en tchèque et, plus particulièrement, de la position de l'adjectif. Comme les adjectifs tchèques sont antéposés, il peut naturellement arriver qu'un ou plusieurs d'entre eux se glissent entre le démonstratif et le nom. Pour ce qui est du nom et de ses compléments, l'ordre canonique du tchèque est le suivant :

1) to	krásn-é	červen-é	jablk-o
ten.N.NOM.SG	beau-N.NOM.SG	rouge-N.NOM.SG	pomme-NOM.SG
ze soused-ova		strom-u	
de du voisin.ADJposs-M.GEN.SG		arbre-GEN.SG	
<i>cette belle pomme rouge de l'arbre du voisin</i>			

Comme nous l'expliquons dans la Section 3.1.3 en nous référant à l'étude de Veselovská (2014), cet ordre est celui de la quasi-totalité des complexes nominaux en tchèque, même si, théoriquement, le démonstratif ainsi que le ou les adjectif(s) peuvent se retrouver postposés au nom (ou, éventuellement, le démonstratif seul peut se retrouver postposé à un ou plusieurs adjectifs au sein de la structure antéposée). Pourtant, ces cas – très marqués stylistiquement – sont d'une grande rareté dans le tchèque contemporain et a fortiori dans la langue parlée informelle, où ils font presque complètement défaut⁶³.

⁶² Voir le parti-pris adopté dans l'Introduction générale, qui sera par la suite développé dans les Chapitres 3, Section 3.1.3.

⁶³ Même pour ce qui est de la fiction du 19^{ème} et du 20^{ème} siècle, et contrairement à ce que l'on pourrait croire (notamment au sujet de certains auteurs tchèques « archaïsants », tel K. Čapek), Veselovská (2014) a montré en travaillant avec le CNT que dans un corpus composé de 11 auteurs, ces cas de figure concernent à peine 6,5 % des complexes nominaux. Dans ORTOFON v1, nous n'avons trouvé aucune occurrence de ce type.

Outre l'adjectif, d'autres éléments adnominaux ressortissant au domaine fonctionnel (possessifs, indéfinis, etc. ; Abney, 1987) peuvent accompagner le démonstratif devant le nom (voir l'Introduction générale et la Section 3.1.3 du Chapitre 3)⁶⁴. Aussi a-t-il fallu anticiper ces possibilités. En même temps, la requête devait éliminer un maximum d'occurrences bruitées (erreurs), dont la probabilité d'apparition va croissant à mesure que la distance entre le démonstratif et le nom s'allonge. Une partie considérable de ce bruit était constituée par des formes du démonstratif indéfini TAKOVÝ (TEL) ou par l'exclamation *ty vole*, où *ty* n'est pas un démonstratif, mais le vocatif du pronom de la deuxième personne du singulier. Ces deux phénomènes ont dû être explicitement écartés. Voici la requête appliquée aux textes d'ORTOFONv1 :

```
[pos="PRO\DEM" & !lemme="takov.*"] [pos="VER|CON|PRE|NOM|PCT"]{0,3}
[pos="N.*" & !word="vole"]
```

Arrêtons-nous à présent sur les différentes séquences de la requête. Commençons d'abord par la séquence [pos="VER|CON|PRE|NOM|PCT"]{0,3}. La partie {0,3} signifie que le nombre de positions occupées par d'autres éléments intercalés entre le démonstratif et le nom/numéral peut aller de zéro à trois. La partie [pos="VER|CON|PRE|NOM|PCT"] signifie que, de cet intervalle, ont été exclus les verbes, conjonctions, prépositions, noms et particules. Cette option prévoit la possibilité d'avoir dans la séquence un ou plusieurs adjectifs ou éléments fonctionnels antéposés au nom/numéral tout en réduisant le nombre d'occurrences bruitées. La séquence [pos="N.*" & !word="vole"] indique que le démonstratif introduit un nom ou un numéral, à l'exception de la forme *vole*, vocatif du nom VŮL (BŔUF) et faisant partie d'une interjection exclamative. Concernant la toute première séquence [pos="PRO\DEM" & !lemme="takov.*"], elle permet de rechercher toutes les formes du démonstratif nominal (Diessel, 1999) défini, en écartant le paradigme du démonstratif indéfini TAKOVÝ (TEL). Nous rappelons que notre requête visait toutes les formes (lemmes) du démonstratif et non seulement TEN⁶⁵.

Cette requête est loin d'être parfaite : elle a certainement échoué à isoler toutes les occurrences du démonstratif adnominal en tchèque ; de plus, malgré toutes les précautions prises, le résultat continuait tout de même à contenir à peu près 15,7 % d'occurrences bruitées⁶⁶, que nous avons ensuite

⁶⁴ Sachant que le démonstratif peut précéder ces autres mots fonctionnels (grammaticaux) et/ou les suivre. Voir à cet égard la Section 3.1.3.

⁶⁵ Bien que notre principale hypothèse de recherche porte sur TEN, nous avons aussi cru utile de regarder la présence des autres lemmes, ne serait-ce que 1) pour comparer leur fréquence avec celle de TEN et, ce faisant, montrer l'immense prévalence de ce dernier au sein des données et 2) pour pointer les différences dans le fonctionnement sémantico-référentiel entre une forme à mi-chemin entre le démonstratif et l'article défini (TEN, en l'occurrence) et des formes qui sont de purs démonstratifs.

⁶⁶ Voir la Section 2.2.4 *infra* pour l'explication de l'obtention de ce chiffre.

dû éliminer manuellement (notamment les occurrences de la forme TEN en emploi pronominal et les occurrences du pronom personnel TY)⁶⁷. Mais elle semble à ce jour la meilleure possible.

Pour le CEFC, dont les annotateurs avaient opéré la distinction entre emplois pronominaux et adnominaux (déterminants) du démonstratif, il était beaucoup plus aisé de formuler la requête :

[pos="DET.DEM"]⁶⁸

Mais même ici, la liste d'occurrences générée avec cette requête contenait 7,15 % d'occurrences bruitées, ces occurrences étant avant tout : 1) des emplois de la forme *ce* pronom et 2) des emplois de CE déterminant, mais dans des contextes de disfluency, où la même forme est répétée plusieurs fois :

- 2) c'est donné euh et puis de enfin tu dois pouvoir avoir **ce ce ce voyage** tu dois pouvoir pour l'Es-pagne c'est comme si tu partais de de Lyon pour euh
CEFC

Dans le concordancier des résultats d'ORTOFON v1, ces occurrences s'affichent sur la même ligne et comptent comme une seule. Par conséquent, en français, nous les avons également comptées comme une seule.

2.2.2 Interroger les corpus : InterCorp v10

Passons maintenant à notre sous-corpus créé dans InterCorp v10. C'est le seul parmi les trois ensembles de données que nous n'avions pas importé dans TXM. Nous l'avons interrogé directement en ligne, en utilisant la requête suivante⁶⁹ :

[tag="P[D].*" & !lemma="takový"] [tag="N.*"]

Expliquons : la séquence tag="P[D].*" cible les « pronoms démonstratifs » (c'est-à-dire tous les démonstratifs : rappelons que le système d'annotation des textes tchèques du CNT ne permet pas de faire la différence entre emplois adnominaux et pronominaux) ; la séquence !lemma="takový" élimine autant que faire se peut les formes du démonstratif indéfini TAKOVÝ ; et, finalement, [tag="N.*"] indique que nous cherchons les cas où le démonstratif introduit un nom. Comme dans les deux cas précédents, du bruit s'est frayé un chemin dans la concordance. De plus, l'inclusion de

⁶⁷ Une part de ces occurrences bruitées sont dues, l'on s'en doute bien, aux défis posés par le traitement de la langue parlée (tels que les difficultés de désambiguïsation de formes polysémiques et homonymes).

⁶⁸ Nous avons dû écarter la possibilité de rechercher le lemme CE, ce dernier incluant aussi les occurrences de la forme pronominale.

⁶⁹ Cela explique les écarts entre cette requête et les requêtes formulées pour les corpus de conversations (cf. « lemme » vs « lemma », etc.) : la langue de requête utilisée par TXM n'étant pas tout à fait la même que celle utilisée par l'interface d'InterCorp (voir *supra*).

positions vides entre le démonstratif et le nom pour trouver les cas où ceux-ci sont séparés par des adjectifs ou des mots fonctionnels s'est ici avérée trop problématique, car elle a généré une quantité insupportable de bruit. Aussi avons-nous préféré ne pas coder ces cas de figure dans la requête : après avoir formulé notre requête, nous avons lu les textes des pièces pour récupérer les cas restants (soit 44 occurrences), que nous avons ajoutés manuellement à la liste Excel des occurrences générées à partir de cette requête. Ici aussi, un tri était nécessaire au départ (en raison d'un nombre considérable d'occurrences où le démonstratif était employé de façon pronominale, de la présence de certains démonstratifs indéfinis, etc.). Lorsqu'une occurrence se répétait à l'identique (une intention de l'auteur), nous ne comptons cette occurrence qu'une seule fois⁷⁰. Nous avons également écarté quelques occurrences apparaissant dans les didascalies.

2.2.3 Annoter les données : catégories et valeurs d'annotation

Nous allons maintenant décrire notre système d'annotation. Le processus d'annotation a eu lieu dans Excel, où nous avons au préalable exporté les résultats des requêtes sous forme de concordances. C'est également là que nous avons éliminé les occurrences bruitées. Si les catégories diffèrent partiellement en fonction de la provenance des textes, une partie d'entre elles reste constante. Pour toutes les occurrences, nous nous sommes intéressé au type d'usage du démonstratif, à savoir au régime d'indexicalité dans lequel le démonstratif apparaissait. Sur ce plan, notre typologie est notamment inspirée par les travaux d'Himmelman (1996), de J. Lyons (1990 [1978]), de Fraser et Joly (1979 et 1980), de Corblin (1985) et de Kleiber (1986a, 1988 et 1989). Dans son article de 1996, Himmelman distingue les emplois suivants du démonstratif : l'usage déictique situationnel ("situational use"), la deixis discursive ("discourse deictic use"), l'usage endophorique ("tracking use") et l'usage mémoriel ("recognitionnal use"). Nous rappelons que la deixis discursive a également été théorisée par d'autres auteurs, tels que Guillot (2006a) ou encore Kara et Wiederspeil (2001). Les deux derniers, se situant dans un paradigme davantage cognitif, parlent d'« anaphore résomptive ». Les deux termes étant synonymes, nous avons choisi le second, pour des raisons de pure préférence personnelle⁷¹. Rappelons-en ici la définition donnée par Himmelman (1996 : 224) : « Discourse deixis is to be understood here as reference to propositions or events. » La spécificité de ce mode indexical est due au fait qu'il se situe à mi-chemin entre l'endophore et la deixis. Son versant endophorique repose sur la mobilisation de sens et de références apportés par le discours ; son versant déictique, sur la mobilisation du texte amont (ou aval) dans sa matérialité, car l'expression démonstrative ne se contente pas ici de reprendre un référent tout prêt : il en forge un nouveau à partir

⁷⁰ Au total, cela s'est produit trois fois dans le sous-corpus.

⁷¹ Un autre synonyme de cet emploi du démonstratif, utilisé par J. Lyons, est celui de « deixis textuelle impure » (1990 [1978]).

de ce qui précède ou suit, ce réfèrent en tant que tel n'existant pas encore. Himmelmann (1996) et Guillot (2006a) montrent que les démonstratifs employés dans ce régime sont le plus souvent des pronoms, mais que ce n'est pas nécessairement le cas. Voici un exemple pour chaque langue :

- 3) L1 : on me disait fais attention quand tu prends le tramway euh prends le tramway accompagnée euh et puis euh si on te parle tu réponds pas parce que si tu parles français euh voilà quoi et ouais ouais

L2 : *à ce point là* quand même

CEFC

- 4) mně to přijd-e úplně divn-ý no nebo to
moi.DAT cela.NOM sembler-PRS.3SG complètement bizarre-N.NOM.SG MOD ou cela.NOM

jak zpíva-l-a Basikov-á v tý latin-ě víš
REL chanter-PST-3SG.F NP-F.NOM.SG en ten.F.LOC.SG latin-LOC.SG MOD

[...].. *tenhle typ modern-ích písni*
tenhle.M.NOM.SG genre.NOM.SG moderne-F.GEN.PL chanson-GEN.PL

*moi, je trouve ça complètement bizarre ; ou le truc que Basiková a chanté en latin, tu vois... ce genre de chansons modernes*⁷²

ORTOFON v1

Deux autres types de deixis ont dû être inclus dans les annotations : la « deixis textuelle » et la « deixis temporelle ». Même si la première n'est pas reconnue par tous les auteurs, il ne saurait y avoir, comme on le verra, aucun doute sur l'utilité de la présence de cette catégorie pour notre matériel empirique. On la trouve par exemple chez J. Lyons (1990 [1978]), qui parle de « deixis textuelle pure » et « impure ». Alors que le second terme est juste un autre synonyme pour l'« anaphore résomptive » (et donc pour « la deixis discursive »), le premier renvoie aux cas où le texte est purement considéré dans sa matérialité. L'acte de référence n'y vise plus un référent tout prêt apporté par le discours (l'endophore) ni un segment plus long et qu'il s'agirait d'élever au statut de nouveau référent synthétique (l'anaphore résomptive) : il vise le texte lui-même. Ce dernier peut être mobilisé en sa qualité d'espace : *ci-dessus*, *ci-dessous*, *infra*, *supra*, etc., ou bien ce sont les signes le constituant qui sont mobilisés dans leur matérialité, en tant que signifiants et/ou en tant que signifiés :

- 5) L1 : Ça, c'est un rhinocéros.

L2 : Un quoi ? tu peux me *l'*épeler ? (J. Lyons, 1990 [1978] : 289)

Quant à la seconde catégorie, c'est-à-dire la deixis temporelle, il convient assurément de la séparer de la deixis situationnelle (spatiale). Rappelons que cette répartition de la deixis (avec, comme troisième catégorie, la deixis personnelle) a une longue histoire et qu'elle est déjà apparue chez les

⁷² Cet exemple montre que malgré la définition d'Himmelmann présentée *supra*, le référent d'une anaphore résomptive n'est pas nécessairement un événement ou une proposition (soit des entités dites « du second et du troisième ordre » par Lyons, 1990 [1978] : 77-81), mais qu'il peut aussi s'agir d'une entité concrète (« du premier ordre »).

auteurs du paradigme énonciatif de l'indexicalité (voir la Section 1.1.1). Comme son nom l'indique, on y range les expressions déictiques référant aux repères temporels de l'énonciation, dont certaines prennent le démonstratif : *aujourd'hui*, *demain*, *ce soir*, *cet été*, etc. Comme nous le verrons plus tard (Sections 4.5.1, 5.2.2.1 et 5.2.3.1), l'une des singularités sur lesquelles nos corpus nous ont permis de mettre le doigt est un emploi bien plus abondant du démonstratif dans ce régime en français qu'en tchèque.

S'agissant des emplois « mémoriels » (“recognitional uses” ; Himmelmann, 1996), le terme d'« exophore mémorielle » de Fraser et Joly (1979 et 1980) peut également être utilisé, quoiqu'il s'inscrive au sein d'une approche essentiellement onomasiologique (voir la Section 1.1.2 du Chapitre 1).

Les emplois strictement endophoriques se répartissent en anaphoriques et cataphoriques. En ce qui concerne l'anaphore, nous avons distingué dans nos annotations entre plusieurs sous-types, à savoir : l'anaphore fidèle, infidèle, indirecte, résomptive, associative et, en plus de cela, un autre type, pouvant difficilement être rangé dans aucune de ces catégories, et que nous avons pris la décision d'appeler l'« anaphore temporelle ». Traditionnellement, l'anaphore fidèle suppose une identité lexicale de l'expression anaphorique avec l'antécédent. Notre propre acception de cette notion est un peu plus large. En effet, le nom anaphorique introduit par le démonstratif est doté à quelques détails près du même sens que le nom de l'expression antécédente :

6) [un film]_i → [***ce film***]_i CEFC

7) [...] l'histoire [du chat]_i → [...] il est devenu quoi [***ce petit chat***]_i CEFC

L'ajout « à quelques détails près » signifie ici, comme cela advient dans l'exemple 7, que de légères modifications du sens peuvent apparaître : dans la reprise anaphorique, le nom peut être assorti d'une épithète ou d'une expansion (ou, au contraire, en être dépourvu), il peut être transformé en diminutif, etc. Nous avons opté pour cet élargissement de la notion afin de la rendre plus opérante pour le travail avec notre matériel empirique. Pourtant, il fallait toujours pour nous que dans la reprise, le substantif de la tête nominale garde la même racine lexicale que celui de l'antécédent. Ainsi, lorsqu'un synonyme était employé dans la reprise, celle-ci était considérée comme un cas d'anaphore infidèle⁷³, à savoir celui où le sens des deux noms n'est pas le même :

8) [un mec]_i → [***ce gars***]_i

⁷³ Dans les travaux tchèques, c'est le terme d'« anaphore coréférentielle indirecte » („nepřímá koreferenční anafora“) qui apparaît. Voir par exemple Pešek (2014).

Le terme d'« anaphore indirecte » – plus ou moins synonyme, dans la tradition linguistique française, de ceux d'« anaphore divergente », « lexicale », « in absentia », etc. (voir à sujet par exemple Corblin (1985), Kleiber (1988 et 1989) ou Decool-Mercier et Akinci (2010)) – représente une catégorie assez complexe et hétérogène. Cependant, tous les cas qui en relèvent ont ceci en commun que soit il n'y a pas du tout de coréférence entre l'antécédent et l'expression anaphorique (exemples 9 et 10) soit que la coréférence n'est que partielle (exemple 11) :

9) Tu as aimé l'exposé de Paul ? J'ai préféré *celui de Patrick*.

10) J'ai vu dix lions et toi tu *en* as vu quinze. (Milner, 1976 : 64)

11) tu prends une bouteille de gaz tu sais les grosses bouteilles de gaz que tu utilises pour [la cuisine],
ou [*ce genre de choses*]_{i+} ? CEFC

Avec le démonstratif adnominal, seul le second cas de figure peut se présenter. À propos de l'exemple 11, nous estimons qu'il y a coréférence partielle dans la mesure où le SN *ce genre de choses* reprend le référent du SN *la cuisine* mais, en plus de cela, il renvoie à toute une classe de référents dont la cuisine fait partie⁷⁴.

Un autre type d'anaphore a dû être introduit dans notre système d'annotation vu sa présence assez élevée parmi nos occurrences tchèques et, surtout, son rôle crucial dans la grammaticalisation de TEN en article défini : il s'agit de l'anaphore dite « associative ». Parmi les nombreux auteurs qui ont travaillé sur ce phénomène, nous avons choisi de nous référer, avant tout, à Löbner (1985, 1998 et 2011), à Schwarz (2009), à Ortmann (2014) et à Czardybon (2017), tous les quatre soulignant le rôle de l'anaphore associative dans l'émergence et la propagation de nouveaux articles définis. Dans une anaphore associative, la définitude du nom est enclenchée en première mention par un lien associatif entre celui-ci et un terme précédemment introduit :

12) Nous sommes arrivés dans une ville inconnue. *La mairie* était un somptueux édifice baroque.

13) Po kin-ě se ještě bavi-l-i o (*tom*) *film-u*.
après cinéma-LOC.SG REFL encore discuter-PST-3PL.M de ten.M.LOC.SG film-LOC.SG
Après être allés au cinéma, ils ont encore discuté du film. (Czardybon, 2017 : 90)

Schwarz (2009) distingue lui-même deux types d'anaphore associative, à savoir celle basée sur le lien partie-tout ("part-whole") et celle basée sur une simple relation entre les deux entités ("relational"). Löbner élargit cette distinction à un troisième type, mobilisant un lien situationnel

⁷⁴ Nous pouvons rapprocher ce cas de ce que Kleiber (1988 : 3-4) appelle l'anaphore « générique », citant l'exemple suivant emprunté à Corblin : [*Un caniche*]_i m'a mordu. [*Ces chiens*]_{i+} sont pourtant très gentils en général.

(“situational definite associative anaphora” ; Czardybon, 2017 : 66). Notre exemple de la mairie relève du rapport partie-tout (la mairie constitue une partie de la ville). L'exemple du film relève du type « situationnel » (le film visionné faisant partie de la situation qu'est une sortie au cinéma). Voici enfin un exemple du type relationnel :

14) J'ai pris un bus l'autre jour. **Le chauffeur** était un homme gaillard.

Nous aurons l'occasion de revenir sur cette distinction, qui s'avère très utile pour la description de la grammaticalisation de TEN adnominal en défini, dans la Section 6.2.2 du Chapitre 6. Le français n'est pas ici concerné, car, comme cela transparaît dans les exemples cités ici, le démonstratif est proscrit dans ce contexte.

L'anaphore que nous appelons « temporelle » concerne presque exclusivement le français (une seule occurrence en tchèque contre 44 occurrences en français)⁷⁵. Dans ce type d'anaphore, le complexe nominal démonstratif, contenant un nom à sémantique temporelle (*fois, moment, jour, temps, année*), s'interprète par rapport au discours antérieur, mais ne reprend aucun référent déjà mentionné en tant que tel. Il a notamment pour objectif de situer l'événement dans le temps, éventuellement aussi de le contraster à d'autres événements du même type (notamment avec le nom *fois*) :

15) L1 : on est allés à une plage c'était une plage pas de sable fin on va dire de petits graviers

L2 : ouais oh ça fait un peu plus mal déjà ça quand tu t'installas

L1 : donc ça j~ j'aime j'aime pas trop mais alors **ce jour-là** il y avait un petit peu de vent

16) ils ont un ami qui est euh qu'ils avaient invité justement là aussi à son anniversaire puis qui a pas pu venir puis la dernière fois qu'il est allé chez eux il a mangé la fondue Gerber puis il leur a dit euh bon je reviens manger la fondue chez vous mais **cette fois** vous me faites vraiment une vraie fondue

Concernant l'endophore cataphorique, nous n'avons pas explicitement distingué, dans notre système d'annotation, de sous-catégories au sein de cette catégorie, étant donné que sa présence parmi les occurrences est beaucoup moins élevée que celle de l'anaphore. Pourtant, nous avons été particulièrement sensible, en tchèque, à ce que nous appelons à l'instar d'autres auteurs la cataphore « structurale » ou « intra-syntagmatique » (voir l'Introduction générale et la Section 1.1.2 du Chapitre 1).

Outre ces catégories, renvoyant à des emplois globalement bien connus du démonstratif, nous avons été forcé d'introduire d'autres étiquettes au fur et à mesure de notre travail avec les textes. Cela

⁷⁵ L'une des raisons de cette rareté est certainement le fait que le tchèque emploie les adverbes déictiques (*pro*)*tentokrát* pour *cette fois*-(ci/là) et *tehdy* ou *tenkrát* pour *à cette époque*-(là) / *en ce temps*-là.

concerne d'abord la catégorie des emplois « indéfinis » du démonstratif. Il s'agit bel et bien d'une réalité dans nos occurrences pour les deux langues, même si la littérature semble se taire sur le sujet :

17) L1 : faut pas se laisser aller

...

c'est ça vu que bon comme tu es à la fac tu es libre tu fais un peu euh tu vas à **ce cours**

L2 : ouais il faut pas se dire à chaque fois qu'il y a une heure de trou bon je vais je vais rien faire
CEFC

18) představ-a že . prostě chod-í-m se ps-em ven a pokaždé
idée-NOM.SG que MOD aller-PRS-1SG avec chien-INS.SG dehors et chaque fois

se boj-í-m abych potka-l-a **tamty**
REFL avoir peur-PRS-1SG que.CONJ+AUX.COND.1SG croiser-PTCP-SG.F tamten.M.ACC.PL

a **tamty** **soused-y** protože se na mě bud-ou
et tamten.M.ACC.PL voisin-ACC.PL parce que REFL sur moi.ACC AUX.FUT-3PL

ksichtit ..

faire la gueule.INF

genre l'idée d'avoir peur; à chaque fois que je sors promener mon chien, de croiser ces voisins-là ou ces voisins-là qui vont me faire la gueule...
ORTOFON v1

La particularité de *ce cours* et *tamty sousedy* est que le démonstratif ne permet pas d'identifier un référent défini, son rôle étant purement de donner un exemple : « tu vas à un/tel cours (ou un/tel autre) » / « j'ai peur de rencontrer tels voisins ou tels autres ».

Un certain nombre d'occurrences se sont avérées impossibles à ranger dans aucun des types d'emploi énumérés *supra*, le plus souvent faute d'un contexte suffisamment large. Elles ont ainsi été marquées comme « ambiguës ».

Enfin, nous avons distingué deux autres types d'emploi, dont l'un concerne uniquement le tchèque, à savoir les « emplois émotionnels » et les « emplois métalinguistiques ». Pour résumer, le premier regroupe toutes les expressions démonstratives où le démonstratif ne participe pas (directement) au processus d'identification du référent (cf. devant un nom propre, un nom à référent unique ou générique). Les emplois que nous avons appelés « métalinguistiques » (n'apparaissant qu'en tchèque) sont ceux où le démonstratif est utilisé comme une sorte de béquille par un locuteur à la recherche du mot ou de l'expression juste. D'habitude, cet usage est signalé par la présence de traits prosodiques, notamment une pause entre le démonstratif et le nom qu'il introduit, suivie d'une accentuation de la forme du démonstratif (voir les Sections 3.3.3 et 5.2.2.5) :

19) no a Páj-e kvet-e ibišek úplně ve velk-ým ..
MOD NP-F.DAT.SG fleurir-PRS.3SG hibiscus.NOM.SG complètement en grand-N.LOC.SG

a ten	můj	by	asi	potřebova-l
et ten.M.NOM.SG	mien.M.NOM.SG	AUX.COND.3SG	MOD	avoir besoin-PTCP.SG.M

ostříhat	všechn-y	ty ..	všechny	větv-e
couper.INF	tout-F.ACC.PL	ten.F.ACC.PL	tout-F.ACC.PL	branche-ACC.PL

Pája, son hibiscus, il est tout en fleurs... ; et le mien, il faudrait lui couper toutes les... toutes les branches
ORTOFON v1

Ce type d'emploi singulier est relevé, entre autres, par Schneiderová (1993) et par Trovesi, qui constate sa présence non seulement en tchèque, mais aussi en haut-sorabe (la forme TÓN) :

Je voudrais encore attirer l'attention sur un usage singulier des pronoms *ten* et *tón* en tant que mots de remplissage, soit sur les cas où le locuteur tente de surmonter une incertitude lexicale lors de la quête de l'expression appropriée [...] ⁷⁶ (2002 : 133)

Après les différents types et sous-types d'emplois du démonstratif adnominal, venons-en à présent aux autres catégories d'annotation. Pour les deux corpus oraux, nous avons relevé les marques de disfluente accompagnant l'emploi du démonstratif adnominal. Ici, il s'agit de plusieurs phénomènes – amorces inachevées du complexe nominal, répétitions de la forme du démonstratif, marques d'hésitation, pauses entre le démonstratif et le nom, interruptions – ou de leur combinaison. Remarquons que dans ORTOFON v1, les pauses sont marquées par les points (un, deux ou trois, proportionnellement à la longueur de la pause) :

20) c'est donné euh et puis de enfin tu dois pouvoir avoir **ce ce ce voyage** tu dois pouvoir pour l'Espagne c'est comme si tu partais de de Lyon pour euh CEFC

21) tady začín-a-j příloh-y což bud-e úplně
ici commencer-PRS-3PL annexe-NOM.PL ce qui.NOM être.FUT-3SG complètement

strašn-ej	masakr	protože	to	jsou
horrible-M.NOM.SG	massacre.NOM.SG	parce que	ce	être.PRS.3SG

ty	různ-ý	stadi-a	zpracováván-í	těch .. těch
ten.N.NOM.PL	divers-N.NOM.PL	stade-NOM.PL	traitement-GEN.SG	ten.M.GEN.PL

lexém-ů

lexème-GEN.PL

ici, ce sont les annexes qui commencent, et ça, ça sera un vrai massacre parce que ce sont les différents stades du traitement des... des lexèmes
ORTOFON v1

La raison de ce choix est qu'il pourrait s'agir d'importants effets pragmatiques en rapport avec le démonstratif adnominal, effets qu'il convient d'explorer davantage. Les deux phénomènes

⁷⁶ Nous traduisons : „Chtěl bych upozornit ještě také na zvláštní užívání zájmen *ten* a *tón* jako slov výplňkových, tedy na případy, kdy se mluvčí snaží překonat lexikální nejistotu při hledání správného výrazu [...]“

précédemment mentionnés nous amènent à un constat très important, à savoir que pour bien évaluer les traits qui sont de nature prosodique, il a fallu travailler avec les enregistrements des conversations : raison de plus pour avoir recours à l'interface en ligne.

Dans les deux corpus oraux, nous nous sommes également intéressé à la position du complexe nominal démonstratif par rapport au verbe de la proposition où il apparaît pour essayer de voir à quel point ces deux phénomènes (c'est-à-dire l'apparition du démonstratif et sa position dans la proposition) pourraient se trouver corrélés⁷⁷. Sept valeurs ont été retenues : position préverbale, position postverbale, complément de phrase, dislocation à gauche, dislocation à droite, clivage (pour le français uniquement) et position inconnue. Dans ce dernier cas, deux situations se présentaient : soit la position ne pouvait être établie, car la réplique n'était pas de nature propositionnelle (cas assez courant dans la langue orale informelle ; cf. Bally, 1951 [1909] ; Blanche-Benveniste, 1990 ; Rioul, Pellat et Riegel, 2016 [1994]) soit la position était trop ambiguë pour qu'il soit possible de trancher. Voici un exemple par langue pour chacune de ces valeurs :

Position préverbale

- 22) *cette cabine* fonctionnera bientôt avec les télécartes CEFC
- 23) tak dlouho **ty** **tapet-y** tam byl-y
 si longtemps ten.F.NOM.PL papier peint-NOM.PL là-bas être.PST-3PL.F
ça faisait si longtemps que les papiers peints étaient là ORTOFON v1

Position postverbale

- 24) et moi j'ai beaucoup apprécié *cet aspect théâtral* CEFC
- 25) tady jsou **ty** *promenád-y* ne ? kolem
 ici être.PRS.3PL ten.F.NOM.PL promenade-NOM.PL MOD autour
ici, il y a les/ces promenades tout autour, non ? ORTOFON v1

Complément de phrase

- 26) à *cette époque* ma tante c'était sa patronne CEFC
- 27) na **ty** *nej-rušnějš-í* *ulic-i* v Los Angeles
 sur ten.F.LOC.SG SUP-plus animé-F.LOC.SG rue-LOC.SG à Los Angeles.LOC.SG
- jsou** **takov-ý-dle** **upoutávk-y**
 être.PRS.3PL tel-F.NOM.PL-SUFFIXE DEICTIQUE panneau-réclame-NOM.PL

⁷⁷ Par « proposition », nous entendons le terme utilisé « [...] pour identifier, dans les phrases complexes, les “phrases constituantes” soit enchâssées dans une autre phrase, soit combinées par juxtaposition ou coordination [...] » (Rioul, Pellat et Riegel, 2016 [1994] : 784).

sur l'avenue la plus animée à Los Angeles, il y a cette espèce de panneaux-réclames

ORTOFON v1

Dislocation à gauche

28) [**ce mec**]_i [ɛ']_i est quand même une blague

29) a [**ten** **olivov-ej** **olej**]_i .. co jsme kupova-l-i ..
et ten.M.NOM.SG d'olive-M.NOM.SG huile.NOM.SG.REL AUX.PST.1PL acheter-PTCP-PL.M

tak . [**ten**]_i je taky už podraž-en-ej
alors celui-ci.M.NOM.SG être.PRS.3SG aussi déjà plus cher-M.NOM.SG

et cette/l'huile d'olive... qu'on achetait.. alors elle aussi, elle coûte déjà plus cher

ORTOFON v1

Dislocation à droite

30) oh mon dieu m' [**en**]_i parle plus de [**cette fiche**]_i

CEFC

31) bab-a mně řek-l-a že .. [**vona**]_i k ní
bonne femme-NOM.SG moi.DAT dire-PST-3SG.F que elle.NOM chez elle.DAT

chod-í na svačink-u . [**ta** **kočk-a**]_i
aller-PRS.3SG à goûter-ACC.SG ten.F.NOM.SG chatte-NOM.SG

la bonne femme, elle m'a dit que... qu'elle venait chez elle à l'heure du goûter, la chatte

ORTOFON v1

Clivage (en français uniquement)

32) c'est **cette semaine** de toute manière hum que je vais avoir la réponse

CEFC

Position inconnue

33) tu sais Marseille les gens les jeunes qui sont pas trop branchés la mode comme il dit latino machin
truc **ce genre de choses**

CEFC

34) no tak co s **tím** **byt-em** teda ?
MOD que.? avec ten.M.INS.SG appartement-INS.SG MOD
du coup, cet appartement ?

ORTOFON v1

Ces valeurs appellent quelques commentaires. Dans l'approche de la structure thème-rhème, la position préverbale est traditionnellement associée au thème, tandis que la position postverbale est associée au rhème (cf. Mathesius, 1947 ; Daneš, 1974). Cette dichotomie formelle est adoptée par Zíková dans son travail étudiant la grammaticalisation de TEN en contexte anaphorique (2017). Comme l'auteure a prouvé la pertinence de ce facteur pour son sujet, nous avons suivi son exemple, reproduisant cette distinction dans notre propre travail. Outre ces deux cas de figure, Zíková (2017)

distingue également la dislocation à droite. Dans ce dernier cas, le constituant est extrait de la proposition (et n'en fait donc plus partie) pour être placé à sa droite. Il est représenté au sein de la proposition par un pronom (exemples 30 et 31). Contrairement à Zíková, nous avons également identifié dans notre matériel des cas relevant de la dislocation à gauche, où le constituant est extrait de la proposition pour être placé à sa gauche (il est également représenté par un pronom intra-propositionnel ; exemples 28 et 29). La dislocation a été théorisée par exemple par Givón (1983), Lambrecht (2001) et par Kalbertodt, Primus et Schumacher (2015). Si, dans les deux types de dislocation, le constituant est thématique, il n'est pas présenté de la même façon et ne joue pas le même rôle par rapport à la structure à l'intérieur de laquelle il s'inscrit : alors que la dislocation gauche semble davantage commencer une nouvelle chaîne de référence, introduite par le constituant disloqué et mis en relief, la dislocation droite sert davantage à maintenir la chaîne en cours tout en mettant l'accent sur l'information nouvelle (le rhème) apportée par la proposition. Ainsi, les constituants des dislocations droites seraient plus thématiques que ceux des dislocations gauches. Zíková (2017 : 237) écrit ensuite que dans les deux cas, l'identification du référent repose sur un mécanisme différent :

[...] les référents des dislocations à gauche sont définis sur la base d'une connaissance discursive (ils se recrutent parmi les référents discursifs que sont, typiquement, les personnages principaux et les objets conceptuellement distincts), les référents des dislocations à droite sont plus souvent identifiables sur la base du contexte situationnel [...]⁷⁸

S'appuyant sur la littérature existante, Zíková (2017 : 236-237) énumère une série de critères qu'elle tient pour définitoires pour la notion de dislocation et que nous reproduisons ici : a) le constituant est extrait de la proposition et remplacé par un pronom au sein de celle-ci⁷⁹, b) le constituant et le pronom qui le remplace s'accordent morphologiquement (en genre, en nombre et en cas), c) le constituant disloqué ainsi que la proposition sont séparés par un contraste intonatif (voire une pause), si bien que le constituant représente une unité prosodique à part⁸⁰ et d) les constituants disloqués sont toujours définis. Cependant, comme nous avons pu le constater en travaillant avec nos données, les critères a), b) et d) sont loin d'être toujours respectés.

Pour ce qui est du premier, il est possible, selon nous, d'identifier des cas en tchèque où le constituant, tout en étant extrait de la proposition, n'est pas représenté dans celle-ci par un pronom, mais par une anaphore zéro (marquée par le symbole Ø en 35) :

⁷⁸ Nous traduisons : „[...] referenty LD jsou určité na bázi znalosti diskursní (rekrutují se z diskursních referentů, typicky hlavních postav a konceptuálně vyhraněných předmětů), referenty PD jsou častěji identifikovatelné na základě situačního kontextu [...]“

⁷⁹ La littérature d'expression anglaise emploie le terme de « clause ».

⁸⁰ Zíková remarque cependant que ce point est souvent source de controverses, voire de positions contradictoires.

- 35) [Ø]_i to úplně vodflák to [ten doctor]_i
 cela.ACC complètement bâcler.PST.3SG cela.ACC ten.M.NOM.SG médecin.NOM.SG
 [il]_i l'a complètement bâclé, [ce médecin]_i ORTOFON v1

Ce qui nous permet ici de soutenir qu'il s'agit d'une dislocation à droite malgré l'absence d'un pronom représentant dans la structure matrice, c'est la réalisation prosodique de l'ensemble : *ten doktor (ce médecin)* constitue justement une unité accentuelle autonomisée et séparée du reste par une pause. D'ailleurs, si l'énoncé est traduit en français, langue qui n'est pas pro-drop (à sujet nul), le pronom apparaît (voir notre traduction de l'exemple). Cependant, ce cas de figure n'est possible que lorsque le référent disloqué occupe la fonction de sujet.

Dans notre corpus d'oral informel français, il est quelques cas où le pronom représentant fait défaut dans la structure matrice. Nous tenons pour légitime l'analyse selon laquelle ce pronom, faisant l'objet d'une ellipse, peut être restitué (comme nous l'avons fait en 36), et que, par conséquent, nous sommes également face à une dislocation, même si nous avons tout à fait conscience que ce parti-pris est controversé⁸¹:

- 36) [cet inter-rail]_i je fais de la pub [pour lui]_i un petit peu CEFC

Ce cas de figure n'est pas rare à l'oral (*Le chocolat, j'adore*). De nouveau, c'est la prosodie qui favorise une analyse en dislocation.

Concernant l'accord morphologique, c'est peut-être le critère le moins respecté de tous, notamment parce que l'élément représentant s'avère souvent être un adverbe :

- 37) [tam]_i je prostě špatn-ý kontakt [mezi těma
 là-bas être.PRS.3SG MOD mauvais-M.NOM.SG contact.NOM.SG entre ten.F.INS.PL
 ližin-ama]_i asi
 glissière-INS.PL MOD
 [là]_i, il y a tout simplement un mauvais contact, entre [les glissières]_i ORTOFON v1

Notamment en français, les dislocations à gauche enfreignent couramment cette règle :

- 38) [Marie]_i, je [lui]_i ai donné un livre.

- 39) La bière, j'en bois beaucoup⁸²

Et, enfin, un énoncé comme 40 remet facilement en question – en tout cas en français – le critère de la nature définie de l'élément disloqué :

⁸¹ Prévost (2003) parle à ce titre de « constructions détachées non liées / sans coindexation ».

⁸² En tchèque, une structure analogue est possible, quoique moins courante et connotée familière : [Marie]_i, [tý]_i jsem dal knihu. Pivo, toho piju hodně.

40) *Du café*, j'en bois beaucoup.

Pour résumer : de même que celle de Zíková (2017), notre propre acception de la dislocation est certainement plus généreuse que celle des auteurs susmentionnés, sachant que le critère le plus important s'avère être le critère prosodique.

Disons à présent un mot au sujet des compléments de phrase. Ce phénomène syntactico-sémantique complexe a fait l'objet de nombreuses études (cf. le numéro 148 de *Langue française* de 2005, consacré aux « adverbiaux cadratifs », représentant un cas particulier des compléments de phrase). Dans les grammaires françaises (cf. Rioul, Pellat et Riegel, 2016 [1994]) – dont la délimitation du phénomène se montre suffisante pour le propos du présent travail –, les termes de « compléments circonstanciels extra-prédicatifs/extra-propositionnels » ou « compléments de phrase » ont tendance à figurer. Rioul, Pellat et Riegel (2016 [1994] : 265) montrent le statut particulier de ce type de circonstant en recourant à une série de critères à la fois syntaxiques et logico-sémantiques. Sur le plan syntaxique, le complément circonstanciel extra-prédicatif français se trouve le plus souvent en position de détachement, « [...] séparé du reste de la phrase par une pause ou par une virgule [...] » (cf. l'exemple 26). En tchèque, en revanche, il n'est, en règle générale, détaché que s'il est placé en fin de phrase. Placé à son début, il est intégré dans la matrice phrastique (cf. l'exemple 27)⁸³. Sur le plan logico-sémantique :

[...] le circonstant spécifie un élément du cadre général (souvent implicite) où se situe l'information véhiculée par le reste de la phrase. L'information véhiculée par ces circonstants périphériques parce que **extraprédicatifs** (extérieurs à la proposition et à l'attitude propositionnelle qui l'affecte) est présupposé et thématisé. C'est pourquoi ils restent hors du champ de la négation totale de la phrase, ne se prêtant pas à l'extraction rhématique par *c'est... que* et, pour la même raison, ne constituent pas une réponse à l'interrogation partielle [...]

Rioul, Pellat et Riegel (2016 [1994] : 265)

Ce comportement est facilement vérifiable :

41) Les jours rallongent **fin décembre**.

- Les jours ne rallongent pas fin décembre (= ils rallongent à un autre moment de l'année)
- C'est fin décembre que les jours rallongent (= et non pas à un autre moment de l'année)
- Quand les jours rallongent-ils ? Fin décembre.

vs **Fin décembre**, les jours rallongent. (Rioul, Pellat et Riegel, 2016 [1994] : 265)

⁸³ Cette différence est due à la différence typologique entre les deux langues. Tandis que le français moderne tolère rarement d'autres configurations que celle sujet-verbe-complément, le tchèque recourt bien plus aisément à des configurations s'écartant de ce schéma, dont, typiquement, celle complément-verbe-sujet. À la différence du français moderne, Marchello-Nizia (1995 : 35-112) explique que cet ordre était parfaitement naturel et courant en ancien français, langue qui maintenait encore une certaine flexion nominale.

Comme cela a déjà été suggéré, le complément de phrase n'est pas nécessairement placé en début de proposition, il peut également être placé à sa fin, voire au milieu⁸⁴ :

42) mon site il marchait pas **ce matin** CEFC

43) tu sais on voulait y aller **cet été** en Corse avec euh les Japonais CEFC

44) a co jako děl-a-j s **těma** **mask-ama**
 et que.ACC MOD faire-PRS-3PL avec ten.F.INS.PL masque-INS.PL
et ils font quoi, avec les masques ? ORTOFON v1

45) Kristýnk-a třeba v tom .. Vranov-ě .. **ty** **tři**
 NP-F.NOM.SG par exemple à ten.M.LOC.SG Vranov-LOC.SG ten.M.ACC.PL trois.ACC
 rok-y byl-a .. a normálně umejva-l-a záchod-y
 an-ACC.PL être.PST-3SG.F et normalement laver-PST-3SG.F toilette-ACC.PL
par exemple Kristýnka, elle est restée les trois ans à... Vranov (comme tu le sais)... et elle y lavait les toilettes et tout ORTOFON v1

Contrairement à la position antéposée, où la syntaxe permet de trancher immédiatement, dans ces deux derniers cas, c'est, de nouveau, le critère prosodique – changement d'intonation, pause –, ainsi que les tests de la négation, du clivage et de l'interrogation qui permettent d'établir (avec plus ou moins de certitude) le statut de complément de phrase du constituant. Pour la position finale, le critère prosodique est seul à même de faire la distinction entre un circonstant extra- et intra-prédicatif (voir les exemple 42 et 44). Pour la position interne à la phrase, ce sont plutôt les tests sémantiques qui permettent de trancher entre les deux interprétations (dans les deux langues, le complément extra-prédicatif occupant cette position semble pouvoir être prosodiquement intégré à la phrase ; voir les exemples 44 et 45).

Ensuite, sur le plan de la fonction syntaxique, force est de constater que les compléments de phrase ne sont pas nécessairement des compléments circonstanciels au sens strict du terme : il peut aussi s'agir de constituants que Rioul, Pellat et Riegel (2016 [1994] : 266) subsument sous le terme de « circonstants de commentaire énonciatif » (*selon moi, évidemment*) ou de connecteurs logiques/argumentatifs structurant le discours. Dans notre corpus français, c'était typiquement le cas de la locution anaphorique résomptive *dans ce cas-là* :

46) oui **dans ce cas là** vaut mieux prendre le bus CEFC

⁸⁴ Fait qui permettrait de distinguer entre un complément de phrase en général et un adverbial cadratif, lequel est toujours placé en position initiale et ouvre ainsi un cadre pouvant s'étendre sur plusieurs propositions, dont il indexe les prédications (cf. Charolles et Vigier, 2005).

Dans le corpus tchèque, il s'agissait, entre autres, de la locution (également anaphorique résomptive) *tím pádem* (du coup, par conséquent) :

47) já	m-á-m	totiž ty	špatn-ý	zád-a
moi.NOM	avoir-PRS-1SG	MOD ten.N.ACC.PL	mauvais-N.ACC.PL	dos-ACC.PL tantum
<i>tím</i>	<i>pád-em</i>	potřeb-uj-u	aby	mě
ten.M.INS.SG	cas-INS.SG	avoir besoin-PRS-1SG	que.CONJ+AUX.COND.3PL	moi.ACC
vopravova-l-i				
réparer-PST-3PL.M				

moi, j'ai un souci avec mon dos (comme tu le sais) ; du coup, j'ai besoin qu'on me le répare

ORTOFON v1

De manière générale, nous estimons que la prise en considération du phénomène des compléments de phrase est incontournable dans nos annotations, et ce non seulement pour le français, où il semble, de manière générale, atteindre des proportions plus grandes qu'en tchèque, mais aussi pour le tchèque (voir les exemples 27, 44, 45 et 47). Cette décision marque une rupture avec la pratique adoptée par Zíková (2017), ne considérant pas ce phénomène comme une catégorie à part. Dans son système d'annotation, les exemples 27, 45 et 47 seraient ainsi annotés comme des cas de position préverbale et l'exemple 44 comme un cas de position postverbale. Or les critères cités *supra* permettent d'établir que dans tous ces exemples, le constituant dont fait partie l'expression démonstrative fonctionne précisément comme un cadre par rapport à l'information prédicative exprimée dans la proposition : en 27, *na tý nejrušnější ulici v Los Angeles* (sur l'avenue la plus animée à Los Angeles) constitue ainsi un cadre pour l'information – rhématique – mentionnée juste après, à savoir la présence des panneaux-réclames.

Compte-tenu de la syntaxe de l'oral informel (interruption, aposiopèses, anacoluthes, nature elliptique des propos), il était souvent difficile de trancher à propos de la position d'une occurrence par rapport au verbe de la proposition. Par exemple, il arrivait que le verbe, quoique absent, soit facilement restituable. Que faire dans des situations comme celle-ci ? Fallait-il adopter un parti-pris très formel et conclure que la position ne pouvait pas être déterminée faute de présence explicite d'un verbe ? Nous ne le pensons pas :

48) tak od	<i>té</i>	<i>dob-y</i>	co	jsem	todle ..
alors depuis	ten.F.GEN.SG	temps-GEN.SG	REL	AUX.PST.1SG	cela.ACC
<i>du coup, depuis le moment où je l'ai...</i>					

ORTOFON v1

49) mais <i>cette fois</i>	euh comme comme	NNAAMMEE	euh je crois qu'elle voudrait rencontrer notre
ami			CEFC

Dans les cas comme 48 et 49, nous sommes parti de notre intuition la plus forte, à savoir que les constituants sont bel et bien des compléments de phrase, dans la mesure où ils fonctionnent comme un cadre pour une prédication que le locuteur s'apprête à introduire, mais qu'il finit par ne pas introduire pour cause d'aposiopèse ou de changement de la structure commencée. Évidemment, la frontière entre ce type de situation et celles où la position a été marquée comme inconnue (car impossible à déterminer) était souvent très ténue, la décision finale reposant en partie sur notre jugement subjectif.

En tchèque, où l'ordre des mots est beaucoup plus souple qu'en français, d'autres cas potentiellement litigieux se sont présentés, tel celui où le complexe nominal démonstratif s'intercale entre un auxiliaire et la forme verbale principale des temps et modes composés⁸⁵ :

50) když bud-e to výběr-o .. tak já bych
quand être.FUT-3SG ten.N.NOM.SG appel d'offres-NOM.SG alors moi.NOM AUX.COND.1SG

ti i ty podmínk-y a propozic-e da-l-a
toi.DAT même ten.F.ACC.PL condition-ACC.PL et proposition-ACC.PL donner-PTCP-SG.F

s'il y a l'appel d'offres... moi, je t'enverrais même les conditions et les propositions

ORTOFON v1

Ici, c'est la position de la forme verbale principale, soit le participe passé *dala*, que nous avons tenue pour décisive : l'expression démonstrative a donc été analysée comme se trouvant en position préverbale.

Quant au français, nous nous sommes intéressé à la présence des particules localisatrices *-ci* et *-là* accompagnant le démonstratif. En revanche, deux autres catégories ont été créées uniquement pour le tchèque, à savoir : 1) le cumul de mots fonctionnels (grammaticaux) et 2) la grammaticalisation (présumée) en cours du démonstratif TEN en article défini⁸⁶.

La batterie des catégories d'annotation appliquées au sous-corpus des pièces de V. Havel est plus réduite. Outre les type et sous-type d'emploi du démonstratif du texte source, nous nous intéressons à la transposition des démonstratifs tchèques en français. À cet égard, six situations peuvent se présenter : 1) le démonstratif est transposé par un démonstratif, 2) le démonstratif est transposé par un article défini, 3) il est transposé par un article indéfini, 4) il est transposé par un déterminant possessif, 5) il est transposé par un autre moyen que les quatre derniers (un déterminant quantificateur, un adverbe, le *de* d'extensité nulle, une absence de déterminant, un pronom personnel,

⁸⁵ Rappelons que BY est, en tchèque, l'auxiliaire du conditionnel.

⁸⁶ Pour trancher à ce propos, nous nous appuyons sur une série de critères qui seront présentés de manière détaillée au sein du Chapitre 6.

etc.) et, enfin, 6) il n'est pas transposé du tout. La dernière situation est due à une traduction française trop libre de l'extrait contenant l'occurrence tchèque. Deux autres catégories sont présentes : la grammaticalisation du démonstratif TEN et les cumuls de celui-ci avec d'autres mots fonctionnels.

Les Tableaux 2.7, 2.8 et 2.9 synthétisent les catégories d'annotation pour chacune de nos trois principales sources de données. Nous y distinguons : 1) la catégorie en tant que telle et 2) ses valeurs. Du point de vue de ces dernières, la situation peut se présenter soit comme binaire – par exemple, la grammaticalisation est présente ou absente – ou comme plus complexe. Les deux premières catégories (« type d'emploi » et « sous-type d'emploi ») doivent être remplies avec une valeur, contrairement au reste, où l'absence de valeur (« vide ») représente en soi une valeur possible :

CEFC (Orféo)				
TYPE D'EMPLOI	SOUS-TYPE D'EMPLOI	PARTICULE LOCALISATRICE	DISFLUENCES	POSITION PAR RAPPORT AU VERBE
Endophore (E)	anaphore fidèle (AF)	-ci	vide	préverbale (P)
	anaphore infidèle (AI)	-là	AMORCE	postverbale (POST)
	anaphore indirecte (AIND)	vide		complément de phrase (CP)
	anaphore résomptive (AR)			dislocation à droite (RD)
	anaphore temporelle (AT)			dislocation à gauche (LD)
	cataphore (C)			clivage (CLIVAGE)
Deixis (D)	deixis situationnelle (S)			inconnue (?)
	deixis temporelle (T)			
	deixis textuelle (TEXT)			
Exophore mémorielle (EM)				
Emploi « émotionnel » (EE)				
Démonstratif indéfini (DI)				
Ambigu (AMB)				

Tableau 2.7 : Catégories et valeurs d'annotation appliquées à l'échantillon de 1 000 occurrences du CEFC (Orféo)

ORTOFON v1 (CNT)

TYPE D'EMPLOI	SOUS-TYPE D'EMPLOI	GRAMMATICALISATION	DISFLUENCES	CUMUL	POSITION PAR RAPPORT AU VERBE
Endophore (E)	anaphore fidèle (AF)	vide	vide	vide	préverbale (P)
	anaphore infidèle (AI)	GRAM	AMORCE	CUM	postverbale (POST)
	anaphore indirecte (AIND)				complément de phrase (CP)
	anaphore résomptive (AR)				dislocation à droite (RD)
	anaphore temporelle (AT)				dislocation à gauche (LD)
	anaphore associative (AA)				inconnue (?)
	cataphore (C)				
Deixis (D)	deixis situationnelle (S)				
	deixis temporelle (T)				
	deixis textuelle (TEXT)				
Exophore mémorielle (EM)					
Emploi « émotionnel » (EE)					
Démonstratif indéfini (DI)					
Emloi « métalinguistique » (MET)					

Tableau 2.8 : Catégories et valeurs d'annotation appliquées à l'échantillon de 1 000 occurrences d'ORTOFON v1 (CNT)

Pièces de V. Havel (InterCorp v10)

TYPE D'EMPLOI	SOUS-TYPE D'EMPLOI	GRAMMATICALISATION	CUMUL	TRANSPOSITION
Endophore (E)	anaphore fidèle (AF)	vide	vide	démonstratif (D)
	anaphore infidèle (AI)	GRAM	CUM	article défini (AD)
	anaphore indirecte (AIND)			article indéfini (AI)
	anaphore résomptive (AR)			déterminant possessif (POS)
	anaphore temporelle (AT)			autre (A ; quantificateurs, adverbess...)
	anaphore associative (AA)			vide (aucune transposition ; dû à une traduction libre)
	cataphore (C)			
Deixis (D)	deixis situationnelle (S)			
	deixis temporelle (T)			
	deixis textuelle (TEXT)			
Exophore mémorielle (EM)				
Emploi « émotionnel » (EE)				
Démonstratif indéfini (DI)				

Tableau 2.9 : Catégories et valeurs d’annotation appliquées à l’échantillon de 241 occurrences des pièces de V. Havel (InterCorp v10 ; CNT)

2.2.4 L’annotation des données : échantillonnage et taille des échantillons

Une étape très importante était celle de la définition de la taille de l’échantillon à annoter pour chaque ensemble de données. Pour les corpus de langue orale informelle, la taille de cet échantillon a finalement été fixée à 1 000 occurrences. Pour le sous-corpus des pièces dans InterCorp, elle a été fixée par le nombre maximal d’occurrences trouvées : 241. Le choix de 1 000 occurrences pour les corpus de langue orale informelle se justifie par des raisons pragmatiques. D’une part, il a fallu que l’échantillon soit assez grand pour nous donner une image représentative des phénomènes, servant ainsi d’appui solide aux conclusions qualitatives et quantitatives que nous allions en tirer. D’autre part, il a fallu que sa taille nous permette de bien en prendre connaissance et de nous y repérer aisément.

Le nombre d’occurrences initialement obtenu pour l’oral informel des deux langues manifestait une forte disproportion : alors que la requête définie pour ORTOFON v1 a donné 24 096 occurrences, la recherche dans le CEFC n’en a généré que 2 015 ! À elle seule, cette différence abyssale témoigne d’un emploi beaucoup plus répandu du démonstratif adnominal en tchèque qu’en français pour le même type de productions langagières, même si – précision importante – ces deux chiffres ne représentaient que le premier jet, c’est-à-dire une sélection préalable au tri ayant pour but d’éliminer le bruit –, et que les deux ensembles de données n’ont pas exactement la même taille (voir *supra*).

Afin d’équilibrer ces échantillons, nous avons procédé de la façon suivante : pour le tchèque, nous avons tout d’abord retenu une liste de 2 409 occurrences, obtenue via le maintien, à partir de la liste originelle exportée dans Excel, d’une occurrence sur dix. Ce nombre, toujours bien plus élevé que le nombre final, se justifiait au départ notamment par la volonté de disposer d’une réserve suffisante eu égard à la présence de bruit. Dans un premier temps, l’annotation a été commencée au début du document Excel contenant cette liste de 2 409 occurrences et elle a été terminée dès que le nombre de 1 500 occurrences valides a été atteint. Cependant, étant donné les difficultés rencontrées lors de l’annotation (nombre élevé de cas ambigus, processus d’évaluation des occurrences très long et laborieux), nous avons finalement décidé de réduire le nombre d’occurrences analysées à 1 000.

C'est également l'échantillon de 2 409 occurrences qui nous a permis de calculer la proportion approximative de bruit au sein de la première liste générée via la requête. Ainsi, nous avons d'abord relevé toutes les occurrences bruitées dans l'échantillon de 2 409 occurrences. À l'aide de l'interface « 1 feature – many samples » du site du CNT (voir la Section 2.1.6 *supra*), nous avons calculé la moyenne du bruit obtenu pour 24 échantillons de 100 occurrences (la liste de 2 409 occurrences avait ainsi été répartie en 24 échantillons avec un résidu de 9 occurrences). C'est de cette façon que nous sommes parvenu au chiffre de 15,71 (soit 15,71 % d'occurrences indésirables dans l'ensemble de la liste initiale). Le module permet aussi de déterminer le degré de précision du résultat obtenu, et ce au moyen de la loi de Student ou de la loi normale. Nous avons opté pour la première, qui est réputée fournir un intervalle de confiance plus précis. Dans notre cas, l'interface affiche un intervalle de confiance pour l'ensemble de la population n (liste des concordances initiale) s'étendant entre 3 391,034 et 4 179,126 occurrences à rejeter et une moyenne égale à 3 785,08. Malheureusement, l'écart-type est plutôt élevé : 3,873 (α étant égale à 0,05). Malgré cela, nous nous sommes contenté de ce résultat, car le bruit manifestait une grande variabilité d'un échantillon à l'autre (de 10 à 23 occurrences, avec une médiane de 16), si bien que l'ajout d'autres échantillons n'auraient guère permis d'arriver à réduire l'écart-type. Connaissant la fréquence absolue moyenne du bruit dans l'ensemble de la population n , nous avons pu calculer les fréquences absolue et relative approximatives du démonstratif en emploi adnominal (voir la Section 5.2.1 du chapitre 5 pour l'ensemble du procédé). La Figure 2.9 représente l'interface « 1 feature – many samples » avec les valeurs saisies et les résultats générés :

Source sample (population) size: 240096

Sample size: 100

Significance level (α): 0.0001 to 0.05

Results of sample analysis:

20	18	14	13	12	23	17	17	10
11	12	14	10	17	15	16	19	21
10	18	16	13	19	22			

Comma separated values (integers).

This module is intended to evaluate the reliability of results of manual analysis of a number of random samples (typically from an extensive concordance that is impossible to examine in detail). The usual problem of such analysis stems from our inability to determine the number and size of the individual samples necessary for a reliable estimate. There is no universal answer to this problem, because it depends on a number of factors (esp. on the frequency and dispersion of the feature), however, based on the *confidence intervals*, we can estimate in what range of frequencies the feature in question appears in the whole concordance. This module shows that by adding samples, the frequency

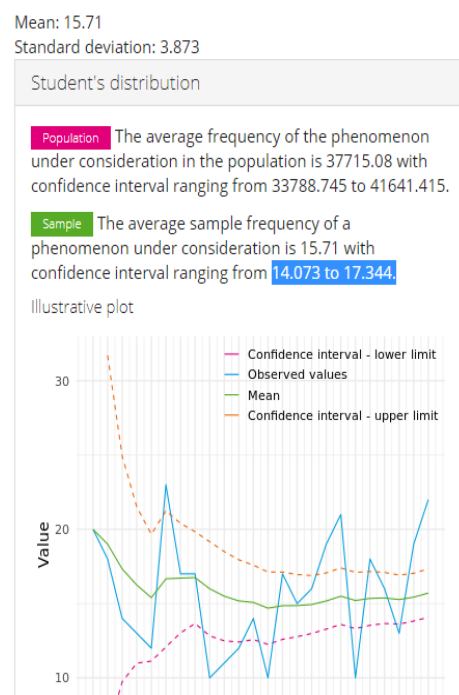


Figure 2.9 : Interface « 1 feature – many samples » (CNT) ; évaluation de la représentativité de 24 échantillons du démonstratif adnominal (choisis au hasard) eu égard à la proportion de bruit dans l'ensemble des occurrences (24 096) générées par la requête

Pour le français, le procédé était plus compliqué, et ce pour deux raisons. Dans un premier temps, nous n'avions pas inclus dans notre corpus les conversations téléphoniques, soit un sous-ensemble de 12 textes contenus dans C-ORAL-ROM dont la taille fait 11 664 positions. Ce n'est qu'après avoir pris connaissance du fait que le corpus ORTOFON v1 contient lui aussi des conversations téléphoniques et par Skype que nous avons corrigé cette erreur. Ainsi, nous avons d'abord travaillé avec la liste Excel obtenue à partir de la requête reproduite dans la Section 2.2.1. Celle-ci contenait 1 991 occurrences (bruit compris), contrairement à la liste finale, qui, après l'ajout des occurrences apparues dans les conversations téléphoniques, en contenait 2 015 (bruit compris). Le nombre d'occurrences fixé au départ pour l'annotation était, ici aussi, de 1 500. Ce chiffre a ensuite été changé à 1 000, afin qu'il corresponde au chiffre retenu pour le tchèque.

Mais le procédé d'annotation n'a pas suivi tout à fait le même schéma, car nous visions un maximum de diversité des sources disponibles. Or dans notre liste initiale, la très grande majorité des occurrences – 1 514 (!) – était fournie par la base TUFS. Il nous a donc paru urgent de contrer par tous les moyens possibles cette hégémonie. C'est la raison pour laquelle nous avons, dans un premier temps, inclus dans notre échantillon toutes les occurrences provenant des autres sous-corpus pour, dans une seconde phase seulement, compléter cet ensemble avec les occurrences de la base TUFS, prises dans l'ordre de leur apparition dans la liste Excel. Malgré cette mesure, seules 465 des 1 000 occurrences proviennent des autres sous-corpus, le reste, c'est-à-dire 535, provenant de la base TUFS.

2.3 Résumé

Dans ce chapitre, nous avons décrit nos données et notre méthode de travail. Tout d'abord, nous avons présenté nos corpus de langue orale informelle CEFC et ORTOFON v1 et le sous-corpus des pièces de V. Havel, créé sur le portail du CNT. Outre la présentation des métadonnées, nous avons prêté une attention toute particulière à la comparaison de ces sources entre elles. Nous avons d'emblée insisté sur la nature à première vue assez différente des deux sources d'oral informel, résultant des projets à leur origine ainsi que d'une situation de départ fort divergente dans les deux milieux linguistiques. Comparé à ORTOFON v1, corpus fait sur mesure et s'inscrivant dans une tradition solide de linguistique de corpus, le projet du CEFC est clairement imparfait. Son principal défaut réside dans sa composition non équilibrée (la majorité des conversations, issues d'une seule source, pèchent par leur profil monolithique). En dépit de cela, nous ne pensons pas que ce point faible aille jusqu'à remettre en question notre projet de recherche. Il s'agit tout de même d'interactions informelles entre personnes qui se connaissent, et où des contraintes en tout genre ne se font que

rarement ressentir. De manière générale, il est très difficile, sinon impossible, de trouver des corpus parfaitement comparables. Mais n'ayant jamais eu pour objectif de confronter les deux langues du point de vue des différents paramètres relatifs aux locuteurs, nous restons persuadé que pour les objectifs de notre comparaison, cette ressemblance approximative entre les deux corpus suffit. Pour ce qui concerne notre sous-corpus défini depuis la partie tchéco-française d'InterCorp v10, là aussi, des faiblesses peuvent être relevées, à savoir, notamment, le sens unidirectionnel des analyses, la datation des textes et la présence d'éléments de langage formel. Comme l'écrit Nádvorníková (2017), il est fortement conseillé, lorsque l'on se sert de corpus parallèles (multilingues), de travailler dans les deux directions afin de minimiser l'impact des interférences. Cependant, notre recours à cet outil joue un rôle exclusivement complémentaire par rapport au noyau de notre projet qui est le travail avec les corpus de langue orale informelle. Compte-tenu de cela, nous ne pensons pas que cette unidirectionnalité s'apparente à un handicap grave.

Concernant les catégories d'annotation, elles ont été définies notamment pour rendre compte des différents emplois du démonstratif adnominal dans les deux langues afin que ceux-ci puissent être analysés et comparés. Une bonne partie de ces catégories restent les mêmes pour les trois sources de données. Le reste diffère en fonction des deux langues et des deux types de productions langagières représentés. Une série de procédés ont été mis en place afin de définir, pour chaque corpus, un échantillon aussi équilibré et représentatif que possible.

Les deux chapitres suivants seront consacrés à un état de l'art concernant la situation relative aux démonstratifs dans les systèmes linguistiques tchèque (Chapitre 3) et français (Chapitre 4).

3 Le démonstratif en tchèque

Dans ce chapitre, nous présenterons le paradigme des démonstratifs tchèques. Cette présentation se fera en plusieurs temps. D’abord, nous traiterons du démonstratif tchèque en synchronie : après avoir donné un aperçu de l’ensemble du paradigme avec la diversité de ses formes ainsi que quelques chiffres indicatifs quant à la fréquence de ces formes dans différents types de textes, nous nous interrogerons sur le modèle d’analyse syntaxique qui sied le mieux à la situation des démonstratifs dans cette langue. Nous aborderons également la question de la polysémie/homonymie en rapport avec le système du démonstratif nominal (Dixon, 2003) tchèque. Dans un deuxième temps, nous regarderons brièvement comment le système actuel s’est constitué en diachronie. Ensuite, nous nous pencherons sur la distribution des différentes formes au sein des différents régimes de référence indexicale, présentés dans le Chapitre 1 (Section 1.1.2) : l’exophore a-mémorielle, l’endophore et l’exophore mémorielle. Il s’ensuivra une section consacrée au potentiel de nominalisation de TEN lorsque celui-ci introduit un adjectif sans nom. Enfin, une introduction aux emplois « émotionnels » clôturera le chapitre.

Une série de critères et de plans seront adoptés tout au long de cette présentation : la morphologie, la phonétique (dans une moindre mesure), la syntaxe, les oppositions diastratiques et diatopiques, mais aussi le critère de la distance du référent par rapport aux participants à l’interlocution et, bien évidemment, celui du régime d’indexicalité. Plusieurs sources de référence – notamment du domaine des études tchèques – seront mobilisées. Nous rappelons que, conformément à l’orientation de ce travail, il ne sera question que des démonstratifs dits « nominaux » dans la typologie de Dixon (2003 ; voir la Section 1.3.3) et presque exclusivement des emplois adnominaux de ceux-ci. Nous veillerons à illustrer les phénomènes traités à l’aide d’exemples, dont une bonne partie sera tirée de nos corpus pour le tchèque, notamment d’ORTOFON v1.

3.1 Le démonstratif tchèque en synchronie

Dans le tchèque parlé de nos jours – même à condition de laisser de côté les formes du démonstratif indéfini TAKOVÝ et ses dérivés, exprimant une propriété, ainsi que les formes qui expriment l’identité (TÝŽ, TENTÝŽ) –, nous nous retrouvons face à un paradigme assez riche et diversifié de formes définies du démonstratif. Leur classement peut et doit être réalisé selon plusieurs axes, dont notamment l’axe morphologique (ou morphophonologique), l’axe de la proximité/distance du référent, les axes des distinctions diastratiques et diatopiques, ainsi que l’axe d’usage, à savoir

le(s) régime(s) d'indexicalité au(x)quel(s) les formes s'associent. Le facteur de la fréquence des formes dans le tchèque contemporain doit lui aussi être pris en compte (voir les Sections 3.1.1 et 3.1.2 *infra* ainsi que la Section 5.2.3 du Chapitre 5). Dans ce qui suit, nous présenterons d'abord le paradigme dans ses grands traits. Ensuite, nous nous focaliserons sur les distinctions existant entre ses membres tout en nous référant à des sources diverses et tout en attirant l'attention du lecteur sur une série de problèmes posés par la situation en tchèque. Ces problèmes seront liés, entre autres, à la distinction entre emplois adnominaux et pronominaux, ainsi qu'au modèle d'analyse syntaxique à choisir parmi les modèles et les approches existants afin de rendre compte du fonctionnement du démonstratif nominal. Enfin, la question – plus complexe qu'il n'y paraît – de la polysémie/homonymie des formes sera abordée. Quant à la distribution des formes dans les différents régimes d'indexicalité, elle sera développée dans la Section 3.3. Deux autres thèmes seront abordés à la fin du présent chapitre : le rôle nominalisateur du démonstratif TEN (Section 3.4) et les emplois que nous appelons les emplois « émotionnels » (Section 3.5).

3.1.1 Le système du démonstratif en tchèque : une diversité (à première vue) vertigineuse

En amont, il semble légitime d'opérer une première distinction entre une série de formes se trouvant au centre du système et celles qui occupent davantage sa périphérie. Il est crucial de préciser que le terme de « périphérie » – certes quelque peu controversé – reflète ici un point de vue essentiellement morphologique et diachronique et non pas celui de la fréquence d'utilisation. Ainsi, les formes qui seront dites « centrales » sont morphologiquement plus simples et plus anciennes que les formes dites « périphériques », qui, elles, ont été créées plus tard, et ce en grande partie par reduplication à partir des formes « centrales », d'où leur plus grande complexité morphologique.

Si l'on consulte plusieurs sources de référence pour la langue tchèque¹ dont les dates de publication s'étendent entre les années cinquante du siècle dernier et nos jours (2020), l'image que l'on acquiert est à peu près la suivante :

les formes « centrales » : TEN, TENTO, TENHLE (avec ses variantes relâchées *tendle* et *tenle*), TAMTEN, éventuellement aussi ONEN

les formes plus ou moins « périphériques » : TENHLETEN (avec ses variantes *tendleten*, *tenleten*, *tenhlecten*, *tendlecten*, *tenlecten*, *tenhlencten*, *tendlencten* et *tenlencten*, TADYTEN (avec sa

¹ Au sein de cet ensemble, il est possible d'opérer une distinction entre les sources à vocation explicitement normative et le reste. C'est notamment le cas des *slovníky* (dictionnaires), se fixant ouvertement l'objectif de servir de référence pour le bon usage. Faute d'espace, nous ne développerons pas ici cette distinction et renvoyons, à ce sujet, le lecteur à la thèse d'habilitation de Berger (1993).

variante *tadycten*), TADYHLETEN (avec ses variantes *tadydleten*, *tadyhlecten*, *tadydlecten* et *tajdleten*), TAMHLETEN (avec ses variantes *tamdleten* et *tamleten*)², TUHLETEN (avec ses variantes *tudleten*, *tuhlecten* et *tudlencten*), TUTEN, TOŤTEN (avec sa variante *tojten*), HENTEN³.

Toutes ces formes peuvent être employées à la fois en emplois adnominaux et pronominaux, même si, pour certaines d'entre elles (notamment TEN), des analyses en termes de polysémie, voire d'homonymie, paraissent s'imposer pour une partie de ces emplois. Nous verrons cela plus loin (Section 3.1.4).

Comme nous pouvons le constater, la classe des formes « périphériques » constitue un ensemble très riche. Néanmoins, il faut insister sur le fait que bon nombre de ces formes sont fortement apparentées à d'autres, dont elles sont issues par des mutations phonétiques. Celles-ci concernent notamment l'amuïssement de la consonne /h/ ou bien son remplacement par la consonne /d/, éventuellement aussi l'apparition de la/des consonne(s) d'appui /(n)c/ :

tenhleten → *tendleten/tenleten/tenhle(n)cten*
tamhleten → */tamdleten/tamleten/tamhle(n)cten*
 ...
 mais aussi : *tenhle* → *tendle/tenle*.

Au vu de ces faits, il est légitime de se poser la question de savoir s'il faut vraiment accorder à ces formes le statut de lexème (lemme) indépendant, une option à laquelle Berger s'oppose :

Es erscheint mir relativ unsinnig, hier von eigenen Lexemen *tendlencten*, *támdleten* o.ä. auszugehen (soweit mir bekannt, wird eine solche Lösung auch nirgends vorgeschlagen), sondern es handelt sich offenbar um mehr oder weniger fakultative V a r i a n t e n der schriftsprachlichen Pronomina [...]. (1993 : 9)

Dans ORTOFON v1 et ORAL v1, les formes où /h/ est remplacé par /d/ ou disparaît complètement (y compris pour *tenhle/tendle/tenle*), ainsi que celles où apparaît /(n)c/, sont subsumées sous le même lemme. Toutefois, ce n'est pas le cas des corpus écrits du CNT.

Les formes « périphériques » sont fortement marquées à la fois sur les plans diastratique et diatopique. Sur le plan diastratique, elles appartiennent toutes à la langue familière, non-soutenue. Pour ce qui est des oppositions diatopiques, nous nous référons ici notamment à l'article de Jodas « Regionální difference v užívání ukazovacích zájmen » (« Différences régionales dans l'usage des pronoms indéfinis », 2010). Dans cet article issu d'une étude de terrain, l'auteur relève l'existence d'une dichotomie importante entre, d'une part, la Bohême et, d'autre part, la Moravie et la Silésie

² Dans toutes les variantes de TAMHLETEN, la voyelle /a/ peut être allongée (cf. *támhleten*).

³ Comme annoncé dans l'Introduction générale, seules les formes constituant des lemmes (dont la plupart, sauf les formes dialectales TOŤTEN et HENTEN, apparaissent dans le corpus ORTOFON v1) sont indiquées en majuscules.

(voir la Figure 3.1 *infra* pour la représentation de ces trois régions historiques de l'actuelle République tchèque et la Figure 3.3 *infra* pour la représentation de la situation dialectale du pays). Dans cette optique, toutes les formes sont avant tout employées en Bohême, à l'exception de TOŤTEN, *tojten* (une variante de TOŤTEN) et HENTEN, qui sont limitées aux territoires morave et silésien. La forme TUTEN et, dans une moindre mesure aussi TUHLETEN, sont, quant à elles, avant tout associées au territoire de la Bohême de l'Ouest (Pilsen, Karlovy Vary). Les formes dites « rédupliquées » (pour la définition, voire *infra*) comme TENHLETEN et TAMHLETEN et a fortiori leurs variantes *tendleten*, *tendle(n)cten*, *tamdleten*, *tamdle(n)cten*, etc. seraient ressenties comme étrangères par les locuteurs moraves et silésiens⁴. Il serait ainsi question de formes relevant, avant tout, de la variante de la langue connue sous le nom de « tchèque commun » („obecná čeština“) – un inter-dialecte utilisé au quotidien notamment par les habitants de la Bohême – et employées avant tout sur le territoire de la Bohême centrale, d'où elles se propageraient progressivement dans d'autres régions de la Bohême et, accessoirement, dans celles de la Moravie-Silésie (voir la Figure 3.2 *infra* pour le découpage régional de l'actuelle République tchèque). Krčmová (2017) écrit dans le *Nouveau dictionnaire encyclopédique du tchèque* (*Nový encyklopedický slovník češtiny*) que même s'il s'agit à l'origine d'une variété marquée sur le plan régional, l'acception majoritaire que reçoit aujourd'hui le terme de « tchèque commun » parmi les bohémistes est le suivant :

La plupart du temps, le tchèque commun désigne pourtant la langue de la communication parlée spontanée et privée, éventuellement semi-publique, peu marquée à la fois sur les plans régional et social et qui est en passe de quitter le statut d'une variété non standard du tchèque parmi d'autres pour celui de son registre substandard général. Les auteurs travaillant avec le terme de tchèque commun sont certes conscients des différences régionales qui touchent la langue au sein de la Bohême même, où les spécificités vont croissant dans les marges du territoire, mais ils se focalisent uniquement sur les traits communs. Ils trouvent le point de départ dans la langue non soutenue la plus nivelée de la Bohême centrale, qui constitue également la base du parler de Prague, premier centre culturel de notre territoire et qui, de ce fait, influe le plus sur les autres variétés du tchèque⁵.

Štícha (2013 : 26) écrit :

Le terme de « tchèque commun », assez solidement établi en linguistique, désigne une variété du tchèque non soutenu, utilisée de manière commune (d'où l'appellation de « commun ») sur l'intégralité du territoire de la

⁴ En effet, nous avons pu constater grâce à notre Questionnaire n°1 que si les variantes avec /d/ et/ou /(n)c/ sont assez naturellement employées dans une conversation avec des amis par tous nos informateurs originaires de Bohême, elles ne le sont pas par ceux originaires de Moravie, même lorsque ces derniers sont depuis longtemps installés en Bohême (cela vaut aussi pour notre informateur originaire de Rosice u Brna, qui vit depuis 24 ans à Prague). En revanche, tous acceptent plus ou moins les formes *tenhleten*, *tamhleten* et *tadyten*.

⁵ Nous traduisons : „Většinou se však chápe o.č. jako jazyk spontánní mluvené soukromé, event. poloveřejné komunikace, regionálně i sociálně málo příznakový, který se z pozice jedné z nestandardních forem č. dostává do postavení jejího substandardu. Autoři, kteří s termínem o.č. pracují, jsou si sice vědomi regionálních rozdílů v nespisovné mluvě ve vlastních Čechách, kde k okrajům území specifika přibývá, zaměřují se však jen na prvky společné; východiskem je pro ně nejvíce nivelizovaný nespisovný jazyk středních Čech, který je základem i mluvy Prahy, přirozeného kulturního centra našeho území, a který také díky tomu nejvíce působí na jiné variety č.“

Bohême et en partie aussi sur celui de la Moravie. Il s'agit, avant tout, de phénomènes morphologiques et phonologiques [...]⁶

Il faut cependant préciser que les formes non modifiées TENHLETEN, TAMHLETEN et TADYHLETEN peuvent également être considérées comme appartenant à la variété parlée du tchèque soutenu („hovorová čeština“)⁷, notamment sur le territoire de la Bohême. Le terme de « hovorová čeština », souvent rejeté de nos jours faute d'une définition claire et unifiée (cf. Nebeská, 2017 ; Short, 1993), désigne globalement le tchèque soutenu dans son usage oral courant, épuré tant d'éléments trop familiers que d'éléments archaisants et livresques⁸.

Une comparaison des fréquences relatives des formes « périphériques » dans ORTOFON v1 eu égard à l'opposition Bohême/Moravie-Silésie nous a apporté les constats suivants : selon les données du corpus, TENHLETEN est beaucoup plus utilisé en Bohême (898,7⁹ contre 186,9 pour la Moravie-Silésie¹⁰). Bien que TADYTEN soit davantage employé en Bohême (202,7), sa fréquence relative en Moravie-Silésie représente pour nous un constat plutôt inattendu : 142,8¹¹. TAMHLETEN jouit d'une fréquence relative de 24,3 en Bohême contre seulement 2,1 en Moravie-Silésie¹². TADYHLETEN atteint 12,2 en Bohême contre seulement 4,2 en Moravie-Silésie¹³. TUHLETEN et TUTEN atteignent, respectivement, 32,4 et 63,5 en Bohême contre 0 et 6,3 en Moravie-Silésie¹⁴. Parmi les locuteurs originaires de Bohême de l'Ouest, sa fréquence relative atteint 333,3 pour TUTEN et 175 pour TUHLETEN – c'est par des locuteurs de cette région que l'écrasante majorité des occurrences de ces

⁶ Nous traduisons : „Termín ‚obecná čeština‘ je v lingvistice dosti pevně zavedený termín, jímž se rozumí čeština nespisovná, která se užívá obecně (odtud název ‚obecná‘) po celém území Čech a zčásti i na území Moravy. Jde především o jevy morfologické a hláskové [...]” Pour une définition semblable de ce terme, voir aussi Vey (1946). Pour la question de savoir dans quelle mesure le « tchèque commun » pourrait être considéré comme étant à l'origine d'une situation de diglossie en Bohême, voir Čermák (2014).

⁷ Ce choix de classement est déjà opéré par exemple par *Slovník spisovného jazyka českého* (1966).

⁸ C'est Short qui, à notre sens, résume avec justesse les tensions entre le « tchèque soutenu » („spisovná čeština“), la variété parlée du tchèque soutenu („hovorová čeština“) et le tchèque commun („obecná čeština“) : “Standard Czech (*spisovná čeština*) is then a semi-artificial creation, archaic in many respects, while the vernacular has continued to evolve since the norms (whether of the sixteenth or the nineteenth century) were set. There is a consequent tension between the modern literary language and the spoken Czech, usually known as Common Czech (*obecná čeština*), in which natural development has culminated. Between these two poles there are transitional strata, notably Colloquial Czech (*hovorová čeština*), an informal spoken version of the standard language, whose existence is often denied) and Commonly Spoken Czech (*běžně mluvená čeština*), basically the everyday speech of the big cities.” (1993 : 455-456)

⁹ Toutes les fréquences relatives (ipm) dans ce chapitre sont calculées relativement à un million de positions (= nombre d'occurrences pour un million de positions), et les chiffres sont arrondis au dixième.

¹⁰ La différence est statistiquement significative, car le score de la fonction de vraisemblance (L) est égal à 280,4558 lorsque nous appliquons celle-ci aux fréquences absolues que la forme atteint dans les deux corpus et que nous tenons compte de leurs tailles respectives ; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$). Dans le cadre de cette comparaison ainsi que celles qui suivent, les deux types d'emploi (adnominal et pronominal) ont été recherchés.

¹¹ $L = 5,977 \rightarrow$ la différence est statistiquement significative (la valeur critique = 3,841).

¹² $L = 11,9267 \rightarrow$ la différence est statistiquement significative (la valeur critique = 3,841).

¹³ $L = 2,2627 \rightarrow$ la différence n'est pas statistiquement significative (la valeur critique = 3,841).

¹⁴ $L = 23,849 \rightarrow$ la différence est statistiquement significative (la valeur critique = 3,841).

deux formes représentées dans le corpus est prononcées¹⁵. Les formes TOŤTEN et HENTEN sont complètement absentes du corpus¹⁶.

Les formes « périphériques » ont toutes un penchant pour la deixis situationnelle et la plupart d'entre elles sont spécialisées dans l'expression de la proximité du référent, sauf *tamhleten*, *tamdleten*, *tamleten*, qui marquent la distance. Concernant la morphologie, toutes les formes « périphériques » ainsi qu'une partie des formes « centrales » sont des formes dites « renforcées », voire « rédupliquées » (Jodas, 2010). Le renforcement se fait par la suffixation de la forme neutre de TEN – *to* – devenue invariable (*ten-to*), de la particule *hle* (*ten-hle*), éventuellement aussi par la préfixation des adverbes démonstratifs locatifs *tady* (*ici* → *tady-ten*)¹⁷, *tam* (*là-bas* → *tam-ten*) ou *tu* (*ici/là* → *tu-ten*). D'autres morphèmes – *toť*, *toj* et *hen* (pour les formes dialectales moravo-silésiennes) – peuvent éventuellement être préfixés à TEN. Dans les cas de reduplication, l'affixation est double : la forme du démonstratif TEN se suffixe aux formes déjà suffixées par *hle* : *ten-hle-ten*, *tady-hle-ten*, *tam-hle-ten*, *tu-hle-ten*¹⁸. Contrairement aux autres formes affixées¹⁹, TEN maintient, au sein de ces formes renforcées, sa déclinaison (voir *infra*)²⁰. Il est intéressant de noter que l'origine de la particule présentative *hle* remonterait à la forme de l'impératif *hled'* du verbe *hleděti* (*regarder* ; Bělič, 1970 ; Vaillant 1958). Le fonctionnement de *hle* est très proche de celui du terme latin *ecce*, lui-même productif dans la formation du démonstratif français (voir la Section 4.2 du Chapitre 4). Tout cela permet ainsi de discerner un certain parallèle entre les évolutions des paradigmes démonstratifs des deux systèmes.

¹⁵ Ayant comparé les fréquences absolues pour ces deux formes dans la partie d'ORTOFON v1 comprenant uniquement des locuteurs originaires de Bohême de l'Ouest et dans celle qui exclut ces locuteurs, la différence entre elles est statistiquement significative : pour TUHLETEN, $L = 80,4988$; pour TUTEN, $L = 138,6185$.

¹⁶ Il faut, bien évidemment, relativiser ces résultats au vu de la taille limitée du matériel sur lequel ils ont été obtenus (ainsi, le sous-corpus composé de locuteurs originaires de Bohême de l'Ouest ne contient que 119 998 positions), mais aussi au vu du fait que seule importe la région où le locuteur a passé son enfance jusqu'à l'âge de quinze ans, sans tenir compte des influences que peut exercer sur son idiolecte son actuel lieu de résidence.

¹⁷ Il est toujours possible d'employer la graphie *tady ten*. Ainsi, *Slovník spisovné češtiny pro školu a veřejnost* (2005 [1994] : 441), une source normative importante et régulièrement rééditée (la dernière réédition date de 2012) cite cette graphie comme la seule possible. Cependant, celle-ci est source d'ambiguïtés dans les textes : s'agit-il – encore – d'un adverbe locatif suivi de TEN ou – déjà – d'une forme composée du démonstratif ? Bien que les annotateurs du CNT fassent la distinction entre les deux graphies, après un examen d'un nombre assez grand d'occurrences, il nous faut conclure que cette distinction formelle ne reflète pas bien la distinction sémantique : d'une part, il est des cas où les deux éléments sont séparés à l'écrit sans l'être à l'oral et vice versa.

¹⁸ Ici aussi, un certain flou enveloppe l'orthographe des formes, qui peuvent également se grafier *tadyhle ten*, etc., variante orthographique qui reste plus fidèle à l'étymologie. À ce titre, Berger (1993 : 6) se demande à quel point il est légitime de traiter TADYHLETEN et TAMHLETEN comme une seule forme : „[...] die Frage diskutiert wird, ob man wirklich von einer Wortform *tamhleten*, *tadyhleten* usw. ausgehen muß [...] oder ob es sich nicht vielmehr um lediglich orthographisch zusammengezogene Verbindungen der Adverbien *tamhle*, *tadyhle* usw. mit dem Pronomen *ten* handelt ”

¹⁹ Il serait peut-être plus juste de parler, comme le fait *Mluvnice češtiny 2* (1986 : 93 et 393), de formes « composées », vu que les éléments qui les constituent peuvent fonctionner comme des lexèmes indépendants.

²⁰ Masculin singulier : *tenhleten*, féminin singulier : *tahleta*, neutre singulier : *tohleto*, etc.



Figure 3.1 : Trois régions historiques de l'actuelle République tchèque - la Bohême (vert), la Moravie (bleu) et la Silésie (orange). Source : https://fr.wikipedia.org/wiki/Sil%C3%A9sie_tch%C3%A8que#/media/Fichier:CZ_Cechy_Morava_kraje.gif.



Figure 3.2 : Découpage régional de l'actuelle République tchèque. Source : <https://www.axl.cefan.ulaval.ca/europe/reptcheque-1General-demo.htm>.

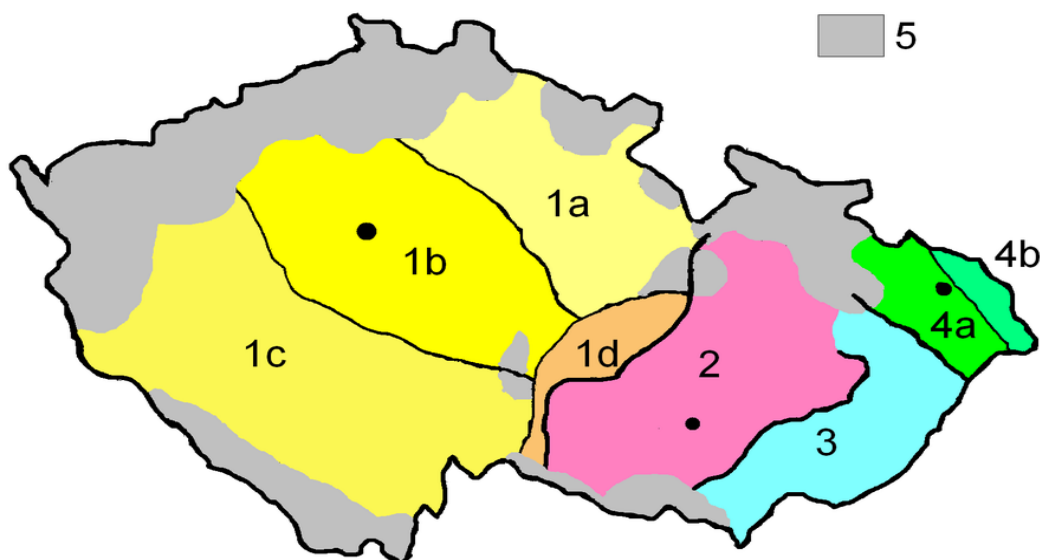


Figure 3.3 : Carte des dialectes de l'actuelle République tchèque.

1. groupe de Bohême : 1a – sous-groupe de Bohême du Nord et de l'Est, 1b – sous-groupe de Bohême centrale, 1c – sous-groupe de Bohême du Sud et de l'Ouest, 1d – sous-groupe de Bohême-Moravie
2. groupe de Moravie centrale
3. groupe de Moravie de l'Est
4. groupe silésien : 4a – sous-groupe moravo-silésien, 4b – sous-groupe siléso-polonais
5. zones de cohabitation de dialectes divers (anciens territoires germanophones)

Source :

https://cs.wikipedia.org/wiki/N%C3%A1%C5%99e%C4%8D%C3%AD_%C4%8De%C5%A1tiny#/media/Soubor:Czech_dialects.PNG ; notre traduction

Venons-en maintenant aux formes « centrales », à savoir TEN, TENTO, TENHLE, TAMTEN et ONEN. Outre la morphologie, il y a trois autres plans fondamentaux sur lesquels ces formes s'opposent : l'expression de la distance du référent par rapport aux personnes de l'interlocution, le niveau diastratique et le niveau diatopique²¹. Commençons par le critère de la distance. À cet égard, la forme TEN est strictement neutre dans le tchèque soutenu ainsi que dans le « tchèque commun », alors qu'elle peut marquer un référent plus proche chez certains locuteurs moraves de l'Est et silésiens (par opposition à TAMTEN ; Jodas, 2010 : 91 et Jodas, communication personnelle)²². TENTO et TENHLE marquent la proximité du référent, tandis que TAMTEN et ONEN marquent son éloignement. Compte-tenu de ces faits, le tchèque se présente clairement comme un système orienté par rapport à la distance

²¹ Un autre plan sur lequel des différences entre les formes apparaissent est celui du régime d'indexicalité. Ce plan sera développé dans la Section 3.3 *infra*.

²² La Moravie de l'Est inclut la majorité de la région de Zlín ainsi que la partie sud-est de la région de Moravie du Sud (voir la Figure 3.3). Ce constat est confirmé par notre informatrice originaire de Silésie. En tant que locuteur originaire de Bohême du Sud, nous n'étions pas au courant de cette réalité, au sujet de laquelle Jodas nous a confirmé dans une communication personnelle qu'elle s'étend aussi aux emplois adnominaux de TEN. Il convient par ailleurs de préciser que même lorsqu'il fonctionne de la sorte, TEN ne saurait, à notre avis, être considéré comme une forme de proximité, de la même manière que l'est TENTO et TENHLE : comme nous le verrons dans la Section 3.2 au sujet du polonais, nous sommes ici face à une configuration où TEN s'oppose, en tant que forme neutre, à TAMTEN, forme spécialisée dans l'expression de l'éloignement. Ainsi, tandis que l'opposition TENTO/TENHLE vs TAMTEN correspondrait à une conception de l'espace « proche vs lointain », l'opposition TEN vs TAMTEN correspondrait, elle, à une conception de l'espace « neutre vs lointain ».

(“distance-oriented” ; Anderson et Keenan, 1985) comprenant trois sphères : les sphères proche, neutre et lointaine²³. Il va sans dire que la notion même de distance est assez problématique (voir la Section 1.3.2 du Chapitre 1) et que si elle reste encore relativement facile à définir dans le régime de l’exophore situationnelle et temporelle, les choses deviennent plus délicates dans les domaines de la référence endophorique et exophorique mémorielle. Mais à ce stade de notre travail, nous nous contenterons de classer les formes sur la base d’une conception assez « naïve » et rudimentaire de ladite opposition.

Sur le plan morphologique, les démonstratifs nominaux tchèques se déclinent. Voici les paradigmes de déclinaison des formes TEN, TENTO et TENHLE en tchèque soutenu :

TEN		Singulier			Pluriel		
Genre		M	F	N	M	F	N
Cas							
Nominatif		<i>ten</i>	<i>ta</i>	<i>to</i>	<i>ti</i>	<i>ty</i>	<i>ta</i>
Génitif		<i>toho</i>	<i>té</i>	<i>toho</i>	<i>těch</i>	<i>těch</i>	<i>těch</i>
Datif		<i>tomu</i>	<i>té</i>	<i>tomu</i>	<i>těm</i>	<i>těm</i>	<i>těm</i>
Accusatif		<i>toho</i>	<i>tu</i>	<i>to</i>	<i>ty</i>	<i>ty</i>	<i>ta</i>
Locatif		<i>tom</i>	<i>té</i>	<i>tom</i>	<i>těch</i>	<i>těch</i>	<i>těch</i>
Instrumental		<i>tím</i>	<i>tou</i>	<i>tím</i>	<i>těmi</i>	<i>těmi</i>	<i>těmi</i>

Tableau 3.1 : Paradigme morphologique de TEN (tchèque soutenu)

TENTO		Singulier			Pluriel		
Genre		M	F	N	M	F	N
Cas							
Nominatif		<i>tento</i>	<i>tato</i>	<i>toto</i>	<i>tito</i>	<i>tyto</i>	<i>tato</i>
Génitif		<i>tohoto</i>	<i>této</i>	<i>tohoto</i>	<i>těchto</i>	<i>těchto</i>	<i>těchto</i>
Datif		<i>tomuto</i>	<i>této</i>	<i>tomuto</i>	<i>těmto</i>	<i>těmto</i>	<i>těmto</i>
Accusatif		<i>tohoto</i>	<i>tuto</i>	<i>toto</i>	<i>tyto</i>	<i>tyto</i>	<i>tato</i>
Locatif		<i>tomto</i>	<i>této</i>	<i>tomto</i>	<i>těchto</i>	<i>těchto</i>	<i>těchto</i>
Instrumental		<i>tímto</i>	<i>touto</i>	<i>tímto</i>	<i>těmito</i>	<i>těmito</i>	<i>těmito</i>

Tableau 3.2 : Paradigme morphologique de TENTO (tchèque soutenu)

²³ Cette série d’observations, figurant dans la plupart des travaux écrits, a été validée par nos dix informateurs natifs via le Questionnaire n°1.

TENHLE		Singulier			Pluriel		
Genre		M	F	N	M	F	N
Cas							
Nominatif		<i>tenhle</i>	<i>tahle</i>	<i>tohle</i>	<i>tihle</i>	<i>tyhle</i>	<i>tahle</i>
Génitif		<i>tohohle</i>	<i>téhle</i>	<i>tohohle</i>	<i>těchhle</i>	<i>těchhle</i>	<i>těchhle</i>
Datif		<i>tomuhle</i>	<i>téhle</i>	<i>tomuhle</i>	<i>těmhle</i>	<i>těmhle</i>	<i>těmhle</i>
Accusatif		<i>tohohle</i>	<i>tuhle</i>	<i>tohle</i>	<i>tyhle</i>	<i>tyhle</i>	<i>tahle</i>
Locatif		<i>tomhle</i>	<i>téhle</i>	<i>tomhle</i>	<i>těchhle</i>	<i>těchhle</i>	<i>těchhle</i>
Instrumental		<i>tímhle</i>	<i>touhle</i>	<i>tímhle</i>	<i>těmihle</i>	<i>těmihle</i>	<i>těmihle</i>

Tableau 3.3 : Paradigme morphologique de TENHLE (tchèque soutenu)

Pour TEN et TENHLE, un décalage considérable existe entre le paradigme du tchèque soutenu et celui du « tchèque commun » (les formes concernées sont marquées en gras). Pour TEN, il s'agit notamment du syncrétisme au nominatif pluriel :

TEN en « tchèque commun » ²⁴		Singulier			Pluriel		
Genre		M	F	N	M	F	N
Cas							
Nominatif		<i>ten</i>	<i>ta</i>	<i>to</i>	<i>ty</i>	<i>ty</i>	<i>ty</i>
Génitif		<i>toho</i>	<i>tý</i>	<i>toho</i>	<i>těch</i>	<i>těch</i>	<i>těch</i>
Datif		<i>tomu</i>	<i>tý</i>	<i>tomu</i>	<i>těm</i>	<i>těm</i>	<i>těm</i>
Accusatif		<i>toho</i>	<i>tu</i>	<i>to</i>	<i>ty</i>	<i>ty</i>	<i>ty</i>
Locatif		<i>tom</i>	<i>tý</i>	<i>tom</i>	<i>těch</i>	<i>těch</i>	<i>těch</i>
Instrumental		<i>tim</i>	<i>tou</i>	<i>tim</i>	<i>těma</i>	<i>těma</i>	<i>těma</i>

Tableau 3.4 : Paradigme morphologique de TEN (« tchèque commun »)

Sur les plans diastratiques et diatopiques, les choses se présentent de la manière suivante : la forme TEN est neutre, c'est-à-dire employée sur l'intégralité du territoire et dans tous les registres. Cependant, il est à noter qu'elle connaîtrait un usage plus extensif chez certains locuteurs de Moravie

²⁴ Les formes qui diffèrent par rapport au tchèque soutenu sont marquées en gras. Pour les quelques différences de nature régionale qui, de nos jours, subsistent dans la déclinaison de TEN en Bohême, nous renvoyons le lecteur à l'article de Jodas (2010).

(notamment de l'Est) et de Silésie, vu qu'elle peut aussi y marquer la proximité du référent (voir *supra*)²⁵.

La forme TENTO n'a pas le même statut en Bohême et en Moravie-Silésie. Sur ce dernier territoire, TENTO constituerait – parfois avec TEN et, dans une plus ou moins grande mesure, les formes régionales – la forme neutre d'expression de la sémantique de proximité. En Bohême, TENTO serait actuellement perçu comme « exclusivement soutenu, voire livresque » („vyhraněně spisovné až knižní“ ; Jodas, 2010 : 91). Ce constat n'est pas sans surprendre dans la mesure où jusqu'à à peu près les années quatre-vingts du siècle dernier, TENTO était globalement décrit par les grammaires de référence comme la forme neutre du tchèque soutenu. La grammaire de Kořenský et Komárek (*Mluvnice češtiny 2*, 1986 : 93) est la première source à caractériser TENTO comme « neutre, voire exclusivement soutenu » („neutrální až vyhraněně spisovné“). La grammaire de Štícha (2013 : 398), se basant en grande partie sur l'usage de la langue documenté par le Corpus national tchèque (CNT), la décrit comme « formelle ». D'autre part, TENHLE est considéré par lui comme la forme « fondamentale » (2013 : 400), même s'il faut préciser que Štícha ne mentionne TENHLE que dans la partie consacrée à la deixis situationnelle²⁶. Jodas (2010 : 90-91) est du même avis, voyant dans TENHLE la forme de proximité « neutre », mais seulement à condition que le jugement se limite à la Bohême²⁷. En Moravie, TENHLE serait toujours ressenti comme étranger par une partie importante des locuteurs²⁸. Dans des sources plus datées telles que *Mluvnice češtiny 2* (1986) et *Slovník spisovné češtiny pro školu a veřejnost* (la première édition date de 1994), TENHLE est encore explicitement associé à la variété parlée du tchèque soutenu („hovorová čeština“). Selon la première source (1986 : 93), TENHLE serait, en plus, doté d'une sémantique déictique plus forte que TENTO. Pour la

²⁵ Dans notre corpus, nous avons détecté une seule occurrence adnominale de TEN de ce type. TEN y est employé en deixis situationnelle pour référer à un référent proche. Autre fait très singulier : la locutrice (originnaire de la région de Moravie du Sud) accentue la forme, alors que le paradigme de TEN en positions adnominales n'est que très rarement accentué en Bohême (voir *infra*).

²⁶ En effet, son plan opère une stricte dissociation des emplois endophoriques („odkazovací“) et exophoriques situationnels („ukazovací“). Si nous accueillons positivement ce choix onomasiologique, nous éprouvons des réticences face à la manière quelque peu violente et arbitraire dont il associe les différentes formes à ces deux types d'emplois. Berger (1993 : 425-428) considère lui aussi qu'au sein de la langue écrite, TENHLE figurerait avant tout en emploi situationnel et que son emploi en anaphore connoterait toujours la langue parlée informelle.

²⁷ En revanche, les variantes avec /d/ (*tendle*) sont décrites par lui comme « marquées à la fois sémantiquement et stylistiquement, éventuellement aussi connotées comme expressives » („významově i stylově příznakové, eventuálně i expresivně zabarvené“). Selon Jodas (2010 : 91), s'appuyant sur des résultats d'enquêtes par questionnaires à l'écrit et à l'oral, TENHLE serait également en train de gagner du terrain (notamment chez les jeunes locuteurs) en Moravie de l'Ouest, territoire incluant la partie est de la région de Vysočina, la partie occidentale de la Région de Moravie du Sud ainsi que la Moravie centrale (la Région d'Olomouc) ; voir la Figure 3.3. Cette tendance est confortée par nos deux informateurs venant de Moravie de l'Ouest. Pour le premier, TENHLE constitue le premier choix pour renvoyer à un référent proche. Pour la seconde, c'est TENTO qui constitue le premier choix, mais TENHLE est jugé comme tout à fait acceptable. Il faut cependant rappeler que tous les deux vivent à Prague depuis 24 et 9 ans respectivement. Nos deux informatrices venant de Moravie de l'Est et de Silésie – elles aussi installées à Prague depuis sept ans – considèrent également TENHLE comme acceptable, même s'il ne représente pas leur premier choix pour renvoyer à un référent proche.

²⁸ En effet, pour nos deux informatrices de Moravie de l'Est et de Silésie, c'est TENTO (pour la première) et TEN (pour la seconde) qui représentent le premier choix pour renvoyer à un référent proche en deixis situationnelle.

seconde source (2009 [1994] : 447), TENHLE ne constitue pas une forme (un lemme) à part, mais juste une variante de TENTO. Telle est aussi la position de *Slovník spisovného jazyka českého* (1966 : 793), considérant TENHLE comme une variante de TENTO propre à la variété parlée du tchèque soutenu et au « tchèque commun »²⁹. Sans mentionner ici d'autres sources s'exprimant au sujet de TENHLE et de l'opposition entre TENHLE et TENTO, il est possible de conclure par un double constat. D'abord, il semble clair que le statut de TENHLE a évolué au cours de la période s'étendant entre les années soixante du siècle dernier et l'époque actuelle en allant progressivement vers une « standardisation ». Ainsi, il serait aujourd'hui possible de le considérer, à la limite, comme une forme neutre du tchèque soutenu, avec cependant une propension nette pour la variété parlée, moins formelle, de ce dernier. Mais il faut aussitôt répéter que ce statut touche avant tout la Bohême.

Dans le corpus de référence pour le tchèque écrit synchronique SYN2015, le lemme TENTO atteint une fréquence relative de 2 472,1 et TENHLE, une fréquence relative de 636,3³⁰. Si ce bilan prouve que TENTO demeure la forme de base du tchèque écrit, il était également l'importance croissante de TENHLE au sein de ce registre. Cependant, il est également intéressant de regarder la situation au sein des trois types de textes dont le corpus est composé. Dans le cadre des textes de fiction, TENHLE atteint une fréquence relative de 1 270,2 contre seulement 860 pour TENTO, mais seulement 191,1 dans le cadre des textes de littérature spécialisée, contre 4 559,9 (sic) pour TENTO. Enfin, dans les textes de presse, c'est également TENTO qui prime, avec une fréquence relative de 2 091,3, alors que celle de TENHLE n'y atteint que 413,8. Il semble du moins possible de conclure à partir de cette comparaison succincte que c'est avant tout dans les textes de fiction que TENHLE gagne du terrain, chose qui est tout à fait compréhensible vu les tentatives, propres à ce groupe de types de textes, de représenter et d'imiter la langue spontanée³¹. À titre de comparaison, nous avons également regardé la fréquence relative de ces deux démonstratifs dans le corpus Diakorp v6, un corpus diachronique de référence contenant des textes écrits ou parlés formels (prières, discours solennels) de la période 1375-1939³². Ce corpus n'étant pas lemmatisé, seules ont été recherchées les formes du nominatif singulier masculin. La différence est tout de même frappante : 670,2 pour *tento* et 6,05 pour *tenhle*, dont la plus ancienne occurrence apparaît seulement dans un texte de 1892.

²⁹ Celui qui va encore plus loin dans la fusion des formes dans un seul lemme, c'est Trávníček (1951 : 1130), pour qui les deux formes TENTO et TENHLE, mais aussi TENHLETEN, TAMTEN, TAMHLETEN, TUTEN et TOTEN appartiendraient toutes au lemme TEN.

³⁰ Rappelons que toutes les fréquences relatives (ipm) ici mentionnées expriment le nombre d'occurrences relativement à un million de positions de la population. Dans SYN2015, les formes affichant un /d/ (*tendle*), un /(n)c/ (*tendle(n)cten*) ou une syncope de la consonne (*tenle*) sont considérées comme des lemmes à part. Contrairement à ce choix, nous suivons ici le choix fait pour les corpus de langue parlée ORAL v1 et ORTOFON v1, où ces formes sont subsumées sous le même lemme que la forme de base (*tenhle*, *tenhleten*).

³¹ Dans la Section 3.1.2 *infra*, nous recourrons à la fonction de vraisemblance pour étayer la significativité statistique de ces différences.

³² Rappelons que la composition n'est pas équilibrée, 48 % des textes datant du 19^{ème} siècle (voir la Section 2.1.4).

Pour ce qui est de la situation diatopique au sein du tchèque parlé informel contemporain, une brève exploration du corpus ORTOFON v1 nous a montré que TENTO possède une fréquence relative de 194,6 auprès des locuteurs originaires de Bohême contre 735 auprès des locuteurs originaires de Moravie ou de Silésie³³. Quant à TENHLE, sa fréquence relative atteint 1 313,6 auprès des locuteurs de Bohême et 550 auprès des locuteurs moraves et silésiens³⁴. Pour l'ensemble d'ORTOFON v1, les fréquences relatives pour TENTO et TENHLE sont de 399,5 et 999,6 respectivement³⁵.

Venons-en maintenant aux deux autres membres du système « central », à savoir TAMTEN et ONEN. Les deux formes marquent l'éloignement du référent. Clairement, ONEN possède aujourd'hui des connotations livresques :

Le pronom *onen* est livresque et sert d'habitude de moyen de stylisation marqué, prêtant souvent à la phrase un caractère ironique ou pathétique. (Štícha, 2013 : 398)³⁶

Gebauer écrivait déjà dans son traité de grammaire historique (*Historická mluvnice jazyka českého, IV, Skladba* ; publié à titre posthume en 1929 mais rédigé au tournant du 19^{ème} et du 20^{ème} siècle) que « le pronom *onen* est en train de sortir de l'usage commun, étant remplacé par *tam-ten, ten-tam* » (1929 : 243)³⁷. Dans le tchèque parlé informel contemporain, cette forme serait représentée de manière très parcimonieuse : le corpus ORTOFON v1 n'en contient aucune occurrence, et le corpus ORAL v1, cinq fois plus grand, n'en contient que sept, dont quatre font partie du nom propre d'un restaurant³⁸. L'on pourrait ainsi être tenté de conclure que TAMTEN serait aujourd'hui devenu la forme neutre pour marquer l'éloignement, et ce notamment dans le régime exophorique situationnel (ses fréquences relatives pour ORTOFON v1 et ORAL v1 étant respectivement de 84,9 et de 129,2)³⁹.

³³ La différence est statistiquement significative : $L = 203,2721$ pour la comparaison des fréquences absolues et compté-tenu de la taille des deux sous-corpus (« 2 words in 2 corpora »). Ce bilan est, bien entendu, approximatif : l'échantillon des locuteurs originaires de Bohême contient 739 963 positions et celui des locuteurs moraves et silésiens, 476 191 positions. De plus, rappelons que le seul critère déterminant est ici la région où les locuteurs ont passé leur enfance.

³⁴ $L = 181,3767 \rightarrow$ la différence est statistiquement significative.

³⁵ $L = 318,4679$ pour la comparaison des fréquences absolues des deux lemmes au sein de la même population n (« two words in one corpus ») $\rightarrow$ la différence est statistiquement significative.

³⁶ Nous traduisons : „Zájmeno *onen* je knižní a obvykle slouží jako výrazný stylizační prostředek, mnohdy dodávající větě ironický či patetický ráz.“

³⁷ Nous traduisons : „Zájmeno *onen* zaniká v usu ob. a za ně nastupuje *tam-ten, ten-tam*.“

³⁸ De plus, son usage semble aujourd'hui complètement limité à l'anaphore et à l'exophore mémorielle, alors que des sources plus anciennes, tel *Slovník spisovného jazyka českého* (1964 : 394), le citent encore comme un démonstratif pouvant fonctionner au sein de la deixis situationnelle.

³⁹ Berger (1993 : 518) exclut TAMTEN de l'emploi endophorique de ce qu'il appelle le « système minimal » des démonstratifs tchèques. Ce terme recouvre, chez lui, le fonctionnement des démonstratifs dans le domaine de la langue écrite („die geschriebene Sprache“) soutenue, comprenant les types de textes suivants : textes administratifs, indications scéniques et, dans une moindre mesure, des textes spécialisés („Fachstil“) et de presse. En revanche, TAMTEN a sa place légitime – à la fois en emploi situationnel et endophorique – au sein du « système minimal élargi », comprenant des textes de presse et spécialisés caractérisés par une norme plus relâchée (1993 : 350 et 518). La langue parlée informelle aurait, quant à elle, tendance à remplacer TAMTEN par TAMHLETEN. Enfin, il ne nous semble pas juste d'affirmer, comme le suggère également Berger (1993 : 432), que TAMTEN soit exclu de l'exophore mémorielle (voir l'exemple 40 de *Slovník spisovného jazyka českého*, cité *infra*).

Or la situation est plus complexe, car dans le corpus SYN2015, c'est ONEN qui est largement supérieur à TAMTEN : 135 occurrences pour un million de positions contre seulement 5,95⁴⁰. Ce constat indique clairement que le registre et le type de textes jouent un rôle crucial dans le choix entre les deux formes, de même que le régime d'indexicalité⁴¹.

Voici maintenant une synthèse récapitulative des formes du démonstratif tchèque comprenant les informations suivantes : l'expression de la distance du référent par rapport au locuteur, le territoire sur lequel la forme est employée avec, le cas échéant, des précisions concernant son emploi dans les différents régimes d'indexicalité, ainsi que la/les variété(s) de langue où il apparaît⁴² :

Forme	Distance	Zone d'emploi	Variété de langue
TEN	neutre	partout	toutes
TENHLE	proximité	la Bohême (notamment)	soutenue parlée « tchèque commun » (soutenue écrite (?))
TENTO	proximité	partout	soutenue écrite (partout ; avec, éventuellement, une connotation archaïsante, voire livresque, en Bohême) soutenue parlée uniquement en Moravie et en Silésie parlée informelle uniquement en Moravie et en Silésie

⁴⁰ La différence est statistiquement significative, car $L = 17\,638,0501$ lorsque nous comparons les fréquences absolues des deux formes au sein de SYN2015. ONEN possède la fréquence relative la plus élevée au sein des textes de littérature spécialisée (180,3). Au sein des textes de fiction et des textes de presse, la fréquence relative atteint les chiffres 155,4 et 68,7 respectivement. TAMTEN est le plus utilisé dans les textes de fiction (ipm = 13,9). Son apparition est très basse dans les textes spécialisés (ipm = 2,03). Dans le corpus Diakorp v6, il n'y a aucune occurrence de TAMTEN contre une fréquence relative de 92,3 pour ONEN (toutefois, les deux formes ont été recherchées uniquement au nominatif singulier du masculin).

⁴¹ Malheureusement, les corpus étant à notre disposition (notamment ceux de langue écrite) semblent généralement peu représentatifs au regard de l'emploi situationnel, ce qui empêche d'avoir une image plus réaliste au sujet de l'usage de TAMTEN.

⁴² Les formes où /h/ est remplacé par /d/ ou disparaît et celles avec /(n)c/ ne sont pas considérés comme autonomes.

TAMTEN	éloignement	partout (avec une affinité pour l'emploi situationnel)	partout
ONEN	éloignement	partout (hormis l'emploi situationnel)	soutenue écrite (connotations archaïsantes et livresques) (soutenue parlée ; mêmes connotations)
TENHLETEN, TADYTEN TADYHLETEN,	proximité	la Bohême (notamment) (avec une affinité pour l'emploi situationnel)	« tchèque commun » (variante parlée du tchèque soutenu)
TAMHLETEN	éloignement	la Bohême (notamment) (uniquement en emploi situationnel)	« tchèque commun » (variante parlée du tchèque soutenu)
TUTEN, (TUHLETEN)	proximité	la Bohême de l'Ouest	trait dialectal de la Bohême de l'Ouest
TOŤTEN, HENTEN	proximité	la Moravie (notamment de l'Est) et la Silésie	dialectes moraves (notamment celui de l'Est) et silésien

Tableau 3.5 : Démonstratif en tchèque : synthèse des paradigmes

3.1.2 Le démonstratif tchèque : comparaison des fréquences des lemmes dans différents types de textes

À présent, pour donner au lecteur davantage de chiffres représentatifs de la fréquence des formes dans le tchèque contemporain, nous présenterons d'abord les données de deux dictionnaires de fréquences, à savoir *A Frequency Dictionary of Czech* (2011) de Čermák et Křen et *Frekvenční slovník mluvené češtiny* (*Dictionnaire des fréquences du tchèque parlé* ; 2007) de Čermák. Dans un second temps, nous comparerons les fréquences de plusieurs lemmes dans les corpus ORTOFON v1, SYN2015 et le PMK.

A Frequency Dictionary of Czech (2011), utilisant les données du CNT, regroupe quatre types de textes, à savoir : la fiction, la non-fiction (littérature spécialisée), les textes de presse et

les conversations⁴³. Bien que les masses de données représentant ces types de textes soient de tailles différentes, les valeurs de référence sont calculées relativement à un million de positions (elles sont « normalisées »). Ainsi, les quatre types de textes participent chacun à hauteur de 25 % au bilan final. Dans l'ensemble de l'échantillon, les lemmes des démonstratifs occupent les positions suivantes :

Lemme	Position (rang) dans le dictionnaire ⁴⁴	Fréquence réduite moyenne normalisée
TEN	3	20 828
TENTO	39	1 554
TENHLE	115	423
TENHLETEN	331	137
ONEN	354	127
TAMTEN	1730	27
TADYTEN	1980	23
TAMHLETEN	4987	7

Tableau 3.6 : Position des 8 lemmes du démonstratif tchèque les plus représentés dans *A Frequency Dictionary of Czech* (2011) avec leur rang et leur fréquence réduite moyenne normalisée indiquée en positions

Ce qui est le plus frappant, c'est, évidemment, la position de TEN, qui serait ainsi le 3^{ème} lemme le plus représenté en tchèque contemporain. Cette donnée est d'autant plus significative qu'elle constitue une preuve importante de la très haute fréquence de TEN dans les conversations. En effet, en 2004, les mêmes auteurs ont publié un autre dictionnaire classant les lexèmes tchèques en fonction de leur fréquence (*Frekvenční slovník češtiny / Dictionnaire des fréquences du tchèque*⁴⁵). Or dans ce dictionnaire, les conversations en tant que type de textes ne figuraient pas encore, en conséquence de quoi le lemme TEN y occupait « seulement » la 6^{ème} position. Ainsi, cette comparaison laisse voir à quel point TEN est sur-représenté dans ce type de textes.

Toutefois, s'il permet d'acquérir une bonne vue d'ensemble de l'usage des démonstratifs dans le tchèque actuel, le dictionnaire publié en 2011 présente quelques inconvénients, à commencer par

⁴³ Les données des conversations, puisées dans les corpus oraux ORAL2006 et ORAL2008, proviennent uniquement de Bohême. Les autres données ne sont pas marquées sur le plan diatopique.

⁴⁴ Il s'agit de la position (rang) dans un index hiérarchique qui liste les lemmes dans l'ordre décroissant en fonction de leur « fréquence réduite moyenne normalisée » (“overall normalized average reduced frequency”). Cela veut dire qu'outre la simple fréquence relative (normalisée) du lemme, c'est également sa dispersion dans le corpus qui est prise en compte (les occurrences sont-elles réparties de manière équilibrée dans le corpus ou apparaissent-elles davantage par faisceaux ?). Čermák et Křen (2011 : 4) précisent au sujet de la fréquence réduite moyenne : “ARF can be viewed as a dispersion-based correction of the frequency: the more even the dispersion, the closer ARF approaches to the frequency and vice versa. For words that occur only in one small cluster it is close to 1 regardless of their frequency. In practice, the ARF of common function words is typically around a half of their frequency, but it is considerably smaller for domain-specific words that occur only in a few documents.”

⁴⁵ Ce dictionnaire se base sur un corpus de trois types de textes écrits : fiction, textes spécialisés et textes de presse. Les auteurs ont également classé les lemmes dans l'ordre décroissant et en fonction de leur fréquence réduite moyenne. Néanmoins, les valeurs obtenues n'ont pas été « normalisées » (calculées relativement à un million de positions).

le fait – commun à tous les dictionnaires des fréquences du tchèque – de ne pas dissocier les emplois adnominaux et pronominaux, car se basant sur le système d’annotation du CNT⁴⁶. Pour ce qui est du lemme TEN, il regroupe également certains emplois qui ont, en réalité, peu de choses à voir avec un démonstratif (voir la Section 3.1.4 *infra*), et ce malgré tous les efforts de désambiguïsation déployés par les auteurs. Ensuite, le dictionnaire s’avère être d’une utilité limitée pour l’objectif du présent travail du fait que l’oral informel n’y est représenté qu’à hauteur de 25 %, et que, de plus, la collecte des données des conversations s’est limitée à la seule Bohême. C’est pour cela que le bilan reproduit *supra* ne reflète pas le profil de ce type de textes, par rapport auquel certaines formes – TENTO et ONEN – sont sur-représentées ou, au contraire, sous-représentées (TENHLE). Malgré cela, le bilan de Čermák et Křen (2011) permet de voir la haute fréquence de TEN dans la langue orale informelle par comparaison avec une langue écrite et (globalement) formelle.

Frekvenční slovník mluvené češtiny (Dictionnaire des fréquences du tchèque parlé ; 2007) est, quant à lui, certainement plus pertinent dans la mesure où il ne contient que des productions orales. Néanmoins, il présente notamment les deux inconvénients suivants : 1) basé sur les données du PMK, les enregistrements ont été faits exclusivement à Prague (mais pas uniquement avec des locuteurs originaires de Prague) et 2) il ne s’agit pas que de conversations, mais aussi d’entretiens dirigés. Enfin, la taille plutôt modeste du corpus (780 322 positions) pourrait elle aussi être perçue comme un inconvénient⁴⁷. Contrairement aux deux sources mentionnées précédemment, les lemmes sont ici classés en fonction de leur seule fréquence absolue (il n’est donc tenu compte ni de la dispersion ni de la fréquence relative). Il importe également d’ajouter que les enregistrements datent de la période 1988-1996⁴⁸. De même que pour les sources précédentes, les fréquences incluent les deux positions syntaxiques du démonstratif. La conception de lemmatisation étant différente pour ces textes de celle appliquée d’habitude par les annotateurs du CNK aux corpus oraux, un grand avantage réside dans l’exclusion des emplois de la particule *to* du lemme TEN (voir la Section 3.1.4). Autres précisions importantes : les auteurs ont considéré les formes avec la consonne /d/ (*tendle*), celles avec la syncope de la consonne (*tenle*) et celles avec /ct/ (*tendlecten* et *tendlencten*) comme des lemmes autonomes. Sont également recensés dans le dictionnaire les cumuls de démonstratifs avec d’autres mots grammaticaux (cf. *takový ten*), ces constructions étant appréhendées comme des lemmes à part. Il s’ensuit de cela que lorsque TEN apparaissait au sein d’une telle construction, cette occurrence n’était

⁴⁶ Ce système se justifie dans la mesure où il est fidèle à la tradition grammaticale tchèque, elle-même basée sur les propriétés typologiques de cette langue (voir la Section 3.1.3 *infra*).

⁴⁷ La version du PMK qui est aujourd’hui accessible sur le site du CNK est légèrement élargie par rapport à la version utilisée dans le cadre du projet du dictionnaire. Elle est aussi non lemmatisée, car aucune lemmatisation automatique de ce corpus n’a jamais été réalisée (dans le cadre du projet du dictionnaire, le corpus a été annoté manuellement, et cette annotation n’a pas été rendue publique par la suite).

⁴⁸ Ce qui crée tout de même un écart temporel non négligeable, pouvant aller jusqu’à 29 ans, entre ces données et les données des corpus de la série ORAL (période 2002-2011) et celles d’ORTOFON v1 (période 2012-2017).

pas comptabilisée dans le cadre du lemme de TEN seul. Mais le constat le plus pertinent pour nous est sans aucun doute la fréquence de TEN, qui occupe la 1^{ère} position (sic). Le Tableau 3.7 répertorie la position (rang) des formes du démonstratif dans le dictionnaire avec, pour chacune de ces formes, la fréquence relative (calculée par nous à partir de la fréquence absolue et la taille du corpus⁴⁹) :

Lemme	Position (rang) dans le dictionnaire ⁵⁰	Fréquence relative
TEN	1	51 608,2
TENDLE	144	585,7
TENHLETEN	161	517,7
TENDLETEN	162	513,9
TENHLE	174	463,9
TENTO	241	296
TENLETEN	356	75,6
TAMTEN	361	69,2
TENLE	360	67,9
TENDLENCTEN	364	62,8
TENDLECTEN	366	60,2

Tableau 3.7 : Position des 7 lemmes du démonstratif tchèque les plus représentés dans *Frekvenční slovník mluvené češtiny* (2007) avec leur rang et leur fréquence relative indiquée en positions

En plus de la position de TEN, le Tableau 3.7 met en évidence la forte position de TENHLE et de TENHLETEN à condition que nous subsumions – ce qui est notre cas – les formes avec /d/, /ct/ et celles avec la syncope de la consonne sous le même lemme que les formes de base. C’est à cette condition que TENHLE atteint une fréquence relative de 1 117,5, et que TENHLETEN atteint une fréquence relative de 1 230,2.

Faisons à présent une brève comparaison des fréquences des lemmes dans les conversations et à l’écrit. Dans les corpus ORTOFON v1 et ORAL v1, les fréquences relatives de TEN sont assez proches : 73 159,3 et 76 209,7 respectivement. Or dans SYN2015, la fréquence relative de TEN atteint uniquement le chiffre de 14 597,4. Les listes des fréquences disponibles sur le site du CNT nous montrent que le lemme TEN est le 6^{ème} lemme le plus représenté dans SYN2015 et le 1^{er} (sic) lemme le plus représenté dans ORAL v1⁵¹. Remarquons que ces positions sont exactement les mêmes que

⁴⁹ Ne sont pas inclus dans le tableau les cumuls du démonstratif avec d’autres mots grammaticaux.

⁵⁰ Il s’agit de la position dans un index hiérarchique qui liste les lemmes dans l’ordre décroissant en fonction de leur fréquence absolue.

⁵¹ Voir <https://www.korpus.cz/lists/>.

celles établies par *Frekvenční slovník češtiny* (Dictionnaire des fréquences du tchèque ; 2004) pour l'écrit et *Frekvenční slovník mluvené češtiny* (Dictionnaire des fréquences du tchèque parlé ; 2007) pour l'oral⁵². Le recours à la fonction de vraisemblance nous permet de déterminer si les différences de fréquence constatées sont statistiquement significatives. Aussi avons-nous recouru à cette fonction pour comparer, dans un premier temps, les fréquences absolues de TEN dans ORAL v1 et dans ORTOFON v1 et, dans un second temps, les fréquences absolues de TEN dans SYN2015 et dans ORTOFON v1⁵³. Pour cela, nous avons utilisé l'interface du CNT appelée Corpus Calculator (voir la Section 2.1.4 du Chapitre 2), plus précisément l'application « 2 words in 2 corpora », où nous avons saisi les fréquences absolues de TEN pour chacun des deux corpus ainsi que les tailles de ces derniers, exprimées en positions. Le score de la fonction de vraisemblance est égal à 138,8287 pour la paire ORAL v1 / ORTOFON v1 et il est égal à 148 230,463 pour la paire ORTOFON v1 / SYN2015. La valeur critique étant égale à 3,841 (degré de liberté = 1 et le seuil de significativité (α) = 0,05), les résultats prouvent la significativité statistique des différences entre les fréquences⁵⁴.

Si nous réalisons exactement le même calcul pour évaluer la significativité statistique des différences entre les fréquences absolues de TENHLE, TENTO et TENHLETEN dans ORTOFON v1 et SYN2015, le score de la fonction de vraisemblance est de 215,6557 pour TENHLE, 3 308,2109 pour TENTO et de 4 791,6368 pour TENHLETEN. Les Tableau 3.8 et 3.9 synthétisent les fréquences absolues et relatives ainsi que les scores de la fonction de vraisemblance. Comme tous les scores sont supérieurs à la valeur critique de 3,841, le test appuie la significativité statistique des différences ($\alpha = 0,05$) :

	ORTOFON v1		SYN2015	
	F abs.	F rel.	F abs.	F rel.
TEN	90 462	73 159,3	1 762 613	14 597,4
TENTO	494	399,5	298 500	2 472,1
TENHLE	1 236	999,6	76 835	636,3
TENHLETEN	757	612,2	815	6,7

Tableau 3.8 : Fréquences absolues et relatives des lemmes TEN, TENTO, TENHLE et TENHLETEN dans les corpus ORTOFON v1 (CNT ; 1 236 508 de positions) et SYN2015 (CNT ; 120 748 715 de positions)

⁵² Pour l'oral, le même constat est fait par Šonková (2008).

⁵³ Plutôt que dans ORAL v1, qui n'est pas un corpus représentatif (voir la Section 2.1.4).

⁵⁴ Il n'en reste pas moins que pour l'opposition ORTOFON v1 et ORAL v1, les valeurs restent assez proches.

ORTOFON v1 vs SYN2015	
TEN	148 230,463
TENTO	3 308,2109
TENHLE	215,6557
TENHLETEN	4 791,6368

Tableau 3.9 : Scores de la fonction de vraisemblance appliquée aux différences entre les fréquences absolues des lemmes TEN, TENTO, TENHLE et TENHLETEN dans les corpus ORTOFON v1 (CNT) et SYN2015 (CNT) ; valeur critique = 3,841

Maintenant, revenons un instant aux disproportions relatives à l'emploi des lemmes TEN, TENTO, TENHLE et TENHLETEN dans les trois parties de SYN2015, représentant trois types de textes appartenant au domaine de l'écrit : la fiction, la non-fiction et les textes de presse (certaines données ont déjà été présentées dans la section précédente). Le Tableau 3.10 récapitule les fréquences absolues et relatives par lemme et par partie :

	Fiction		Non-fiction		Presse	
	F abs.	F rel.	F abs.	F rel.	F abs.	F rel.
TEN	866 333	20 815,5	388 106	9 854,2	508 174	12 786,1
TENTO	35 791	860,0	179 590	4 559,9	83 119	2 091,3
TENHLE	52 863	1 270,2	7 525	191,1	16 447	413,8
TENHLETEN	688	16,5	80	2,0	47	1,2

Tableau 3.10 : Fréquences absolues et relatives des lemmes TEN, TENTO, TENHLE et TENHLETEN dans les trois parties du corpus SYN2015 (CNT) : fiction (41 619 587 de positions), non-fiction (39 384 709 de positions) et presse (39 744 419 de positions)

Nous avons de nouveau appliqué la fonction de vraisemblance⁵⁵ dans le but de calculer la significativité statistique de ces différences. Concrètement, nous avons comparé les fréquences absolues des lemmes obtenues pour les parties non-fiction et textes de presse avec les fréquences absolues obtenues pour la partie fiction. Le Tableau 3.11, synthétisant les résultats des calculs, prouve la significativité statistique de toutes les différences en matière de fréquence des quatre lemmes au sein des trois principaux types d'écrit tels qu'ils sont représentés dans SYN2015 :

	Non-fiction vs fiction	Presse vs fiction
TEN	164 037,8484	79 964,1286
TENTO	113 216,3645	21 645,0472
TENHLE	35 869,5049	18 501,0852
TENHLETEN	518,4623	640,3683

⁵⁵ La valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

Tableau 3.11 : Scores de la fonction de vraisemblance appliquée aux différences entre les fréquences absolues des lemmes TEN, TENTO, TENHLE et TENHLETEN dans a) la partie non-fiction du corpus SYN2015 et la partie fiction du corpus SYN2015 et b) la partie presse du corpus SYN2015 et la partie fiction du corpus SYN2015 ; valeur critique = 3,841

Partant des résultats affichés dans les Tableaux 3.8-3.11, nous constatons que si l’oral informel emploie, de manière générale, plus souvent TEN que l’écrit, le clivage est, selon nos attentes, le plus grand entre l’oral informel et les écrits appartenant à la non-fiction. TENTO est, de manière générale, plus employé à l’écrit que dans les conversations, notamment dans les textes de non-fiction. Quant à TENHLE, sa place est plus importante au sein de l’oral informel qu’au sein de l’écrit, à l’exception des textes de fiction, où il serait plus présent que dans l’oral informel⁵⁶. TENHLETEN semble très peu représenté au sein de l’écrit, où il n’apparaît quasiment qu’en fiction.

Pour ce qui est de l’oral, le lecteur aura par ailleurs remarqué la différence entre les fréquences relatives de TEN dans ORTOFON v1 et ORAL v1 d’une part et dans les données utilisées par *Frekvenční slovník mluvené češtiny* (Dictionnaire des fréquences du tchèque parlé ; 2007) d’autre part. Cela pourrait s’expliquer par la présence d’entretiens parmi ces dernières, type de textes où l’usage de TEN serait moindre du fait de sa moins grande spontanéité. Ce facteur réduirait ainsi la fréquence relative pour l’ensemble. Autre constat intéressant : en comparant les fréquences relatives de TENTO, TENHLE et TENHLETEN dans ORTOFON v1 et dans les données utilisées par le dictionnaire de fréquences susmentionné, nous nous apercevons que pour TENTO, cette fréquence est la plus élevée pour l’ensemble d’ORTOFON v1 (399,5) et la moins élevée pour la partie d’ORTOFON v1 comprenant uniquement des locuteurs de Bohême (194,6). Dans les données du dictionnaire de fréquences, elle se situe entre les deux : 296. Dans le cas de TENHLE, c’est l’inverse qui se produit : la fréquence relative est de 1 313,6 pour la partie Bohême d’ORTOFON v1, contrairement à 999,6 pour l’ensemble. Dans les données du dictionnaire, elle est égale à 1 117,5. Du moins à première vue, cette gradualité pourrait refléter la différence de nature entre les trois sources de données : langue entièrement spontanée de locuteurs de l’ensemble du pays vs langue semi-spontanée de locuteurs majoritairement pragoais (donc de Bohême) vs langue entièrement spontanée de locuteurs de Bohême⁵⁷. Le Tableau 3.12 synthétise cette différence ; le Tableau 3.13 contient, quant à lui, les résultats de l’application de la fonction de vraisemblance aux différences entre, d’une part, les fréquences absolues pour l’ensemble d’ORTOFON v1 et celles relatives uniquement à la partie

⁵⁶ Il nous semble légitime de faire ce constat dans la mesure où ORTOFON v1 est un corpus oral représentatif. Mais même dans ORAL v1 – où les locuteurs de Bohême sont sur-représentés –, la fréquence relative de TENTO est de 1 244,6, soit également (légèrement) plus basse que dans la partie fiction de SYN2015.

⁵⁷ Seul TENHLETEN défie un peu cette conclusion, car si son emploi est moins répandu dans l’ensemble d’ORTOFON v1 que dans les deux autres sources, il est plus répandu au sein de la source du dictionnaire de fréquences que dans la partie Bohême d’ORTOFON v1. C’est la nature des données qui pourrait être à l’origine de ce phénomène : la source du dictionnaire favoriserait davantage l’emploi situationnel, qui est le domaine de prédilection de TENHLETEN (voir la Section 3.1.1).

Bohême d'ORTOFON v1 et, d'autre part, les fréquences absolues pour l'ensemble d'ORTOFON v1 et celles relatives aux données du dictionnaire de fréquences⁵⁸ :

	ORTOFON v1		ORTOFON v1 - Bohême		Frekvenční slovník (2007)	
	F abs.	F rel.	F abs.	F rel.	F abs.	F rel.
TEN	90 462	73 159,3	54 690	73 909,1	40 271	51 608,2
TENTO	494	399,5	144	194,6	231	296,0
TENHLE	1 236	999,6	972	1 313,6	872	1 117,5
TENHLETEN	757	612,2	665	898,7	960	1 230,3

Tableau 3.12 : Fréquences absolues et relatives des lemmes TEN, TENTO, TENHLE et TENHLETEN dans l'ensemble du corpus ORTOFON v1 (CNT ; 1 236 508 de positions), la partie Bohême du corpus d'ORTOFON v1 (CNT ; 739 963 positions) et le corpus du *Frekvenční slovník mluvené češtiny* (2007 ; 780 322 positions)

	ORTOFON v1 – Bohême vs ORTOFON v1	Frekvenční slovník (2007) vs ORTOFON v1
TEN	3,8252	3 665,4599
TENTO	60,2343	14,2563
TENHLE	40,9002	6,369
TENHLETEN	17,5292	214,8436

Tableau 3.13 : Scores de la fonction de vraisemblance appliquée aux différences entre les fréquences absolues des lemmes TEN, TENTO, TENHLE et TENHLETEN dans a) la partie Bohême du corpus d'ORTOFON v1 et l'ensemble du corpus ORTOFON v1 et b) le corpus du *Frekvenční slovník mluvené češtiny* (2007) et l'ensemble du corpus ORTOFON v1 ; valeur critique = 3,841

Tous les scores se situent au-dessus du seuil de significativité statistique sauf celui relatif à la comparaison des fréquences absolues de TEN dans la partie Bohême d'ORTOFON v1 et dans l'ensemble de ce corpus : ce score est légèrement inférieur à la valeur critique de 3,841.

3.1.3 Quel modèle d'analyse syntaxique pour les démonstratifs adnominaux tchèques ?

Dans la tradition grammaticale tchèque, l'habitude est solidement établie de traiter tous les emplois des démonstratifs appartenant au type nominal comme des « pronoms ». Cette pratique – où le terme de « pronom » devient ainsi une désignation générique – remonte déjà aux sources aussi anciennes que les traités de grammaire historique de Gebauer (1929 et 1960 [1896]) ou l'article pionnier de Mathesius (1926) consacré aux différentes fonctions de TEN. Une bonne part des sources postérieures restent fidèles à cette pratique, comme par exemple *Slovník spisovného jazyka českého* (1966) ou *Česká mluvnice* (1960), mais aussi des sources bien plus modernes comme *Akademická*

⁵⁸ La valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

gramatika spisovné češtiny (2013) ou *Mluvnice současné češtiny I* (2015 [2010]). Ce syncrétisme terminologique est en partie dû au fait déjà mentionné, à savoir que les mêmes formes apparaissent à la fois dans les emplois pronominaux et adnominaux.

En même temps, des voix se sont toujours levées pour remettre en question, ne serait-ce que timidement, cette tendance, à commencer encore par Mathesius qui, dans ce même article de 1926, parle certes de « pronom », mais de « pronom épithète » („přívlastkové zájmeno“) pour les emplois adnominaux. Dans *Mluvnice spisovné češtiny II*, Trávníček, bien qu'il parle lui-aussi de « pronoms », opère la distinction entre ce qu'il appelle les emplois « nominaux » et « adjectivaux » (dans ce dernier cas, le démonstratif est traité comme l'épithète („přívlastek“) du nom qu'il introduit dans le discours). Plus tard, cette même position est encore adoptée par Adamec dans son étude comparant l'usage des démonstratifs en tchèque et en russe (1983) ainsi que par Berger (1993). Dans son article publié en 1999, Štícha cherche à éviter la controverse en employant le terme neutre de « lexème »⁵⁹. Le même choix est opéré par Zíková (2017), employant les termes de « lexème » et d'« expression ». Schneiderová (1993), même si elle continue à employer le terme de « pronom », est bien consciente de sa nature problématique :

Les manuels de langue courants rangent l'expression *ten* (avec ses formes genrées *ta* et *to*) parmi les pronoms, caractérisés comme les mots qui ne font que se substituer aux noms ou aux adjectifs en question. [...] Si des travaux plus récents [...] ne cherchent pas à éviter l'appellation traditionnelle de **pronom**, ils préfèrent les termes de **mot déictique** ou **identificateur**. En effet, ces termes ne font plus pencher vers une restriction de cette classe de mots aux seuls noms et ne nous font plus considérer la fonction de substitution (pro-nom) comme un trait caractéristique des mots déictiques. (1993 : 31)⁶⁰

Concrètement, les termes de « mots déictiques » („deiktická slova“) et « identificateurs » („identifikátory“) apparaissent à la fois dans les travaux de Hirschová (1988) et Komárek (1978). Ce dernier exprime sa réticence face à l'usage du terme « pronom » en écrivant que :

La capacité de substitution ne constitue que l'une des fonctions des mots déictiques (lors de ce que l'on appelle la pronominalisation) et dans le cas de certains types de mots déictiques (tels les pronoms de la 1^{ère} et 2^{ème} personne), celle-ci n'est nullement en jeu. (1978 : 5)⁶¹

⁵⁹ « Nous employons ici l'appellation générique de *lexème* et non pas celle, traditionnelle, de *pronom*, car ce lexème possède une fonction pronominale uniquement dans ses emplois substantivaux et non pas dans ses emplois adjectivaux, auxquels cette étude est exclusivement consacrée. » (1999 : 134) Nous traduisons : „Užíváme tu obecného označení *lexém*, nikoli tradičně *zájmeno*, neboť tento lexém má zájmennou funkci pouze při užití substantivním, nikoli však při užití adjektivním, o které nám v tomto příspěvku výhradně jde.“

⁶⁰ Nous traduisons : „Výraz *ten* (s rodovými podobami *ta*, *to*) je v běžných jazykových učebnicích řazen mezi **zájmena**, jež jsou charakterizována jako slova, která jen zastupují příslušná jména podstatná nebo přídavná. [...] Novější jazykovědné práce [...] se sice tomuto tradičnímu označení *zájmeno* nevyhýbají, dávají však přednost termínům **deiktické slovo** nebo **identifikátor**. Tyto termíny totiž již nespádají k omezování třídy slov jen na tzv. jména a k tomu, abychom za charakteristický rys deiktických slov považovali funkci zástupnou (tj. za-jméno).“

⁶¹ Nous traduisons : „Zástupnost je pouze jedna z funkcí deiktických slov (při tzv. pronominalizaci) a u některých druhů deiktických slov (např. zájm. 1. a 2. os.) s ní není vůbec možno počítat.“

Une position toute particulière est adoptée par *Mluvnice češtiny 2* (1986 : 81-104). Ces auteurs, bien qu'ils ne rejettent pas le terme de « pronom » dans son acception générique, emploient assez souvent et y compris pour les démonstratifs nominaux le terme de « mots déictiques ». Ensuite, les démonstratifs TEN, TENTO et TENHLE sont explicitement désignés par eux comme des « pronoms adjectivaux » („adjektivní zájmena“), dont la fonction dite « première » est celle d'introduire un nom. Par conséquent, les cas où le démonstratif apparaît seul sont analysés comme ressortissant à la « fonction secondaire » de celui-ci, sachant que ces emplois relèveraient d'une transposition et feraient automatiquement apparaître un phénomène d'ellipse du nom attendu. Ce parti-pris encourage une vision en termes de polysémie, comme nous essaierons de l'expliquer dans la Section 3.1.4 *infra*.

Certains auteurs vont même jusqu'à parler de « déterminants » : Uhlířová (1992) emploie le terme de « déterminants » („determinátory“), Hlavsa (1975, 1987) parle de « déterminants identifiants » („delimitátory identifikační“). Veselovská (2014), dont l'analyse fait sienne l'hypothèse du syntagme déterminatif (“determiner phrase”) formulée au sein du paradigme de la grammaire générative (Abney, 1987), utilise également ce terme. Son objectif est de démontrer l'applicabilité de cette hypothèse au tchèque en particulier et à toutes les langues dépourvues d'articles en général. Veselovská utilise des données du CNT (textes de fiction du 19^{ème} et du 20^{ème} siècle) pour prouver que malgré les possibilités théoriques du tchèque et l'image qu'en ont souvent les locuteurs natifs, le complexe nominal⁶² de cette langue obéirait globalement aux mêmes règles d'agencement strict que celui de l'anglais. Elle écrit littéralement ceci :

Although Czech speakers often believe in a free word order with no limits in their mother tongue, I have demonstrated several results of corpora searches which illustrate relatively strict rules attested in the ordering inside complex noun phrases⁶³. (2014 : 28)

Loin de contester l'ensemble des conclusions de cette étude éclairante à bien des égards, il nous semble que l'auteure minimise tout de même deux phénomènes assez importants. D'abord, nous pensons justement à la possibilité théorique (quoiqu'extrêmement peu sollicitée en pratique⁶⁴) de postposer au nom les éléments que les langues telles que l'anglais et le français classent dans la catégorie des déterminants⁶⁵. Mais, surtout, il y a la possibilité d'apparition simultanée de plusieurs de ces éléments devant le nom. Ainsi, Veselovská a beau souligner que les éléments dans

⁶² C'est également elle qui emploie ce terme, que nous allons nous-même utiliser pour le tchèque dans l'ensemble de ce travail.

⁶³ Ainsi, l'auteure montre que la postposition des démonstratifs, des possessifs et des adjectifs non complétés est très rare dans son corpus (contenant également des textes d'auteurs traditionnellement considérés comme archaïsants) : elle n'apparaît, en moyenne, que dans 6,4 % des complexes nominaux (2014 : 19-20).

⁶⁴ Dans l'Introduction générale et le Chapitre 2, nous avons expliqué à quel point cet ordre est marqué comme vieilli et minoritaire.

⁶⁵ Il s'agit d'une classe fermée d'éléments qui tendent à s'exclure mutuellement sur une base de critères distributionnels, et qu'Abney (1978) appelle « the functional domain ». Dans les langues comme l'anglais et le français, les éléments de cette classe ne peuvent pas, bien entendu, se postposer au nom (**maison cette*).

l'antéposition du nom tête (pré-déterminants, déterminants centraux, post-déterminants, adjectifs sans compléments) suivent globalement le même ordre d'apparition qu'en anglais⁶⁶, cette dernière langue n'autorise pas (pas plus que le français) la coprésence, par exemple, d'un démonstratif et d'un possessif ou d'un indéfini dans la position réservée au déterminant central ("the central determiner slot") :

1) **tenhle** **můj** **znám-ý**
 tenhle.M.NOM.SG mien.M.NOM.SG connu-M.NOM.SG
 *cette ma connaissance / cette connaissance à moi

nějak-é **to** **piv-o**
 quelque-N.NOM.SG ten.N.NOM.SG bière-NOM.SG
 *quelque/un peu cette/la bière / un peu de (cette/la) bière (voire : quelques bières)⁶⁷

vs *this my acquaintance
 *some this/the beer

Pour revenir à Uhlířová (1992 : 247), celle-ci écrit à propos de l'ensemble de ces éléments :

Leur position initiale reflète leur fonction au sein du SN. Ce ne sont pas – contrairement à l'écrasante majorité des adjectifs – des expressions spécifiant le sens du substantif dominant au sein du SN en question tout en exprimant une propriété, une qualité ou un trait le concernant, mais des expressions caractérisant la façon dont le SN est employé dans l'acte de communication⁶⁸.

Autrement dit, l'on pourrait résumer en disant que le trait commun à tous ces éléments est celui d'actualiser le nom en discours. Ensuite, il n'est pas non plus convenable d'en faire une sous-catégorie des adjectifs. La raison qui est ici avancée par Uhlířová est d'ordre sémantique : contrairement aux adjectifs, les éléments en question ne sont pas porteurs d'une sémantique qualificative ou relationnelle⁶⁹. Et pourtant, la prédilection qu'a Uhlířová pour le terme de « déterminant » pourrait sembler quelque peu étonnante, voire contradictoire, compte-tenu justement du phénomène que son article étudie, à savoir le cumul de ces éléments en tchèque⁷⁰. En effet, l'existence de ce cumul

⁶⁶ Par exemple dans :

quantificateur	déterminant central	numéral	adjectif	tête nominale
<i>all</i>	<i>these</i>	<i>four</i>	<i>young</i>	<i>people</i>
<i>všichni</i>	<i>tihle</i>	<i>čtyři</i>	<i>mladí</i>	<i>lidé.</i>

⁶⁷ Cet exemple montre, entre autres, que les seuls contextes en français qui autorisent la contiguïté de deux éléments pouvant fonctionner comme déterminants centraux sont ceux d'une expression partitive (et donc de la présence de deux SN) : « une certaine quantité d'une bière préalablement déterminée » (le premier élément est un pronom). Si cette lecture est également possible en tchèque, elle n'est pas la seule.

⁶⁸ Nous traduisons : „Počátková poloha odpovídá jejich funkci v NP. Nejsou to — na rozdíl od naprosté většiny adjektiv — výrazy, které by blíže specifikovaly význam dominujícího substantiva v příslušné NP tím, že by vyjadřovaly nějakou jeho vlastnost, kvalitu, příznak apod., nýbrž výrazy, které charakterizují způsob užití NP v komunikačním aktu.“

⁶⁹ Les raisons du refus de cette étiquette pour le démonstratif sont, comme nous pourrions encore le voir dans le Chapitre 4 (Section 4.1.2), consacré au démonstratif adnominal français, plus nombreuses.

⁷⁰ Il faut cependant préciser que l'auteure laisse de côté les possessifs. Nous renverrons à cette étude dans les Chapitres 6 et 7, car elle contient une série de remarques très révélatrices au sujet des cumuls du démonstratif défini avec d'autres éléments du « domaine fonctionnel ».

constitue une violation de la règle distributionnelle traditionnellement mise en avant en rapport avec la classe des déterminants, à savoir la présence d'un seul déterminant central par SN (cf. Abney, 1987 ; Jackendof, 1977).

Au vu de tout cela, nous préférons, comme nous l'avons déjà évoqué, ne pas souscrire à la notion de déterminant pour le tchèue. La position que nous suivons est celle de Himmelmann (1997). En effet, Himmelmann remet en question l'usage de cette notion pour les langues dans lesquelles le processus de grammaticalisation aboutissant à l'apparition d'un article défini n'a pas encore eu lieu ou n'est pas achevé :

Zweitens ist die Universalität der Kategorie Determinator [...] äußerst zweifelhaft. Es scheint vielmehr so zu sein, dass diese Kategorie erst im Zug der adnominalen Grammatikalisierung von D-Elementen entsteht.
(1997 : 110)

In all den Fällen, in denen es genügend empirische Evidenz für die Annahme einer Kategorie *Determinator* gibt (z.B. im Deutschen, Englischen oder den romanischen Sprachen), gibt es auch immer mindestens ein grammatikalisches D-Element (einen Definitartikel oder einen spezifischen Artikel). Das heißt, es sind mir keine Fälle bekannt, wo z.B. ein Indefinitartikel und Possessiva oder Demonstrativa und Quantoren sich morphosyntaktisch so ähnlich verhalten, dass sie unter eine morphosyntaktische Kategorie subsumiert werden können, ohne dass in der betreffenden Sprache nicht auch mindestens ein grammatikalisches D-Element existierte. (1997 : 133)

Autrement dit, ce que montre Himmelmann, c'est qu'une fois que l'article défini (ou spécifique) fait son apparition dans le système, d'autres éléments adnominaux du domaine fonctionnel (Abney, 1987) – indéfinis, possessifs, quantificateurs et démonstratifs – en viennent à adopter le même comportement syntactico-distributionnel, à savoir à occuper la position extrême du complexe nominal („auf den linken oder rechten Rand des nominalen Syntagmas“ (1997 : 188)). Littéralement, il souligne le fait que les autres éléments du domaine fonctionnel („D-Elementen“) adoptent le même comportement syntaxique que l'article produit par le processus de grammaticalisation. Ce constat vaut naturellement pour les démonstratifs adnominaux, qui, auparavant, jouissaient d'une mobilité beaucoup plus grande à l'intérieur du complexe nominal⁷¹ :

Es sei denn, in der betreffenden Sprache gibt es eine stark grammatikalisierte Konstruktion mit einem D-Element, an deren Syntax die Syntax der Konstruktionen mit Demonstrativa angeglichen ist (wie das etwa in westeuropäischen Sprachen überwiegend der Fall ist, wo sich die Morphosyntax der Demonstrativa weitestgehend an die der Definitartikel angeglichen ist). Gegenbeispiele zu dieser Behauptung – also z.B. eine Sprache, in der es keinen definiten oder spezifischen Artikel gibt, die Stellung der Demonstrativa in nominalen Ausdrücken aber rigoros festgelegt ist und die Demonstrativa in einigen syntaktischen Kontexten

⁷¹ Diessel (1999 : 65) illustre ce même phénomène sur le cas de l'anglais : “Though articles and demonstratives are not always members of the same category, there is good evidence that English *the* and *this/that* have the same categorial status. English is one of the few languages in my sample in which adnominal demonstratives are in paradigmatic relationship with articles, possessives and other noun operators (e.g. *every*), which share a number of syntactic features.” Remarquons qu'un synonyme du terme « élément fonctionnel » apparaît dans cette citation : celui d'« opérateur nominal » (“noun operator”).

(z.B. Adpositionalphrasen) prinzipiell nicht verwendet werden können – sind mir nicht bekannt. (1997 : 156)

Mais Himmelmann ne s'arrête pas là. Sur son matériel empirique composé d'une série de langues appartenant à des familles différentes à la fois sur les plans typologique et génétique, il montre que ce n'est pas uniquement le déterminant en tant que partie du discours qui émerge au bout de la grammaticalisation du démonstratif en article défini, mais l'ensemble de l'unité syntaxique appelé « syntagme/groupe nominal » :

Phrasale Gestalten, so die hier entwickelte Hypothese, entstehen im Rahmen von Grammatikalisierungsprozessen; die für phrasale Gestalten typischen Gestalteeigenschaften sind mithin Charakteristika von grammatikalisierten Konstruktionen. Diese Hypothese [...] ist empirisch falsifizierbar. (1997 : 156)

Cette unité est caractérisée par des contraintes d'agencement beaucoup plus strictes qu'un simple complexe nominal, ce qui explique pourquoi un déterminant ne peut y changer de place, pourquoi plusieurs éléments ne peuvent y occuper la position du déterminant central, mais aussi, par exemple, pourquoi la présence d'un déterminant devient obligatoire pour une grande partie (sinon la majorité) des noms actualisés en discours⁷². La justesse de l'hypothèse de Himmelmann est étayée par des travaux empiriques comme *Émergence et évolution du syntagme nominal en français* (1994) de Goyens, comparant le français actuel à l'ancien français et au latin, ou bien l'étude de Traugott (1992) portant sur l'évolution de la syntaxe anglaise. Moyennant un corpus composé de textes latins et de leurs traductions en ancien français et en français moderne, Goyens montre comment le SN a vu le jour en français et comment il y a acquis une existence stable et solide (voir la Section 4.1.2). Traugott, mobilisant également des exemples authentiques, montre que l'ancien anglais autorisait – exactement comme le tchèque d'aujourd'hui – la co-occurrence du démonstratif adnominal avec d'autres éléments fonctionnels, dont le possessif.

Soulignons en passant que pour éviter de parler de déterminants – catégorie non-universelle selon lui –, Himmelmann utilise le terme de « D-Elementen ». Ce terme ne désigne pas uniquement les éléments du domaine fonctionnel engagés dans le processus de grammaticalisation (donc, dans l'écrasante majorité des cas, les démonstratifs), mais aussi les produits de ce processus (articles

⁷² Voir l'étude de Carlier et Combettes (2015) pour des conclusions fort semblables, bien que concernant plus spécifiquement la différence typologique entre le latin et le français. Les auteurs emploient le terme de « configurationnalité », emprunté à la grammaire générative (Hale, 1983) : « Il semble ainsi pertinent de s'appuyer sur la notion de configurationnalité qui permet d'opposer les langues configurationnelles, à structure fortement hiérarchisée, et les langues non configurationnelles ou moins configurationnelles, à structure "plate". Dans les langues configurationnelles, les relations entre les constituants relèvent de l'hypotaxe, de la dépendance syntaxique, alors que dans le cas d'un système linguistique à structure "plate", les relations entre les syntagmes sont davantage de l'ordre de la parataxe. » (2015 : 31) Cette distinction tient autant pour le verbe et ses compléments que pour le nom et ses modificateurs/actualisateurs. Dans cette perspective, le français est une langue clairement configurationnelle, contrairement au tchèque, qui, bien plus proche du latin, est (plutôt) non configurationnel.

définis et spécifiques) ainsi que, par extension, tous les autres éléments fonctionnels, dont le comportement syntaxique s'aligne plus tard sur celui des articles. Comme l'écrivent également Carlier et Combettes à propos des changements morphosyntaxiques ayant accompagné le passage du latin au français :

Prenant en compte l'impact de la fréquence comme moteur de la grammaticalisation et force créatrice de nouvelles constructions grammaticales [...], on peut supposer que la haute fréquence de la séquence < Article défini + N (+ Modifieur) > a engendré une nouvelle structure syntagmatique, qui prévoit une position structurale pour le déterminant. Une fois que cette structure existe, d'autres morphèmes grammaticaux ou lexicaux peuvent s'y insérer. C'est ainsi que se forment d'autres articles, tels l'article indéfini singulier dérivé du numéral de l'unité et l'article partitif formé à partir de la préposition *de* suivi de l'article défini, et que se forme donc le paradigme des articles. (2015 : 52)

Dans ce travail, le terme d'« éléments fonctionnels » sera également retenu pour le tchèque. En revanche, si les mots appartenant à cette catégorie n'y sont, sur le plan syntaxique, ni des pronoms *stricto sensu*, ni des adjectifs, ni même des déterminants, que sont-ils alors ? Un élément de réponse important à cette question est apporté par plusieurs travaux, dont Himmelmann (1997) et Diessel (1999). Ce dernier passe en revue toutes les solutions possibles. La première postule l'existence de déterminants comme catégorie universelle, mais, comme nous venons de le voir, cette hypothèse n'est pas satisfaisante. Elle est donc rejetée par Diessel. La seconde est élaborée par Abney (1978) lui-même. Fidèle au formalisme de la théorie X-barre, Abney cherche à unifier la catégorie des déterminants avec celle des pronoms. Ainsi, pour lui, il n'y aurait, *in fine*, aucune différence substantielle entre un démonstratif en emploi adnominal et pronominal, les deux représentant la tête d'un syntagme déterminatif, respectivement pourvue ou dépourvue d'une projection⁷³. Cette approche est elle aussi rejetée par Diessel faute de preuves empiriques. Paradoxalement, la solution qui est retenue par lui (et par Himmelmann aussi) nous ramène à la notion de pronom :

[...] there are languages in which adnominal demonstratives are independent pronouns that are adjoined to a coreferential noun in apposition, and there are other languages in which adnominal demonstratives are determiners, which require a cooccurring noun. (1999 : 65)

En effet, Diessel cite des exemples d'autres langues où, comme en tchèque, le démonstratif affiche un comportement syntaxique (beaucoup) plus souple que dans les langues telles que l'anglais. Ainsi, dans ces langues – majoritaires dans son corpus ⁷⁴ –, les démonstratifs adnominaux seraient

⁷³ “Abney’s analysis is largely motivated by conceptual and theory internal considerations, as he himself points out (Abney 1987: 351). He abandons the distinction between determiners and pronouns primarily because this strengthens the ‘Det-as-head analysis’, which, in turn, allows him to maintain that the structure of NP/DP is basically the same as the structure of IP. (Diessel, 1999 : 67)

⁷⁴ Le corpus de Diessel comprend 85 langues, dont 61 usent de la même forme de démonstratif dans les deux types d'emploi. À cela s'ajoute une précision cruciale : “In most of these languages there is no evidence that pronominal and adnominal demonstratives belong to different categories.” (1999 : 60)

tout de même des pronoms, mais des pronoms assez singuliers du fait d’être apposés/juxtaposés au nom avec lequel ils apparaissent en discours. L’auteur choisit ensuite l’exemple du tuscara, où le nom comme le démonstratif peuvent tous les deux et indépendamment de la présence de l’autre constituer un SN, sachant que la position vis-à-vis de l’autre est libre. Dans cette description, nous reconnâtrons facilement la situation qui est celle du tchèque (notamment pour ce qui est du premier critère)⁷⁵ :

2)	Mně	se	lib-í	tenhle	dům.	
	moi.DAT	REFL	plaître-PRS.3SG	tenhle.M.NOM.SG	maison.NOM.SG	
	<i>Cette maison me plaît.</i>					
	Mně	se	lib-í	tenhle.		
	moi.DAT	REFL	plaître-PRS.3SG	celui-ci.M.NOM.SG		
	<i>Celle-là me plaît.</i>					
	Mně	se	lib-í	dům.		
	moi.DAT	REFL	plaître-PRS.3SG	maison.NOM.SG		
	<i>Une/la maison me plaît.</i>					
	tento		dům	/	? dům	tento
	tento.M.NOM.SG		maison.NOM.SG		maison.NOM.SG	tento.M.NOM.SG

Evidemment, il ne s’agit en aucun cas de conclure à une quelconque uniformité, sur le plan des démonstratifs, des langues où le démonstratif adnominal manifeste un degré d’indépendance et de liberté similaire. Ces langues font toutes apparaître des idiosyncrasies et des écarts plus ou moins saillants les unes par rapport aux autres. Toujours est-il que ces comparaisons font apparaître une certaine dichotomie entre, d’une part, les langues comme le français et l’anglais et, d’autre part, les langues comme le tuscara et la majorité des langues slaves. Dans les premières, le démonstratif adnominal appartient à la classe des déterminants et, le cas échéant (comme en français), il peut même être formellement différent du démonstratif pronominal. Dans les secondes, il se laisserait davantage analyser comme un pronom apposé ou juxtaposé au nom (et ses compléments), même si la nature de cette apposition/juxtaposition demeure en partie floue⁷⁶. Cette dichotomie paraît d’autant plus convaincante qu’elle recouvre l’opposition entre les systèmes linguistiques se situant, respectivement, après et avant le processus de grammaticalisation étudié par Himmelmann (1997). En effet, celui-ci confirme qu’au début du processus, les constituants d’un complexe nominal sont liés entre eux par des liens beaucoup moins étroits que les constituants d’une construction déjà grammaticalisée⁷⁷, les

⁷⁵ En plus de ces deux constats, Diessel en ajoute un troisième pour le tuscara, à savoir que les deux éléments sont souvent séparés par une pause dans l’intonation. À notre connaissance, ce dernier phénomène ne se produit pas en tchèque.

⁷⁶ En des termes délibérément vagues, Diessel (1999 : 62) parle de « some kind of appositional structure ».

⁷⁷ „[...] die lockeren Gefüge (Gruppen), die am Beginn der Grammatikalisierung einer Konstruktion stehen, generell keine (oder weniger starke) Gestalteigenschaften aufweisen als grammatikalisierte Konstruktionen.“ (1997 : 156) De

démonstratifs ne représentant pas, à ce moment-là, des éléments fermement intégrés dans ces structures („[...] sie nicht integraler Bestandteil einer phrasalen Gestalt sind.“ (1997 : 156)).

Pour le tchèque, il serait, par ailleurs, fort intéressant de pouvoir mener une étude sur corpus qui concernerait l'évolution de l'ordre démonstratif-nom vs nom-démonstratif. Malheureusement, une telle étude ne nous semble pas envisageable pour le moment, les corpus de tchèque diachronique n'étant pas lemmatisés. Si cette étude permettait de prouver qu'un figement important s'est produit, au cours du temps, quant à la position du démonstratif (avec, bien entendu, une nette préférence pour l'antéposition) contrairement à des états plus anciens de la langue, où les deux ordres semblent avoir été tout à fait acceptables⁷⁸, cela pourrait constituer une preuve que le démonstratif adnominal tchèque tend aujourd'hui à se rapprocher du statut de déterminant.

3.1.4 Les emplois adnominaux et pronominaux : peut-on parler de polysémie/homonymie ?

Dans la Section 3.1.1 *supra*, nous écrivions que tous les membres du paradigme des démonstratifs nominaux définis connaissent deux modes d'emploi : l'emploi adnominal et pronominal. À présent, le moment est venu de mieux examiner cette affirmation. Plus exactement, il s'agit de se demander si, dans ces deux types d'emploi, nous avons encore affaire au même sens (à la même acception) du lexème. Les interrogations suivantes de Berger (1993 : 261) nous serviront d'introduction au problème :

Sowohl die Aufgliederung der Beschreibung nach rein syntaktischen Verwendungen der Demonstrativpronomina und unterschiedlichen Verweisarten (Anapher, Katapher, Deixis) als auch die Untergliederung nach verschiedenen syntaktischen Funktionen (substantivische vs. adjektivische Verwendung u.ä.), die ich beide aus der bisherigen Literatur übernommen (z.T. Auch „herausgefiltert“) habe, **legen allerdings die Frage nahe, ob die rein morphologisch abgegrenzten Einheiten *ten*, *tento*, *tenhle* usw. nicht weiter untergliedert werden müssen**⁷⁹. Dabei betrifft die erste Gliederung die semantischen Merkmale der Demonstrativpronomina und wirft die Frage nach H o m o n y m i e und P o l y s e m i e dieser Pronomina auf, die zweite Gliederung betrifft die grammatischen Merkmale der Pronomina, also ihre Zuordnung zu einzelnen Wortarten.

Autrement dit, se pose ici la question de savoir à quel point les différences et asymétries identifiées dans les emplois d'une même forme sur le plan morphologique pourraient pointer vers un

manière générale, ladite dichotomie semble également recouper l'opposition entre systèmes configurationnels et systèmes non configurationnels (Hale, 183 ; Carlier et Combettes, 2015).

⁷⁸ Juste à titre indicatif, nous avons recherché dans les corpus SYN2015 et Diakorp v6 les complexes nominaux suivants : *tento muž* (cet homme) / *muž tento* (*homme ce) et *toto město* (cette ville) / *město toto* (*ville cette). Si dans SYN2015, l'antéposition est majoritaire (827 occurrences de *tento muž* contre 57 occurrences de *muž tento* et 120 occurrences pour *toto město* contre seulement 3 occurrences pour *město toto*), dans Diakorp v6, *muž tento* apparaît 12 fois contre 7 occurrences de *tento muž* et *město toto* apparaît 17 fois contre 15 occurrences de *toto město*.

⁷⁹ Nous soulignons.

phénomène de polysémie, voire d'homonymie sur le plan sémantique. Dans la littérature, il semble y avoir à cet égard un consensus : certes les formes peuvent exercer plusieurs fonctions, voire être dotées de sens différents (situation de polysémie), mais ces différences n'atteignent pas de telles proportions qu'il soit nécessaire de conclure à l'existence de différents lexèmes (situation d'homonymie) :

Was die Frage nach der Semantik der Demonstrativpronomina angeht, besteht in der Literatur und der lexikographischen Praxis Einigkeit, daß die Pronomina jeweils mehrere Funktionen (oder Bedeutungen?) haben, deren Abweichungen aber nicht soweit gehen, daß es nötig wäre, von verschiedenen Lexemen zu sprechen.
 Berger
 (1993 : 261)

En vérité, cette problématique est plus complexe qu'il n'y paraît, les choses étant particulièrement délicates dans le cas de la forme TEN. Cette forme peut assumer la fonction pronominale ou, lorsqu'elle est utilisée de façon adnominale, celle d'un identificateur, à savoir d'un élément fonctionnel introduisant dans le discours un nom tout en renseignant l'interlocuteur sur la possibilité d'identifier son référent⁸⁰ ; en 3b, l'identification se fait par la voie anaphorique :

- 3) a) Face à un chien : **Ten** je roztomil-ý!
 celui.ci.M.NOM.SG être.PRS.3SG mignon-M.NOM.SG
Il est mignon, lui !
- b) Příběh-l k nám [toulav-ý pes].i.
 arriver en courant-PST.3SG.M vers nous.DAT errant-M.NOM.SG chien.NOM.SG
- [**Ten pes**].i byl opravdu roztomil-ý.
 ten.M.NOM.SG chien.NOM.SG être.PST.3SG.M vraiment mignon-M.NOM.SG
- Un chien errant est arrivé vers nous en courant. Ce chien était vraiment mignon.*

Quant à la forme neutre du même paradigme (*to*), celle-ci apparaît souvent, par exemple, dans les constructions présentatives (de même que le pronom démonstratif neutre *ce* en français) :

- 4) **To** je Petr.
 ce être.PRS.3SG NP.M.NOM.SG
C'est Pierre.

Cependant, il est un grand nombre d'énoncés où l'emploi de cette même forme ne peut être assimilé à celui qui apparaît en 4⁸¹ :

- 5) **To** je ale horko!
 ce être.PRS.3SG MOD chaud
Qu'est-ce qu'il fait chaud !

⁸⁰ Nous simplifions ici, laissant volontairement de côté les emplois non référentiels de TEN (voir notamment le Chapitre 6).

⁸¹ Voir à cet égard par exemple Trávníček (1951 : 1147).

La forme neutre *to* qui apparaît en 5 ne peut plus être rangée parmi les démonstratifs nominaux. Il s'agit en vérité d'une particule (cf. Loucká, 1990 : 74-75, utilisant le terme de « particule connective emphatique » („konektorová zřetelová částice“)). Du point de vue de la fonction syntaxique, *to* ne remplit pas la fonction sujet, car il s'agit d'une phrase sans sujet („věta jednočlenná“). Ainsi, le tchèque abonde d'emplois – plus ou moins similaires à celui en 5 – d'un *to* qui, en synchronie, n'a plus rien à voir avec le paradigme du démonstratif TEN⁸². L'écart entre les emplois en 3 et 4 et en 5 est aujourd'hui tel qu'il est sans doute acceptable de parler d'homonymie (Berger, 1993 : 27-35), même si, au vu du fait que les deux lexèmes restent liés en diachronie, le terme de polysémie ne saurait non plus être complètement écarté selon nous⁸³.

Néanmoins, rappelons que dans l'approche de Diessel (1999 : 27), même la forme pronominale neutre de TEN (*to*) doit être considérée à part lorsqu'elle est employée dans les emplois comme exemplifié en 4. Comme nous l'avons vu dans la Section 1.3.3 du Chapitre 1, Diessel se rend compte que le pronom démonstratif apparaissant dans ce type de constructions (“the identificational demonstrative”) se démarque souvent de la forme pronominale apparaissant dans d'autres contextes. Nous avons vu que c'est clairement le cas en français, où le *ce* des présentatifs est même complètement différent sur le plan morphologique des formes pronominales « libres » *ceci*, *cela* et *ça*. En tchèque, si la distinction ne va pas aussi loin, force est de reconnaître qu'alors que le pronom TEN en emploi « libre » s'accorde en genre et en nombre, il n'en va pas ainsi pour la forme neutre *to* des constructions présentatives, qui reste invariable :

- 6) a) **Ten** /*ta* /*to* se mi líb-í.
celui-ci.M.NOM.SG /celui-ci.F.NOM.SG /cela.NOM REFL moi.DAT plaître-PRS.3SG
J'aime bien celui-là/celle-là/ça.
- vs b) **To** je /jsou Petr /Mari-e /moj-i rodič-e.
ce être.PRS.3SG /3PL N.P.M.NOM.SG /NP-F.NOM.SG /mien-M.NOM.PL parent-NOM.PL
C'est/ce sont Pierre/Marie/mes parents.

La même opposition s'applique aux autres démonstratifs tchèques (TENTO, TENHLE, TAMTEN, etc.), dont les formes neutres connaissent aussi le même emploi présentatif. Même si Diessel

⁸² Un autre exemple typique serait le *to* apparaissant dans un type de construction à valeur rhématique qui, en français, trouve son équivalent dans le clivage ou dans la phrase complexe contenant une subordonnée relative attributive :

To Petr volal (a ne Pavel).	C'est Pierre qui a appelé (et non pas Paul).
Co se děje? To přišel Petr.	Qu'est-ce qui se passe? Il y a Pierre qui est arrivé.

⁸³ Malgré cela et comme nous l'avons écrit *supra*, les annotateurs du CNT ont fait le choix de subsumer tous ces emplois de *to* sous le lemme de TEN.

n'emploie nulle part le terme de polysémie, nous pensons qu'une conclusion en ces termes s'impose ici si nous suivons son argumentation jusqu'au bout.

Une autre position de départ pouvant déboucher sur une analyse en polysémie est celle, déjà mentionnée, de *Mluvnice češtiny 2*. Si, en effet, nous partons de l'apriori que tous les emplois pronominaux (« nominaux ») sont le fruit d'une transposition à partir d'emplois adnominaux (« adjectivaux »), rien ne nous empêche, à première vue, d'appréhender les premiers non seulement comme une fonction dérivée, mais clairement comme un sens dérivé du même lexème. Mais une telle conclusion ne nous convainc pas, car, d'abord, le primat des emplois adnominaux, postulé par les auteurs, nous semble arbitraire, ne s'appuyant, à notre connaissance, sur aucune preuve d'ordre étymologique, morphologique, syntaxique ou autre. Et quand bien même l'un des deux emplois aurait le primat sur l'autre, cela ne nous obligerait pas pour autant à conclure à l'existence d'une polysémie : il serait au moins tout aussi satisfaisant de parler d'une simple dualité d'emploi, sans plus⁸⁴. Ainsi, si polysémie il y a, celle-ci touche, d'après nous, exclusivement les emplois des formes neutres apparaissant dans les présentatifs (exemples 4 et 6b), éventuellement aussi tous les emplois où *to* n'est plus un pronom (exemple 5). Dans ce dernier cas, l'analyse en homonymie semble aussi justifiable.

Un dernier phénomène à ne pas négliger est le fait que dans certaines situations (certains régimes d'indexicalité), les deux emplois d'un lexème ne soient pas également sollicités ou acceptables, voire n'aient pas tout à fait la même valeur. Ainsi, TEN peut être employé de façon aussi bien adnominale que pronominale. Cependant, sa sémantique déictique semblerait plus affaiblie dans le premier des cas, car c'est là qu'il est en train d'évoluer vers un article défini (voir la Section 3.3.1 ainsi que le Chapitre 6)⁸⁵. Quant à TENTO, sa présence est depuis longtemps combattue par les sources normatives dans les contextes anaphoriques où il se substitue au pronom de la troisième personne (ON). Cet emploi, probablement un calque de l'emploi du démonstratif allemand DIESER dans le même contexte, est perçu comme étant en contradiction brutale avec la norme du tchèque („hrubě nečesky“ ; Trávníček, 1951 : 1132) :

⁸⁴ Grâce à Diessel (1999), nous avons effectivement pu voir dans la section précédente que le tchèque, contrairement par exemple au français et à l'anglais, appartient aux langues où les deux types d'emploi semblent faire apparaître, sur le plan morphosyntaxique, le même démonstratif de nature pronominale, seul ou apposé/juxtaposé au nom.

⁸⁵ Même si, dans ses emplois pronominaux, TEN ne fonctionne d'habitude pas non plus comme un démonstratif en deixis situationnelle. Dans les contextes anaphoriques, il semble également assez affaibli, dans la mesure où il peut souvent commuter avec l'anaphore zéro ou le pronom personnel ON (IL) ; son fonctionnement serait, à cet égard, assez proche de celui du démonstratif pronominal allemand DER :

CZ : *A ten druhý, toho se to nechytlo*. DE : *Und der zweite, den packte es nicht* (Berger, 1993 : 55). vs FR : *Et le second, lui, il n'a pas été emporté par ça*.

7) Co	kdyby	klavír-u	uje-l-y	noh-y
MOD	si.CONJ+AUX.COND.3PL	piano-DAT.SG	fuir-PTCP-PL.F	pied-NOM.PL
a	<i>tento</i>	mne	zavali-l?	
et	tento.M.NOM.SG	moi.ACC	encombrer-PTCP.SG.M	

Et si les pieds du piano glissaient et que celui-ci me tombait dessus ? (emprunté à Štícha, 2013 : 398)

Cependant, comme les emplois pronominaux ne font pas l'objet de ce travail, nous ne rentrerons pas ici dans les détails de ces asymétries.

3.2 La constitution du système en diachronie

Dans cette section, nous essaierons de retracer, d'une manière aussi succincte et synthétique que possible, les origines du système du tchèque actuel. Pour l'histoire plus ancienne, nous nous appuierons essentiellement sur les données fournies par Gebauer dans sa grammaire historique du tchèque (1929 et 1960 [1896]) et en partie aussi sur celles de la *Grammaire comparée des langues slaves* (1958) de Vaillant⁸⁶. Pour l'histoire plus récente, nous nous en tiendrons notamment à une étude de Bělič (1970) et en partie aussi au travail de Berger (1993). Une brève comparaison avec les autres langues slaves occidentales nous paraît nécessaire par ailleurs.

Gebauer (1929 : 230) écrit que le vieux tchèque possédait les démonstratifs suivants : JEN (*jb*), SEN (*sb*), TEN (*tb*) et ONEN (*on*)⁸⁷. Comme aujourd'hui, le système exprimait des oppositions de distance. Sur ce plan, la forme JEN aurait été neutre, dotée d'un « sens simplement démonstratif, sans d'autres nuances » (1929 : 230)⁸⁸. Plus loin (1929 : 231), Gebauer compare ce démonstratif au paradigme *is, ea, id* en latin ; il cite également de nombreux exemples où JEN sert d'antécédent à une proposition relative périphrastique et un exemple où il introduit un nom rendu défini par la présence d'une relative. Le fait le plus remarquable concernant cette forme est le suivant : bien que le nominatif ait disparu de la langue (la majorité des exemples cités par Gebauer proviennent de textes du 15^{ème} siècle), les autres cas ont survécu (avec certaines modifications) sous la forme des paradigmes des pronoms personnels et possessifs de la troisième personne, ainsi que de celui du pronom relatif JENŽ.

La forme SEN marquait la proximité du référent. Littéralement, Gebauer écrit qu'il s'agissait d'un démonstratif « [...] désignant l'objet présent ou le plus proche sur les plans local, temporel ou mental⁸⁹. » (1928 : 233) Cette forme aurait cependant disparu assez vite de la langue, car elle aurait

⁸⁶ Mentionnons ici également l'existence de grammaires historiques du tchèque plus modernes comme celle de Lamprecht, Bauer et Šlosar (1986 : 182-184). Cette dernière ne consacre cependant que trois pages aux démonstratifs. De précieux éléments sur la diachronie des langues slaves en général et du tchèque en particulier se trouvent également chez Corbett et Comrie (1993), Horálek (1962) ou chez Petr (1984).

⁸⁷ Les formes entre parenthèses sont à considérer comme des variantes des formes en majuscules.

⁸⁸ Nous traduisons: „význam prostě demonstrativní bez další nuance“.

⁸⁹ Nous traduisons : „[...] ukazující na předmět místem, časem nebo pomyslně přítomný nebo nejbližší.“

déjà été ressentie comme archaïsante au 15^{ème} siècle. Gebauer précise ensuite que lorsqu'elle continue à s'utiliser, c'est notamment à l'intérieur de locutions figées contenant des noms dotés d'un sens temporel (cf. *moment*, *temps*, *année*, les jours de la semaine, les saisons, les différents moments de la journée) ou locatif (*côté*, *monde*), éventuellement aussi là où elle entre dans une opposition binaire avec la forme ONEN, marquant l'éloignement : *sen svět* vs *onen svět* (*ce monde-ci* = *l'ici-bas* vs *ce monde-là* = *l'au-delà*). De nos jours, elle subsiste uniquement dans les adverbes déictiques temporels *dne-s* (*aujourd'hui*) et *leto-s* (*cette année*)⁹⁰.

Le sort de la forme TEN est particulièrement intéressant, et ce à plusieurs égards. Contrairement aux deux autres formes susmentionnées, c'est la seule à avoir été conservée jusqu'à nos jours. Cependant, son sens a traversé une évolution mouvementée. Dans un premier temps, TEN aurait désigné « [...] l'objet se trouvant en ma présence en face de moi (personne parlante) [...] » et son sens aurait (en partie) correspondu « [...] au sens de *iste* en latin. » Cette valeur dans les emplois situationnels aurait ensuite eu pour corollaire une valeur proche dans les emplois discursifs (anaphoriques), où la forme aurait renvoyé à ce/celui « dont on a parlé ou dont on est en train de parler »⁹¹. » (1928 : 237) Suite à la disparition des formes JEN et SEN, TEN aurait pris leur place. Dans le premier des cas, TEN remplace JEN en position de proforme pronominale antécédent d'une subordonnée relative périphrastique (*celui qui...*) et peut-être aussi – quoique Gebauer ne donne ici aucun exemple concret – en position adnominale où JEN introduisait un nom rendu défini par la présence d'une relative, de même qu'à l'époque du latin tardif, ILLE remplace, dans ces positions, la forme IS disparue (cf. Orlandini 1981 ; Carlier et De Mulder, 2007). Dans le second des cas, Gebauer écrit que

[l]e pronom *ten* prend la place de *sb*, *sen*, en voie de disparition ; c'est là qu'il efface son propre sens pour prendre le sens de ce dernier, désignant un objet présent. C'est là que la forme renforcée *ten + to* est particulièrement en vogue⁹². (1928 : 237)

C'est ainsi que TEN, à l'origine une forme, semble-t-il, médiane⁹³ (ou identifiable avec l'espace de l'interlocuteur), en vient également à combler le vide créé par la sortie de l'usage de SEN. Mais

⁹⁰ Le paradigme HIC du latin classique a connu un sort semblable : lat. **hoc die** (*en ce jours*) → *hodie* → an. fr. *hui* → fr. moderne *aujourd'hui*.

⁹¹ Nous traduisons : „Zájmeno toto ukazuje na předmět, který jest v přítomnosti naproti mně (mluvícímu); shoduje se s významem lat. *iste*. S významem tímto souvisí jiný: *ten* = o němž byla n. je řeč.“

⁹² Ibidem : „Zájmeno *ten* vstupuje na místo hynoucího *sb*, *sen*; tu stírá význam svůj, přejímá význam jeho a ukazuje na předmět přítomný. Tu pak zvláště jest oblíben tvar sesílený *ten + to*.“

⁹³ « Médiane » plutôt que « neutre », car pour nous, tout semble indiquer que la distinction entre « formes médianes » et « formes neutres » procèderait de la configuration du système. Ainsi, la forme dite « médiane » serait celle désignant la sphère de l'interlocuteur dans les systèmes orientés par rapport à la personne, tandis que la forme dite « neutre » serait, en principe, celle parmi les formes des systèmes orientés par rapport à la distance en laquelle cette opposition est partiellement ou complètement neutralisée. Dans le tchèque contemporain, par conséquent, il est plus juste, d'après nous, de désigner TEN comme « forme neutre », car ne faisant plus partie d'un système orienté par rapport à la personne et

aussitôt installé dans ce nouveau rôle, TEN se renforce par l'ajout du morphème invariable *-to*. En l'occurrence, ce renforcement a eu lieu assez tôt, car le plus ancien exemple d'emploi de TENTO cité par Gebauer provient d'un manuscrit datant du 14^{ème} siècle (*na tompo svěťě*, soit *dans ce monde-ci, ici-bas*). Pour ce qui est ensuite de la forme non-renforcée, il n'est pas moins remarquable que l'auteur cite des occurrences du TEN adnominal (dont la plus ancienne remonte au 15^{ème} siècle) où la forme apparaît déjà a) dans des contextes où elle se rapproche d'un article défini et b) dans des contextes où elle est porteuse d'une valeur « émotionnelle ». Pour ce qui est du premier phénomène, il s'agit de TEN introduisant un nom modifié par un adjectif au superlatif ainsi que par des compléments prépositionnels établissant son unicité dénotative (voir notamment la Section 6.2.2 du Chapitre 6). Concernant le second phénomène, il s'agit de TEN introduisant un nom employé de façon générique (voir *infra*). Les deux phénomènes s'inscrivent au sein d'une catégorie assez large, regroupant à la fois emplois adnominaux et pronominaux de TEN, et pour laquelle Gebauer utilise le terme de « significations emphatiques » („významy vytýkací“ ; 1929 : 238). Il est légitime de présumer que c'est également à ce moment-là que la forme simple TEN, non renforcée par *-to*, commence à perdre sa sémantique déictique pour entamer sa mutation en article défini. En effet, cette mutation est enclenchée très tôt, comme l'écrivent par exemple Machač et Zíková (2014 : 99) :

[...] it seems reasonable to argue that the Czech demonstrative *ten* is on its way to becoming a definite article – **indeed, on a way which stretches over hundreds of years**⁹⁴.

Le renforcement morphologique de TEN par *-to* se limite au tchèque et au slovaque (cf. Bělič, 1970), et s'avère, pour cette raison, particulièrement important dans une optique de comparaison de ces deux langues avec d'autres langues slaves occidentales telles que le polonais et le sorabe. Le polonais actuel possède un système à deux membres, à savoir TEN et TAMTEN. Alors que TAMTEN y est doté de la même valeur qu'en tchèque, le statut de TEN est différent. Comme l'écrit Czardybon, cette forme est également en voie de grammaticalisation vers un article défini. Cependant, dans les emplois situationnels, elle connaît deux réalisations : une non-accentuée et une accentuée. Dans le premier des cas, TEN se comporte comme un proto-article, tout à fait comme en tchèque ; dans le second cas, il garde toutefois son statut de démonstratif⁹⁵, permettant de contraster le référent à d'autres, plus lointains et identifiés par TAMTEN :

apparaissant dans des emplois typiques, dans les langues du monde, pour les formes à la fois proximales (emplois endophoriques) et distales (emplois mémoriels et déterminatifs ; voir la Section 3.3 *infra*).

⁹⁴ Nous soulignons.

⁹⁵ Selon Vaillant (1958 : 390), TEN pourrait, dans cet emploi, être facultativement renforcé par l'adverbe *tu* : *ten tu* ; cette possibilité n'est pas mentionnée par Czardybon. Remarquons que cette situation correspond exactement à la situation des dialectes de Moravie de l'Est et de Silésie, décrite dans la Section 3.1.1 *supra*, où le contexte situationnel est structuré de la même façon : TEN (accentué) pour référer aux objets non éloignés (plutôt que proches) vs TAMTEN pour référer aux objets éloignés. Cela semble s'expliquer par une continuité géographique, dans la mesure où les régions concernées par ce phénomène sont limitrophes de la Pologne (voir les cartes des Figures 3.2 et 3.3 *supra*).

One has to distinguish between two types of deictic contexts which differ in the occurrence of the determiner *ten* in Polish. There are deictic contexts in which more than one potential referent can fit the definite NP. In such a context, the determiner *ten* cannot be omitted and has to be stressed. Furthermore, a pointing gesture is required in order to achieve unique reference. [...] The second type of deictic contexts is characterized by the fact that there is accidentally only one referent in the context that fits the definite NP. In such a context, the determiner *ten* is optional and unstressed. Moreover, a pointing gesture is not required to achieve a unique reference because unique reference is already created by the context itself. (Czardybon, 2017 : 55-56)

En haut sorabe, l'on retrouve également une opposition à deux membres : TÓN et TÓNE. Le premier est neutre, le second marque la proximité. De même qu'en polonais, TÓN possède un double statut, allant de pair avec deux réalisations phonétiques possibles (cf. Czardybon, 2017 : 84 ; Nédolužko, 2004).

Quant à la forme ONEN, elle est utilisée, dans le vieux tchèque, pour désigner un référent lointain ou absent dans la situation d'interlocution⁹⁶. Dès le début, il tend également à introduire des noms dotés d'un sens temporel pour renvoyer à des époques reculées du passé ou des noms dotés d'un sens locatif pour renvoyer à des endroits autres que la situation d'énonciation (et donc pas visibles). Aujourd'hui, cet usage subsiste, dans une certaine mesure, dans la langue soutenue (archaïsante) à la fois dans les régimes anaphorique et exophorique mémoriel : *za oněch časů* (au temps jadis ; lit. dans ces temps-là), *onen svět* (l'au-delà ; lit. ce monde-là). Comme nous l'avons déjà écrit *supra*, ONEN est relativement vite remplacé par TAMTEN, notamment dans les emplois situationnels, dont il semble avoir complètement disparu aujourd'hui (voir aussi la Section 3.1.1). Néanmoins, sa variante plus courte ON a produit le nominatif des pronoms personnels de la troisième personne :

	Masculin	Féminin	Neutre
Singulier	<i>on</i>	<i>ona</i>	<i>ono</i>
Pluriel	<i>oni</i>	<i>ony</i>	<i>ona</i>

Tableau 3.14 : Formes nominatives des pronoms personnels tchèques de la 3^{ème} personne

Ainsi, le paradigme de ces pronoms est un produit syncrétique, issu de la fusion des paradigmes de JEN et de ON, celui-ci fournissant le cas nominatif et celui-là, les autres cas.

Dans la description de l'ancien système faite par Gebauer, un aspect attire nécessairement l'attention, à savoir les parallèles que l'auteur dresse entre ce système et le système du latin classique.

⁹⁶ Lamprecht, Bauer et Šlosar (1986 : 183) écrivent que ONEN, ne connaissant au début que la forme neutre ONO, était à l'origine une forme de présentatif équivalente au latin *ecce*. Ce cas de figure invaliderait ainsi l'hypothèse de Diessel (1999 : 152), selon laquelle les démonstratifs n'évoluent jamais à partir de mots grammaticaux (dont ONEN fait partie de fait, en tant que présentatif/particule).

En effet, non seulement TEN est comparé tantôt à ISTE, tantôt à IS, mais SEN et ONEN sont comparés, respectivement, à HIC et à ILLE. En traitant du sens de ONEN, Gebauer écrit :

Il sied ensuite d'attirer l'attention sur la signification de ce pronom, qui est [celle de] *ille*, par opposition à *ten* = *is* et *sen* = *hic*⁹⁷. (1960 : 451)

Inévitablement, la question qui se pose est ainsi celle de la nature du système du vieux tchèque. Car si ce parallélisme avec le système du latin va au-delà d'une simple comparaison jugée opportune pour la compréhension du propos à une époque où le latin continue à représenter une référence importante chez les linguistes⁹⁸, il faudrait en conclure que le vieux tchèque, contrairement au tchèque actuel, possédait un système de démonstratif nominal fondé sur la notion de personne. Or cette conclusion est loin de paraître incongrue à condition de remonter plus loin dans l'histoire des langues slaves. Vaillant (1958 : 379) explique que le vieux slave présentait un système à trois formes, à savoir : « [...] *sī* “celui-ci” désignant l'objet tout proche, *tŭ* “ce” pour l'objet proche plus vaguement localisé, *onŭ* “celui-là” pour l'objet éloigné. » Aussitôt après, il précise que « [l]e rapport des trois démonstratifs avec les trois personnes des pronoms personnels est net en slave comme en latin et en grec. » (1958 : 379)

Pour ce qui concerne l'évolution ultérieure du système, elle est notamment marquée par l'apparition de la forme TENHLE, attestée à partir du 16^{ème} siècle ; dès son apparition, TENHLE tend à remplacer progressivement TENTO sur le territoire de la Bohême, contrairement à la Slovaquie et le reste des pays tchèques, où TENTO maintient sa position forte jusqu'à aujourd'hui (Bělič, 1970 : 234 ; Jodas, 2010 : 90-91). C'est également plus tard qu'émergent, en Bohême, les autres formes renforcées et redupliquées de la langue parlée non-soutenue (TENHLETEN, TADYTEN, TADYHLETEN, TAMHLETEN, etc.), se spécialisant, dès le début, dans les usages situationnels. En « tchèque commun », la popularité de certaines de ces formes – notamment TENHLETEN et TAMHLETEN – irait croissant selon Berger, si bien que celui-ci se demande (déjà au début des années quatre-vingt-dix du siècle dernier) si elles ne pourraient pas finir par remplacer un jour TENHLE et TAMTEN dans le régime de l'exophore situationnelle :

In der gesprochenen Sprache könnte dieser Entwicklung in einem System resultieren, in dem neben dem universalen und daher zur Verwendung als Deiktikum nicht geeigneten Pronomen *ten* ein Nahdeiktikum *tenhleten* und ein Ferndeiktikum *tamhleten* stehen. (1993 : 510)

À cet égard, le bilan fourni par le corpus ORAL v1 incite toutefois à la prudence, notamment pour TAMHLETEN, dont la fréquence relative est seulement de 31,8 contre celle de TAMTEN, qui est égale à

⁹⁷ Ibidem : „Dále jest upozorniti na význam zájmena tohoto, kterýž jest = *ille*, proti *ten* = *is* a *sen* = *hic*.“

⁹⁸ À la fin du siècle précédent, J. Dobrovský, fondateur des études slaves et l'un des prédécesseurs de Gebauer, publie encore une partie de ses travaux en latin.

129,2⁹⁹. S'agissant de TENHLETEN, le clivage reste également important : imp = 755,4 contre 1 244,6 pour TENHLE¹⁰⁰. Le bilan est légèrement plus encourageant à l'égard des formes longues si la recherche est limitée à la seule Bohême : imp/TENHLE = 1 475,7 ; imp/TENHLETEN : 942,9¹⁰¹ ; imp/TAMTEN : 153 et imp/TAMHLETEN : 42,4¹⁰².

3.3 La distribution des démonstratifs adnominaux tchèques dans les différents régimes d'indexicalité

Comme cela a déjà été évoqué, le classement des démonstratifs nominaux tchèques proposé par Štícha (1999 et 2013) se base sur une distinction première, onomasiologique, entre emplois endophoriques et emplois exophoriques situationnels. En vérité, cette voie d'approche n'a rien de nouveau : adoptée par des sources aussi anciennes que *Průruční slovník jazyka českého* (1935-1957), *Mluvnice spisovné češtiny II* (1951) ou *Česká mluvnice* (1960), elle remonte déjà à Mathesius, qui, dans son article de 1926, introduit l'opposition entre la « deixis intérieure » („deixe/ukazování vnitřní“) et la « deixis extérieure » („deixe/ukazování vnější“). L'approche est également reproduite dans d'autres sources parues plus récemment, telle *Mluvnice současné češtiny I* (2015). Mais au-delà de cet accord terminologique assez bien établi dans la littérature tchèque, il est intéressant de regarder la distribution des différentes formes du démonstratif adnominal au sein des différents domaines d'indexicalité. Ainsi, dans un premier temps, nous nous pencherons sur les emplois exophoriques a-mémoriels (situationnels, surtout, mais aussi temporels et textuels) et endophoriques. Dans un second temps, il nous semble nécessaire d'élargir cette distinction première par l'inclusion des emplois exophoriques mémoriels. Nous précisons qu'il ne s'agit ici que d'une vue d'ensemble : une bonne partie des phénomènes passés en revue – notamment ceux où TEN possède un potentiel de grammaticalisation – seront ensuite étudiés en détail dans les chapitres à venir, et ce à la fois sur une base quantitative et qualitative du matériel empirique se trouvant à notre disposition. Eu égard à l'orientation générale du présent travail, le lecteur ne s'étonnera pas du fait que la plus grande attention soit accordée à TEN.

3.3.1 Les emplois exophoriques situationnels, temporels et textuels

Regardons les exemples suivants, contenant TEN adnominal en usage exophorique situationnel :

⁹⁹ $L = 402,5827$ pour la comparaison des fréquences absolues des deux lemmes au sein de la même population n (« 2 words in 1 corpus ») → la différence est statistiquement significative ; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

¹⁰⁰ $L = 769,6619$ → la différence est statistiquement significative.

¹⁰¹ $L = 534,4317$ → la différence est statistiquement significative.

¹⁰² $L = 300,1323$ → la différence est statistiquement significative.

8) Pod-ej mi prosim **ty** **modr-é** **desk-y**.
 passer-IMP.2SG moi.DAT s'il te plaît ten.F.ACC.PL bleu-F.ACC.PL pochette-ACC.PL tantum
*Passe-moi **la pochette bleue**, s'il te plaît.*

9) L1 : no jo já to m-á-m proti východ-u
 MOD moi.NOM cela.ACC avoir-PRS-1SG contre Est-DAT.SG

jenže když je tam na **tom** **stolečk-u** u
 mais quand être.PRS.3G là-bas sur ten.M.LOC.SG petite table-LOC.SG auprès de

toho **radiator-u** ..
 ten.M.GEN.SG radiateur-GEN.SG

L2 : tak oni ty orchidej-e ne-m-a-jí takov-ý
 MOD eux.NOM ten.F.NOM.PL orchidée-NOM.PL NEG-avoir-PRS-3PL tel-N.ACC.SG

ne ne-m-a-jí rád-y takov-ej .. žár .. víš ?
 non NEG-avoir-PRS-3PL content-F.NOM.PL tel-M.ACC.SG chaleur.ACC.SG MOD

*L1 : bon, je l'ai placée contre l'Est, mais quand elle est là, sur **la petite table** près **du radiateur**...*

L2 : bon, les orchidées n'aiment pas... je veux dire elles n'aiment pas autant de chaleur, tu vois ?

ORTOFON v1

Ce que nous pouvons directement observer dans les exemples 8 et 9, c'est un affaiblissement de la sémantique déictique de TEN : vu que la situation contient, respectivement, une seule pochette bleue (exemple 8) et une petite table et un radiateur (exemple 9), lors de la traduction des phrases en français, nous avons été obligé de remplacer TEN par l'article défini. De plus, la présence de TEN semble plus ou moins obligatoire dans ces contextes (voir la Section 6.2.1.1.). Au vu de ces faits, Berger (1993 : 463) conclut que TEN adnominal a cessé de fonctionner comme un démonstratif en emploi exophorique situationnel, se rapprochant, dans ce régime, de l'article défini des langues qui en disposent¹⁰³. Tel est également le constat d'autres chercheurs comme Adamec (1983), Hlavsa (1975), Mathesius (1926), Vey (1946) et Zíková (2017). Mathesius est, à notre connaissance, le premier à avoir relevé ce fait :

Tout d'abord, force est de souligner que dans le tchèque parlé, le pronom épithète *ten* n'a plus la capacité à pointer de manière autonome dans une situation extérieure. Dans cette fonction, il a été complètement évincé par les formes composées (*tenhle ten*, *tuhle ten*, *tadyhle ten*, y compris, dans les dialectes, *tu ten*, *tam ten*, *tamhle ten*)¹⁰⁴. (1926 : 39)

¹⁰³ Berger écrit littéralement : „Wenn wir *ten* außer Acht lassen, das, wie ich gezeigt habe, nie wirklich deiktische Funktion hat [...]“ Ici, nous croyons légitime de rectifier légèrement cette formulation en nous appuyant sur la précieuse distinction, introduite par Kleiber (1991b), entre le sens et les emplois des expressions (voir la Section 1.1.4.). Ainsi, ce n'est pas qu'en contexte de deixis situationnelle, TEN ne soit pas doté d'une *fonction* déictique, mais qu'il a perdu son *sens* déictique, dans la mesure où il ne se comporte plus comme un démonstratif mais davantage comme un article défini (qui, comme nous le savons, connaît lui aussi des emplois déictiques).

¹⁰⁴ Nous traduisons : „Předně je důležité vytknouti, že v hovorové češtině přívlastkové zájmeno *ten* nemá už schopnosti samostatně ukazovati ve vnější situaci. V tom úkolu je zcela zatlačily tvary složené (*tenhle*, *tenhle ten*, *tuhle ten*, *tadyhle*

La question se pose ensuite de savoir dans quelle mesure l'on peut parler ici d'un véritable article défini, question autour de laquelle les avis des chercheurs divergent. Dans tous les cas, cette thèse se voit corroborée par le fait qu'en Bohême, TEN adnominal ne serait jamais accentué en emploi situationnel¹⁰⁵ ainsi que par l'existence des formes renforcées TENTO, TENHLE, TENHLETEN, TADYTEN, TADYHLETEN, etc. Ces formes, elles, se prêtent facilement à l'accentuation, appellent explicitement un geste ostensif, et sont ainsi aptes à contraster le référent en question avec d'autres référents de la même classe. Nous pouvons observer cela dans l'exemple suivant :

10) no MOD	<i>tahle</i> tenhle.F.NOM.SG	<i>sukn-ě</i> jupe-NOM.SG	už déjà	mi moi.DAT
je être.PRS.3SG	ted'ka maintenant	velk-á grande-F.NOM.SG		
<i>tu vois, maintenant, cette jupe-là, elle est déjà trop grande pour moi</i>				
ORTOFON v1				

Comme la locutrice parle d'une jupe particulière qu'elle compare (implicitement) à d'autres jupes, elle emploie la forme TENHLE, plus étoffée sur les plans morphologique et phonétique. Le fait que TEN ne se comporte plus comme un démonstratif dans les contextes situationnels en Bohême a évidemment des conséquences importantes sur la configuration de l'ensemble du système au sein de ce régime d'indexicalité. En effet, d'un système à trois membres et distinguant entre trois degrés de distance (TENTO/TENHLE vs TEN vs ONEN/TAMTEN, la langue est passée à un système à deux membres, la nouvelle opposition se voyant réduite à deux valeurs seulement, à savoir la proximité et l'éloignement. De plus, autant la position neutre serait perdue, autant la langue orale informelle multiplie les formes marquant la proximité et l'éloignement (voir la Section 3.2., consacrée à l'évolution du système en diachronie).

Concernant la deixis temporelle, TEN semble exclu de ce régime, hypothèse que les données de nos corpus valident¹⁰⁶. Par conséquent, seules les formes renforcées (notamment TENTO/TENHLE) se prêteraient à cet usage :

ten, dial. i *tu ten*, *tam ten*, *tamhle ten*).“ Remarquons, chez Mathesius, la présence des graphies séparées des formes composées du démonstratif, conformes à une norme orthographique plus ancienne.

¹⁰⁵ Contrairement aux dialectes moraves et silésiens où il a gardé sa capacité à référer à un objet plus proche que TAMTEN (Jodas, communication personnelle).

¹⁰⁶ Hormis nos échantillons, nous avons recherché la combinaison d'un démonstratif défini (quel qu'il soit) suivi du lemme TÝDEN (SEMAINE) dans ORAL v1. Dans aucune des occurrences trouvées, TEN ne peut être dit explicitement doté d'une valeur déictique temporelle, mis à part deux cas, où le contexte ne permet pas de trancher.

- 11) *perníčk-y* *se* *mus-í* *už* *začít* *píct* *asi*
 pain d'épices-NOM.PL REFL devoir-PRS.3PL déjà commencer.INF cuire.INF MOD
- tenhle* *tejden ..*
tenhle.M.ACC.SG *semaine.ACC.SG*

je crois que c'est cette semaine qu'il faut commencer à préparer des petits pains d'épices...
 ORTOFON v1

Dans l'exemple 11, la lecture déictique temporelle semble devenir impossible si nous remplaçons *tenhle* par *ten*¹⁰⁷. Trávníček (1951 : 1153) écrit : « En tant qu'épithète des noms porteurs d'un sens temporel, *tento* signifie la même chose que *actuel*, *le plus proche dans le temps* [...]»¹⁰⁸. Les exemples que l'auteur cite sont *tuto středu* (*ce mercredi*) et *tento týden* (*cette semaine*). Il faut ici insister sur le fait qu'une expression démonstrative possède une référence déictique temporelle uniquement lorsqu'aucun antécédent ne peut être identifié dans le discours antérieur. De ce fait, des exemples comme 12 (ou TEN est lui aussi communément admis) ne relèvent pas de la deixis (de l'exophore), mais de l'anaphore. En 12, la locutrice parle de ses problèmes avec le genou ; le complexe nominal *ta doba* (*ce temps-là*) reprend la référence de l'adverbe temporel *tenkrát* (*autrefois*, *à l'époque*)¹⁰⁹ :

- 12) *mně* *fakt* [*tenkrát*]_i *pomoh-l-i [...]* *to* *napíchán-í*
 moi.DAT vraiment à l'époque aider-PST-3PL.M ten.N.NOM.SG piqûre-NOM.SG
- té* *kyselin-y* *hyaluronov-é*
ten.F.GEN.SG acide-GEN.SG hyaluronique-F.GEN.SG
- ...
takže fakt *od* [*tý* *dob-y*]_i *m-á-m* *pokoj*
 MOD vraiment depuis ten.F.GEN.SG époque-GEN.SG avoir-PRS-1SG calme.ACC.SG
- ce qui m'a vraiment aidé [à l'époque]_i, c'était les piqûres d'acide hyaluronique*
mais vraiment, depuis [ce temps-là]_i, je suis tranquille
 ORTOFON v1

S'agissant de la deixis textuelle, ce régime admet à la fois TENTO/TENHLE et TEN :

- 13) *Tato* *prác-e* *se* *zabýv-á* *demonstrativ-y*.
tento.F.NOM.SG travail-NOM.SG REFL traiter-PRS.3SG démonstratif-INS.PL
- Ce travail traite des démonstratifs.*

Voici pourtant l'unique exemple de TEN employé de cette manière que nous avons identifié dans nos corpus. À notre avis, il est ici légitime de parler de deixis textuelle, car le complexe nominal *ten*

¹⁰⁷ Un exemple semblable a été soumis à l'évaluation de nos informateurs qui portait sur la météo au moment de l'énonciation (le contexte avait été précisé). La variante avec TEN a été rejetée à l'unanimité.

¹⁰⁸ Nous traduisons : „Jako přívlastek substantiv významu časového znamená *tento* tolik co *nynější, časově nejbližší*; na př.: *tuto středu [...]* *tento týden [...]*“

¹⁰⁹ Remarquons que l'adverbe *tenkrát* contient lui aussi *ten*. Comme nous écrivions plus haut, afin de marquer qu'il s'agit d'un moment ancré dans un passé lointain et situé loin du moment de l'énonciation, la langue (très) soutenue peut aussi employer, dans ce type d'anaphore, la forme ONEN ; par ex. *za oněch časů* (*au temps jadis* ; lit. *dans ces temps-là*).

tón (*le/ce ton*) renvoie au ton sur lequel des propos ont été dits par une parente de la locutrice. La locutrice use de cette expression après avoir elle-même reproduit le ton en question, qu'elle juge inapproprié. Le langage est ainsi utilisé pour référer à lui-même :

- 14) *ten* *tón* ano . to je *ten* *tón* ..
ten.M.NOM.SG ton.NOM.SG oui ce être.PRS.3SG ten.M.NOM.SG ton.NOM.SG
[c'est] le ton [sur lequel elle dit ça], oui, c'est le ton... ORTOFON v1

3.3.2 Les emplois anaphoriques

Voici maintenant deux exemples de TEN adnominal en emploi anaphorique (anaphore coréférentielle fidèle) :

- 15) si [na štět-ec]_i vezm-e-š barv-u ..
REFL sur pinceau-ACC.SG prendre-PRS-2SG peinture-ACC.SG

[ten *štět-ec]*_i co nej-víc votř-e-š
ten.M.ACC.SG pinceau-ACC.SG autant SUP-plus essuyer-PRS-2SG

vo papírov-ej ubrousek ..
dans de papier-M.ACC.SG serviette.ACC.SG

tu mets de la peinture sur [un pinceau]_i... puis tu essuies [ce/le pinceau]_i autant que tu peux dans une serviette en papier... ORTOFON v1

- 16) Koupi-l jsem si [aut-o]_i a [motork-u]_{ii}
acheter-PTCP.SG.M AUX.PST.1SG REFL voiture-ACC.SG and moto-ACC.SG

[(To) *aut-o]*_i jezd-í výborně, ale s *[(tou)*
ten.N.NOM.SG voiture-NOM.SG rouler-PRS.3SG parfaitement mais avec ten.F.INS.SG

*motork-ou]*_{ii} jsou pořád problém-y.
moto-INS.SG être.PRS.3PL toujours problème-NOM.PL

J'ai acheté [une voiture]_i et [une moto]_{ii}. [La voiture]_i roule parfaitement, mais il y a tout le temps des soucis avec [la moto]_{ii}. (exemple inspiré par Pešek, 2014 : 156)

Ce qui est certain, c'est que dans le régime anaphorique, TEN est déjà moins proche d'un article défini que dans le régime exophorique situationnel. Quant à 15, si sa transposition par le défini français nous semble acceptable, TEN peut aussi y être rendu par le démonstratif CE. En revanche, en 16, seul le défini est acceptable en français, car l'énoncé met en place un contraste externe entre deux référents dont chacun a une « référence virtuelle » différente (cf. Corblin, 1987 ; Milner, 1976)¹¹⁰.

¹¹⁰ Après que cet exemple eut été soumis au jugement de nos informateurs, qui devaient choisir entre une version où les noms *auto* et *motorka* étaient employés nus et celle où ils étaient précédés de TEN, la version avec TEN a été choisie par 4 personnes sur 10, alors que 5 personnes ont choisi la version sans TEN. Une personne nous a signalé qu'elle emploierait TEN uniquement avec le nom *motorka*. Ce choix intéressant s'explique certainement par une volonté de mettre l'accent sur ce référent en tant que celui-ci pose problème dans les faits : l'on voit ainsi que le choix d'employer TEN peut, à ce

D'une part, des études (cf. Berger, 1993 ; Czardybon, 2017 ; Křížková 1971 ; Nědolužko, 2004 ; Štícha, 1999 ; Souček, 2011 ; Zíková, 2017 et 2018) montrent que la fréquence élevée avec laquelle TEN adnominal apparaît avec l'anaphore textuelle – notamment dans les productions orales du registre informel – l'approche du défini. D'autre part, comme nous le verrons, ce constat est aussitôt contré par un autre, à savoir que TEN est loin d'être toujours obligatoire dans ces contextes.

Un nombre non-négligeable de travaux (cf. Pešek, 2014 ; Štícha, 2001 ; Trávníček, 1951 ; Uhlířová, 2006 ; Voráč, 1946 ; Zimová, 1994) s'intéressent à la possibilité de commutation et à la concurrence, dans le régime de l'anaphore textuelle coréférentielle, entre a) un nom anaphorique nu, b) un nom introduit par TEN et b) un nom introduit par TENTO/TENHLE, éventuellement par d'autres formes du démonstratif. Outre le fait que cette problématique est extrêmement complexe, et que personne n'a, jusqu'à présent, réussi à en proposer un modèle explicatif satisfaisant, deux constats importants se dégagent de tous ces travaux : 1) dotées d'une sémantique indexicale plus forte, les formes TENTO/TENHLE sont davantage valorisées dans le cas de l'anaphore infidèle, où il est plus important d'insister sur le lien de coréférence entre deux expressions lexicalement différentes et 2) du fait de cette indexation plus forte du référent tout comme de la sémantique proximale des formes, des effets peuvent être produits qui n'accompagneraient pas, dans la même situation, l'emploi de TEN. Ainsi, selon Štícha (2001 : 92), moyennant l'usage de TENTO/TENHLE, la réalité nommée par l'antécédent est, au sein du nouvel énoncé, identifiée et accentuée comme le thème explicite de celui-ci, thème au sujet duquel un nouveau jugement sera aussitôt formulé¹¹¹. Pešek tente d'unir les deux constats susmentionnés en un seul en partant de l'analyse de l'exemple suivant :

17) Z cest	po jižn-í	Americ-e	si	přivez-l	
de voyage.GEN.PL	après du sud-F.LOC.SG	Amérique-LOC.SG	REFL	ramener-PST.3SG.M	
[dv-ě	žáb-y	rod-u		dendrobates] _i	
deux-F.ACC.PL	grenouille-ACC.PL	genre-GEN.SG		dendrobates.NOM.SG	
[<i>Tito</i>	<i>obojživelníc-í</i>] _i	m-a-jí	ve	sv-ých	žláz-ách
tento.M.NOM.PL	batracien.NOM.PL	avoir-PRS-3PL	dans	leur.F.LOC.PL	gland-LOC.PL
smrteln-ý	jed	kurare.			
mortel-M.ACC.SG	poison.ACC.SG	curare.ACC.SG			

De ses voyages en Amérique du Sud, il a ramené [deux grenouilles du genre dendrobates]_i. Les glands de [ces batraciens]_i contiennent le poison mortel curare. (Pešek, 2014 : 159)

stade de sa grammaticalisation, constituer une voie d'expression de la subjectivité du locuteur (voir à cet égard la Section 6.1.3.2 du Chapitre 6).

¹¹¹ „[Identifikátorem *tento/tenhle* se] skutečnost pojmenovaná antecedentem v nové výpovědi aktuálně vytýká jako její zdůrazněné a vytčené téma pro nový soud.“

Pešek explique que TENTO constitue ici le premier choix au sein de l'expression anaphorique, par laquelle le référent est réévalué comme appartenant à la classe des batraciens¹¹². Cette réévaluation (recatégorisation) représente une information nouvelle, un jugement nouveau formulé au sujet du référent. Littéralement, Pešek écrit : « Le SN fonctionne ainsi comme une prédication condensée du type “les grenouilles du genre dendrobates sont des batraciens”¹¹³. » c'est, entre autres, grâce à cette réévaluation du référent¹¹⁴ que le complexe nominal est, au sein du nouvel énoncé, perçu comme porteur d'un thème accentué et appelant un jugement nouveau.

Mais il y a aussi des cas où la différence entre TENTO/TENHLE et TEN est minimisée. Tel est aussi l'avis de Zimová (1994 : 33), qui fait aussitôt remarquer que dans certaines situations, le choix de l'une et ou de l'autre forme reflète une intention du locuteur/auteur de représenter le référent comme davantage proche (présent) ou lointain (absent). C'est TENTO/TENHLE qui, on l'aura compris, apporterait cette connotation de proximité, du fait de servir également de démonstratif de proximité en deixis situationnelle. Toutefois, Štícha et Pešek présentent une analyse différente des choses. Ces deux auteurs mettent au contraire l'accent sur le fait que c'est TEN qui évoquerait une certaine familiarité/complicité, d'où sa présence moins élevée dans les textes soutenus, où cet effet n'est pas désirable. À cet égard, Štícha (2001) remarque que l'usage de TEN entraîne la présence d'un sentiment de « rétrospection partagée » („sdílená retrospekce“) concernant le référent :

18) Jan	si	koupi-l	[nov-é	aut-o] _i
NP.M.NOM.SG	REFL	acheter-PST.3SG.M	nouveau-N.ACC.SG	voiture-ACC.SG
Já	bych	ale do [tý	popelnic-e]_i	
moi.NOM	AUX.COND.1SG	mais dans ten.F.GEN.SG	poubelle-GEN.SG	
nikdy	ne-vlez.			
jamais	NEG-rentre.PTCP.SG.M			

Jan s'est acheté [une nouvelle voiture]_i. Mais moi, je ne montera jamais [dans cette poubelle]_i.
(Pešek, 2014 : 153 et 161)

Ici, la présence de TEN permettrait, selon l'auteur de l'étude, la mise en place d'une réévaluation du référent (la voiture) partagée entre le locuteur et l'interlocuteur, d'une complicité du type « moi, j'appelle sa voiture une poubelle et toi, tu partages mon point de vue » (2014 : 161)¹¹⁵. Bien

¹¹² Nous avons soumis cet exemple au jugement de nos informateurs, à qui nous avons demandé de choisir entre trois versions : celle où *oboživelníci* est un nom nu, celle où il est introduit par TEN et celle où il est introduit par TENTO. Neuf parmi les dix personnes sollicitées ont choisi TENTO et une seule a choisi TEN, ce qui conforte le sentiment de Pešek et le nôtre propre.

¹¹³ Nous traduisons : „NS tak funguje jako kondenzovaná predikace typu „žáby rodu dendrobates jsou oboživelníci“ “.

¹¹⁴ Cette capacité qu'a le démonstratif de recatégoriser le référent est bien connue (voir la Section 1.2.2 du Chapitre 1) ; répétons que selon Kleiber (1984a : 65), toute description démonstrative « [...] doit être analysée comme étant l'abréviation d'une structure attributive classificatoire [...] » du type *c'est un x*.

¹¹⁵ Nous traduisons : „já jeho auto říkám popelnice a ty tuto mou perspektivu sdílíš“.

évidemment, avec TENTO/TENHLE, l'opération de réévaluation/recatégorisation du référent est aussi présente, mais sans cette teinte de rétrospection partagée.

Les exemples 17 et 18 permettent de rapprocher TENTO/TENHLE du CE anaphorique en français dans la mesure où ce dernier est, lui aussi, souvent nécessaire pour maintenir la coréférence dans les cas d'anaphore infidèle. En revanche, la situation du tchèque est plus complexe, car (sans tenir compte des autres moyens de l'anaphore textuelle), il dispose de trois possibilités systémiques (Ø/TEN/TENTO) là où le français n'en garde que deux (LE/CE). Le statut de TEN est ambigu : d'une part, il peut (comme en 16), voire doit (comme en 18) être rendu par CE ; d'autre part, nous verrons plus tard qu'il est souvent plus naturellement rendu par LE. Mais tandis que la présence du défini français est une contrainte grammaticale, TEN peut, dans une grande partie des contextes, ne pas apparaître au profit d'une anaphore nominale nue (une possibilité théorique en 16).

Il est cependant des cas où TEN est lui aussi obligatoire (cf. notamment Štícha, 1999), comme l'apparition du complexe nominal anaphorique en position finale (rhématique). Ici, TEN est requis par une lecture coréférente :

19) Když	jste	odeš-l-i	s	Old-ou	do toho
quand	AUX.PST.2PL	partir-PTCP-PL.M	avec	NP-M.INS.SG	à ten.N.GEN.SG
kin-a,		uděla-l-i	jsme	si	s
cinéma-GEN.SG		faire-PTCP-PL.M	AUX.PST.1PL	REFL	avec
Luc-y		<i>ta</i>	<i>játr-a.</i>		
NP-F.INS.SG		ten.N.ACC.PL	foie-ACC.PL	tantum	

Quand vous êtes partis au cinéma avec Olda, Lucy et moi, on s'est fait le foie à manger.

(V. Havel, emprunté à Štícha, 1999 : 128)

Ici, *le foie (játra)*, constitue un rhème ancré contextuellement¹¹⁶, à savoir un référent dont il a déjà été question auparavant dans le discours, mais qui apparaît en fin d'énoncé et qui porte la charge informationnelle principale. Aussi doit-il être introduit par TEN, faute de quoi le lien de coréférence serait rompu : le lecteur comprendra qu'il s'agit d'un autre foie, mais pas de celui dont on a parlé avant. D'autres contextes où TEN est obligatoire seront présentés dans la Section 6.2.1.2 du Chapitre 6.

En ce qui concerne l'anaphore résomptive – un régime indexical à la charnière entre l'anaphore et la deixis –, ce qui vient d'être dit jusque-là pourrait laisser croire que c'est nécessairement TENTO/TENHLE qui prime dans cet usage. En effet, au sein de l'anaphore résomptive, le référent est institué moyennant un résumé des propos précédemment mentionnés. Or cette hypothèse s'avère juste uniquement pour le corpus des pièces de Havel (30 occurrences de TENTO/TENHLE contre 17

¹¹⁶ „Kontextově zapojené réma“ (Štícha, 1999).

occurrences de TEN), alors que dans notre échantillon d'ORTOFON v1, nous avons identifié 7 complexes nominaux introduits par TEN contre seulement 5 introduits par TENTO/TENHLE¹¹⁷ :

20) mně to přijd-e úplně divn-ý no nebo to
moi.DAT cela.NOM sembler-PRS.3SG complètement bizarre-N.NOM.SG MOD ou cela.NOM

jak zpíva-l-a Basikov-á v tý latin-ě víš
REL chanter-PST-3SG.F NP-F.NOM.SG en ten.F.LOC.SG latin-LOC.SG MOD

[...] .. *tenhle typ modern-ích písni*
tenhle.M.NOM.SG genre.NOM.SG moderne-F.GEN.PL chanson-GEN.PL

moi, je trouve ça complètement bizarre ; ou le truc que Basiková a chanté en latin, tu vois, ce genre de chansons modernes ORTOFON v1

21) já m-á-m totiž ty špatn-ý zád-a
moi.NOM avoir-PRS-1SG MOD ten.N.ACC.PL mauvais-N.ACC.PL dos-ACC.PL tantum

tím pád-em potřeb-uj-u aby
ten.M.INS.SG cas-INS.SG avoir besoin-PRS-1SG que.CONJ+AUX.COND.3PL

mě vopravova-l-i
moi.ACC réparer-PTCP-PL.M

moi, j'ai un souci avec mon dos (comme tu le sais) ; du coup, j'ai besoin qu'on me le répare ORTOFON v1

En 20, *tenhle typ moderních písni* (ce genre de chansons modernes) a pour antécédent non seulement l'ensemble du segment souligné mais aussi toute une partie qui a précédé l'extrait ici reproduit, et où il est question d'un certain type de chansons que la locutrice n'apprécie pas. La situation est analogue en 21, où l'expression figée *tím pádem* (du coup, par conséquent) fonctionne comme un connecteur logique à valeur consécutive.

Comme nous l'avons déjà signalé dans la Section 2.2.3, l'anaphore dite « associative » est un type d'anaphore tout à fait singulier. Il s'agit de cas où une entité est d'emblée introduite comme définie en première mention. Cette définitude repose sur un lien conventionnel plus ou moins solide entre l'entité ainsi introduite et une autre entité introduite précédemment (“[...] a hidden link or anchor which has to be introduced earlier.” (Löbner, 1998 : 1)). Seul TEN peut fonctionner avec certain types d'anaphore associatives, les autres formes étant exclues de ce contexte :

22) Po kin-ě se ještě bavi-l-i o (tom) film-u.
après cinéma-LOC.SG REFL encore discuter-PST-3PL.M de ten.M.LOC.SG film-LOC.SG
Après être allés au cinéma, ils ont encore discuté du film. (Czardybon, 2017 : 90)

¹¹⁷ Naturellement, les échantillons sont trop petits pour qu'une conclusion plus solide puissent en être tirée. Malheureusement, étant donné que les annotations des corpus du CNT ne font pas de distinctions entre les différents régimes d'indexicalité, vouloir comparer la distribution des trois formes avec l'anaphore résomptive dans l'ensemble d'un corpus (oral ou écrit) dépasserait les limites du présent travail.

Rappelons qu'en 22, le lien en question est celui qui existe entre une séance de cinéma et le film visionné lors de celle-ci (voir la Section 2.2.3). Lorsque, dans le Chapitre 6 (Section 6.2.2.2), nous étudierons plus en détail la grammaticalisation de TEN, nous analyserons les différents types d'anaphore associative eu égard à la possibilité de l'apparition de TEN.

3.3.3 Les emplois cataphoriques

TEN adnominal connaît également des emplois cataphoriques. Or ces emplois cataphoriques semblent relever (presque) exclusivement de la cataphore intra-syntagmatique, contrairement à TENTO (éventuellement TENHLE), qui renvoie couramment au-delà de la frontière d'un complexe nominal :

- 23) Lidsk-é těl-o se sklád-á z
 humain-N.NOM.SG corps-NOM.SG REFL composer-PRS.3SG de

těchto *úd-ů:* hlav-y, z krk-u...
 tento.M.GEN.PL membre-GEN.PL tête-GEN.SG de cou-GEN.SG

Le corps humain est composé des membres suivants : la tête, le cou... (Trávníček, 1951 : 1144)

Quant à TEN, il introduit couramment dans le discours des noms modifiés par une proposition subordonnée relative ou complétive, ainsi que par un superlatif, un numéral ordinal et quelques autres termes qui enclenchent l'unicité dénotative du nom. Un premier aperçu de ces phénomènes est donné dans les exemples 24-27 :

- 24) *ta* *žen-a,* *kter-ou* *jsem* *si* *vza-l*
 ta.F.NOM.SG femme-NOM.SG qui-F.ACC.SG AUX.PST.1SG REFL épouser-PTCP.SG.M
la femme que j'ai épousée
- 25) *ale* *on* *samozřejmě* *m-á* *potom* *ten* *psychick-ej*
 mais lui.NOM évidemment avoir-PRS.3SG après ten.M.ACC.SG psychologique-M.ACC.SG

problém *že* *jako* *si* *znova* *příprav-uj-e* *kdyby*
 problème.ACC.SG que MOD REFL de nouveau préparer-PRS-3SG si.CONJ+AUX.COND.3SG

byl *náhodou* *dalš-í* *hlad ..*
 être.PTCP.SG.M par hasard autre-M.NOM.SG famine.NOM.SG

mais après, lui, le problème mental [qu'il a], c'est qu'il continuer à se préparer au cas où il y aurait une autre famine ORTOFON v1

- 26) a *ten* *druh-ej* *film* *ten* *dáva-l-i*
 et ten.M.ACC.SG second-M.ACC.SG film.ACC.SG celui-ci.M.ACC.SG donner-PST-3PL.M

včera *na dvoj-ce*
 hier sur deux-LOC.SG

le second film, lui, il est passé hier sur la deuxième chaîne... ORTOFON v1

27) jo a todleto jsi vidě-l-a ? .. jak v tý na
 MOD cela.ACC AUX.PST.2SG voir-PTCP-SG.F comment dans ten.F.LOC.SG sur

tý *nej-rušnějš-í* *ulic-i* v Los Angeles
 ten.F.LOC.SG SUP-plus animé-F.LOC.SG avenue-LOC.SG à Los Angeles.LOC.SG

jsou takov-ý-dle upoutávk-y
 être.PRS.3PL tel-F.NOM.PL-SUFFIXE DEICTIQUE panneau-réclame-NOM.PL

et ça, tu l'as vu ? ces espèces de panneaux-réclames qu'il y a dans l'avenue la plus animée de Los Angeles ?
 ORTOFON v1

En 24, la présence de TEN est entraînée par celle d'une relative « établissante » ("referent-establishing relative clause" ; cf. Hawkins, 1978 : 131), à savoir une relative dont la seule présence suffit pour rendre le nom tête du complexe nominal défini. Cette nature de la relative est, bien évidemment, conditionnée par un contexte large (celui d'une société monogame) ainsi que, le cas échéant, par un contexte plus étroit (si le locuteur s'est remarié, il est question de son épouse au moment de l'énonciation). En 25, TEN apparaît suite à la présence de la subordonnée conjonctive pure complétive (malheureusement, le français ne permet pas que la même structure soit maintenue dans la traduction). Les énoncés 26 et 27 contiennent, respectivement, un TEN appelé par le numéral ordinal (*druhej*) et l'adjectif superlatif (*nejrušnější*). De nombreux auteurs (cf. Czardybon, 2017 ; Mathesius, 1926 ; Zubatý, 1916) prétendent que dans ces deux derniers cas, TEN est également doté d'une valeur « émotionnelle ». Nous nous pencherons sur tous ces cas dans les Sections 6.2.2.1 et 6.2.2.3.

Il est un autre phénomène, fort singulier, qu'il est possible, à notre avis, de rattacher à la cataphore intra-syntagmatique. Ici, TEN sert également de support à une subordonnée relative (« établissante ») ; seulement, cette relative restreint l'extension d'un nom employé le plus souvent de façon non-spécifique (très souvent générique). De plus, dans ce contexte, TEN possède la particularité de pouvoir être accentué. À ce titre, il faut rappeler ici que TEN adnominal n'est que très rarement accentué dans les tchèques soutenu et « commun »¹¹⁸. Outre le cas que nous venons d'évoquer, un seul type d'emploi de TEN identifié dans ORTOFON v1 fait systématiquement exception à la règle. Nous avons fait le choix d'y référer par le terme d'emplois « métalinguistiques ». Il s'agit de situations où le locuteur utilise TEN, car il est dans l'incapacité de trouver le terme lexical approprié¹¹⁹. Ici, l'accentuation de la forme peut facilement être expliquée par des contraintes

¹¹⁸ Voir par exemple Ziková et Volín (2012 : 33) : « En règle générale, *ten* est non-accentué. C'est un mot grammatical et, dans la plupart de ses formes flexionnelles, monosyllabique, qui, dans un contexte prosodique habituel, ne fonctionne pas comme la tête d'un groupe accentuel [...] » Nous traduisons : „*Ten* je zpravidla nepřízvučné. Je to gramatické, ve většině pádů jednoslabičné slovo, které v běžném prozodickém kontextu neplní roli hlavy mluvnického taktu [...]“

¹¹⁹ Il serait possible de voir dans cet emploi un type de déixis textuelle cataphorique, dans la mesure où TEN anticipe le terme à venir (lequel est considéré en sa qualité de signe).

prosodiques et pragmatiques – après la prononciation de la forme du démonstratif, le locuteur, en quête du nom juste, est amené à faire une pause :

- 28) no a Páj-e kvet-e ibišek úplně ve velk-ým ..
 MOD NP-F.DAT.SG fleurir-PRS.3SG hibiscus.NOM.SG complètement en grand-N.LOC.SG
- a ten můj by asi potřebova-l
 et ten.M.NOM.SG mien.M.NOM.SG AUX.COND.3SG MOD avoir besoin-PTCP.SG.M
- ostříhat **všechn-y** **ty** .. **všechn-y** **větv-e**
 couper-INF tout-F.ACC.PL ten.F.ACC.PL tout-F.ACC.PL branche-ACC.PL

Pája, son hibiscus, il est tout en fleurs... ; et le mien, il faudrait lui couper toutes les... toutes les branches
 ORTOFON v1

En 28, la locutrice accentue la forme, car elle marque une pause d'hésitation avant le mot *větv-e*. Plus tard, dans la Section 5.2.3.5, nous analyserons ce type d'emploi de façon plus détaillée pour voir, entre autres, qu'il est également possible dans des contextes où le complexe nominal ainsi introduit par TEN n'est ni référentiel ni défini.

À présent, revenons-en aux emplois qui nous intéressent ici, faisant également exception à la règle susmentionnée. Dans ces emplois, TEN introduit un « référent » non-spécifique (ou d'une spécificité faible, reléguée au second plan), modifié par une subordonnée relative exprimant sa seule appartenance à une classe/catégorie¹²⁰. Comme l'expliquent Volín et Zíková (2012), à qui nous empruntons ici les deux exemples suivants, une relation structurelle est établie entre TEN et la subordonnée relative dite « classificatoire » („klasifikující“)

- 29) Obslouž-í jen **ty** **zákazník-y**, **kter-í** **potřeb-uj-í**
 servir-PRS.3SG seulement ten.M.ACC.PL client-ACC.PL qui-M.ACC.PL avoir besoin-PRS-3PL
- lék-y bez předpis-u.
 médicament-ACC.PL sans ordonnance-GEN.SG

*Il sert uniquement **les clients** qui viennent chercher des médicaments sans ordonnance.*

- 30) Vítěz-em se stáv-á **ten** **účastník**,
 gagnant-INS.SG REFL devenir-PRS.3SG ten.M.NOM.SG participant.NOM.SG
- kter-ý zůstan-e jako posledn-í v nov-ém
 qui-M.NOM.SG rester-FUT.3SG comme dernier-M.NOM.SG dans nouveau-M.LOC.SG
- voz-e nej-méně po dob-u 120 minut.
 véhicule-LOC.SG SUP-moins pour durée-ACC.SG 120 minute.GEN.PL

***Le participant** qui sera le dernier à quitter le véhicule après une durée d'au moins 120 minutes gagna la compétition.*

¹²⁰ Il s'agit ainsi d'une lecture *de dicto* – opaque, non référentielle – et non pas *de re* (cf. Ortmann, 2014 : 313).

L'information contenue dans la proposition subordonnée possède ici une signification de catégorisation ou de classification. Elle détermine le référent en question tout en l'extrayant d'un ensemble plus large de référents potentiels. D'habitude, le référent ainsi déterminé est d'une nature différente comparé à celui qui est identifié par une subordonnée relative avec *ten* non-accentué : ce n'est pas un sujet singularisé et pleinement individualisé mais un sujet dépourvu de spécification plus détaillée, qui ne se distingue que par un certain ensemble de propriétés précisées au sein de la subordonnée ; toutefois, son identité est souvent inconnue ou non-pertinente [...] ¹²¹ (Zíková et Volín, 2012 : 36)

En vérité, cet emploi de TEN semble posséder toutes les propriétés de la classe des « déterminatifs » dont parle Diessel (1999), classe déjà reconnue par Quirk (1972) et par Himmelmann (1997) :

Similar to the recognitional use, the demonstratives [...] occur with a subsequent relative clause. However, they do not appeal to private shared knowledge, nor do they focus the hearer's attention on entities in the speech situation or discourse. These demonstratives do not fit any of the uses discussed thus far. Himmelmann (1997:78) characterizes them as semantically empty proforms that provide an anchorage point for the relative clauses that follow. (Diessel, 1999 : 108)

31) The true artist is like one of *those scientists* who, from a single bone, can reconstruct an animal's entire body. (Diessel, 1999 : 108 ; emprunté à Himmelmann, 1997 : 78)¹²²

On voit pourquoi ce cas de figure doit être dissocié de celui où TEN participe également d'une relation de cataphore intra-syntagmatique avec la subordonnée relative, mais où l'ensemble du complexe nominal permet d'identifier un référent spécifique, comme en 32 :

32) už to	není	ta	<i>Srí Lank-a</i>
dějà ce	être.PRS.3SG.NEG	ten.F.NOM.SG	Sri Lanka-NOM.SG
co	to	byl-a	třeba
REL	ce	être.PST-3SG.F	par exemple
		před pět-i	rok-ama jo
		an-INS.PL	MOD

ce n'est plus le Sri Lanka que c'était il y a par exemple cinq ans, tu vois ORTOFON v1

Cependant, un examen de notre échantillon de 1 000 occurrences d'ORTOFON v1 a révélé que TEN n'y est jamais accentué lorsqu'il apparaît dans une relation de cataphore intra-syntagmatique avec une subordonnée relative de nature « établissante » (23 occurrences au total), et ce même lorsqu'une lecture « classificatoire » de celle-ci ne peut être exclue¹²³.

¹²¹ Nous traduisons : „Informace obsažená ve vedlejší větě zde má kategorizační či klasifikující význam. Určuje příslušný referent jeho vydělením z širší množiny potenciálních referentů. Takto určený referent je zpravidla odlišné povahy než referent identifikovaný přívlastkovou větou s nepřízvučným *ten*: není to jedinečný, plně individualizovaný subjekt, ale subjekt blíže nespecifikovaný, který je vymezen pouze určitým souborem vlastností upřesněných ve vedlejší větě; jeho identita je ovšem často neznámá nebo není relevantní [...]“

¹²² En allemand, c'est la forme *derjenige* qui est réservée à ces emplois. D'ailleurs, un « déterminatif » n'est pas nécessairement un démonstratif en emploi adnominal, il peut aussi s'agir d'emplois pronominaux. En français, seuls ces derniers sont possibles : *Ceux qui fument nuisent à leur santé.*

¹²³ Cependant, ces cas y sont très rares (voir la Section 6.2.2.1 du Chapitre 6). De plus, l'expérience menée par Zíková et Volín (2012 : 33) montre que l'accentuation de TEN dans ces contextes n'est pas systématique. Si les cas en question sont plus nombreux dans le corpus composé des pièces de Havel, c'est également là que, malheureusement, l'accentuation de la forme ne peut être examinée.

À présent, un constat très important – tant du point de vue contrastif que du point de vue de la grammaticalisation de TEN en tchèque – doit être fait au sujet de l'anaphore associative et de la cataphore intra-syntagmatique. Comme le lecteur l'aura certainement remarqué, dans les exemples appartenant à ces deux cas de figure, le français, langue pourvue d'articles définis, admet plus difficilement, voire pas du tout, la présence d'un démonstratif. Ainsi, le premier, voire l'unique choix y sera toujours l'article défini. En 26 et 27, CE serait très difficilement acceptable ; dans les exemples 22, 24, 29, 30 et 32, sa présence est envisageable, mais elle changerait de façon radicale l'interprétation des énoncés : 22 ne s'interpréterait plus comme une anaphore associative, l'identification du film ne reposant plus sur le cadre situationnel posé par le discours antérieur :

33) Après être allés au cinéma, ils ont encore discuté de *ce film*.

En 24 et 32, le remplacement du défini par le démonstratif fera en sorte que l'identification du référent ne s'appuiera plus exclusivement sur la présence de la relative, relative dont le statut restrictif/déterminatif n'apparaîtra plus, d'ailleurs, avec la même évidence :

34) *Cette femme* que j'ai épousée.

35) Ce n'est plus *ce Sri Lanka* que c'était il y a par exemple cinq ans, tu vois.

En 29 et 30, c'est la nature de tout l'énoncé qui changera complètement, l'énoncé passant du non-spécifique au spécifique :

36) Il sert uniquement *ces clients*(,) qui viennent chercher des médicaments sans ordonnance.

37) *Ce participant*(,) qui sera le dernier à quitter le véhicule après une durée d'au moins 120 minutes(,) gagnera la compétition.

Tout cela plus le constat – crucial – que dans aucun de ces énoncés tchèques, la présence de TEN n'est ressentie comme obligatoire nous donne déjà une certaine idée du degré de grammaticalisation subi par ce démonstratif. À cela s'ajoute le fait que, mis à part les cas où TEN fonctionne comme un « déterminatif » (exemples 29 et 30), sa présence dans ces contextes est justement bien plus élevée dans la langue parlée informelle que dans la langue (écrite) soutenue (cf. Czardybon, 2017 : 81, 88 et 92).

3.3.4 Les emplois exophoriques mémoriels

La forme de loin la plus sollicitée dans le régime de l'exophore mémorielle est TEN. De même que dans le cas des emplois anaphoriques, ici aussi, TEN adnominal a un statut fort équivoque, car il est plus ou moins proche de l'article défini français. Dans l'exemple suivant, les locuteurs mobilisent

leur mémoire pour se rappeler des référents qui, en même temps, se trouvent représentés sur une photo devant eux (pourtant, l'identification de ces référents se fait bel et bien via le souvenir d'un vécu partagé – les locuteurs regardent des photos prises lors de leurs dernières vacances, passées dans un pays exotique, tout en s'en rappelant les moments forts) :

- 38) L1 : tady jsou ty **promenád-y** ne ? kolem
ici être.PRS.3PL ten.F.NOM.PL promenade-NOM.PL MOD autour
- ty chodník-y**
ten.M.NOM.PL trottoir-NOM.PL
- L2 : jo tady-hlenc¹²⁴ je **ta promenád-a**
MOD ici-SUFFIXE DEICTIQUE être.PRS.3SG ten.F.NOM.SG promenade-NOM.SG
- L1 : ici, il y a **les/ces promenades** et **les/ces trottoirs** tout autour, non ?
L2 : ouais, là, il y a **la/cette promenade** ORTOFON v1

Nous estimons que les deux façons de rendre TEN en français sont ici possibles, le choix du défini ou du démonstratif dépendant essentiellement du coût d'activation du référent pour le locuteur ainsi que du coût d'activation que le locuteur anticipe chez son interlocuteur. De même que dans le régime anaphorique, ici aussi, la double transposition possible de TEN dans une langue à articles définis nous amène, d'ores et déjà, à apercevoir une des difficultés que pose l'étude de la grammaticalisation du démonstratif et le passage d'un stade à l'autre. Pour les deux domaines d'usage, des critères d'analyse seront détaillés dans le Chapitre 6 (notamment dans les Sections 6.2.1.2 et 6.2.1.3). Ces critères s'appuieront, avant tout, sur les travaux de Laury (1997) et de Zíková (2017).

Plus haut, nous écrivions que d'autres formes du démonstratif tchèque adnominal connaissent des emplois mémoriels. Les dictionnaires et les grammaires citent notamment les formes ONEN et TAMTEN. Si ONEN est typique de la langue (très) soutenue, où il fait entrer, qui plus est, une teinte d'archaïsme (voir la Section 3.1.1 *supra*), TAMTEN semble évoquer, au contraire, plutôt un registre informel :

- 39) Druh-á alternativ-a, německ-á,
second-F.NOM.SG alternative-NOM.SG allemande-F.NOM.SG
- je protislovensk-á. Je variant-ou
être.PRS.3SG anti-slovaque-F.NOM.SG être.PRS.3SG variante-INS.SG
- oné předválečn-é snah-y přenechat**
onen.F.GEN.SG d'avant-guerre-F.GEN.SG tentative-GEN.SG abandonner-INF

¹²⁴ La forme TADYHLE, contenant la particule *hle*, ainsi que ces variantes familières comme *tadydle*, *tadyhlenc*, *tajdle*, etc. sont marquées sur le plan diatopique comme provenant de Bohême. Dans ORAL v1, la fréquence relative de ces formes est de 113,8 parmi les locuteurs originaires de Bohême et 24,2 pour ceux originaires de Moravie-Silésie.

Slovensk-o **velk-ému** **východn-ímu** **bratr-ovi.**
 Slovaquie-ACC.SG grand-M.DAT.SG oriental-M.DAT.SG frère-DAT.SG

La seconde alternative, l'alternative allemande, est anti-slovaque. Il s'agit d'une variante de cette tentative d'avant-guerre d'abandonner la Slovaquie au grand frère oriental.
 (emprunté à Berger, 1993 : 423)

40) O **tamtěch** **věc-ech** ne-v-í-š nic?
 de tamten.F.LOC.PL chose-LOC.PL NEG-savoir-PRS-2SG rien.ACC
De ces choses-là, tu n'en sais rien ? (Slovník spisovného jazyka českého, en ligne)

Berger (1993 : 422-423) souligne également la propension de ONEN à l'introduction d'un nom modifié par une proposition relative. Ce même cas de figure se rencontre également avec TEN. Mais contrairement aux occurrences avec une relative « établissante », où TEN ne réfère qu'en aval et à l'intérieur du complexe nominal, ici, il possède également une référence en amont :

41) a **ten** **olivov-ej** **olej ..** co **jsme** kupova-l-i ..
 et ten.M.NOM.SG d'olive-M.NOM.SG huile.NOM.SG REL AUX.PST.1PL acheter-PTCP-PL.M

 tak ten je taky už podražen-ej
 alors celui-ci.M.NOM.SG être.PRS.3SG aussi déjà plus cher-M.NOM.SG

et cette/l'huile d'olive... qu'on achetait..., alors elle aussi, elle coûte déjà plus cher
 ORTOFON v1

Même si les deux cas de figure ne sont pas toujours faciles à dissocier, nous sommes d'avis qu'en 41, contrairement à 24 et 32, la relative ne serait pas une relative « établissante », c'est-à-dire que sa seule présence ne suffirait pas pour instaurer l'unicité dénotative du complexe nominal sur le plan sémantique et pour identifier son référent sur le plan référentiel : le référent serait ainsi identifié non seulement grâce à l'information contenue dans la relative mais aussi grâce à la mobilisation d'un souvenir partagé (*tu te rappelles l'huile d'olive qu'on achetait ?*)¹²⁵. Trovesi (2002 : 132) note très justement à cet égard que

[d]ans certains cas de renvois anaphoriques et notamment lors de l'introduction d'un nouvel objet dans la communication, l'interprétation de la valeur fonctionnelle du démonstratif du type « ten » est équivoque, la sémantique de l'article défini et celle du pronom démonstratif se chevauchant ici. Pour l'italien parlé, Beretta (1992 : 21) démontre un emploi plus fréquent du pronom démonstratif à la place de l'article défini en tant que « moyen plus fort pour exprimer la définitude », et ce par besoin d'une clarté et netteté plus grandes de la langue parlée par comparaison aux autres types de textes¹²⁶.

¹²⁵ D'autant que le locuteur marque une pause entre le nom et la relative, celle-ci paraissant ainsi ajoutée « après coup ».
¹²⁶ Nous traduisons : „V několika případech anaforického odkazu a zvlášť při uvedení nového předmětu do komunikace je interpretace funkční hodnoty demonstrativa typu „ten“ dvojsmyslná, sémantika určitého členu a ukazovacího zájmena se tu překrývá. Pro mluvenou italštinu dokazuje Beretta (1992 :21) častější užití ukazovacího zájmena místo určitého členu jako „silnější prostředek k vyjádření určenosti“ a to pro potřebu větší jasnosti a zřetelnosti mluveného jazyka oproti jiným typům textů.“

Par « renvois anaphoriques », Trovesi entend l'emploi endophrorique en général et, plus précisément encore, l'emploi cataphorique intra-syntagmatique. Si TEN possède la valeur fonctionnelle d'un démonstratif, il y a également, en plus de ce dernier, un emploi mémoriel. La remarque faite au sujet de l'italien est elle aussi utile, car le même phénomène se produit en français lorsque LE est remplacé par CE dans ce type de contextes (voir les exemples 34 et 35 *supra* et la Section 4.5.3 du chapitre suivant).

Force est de constater que s'agissant des autres formes que TEN en emplois mémoriels, nous n'avons trouvé, dans nos données tchèques, qu'une seule occurrence de TENHLE, occurrence qui, de surcroît, se présente comme ambiguë, car TENHLE s'y combine avec TEN. Il y est question du nettoyage des différentes parties d'un terrarium pour hamster (l'ensemble de la situation fait croire que l'objet n'y est pas physiquement présent) :

- 42) a jinak **tohle** jinak **to ..** **to** **hnízd-o**
 et sinon tenhle.N.ACC.SG sinon ten.N.ACC.SG nid-ACC.SG
 bych necha-l
 AUX.COND.1SG laisser-PTCP.SG.M
 et, sinon, ce... sinon, le... le nid, je le laisserais (comme il est) ORTOFON v1

En parlant des emplois mémoriels, nous trouvons opportun de mentionner un cas que la littérature existante semble globalement négliger, mais que nous avons détecté dans les deux langues. Il s'agit d'emplois laissant à première vue croire que le démonstratif y remplit sa fonction habituelle, à savoir qu'il permet à l'interlocuteur d'identifier un référent. Or, à y regarder de plus près, on s'aperçoit qu'il n'en est rien : en réalité, l'expression démonstrative ne renvoie pas à un référent existant et identifiable, mais il sert uniquement à renseigner sur une alternative purement hypothétique. Si nous avons choisi le présent moment pour introduire ce phénomène, c'est que celui-ci semble justement être le plus proche des emplois mémoriels :

- 43) představ-a že prostě chod-í-m se ps-em ven a pokaždě
 idée-NOM.SG que MOD aller-PRS-1SG avec chien-INS.SG dehors et chaque fois
 se boj-í-m abych potka-l-a **tamty**
 REFL avoir peur-PRS-1SG que.CONJ+AUX.COND.1SG croiser-PTCP-SG.F tamten.M.ACC.PL
 a **tamty** **soused-y** protože se na mě bud-ou
 et tamten.M.ACC.PL voisin-ACC.PL parce que REFL sur moi.ACC AUX.FUT-3PL
 ksichtit ..
 faire la gueule.INF
 genre l'idée d'avoir peur, à chaque fois que je sors promener mon chien, de croiser ces voisins-ci ou ces voisins-là qui vont me faire la gueule... ORTOFON v1

Ce type d'emploi semble admettre toutes les formes du démonstratif. Parfois, une forme de proximité apparaît à côté d'une forme d'éloignement pour créer un contraste :

- 44) Tak za prvé: nikdy nebudete mít jistotu, zda **ten** či **onen doklad** Peachumovy zrady není jen mnou inspirovaná zkouška vaší spolehlivosti [...]

*Tout d'abord, vous ne serez jamais certain que **telle preuve de la trahison de Peachum** ou **telle autre** ne soit juste une épreuve inspirée par moi dans le but de tester votre loyauté [...]*

V. Havel, *La Grande roue*, 1975, InterCorp v10

En 44, la traduction est faite par nous-même, car la version du traducteur I. Palec ne reproduit pas le phénomène en question¹²⁷ (notre traduction met aussi en valeur le fait que le français use parfois, dans ces contextes, de la forme du démonstratif indéfini TEL).

3.4 La fonction de substantivation de TEN adnominal

Un autre phénomène relatif à TEN adnominal qu'il convient de mentionner ici est le rôle que ce démonstratif joue dans la nominalisation (substantivation) des adjectifs. Ce rôle n'a pas fait l'objet de beaucoup d'attention dans la littérature. Il est brièvement mentionné par Mathesius (1926) et par Vey (1946), ce dernier écrivant que TEN peut s'utiliser « devant un adjectif employé comme substantif » (1946 : 61). *Mluvnice češtiny 2* (1986 : 87) met en avant le fait que TEN a tendance, notamment dans le discours parlé („v mluvených projevech“), à précéder les identificateurs possessifs. Ce phénomène est fort intéressant du point de vue de la comparaison entre nos deux langues, car un parallèle se laisse facilement apercevoir entre cette construction et la construction française avec l'article défini introduisant un adjectif possessif, le tout ayant la même distribution syntaxique qu'un pronom (un SN) :

45) ten	tvůj
ten.M.NOM.SG	tien.M.NOM.SG
<i>le tien</i>	

En français, les formes du paradigme de MIEN appartiennent à la classe des adjectifs (éventuellement à celle des post-déterminants ; voir la Section 4.1.2), contrairement aux formes du paradigme de MON, qui sont des déterminants, et donc en distribution complémentaire avec l'article défini. En tchèque, il n'est guère surprenant que cette opposition fasse défaut, le paradigme des possessifs pouvant s'employer a) en emploi adnominal devant un substantif (*tvůj kolega* ; *ton*

¹²⁷ Sa traduction est la suivante : « Eh bien, d'abord, vous ne serez jamais sûr, vous, que Volland me trahit vraiment et que ce n'est pas moi qui mets à l'épreuve votre fidélité – ce que je ne manquerai pas de faire. »

*collègue/un collègue à toi*¹²⁸), b) en emploi substantivé avec TEN comme en 45 ou c) en emploi purement pronominal, notamment dans certains contextes figés où il est question de référer à des personnes de la sphère intime du locuteur :

- 46) *naš-i*
 nôtre-M.NOM.PL
mes parents

Cette fonction de substantivation est ensuite abordée par Adamec (1983 : 155), pour qui de tels emplois de TEN occuperaient une position quelque peu ambiguë, à la charnière entre emplois adnominaux et pronominaux (littéralement : « adjectivaux » et « nominaux ») :

- 47) L1 : Kter-ý kousek chc-e-š?
 quel-M.ACC.SG morceau.ACC.SG vouloir-PRS-2SG
- L2 : **Ten** **menš-í.**
 ten.M.ACC.SG plus petit-M.ACC.SG
- L1 : *Quel morceau veux-tu ?*
 L2 : **Le plus petit** (*des deux*). (Adamec, 1983 : 155)¹²⁹

Il en conclut cependant que TEN reste tout de même plus proche d'un adjectif. Nous croyons d'ailleurs que ce qui le fait avant tout douter quant à la nature de TEN, c'est le fait que TEN apparaît aussi, en tchèque, avec des expressions substantivées contenant des adverbes/locutions adverbiales ou des subordonnées, là où le français emploie partout la proforme pronominale CELUI :

- 48) L1 : Kter-ý z těch obraz-ů se ti nej-víc
 lequel-M.NOM.SG de ten.M.GEN.PL tableau-GEN.PL REFL toi.DAT SUP-plus
- líb-í?
 plaître-PRS.3SG
- L2 : **Ten** **vpravo.**
 celui.M.NOM.SG à droite
- L1 : *Lequel des tableaux te plaît le plus ?*
 L2 : **Celui à droite.** (Adamec, 1983 : 155)

L'auteur qui semble le premier prendre la mesure de l'importance de ce phénomène est de nouveau Berger (1993), pour qui la fonction de substantivation („die substantivierende Funktion“) constitue l'une des principales fonctions de TEN. En guise d'introduction, il écrit que

¹²⁸ Ici, l'on remarquera que contrairement au français, où le déterminant possessif, de par son alignement sur le défini, est porteur du présupposé d'unicité/maximalité (voir la Section 3.1.3), cela n'est pas le cas en tchèque, où un complexe nominal possessif peut toujours, du moins sur le plan strictement sémantique, recevoir une double lecture : définie et indéfinie.

¹²⁹ Cet énoncé a été soumis à l'évaluation de nos informateurs, à qui nous avons demandé de choisir entre *ten menší* et *menší* (nom nu). Une seule personne sur dix a préféré le nom nu.

[...] in Artikelsprachen dem bestimmten Artikel genau diese Funktion zukommt und entsprechende Verwendungen daher eher als Spezialfälle der artikelähnlichen Funktionen von *ten* angesehen werden. (1993 : 97)

Autrement dit, TEN introduisant un adjectif avec lequel il construit un complexe nominal en l'absence de nom relève des cas où le démonstratif est en voie de grammaticalisation. Ce constat est vérifiable à la fois au sein même du tchèque et par comparaison avec d'autres langues. Au sein du tchèque lui-même, nous pouvons observer que dans les exemples comme 45 et 47, TEN est ressenti comme quasi-obligatoire¹³⁰. Ces emplois ressortissent, en l'occurrence, à un contexte exophorique situationnel pour 47 et, le plus probablement, à un contexte anaphorique pour 45 (cf. *Je trouve que mon père est trop sévère, le tien est plus sympa.*). Mais, outre cela, nous constatons que le complexe nominal peut également se prêter à un usage générique, indépendant de tout régime d'indexicalité :

49) *ti* *bohat-i*
 ten.M.NOM.PL riche-M.NOM.PL
 les riches

Ici, il est donc question des gens riches en général. S'il est possible de trouver des contextes où ce même emploi générique de l'adjectif ainsi substantivé se passerait plus aisément de TEN, la présence de ce dernier est perçue comme très naturelle et neutre par un locuteur natif, contrairement à la situation où le nom serait présent :

50) *ti* *bohat-i* *lid-é*
 ten.M.NOM.PL riche-M.NOM.PL gens-NOM
 les gens riches

En 50, pourvu que la lecture générique soit gardée, TEN sera forcément à mettre au compte d'un emploi « émotionnel » (voir la Section 3.5 *infra*). Une différence analogue peut également être observée dans une langue comme l'anglais, où l'article défini n'accompagne pas d'habitude le nom dans ses emplois génériques au pluriel, mais où il sera pourtant ressenti comme parfaitement grammatical (et obligatoire) lorsque la catégorie est désignée par un adjectif substantivé :

51) *the rich* and *the poor*
 the English and *the French*

Fort malheureusement, la requête que nous avons formulée pour ORTOFON v1 n'est pas appropriée à la recherche des occurrences contenant un adjectif substantivé, car pour identifier le plus d'occurrences de TEN apparaissant en emploi adnominal, nous avons été obligé de demander que le

¹³⁰ Au sein du groupe slave, il le serait encore davantage en slovène. Dans cette langue, le démonstratif adnominal TA va jusqu'à perdre son autonomie morphologique (sa déclinaison) lorsqu'il est employé dans ce contexte (Tovesi, 2002).

démonstratif soit suivi d'un nom, éventuellement d'un numéral (voir la section 2.2.1.)¹³¹. Par conséquent, il n'y a que 30 occurrences de ce type dans nos données, dont 3 dans le sous-corpus des pièces de V. Havel. En voici quelques exemples. En 52, la locutrice raconte ses péripéties avec une paire de martres. Après y avoir référé par le nom *kuna* (*martre*), elle se contente de reprendre les adjectifs *malá* (*petite/jeune*) et *stará* (*vieille*), précédés de TEN :

52) řek-l	tak už to	prostě dál	ne-jd-e	my
dire-PST.3SG.M	ainsi plus cela.NOM	MOD plus loin	NEG-aller-PRS.3SG nous.NOM	
mus-í-me	dostat ..	v dob-ě	kdy	ta star-á
devoir-PRS-1PL sortir.INF	à moment-LOC.SG	où	ten.F.NOM.SG	vieux-F.NOM.SG
je .	na lov-u .	mus-í-me	tu mal-ou	dostat . pryč ..
être.PRS.3SG	à chasse-LOC.SG	devoir-PRS-1PL	ten.F.ACC.SG petit-F.ACC.SG	sortir dehors
a ta	star-á	už ne-bud-e	mít	prostě důvod .
et ten.F.NOM.SG	vieux-F.NOM.SG	plus NEG-AUX.FUT-3SG	avoir.INF	MOD raison.ACC.SG
tam lézt ..	tak nej-dřív	zača-l-i	vobíjet ty	
là-bas grimper.INF	alors SUP-plus tôt	commencer-PST-3PL.M	garnir.INF	ten.M.ACC.PL
vokap-y ..	jenomže ta	star-á	prostě mě-l-a	
avant-toit-ACC.PL	mais ten.F.NOM.SG	vieux-F.NOM.SG	MOD avoir-PST-3SG.F	
důvod	se tam	dostat		
raison.ACC.SG	REFL là-bas	parvenir.INF		

il a dit c'est plus possible comme ça, on doit arriver... au moment où la vieille est à la chasse, on doit arriver à faire dégager la jeune... et la vieille n'aura plus de raison pour grimper là-haut ; du coup, ils ont commencé par garnir les avant-toits... seulement, la vieille avant tout simplement une raison pour rentrer là-bas
ORTOFON v1

En 53, la locutrice parle de ses voisins âgés. Plutôt que d'employer un nom modifié par un adjectif, elle n'emploie que l'adjectif *staří* (*vieux*) :

53) takže před těma	řekněme dvacet-i	let-y	když jste	se sem
alors avant ten.M.INS.PL	MOD vingt-INS	an-INS.PL	quand AUX.PST.2PL	REFL ici
přistěhova-l-i	tak tu byl-i	ještě fajn-í	lid-é	
emménager-PTCP-PL.M	alors ici être.PST-3PL.M	encore sympa-M.NOM.PL	gens-NOM	
vesměs takže já	už já	už je	jako všichn-i	ti
en général alors moi.NOM	plus moi.NOM	plus eux.ACC	MOD tout-M.NOM.PL	ten.M.NOM.PL
stař-í	už vlastně umře-l-i ..	nebo je	ode* odeš-l-i	no
vieux-M.NOM.PL	déjà MOD mourir-PST-3PL.M	ou eux.ACC	partir-PST-3PL.M	MOD
ted' ti	stař-í	už tu začín-á-me	být my	
maintenant ten.M.NOM.PL	vieux-M.NOM.PL	déjà ici commencer-PRS-1PL	être.INF	nous.NOM

¹³¹ Si nous ne l'avions pas fait, la quantité de bruit eût été insupportable.

*du coup, il y a vingt ans (comme je disais), quand vous avez emménagé ici, il y avait encore des gens sympas en général ; du coup, moi, je les... genre **tous les anciens** sont déjà décédés ou ils sont partis, tu vois ; maintenant, c'est nous qui commençons à être **les anciens** ici* ORTOFON v1

Dans l'exemple suivant, il est question de deux référents mentionnés auparavant (deux Français). La reprise se fait uniquement par la combinaison TEN + numéral. Précisons que si nous parlons de substantivation, c'est que le numéral possède ici davantage la distribution d'un adjectif (ou d'un post-déterminant¹³²) que d'un pronom (contrairement à la seconde occurrence de *jeden* : **jeden už pracuje**). Il serait ainsi proche de *un* dans *l'un et l'autre* :

54) Katk-a	[...] mě-l-a	tady na návštěv-ě	[dv-a	Francouz-e]	i + ii
NP-F.NOM.SG	avoir-PST-3SG.F	ici sur visite-LOC.SG	deux-M.ACC	Français-ACC.PL	
 jak	 byl-a	 ve Franci-i	 tak tam	 chodi-l-i	
comme	être.PST-3SG.F	en France-LOC.SG	alors là-bas	aller-PST-3PL.M	
 lézt	 na	 nějak-ou	 stěn-u		
escalader.INF	sur	quelque-F.ACC.SG	mur-ACC.SG		
...					
[ten	jeden] _i	je	student		
ten.M.NOM.SG	un.M.NOM.SG	être.PRS.3SG	student.NOM.SG		
 a	 [jeden] _{ii}	 už	 prac-uj-e		
et	un.M.NOM.SG	déjà	travailler-PRS-3SG		

*Katka [...] a accueilli chez elle [deux Français]_{i + ii} ; au moment où elle était en France, ils allaient faire des escalades sur un mur d'escalade [...] [**l'un**]_i (d'entre eux) est étudiant et [**l'autre**]_{ii} travaille déjà.* ORTOFON v1

En 55, le complexe nominal tchèque *ty neodbytné* (littéralement : *les collantes/les pressantes*) renvoie à une certaine catégorie de femmes dans la vision du locuteur. Nous remarquerons que nous sommes face à un emploi anaphorique partiellement coréférentiel (indirect) : l'anaphorique reprend uniquement la référence d'une partie de son antécédent (= les femmes en général)¹³³ :

55) JIM Zajímala by mě ještě jedna věc: dosud jsme mluvili převážně o tom, jak nejlépe [ženské]_{i + ii} získávat. Neméně důležitou se mi však zdá být otázka, jak se [jich]_{i + ii} potom zbavovat – MACHEATH Myslíš [**těch neodbytných**]_{ii}?

*JULES Il y a encore quelque chose qui m'intéresse. Jusqu'à maintenant on a surtout parlé de la méthode pour conquérir [les petites femmes]_{i + ii}. Comment s'[en]_{i + ii} débarrasser après, ça m'a l'air aussi une question très importante ... MAX Tu veux parler [**des pots de colle**]_{ii}, comme on les appelle, n'est-ce pas ?* V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

¹³² Précisons qu'il est ici question uniquement de sa distribution syntaxique et non pas de sa nature, qui, elle, se rapprocherait le plus d'un pronom (voir la Section 3.1.3).

¹³³ Si tant est que les emplois génériques peuvent être considérés comme référentiels. Nous essaierons d'aborder cette question épineuse dans la Section 4.3 du Chapitre 4.

Résumons : en 52, 54 et 55, le complexe nominal possède une référence anaphorique (indirecte en 54 et 55, car ne reprenant qu'une part de la référence de son antécédent). En 53, le complexe nominal semble référer de façon mémorielle, du moins dans sa première occurrence. Sa seconde occurrence pourrait, quant à elle, être analysée comme un cas d'anaphore associative du type situationnel (*les anciens de cet endroit, quels qu'ils soient* ; voir la Section 6.2.2.2¹³⁴). Dans tous les cas susmentionnés, en plus de son rôle dans la nominalisation, la présence de TEN est entraînée par des facteurs contextuels¹³⁵.

Une requête nous a permis de vérifier que l'emploi nominalisateur est loin d'être rare dans l'ensemble d'ORTOFON v1. Comme il est difficile de viser ce type d'occurrences dans le corpus, nous nous sommes limité à la recherche de TEN à l'accusatif suivi d'un adjectif substantivé (non suivi d'un nom) lui aussi à l'accusatif¹³⁶. Après l'élimination des occurrences bruitées, nous avons obtenu une fréquence relative égale à 327. Ce chiffre n'est pas négligeable, surtout au vu du fait que le système de la déclinaison tchèque contient sept cas.

3.5 Les emplois « émotionnels »

Vu leurs singularités patentes, nous avons décidé de traiter ces emplois à part, dans une section qui leur est exclusivement consacrée, même si, comme nous le verrons, des liens existent entre les emplois « émotionnels » et les régimes de référence indexicale habituellement reconnus. Les emplois en question sont de loin le plus courants avec TEN adnominal. Il s'agit avant tout de la présence d'un démonstratif avec un nom dont le référent est identifiable sans démonstratif. De ce fait, ce dernier apparaît comme étant « de trop » sur le plan de l'identification du référent, mais véhicule des nuances sémantico-pragmatiques importantes qui renseignent sur l'attitude du locuteur vis-à-vis du référent. Néanmoins, il existe aussi des cas où le démonstratif est porteur d'une subjectivité accrue tout en gardant son rôle direct dans le processus d'identification du référent. Ce rôle fondamental du démonstratif, nous y référerons en utilisant le terme de « fonction désignative »¹³⁷. Ainsi, après une introduction très générale, où il sera question de montrer que le démonstratif et la subjectivité vont de pair, ce seront précisément ces cas-ci que nous aborderons en premier, et qui nous serviront de

¹³⁴ C'est dans la Section 6.2.2.3 du Chapitre 6 que nous verrons qu'une grande partie de ce type d'occurrences relèvent également de la cataphore intra-syntagmatique du fait de contenir un adjectif superlatif, un numéral ordinal ou d'autres termes semblables.

¹³⁵ Comme ces facteurs sont tels que même si le substantif était restitué, TEN aurait de fortes chances de rester en place, il est possible d'en conclure que sa présence relève doublement de la grammaticalisation : une fois de par son rôle joué dans la nominalisation et une fois grâce à sa valeur contextuelle (discursive ou extra-discursive).

¹³⁶ [tag="PD..4-----.*" & !lemma="takový" & !lemma="tento" & !lemma="tenhle" & !lemma="tamten"]
[tag="A...4.---.---.*"] [!tag="N.*"]

¹³⁷ Synonyme de ce que Kleiber (1991a : 85) appelle le « sens instructionnel » du démonstratif, elle procède directement de la sémantique indexicale de cette unité linguistique.

transition entre les emplois « désignatifs » (soit les emplois prototypiques du démonstratif) et les emplois « émotionnels » au sens strict du terme.

3.5.1 Introduction : le démonstratif et la subjectivité par défaut

Pour commencer, un constat assez trivial s'impose : le démonstratif est une unité de la langue intrinsèquement liée à la subjectivité, ne serait-ce que pour la simple raison qu'il participe de la manière dont le locuteur organise l'espace (dont il constitue souvent le centre déictique ; voir la Section 1.1.1) et, plus généralement, de la manière dont il dirige l'attention de l'interlocuteur sur le bon référent. Le démonstratif véhicule donc une subjectivité par défaut, quels que soient ses emplois. Ce constat est d'ailleurs étendu par Kerbrat-Orecchioni à l'ensemble des déictiques, qu'elle décrit comme des

[...] unités, de quelque nature et de quelque niveau qu'elles soient, qui fonctionnent comme indices de l'inscription dans l'énoncé du sujet d'énonciation. (2014 [1999] : 36)

Dans les systèmes des différentes langues, diverses possibilités sont mises à la disposition du locuteur pour marquer son point de vue à travers les emplois du démonstratif. Ces stratégies exploitent assez souvent la notion de distance, qui peut être entendue au sens littéral au figuré, comme l'illustre en latin la possibilité d'un emploi péjoratif de la forme *ISTE* et celle d'un emploi mélioratif, honorifique, de la forme *ILLE*. Le concept de la « sphère personnelle du locuteur », forgé par Marchello-Nizia (2003) pour rendre compte du fonctionnement des démonstratifs en ancien français, en constitue un autre exemple éloquent. Ce qui nous semble ici important à relever, c'est que ces effets pragmatiques semblent toujours être dérivés du sémantisme originel des démonstratifs (ainsi, une forme qui, au départ, dénote un absent de l'interlocution en vient à véhiculer le respect pour ce qui est distant, hors d'atteinte, etc.).

Ensuite, la recatégorisation d'un référent déjà introduit en discours constitue une autre preuve puissante de la subjectivité propre à l'usage du démonstratif. Selon Gary-Prieur, cette reclassification s'accompagne d'un rejet au second plan de la fonction référentielle du démonstratif au profit de sa fonction subjective :

On assiste là à une renomination d'un référent introduit précédemment [...] le référent du GN est déjà connu, et la fonction référentielle passe au second plan, derrière la renomination qui vise à exprimer le point de vue du locuteur [...]. (2011 : 74)

Ainsi, parmi les trois fonctions du langage définies par Bühler (2008 [1934]), de tels usages du démonstratif ressortiraient davantage à la fonction *expressive* qu'à la fonction *référentielle*. À notre

avis, cette idée pourrait s'avérer cruciale pour l'analyse de tous les emplois du démonstratif qui font apparaître une subjectivité accrue du locuteur.

3.5.2 Les emplois « désignatifs » porteurs d'une subjectivité accrue

Des effets de subjectivité renforcée peuvent accompagner certains usages du démonstratif adnominal du fait que ceux-ci enfreignent, en l'occurrence, les conditions d'emploi prototypiques de cette unité linguistique, détaillées dans la Section 1.2 du Chapitre 1. C'est que dans ces cas particuliers, comme le dit Corblin, le référent est complètement « isolé sur des critères externes » (1987 : 217), indépendamment du sens descriptif du nom que le démonstratif introduit. Charolles commente ce cas de figure en écrivant que « [...] le référent est exhibé ici et maintenant comme quelque chose qui existe dans les conditions matérielles d'énonciation avant d'être catégorisé au niveau de base [...] » (2002 : 120). Concrètement, c'est ce qui se produit dans les deux exemples suivants. Les deux énoncés sont prononcés dans une situation où il n'y a, respectivement, qu'un seul morceau de gâteau et qu'une seule casquette, et où c'est donc l'article défini qui devrait normalement être employé :

56) Passe-moi *ce gâteau* !¹³⁸

Pose-moi *cette casquette*. (Charolles, 2002 : 121-122)

Charolles (2002 : 121) explique que

Si plusieurs gâteaux trônent au milieu de la table, le démonstratif, accompagné d'un geste de monstration sélectif, remplit sa mission sans effet particulier. S'il n'y a, par contre, notoirement qu'un seul gâteau dans l'environnement partagé, l'usage du démonstratif ne paraît motivé que s'il s'accompagne d'une **nuance d'impatience**.

Quant à l'énoncé comprenant le SN *cette casquette*, il serait concevable uniquement « [...] dans le contexte particulier où quelqu'un s'affiche soudainement avec une casquette notoirement incongrue sur la tête. » (Charolles, 2002 : 121-122).

Cependant, l'effet subjectif ainsi véhiculé par le démonstratif ne se réduit pas uniquement à une « nuance d'impatience ». Bonnard (1972 : 1207) écrit que dans ce type d'emplois, « [...] le démonstratif fait imaginer une qualification, favorable ou défavorable, selon les cas. » Par exemple, il est possible d'imaginer la situation où quelqu'un arrive avec une casquette toute neuve sur la tête,

¹³⁸ La valeur subjective (expressive) de cet énoncé n'est cependant pas due au seul démonstratif : elle provient de la synergie de trois phénomènes, à savoir : le démonstratif, la modalité phrastique exclamative et la présence du datif éthique. Nous remercions à Mme Nádvorníková d'avoir attiré notre attention sur ce fait.

suscitant chez le locuteur une réaction de forte admiration, en conséquence de quoi celui-ci s'exclame :

57) Regarde-moi *cette casquette* !

Pour résumer, des auteurs tels que Corblin (1987), Charolles (2002) et Danon-Boileau (1990) soutiennent que l'effet subjectif en question est dû au fait que le référent soit présenté à partir de la situation d'énonciation immédiate et les circonstances d'évaluation qui en dépendent. Ainsi, l'existence de ce type d'emplois du démonstratif semble être bien reconnue dans la littérature, à commencer par Fillmore (1982), qui en parle en utilisant le terme de « acknowledging uses », et pour qui il s'agit d'un usage dérivé du démonstratif, contrairement à ses usages prototypiques, dits par lui « informing » et « identifying » :

Elements of demonstrative systems that have primarily the Informing function are the demonstrative adverbs [...] those that have mainly the Identifying function may be thought of as the prototypic uses of [nominal] demonstratives, the Acknowledging use being derivative and secondary. (Fillmore, 1982 : 45)

Les emplois dits « informing » semblent réservés aux démonstratifs adverbiaux : *She lives here* ; les emplois dits « identifying » seraient les emplois premiers et prototypiques des démonstratifs nominaux : *Don't sit in this chair, but that one* ; les emplois dits « acknowledging » seraient, quant à eux, des emplois dérivés et secondaires des démonstratifs nominaux : *Susan, get that snake out of this house!* (Fillmore, 1982 : 42-44)¹³⁹

Gary-Prieur (2011) parle d'« effets de sens » du démonstratif, terme qui suggère, de même que chez Fillmore, que ces emplois ont leur origine dans le sémantisme de base de cette classe linguistique. De tels emplois existent également en tchèque. C'est Adamec (1983) qui en cite plusieurs exemples, dont un que nous reproduirons ici. Il est question d'une situation où le locuteur s'aperçoit qu'un groupe de personnes est en train de vandaliser sa voiture. Remarquons que 58 est assez proche de 56 et de 57 :

58) Co	to	prosím vás	děl-á-te?	Okamžitě	nech-te
que.ACC		MOD	faire-PRS-2PL	immédiatement	laisser-IMP.2PL
být	<i>to</i>		<i>aut-o</i> !		
être.INF	ten.N.ACC.SG		voiture-ACC.SG		

Qu'est-ce que vous faites ? Laissez cette voiture tranquille ! (Adamec, 1983 : 159)

¹³⁹ Notons que THIS et THAT ne seront pas ici accentués, contrairement à ce qui semble être la tendance des emplois « identifying ».

TEN apparaît ici dans un emploi situationnel, où le référent n'est pas contrasté avec d'autres référents (de même qu'en 56 et 57). En même temps, il ne s'agirait sans doute pas d'un emploi grammaticalisé de TEN, car l'énoncé relève de la modalité exclamative et il est prononcé dans un contexte très semblable à celui des exemples commentés par Charolles (2002)¹⁴⁰. Notre corpus oral tchèque contient d'autres occurrences de ce type. L'exemple 59 met en scène une femme qui vient d'agresser des policiers en public, au point que ces derniers ont été obligés de la menotter. La locutrice décrit la réaction des témoins de cette scène :

59) ted' všichn-i lid-i na ni kouka-l-i
 maintenant tout.M.NOM.PL gens-NOM sur elle.ACC.SG regarder-PST-3PL.M

říka-l-i páni ta žensk-á
 dire-PST-3PL.M MOD ten.F.NOM.SG femme-NOM.SG

et là tout le monde la regardait en disant mais alors cette femme ORTOFON v1

Si, dans les deux énoncés précédents, nous avons rendu TEN par un démonstratif dans la traduction française, ce choix est loin d'être anodin : il reflète notre conviction que TEN remplit ici – malgré l'affaiblissement qu'il affiche d'habitude dans les emplois situationnels (Berger, 1993 : 463) – la fonction d'un démonstratif avec la sémantique déictique qui lui incombe, exhibant le référent « [...] ici et maintenant comme quelque chose qui existe dans les conditions matérielles d'énonciation [...] » (Charolles, 2002 : 120). Les emplois comme 58 et 59 compliquent ainsi l'analyse de TEN adnominal dans les emplois situationnels, car ils semblent défier la vision globale d'un TEN en voie de grammaticalisation et proche de l'article défini (voir la Section 3.3.1 *supra*). Nous reviendrons à ce cas dans la Section 6.2.1.1. Par souci d'exhaustivité, rappelons que les emplois où TEN introduit un nom modifié par un superlatif ou un ordinal s'accompagneraient également d'un effet de subjectivité accrue (voir la Section 3.3.3 *supra* et la Section 6.2.2.3).

3.5.3 Les emplois « émotionnels » propres

La différence entre les usages dont il vient d'être question et les emplois à proprement parler « émotionnels » réside, selon nous, dans la nature de l'acte référentiel auquel le démonstratif se prête. C'est là précisément que deux cas de figure doivent être distingués. D'une part, il y a les emplois que l'on observe dans les exemples 56-59, où l'effet subjectif renforcé cohabite avec la fonction première (« désignative ») du démonstratif, qui est celle d'identifier le référent via son indexation sur le contexte. D'autres part, il y a les emplois où le démonstratif ne saurait être analysé comme participant de cette opération, et où, par conséquent, cette fonction première céderait entièrement la place à l'effet pragmatique subjectif. Nous avons pris la décision d'appeler ces emplois « émotionnels », en

¹⁴⁰ Mathesius (1926) cite également des exemples de ce type.

traduisant le terme de „funkce emocionální“, utilisé par Mathesius (1926), cependant qu'en fonction des auteurs des deux traditions linguistiques, ils sont aussi appelés « secondaires » (Adamec, 1983 ; Štícha, 1999), « expressifs » ou « additionnels » (Štícha, 1999), « affectifs » (Kleiber, 1991 ; Trávníček, 1951 ; Voráč, 1946), « de familiarité/distanciation » (Kleiber, 1991), etc.

Proches en cela de la position de Fillmore (1982), les chercheurs tchèques semblent s'accorder sur le constat que ces emplois occupent la périphérie du système du démonstratif. Cette vision des choses est bien résumée par la citation suivante de Štícha :

L'on peut, à mon avis, considérer comme généralement accepté (ou acceptable) le fait qu'il est possible d'attribuer au lexème *ten* trois ensembles de fonctions : 1) les fonctions déictiques (monstratives) dénotatives (d'identification), 2) les fonctions de renvoi contextuel et discursif (anaphoriques et cataphoriques) et 3) les fonctions expressives, dites « additionnelles ». Les deux premiers ensembles se trouvent au centre du système, le troisième, à sa périphérie. (1999 : 124)¹⁴¹

Voici quelques exemples (une partie de ceux-ci figurent déjà dans l'Introduction générale) :

60) Jen kdyby mne **ta** **hlav-a**
seulement si.CONJ+AUX.COND.3SG moi.DAT ten.F.NOM.SG tête-NOM.SG

přesta-l-a bolet!
cesser-PTCP-SG.F faire mal.INF

Si seulement je n'avais plus mal à la tête ! (Mathesius, 1926 : 40)

61) Une employée de banque parle de sa journée au travail :

jako dneska jsem mě-l-a devět lid-í vid'..
MOD aujourd'hui AUX.PST.1SG avoir-PTCP-SG.F neuf.ACC gens-GEN MOD

prostě **ty** **lid-i** ne-chod-ěj vůbec..
MOD ten.M.NOM.PL gens-NOM NEG-vont-PRS.3PL du tout

par exemple aujourd'hui, j'ai reçu neuf clients [en tout et pour tout], tu vois... genre les gens, ils viennent pas chez nous du tout... ORTOFON v1

62) **Ten** **náš** **Vilk-a** nás dnes
ten.M.NOM.SG notre.M.NOM.SG NP-M.NOM.SG nous.ACC aujourd'hui

všechn-y předhon-í!
tout-ACC.PL dépasser-PRS.3SG

À présent, notre Vilka est plus rapide que nous tous ! (Mathesius, 1926 : 40)

¹⁴¹ Nous traduisons : „Lze myslím pokládat za obecně akceptovaný (resp. akceptovatelný) fakt, že lexému *ten* lze přisoudit trojí okruh funkcí: 1. deiktické (ukazovací) funkce denotační (identifikační); 2. kontextově-diskurzivní (anaforické a kataforické) funkce odkazovací; 3. expresivní funkce „přídátečné“. První dva okruhy funkcí jsou systémově centrální, třetí je systémově periferní.“

63) L1 : pak jsou holky tady z toho pokoj-e
après être.PRS.3PL fille-NOM.PL ici de ten.M.GEN.SG chambre-GEN.SG

L2 : tady z tohoble prvn-ího ? .. to
ici de tenhle.M.GEN.SG premier-M.GEN.SG ce.NOM

jsou všechno čtvrtáčk-y ?
être.PRS.3PL tout.N.NOM.SG élève de terminale-NOM.PL

L1 : no Elišk-a .. Míš-a a **ta** **Kamč-a**
MOD NP-F.NOM.SG NP-F.NOM.SG et ten.F.NOM.SG NP-F.NOM.SG

L1 : après, y a les filles de cette chambre-là

L2 : de cette première chambre ? ce sont toutes des filles de terminale ?

L1 : bah, Eliška... Míša et (cette) Kamča ORTOFON v1

64) to ne v Německ-u jed-u určitě cel-ou dob-u
MOD en Allemagne-LOC.SG roule-PRS.1SG certainement tout-F.ACC.SG temps-ACC.SG

po dálnic-i tak to není že jo problém
sur autoroute-LOC.SG MOD ce être.PRS.3SG.NEG MOD problème.NOM.SG

do **toho** **Hannover-u**
à ten.M.GEN.SG Hanovre-GEN.SG

en Allemagne, c'est sûr que je vais tout le temps rouler sur l'autoroute ; du coup, ça ne sera pas un problème d'aller à Hanovre ORTOFON v1

65) voni .. tu vobjednávk-u tím pádem v **tom**
eux.NOM ten.F.ACC.SG commande-ACC.SG du coup à ten.M.LOC.SG

Amazon-u jak jsem pochopi-l-a .. zřejmě ..
Amazon-LOC.SG comme AUX.PST.1SG comprendre-PTCP-SG.F apparemment

rozděl-í .. na víc část-í podle toho jak-ýho m-a-jí
diviser-FUT.3PL sur plus partie-GEN.PL selon cela.GEN quel-M.ACC.SG avoir-PRS-3PL

dodavatel-e ..
fournisseur.ACC.SG

du coup... comme je l'ai compris, moi, chez Amazon... ils vont apparemment... diviser... la commande en plusieurs parties, en fonction du fournisseur qu'ils ont... ORTOFON v1

Dans tous ces énoncés, le démonstratif introduit un nom dont l'identification ne recourt pas au contexte immédiat, comme c'est normalement le cas avec cette unité linguistique. Les substantifs introduits par TEN dans les exemples 60-65 peuvent, selon un premier tri, être répartis dans quatre catégories. En 60, le référent de *hlava* (*la tête*) est « sémantiquement » défini (cf. Carlier et De Mulder, 2006 et 2011 ; Himmelmann, 1997 ; Löbner, 1985 et 2011 ; Ortmann, 2014)¹⁴², car le nom

¹⁴² Les notions de « définitude sémantique » et « pragmatique » seront présentées de manière détaillée dans le Chapitre 6 (Section 6.1.3.4). Pour l'instant, contentons-nous de rappeler que contrairement à la définitude « pragmatique », établie moyennant une indexation du référent sur le contexte (situationnel, discursif ou mémoriel) immédiat, la définitude

exprime un concept dit « relationnel fonctionnel » (une personne a une seule tête) dont l'argument possessif est saturé par le lien associatif de propriété inaliénable liant la partie (la tête) au tout (la personne du locuteur). En 58, *lidi* (*les gens*) est un nom en emploi générique. En 62 et 63, TEN introduit un nom propre de personne. Dans la théorie de Löbner (1985 et 2011), les noms propres de personnes sont analysés comme des « concepts individuels », et relèvent donc de la définitude « sémantique ». En effet, leur sens instructionnel de désignateurs rigides (cf. Gary-Prieur, 1994 ; Jonasson, 1994 ; Kripke, 1980 ; Kleiber, 1991a) permet à l'interlocuteur d'identifier infailliblement leur porteur à l'intérieur d'un univers du discours plus ou moins large ou restreint. Kleiber résume la différence entre le démonstratif et le nom propre (et donc le paradoxe apparent de l'association des deux) comme suit : « Ses propriétés de désignateur rigide le conduisent à l'opposé du démonstratif en l'asservissant de toute contrainte de présence. » (1991a : 84) En 62, TEN apparaît en plus avec le possessif (*náš*), lui non plus n'ayant aucune incidence sur l'identification du nom propre (ces cumuls de mots fonctionnels seront examinés de plus près dans les Chapitres 6 et 7). Selon Gary-Prieur, le démonstratif serait ici dépourvu de fonction référentielle :

Combinés à un nom propre, il est clair que les démonstratifs [...] n'ont pas de fonction référentielle, l'identification du référent étant assuré par le nom lui-même. Ces démonstratifs servent dans ce cas à manifester les sentiments que le référent du nom propre inspire au locuteur. (2011 : 75)¹⁴³

Nous ne souscrivons pas entièrement à cette formulation pour la raison suivante : dire que le démonstratif n'a pas de fonction référentielle pourrait facilement laisser croire qu'il est dépourvu de référence tout court, ce qui n'est pas acceptable. Car s'il ne participe pas (directement) à l'identification du référent¹⁴⁴, cela ne veut pas du tout dire qu'il soit référentiellement vide. Comme l'explique Bonnard (1972) – que nous nous permettons ici de paraphraser –, le démonstratif effectue ici un renvoi déictique aux circonstances qui ont permis l'évaluation du référent, l'effet émotionnel surgissant précisément en tant que le référent est présenté sur fond de ces circonstances. Cette évaluation peut d'ailleurs être explicitée via la présence, au sein de l'expression démonstrative, d'un adjectif du type évaluatif. Comme le résume Kleiber, adhérant à l'analyse de Bonnard :

H. Bonnard (1972, p. 1207) indique de quel côté chercher la solution. Dans des expressions comme *Ce cher Pellegrin !*, *Cette pauvre Josette !* et *Ce sacripant de Victor !*, le « démonstratif rappelle, écrit-il, la réalité antérieurement constatée par la qualité qu'expriment les adjectifs *cher* et *pauvre*, ou le nom *sacripant* ». Autrement dit, l'énonciation de *ce cher Pellegrin* ou de *cette pauvre Josette* s'appuie sur une situation antérieure qui justifie la qualification opérée par *cher* et par *pauvre*. Il reste que l'interlocuteur doit aussi avoir accès à

« sémantique » « [...] suppose que le référent peut être identifié à partir du seul contenu descriptif du SN, compte tenu des associations stéréotypées qu'il entretient avec un cadre ou scénario à restituer. » (Carlier et De Mulder, 2006 : 102)

¹⁴³ Nous trouvons la même position chez Trovesi (2002 : 133), qui va jusqu'à refuser au démonstratif employé dans ces contextes son statut de démonstratif.

¹⁴⁴ C'est à la fin du Chapitre 7 (Section 7.3) que nous essaierons d'apporter quelques nuances à ce constat.

une telle situation et, puisqu'il s'agit d'une identification démonstrative, qu'il y ait dans la situation d'énonciation des éléments qui y conduisent. (Kleiber, 1991a : 94)

En français, la combinaison en question serait toujours à l'origine d'un effet d'affectivité, de familiarité ou de prise de distance avec le référent du nom propre (cf. Kleiber, 1991a et Jonasson 1994 et 1998a)¹⁴⁵. Pour ce qui est du tchèque, la situation est plus complexe, car, si ces effets sont également présents, il en existe d'autres. L'emploi du démonstratif en 62 s'explique, vu également la présence du possessif, certainement par l'intention du locuteur de marquer son affection vis-à-vis de l'enfant Vilka. En 63, c'est une interprétation en termes d'effet de distance qui semble la plus plausible : la locutrice utilise la simple forme du nom propre pour référer aux deux colocataires précédentes (*Eliška, Miša*), mais elle a aussitôt recours à la combinaison TEN + nom propre pour parler de la troisième (*ta Kamča*). Comme l'explique de nouveau Kleiber pour le français :

Ainsi le locuteur peut vouloir indiquer que la connaissance qu'il a du référent n'est pas parfaite ou totale, que celui-ci présente des aspects inconnus, troubles, mystérieux, etc. L'emploi du SN démonstratif *Ce + Nom propre* lui permet alors de souligner une telle distance là où l'emploi d'un nom propre non modifié irait dans la direction opposée, celle de la familiarité. (1991a : 92)

Malheureusement, il faudrait, afin de pouvoir trancher avec certitude au sujet de cet emploi, avoir accès à un contexte plus large. Or le nom propre *Kamča* fait son apparition au début de la conversation et il s'agit de sa seule occurrence au sein de celle-ci.

Mais voici maintenant un autre exemple, où l'effet produit par TEN ne rentre pas les deux cases susmentionnées :

66) L1 : když	by	byl	[Jindr-a] _i	voln-ej .	tak
si	AUX.COND.3SG	être.PTCP.SG.M	NP-M.NOM.SG	libre-M.NOM.SG	alors
	by	řek-l	[Jindr-ovi] _i		
	AUX.COND.3SG	dire-PTCP.SG.M	NP-M.DAT.SG		
L2 : no	protože	von	říka-l	jakoby že že	
MOD	parce que	lui.NOM	dire-PST.3SG.M	MOD que	
[<i>ten</i>	<i>Jindr-a</i>] _i ..	von	vo [něm] _i	mluvi-l	
ten.M.NOM.SG	NP-M.NOM.SG	lui.NOM	de lui.LOC	parler-PST.3SG.M	

L1 : si *Jindra* était libre il demanderait à *Jindra*

L2 : ouais parce que lui, il disait, genre, que *Jindra*... il a parlé de lui ORTOFON v1

¹⁴⁵ Nous laissons ici de côté les cas où le nom propre est utilisé dans un emploi dit « dénominatif » ainsi que tous ses autres usages moins habituels, tel le « fractionnement du référent du nom propre » ou les emplois métaphoriques et métonymiques. Pour l'explication de ces emplois, voir Kleiber 1991a et 2004. Ce qu'ils ont tous en commun, c'est le rejet complet ou la mise à l'écart du sens instructionnel (chercher le x appelé /ainsi/) au profit d'un sens dénominatif ou conceptuel.

Dans ce type d'emplois, TEN signale qu'il a déjà été question du référent auparavant. En 66, le locuteur rapporte les paroles d'une autre personne, qui, elle, a fait savoir que si un homme prénommé Jindra était libre, elle le contacterait pour lui demander un service. Comme il a déjà été question de Jindra avant, la répétition du nom propre s'accompagne de TEN. Comme l'explique Adamec (1983 : 157 et 165), qui parle à ce propos d'une « fonction de rappel » („připomínací funkce“), TEN joue ici le rôle d'un commentaire mémoriel condensé, d'une sorte de « connecteur ad-propositionnel » („adpropoziční konektor“) équivalant à une « adjonction parenthétique » („parentetický výraz“) du type *comme on en a parlé, comme tu le sais*, etc¹⁴⁶.

En 64 et 65 *supra*, nous sommes face à un TEN introduisant des noms à référence unique (*unica*). Comme les noms propres de personnes, cette catégorie appartient aux « concepts individuels », et donc « sémantiquement » définis. Néanmoins, il y a intérêt à distinguer les deux, car contrairement aux noms propres de personnes (et, avant tout, aux prénoms), dont l'unicité référentielle tend à se limiter à un univers du discours beaucoup plus restreint, les *unica* permettent de traquer le référent au sein de l'ensemble large d'une communauté de locuteurs qui, dans le cas des représentants les plus forts de la catégorie, s'identifie à l'échelle planétaire¹⁴⁷. Dans les énoncés de 64 et 65, vu le cadre général de la conversation, l'explication la plus plausible de l'usage de TEN est également celle d'un effet de rappel, comme en parle Adamec (1983). Pour 64, cette analyse est d'autant plus convaincante que le nom de la multinationale a déjà été mentionné en amont par la locutrice et, qui plus est, dans un contexte de rappel lui aussi :

67) (précédant l'énoncé 65)

dokonči-l	jsi	tu	vobjednávk-u	..	na
terminer-PTCP.SG.M	AUX.PST.2SG	ten.F.ACC.SG	commande-ACC.SG		sur
ten		Amazon ?			
ten.M.ACC.SG		Amazon.ACC.SG			

tu as déjà terminé la/cette commande... sur Amazon (comme on en a déjà parlé) ?

ORTOFON v1

¹⁴⁶ Šimík (2015) parle ici d'« anaphore discursive pragmatique » (“pragmatic discourse anaphora”). À certains égards, ces emplois de TEN ne sont pas sans rappeler l'emploi de *ISTE* en latin tardif lorsque celui-ci apparaît avec un référent « sémantiquement » défini dont il a déjà été question dans le discours (cf. Carlier et De Mulder, 2010).

¹⁴⁷ Certains auteurs (cf. Erbert, 1971 ; Witschko, 2012 et Hofherr, 2014) opèrent une différence entre ce qu'ils appellent « absolute uniqueness » et « situational uniqueness ». *Hanovre*, *Amazon* et *la Tour Eiffel* seraient des représentants de la première catégorie, contrairement aux termes comme *la gare* et *la mairie*, employés par les habitants d'une ville, ou *la salle de bain* et *la cuisine*, employés par les locataires d'un appartement. Ces derniers sont appelés « permanently established individual concepts » dans la terminologie de Gerland et Horn (2010). Notre propre acception du terme *unica* se limite à la notion de « absolute uniqueness ». Voir aussi la distinction de Hawkins (1978) entre « larger situational use based on general knowledge » et « large situational use based on specific knowledge ».

L'on remarquera également que le complexe nominal *ta objednávka* (la/cette commande) réfère de façon mémorielle. Dans les faits, les emplois « émotionnels de rappel » de TEN et ses emplois mémoriels (où TEN, rappelons-le, aide à identifier un référent) apparaissent souvent côte à côte (Adamec, 1983 : 165-168). Cette proximité n'a rien d'étonnant, car dans les deux cas, le locuteur en appelle à une mémoire partagée par l'interlocuteur et lui-même (voir pour cela la Section 7.2.3 du Chapitre 7). Néanmoins, dans les énoncés comme 67, il faut également reconnaître que de tels cumuls d'occurrences de TEN semblent être, au moins dans une certaine mesure, le produit d'un phénomène d'attraction, où les occurrences déjà présentes en appellent naturellement et presque mécaniquement d'autres¹⁴⁸. Qui plus est, nous admettons que le choix du terme commun d'« émotionnel » pour désigner tous ces cas de figure peut paraître un peu forcé – l'on pourrait par exemple se demander en quoi réside la supposée valeur émotionnelle des exemples 64, 65 et 66. En effet, il n'y est question ni d'affectivité ni de prise de distance avec le référent. Au vu de ce fait, nous reconnaissons que le choix du terme est en partie arbitraire et contestable. Si nous le maintenons tout de même, c'est pour deux raisons : d'abord, dans les emplois en 64, 65 et 66, TEN est bel et bien à l'origine d'un effet de connivence en ce qu'il en appelle à une connaissance privée et partagée au sujet du référent ; ensuite, ce choix est pratique, car il nous permet de rassembler sous la même étiquette tous les emplois de TEN où celui-ci semble ne pas jouer un rôle direct dans l'identification du référent. Plus tard, dans le Chapitre 7, nous entreprendrons cependant une analyse plus raffinée et plus nuancée de tous ces emplois. C'est également là que nous serons davantage attentif aux autres facteurs contribuant à la valeur « émotionnelle » de l'énoncé tels que la modalité de la phrase et le lexique, mais aussi certains phénomènes syntaxiques comme la séquence ON (IL) + TEN en emploi adnominal (cf. l'exemple 9 : **oni ty orchideje**)¹⁴⁹.

À notre connaissance, c'est précisément Adamec qui, dans son étude contrastive entre le tchèque et le russe (« České zájmeno *ten* a jeho ruské ekvivalenty », 1983) a développé l'analyse la plus fine des emplois « émotionnels » de TEN. Au départ, l'auteur sépare ce qu'il appelle les « fonctions d'identification » („funkce identifikační“) des « fonctions secondaires » („funkce sekundární“). Au sein des premières – considérées comme fondamentales –, il opère la distinction entre les emplois « déictiques », « anaphoriques » et « cataphoriques ». À l'intérieur des secondes, il distingue entre ce qu'il appelle, respectivement, la « fonction emphatique » („funkce zvýrazňovací“) et la « fonction de rappel » („funkce připomínací“). Quant à la « fonction emphatique », celle-ci

¹⁴⁸ Nous remercions Mme Nádvorníková d'avoir attiré notre attention sur ce fait. L'étude d'Adamec fournit une série d'exemples de ce type : *Milane, kdy už začneš dělat ten úkol do té matiky na ten čtvrtek?* (1983 : 157) Littéralement : **Milan, quand est-ce que tu vas enfin commencer à faire ce devoir de ces maths pour ce jeudi ?*

¹⁴⁹ Cvrček et al. (2015 [2010] : 175) considèrent la séquence ON + TEN comme une forme composée du démonstratif, typique du tchèque parlé.

connaît une sous-division en « variante d'émotionnalisation » („varianta emocionalizační“) et la « variante d'intensification » („varianta intenzifikační“). L'exemple 60 *supra*, qu'il emprunte à Mathesius (1926), est classé par lui dans la première. L'exemple suivant relèverait de la seconde :

- 68) At' *ten* *Jiříček* ne-vypadn-e *z toho* *okn-a!*
 MOD ten.M.NOM.SG NP.M.NOM.SG NEG-tomber-FUT.3SG de ten.N.GEN.SG fenêtre-GEN.SG
Fais attention à Jiříček, pour qu'il ne tombe pas de la fenêtre ! (1983 : 156)

La différence entre les deux est de nature plutôt subtile et pas toujours facile à discerner. Tandis que la « variante d'émotionnalisation » sert à « exprimer des émotions diverses du locuteur, telles que l'admiration, l'étonnement, la satisfaction, l'indignation, le désir, etc. »¹⁵⁰, la « variante d'intensification » serait motivée par « l'effort d'attirer davantage l'attention, d'amplifier l'effet de l'information communiquée, de l'urgence de l'appel, de l'avertissement, etc., [le tout étant] souvent également sous-tendu par des émotions, telles que l'appréhension, la peur, l'impatience, etc. » (1983 : 156)¹⁵¹ Contrairement à la « variante d'émotionnalisation », entièrement centrée sur la subjectivité du locuteur, il semblerait que la « variante d'intensification » serait davantage orientée sur l'interlocuteur dans la mesure où le locuteur y viserait davantage à produire un effet sur ce dernier, voire même à le pousser à agir : en 68, le locuteur incite l'interlocuteur à entreprendre quelque chose pour empêcher l'enfant de tomber de la fenêtre)¹⁵².

Un autre aspect intéressant de l'approche d'Adamec, c'est le lien de filiation que l'auteur établit entre les fonctions premières d'identification et les fonctions secondaires : « De même que la fonction emphatique a évolué à partir de la fonction déictique, de même la fonction de rappel a progressivement évolué à partir de la fonction anaphorique. » (1983 : 165)¹⁵³ La motivation de ce parti-pris n'est pas difficile à voir : pour ce qui est de la « fonction emphatique », le référent est souvent présent dans la situation d'énonciation (voir les exemples 60 et 68) ; concernant la « fonction de rappel », sa contiguïté (à la fois en termes de sens et d'apparition dans le discours) avec les emplois exophoriques mémoriels – eux-mêmes parfois considérés comme un cas particulier de l'anaphore¹⁵⁴ – de TEN vient ici d'être relevée. Pourtant, cette analyse présente, selon nous, deux failles. D'abord, il n'est pas vrai que dans tous les emplois dits « emphatiques » par l'auteur, le référent soit présent dans la situation d'énonciation. Il en cite lui-même un bon nombre où cela n'est pas le cas, notamment lorsqu'il s'agit de noms en emploi générique (cf. 70), mais pas uniquement :

¹⁵⁰ Nous traduisons : „[...] vyjadřování různých emocí mluvčího, jako obdiv, překvapení, spokojenost, rozhořčení, touha apod.“

¹⁵¹ Idem : „[...] snaha o upoutání větší pozornosti, zvyšování působivosti sdělení, naléhavosti výzvy, varování apod., spojené ovšem často rovněž s určitými emocemi, jako obava, strach, netrpělivost apod.“

¹⁵² Nous remercions Mme Guillot pour cette remarque.

¹⁵³ Nous traduisons : „Tak jako se zvýrazňovací funkce vyvinula z funkce deiktické, připomínací funkce se postupně vyvinula z funkce anaforické.“

¹⁵⁴ Voir le terme de « pseudo-anaphore », employé par Berger (1993).

- 69) *Ta Ivet-a, ta mi už jd-e*
 ten.F.NOM.SG NP-F.NOM.SG celui-ci.F.NOM.SG moi.DAT déjà aller-PRS.3SG
 na nerv-y.
 sur nerf-ACC.PL

(*Franchement*), (*cet*) *Iveta*, elle commence à me casser les pieds. (Adamec, 1983 : 162)

- 70) *Ty na ty peníz-e jednou zdechn-e-š.*
 toi.NOM. sur ten.M.ACC.PL argent-ACC.PL tantum une fois crever-FUT-2SG
Toi, ton argent, tu finiras par en crever. (Adamec, 1983 : 165)

Ensuite, il nous semble qu'Adamec colle un peu trop vite l'étiquette de « fonction de rappel » à tous les emplois de TEN qui en appellent, chez les participants à l'interlocution, à une mémoire commune, y compris ceux où TEN participe directement à l'identification d'un référent¹⁵⁵. Par conséquent, tous les emplois mémoriels (les « recognitional uses » de Himmelmann (1996)) de TEN ressortiraient, *in fine*, à la catégorie des emplois « émotionnels », conclusion que nous ne pouvons accepter, car nous tenons à séparer rigoureusement les emplois « émotionnels » des emplois « désignatifs » (ceux où l'identification du référent par l'interlocuteur reste la raison première de l'usage du démonstratif).

3.6 Résumé

Dans ce chapitre consacré au démonstratif en tchèque, nous croyons avoir suffisamment préparé le terrain pour de futures analyses plus poussées. Ces analyses figureront notamment dans les Chapitres 6 et 7. Nous avons commencé par esquisser l'organisation du système du démonstratif en tchèque contemporain, accordant une attention particulière aussi bien aux oppositions en matière de structuration de l'espace qu'en matières diatopique et diastratique. A également été présentée une série de comparaisons succinctes effectuées sur des corpus du CNT et qui, nous l'espérons, ont permis au lecteur d'acquérir une certaine image de la situation des formes existantes dans différents types de textes (les chiffres ont été soumis au test de la fonction de vraisemblance, qui, dans la majorité des cas, a prouvé leur significativité statistique). Nous nous sommes ensuite attardé sur la question du modèle d'analyse syntaxique à appliquer au démonstratif adnominal tchèque et, par extension, à tous les éléments fonctionnels/actualisateurs du nom. Nous avons pu voir, les travaux de Himmelmann (1997) et de Diessel (1999) à l'appui, que le tchèque fait partie des systèmes où les démonstratifs adnominaux se comportent comme des pronoms apposés/juxtaposés au nom, à la différence de ceux où ils se comportent comme des déterminants. Ont également été abordés certains emplois particuliers

¹⁵⁵ En effet, il admet l'existence d'emplois « de rappel » où TEN joue un rôle incontestable dans l'identification du référent. (1983 : 168)

du démonstratif TEN en usage pronominal dont le sens et la fonction semblent, de nos jours, assez éloignés de ceux du démonstratif d'origine. À cette occasion, la question de savoir dans quelle mesure il semble juste de parler de polysémie, voire d'homonymie, a été posée.

Dans le second volet, nous nous sommes employé à retracer les origines du système actuel du démonstratif tchèque. Pour ce faire, nous avons dû remonter à des époques antérieures, soit celles du vieux tchèque et du vieux slave. Les faits suggèrent qu'au cours de son histoire, le système aurait subi une réfection en passant d'une configuration orientée par rapport aux personnes de l'énonciation à une configuration structurée autour de la distance du référent par rapport au centre déictique. Certaines tendances orientant l'évolution du système à l'heure actuelle ont été aussi signalées. Le troisième volet a été consacré aux différents types d'emploi du démonstratif. Comme le présent travail traite avant tout de TEN – la forme de loin la plus fréquente –, c'est cette forme qui était au centre de notre propos. Dans le cadre de cette introduction, nous avons tenu à relever les emplois où TEN encourage une analyse en article défini en devenir. Rappelons que tous ces emplois feront l'objet d'analyses détaillées dans le Chapitre 6 (Section 6.2). C'était également l'occasion pour nous de relever un certain nombre d'emplois hors du commun, dont notamment les emplois « métalinguistiques », indéfinis et ceux qui annoncent une relative « établissante » du type « classificatoire ». Le rôle de TEN au sein de la substantivation d'un adjectif en l'absence de nom, présenté dans la Section 3.4, sera lui aussi davantage développé en rapport avec la grammaticalisation de TEN en article défini, car nos recherches ont montré que l'absence de nom contribue à accélérer ce processus.

Enfin, le dernier volet a été réservé au sujet des emplois « émotionnels ». Nous y avons insisté sur les singularités sémantico-référentielles au travers desquelles ce type d'emploi se démarque des autres. Dans le Chapitre 7, nous poursuivrons l'examen de ce phénomène assez complexe en nous appuyant à la fois sur un relevé quantitatif et des analyses plus poussées de nos données. C'est également là que nous nous efforcerons d'en proposer un traitement unifié, articulé autour des concepts d'« implicature conversationnelle » de Grice (1975) et de « sous-entendu » de Ducrot (1980).

4 Le démonstratif en français

Ce quatrième chapitre s'inscrivant dans la continuité du chapitre précédent, son but est de présenter le paradigme des démonstratifs nominaux en français. Ainsi, le plan adopté reprendra globalement celui du chapitre précédent. Nous commencerons par présenter le paradigme du démonstratif français en synchronie, tout en le décrivant sur les plans morphophonologique et syntaxique et quant à sa capacité à exprimer la distance du référent. Ici aussi, en plus de commenter les différentes formes, nous réitérerons la question du modèle d'analyse syntaxique à faire valoir, en nous référant essentiellement aux travaux de Himmelmann (1997) et de Diessel (1999). Suivra un rappel synthétique du processus complexe que fut la constitution du paradigme en diachronie. Nous examinerons ensuite la thèse de certains auteurs enclins à rapprocher le déterminant démonstratif du français moderne de l'article défini. Nous nous arrêterons brièvement sur la constitution du SN en français, en nous appuyant notamment sur le travail de Goyens (1994). Comme dans le chapitre précédent, nous étudierons les rapports entre les formes du déterminant démonstratif et les régimes d'indexicalité. Dans un dernier temps, un mot sera dit au sujet des emplois « émotionnels » du démonstratif français. L'ensemble de l'exposition sera étayé par des exemples authentiques, puisés, avant tout, dans le CEFC.

4.1 Le démonstratif français en synchronie

Par contraste avec le système du tchèque, le paradigme du démonstratif français se présente comme plus transparent et plus facile à circonscrire. Au départ, pour résumer la situation d'une manière certes approximative et simplificatrice, nous constaterons que le français moderne possède un système démonstratif articulé autour de l'axe de la distance du référent par rapport au locuteur, distance qui connaît trois valeurs (voir la Section 1.3.1 du Chapitre 1). Un peu plus tard dans le présent chapitre, nous reviendrons cependant sur ce constat, qui n'est pas à l'abri de controverses. Contrairement au tchèque, la situation en français ne semble pas présenter d'importantes dichotomies articulées autour des axes diastratique et diatopique¹.

¹ Seule variable diastratique patente : l'éviction de la particule localisatrice *-ci* par la particule localisatrice *-là* (voir la Section 4.1.3 *infra*).

4.1.1 Remarques morphosyntaxiques et sémantiques nécessaires

Le système du français actuel se démarque par le fait qu'il contient deux séries de formes dont la singularité est que les unes sont employées exclusivement en emploi adnominal et les autres, exclusivement en emploi pronominal :

les formes du paradigme pronominal : *ce, ceci, cela, ça, celui(-ci/-là), celle(-ci/-là), ceux(-ci/-là), celles(-ci/-là)*

et les formes du paradigme adnominal : *ce, cet, cette, ces*.

Comme nous l'avons déjà souligné dans le Chapitre 1 (Section 1.3.2) en nous référant à la typologie de Diessel (1999), au sein du paradigme pronominal, il faut distinguer trois cas de figure. Les formes *celui(-ci/-là), celle(-ci/-là), ceux(-ci/-là), celles(-ci/-là)* sont morphologiquement marquées pour le genre – masculin ou féminin – et le nombre – singulier ou pluriel. Ces deux catégories sont, rappelons-le, rangées parmi les traits « qualitatifs » par Diessel (1999). L'association avec les particules localisatrices *-ci/-là* permet également aux formes d'être marquées pour la distance. Les formes *ceci, cela* et *ça* sont morphologiquement marquées pour la seule distance². Sémantiquement, pourtant, elles sont également marquées pour le genre – neutre –, le nombre – singulier –, ainsi que pour la nature ontologique du référent en ce qu'elles servent, avant tout, à référer à des entités inanimées³. Et, enfin, la forme neutre *ce* doit être tenue à part, car elle apparaît d'abord en emploi « identificatoire » (Diessel, 1999), comme sujet clitique d'un tour présentatif avec le verbe copule ETRE⁴ :

- 1) **Ce** sont mes parents.
- 2) [**Paul**]_i, [**c'**]_i est un monstre. (Kleiber, 1984a : 71)

Outre cet emploi, *ce* fonctionne également comme un pronom substitut en fonction d'antécédent inanimé d'une relative périphrastique ou d'une interrogative indirecte partielle portant sur un complément d'objet direct :

- 3) **Ce** qui se passe ne m'intéresse pas.
- 4) Je ne comprends pas **ce** que tu veux dire.

² Marquage qu'il faudrait peut-être relativiser pour *ça*, forme réduite car issue de *cela* par syncope et qui, selon Harris (1977), voyant en elle la variante disjointe du pronom *ce*, serait une forme non-marquée pour la distance.

³ « Les formes composées neutres *ceci, cela* et *ça* (ce dernier, morphologiquement simple mais issu de *cela*, se comporte comme une forme composée) servent à désigner déictiquement des référents non catégorisés [...], voire à décatégoriser péjorativement un référent en lui refusant sa dénomination usuelle [...] » (Riegel, Pellat et Rioul, 2016 [1994] : 377).

⁴ Riegel, Pellat et Rioul (2016 [1994] : 377) de préciser que le verbe ETRE peut, le cas échéant, être modalisé par les verbes POUVOIR et DEVOIR (*Ce devrait être facile*).

Dans ce type de structures et dans d'autres encore, *ce* ne serait plus, à en croire Harris (1977) et Diessel (1999), un vrai démonstratif :

Ce does not generally function as a demonstrative. In fact, in most instances *ce* is non-deictic and functions either as a third person pronoun or as an expletive [...]. However, in those cases in which *ce* has a deictic interpretation, it can be analyzed as a demonstrative identifier. (Diessel, 1999 : 87-88)

Le clivage ferait partie de ces emplois selon Diessel (1999 : 149-150), emplois où *ce* serait devenu un pur « explétif », à savoir une forme vide et dépourvue de référence, comparable au pronom neutre *il* servant de sujet grammatical dans les phrases construites autour de verbes a-valents (**Il pleut**)⁵. Harris (1977 : 257) adhère à un point de vue semblable :

Ce (< OFr *ço* < VL ECCE HOC), despite its evident demonstrative origin, cannot be labelled as such even in OFr within the terminology adopted here, since its use involves no opposition of proximity. It is used in effect (i) as the 'neuter' equivalent of the anaphoric or determinative pronouns discussed above (.,je ne sais pas *ce* que vous voulez dire"), a usage which of course survives, and (ii) as a third person singular 'personal' pronoun.

Contrairement à Diessel, qui reconnaît tout de même à la forme des emplois démonstratifs, Harris considère qu'en dehors des structures périphrastiques susmentionnées, *ce* se comporte effectivement comme un pronom neutre de la 3^{ème} personne, entretenant ainsi des rapports paradigmatiques avec les pronoms personnels *il* et *elle*. Son argumentation s'appuie notamment sur l'hypothèse selon laquelle, au cours de l'évolution du français, *ce* aurait remplacé le pronom *il* à valeur neutre : **C'est vrai** de la langue moderne vs **Il est vrai** de la langue classique. Nous trouvons cet argument fallacieux : ce n'est pas *ce* qui serait devenu pronom personnel neutre de la 3^{ème} personne en remplaçant *il* dans ce type de structures ; au contraire : c'est le pronom *il* qui connaissait jadis une valeur de démonstratif (faible), lui permettant de se prêter à ce type d'emplois (cf. les pronom *it* et *es* en anglais et en allemand modernes). Cette valeur n'a, du reste, rien de surprenant, étant donné l'évolution de IL à partir de ILLE. Plus tard, IL a cependant perdu cette valeur en français pour devenir un pur pronom personnel, et c'est à ce moment-là qu'il fut intégralement remplacé, dans ce type de structures (**Il est vrai**), par *ce*⁶.

De plus, Harris dresse un parallèle explicite entre cette prétendue évolution et une autre hypothèse, selon laquelle le paradigme du déterminant démonstratif CE aurait remplacé celui de l'article défini LE, devenant ainsi lui-même article défini une fois que LE serait devenu un pur marqueur morphologique du nom. Cette conclusion est également à rejeter selon nous (voir la Section 1.3.2 du Chapitre 1 et la Section 4.3 du présent chapitre).

⁵ De nos jours, cette valeur d'explétif s'étend aussi, dans la langue parlée informelle, à des emplois qui, il y a encore quelques années, étaient réservés à *il* : **C'est quelle heure ?**

⁶ Il a cependant gardé cette valeur de démonstratif faible dans sa fonction de pronom neutre COD : **Je le sais** ≈ **Je sais cela**.

À cette analyse, il faut, par ailleurs, opposer le fait que *ce* a toujours la capacité à fonctionner, dans le registre soutenu, comme un pronom démonstratif neutre indépendant (c'est-à-dire en dehors des constructions présentatives et périphrastiques) possédant une référence anaphorique résomptive :

- 5) Plus tard dans ce travail, des données statistiques portant exclusivement sur notre corpus seront présentées, et *ce* en comparaison avec la situation dans les textes écrits.

Quant aux formes du paradigme CELUI, lorsque celles-ci sont privées de la particule localisatrice *-ci* ou *-là*, elles ne peuvent être utilisées que comme des « pronoms déterminatifs » (cf. Harris ; 1977 : 257 ; également appelés « substituts » ou « proformes » par Riegel, Pellat et Rioul, 2016 [1994] : 358⁷), à savoir comme des supports pour une expansion caractérisante à droite, le plus souvent une proposition subordonnée relative périphrastique ou un syntagme prépositionnel⁸ :

- 6) **Ceux** qui fument nuisent à leur santé.

- 7) Les réponses sont toutes arrivées sauf **celle** de Jean. (Riegel, Pellat et Rioul, 2016 [1994] : 358)

Les formes adnominales sont morphologiquement marquées pour le genre – masculin ou féminin – et le nombre – singulier ou pluriel. Complétées par les particules localisatrices, elles peuvent aussi l'être pour la distance. Voici la synthèse de l'ensemble du paradigme proposée par Bonnard (1972) :

⁷ Ces derniers accentuent le fait que de tels pronoms n'ont pas la distribution d'un SN, car c'est le tout composé du pronom et de ses compléments qui a cette distribution (2016 [1994] : 358). Comme nous le voyons dans l'exemple 7, la forme peut se prêter à un fonctionnement anaphorique indirect, reposant sur la reprise du sens lexical d'un terme introduit par le discours antérieur. Au sujet des déterminatifs en français, voir également Pierrard, 1989, 1990 et 1991.

⁸ Au sujet des démonstratifs déterminatifs dans d'autres langues, voir par exemple Quirk, 1972 : 217 ; Harris, 1977 : 257 ; Himmelmann, 1997 : 77-80 ou Diessel, 1999 : 135-137. Comme le résume Himmelmann (1997 : 78) : „Das Demonstrativum scheint [...] eine rein grammatische Platzhalterfunktion zu haben, zumindest lässt sich seine Verwendung in dieser Konstruktion m.E. keiner der typischen Gebrauchsweisen für Demonstrativa zuordnen.“ Selon Harris, les formes en question ne seraient plus du tout des démonstratifs.

	ADJECTIFS		PRONOMS	
	MASCULINS ET FÉMININS		NEUTRES	
FORMES SIMPLES	<i>ce (cet), cette ces</i>	<i>celui, celle ceux, celles</i>	<i>ce</i>	
FORMES COMPOSÉES	<i>ce ...-ci (cet ...-ci) cette ...-ci ces ... -ci ce ... -là (cet ... -là) cette ...-là ces... -là</i>	<i>celui-ci, celle-ci ceux-ci, celles-ci celui-là, celle-là, ceux-là, celles-là</i>	<i>ceci cela</i>	

Tableau 4.1 : Système du démonstratif en français (Bonnard, 1972 : 1206)

S’agissant des plans morphologique et phonétique, une série de remarques s’imposent. D’abord, quelle est la nature de *-ci* et *-là* ? Riegel, Pellat et Rioul (2016 [1994] : 375) parlent de « particules localisatrices » et Diessel (1999 : 88), lui, de « particules déictiques », termes que nous reprenons ici. En effet, le terme de « particule » nous semble plus approprié que celui d’« adverbe » dans la mesure où il s’agit d’un élément fortement dépendant. En même temps – et contrairement à ce que pourrait faire croire la morphosyntaxe –, il est important d’explicitier qu’il ne s’agit pas d’un enclitique, car c’est justement la particule qui, sur le plan phonétique, porte l’accent du groupe accentuel qu’est le SN contenant le démonstratif. Les formes *ceci* et *cela* font exception dans la mesure où la particule y a fusionné avec la forme du démonstratif, devenant ainsi un suffixe⁹.

En situation normale, les formes du paradigme des déterminants *ce*, *cet*, *cette* et *ces* ne sont donc pas accentuées, l’accent étant porté soit par le nom qui suit soit par la particule déictique *-ci* ou *-là*. L’un des rares contextes où ces déterminants peuvent l’être sont ceux où la réalisation de la forme à l’oral est suivie d’une interruption involontaire du discours, due à des hésitations du locuteur, son incapacité à trouver le terme juste, etc¹⁰. En voici un exemple :

- 8) C'est donné euh et puis de enfin tu dois pouvoir avoir **ce ce ce** voyage tu dois pouvoir pour l'Espagne c'est comme si tu partais de de Lyon pour euh CEFC

⁹ Ce qui ne change rien au fait, bien entendu, que *-ci* et *-là* continuent à être accentués.

¹⁰ Ces cas ont été annotés dans notre corpus pour le français parlé informel (voir la Section 2.2.1 du Chapitre 2).

4.1.2 Les démonstratifs adnominaux français sont des déterminants

Sur le plan de l'analyse syntaxique, il n'y a aucun doute sur le fait que les formes du paradigme CE (*ce, cet, cette* et *ces*) sont des déterminants. Rappelons ici le raisonnement qui justifie ce point de vue, et que l'on retrouve à la fois chez Himmelmann (1997) et chez Diessel (1999) :

In all den Fällen, in denen es genügend empirische Evidenz für die Annahme einer Kategorie *Determinator* gibt (z.B. im Deutschen, Englischen oder den romanischen Sprachen), gibt es auch immer mindestens ein grammatikalisches D-Element (einen Definitartikel oder einen spezifischen Artikel). Das heißt, es sind mir keine Fälle bekannt, wo z.B. ein Indefinitartikel und Possessiva oder Demonstrativa und Quantoren sich morphosyntaktisch so ähnlich verhalten, dass sie unter eine morphosyntaktische Kategorie subsumiert werden können, ohne dass in der betreffenden Sprache nicht auch mindestens ein grammatikalisches D-Element existierte. (Himmelmann, 1997 : 133)

Système linguistique possédant un paradigme d'articles définis, le français moderne est doté d'une catégorie d'éléments fonctionnels (Abney, 1987) dont le paradigme CE fait partie et qui respectent tous une série de critères en vertu desquels ils peuvent être considérés comme des déterminants. La situation est ainsi très semblable à celle de l'anglais, décrite par Diessel (1999), à ceci près que dans cette langue, le paradigme des déterminants démonstratifs et celui des pronoms démonstratifs sont formellement identiques :

Though articles and demonstratives are not always members of the same category, there is good evidence that English *the* and *this/that* have the same categorial status. English is one of the few languages in my sample in which adnominal demonstratives are in paradigmatic relationship with articles, possessives and other noun operators (e.g. *every*), which share a number of syntactic features. (Diessel, 1999 : 65)¹¹

De même que son homologue anglais, le démonstratif adnominal français entretient des rapports visibles avec d'autres éléments employés comme déterminants, dont notamment : l'article défini – avec ce dernier, il affiche, en plus, une forte ressemblance formelle (cf. Bonnard, 1972 ; Wilmet, 1986 ; Marchello-Nizia, 1995), l'article indéfini (y compris ses formes partitives *du, de la* et *des*, employées, respectivement, avec les noms indénombrables et dénombrables pluriels), l'article *de* d'extensité nulle¹², les quantificateurs universels TOUT et *chaque*, les quantificateurs existentiels comme *quelques, plusieurs* ou *beaucoup de*, le déterminant interrogatif et exclamatif QUEL et le déterminant possessif¹³.

¹¹ Voir également l'étude de Carlier et Combettes (2015).

¹² Wilmet 1986 et 2010 [1997].

¹³ Il convient cependant de rappeler ici que chez certains auteurs, le terme de déterminant est étendu à tous les éléments entourant la tête nominale (y compris donc aux épithètes, compléments du nom et propositions subordonnées) et formant avec celle-ci un SN. Tel est notamment la position prise par Wilmet (1986) et dont s'est beaucoup inspirée Goyens (1994). Même si nous comprenons les motivations se trouvant derrière ce choix, nous ne l'adopterons pas dans ce travail pour des raisons pratiques, car une telle conception des déterminants abolit toute distinction entre, d'une part, la classe des déterminants telle qu'elle est traditionnellement définie et telle qu'elle apparaît uniquement dans certaines langues et, d'autre part, tous les autres modificateurs du nom. Or cette distinction s'avère cruciale pour la comparaison de nos deux langues.

Le cas de ce dernier est particulièrement intéressant. Grâce à son alignement sur l'article défini, le déterminant possessif français est également porteur du présupposé d'unicité (de maximalité), contrairement au possessif adnominal tchèque et, en général, celui d'autres langues sans articles définis. Un constat typologique intéressant peut à cet égard être formulé, à savoir que dans les langues à articles définis, le possessif est souvent – mais pas toujours¹⁴ – un déterminant portant le présupposé de maximalité relatif au concept dénoté par le nom, tout en étant aussi en distribution complémentaire avec le défini et d'autres déterminants. Outre le français, l'on retrouve la même situation par exemple en anglais. Un exemple de langue défiant cette tendance est l'italien, langue à articles définis où le possessif en emploi adnominal se comporte davantage, sur le plan syntaxique, comme un adjectif, car il apparaît couramment avec un représentant de la classe des déterminants au sein d'un SN¹⁵, ainsi qu'en position d'attribut :

- 9) IT : la mi-a cas-a
 le.F.NOM.SG mien-F.NOM.SG maison-NOM.SG
 vs FR : *ma maison*
- 10) IT : Quest-a cas-a è mi-a.
 ce-F.NOM.SG maison- NOM.SG être.PRS.3SG mien- F.NOM.SG
 vs FR : *Cette maison est à moi.*

La situation de l'italien d'aujourd'hui rappelle fortement celle de l'ancien français, avec ses formes adjectivales du possessif, elles aussi couramment précédées, au sein d'un SN, par un déterminant :

- 11) [...] si salvarai eo *cist meon* fradre Karlo [...]
 ? *je secourrai ce mien frère Charles* Les serments de Strasbourg

En français moderne, ce phénomène survit dans la forme du paradigme possessif adjectival MIEN, de loin le plus souvent employé dans les SN possessifs à fonctionnement majoritairement anaphorique indirect et issus, en diachronie, de l'ellipse de la tête nominale. Remarquons, par ailleurs,

¹⁴ Ainsi, l'allemand possède deux paradigmes du possessif : le paradigme déterminatif et le paradigme adjectival/pronominal. Le premier – plus tenu sur le plan morphologique – n'est employé que pour introduire un nom dans le discours et il est en distribution complémentaire avec les autres déterminant (*mein Haus* vs **dies mein Haus*), tandis que le second peut apparaître a) avec certains déterminants (*diese meine Träume* – tour archaïque/livresque), mais le plus souvent en emplois anaphoriques avec ellipse du nom et précédée de l'article défini : *Die meine/meinige ist besser*, b) en position attributive (*Dieses Haus ist meins*) et c) comme pronom indépendant (*Meiner ist schöner*). En anglais, la forme forte du possessif (*mine*), outre qu'elle fonctionne comme un pronom indépendant (*Mine is better than yours*), peut également être employée en attribut (*This house is mine*), mais jamais en emploi adnominal.

¹⁵ L'exception la plus saillante, c'est la présence du possessif avec les termes de parenté (*mia mamma*), où le possessif se comporte comme un déterminant. S'agissant des emplois du type *la mia casa*, il serait également possible de parler de « post-déterminant ».

que ces SN sont quasiment toujours introduits par l'article défini¹⁶, contrairement au tchèque, où un mot fonctionnel indéfini peut apparaître :

12) Je trouve que mon père est trop sévère, **le tien** est plus cool.

*Si ta chemise est sale, prends *une mienne*.

vs CZ : Pokud m-á-š	špinav-ou	košil-i,	vez-mi	si
si avoir-PRS-2SG	sale-F.ACC.SG	chemise-ACC.SG	prendre-IMP.2SG	REFL
<i>nějak-ou</i>	<i>moj-í.</i>			
quelque-F.ACC.SG	mien-F.ACC.SG			

Même si des auteurs (cf. Kleiber et Bonnard) utilisent pour le démonstratif français en emploi adnominal le terme d'« adjectif », au moins trois arguments parlent en faveur du terme de « déterminant ». D'abord, l'argument distributionnel : les formes sont en distribution complémentaire avec d'autres éléments fonctionnels occupant la position de déterminant central ("the central determiner slot"), évoqués *supra*. Les seuls cas où cette règle distributionnelle ne s'applique pas, ce sont les combinaisons de CE avec certains éléments fonctionnels pouvant agir comme pré-déterminants ou post-déterminants, tels que le quantificateur universel TOUT et les numéraux¹⁷ :

13) je n'ai prévu qu'un guide parce que tu as *Christina* et *Emeline* qui forment un groupe qui s'occupe par exemple de la villa des Mystères et de *toute cette partie-là* CEFC

14) ça aussi ça donne un caractère particulier à *ces deux visages* CEFC

Inutile de s’attarder ici sur le fait que cette restriction ne s’applique pas aux adjectifs, qui peuvent, du moins théoriquement, se cumuler à volonté au sein d’un SN. Le deuxième argument, nous allons l’appeler « syntaxique » (appellation qui relève de notre choix personnel). En réalité, il s’agit d’un argument double. D’abord, CE a une position fixe au sein du SN, occupant toujours son extrémité gauche (voir à ce sujet les remarques de Himmelmann (1997), déjà présentées dans la Section 3.1.3). Ensuite, contrairement aux adjectifs, le démonstratif adnominal français ne peut occuper la position syntaxique d’attribut¹⁸:

15) Pierre est *gentil*.

vs *Pierre est *ce*.

¹⁶ Exception faite du tour *faire des siennes*. Mis à part ces emplois, de loin les plus courants, le paradigme adjectival du possessif survit dans deux autres emplois en marge du système. Le premier est un emploi adnominal au sein d'un SN dont la tête nominale est exprimée (l'exemple 11) et le second est un emploi attributif : *La Commission pourrait faire sienne cette recommandation et prendre des mesures pour la faire appliquer.* (MultiUN : <http://opus.lingfil.uu.se/MultiUN.php>.)

¹⁷ Ceux-ci peuvent aussi être considérés comme un type d'adjectifs (voir *infra*).

¹⁸ Un pronom démonstratif, lui, peut occuper la position attributive, mais il continue à s'y comporter, bien évidemment, comme un SN et non pas comme un adjectif (*C'est ça*).

Et, enfin, un troisième critère, que nous appellerons ici « sémantique » : contrairement aux adjectifs, exprimant le plus couramment un sens qualificatif (*beau, gentil*) ou relationnel (*aérien, aquatique*), le démonstratif adnominal ne saurait être rangé dans aucun de ces deux tiroirs¹⁹.

4.1.3 Le système du démonstratif français et l'expression de la distance

Regardons à présent de plus près le système des démonstratifs français sur le plan de l'expression de la distance. Vu l'orientation de notre travail, nous nous limiterons ici au paradigme des formes employées comme déterminants. À première vue, nous sommes en présence de trois formes qui partagent l'espace déictique d'une manière symétrique (voir le Tableau 4.1 *supra*, emprunté à Bonnard (1972)). En réalité, cette symétrie n'est qu'apparente. D'abord, parce que dans les faits, c'est la forme CE toute seule, dépourvue de particule localisatrice, qui est le plus couramment utilisée. En effet, c'est ce que confirment les données de notre corpus CEFC, où seuls 202 des 1 000 SN démonstratifs annotés contiennent une particule localisatrice (voir la Section 5.2.2.1 du Chapitre 5). Mais aussitôt ce premier chiffre avancé, il en faut minimiser l'importance dans la mesure où il contient tous les types d'emploi du démonstratif, y compris des emplois endophoriques, mémoriels et exophoriques temporels, où les notions de distance et de centre déictique sont par définition plus problématiques (voir la Section 1.3.1). Pour cette raison, limitons-nous ici uniquement aux occurrences exophoriques situationnelles. Celle-ci sont au nombre de 93, mais seules 25 contiennent *-ci* (5 occurrences) ou *-là* (20 occurrences). Ce constat (même si, il faut le reconnaître, il ne s'appuie que sur un nombre assez limité d'occurrences), est tout à fait en phase à la fois avec ce que dit sur ce sujet la littérature (cf. Anderson et Keenan, 1985) et avec ce que nous observons tous les jours en tant que personne vivant depuis plus de cinq ans en France. Ainsi, dans une grande partie des cas, la forme CE toute seule s'avère aujourd'hui tout à fait suffisante pour référer à des objets indépendamment de leur distance par rapport au locuteur, le système du français moderne étant caractérisé par une neutralisation considérable de la notion de distance spatiale, si bien qu'Anderson et Keenan (1985 : 280) envisagent la possibilité de le considérer comme un système qui ne contient qu'un seul terme déictique :

Another possible example of a 'one-term' deictic system is French *ce* (/cette/cet), which does not encode any sort of distance distinction.

¹⁹ Force est de reconnaître que certains adjectifs échappent à ce classement binaire, tels les adjectifs *actuel, ancien, futur, précédent, suivant*, etc. De plus, ces adjectifs – appelés « adjectifs du troisième type » par Schnedecker (2002) – sont, eux aussi, très souvent bannis de la fonction attribut. On peut également ranger dans cette classe d'adjectifs les numéraux cardinaux et ordinaux précédés par un déterminant central (Riegel, Pellat et Rioul, 2016 [1994] : 635) : *le premier élève, ces deux livres*.

Toutefois, une seconde raison est également à l'origine de la remise en cause de la prétendue symétrie du système. Cette raison est à chercher du côté de la fréquence des formes avec *-ci* et *-là*. Là aussi, contrairement à ce qu'un premier aperçu pourrait nous laisser croire, nous sommes face à une situation de dissymétrie abyssale, car les formes en *-là* dépassent de loin, dans la langue parlée informelle actuelle, celles en *-ci*. Cette tendance est aujourd'hui bien connue²⁰ et même reconnue par des sources normatives comme la grammaire de Riegel, Pellat et Rioul (2016 [1994] : 376-377) :

En emploi contrastif, les formes en *-ci* sont censées renvoyer à ce qui est le plus proche dans l'espace référentiel ou dans le texte et la forme en *-là* à ce qui est le plus éloigné. Mais cette opposition n'est pas toujours respectée dans l'usage contemporain qui n'emploie plus guère les formes en *-ci*.

Ce constat est conforté à la fois par nos données du corpus CEFC – seuls 15 des 204 SN démonstratifs dotés d'une particule localisatrice contiennent *-ci* (voir la Section 5.2.2.1) –, mais aussi par nos informateurs francophones natifs. Parmi ces derniers, la plupart éprouvent une certaine réticence à employer les formes en *-ci* dans la langue parlée informelle. Les formes en *-ci*, de plus en plus marginalisées à l'oral, seraient-elles ressenties comme hypercorrectes et auraient-elles fini par acquérir une valeur d'insistance sur la proximité, absente au départ ? Pourtant, il est un contexte où leur position semble encore bien solide, à savoir l'emploi avec les noms dotés d'une sémantique temporelle, que ce soit dans le régime déictique temporel ou anaphorique :

16) je pense qu'il y a pas beaucoup de monde qui a progressé en espagnol *cette année-ci*
CEFC

17) L1 : oh évidemment ma voiture je l'ai encore amenée chez le garagiste

L2 : pourquoi *cette fois-ci*

L1 : un jour je vais pour démarrer j'enclenche la première et la roue arrière elle était bloquée
CEFC

En emploi situationnel, l'abandon progressif des formes en *-ci* a entraîné des reconfigurations du système, pas forcément très faciles à analyser, d'ailleurs. Mis à part l'emploi des seules formes de CE, dépourvues de particule localisatrice, nos données semblent mettre en évidence l'existence de deux autres scénarios possibles. Dans le premier, les locuteurs emploient les formes en *-là* partout, comme nous le voyons dans l'exemple 18 :

18) Lors de la visite d'un musée :

L1 : alors vous pouvez vous tourner regardez bien tout autour de vous et me dire quelles sont les œuvres qui dialoguent entre elles et pour quelles raisons

L2 : euh ben *ces deux là* pour les formes géométriques et *ces deux là* pour le nu CEFC

²⁰ Cf. Foulet (1954) ou Gougenheim (1964 : 216), qui conclut à partir d'une étude d'enregistrements de locuteurs issus de tous milieux sociaux, géographiques et de tous âges confondus : « Il faut reconnaître que les formes en *-ci* sont rares. »

Dans ce type de situations, il semble difficile de soutenir que la forme en *-là* serait tantôt porteuse d'un sens de proximité et tantôt d'un sens de distance, dans la mesure où une telle conclusion défie la bonne logique structuraliste : une opposition sur le plan sémantique devrait être marquée sur le plan formel, surtout au vu du fait que les deux occurrences apparaissent au sein d'un même énoncé. Ainsi, il faudrait plutôt voir dans cet emploi des formes en *-là* non pas une réfection formelle de l'expression de l'opposition de distance, mais un simple renforcement de la composante déictique du démonstratif, accompagné de l'abandon de cette opposition : *-là* ne serait pas ici porteur d'un sens de distance (à moins qu'il s'agisse de référer à deux objets qui se situent tous les deux loin du locuteur), mais aboutirait uniquement à une intensification du sens déictique du démonstratif²¹.

Le second scénario est celui où les locuteurs vont employer CE tout seul, sans particule, pour renvoyer à un référent plus proche, mais où ils vont recourir à CE N-*là* pour renvoyer à un référent plus distant. Dans notre Questionnaire n°2, nous avons présenté à nos informateurs les deux dessins suivants, tout en les contraignant à choisir une seule forme pour référer, chaque fois, à un référent plus proche ou plus éloigné :



Figure 4.1 : « Quelle forme emploieriez-vous si vous étiez la personne à gauche et que vous référiez a) à la boîte rouge et b) à la boîte verte ? »

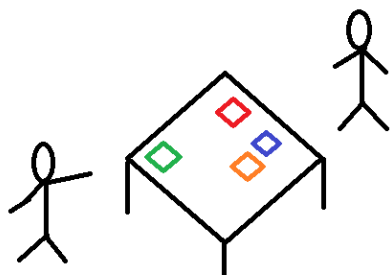


Figure 4.2 : « Quelles formes emploieriez-vous si vous étiez la personne à gauche et que vous référiez, dans un même énoncé, au livre vert et au livre bleu ? »

La différence entre les situations est la suivante : tandis que dans la situation représentée par la Figure 4.1, le locuteur demande soit la boîte rouge soit la boîte verte (deux énoncés indépendants),

²¹ Voir la Section 1.3.2 du Chapitre 1, où, à l'instar d'autres auteurs, nous essayons de défendre l'idée que l'expression de la distance n'est pas une condition nécessaire à la définition de la classe des démonstratifs. Nous remercions B. Meunier pour cette interprétation de *-là* dans les contextes comme celui de 18.

dans celle représentée par la Figure 4.2, il demande, dans un même énoncé, les livres vert et bleu (l'on suppose que le contexte est celui du comptoir d'une librairie).

Pour le premier cas de figure, les informateurs ont préféré, le plus souvent, ou bien employer *cette* sans particule dans les deux situations (pour référer aussi bien à la boîte rouge qu'à la boîte verte) ou bien employer *cette* sans particule pour référer à la boîte rouge et *cette boîte-là* pour référer à la boîte verte. Deux personnes ont cependant suggéré la possibilité d'employer *-là* dans les deux situations. La réticence à utiliser *-ci* était ici plus grande que dans la seconde situation (Figure 4.2), où les référents sont explicitement contrastés au sein d'un même énoncé. Là, le choix qui revenait le plus souvent était l'emploi de *ce livre* pour le livre vert et *ce livre-là* pour le livre rouge. L'opposition *ce livre-ci* vs *ce livre-là* est également apparue plusieurs fois, et un informateur a opté pour *ce livre-là* dans les deux cas²². Que ce soit dans l'un ou l'autre cas, l'usage des formes en *-ci* était quasiment toujours assorti d'un commentaire sur leur caractère moins usité et soutenu. Plus que nous ne l'avions présumé, les locuteurs semblent ainsi avoir conscience de la marginalisation des formes en *-ci* au sein de l'oral informel. Précisons encore que les choix de nos informateurs se présentent à nous comme complètement indépendants de leur lieu d'origine – les mêmes tendances se retrouvent à travers l'ensemble de l'échantillon, ce qui suggère qu'il n'y aurait pas, à cet égard, de différences diatopiques. De la même façon, les choix des francophones installés depuis des années dans un pays non francophone (ici, la République tchèque, en l'occurrence) ne se démarquent pas non plus des choix de ceux qui résident dans un pays francophone.

Dans la Section 4.5, nous présenterons, sur la base de notre échantillon du CEFC, une analyse plus détaillée de la distribution des formes en *-ci* et en *-là* par comparaison avec les formes non-marquées au sein de tous les régimes d'indexicalité, analyse qui sera poursuivie et développée sur une base quantitative dans la Section 5.2.2 du Chapitre 5.

4.1.4 Le déterminant démonstratif français et sa fréquence dans différents types de textes

Nous essaierons de donner ici quelques indices concernant la fréquence avec laquelle le paradigme du déterminant démonstratif apparaît dans le français contemporain. À l'opposé de la situation en tchèque, qui ne connaît pas la distinction entre pronoms et déterminants (voir la Section 3.1.3), nous avons ici l'avantage de pouvoir nous focaliser uniquement sur ces derniers²³.

²² Comme nous l'avons écrit *supra*, *-là* semble dans ce cas – de même que lorsqu'il est employé pour référer à la boîte rouge dans la situation précédente – avoir perdu sa capacité à marquer la distance, son emploi débouchant sur un renforcement du sens indexical de CE.

²³ Malheureusement, il s'ensuit de cela que les deux séries de fréquences obtenues pour les deux langues ne sont pas très comparables entre elles : elles permettent, avant tout, de se faire une idée pour chacun des deux systèmes séparément.

Comme pour le tchèue, il importe de distinguer entre différents types de textes. Pour ce qui est des sources de données, nous nous en remettons d'abord à notre corpus défini à partir du CEFC, mais aussi au *Dictionnaire des fréquences : vocabulaire littéraire des XIXe et XXe siècles* d'Imbs et de Martin (1971), à Frantext et à la plateforme ScienQuest, rassemblant notamment des textes scientifiques, mais pas uniquement (voir la Section 2.1.4).

Le *Dictionnaire des fréquences*, publié en plusieurs volumes en 1971, représente certes une référence importante en la matière. Malheureusement, sa publication, tout comme la période historique qu'il couvre – 1789-1964 (répartie en quinze tranches chronologiques) – sont déjà assez datées à l'heure actuelle. Il faut ajouter à cela le fait que, comme son nom le suggère, l'ensemble des textes constituant le corpus sur lequel se base le dictionnaire sont d'une nature essentiellement (mais pas exclusivement) littéraire. Plus précisément, il s'agit d'à peu près 70 millions de mots, ressortissant à des textes de prose, de poésie et de prose poétique (poèmes en prose)²⁴. Si les quinze périodes couvertes par le dictionnaire sont représentées par des volumes de données d'une taille relativement homogène²⁵, leurs durées ne sont pas les mêmes. Le corpus a bénéficié d'un traitement numérique, lequel a inclus une lemmatisation (les lemmes étant appelés « vedettes ») et une répartition des homographes (constituant un volume à part). Il a servi de documentation au projet du *Trésor de la langue française*, publié entre 1971 et 1996 et informatisé par la suite²⁶. La lemmatisation fait cependant défaut en ce qui concerne les mots grammaticaux, dont le déterminant démonstratif²⁷. Ainsi, le lecteur, ne trouvant pas de lemme CE qui subsume toutes les formes existantes, se voit obligé de regarder les fréquences et les rangs pour chacune de ces formes individuellement. Les fréquences relatives sont calculées pour 10 millions de mots – aussi les avons-nous divisées par dix afin d'arriver aux fréquences relatives rapportées à un million de mots.

Si nous ne regardons que la tranche temporelle la plus récente, soit la période 1946-1964 (celle qui s'approche le plus de l'époque actuelle, et qui est donc la plus à même de nous donner une idée approximative de la situation en synchronie), nous constatons les fréquences relatives et les rangs suivants :

²⁴ La partie prose inclut un ensemble de textes dont la nature se situe entre la fiction et la non-fiction (philosophie, sciences, économie politique), dans la plupart des cas des essais ou des traités. Par ailleurs, le dictionnaire opère aussi la distinction entre textes à caractère de dialogue, textes à caractère de soliloque et le reste.

²⁵ Entre 4 et 6 millions de mots.

²⁶ <http://atilf.atilf.fr/>.

²⁷ « Les variantes fléchies des mots grammaticaux [...] ne sont pas regroupées ; il n'est pas sans intérêt de connaître les fréquences de chacune d'elles, et l'on évite de cette manière le problème du choix de la vedette » [...] (Imbs et Martin, 1971, I : XIV).

Forme	Rang ²⁸	Fréquence relative en mots
<i>Ce</i>	23	3 715,5
<i>Cet</i>	142	352,0
<i>Cette</i>	47	1 661,7
<i>Ces</i>	62	1 057,1
Total		6 786,3

Tableau 4.2 : Rang et fréquence relative des formes du déterminant démonstratif CE pour la période 1946-1964 dans le *Dictionnaire des fréquences : vocabulaire littéraire des XIXe et XXe siècles* (1971)

Regardons à présent la fréquence de CE dans ScienQuest. Dans le corpus composé de textes scientifiques rédigés en français, le lemme CE atteint une fréquence relative de 7 903,7 occurrences pour un million de mots. Dans le corpus de l'*Est républicain*, il atteint une fréquence relative de 5 005,6 occurrences pour un million de mots. La différence entre les deux fréquences est statistiquement significative : le score de la fonction de vraisemblance est égal à 7 455,7464 si nous comparons les fréquences absolues de CE dans les deux corpus et que nous tenons compte de la taille de ces derniers²⁹.

Quant à Frantext, nous avons regardé la fréquence de CE dans notre échantillon de textes de fiction publiés entre 1961 et 2014 (voir la Section 2.1.4 pour la présentation du corpus). Cette fréquence atteint 6 383,2 occurrences pour un million de positions et 7 512,6 occurrences pour un million de mots. Compte-tenu à la fois de la plus grande homogénéité et modernité de l'échantillon en question, c'est ce chiffre qui nous servira de chiffre de référence pour la fiction contemporaine et non pas le chiffre fourni par le *Dictionnaire des fréquences* pour la période 1946-1964, car a) cette dernière source ne se base pas exclusivement sur des textes de fiction, et b) les textes en faisant partie sont plus datés (voir *supra*)³⁰.

Il convient désormais de comparer entre elles ces fréquences relatives obtenues pour les différents ensembles de données, mais aussi – ou surtout – avec la fréquence relative constatée pour notre corpus de conversations défini à partir du CEFC. Cette dernière est de 2 003 occurrences pour un million de positions³¹. La comparaison est compliquée du fait que pour le CEFC, c'est uniquement de la taille en positions que nous disposons, tandis que pour ScienQuest, nous ne disposons que de la taille en mots. Malgré cela, il apparaît que ce sont les textes scientifiques qui emploient le plus de déterminants démonstratifs. Ensuite, les textes de presse, les textes littéraires et, enfin, les

²⁸ Il s'agit de la position dans un index hiérarchique qui liste les formes dans l'ordre décroissant en fonction de la fréquence absolue qu'elles atteignent au sein du sous-corpus représentant la tranche chronologique donnée.

²⁹ La valeur critique est de 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

³⁰ De plus, nous ne disposons que d'une taille approximative de cet ensemble de données et nous ne disposons pas de la fréquence absolue de CE.

³¹ Voir la Section 5.2.1 pour le détail du procédé qui nous a permis d'arriver à ce chiffre.

conversations. Pour nous, ce bilan est fort surprenant dans la mesure où il est complètement aux antipodes de la situation en tchèque (voir la Section 3.1.2 ; bien que, rappelons-le, les fréquences pour le tchèque tiennent également compte des emplois pronominaux). Ainsi, en français, l’oral informel ne serait pas, tant s’en faut, un terrain de prédilection pour l’emploi du déterminant démonstratif, ce dernier apparaissant notamment dans les textes scientifiques, où sa présence pourrait s’expliquer par un rôle important joué dans les procédés d’exposition et d’argumentation (cf. Adam, 2005 [1992] et 1999). La différence reste également frappante entre les textes de presse, la fiction et les conversations. Il ne nous est pas possible d’appliquer la fonction de vraisemblance aux fréquences absolues entre les textes scientifiques et/ou les textes de presse et les conversations, car nous disposons uniquement de la taille en mots pour les premiers et de celle en positions pour les seconds. Nous pouvons toutefois appliquer ce test 1) aux fréquences absolues pour les textes scientifiques et pour les textes de fiction en nous appuyant sur les fréquences en mots, 2) à celles pour les textes de presse et pour les textes de fiction en nous appuyant sur les fréquences en mots et, enfin, 3) à celles pour les conversations du CEFC et pour les textes de fiction en nous appuyant sur les fréquences en positions.

Appliquons d’abord la fonction de vraisemblance à la différence entre la fréquence absolue de CE dans l’échantillon de textes de fiction et sa fréquence absolue dans le corpus de textes scientifiques. Le score obtenu – 29,4396 – atteste la significativité statistique de cette différence. Si nous faisons de même pour les fréquences absolues de CE dans les textes de fiction et dans le corpus de textes de presse, nous obtenons le score de 2 138,0211. Enfin, nous allons appliquer la fonction de vraisemblance aux fréquences absolues de CE dans le corpus de fiction et dans notre corpus de conversations du CEFC. Nous arrivons ici au score de 3 006,3829³².

Le Tableau 4.3 apporte une synthèse des fréquences absolues et relatives de CE dans les différents ensembles de données. Le Tableau 4.4 fournit, quant à lui, une synthèse des scores de la fonction de vraisemblance présentés *supra*. Comme nous pouvons le constater, tous les scores prouvent la signification statistique des différences observées. Cela nous permet ainsi de conclure – dans les limites, bien entendu, qui sont les nôtres (tailles des ensembles et métadonnées différentes, etc.) – que la tendance de CE à apparaître avec ces différents types de textes en français contemporain va décroissant, tout en suivant l’ordre 1) textes scientifiques (non-fiction) > 2) textes de fiction > 3) textes de presse > 4) conversations :

³² Avec la valeur critique de 3,841 (degré de liberté = 1 et $\alpha = 0,05$). La comparaison des corpus de ScienQuest avec le sous-corpus de Frantext emploie la taille en mots. Celle du sous-corpus de Frantext avec le corpus de conversations emploie la taille en positions.

Corpus	Fréquence absolue	Fréquence relative en mots
<i>Dictionnaire des fréquences</i> (1971) : période 1946-1964 ; entre 4 et 6 millions de mots		6 786,3
Écrits scientifiques en français (ScienQuest ; 5 596 492 de mots)	44 233	7 903,7
<i>L'Est républicain</i> (ScienQuest ; 87 341 867 de mots)	437 202	5 005,6
Frantext – échantillon de fiction (1961-2014 ; 2 008 222 de mots)	15 087	7 512,6
		Fréquence relative en positions
Frantext – échantillon de fiction (1961-2014 ; 2 363 566 de positions)	15 087	6 383,2
Conversations du CEFC (Orféo ; 934 102 positions)	1 871	2 003

Tableau 4.3 : Fréquences absolue et relative du déterminant démonstratif CE dans la tranche 1946-1964 du corpus du *Dictionnaire des fréquences* (1971), le corpus Écrits scientifiques en français (ScienQuest), le corpus de *l'Est républicain* (ScienQuest), l'échantillon de textes de fiction publiés entre 1961 et 2014 (Frantext) et le corpus de conversations du CEFC (Orféo)

Corpus	Score de la fonction de vraisemblance
Écrits scientifiques en français (ScienQuest) vs <i>l'Est républicain</i> (ScienQuest)	7 455,7464
Écrits scientifiques en français (ScienQuest) vs Frantext – échantillon de fiction (1961-2014)	29,4396
<i>l'Est républicain</i> (ScienQuest) vs Frantext – échantillon de fiction (1961-2014)	2 138,0211
Conversations du CEFC vs Frantext – échantillon de fiction (1961-2014)	3 006,3829

Tableau 4.4 : Scores de la fonction de vraisemblance appliquée aux différences entre les fréquences absolues du lemme CE dans a) le corpus Écrits scientifiques en français (ScienQuest) et le corpus de *l'Est républicain* (ScienQuest), b) le corpus Écrits scientifiques en français (ScienQuest) et l'échantillon de textes de fiction publiés entre 1961 et 2014 (Frantext), c) le corpus de *l'Est républicain* (ScienQuest) et l'échantillon de textes de fiction publiés entre 1961 et 2014 (Frantext) et d) le corpus de conversations du CEFC (Orféo) et l'échantillon de textes de fiction publiés entre 1961 et 2014 (Frantext) ; valeur critique = 3,841

Ce bilan est pour nous assez inattendu dans la mesure où il révèle une tendance totalement inverse à celle relative à l'apparition du lemme TEN en tchèque contemporain (voir la Section 3.1.2),

où ce sont les conversations et les entretiens qui contiennent le plus de TEN, suivis des textes de fiction, puis des textes de presses et, enfin, des textes scientifiques.

4.2 Le démonstratif français en diachronie

Dans la présente section, nous allons rappeler la genèse du système du démonstratif français du latin jusqu'au français moderne. Vu la complexité du sujet et le fait qu'il ne s'agit pas d'une problématique centrale de notre travail, nous nous contenterons, dans ce qui suit, d'une présentation succincte des grands tournants et bouleversements qui se sont produits dans le système du démonstratif et qui ont abouti à la configuration actuelle. Cette évolution étant assez bien décrite dans la littérature (même si des questionnements et des points obscurs subsistent), nous nous appuierons sur une série de sources de référence, dont notamment : Carlier et Guillot (2015), Carlier et De Mulder (2006, 2010), Fruyt (2003, 2010a, 2010b), Goyens (1994), Guénette (1995), Guillot (2010b, 2017), Kleiber (1984b), Marchello-Nizia (1995, 2004, 2006a, 2006b) et Price (1968, 1969).

Le premier moment clé dans l'évolution du système du démonstratif entre le latin et le français moderne est la réfection d'un système articulé essentiellement autour de la personne grammaticale en un système articulé essentiellement autour de la distance du référent par rapport au locuteur. Il faut cependant dire que même si le latin (à la fois classique et tardif) – avec ses trois paradigmes de formes renvoyant (plus ou moins) aux sphères du locuteur (HIC), de l'interlocuteur (ISTE) et du délocuté (ILLE) – est traditionnellement décrit comme un système fondé sur la notion de personne, cette analyse a déjà été relativisée et nuancée par Fruyt (2010a et 2010b). Cette dernière montre que l'opposition entre les trois formes est en grande partie dépendante de la façon dont le locuteur construit l'énonciation. Deux cas de figure peuvent ainsi se présenter, selon que le locuteur fait le choix d'inclure l'interlocuteur dans sa sphère ou de l'en rejeter. Seul le second scénario débouche sur l'opposition notoire à trois termes, comme le résumant Carlier et Guillot (2015 : 346) :

Pour opposer clairement les deux sphères, le latin classique utilise *hic* pour marquer la sphère personnelle du locuteur et *iste* pour marquer celle de son allocutaire [...]. L'allocutaire peut donc servir de repère, ce qui conduit à dire que la deixis est structurée d'après la personne en latin. *Ille* se rapporte à ce qui est externe à la sphère du locuteur et à celle du destinataire.

Le premier scénario, en revanche, aboutit à une opposition binaire entre seulement HIC et ILLE, celui-là relevant de l'espace partagé par le locuteur et l'interlocuteur et celui-ci, de tout ce qui se situe en dehors de cet espace. Ce constat nous semble parfaitement conforme à la relativisation de l'opposition entre systèmes basés sur la personnes et systèmes basés sur la distance faite par Meira (2003 ; voir la Section 1.3.1 du Chapitre 1).

Contrairement au système ternaire du latin, l'ancien français emploie un système binaire, contenant les paradigmes des formes CIST et CIL, issues, respectivement, de la coalescence de la forme ISTE et de la forme ILLE avec la forme présentative *ecce*³³. En face d'une évolution morphologique transparente, les choses sont moins claires sur le plan sémantique, comme le résumant de nouveau Carlier et Guillot, à qui nous empruntons le tableau suivant :

	Évolution morphologique			Évolution sémantique		
Latin class.	(<i>ecce</i>) <i>hic</i>	(<i>ecce</i>) <i>iste</i>	(<i>ecce</i>) <i>ille</i>	<i>hic</i>	<i>iste</i>	<i>ille</i>
Anc. fr.	Ø (pr. n. <i>ce</i>)	<i>cist</i>	<i>cil</i>	<i>cist</i>	<i>cil</i> ?	<i>cil</i>

Tableau 4.5 : Passage du système du démonstratif latin au système du démonstratif en ancien français (Carlier et Guillot, 2015 : 358)

La valeur du paradigme de HIC est reprise, dans le nouveau système, par le paradigme de CIST. Quant à la forme CIL, outre qu'elle perpétue la valeur de ILLE, elle semble également reprendre celle de ISTE. Se pose donc automatiquement la question de savoir pourquoi la nouvelle forme dotée de la valeur de HIC n'est pas dérivée de HIC mais de ISTE. C'est que le paradigme de la forme HIC, faible sur le plan morpho-phonétique, a disparu de la langue, de même que le paradigme de la forme IS, réservé, dans le latin classique, aux emplois endophoriques (cf. Harris, 1978 ; Carlier et De Mulder, 2010)³⁴. Ces deux formes vont progressivement disparaître entre le latin et son passage au français. À cela s'ajouterait, selon Carlier et Guillot, l'avantage d'une régularité morphologique détenu par les formes des paradigmes ISTE et ILLE contrairement aux formes des paradigmes HIC et IS, plus irréguliers :

Contrairement aux formes *hic* [...] et *is* ainsi que son composé *i(s)dem*, sans descendant dans les langues romanes, les trois formes qui se maintiennent, à savoir *iste*, *ille* et *ipse*, ont une propriété en commun : elles sont dotées d'un morphème lexical invariant auquel se joint la flexion suffixale. [...] Ces formes présentent ainsi une plus grande iconicité ou un plus grand isomorphisme syntagmatique. Elles impliquent donc un appariement plus simple et plus univoque de la forme et du sens et supposent ainsi un moindre coût cognitif pour l'encodage et le décodage [...]. (Carlier et Guillot, 2015 : 367-368)

Ainsi, comme l'indique le Tableau 4.2, la seule forme de l'ancien français et du français moderne ayant son origine dans HIC est la forme pronominale neutre *ce*, issue de *ecce hoc*.

Le nouveau système, réparti entre CIST et CIL, apparaît dès le très ancien français (9^{ème} siècle). Les formes affichent une déclinaison « simplifiée »³⁵, distinguant entre les cas sujet, régime direct et

³³ Les formes renforcées du démonstratif *ecce iste* et *ecce ille* existaient déjà en latin préclassique.

³⁴ Non seulement anaphoriques, comme on le dit souvent, mais également cataphoriques intra-syntagmatiques, fonction qui est assurée, dans le français moderne, par CELUI ou par l'article défini et, dans le tchèque moderne, par TEN ; l'ancien tchèque utilisait la forme JEN (Gebauer, 1928 : 231 ; voir la Section 3.2).

³⁵ Par comparaison avec celle du latin ou du tchèque.

régime indirect. Hormis la forme issue de *ecce hoc*, toujours pronominale, elles s'emploient toutes deux à la fois en emploi pronominal et adnominal³⁶ :

		CIST		CIL	
		masculin	féminin	masculin	féminin
Sing.	sujet	<i>cist</i>	<i>ceste</i>	<i>cil</i>	<i>cele</i>
	rég. 1	<i>cest</i>	<i>ceste</i>	<i>cel</i>	<i>cele</i>
	rég. 2	<i>cestui</i>	<i>cesti</i>	<i>celui</i>	<i>celi</i>
Plur.	sujet.	<i>cist</i>	<i>cestes</i>	<i>cil</i>	<i>celes</i>
	régime	<i>cez</i>	<i>cestes</i>	<i>cels</i>	<i>celes</i>

Tableau 4.6 : Système des démonstratifs au début de l'ancien français

À partir de cet état initial, le système traverse une série de remaniements pour aboutir au système du français actuel, caractérisé à la fois par une neutralisation sémantique (celle de l'opposition de distance) et une spécialisation syntaxique (l'existence de deux paradigmes distincts pour les emplois adnominaux et pronominaux ; voir la Section 4.1.1 *supra*). Pour commencer par un résumé de ce qui s'est passé, citons ici Marchello-Nizia (1995 : 115) :

Dans la morpho-syntaxe du français, le système des démonstratifs est sans doute celui qui a subi le changement le plus radical. On est passé d'un système formel parfaitement symétrique en ancien français à un système absolument dissymétrique en français classique et moderne³⁷.

Le changement qui semble avoir enclenché cette longue évolution est l'apparition d'une nouvelle forme, *ces*, remplaçant les formes du cas régime pluriel masculin *cez* et *cels* en emploi adnominal (fin du 12^{ème} siècle). Les raisons de cette apparition le plus fréquemment mises en avant sont, d'une part, des raisons phonétiques et, d'autre part, des raisons d'analogie. Price (1964 : 494) résume que « [...] *cels* en position faible (c.-à-d. *cels* adjectif, qui précédait le nom) serait devenu **ces*

³⁶ Nous laissons ici de côté la nature de l'opposition entre ces deux paradigmes, à laquelle bon nombre d'auteurs ont essayé d'apporter une explication satisfaisante. Rappelons juste qu'outre les théories fondées sur la notion de distance du référent (cf. Price, 1968), il en existe d'autres, telles les théories dites « étymologique » (Moignet, 1973) et « rhétoriques » (cf. Giraud, 1967), ainsi que la théorie de Kleiber (1984b), articulée autour de la notion de *token-réflexivité*, et celle de Marchello-Nizia (2003), articulée autour de la notion de « sphère personnelle du locuteur ». Selon Kleiber, CIL est la forme non marquée, car, contrairement à CIST, elle ne demande pas à ce que l'appariement entre son occurrence et son référent s'opère par l'entourage spatio-temporel immédiat de cette occurrence. Selon Marchello-Nizia, CIST permet d'inclure le référent dans la sphère du locuteur (notion pragmatique définie dans la Section 1.3.1 du Chapitre 1), contrairement à CIL, qui l'en rejette. Carlier et Guillot (2015) voient dans cette situation un reflet de celle qui aurait existé en latin : CIST aurait ainsi pu s'employer pour la seule zone du locuteur ou pour la zone partagée entre celui-ci et l'interlocuteur ; CIL, quant à lui, aurait soit référé uniquement à tout ce qui était en dehors de cette zone partagée soit également à la zone de l'interlocuteur. Ajoutons que les formes des deux paradigmes pouvaient être renforcées moyennant une préfixation par *i-* : *icist*, *icel*, *icelui*, etc.

³⁷ Rappelons que cette dissymétrie n'est pas seulement quelque chose d'inédit à l'échelle des langues du monde (cf. Dixon, 2003 : 63), mais aussi au sein même de la famille des langues romanes (cf. Marchello-Nizia, 1995 : 117).

(comme *dels* est devenu *des*) [...] ». Quant à *cez*, jusqu'alors prononcé /tsets/, il serait lui aussi devenu *ces*, prononcé /ses/, du fait de la réduction de la consonne affriquée. Aussitôt émergée, la nouvelle forme voit son emploi s'étendre (toujours dans les emplois adnominaux uniquement) aux cas régime et sujet du pluriel féminin. Dès lors, le cas régime pluriel des deux genres tout comme le cas sujet pluriel féminin affichent un syncrétisme morphologique, la forme *ces* y ayant évincé les formes utilisées jusqu'alors. Au début du 13^{ème} siècle, la forme *ce* apparaît en position antéconsonantique au cas régime masculin singulier des emplois adnominaux. Il existe de fortes raisons pour croire que l'apparition de cette forme a pour principale motivation l'analogie avec les formes de l'article défini, *ce* permettant ainsi d'établir une symétrie : *ce/ces*, *le/les*. Ensuite, les vestiges de la déclinaison, hérités du latin et déjà plutôt instables lors de la période de l'ancien français, disparaissent définitivement à la fin de la période (vers l'an 1300). Cette disparition entraîne la généralisation de l'ancien cas régime dans tous les emplois adnominaux du démonstratif.

Pour ce qui concerne les emplois pronominaux, le paradigme encourt lui aussi plusieurs changements dès l'ancien français : a) la disparition des formes *cesti* et *celi* du cas régime indirect féminin, b) la spécialisation des formes *cest* et *cel* du cas régime direct masculin singulier dans les emplois adnominaux (tendant à quitter le domaine des emplois pronominaux), c) l'apparition de *ces* en emploi adnominal du pluriel (voir *supra*). C'est ce dernier changement qui aurait eu pour conséquence l'apparition de deux nouvelles formes – *ceus-ci* et *ceus-la* –, toutes deux assorties d'une particule localisatrice et remplaçant, respectivement, les anciennes formes *cez* (cas régime masculin pluriel de *cist*) et *cels* (cas régime masculin pluriel de *cil*)³⁸. Et, enfin, d) la disparition déjà mentionnée des formes du cas sujet.

Les changements se poursuivent en moyen français et au-delà. Concernant le domaine des formes adnominales, celles issues de CIL sont peu à peu éliminées au profit des seules formes issues de CIST et de celles constituant, dès le 13^{ème} siècle, un troisième paradigme, tout à fait à part. En effet, comme l'explique Marchello-Nizia (1995 : 123-124), s'inspirant de Dees (1971), seules les formes *ceste* et *cest* appartiennent au paradigme de CIST, tandis que la forme *ces* appartient probablement aux deux paradigmes à la fois (voir *supra*), et la forme *ce* est créée plus tard par analogie, n'appartenant, de ce fait, ni à l'un ni à l'autre. Dès le début du 16^{ème} siècle, nous retrouvons en position de déterminant les mêmes formes que nous y retrouvons aujourd'hui³⁹. Force est cependant de constater que pendant un certain moment, outre ces formes-ci, l'on continue encore à rencontrer dans cette

³⁸ Comme l'explique Marchello-Nizia (1995 : 124) en reprenant les arguments de Dees (1971) : « [...] en effet, *cez* pronom ou déterminant étant devenu *ces*, l'emploi déterminant, et sémantiquement indifférencié, de cette forme ne permettait pas de continuer à l'employer comme pronom et de l'opposer à *ceus*. Une opposition sémantique a été reconstruite grâce à *ci* et *la* : *ceus-la* correspondant à l'ancienne valeur de *ceus*, et *ceus-ci* à la valeur de l'ancienne forme *cez*. »

³⁹ Mis à part quelques changements mineurs : *cest* (masculin singulier devant une voyelle) devient *cet* et *ceste* (féminin singulier) devient *cette*.

position des formes plus longues, appartenant à la fois aux paradigmes de CIST et de CIL, y compris des formes préfixées. En effet, ces formes sont encore très nombreuses chez des auteurs de la première moitié du 16^{ème} siècle, tel Rabelais : *en cestui temps, pour icelle journée, durant icelui repas*, etc. (Lagarde et Michard, 1964)

Du côté des formes pronominales, nous assistons à un usage de plus en plus extensif des formes complétées par *-ci* (démonstratif adverbial issu de *ecce hic*) et *-la* (démonstratif adverbial issu de *illac*). De manière générale, ce dernier va de pair avec le déclin de l'opposition sémantique entre CIST et CIL. Comme le résume Guénette (1995 : 37) :

La différence sémantique entre *cist* et *cil* s'estompant, on a fait appel aux adverbes les plus appropriés sémantiquement pour signifier la proximité et l'éloignement par rapport au locuteur.

Le paradigme des pronoms est, quant à lui, globalement achevé au 17^{ème} siècle. Dans cette nouvelle configuration, qui est également celle du français moderne, les formes de CIST sont progressivement éliminées au profit des seules formes ressortissant à CIL. Le facteur qui aurait joué un rôle central dans cette évolution est le rapprochement analogique entre lesdites formes de CIL et les formes toniques du pronom personnel LUI : *luy/celui, ele/cele, eus/ceus, eles/celes*. Cette analogie est particulièrement valorisée par Marchello-Nizia (cf. 1995). L'auteure souligne l'importance, en moyen français (14^{ème} siècle), du partage des formes pronominales personnelles en toniques et atones (clitiques), ayant pour origine le changement dans l'accentuation de la langue : « [...] à un accent tonique de mots encore largement conservé succède un accent de fin de groupe de mots qui cesse d'être distinctif [...] » (1995 : 174). Ainsi, dès que le pronom personnel sujet *il* devient exclusivement atone et se voit remplacé par *luy* en position tonique, *celui* se développe en position de pronom, par analogie avec ce dernier.

Concernant la présence des particules localisatrices, Marchello-Nizia (1995 : 175) précise qu'au 17^{ème} siècle,

[L]es pronoms démonstratifs sont désormais nécessairement suffixés en *-ci/-la*, sur le modèle à la fois des déterminants correspondants et des emplois « déterminatifs » *celui qui/de SN*, où le démonstratif est nécessairement suivi d'un élément.

S'agissant de la présence des formes issues de CIL dans les emplois « déterminatifs », elle n'est guère surprenante dans une perspective typologique, où cette fonction semble toujours être dévolue aux formes du démonstratif non marqué pour la distance (cf. IS en latin classique, JEN en ancien tchèque

et TEN en tchèque actuel ou encore DER/DERJENIGE en allemand actuel et DEN en suédois actuel) ou marquées pour l'éloignement (cf. THAT en anglais⁴⁰).

Du reste, il est utile de rappeler ici que la répartition des formes du démonstratif nominal français en deux paradigmes distincts s'inscrit, selon Marchello-Nizia (1995), dans un macro-phénomène ayant touché, de toutes les langues romanes, exclusivement le français. Un macro-phénomène, car cette dichotomie ne se rencontre pas uniquement dans le domaine des démonstratifs et des pronoms personnels, mais également dans celui des indéfinis et, dans une moindre mesure, dans celui des possessifs (cf. la Section 4.1.2). C'est le changement du modèle d'accentuation qui aurait joué un rôle important dans cette séparation des paradigmes, changement que Marchello-Nizia (1995 : 171) résume comme suit :

On sait que dans le passage du latin aux états anciens des langues romanes, un certain nombre de mots, qui jusque-là possédaient un accent tonique autonome, sont devenus des clitiques, pro- ou post-tonique : c'est le cas des déterminants du nom, ainsi que des pronoms personnels régimes atones. On sait aussi qu'en français moderne, chaque mot autonome ayant un accent sur la dernière syllabe « masculine », cet accent n'est pas distinctif ; en outre, dès que les mots s'enchaînent dans la syntaxe de l'énoncé, c'est le groupe de mots, syntagme ou phrase, qui porte l'accent.

S'agissant de la restructuration du modèle accentuel tonique en modèle accentuel de groupe, située à l'époque du moyen français, l'auteure (1995 : 173-174) rappelle ceci :

En effet, en moyen français se développent des formes de déterminants nouvelles, telles *chaque* à côté de *chascun*, forme jusque-là unique pour le pronom et le déterminant [...]. Et ainsi, à partir de cette époque, il existe en français un véritable macro-système des déterminants atones à trois termes : deux formes au singulier, et au pluriel une seule neutralisée en genre. En font partie dès le XVe siècle : l'article (ou déterminant) défini *le/la/les*, l'article (ou déterminant) indéfini *un/une/des* (*des* est une innovation du moyen français), l'adjectif (ou déterminant) possessif *mon/ma/mes*. En utilisant dans cet emploi, pour compléter le paradigme défectif *ce/ces*, la forme *ceste* [...], le moyen français complète la série : *ce/cette/ces*.

Sans vouloir entrer ici dans trop de détails⁴¹, il paraît incontestable que ce changement du modèle d'accentuation – dont la datation exacte demeure pourtant difficile et incertaine – a joué un rôle important dans la réfection du système du démonstratif français pour aboutir à la situation actuelle⁴².

⁴⁰ Pour l'anglais, la question se pose cependant de savoir s'il ne faudrait plutôt voir dans THAT la forme neutre du système (cf. J. Lyons 1990 [1978]), n'insistant sur l'éloignement que lorsqu'elle est employée de manière contrastive avec THIS. Au fond, la situation ne serait peut-être pas loin de celle décrite par Kleiber (1984b) pour l'ancien français, où CIL apparaît également avoir parfois été la forme neutre. Rappelons également qu'en ancien français, les mêmes formes que celles de l'article défini pouvaient être employées comme déterminatifs (cf. l'espagnol actuel, où, cependant, une forme déterminative ne peut être employée comme article défini, à savoir la forme *lo* : *los que saben* = *ceux qui savent*, en *lo que se refiere a* = *en ce qui concerne*).

⁴¹ À cet égard, nous renvoyons le lecteur notamment à l'étude ici mentionnée de Marchello-Nizia (1995).

⁴² En partie, il expliquerait même l'apparition des formes pourvues des particules localisatrices, dans la mesure où c'est la particule qui porte l'accent de groupe. Nous remercions Mme Guillot pour cette remarque.

4.3 Le déterminant démonstratif français : un article défini fort ?

Dans la présente section, nous tenons à revenir pour mieux l'examiner à la thèse en vertu de laquelle le déterminant démonstratif du français moderne serait proche d'un article défini. Rappelons que ce point de vue est défendu par Harris (1977, 1978 et 1980). Il reçoit également un certain soutien chez Carlier et De Mulder (2006). Rappelons également que chez Harris, cette thèse trouve un corollaire dans l'analyse du pronom démonstratif neutre *CE* comme pronom personnel de la 3^{ème} personne. Une part des arguments corroborant les deux thèses ont déjà été présentés (et réfutés) dans la Section 1.3.2 du Chapitre 1 et dans la Section 4.1.1 du présent chapitre. Dans cette dernière section, il était question en particulier de l'incapacité du démonstratif français (à la fois déterminant et pronom) employé sans particule localisatrice à exprimer des oppositions de distance et du remplacement par *ce* de *il* dans les énoncés comme *Il est vrai*.

Dans une tentative d'articuler les deux thèses, Harris introduit son étude de 1977 en écrivant ceci :

Specifically it will be argued that the development and use of the so-called 'démonstratif neutre' must be interpreted against the background of the changing role of the definite article and this, in turn, is a reflection of certain profound changes in the phonological and grammatical patterns of French. (1997 : 249)

Quant au changement, pointé par l'auteur, qui se serait produit dans le rôle de l'article défini français, celui-ci peut être résumé comme suit : une fois que le nom français a cessé d'être morphologiquement et phonétiquement marqué pour le genre et le nombre, l'article défini aurait peu à peu repris ce rôle, jadis assumé par un système de suffixation. L'érosion de ce système a commencé, sur le territoire de la Gaule, bien avant l'émergence du français par l'apocope des syllabes finales non-accentuées et par la neutralisation des oppositions entre les voyelles finales via leur convergence en /ə/. Bien plus tard, elle s'est poursuivie moyennant l'amuïssement des consonnes finales (dans la langue populaire, le /s/ final s'efface de la prononciation dès le 12^{ème} siècle). Comme le souligne Harris, un parallèle frappant existe ici avec le système verbal, où l'expression du pronom personnel, désormais seul trait permettant de distinguer entre plusieurs personnes grammaticales, devient obligatoire en moyen français (cf. également Marchello-Nizia, 1995). Dans les deux cas, l'on assisterait à une réfection du système du marquage morphologique, quittant le domaine de la suffixation pour celui de la préfixation :

Within noun phrases, information as to gender and number is in principle no longer conveyed by suffixes but by mandatory prefixes (traditionally called determiners), the whole functioning as one unit. (Harris, 1977 : 253)

Outre qu'il ne va pas de soi de parler de préfixes pour des éléments qui, en l'occurrence, se comportent sur les plans phonétique et syntaxique comme des clitiques, il s'agit d'une analyse à

première vue convaincante et en partie juste. Mais en partie seulement. Pour le domaine verbal, il n'y a pas grand-chose à ajouter. Pour ce qui concerne le domaine nominal, en revanche, l'analyse comporte, selon nous, des failles considérables. D'abord, si nous regardons les autres langues romanes, tels l'italien et l'espagnol, nous voyons que dans ces langues, l'usage du défini est globalement tout aussi grammaticalisé et répandu qu'en français, alors même que le nom continue à y être marqué pour les oppositions grammaticales :

19) ES : los libr-os
 le.M.NOM.PL livre-NOM.PL
 les livres

L'existence de ce fait est d'ailleurs bien reconnue par Carlier et De Mulder (2006 : 97) dans une note de bas de page :

L'hypothèse, défendue par Harris (1977), que l'extension de l'article défini dans le domaine du générique en français serait due au fait que cet article est utilisé pour marquer le genre et le nombre, à la place du système antérieur de la flexion suffixale en genre et nombre, est infirmée par des langues comme l'espagnol et l'italien : quoique ces langues maintiennent la flexion suffixale en genre et en nombre et fassent encore un emploi large du degré zéro de la détermination, elles conceptualisent néanmoins le générique en termes de définitude.

Voici un autre grand défaut que nous trouvons à la thèse de Harris. Celle-ci « renvoie » indirectement au schéma de grammaticalisation du démonstratif proposé par Greenberg un an plus tard (1978a). Nous reproduisons ici ce schéma traduit en français, tel qu'il apparaît chez Carlier et De Mulder (2006 : 97) :

Stade 0	Stade I	Stade II	Stade III
DÉMONSTRATIF > ARTICLE DÉFINI > ARTICLE NON GÉNÉRIQUE > MARQUEUR NOMINAL			

Ainsi, dans cette optique, la thèse de Harris postule que le français aurait atteint le stade III, le démonstratif latin *ILLE* de départ étant devenu un pur marqueur de la morphologie nominale. Toutefois, il s'avère que cette conclusion est en contradiction avec le schéma de Greenberg en ce que ce dernier maintient que pour en arriver là, la forme du démonstratif doit d'abord passer par les deux stades précédents. Or l'article défini français n'a jamais épousé des emplois relevant du stade II, consistant à introduire dans le discours tous les référents spécifiques, indépendamment de leur statut de défini ou d'indéfini⁴³. Et pour cause, car dès l'ancien français, le rôle d'introduire les indéfinis

⁴³ Au passage, remarquons que nous ne trouvons pas que le terme de « non générique », utilisé pour le stade II par Greenberg, soit un terme approprié, car tous les emplois non génériques d'un nom ne sont pas, tant s'en faut, spécifiques (voir notre définition de la spécificité dans la Section 1.2.1 du Chapitre 1). Mais il s'agit ici d'un malentendu dû à la définition différente que le terme de « générique » reçoit chez Greenberg. Pour ce dernier, se trouve en emploi générique tout nom qui n'entraîne pas le présupposé de référence, y compris les noms en emplois attributifs et en emplois de négation absolue (cf. *je n'ai pas d'eau et il est tailleur* ; 1978a : 64) : "We may define Stage II as the stage in which we have an article which includes, along with possibly other uses, both definite determination and non-definite specific uses. Specific,

spécifiques est dévolu à l'article UN. Il est ainsi prématuré de prétendre que l'article défini français a perdu son contenu sémantique propre (sa « raison d'être »), consistant, avant tout, à marquer l'unicité (la maximalité) du concept, avis par lequel Carlier et De Mulder concluent leur étude de 2006 :

Tant que l'article défini se maintient avec une substance sémantique propre, le démonstratif *ce* ne peut acquérir le statut d'article défini. Tel est certainement encore le cas dans la langue actuelle : quoique l'article défini présente des symptômes d'affaiblissement sémantique, comme le montrent par exemple les alternances du type *avoir peur* / *avoir la trouille*, nous sommes d'avis, contrairement à Harris (1977, 1980a-b), que l'article défini *le* est loin d'être un simple marqueur du statut nominal et des catégories morphologiques du genre et du nombre qui y sont associées. (2006 : 111)

L'exemple des alternances *avoir peur* / *avoir la trouille* est cependant révélateur. D'autres cas comme celui-ci peuvent être cités, tels *prendre la fuite* ou *avoir la flemme*. Dans toutes ces constructions à verbe support⁴⁴, le nom est couramment analysé comme étant en emploi intensionnel (cf. Wilmet, 1986, 2010 [1997]). Ici, l'article ne marque pas l'extensivité extensive (Wilmet, 1986) du nom, n'actualise pas le nom en discours : son emploi pourrait donc être classé comme relevant du stade III décrit par Greenberg. Mais ce phénomène, bien qu'il semble en hausse dans la langue informelle contemporaine⁴⁵, reste tout de même marginal comparé à tous les autres usages du défini.

Un autre emploi du défini au sujet duquel l'on peut éventuellement se demander s'il y réalise encore sa fonction première, ce sont ses emplois génériques, déjà mentionnés ici :

20) *Le chien* est le meilleur ami de *l'homme*.

Les baleines sont des mammifères.

Ces emplois sont sujets à controverse, à commencer par leur analyse en tant qu'emplois référentiels (cf. Wilmet, 1986) ou non (cf. Löbner, 2011). Aussi l'approche de la définitude de Löbner laisse-elle de côté les emplois génériques, supposément non-référentiels :

The proposal is confined to NPs in referential use, be it nonquantificational or quantificational, but it excludes predicative and generic NPs. This is in line with the basic idea that the function of determination is essentially to fix the type of reference for the NP." (2011 : 279-280)

Ce point de vue est justifiable et, de fait, nous devrions peut-être le faire nôtre si nous continuons – comme nous l'avons fait dans la Section 1.2.1 en nous référant à Donnellan (1966), à Prince (1981)

opposed here to generic, is the use of such an article in contexts in which a specific but unidentified item is referred to, that is, there is a presupposition of reference." (1978a : 62)

⁴⁴ Où l'essentiel de l'information prédicative est porté par le nom et non pas par le verbe, et où l'ensemble verbe + nom est souvent paraphrasable par une seule forme verbale : *prendre la fuite* → *s'enfuir*.

⁴⁵ Ce qui est encore plus certain, c'est qu'il s'agit d'une innovation de la langue par rapport à ses états plus anciens, où l'article n'apparaissait pas dans ces contextes. Dans l'ensemble de Frantext, la construction *avoir peur* apparaît ainsi dès le 12^{ème} siècle, tandis qu'*avoir la flemme* et *avoir la trouille* n'apparaissent qu'à partir du 18^{ème} et du 19^{ème} siècles respectivement. Quant à la construction *prendre la fuite*, elle apparaît dès le 15^{ème} siècle.

et à Fodor et Sag (1982) – à indexer la référence sur la spécificité. En effet, les emplois génériques du nom ne sont pas spécifiques : les SN *le chien*, *l'homme* et *les baleines* de l'exemple 20 ci-dessus ne permettent pas d'identifier un particulier, ils saisissent la classe dans son ensemble. D'autre part, il ne nous semble pas non plus indéfendable de postuler que l'acte linguistique qu'est l'emploi d'un générique s'apparente, tout simplement, à un autre type d'emploi référentiel, dans lequel une référence est faite à l'ensemble de la classe en question⁴⁶. Pour le dire en un mot, pour nous, le statut référentiel ou non des génériques reste une question ouverte et que nous n'avons pas l'intention de trancher ici.

Une autre difficulté liée aux génériques a trait à l'instabilité qui existe, justement, en ce qui concerne leur marquage par l'article défini :

21) FR : **Les** baleines sont des mammifères.

ENG : Whales are mammals.

DE : (**Die**) Wale sind Säugetiere.

La présence du défini avec les noms génériques (et, a fortiori, avec les noms génériques pluriels et indénombrables) constitue l'un des indicateurs majeurs du degré de grammaticalisation de ce dernier :

From a diachronic viewpoint, definite determiners indeed spread along the scale [...] from deictic and anaphoric uses to semantically-unique uses like associative anaphora, individual nouns (*the sun*) and **generic NPs**⁴⁷ (*the Panda*). (Hofherr et Zribi-Hertz, 2014 : 9)

En ce sens, le français moderne (et les langues romanes en général) représente un stade de grammaticalisation du défini nettement plus avancé que l'anglais (et les langues germaniques en général). L'allemand, où la présence du défini est autorisée mais en aucun cas exigée, semble occuper une position intermédiaire⁴⁸. Mais que les génériques soient ou non marqués par l'article défini, ils valident le test de la négation proposé par Löbner (1985) et introduit dans la Section 1.2.1. Sur ce point, ils manifestent une propriété uniforme à travers les langues du monde en ce qu'ils se comportent comme des expressions « sémantiquement » définies⁴⁹.

⁴⁶ Contrairement aux emplois non-spécifiques et non-génériques (cf. *John wants to marry a Norwegian* entendu au sens de n'importe quelle Norvégienne ; Prince, 1981 : 237), où ni une entité particulière ni un ensemble ne peuvent être identifiés.

⁴⁷ Nous soulignons.

⁴⁸ Comme le montrent Barton, Kolb et Kupisch dans leur étude empirique (2015), même si la situation décrite comme neutre pour l'allemand standard est celle de l'absence de défini, sa présence est également autorisée ; le degré de son acceptabilité dépend ensuite d'une série de facteurs ayant trait à la fois à la nature sémantique du nom et au profil du locuteur.

⁴⁹ "The common logical behaviour of characterizing NPs and definite NPs might be considered a motivation why in many languages, for example, French, characterizing NPs receive singular definiteness marking (Lyons 1999 : 51)." (Löbner, 2011 : 296) Par "characterizing NPs", Löbner entend les SN génériques.

Regardons à présent de plus près les arguments avancés par Carlier et De Mulder (2006), dont la thèse est beaucoup plus nuancée que celle de Harris (cf. 1977), les auteurs soutenant que le déterminant démonstratif CE serait seulement « un article défini en devenir » (cf. 2006 : 110). Ils écrivent, entre autres, que CE déterminant s'emploie

[...] pour une reprise anaphorique d'un référent introduit dans le contexte antérieur sous la forme d'un SN, en particulier quand cette reprise anaphorique va de pair avec une reclassification [...]

[...] en emploi cataphorique, quand le nom tête auquel se rapporte le démonstratif est suivi d'une relative établissant l'existence du référent, en situant celui-ci dans le temps et / ou dans l'espace :

Il faudra que je vous montre **ce tableau que j'ai acheté**. (L. Aragon, *Les beaux quartiers*, Frantext)
(2006 : 103)

En guise de récapitulatif, les auteurs écrivent au sujet du déterminant démonstratif CE du français moderne :

Outre le fait qu'en tant que démonstratif, il a perdu la composante déictique permettant d'opposer proximité et distance, il s'est approprié les contextes pragmatiquement définis, où le référent est accessible parce qu'il est présent dans le contexte d'emploi, ce qui nous a conduits à le rapprocher de l'article défini « fort » attesté dans certains parlers germaniques. Il s'est par ailleurs aussi installé dans l'emploi mémoriel qui se situe sur les confins de la définitude pragmatique et peut ainsi assurer la transition vers la définitude sémantique.
(2006 : 110)

Pour notre part, aucun de ces arguments ne nous convainc. Quant au premier, à savoir la perte de la capacité à marquer la distance, nous avons vu dans la Section 1.3.2 que celle-ci n'est pas une propriété nécessaire du démonstratif. Ensuite, dire qu'un démonstratif « [...] s'est approprié les contextes "pragmatiquement" définis, où le référent est accessible parce qu'il est présent dans le contexte d'emploi [...] » est tout simplement une contradiction dans la mesure où ce fonctionnement constitue le fonctionnement de base du démonstratif, et ce depuis toujours et dans toutes les langues. Là encore, nous sommes d'avis qu'il ne faut pas renverser la charge de la preuve : c'est l'article défini « fort » (cf. Ebert, 1971 ; Schwarz, 2009 ; Löbner, 2011 ; Ortmann 2014 ; Studler, 2014)⁵⁰, c'est-à-dire celui dont l'usage est confiné au domaine de la définitude « pragmatique », qui – parce qu'il a évolué du démonstratif dans les premières phases du processus de grammaticalisation – continue à partager avec ce dernier les mêmes contextes d'emploi. S'agissant des emplois mémoriels, ils sont, à n'en pas douter, cruciaux dans la transition entre les deux types de définitude (cf. Himmelmann 1997 et 2001 ; Carlier et De Mulder, 2006, 2010 et 2011). En même temps, il semble que dans toutes les langues du monde, des démonstratifs puissent se prêter à ce type de référence indexicale, au point que Himmelmann (1996) le recense comme l'un des emplois de base de cette unité linguistique. Dans

⁵⁰ Le TEN adnominal du tchèque (informel) de nos jours emprunte souvent ce fonctionnement (voir notamment le Chapitre 6).

cette conjoncture, le fait que CE connaisse des emplois mémoriels ne saurait être considéré comme un argument étayant la thèse de son affaiblissement⁵¹.

À présent, focalisons-nous sur les emplois anaphoriques. Là aussi, l'usage du démonstratif est une norme dans les contextes mentionnés par Carlier et De Mulder dans la citation *supra* : le démonstratif est beaucoup plus naturel que le défini (ou l'absence de détermination dans les langues ne disposant pas d'articles définis) en reprise immédiate d'un référent précédemment introduit (cf. Christophersen, 1939 ; Corblin, 1983 ; Kleiber, 1986c et Himmelmann, 1996⁵²), a fortiori si cette reprise s'accompagne d'une (re)classification⁵³.

Disons maintenant un dernier mot au sujet de l'exemple d'Aragon susmentionné :

22) Il faudra que je vous montre **ce tableau** que j'ai acheté. (emprunté à Carlier et De Mulder, 2006 : 103)

Même si les auteurs semblent analyser cet emploi de CE comme purement cataphorique, nous ne partageons pas cette lecture. Pour nous, le démonstratif est à la fois cataphorique en ce qu'il renvoie en aval, à la subordonnée, et mémoriel en ce qu'il renvoie en amont, à une connaissance censée être partagée par le locuteur et l'interlocuteur⁵⁴. L'on remarquera aussitôt la différence si *ce* est remplacé par *le*. Dans ce dernier cas, la relative pourra, à notre avis, davantage être analysée comme une relative « établissante », quoique là aussi, le renvoi à une connaissance (censée être) partagée ne puisse être exclu.

Tout compte fait, la thèse de Carlier et De Mulder, d'après laquelle le déterminant démonstratif français serait un article défini en devenir, ne nous convainc pas plus que la thèse, beaucoup plus radicale, de Harris, même si nous comprenons une partie des motivations à l'origine de ces deux thèses.

4.4 L'émergence du SN en français

En français, l'évolution du système de la détermination nominale va sans conteste de pair avec l'évolution de l'unité syntaxique couramment appelée par les linguistes et les grammairiens

⁵¹ Les démonstratifs anglais connaissent également ces emplois, etc.

⁵² "This strategy for the introduction of a new participant seems to be very common in languages where no definite article exists. [...] But in languages such as English where a definite article exists, there is as well 'a certain aversion to the use of a the-form immediately after the word is introduced; a demonstrative is more usual in such cases' (Christophersen, 1939: 29). Hence, it is more common to find sequences such as *Once upon a time there was a king. This king had...* instead of *The king had...*" (Himmelmann, 1996 : 229).

⁵³ Voir les Sections 1.2.2 et 4.5.2 pour des exemples français et la Section 3.3.2 pour des exemples tchèques.

⁵⁴ Si, de fait, la connaissance n'est pas partagée, cela ne change pas grand-chose au fonctionnement du démonstratif, l'essentiel étant que le locuteur fait comme si elle l'était. Nous nous sommes positionné vis-à-vis des cas où le démonstratif introduit un nom modifié par une relative dans la Section 1.2.3.

« syntagme nominal »⁵⁵ (voir la Section 3.1.3 pour l'analyse de Himmelmann (1997) appliquée aux langues contenant des articles définis). Nous trouvons important de nous y attarder un moment, car, comme cela a déjà été suggéré dans le Chapitre 3 (Section 3.1.3), des similitudes considérables semblent exister entre la situation antérieure au français moderne (notamment celle du latin) et la situation du tchèque actuel. Nous nous référerons ici essentiellement au travail de Goyens (1994), qui a le mérite d'être une étude empirique. L'auteure a procédé à la comparaison d'un texte latin classique (*De Inventione* de Cicéron) avec sa traduction par Jean d'Antioche en ancien français (fin du 13^{ème} siècle) ainsi qu'avec plusieurs de ses traductions en français classique et moderne (allant du 17^{ème} au 20^{ème} siècle). En guise de résumé, citons un passage de la Conclusion de l'étude de Goyens (1994 : 279) :

Ce sont en effet les déterminants, et plus en particulier les actualisateurs, qui permettent en français moderne de préciser la référence du N, un élément indispensable à toute communication linguistique si celle-ci veut réussir. Toutefois, en latin, un tel système référentiel n'existait pas encore : la référence y est le plus souvent implicite, avec l'aide éventuelle de certains facteurs contextuels. En ancien français, un stade intermédiaire entre le latin et le français moderne, le système des actualisateurs est déjà plus développé, mais pas encore de façon aussi symétrique qu'en français moderne⁵⁶.

Quelques précisions s'imposent. S'agissant du terme d'« actualisateur », il recouvre globalement celui de « déterminant central »⁵⁷. Dès l'ancien français, le « déterminant zéro » (Ø) figure parmi les actualisateurs. Ce choix est justifié par l'auteure :

Ce signifié sans signifiant doit être vu en relation avec les signifiants avec lesquels il alterne dans la même classe et qui sont effectivement réalisés : sa fonction est essentiellement différentielle, c'est-à-dire qu'il signifie ce que les autres actualisateurs ne signifient pas. (1994 : 43)

Il se dégage de ce constat qu'il est légitime de parler de déterminant zéro uniquement pour les systèmes où l'actualisation du nom dans le discours au moyen de déterminants est devenue une contrainte grammaticale, l'expression de la détermination nominale étant ainsi devenue systématique⁵⁸. D'où s'ensuit qu'un tel terme se révèle sans pertinence pour les langues comme le

⁵⁵ Ou « groupe nominal », éventuellement aussi « syntagme déterminatif » (“determiner phrase”) dans les approches adhérent à l'hypothèse selon laquelle la tête du syntagme n'est pas le nom (ou autre élément substantivé) mais le déterminant (cf. Abney, 1987) ; dans le milieu français, cette position est également défendue par Fraser et Joly (1979 et 1980), même si ces derniers s'inscrivent dans une tradition linguistique très différente de la grammaire générative.

⁵⁶ Précisons que pour nous, un nom actualisé en discours – via la présence d'un actualisateur ou pas – ne se prête pas nécessairement à un acte de référence. Sur ce point, notre conception de la référence est plus restreinte et plus sélective que celle de Goyens (voir la Section 1.2.1 du Chapitre 1).

⁵⁷ « [...] nous distinguons les actualisateurs définis, à savoir les formes de l'article défini *le*, du démonstratif *ce* et du possessif *mon* et les actualisateurs indéfinis. Ces derniers se caractérisent tout d'abord par leur relation d'exclusion avec les actualisateurs définis [...] » (1994 : 32). Nous avons déjà vu *supra* que le terme de « déterminant », lui, reçoit chez Goyens l'acception wilmetienne.

⁵⁸ Encore faut-il distinguer la détermination zéro d'une simple absence de détermination, à savoir des cas où le nom est employé en pur emploi intensionnel, non-actualisée dans son extension (cf. Wilmet, 1986). On considère comme relevant de ce phénomène par exemple les emplois où le nom est « satellisé » : *étoile de mer*, *fil de paysan* ou les emplois attributifs du type *Pierre est pompier*. Comme le résume Goyens elle-même (1994 : 44) : « Concrètement, en français

latin et le tchèque. Goyens est bien consciente de cela, qui ne commence à l'employer, justement, qu'à partir de l'ancien français. Et même dans le cas de ce dernier, nous ne trouvons nullement exagéré d'interroger sa pertinence : étant donné le système beaucoup plus rudimentaire et parcimonieux de la détermination nominale en ancien français (une quasi-non-existence du partitif, une absence de l'article défini avec la plupart des noms génériques au pluriel, une quasi-absence de l'article indéfini UN avec les noms en emploi indéfini non-spécifique), est-il approprié de parler de déterminant zéro ? D'autre part, force est de constater que le complexe nominal⁵⁹ de l'ancien français affiche déjà une structure figée, obéissant à un agencement strict : c'est donc déjà un SN tel que le définit Himmelmann (1997 ; voir la Section 3.1.3). Ensuite, entre l'ancien français et le français moderne, cette unité syntaxique ne fera que se consolider à mesure que la détermination nominale se développera davantage (notamment à l'époque du moyen français).

En latin, en revanche, la situation est tout autre : « Le SN latin semble donc caractérisé par une assez grande liberté distributionnelle, qui est encore loin de la rigidité du SN du français moderne. » (Goyens, 1994 : 110) Quant à la possibilité des mots fonctionnels de s'y combiner entre eux, elle n'est pas sans rappeler le tchèque actuel. Voici deux exemples que nous empruntons à l'auteure. Chaque exemple est assorti de son équivalent en tchèque :

23) null-us nul-M.NOM.SG <i>aucun de mes livres</i>	me-us mien-M.NOM.SG	liber livre.NOM.SG
žádn-á nul-F.NOM.SG <i>aucun de mes livres</i> ⁶⁰	m-á mien-F.NOM.SG	kniha livre.NOM.SG
24) nostr-um nôtre-M.ACC.SG <i>notre célèbre Caton</i>	ill-um ce-M.ACC.SG	Caton-em NP-M.ACC.SG
toho ten.M.ACC.SG <i>notre célèbre Caton</i>	naš-eho nôtre-M.ACC.SG	Caton-a NP-M.ACC.SG

Comme nous le constatons en 23, la seule structure autorisée par le français pour rendre la structure latine est explicitement partitive : nous sommes face à deux SN (*aucun* et *de mes livres*),

moderne, on peut distinguer essentiellement deux types d'emploi où l'actualisateur Ø est utilisé, à savoir les cas où la référence est marquée par un autre moyen, et ceux où l'actualisateur Ø signifie, à proprement parler, la non-actualisation. » Nous tenons toutefois à dire que l'appartenance de certains phénomènes syntaxiques à l'un ou à l'autre cas de figure peut faire l'objet de choix interprétatifs et de désaccords.

⁵⁹ Pour éviter ici toute confusion : contrairement à nous, Goyens emploie ce terme exclusivement pour les SN comprenant d'autres SN (cf. *mes voisins* dans la maison de **mes voisins** ; 1994 : 9).

⁶⁰ De même en latin : *quis meus liber* et en tchèque : *která má kniha* → en français : *lequel de mes livres* (Goyens, 1994 : 105).

dont le premier prélève une quantité nulle sur un ensemble défini spécifique circonscrit par le second. Mais rien de tel dans la structure grammaticale latine, où *meus liber* est au nominatif singulier et non pas au génitif pluriel. Il en va exactement de même pour le tchèque. Du moins dans cette dernière langue, la lecture partitive n'est pas exclue sur le plan sémantique, mais elle n'est pas imposée. En 24, aussi bien en latin qu'en tchèque, le démonstratif se combine avec un possessif. Ce qui oppose en revanche les deux langues, c'est l'ordre d'apparition de ces deux éléments : en tchèque contemporain, c'est toujours le démonstratif qui précède le possessif et jamais l'inverse. Quant à la traduction française, si elle refuse de se débarrasser complètement de ILLE, elle est obligée de rendre la seule connotation de notoriété qu'il apporte, d'où l'adjectif *célèbre*.

Par ailleurs, le complexe nominal latin semble afficher une structure encore plus souple que le complexe nominal tchèque, dans la mesure où, par exemple, les possessifs et les démonstratifs peuvent plus facilement s'y retrouver postposés au nom⁶¹. Au vu de ces constats, il est étonnant que Goyens emploie le terme de SN pour le latin. Ce choix nous semble d'autant plus étonnant qu'elle s'applique justement à montrer comment le SN a vu le jour en français, à l'issue d'un figement syntaxique⁶² et d'une forte grammaticalisation de la détermination (actualisation) nominale. Sans doute faut-il voir dans cette appellation le résultat de l'alignement sur une certaine tradition de grammairiens français, telle Fugier (1983), qui continuent à appliquer l'appareil d'analyse propre à leur langue maternelle à la réalité d'un système qui, du point de vue typologique, n'a que très peu à voir avec celle-ci.

Pour conclure : d'abord, les complexes nominaux non-actualisés par un mot fonctionnel, majoritaires dans le texte originel latin, font de plus en plus place, dans les traductions, à des SN introduits par un mot fonctionnel⁶³. Goyens (1994 : 281) de résumer que :

⁶¹ Cf. *liber meus* vs *tuum patrem uidisti* (Goyens, 1994 : 73) pour le possessif et *homines illos quos Waiofarius ad defendendam ipsam ciuitatem dimiserat* vs *illam parentem Francorum* (Carlier, De Mulder, 2007 : 97) pour le démonstratif. Rappelons que si cet ordre n'est pas proscrit en tchèque, il reste, de nos jours, extrêmement minoritaire et très marqué pour ses connotations livresques (cf. Veselovská, 2014). Pour plus d'informations sur l'ordre des mots en latin classique, voir notamment Spevak (2010). Nous retrouvons ici l'opposition entre langues configurationnelles (français) et celles qui le sont beaucoup moins (latin, tchèque). D'autres phénomènes témoignent de cette dichotomie, tels que la possibilité de coordonner, en latin et en tchèque, le possessif avec un adjectif ou un complément du nom ou la facilité avec laquelle ces deux langues recourent à l'ellipse, contrairement au français (Carlier et Combettes, 2015).

⁶² Figement qui fait visiblement défaut au latin, car l'auteure constate : « En ce qui concerne la distribution des déterminants, la distinction [...] que nous avons établie pour l'ancien français et le français moderne, n'est guère applicable au latin, car tous les déterminants peuvent figurer devant ou derrière le N. On peut toutefois cerner les tendances suivantes : les démonstratifs [...], les possessifs [...] et les numéraux [...] précèdent le plus souvent le N, et l'expansion le suit en général ; mais il ne s'agit que d'une tendance et non d'une contrainte stricte. » (1994 : 207) Il est toutefois des auteurs qui restent bien plus prudents quant à l'usage du terme de « syntagme nominal » en rapport avec le latin (cf. Spevak, 2010 ; Carlier et Combettes, 2015). Ainsi, ces derniers écrivent : « Tout comme le verbe dans sa relation au VP, la tête nominale est peu dominante dans sa relation au NP, au point qu'il est parfois difficile voire non pertinent de l'identifier. [...] Par ailleurs, comme le montre l'étude d'O. Spevak (2010 : chap. 6), les différents types de modificateurs du nom n'ont pas de position structurelle fixe au sein du NP. » (2015 : 43).

⁶³ En ce qui concerne les noms dépourvus d'épithètes et d'expansions caractérisantes, ils étaient plus de 70 % à ne pas être introduits par un mot fonctionnel dans le texte latin, contrairement à environ 30 % en ancien français et seulement 11 % en français moderne.

L'explication de ce phénomène doit être cherchée dans le système de référence implicite du latin : la référence est en effet basée essentiellement sur des données contextuelles ; le nombre important d'interprétations différentes du texte-source de la part de deux traducteurs montre que ce système de référence implicite est souvent ambigu ; de là qu'un système de référence explicite, à l'aide d'actualisateurs précis, se développe en français.

Et, ensuite, les mots fonctionnels, désormais soumis à des contraintes distributionnelles de plus en plus strictes, en viennent à acquérir un fonctionnement de déterminants, la structure du complexe nominal se figeant pour devenir un véritable SN. De fortes ressemblances existent ici avec la situation du tchèque contemporain, langue où 1) les éléments fonctionnels peuvent commuter entre eux et où le complexe nominal garde une souplesse certaine et où 2) de nombreux noms apparaissent sans actualisateur aucun, leur interprétation sémantico-référentielle étant implicite et tributaire du contexte⁶⁴.

4.5 La distribution des déterminants démonstratifs français dans les différents régimes d'indexicalité

De même que dans le Chapitre 3 (Section 3.3), nous adopterons ici un plan basé sur des distinctions essentiellement onomasiologiques pour donner un premier aperçu de l'usage du déterminant démonstratif en français. Loin que nous pensions que l'approche onomasiologique soit la meilleure pour décrire le fonctionnement sémantico-référentiel du démonstratif (voir à cet égard la Section 1.1), il nous semble néanmoins qu'elle permet d'aborder de façon claire et systématique les différents types d'emplois dans lesquels le (déterminant) démonstratif est susceptible d'apparaître. Dans le domaine de la linguistique francophone, ce point de vue est adopté, entre autres, par Fraser et Joly dans leurs deux articles intitulés « Le Système de la deixis. Esquisse d'une théorie d'expression en anglais » et « Le système de la deixis. Endophore et cohésion discursive en anglais » (1979 et 1980). Nous avons déjà partiellement recouru à la terminologie de ces deux auteurs pour esquisser la situation du tchèque, nous continuerons donc à y recourir dans la présente section, notamment en employant l'opposition entre « exophore » et « endophore ».

4.5.1 Les emplois exophoriques situationnels, temporels et textuels

Pour ce qui concerne les emplois situationnels, une ligne de démarcation claire et nette semble exister entre l'emploi du démonstratif et du défini. Comme nous l'avons déjà montré dans la Section 1.2.2, cette ligne de démarcation est celle qui sépare la référence par « contraste externe » et la référence par « contraste interne ». Ainsi dira-t-on, dans une situation où seule une pochette bleue est présente :

⁶⁴ Cf. Šimík et Demian, 2020.

25) Passe-moi *la pochette bleue*, s'il te plaît.

Contrairement à celle où il s'agit de choisir parmi plusieurs pochettes bleues :

26) Passe-moi *cette pochette bleue*, s'il te plaît.

De même dans l'exemple suivant – qui est une traduction française d'une séquence de dialogue du corpus ORTOFON v1 citée dans la Section 2.3.1 – est-ce l'article défini qui est employé, les référents des SN *la petite table* et *le radiateur* étant identifiés via un recours à la situation d'énonciation, où ils apparaissent comme les seuls représentants de leur classe :

27) L1 : bon, je l'ai placée contre l'Est, mais quand elle est là, sur *la petite table* près *du radiateur*

L2 : bon, les orchidées n'aiment pas [...] je veux dire elles n'aiment pas autant de chaleur, tu sais ?

ORTOFON v1, notre traduction

S'agissant de la présence des particules localisatrices *-ci* et *-là*, nous renvoyons le lecteur à la fin de la Section 4.1.1 pour une introduction et à la Section 4.1.3 pour une analyse plus poussée. Outre que notre corpus d'oral informel confirme l'idée généralement admise de la neutralisation de l'opposition de distance (la plupart des occurrences ne contiennent ni *-ci* ni *-là*), il étaye également le constat de la disparition des formes en *-ci* au profit des formes en *-là*, tendance qui aboutit tantôt à une alternance des deux formes au sein d'un même contexte d'usage, tantôt à une éviction complète de *-ci* par *-là*. Le second scénario est illustré par l'exemple 18 *supra*. Voici maintenant un exemple du premier scénario. Ici, deux locutrices planifient leur voyage en Italie. Assises devant le plan d'une ville, elles parlent de différents endroits qui y sont représentés et qu'elles s'appêtent à visiter. L'on peut remarquer qu'elles emploient tantôt des formes en *-ci* et tantôt des formes en *-là*, mais, vu la situation, il semble raisonnable de supposer que c'est toujours pour renvoyer à un référent proche :

28) ah voilà donc il y aurait *ce quartier-ci* mais on ne saura pas tout faire hein CEFC

29) on va rester dans *ce coin-là* hein

CEFC

Cependant, alternance ne veut pas dire équivalence de sens : comme cela a déjà été suggéré dans la Section 4.1.1, alors que *-là* n'exprimerait ici ni une distance proche ni une distance lointaine, *-ci* insisterait sur la proximité.

Par ailleurs, nous avons remarqué qu'il y a une tendance à l'apparition de *-là* lorsque le locuteur marque une distance psychologique, pragmatique vis-à-vis du référent du SN démonstratif : en 30, la locutrice reproduit a posteriori et sous la forme d'un discours direct ses pensées du moment où sa voiture est tombée en panne, et où un individu masculin inconnu lui a proposé de l'emmener, dans sa voiture à lui, acheter une pièce manquante pour la réparer. La locutrice fait part d'une certaine méfiance à l'égard de l'inconnu :

- 30) il me dit écoute moi ce que je te propose la pièce moi je sais ce que c'est par contre c'est une pièce qui coûte cinq cents balles mais euh moi je t'emmène on va l'acheter je te la mets puis tu vois là j'étais c'était maintenant quoi c'était maintenant ou jamais puis je fais attends **ce gars-là** et tout quand même je le connais pas CEFC

Comme il s'agit d'un discours direct (introduit par le verbe FAIRE, faisant ici office de verbe de parole), l'occurrence de *ce gars-là* doit être analysée comme exophorique situationnelle et non pas comme anaphorique, bien que, pour l'interlocutrice à qui l'histoire est racontée a posteriori, l'identification du référent de ce SN se fasse grâce au contexte antérieur. Un peu plus tard, la même méfiance subsiste chez la locutrice après qu'elle s'est décidée à monter dans la voiture de l'inconnu pour aller chercher avec lui la pièce manquante :

- 31) il me fait ben écoute on va aller en acheter une parce que là moi j'en ai vraiment pas faut aller en acheter une alors nous voilà partis direction Toulouse là enfin sur les pa sur les panneaux il y avait écrit Toulouse alors moi je fais qu'est-ce que c'est que **cette histoire-là** CEFC

Souvent, le déterminant démonstratif apparaît aussi, en français, en emploi exophorique temporel. Cet emploi représente une part considérable de notre échantillon pour l'oral informel (282 sur 1 000, contre seulement 6 occurrences sur 1 000 pour le tchèque ; voir les Sections 5.2.2.1 et 5.2.3.1). Disons d'emblée que l'une des principales raisons de cette disproportion est le fait que le français emploie le démonstratif avec bien plus de noms à sémantique temporelle que le tchèque. Ainsi, les expressions *ce matin*, *ce midi*, *cet après-midi*, *ce soir*, *cette nuit*, *ce week-end* et *cette année* sont toutes rendues sans démonstratif en tchèque : *ráno*, *v poledne*, *odpoledne*, *večer*, *v noci*, *o víkendu*, *letos*⁶⁵. Cette différence structurelle est incontestablement lourde de conséquences sur le bilan de la comparaison entre les deux langues. Sur le plan syntaxique, ces SN démonstratifs tendent à apparaître en position détachée (207 sur 282), soit en début soit en fin d'énoncé. Sur le plan sémantique, ils agissent comme des compléments de l'ensemble de l'énoncé, renseignant sur des circonstances temporelles de la prédication véhiculée par celui-ci (voir la Section 2.2.3 du Chapitre 2 et la Section 5.2.2.2 du Chapitre 5) :

- 32) moi **ce matin** ah j'avais vraiment pas envie de me lever CEFC

- 33) ouais il faut que je fasse tous les exos **ce soir** CEFC

CE apparaît aussi en deixis textuelle, qui, rappelons-le, consiste à référer soit au discours soit à la langue envisagés dans leur matérialité :

- 34) L1 : ils avaient pas décidé de passer à l'euro les Anglais fierté euh fierté du fierté britannique c'est ça hein tout ce qui tou-

⁶⁵ Rappelons que le mot *letos* comporte l'ancienne forme de démonstratif SEN. Cette forme est suffixée au mot *léto* (été) ; voir la Section 3.2.

L2 : fierté britannique an~ anti-european euro européenisme an comment euh anti-europ comment
 ça se va se dire *ce mot* CEFC

- 35) L1 : tu me fends le cœur ça vient de ce film
 L2 : ah je connais pas *cette réplique* CEFC

4.5.2 Les emplois endophoriques

Pour ce qui est de l'usage du déterminant démonstratif français en anaphore, ce travail n'aspire évidemment pas à un traitement exhaustif de la problématique. Cette dernière a fait l'objet d'une série d'études (cf. Corblin, 1983 ; Kleiber, 1986c), articulées notamment autour de la concurrence entre le démonstratif et le défini dans l'anaphore coréférentielle. D'une manière générale, il s'avère qu'en reprise immédiate d'un référent nouvellement introduit dans le discours, c'est le démonstratif qui est (de loin, semble-t-il) préféré au défini :

- 36) a) J'ai vu une voiture. *La voiture* roulait vite.
 b) J'ai vu une voiture. *Cette voiture* roulait vite. (Kleiber, 1986c : 54)

En 36, b) est de loin préféré à a). En même temps, ce qui a suscité beaucoup de commentaires est la possibilité, pour le SN démonstratif, de commuter ici avec le pronom personnel. En effet, comme le fait remarquer Kleiber, *Elle roulait vite* est tout aussi bien formulé. Depuis Corblin (1983), on parle ainsi du « paradoxe de la reprise immédiate ». Ce paradoxe est dû au fait que la commutation soit possible entre deux expressions anaphoriques reflétant un degré d'accessibilité référentielle très différent, le pronom personnel marquant une accessibilité bien plus élevée que le démonstratif (cf. Ariel, 1988), plus grande encore que le SN défini anaphorique (cf. Yule, 1981). Dans les reprises ultérieures d'un référent établi par le discours, en revanche, c'est le défini qui est préféré au démonstratif (voir la Section 1.1.3). Cette tendance est illustrée par l'exemple suivant :

- 37) et euh il y a une affaire d'[un tueur en série qui euh qui qui dépèce ses victimes en fait qui leur qui leur enlève la peau]_i
 ...
 et hum et en fait il accepte de lui donner des renseignements sur euh sur [*ce* sur *ce tueur en série*]_i
 si elle en échange
 ...
 bon et au final en fait on s'aperçoit que [*le mec*]_i là [*le le tueur*]_i en série il se fait un manteau en peau de femme CEFC

La première reprise du référent emploie le démonstratif⁶⁶, contrairement à la deuxième, qui emploie le défini⁶⁷. Toujours conformément aux constats des approches cognitives de l'anaphore, il s'avère ensuite que si le référent a été désactivé (il n'a pas été question de lui depuis un certain temps), il est de nouveau repris par le démonstratif à son retour dans le discours :

- 38) L1 : oh j'ai lu un truc qui m'a choquée quoi dans un livre c'est un roman c'est euh Anna Karénine Tolstoï tu vois c'est russe ça se situe au dix-neuvième siècle et euh c'est euh euh [un jeune un jeune homme là riche qui monte euh qui aime faire des courses à cheval]_i
 L2 : ouais
 L1 : donc euh là [*il*]_i est en train de monter et de faire une course d'obstacles et [*il*]_i est mal retombé sur *son* cheval et [*il*]_i lui a brisé les reins
 ...
 L2 : mais il pesait combien [*ce cet homme*]_i CEFC

En 38, le référent est d'abord repris par *il* à partir de sa deuxième mention, reprise qui marque son statut activé, de même que la présence du possessif *son*, renvoyant également au jeune aristocrate⁶⁸. Le sujet de la conversation s'éloigne ensuite de l'épisode pour embrasser une réflexion plus générale au sujet de la possibilité pour ce type d'événements de se produire. Quatorze prises de paroles séparent ainsi la dernière mention de l'homme de sa réintroduction par le démonstratif⁶⁹.

La règle que nous venons ici de reproduire doit cependant être considérée avec beaucoup de circonspection, voire de méfiance, comme le travail avec nos données nous l'a fait voir. En réalité, la concurrence entre les différentes expressions de l'anaphore coréférentielle fidèle, y compris celle entre le démonstratif et le défini, s'avère plus complexe, un locuteur pouvant recourir à toutes sortes de stratégies qui lui permettent de mettre en relief un référent, le comparer explicitement à d'autres référents, etc. C'est ce qui semble se passer en 39, où *un festival* est d'abord repris par *le festival*⁷⁰ pour seulement ensuite être repris par *ce festival* :

- 39) L1 : hm donc là ça va ça reste correct mais euh je pensais que ça serait moins vu que [*c'*]_i est un festival tu vois
 L2 : ben ouais mais en fait c'est déjà bien parce que regarde ils font une tournée en France et
 L1 : pas ben c'est les plus c'est les plus chers [*du festival*]_i hein je sais pas si tu as vu

⁶⁶ Il faut cependant remarquer que cette reprise ne suit pas immédiatement l'introduction du référent.

⁶⁷ Et cela même si, dans le cas de *le mec*, il s'agit d'une anaphore infidèle. Toutefois, le novum informationnel est ici minimal, la recatégorisation du référent n'étant pas du tout radicale.

⁶⁸ Relevons les deux occurrences du pronom relatif qui aident également à installer le référent dans le focus de l'attention.

⁶⁹ De plus, le SN démonstratif apparaît dans une dislocation à droite, construction qui insiste, comme nous l'avons vu dans la Section 2.2.3 du Chapitre 2, sur la continuité de la chaîne référentielle et la nature thématique du référent : ce statut de thème de la conversation est donc encore plus accentué que si le démonstratif apparaissait en dehors de la dislocation.

⁷⁰ En vérité, parler de reprise ne va pas de soi, car le SN *un festival* apparaît en position d'attribut, donc non-référentielle. Faut-il donc considérer que ce n'est que l'occurrence du SN *le festival* qui instaure la référence ? Ou bien ce SN défini reprend-t-il plutôt le référent du démonstratif neutre *ce* de la construction *c'est un festival* ? Ou encore : le SN *le festival* ne serait-il pas du tout porteur d'une référence anaphorique mais mémorielle ? Voilà les difficultés auxquelles nous nous sommes heurté lors de l'annotation et que nous exposerons plus en détail au début du Chapitre 5.

L2 : oui je sais mais euh ils pa~ enfin dans toutes les villes où ils passent ils passent dans dans des Zénith dans des trucs très chers et là NNAAMMEE le fait qu'ils soient dans [*ce festival*]_i c'est beaucoup moins cher

L1 : aucune idée mais je crois que c'est euh les artistes les plus euh attendus [*du festival*]_i un truc comme ça
CEFC

La raison de l'apparition du démonstratif est sans doute à chercher du côté d'un contraste que le locuteur vise à évoquer, à travers la seconde reprise du référent, entre le festival en question et d'autres lieux où le groupe se produit sur scène pendant sa tournée. Une fois que cette mise en contraste est terminée, le référent peut de nouveau être repris par un SN défini. Si nous devons traduire le texte du dialogue en tchèque, nous opterions spontanément pour *ten festival* là où le français emploie le festival et pour *tenhle festival* là où il emploie *ce festival*. Cela nous ramène au constat de la Section 3.3.2, à savoir que la situation du tchèque est plus complexe du fait de connaître une opposition à trois termes (Ø/TEN/TENTO (TENHLE)) tandis que le français n'en contient que deux⁷¹. Dans ces cas, où il s'agit d'établir une opposition ou un contraste, TENTO/TENHLE⁷² sera utilisé en tchèque et non pas TEN, qui, quant à lui, sera susceptible d'apparaître à la place du défini français. Nous pouvons ainsi observer, pour ce type de situations, un décalage symétrique dans les valeurs sémantico-référentielles et leur expression : LE ↔ Ø/TEN vs CE ↔ TEN/TENTO(TENHLE).

À présent, il faut préciser que les constats que nous avons faits jusqu'ici s'appliquent uniquement aux cas de l'**anaphore coréférentielle fidèle**. Si le nom de la reprise anaphorique est autre que celui de l'expression antécédente, le référent se voit recatégorisé, et la probabilité de l'apparition de CE en deviendra encore plus élevée :

40) mais moi hier j'ai regardé euh je savais pas quoi foutre du coup j'ai regardé la télé il y avait euh [Coco]_i il est vraiment bidon [*ce film*]_i
CEFC

41) [Leipzig]_i [c']_i est une ville qui est assez euh enfin pff il y a encore l'empreinte de de l'ex Allemagne de l'Est et tout et [c']_i est une ville qui est vachement sombre tu vois qui est vachement ouais ouais tous les monuments c'est c'est c'est masse

...
c'est gigantesque mais c'est c'est très enfin c'est pesant comme atmosphère [*cette ville*]_i [elle]_i est pesante
CEFC

42) ils se sont cachés dans [les montagnes qu'il y a vers euh Santiago de Cuba à l'Est de Santiago]_i je sais plus comment [elles]_i s'appellent putain mais [elles]_i sont magnifiques c'est [là]_i où ils ont fait les premiers plants de café c'est le café c'est les français qui ont été chassés de Haïti tu sais quand Haïti a acquis son indépendance mh mh attends c'est pas son indépendance enfin il y a eu un moment où ils ont été chassés par euh par un un nationaliste haïtien tous les colonialistes français

⁷¹ Tous les autres moyens de l'anaphore coréférentielle étant mis de côté.

⁷² Éventuellement aussi, bien entendu, les autres formes de proximité comme TENHLETEN, TADYHLETEN, etc. (voir la Section 3.1.1 du Chapitre 3).

et ils ont fui ils ont atterri à Cuba à l'Est donc vers Santiago de Cuba et euh c'est eux qui ont fait les premières plantations de café dans ces dans [*ces massifs de montagnes*]_i CEFC

Il va de soi que dans de tels cas, peu importe que le référent soit solidement établi dans le discours comme en 41 et 42, le démonstratif apparaîtra tout de même, car c'est lui qui garantit le lien de coréférence, « fragilisé » par la recatégorisation. Rappelons qu'en tchèque, l'anaphore infidèle est un domaine important de TENTO/TENHLE (voir la Section 3.3.2).

Contrairement au tchèque, le français s'oppose à l'emploi du démonstratif dans la référence par contraste externe (cf. Corblin, 1987). C'est le défini qui, cette fois-ci, apparaîtra beaucoup plus naturellement :

- 43) J'ai acheté [une voiture]_i et [une moto]_{ii}. [*La voiture*]_i roule parfaitement, mais il y a tout le temps des soucis avec [*la moto*]_{ii}. (exemple inspiré par Pešek, 2014)

Selon Corblin (1983), la notion de « contraste externe » permettrait également de rendre compte de la présence du défini dans certains contextes où l'on avait, auparavant, du mal à l'expliquer :

- 44) Il était une fois [un prince]_i très malheureux : [*le prince*]_i aimait une belle princesse qui ne l'aimait pas. (Corblin, 1983 : 131)

En 44, le défini apparaît, selon Corblin, du fait d'un contraste externe entre les référents des SN *un prince* et *une belle princesse*. Il s'avère ainsi que les deux référents entre lesquels le contraste s'établit ne sont pas nécessairement introduits dans la même proposition⁷³.

S'agissant de l'anaphore résomptive, elle est abondamment représentée dans notre corpus d'oral informel (170 occurrences) :

- 45) L1 : ils finissent menottés en fait
L2 : mouais
L1 : elle à lui à Hannibal Lecter mh et il prend un couteau de boucher et il dit soit vous vous soit vous enlevez la clé enfin soit vous enlevez les menottes avec la clé soit je vous coupe la main donc là la meuf elle dit non jamais jamais je vous relâche quoi et en fait c'est lui qui va se couper sa main à lui pour pouvoir partir
L2 : purée
L1 : et c'est trop intense *ce passage* CEFC

- 46) euh je sais pas je sais que à Marseille ils font euh espagnol allemand et italien *ce genre de langues* CEFC

Une série de noms sont particulièrement susceptibles de se prêter à ce type d'emploi anaphorique, dont notamment les noms *cas*, *genre* et *moment*⁷⁴. Il faut également relever que c'est

⁷³ Comme le résume Kleiber (1986c : 54) : « L'article défini anaphorique, opérant par contraste externe de domaine à domaine, nécessite la présence d'au moins un autre SN dans le contexte antérieur ou postérieur immédiat. »

⁷⁴ Nous examinerons mieux ce résultat et d'autres encore dans la Section 5.2.2.1 du Chapitre 5.

dans ce type d'anaphore que le démonstratif apparaît le plus souvent avec la particule *-là* (60 occurrences) :

- 47) elle arrive et elle me fait euh c'est quoi la plage la plus proche de NNAAMMEE je fais euh ce sera une plage à Marseille dans *ce cas-là* CEFC

Les contextes d'anaphore associative sont réservés, en français, à l'article défini. Aussi aucune occurrence de ce type n'est-elle présente dans notre corpus. Quant aux emplois cataphoriques, le démonstratif français s'y prête, évidemment :

- 48) Pendant mon séjour en Italie, j'ai visité *ces villes-ci* : Milan, Naples et Padoue.

Mais dans notre corpus d'oral informel, nous n'avons identifié que deux emplois qui peuvent, selon nous, être analysés de la sorte :

- 49) à côté de chez lui il y a un espèce d'immense Carrefour et euh il y a plein de de de de restaurants pas chers euh genre vietnamiens coréens et tout et ils sont allés manger des sushi dans [*ce restaurant*]_i et où il avait [un restaurant de sushi]_i CEFC

- 50) j'ai prié aussi pour ça euh pour que à *ce moment-là* si il veut que je reste toute seule que que je souffre plus CEFC

En 49, le SN *ce restaurant* peut s'analyser comme un emploi cataphorique, car sa référence s'interprète moyennant le SN indéfini qui suit : *un restaurant de sushi*. Reconnaissons néanmoins qu'une interprétation mémorielle n'est pas non plus à écarter complètement, dans la mesure où la locutrice semble, dans un premier temps, en appeler à une connaissance partagée et seulement après, lorsqu'elle réalise que l'interlocuteur ne la partage pas, elle recourt à une explication : *il y avait un restaurant de sushi*. En 50, à *ce moment-là* réfère de manière résomptive, résumant le contenu référentiel de la proposition *si il veut que je reste toute seule*. Les trois exemples 48, 49 et 50 sont des cataphores extra-syntagmatiques, renvoyant au-delà de la frontière d'un SN. La cataphore intra-syntagmatique avec le démonstratif, en revanche, ne se trouve pas en français, du moins pas dans sa version pure (contrairement au tchèque, où c'est l'un des contextes d'apparition de prédilection de TEN, comme nous l'avons vu dans la Section 3.3.3). Si cataphore intra-syntagmatique il y a, celle-ci se combine toujours avec une valeur mémorielle :

- 51) Tu te rappelles *cette vieille dame* qui roulait toujours à vélo en jupe ?

Rappelons que dans les exemples du type 51, la sémantique indexicale du démonstratif renvoie à la fois en aval (à la subordonnée) et en amont (à une connaissance partagée entre les personnes de l'interlocution). Dans notre corpus, ces occurrences du démonstratif ont été analysées comme mémorielles, car nous estimons que c'est la connaissance partagée qui joue un rôle plus important

dans l'identification du référent, l'information apportée par l'expansion caractérisante jouant un rôle plus complémentaire au sein de ce processus⁷⁵.

Pour revenir ici brièvement à la présence des particules localisatrices, des 488 occurrences que nous avons analysées comme relevant du domaine de l'endophore, seules 4 contiennent *-ci* contre 146 qui contiennent *-là*. Comme déjà indiqué *supra*, la plupart des occurrences de *-là* apparaissent avec l'anaphore résomptive (60 occurrences), suivie de l'anaphore fidèle (31 occurrences), de l'anaphore infidèle et de ce que nous avons pris la décision, dans la Section 2.2.3 d'appeler, l'« anaphore temporelle » (25 occurrences dans les deux cas). Cette dernière mérite que l'on s'y attarde un peu. Voici deux exemples :

52) ils avaient pas décidé de passer à l'euro les Anglais fierté euh fierté du fierté britannique c'est ça
...
c'est pas genre attends que je dise pas de bêtises euh il y avait encore Thatcher ou pas à **cette époque-là** CEFC

53) L1 : on est allés à une plage c'était une plage pas de sable fin on va dire de petits graviers
L2 : ouais oh ça fait un peu plus mal déjà ça quand tu t'installés
L1 : donc ça j~ j'aime j'aime pas trop mais alors **ce jour-là** il y avait un petit peu de vent
CEFC

Ces emplois du déterminant démonstratif ont ceci de particulier que leur interprétation dépend nécessairement du contexte amont, mais qu'en même temps, aucun contenu mentionné par ce dernier n'est directement repris par le SN démonstratif. C'est pour cela que ce type d'emploi ne saurait non plus être subsumé sous l'anaphore résomptive, car le SN démonstratif ne crée pas un nouveau référent par la synthèse d'un contenu mentionné. Ajoutons que les 4 occurrences anaphoriques de CE assorties de *-ci* relèvent de ce type d'emploi :

54) L1 : ouais ouais qu'est-ce que tu as fait ce week-end
L2 : j'ai été à la Japan Expo toi aussi
L1 : non ah ouais tu as vu quoi ouais ouais j'y j'y ai été
...
L1 : ouais c'était pas mal hein franchement c'était vraiment bien ils ont ils ont c'était mieux que l'année dernière pour moi
L2 : oui c'était il y avait plus de choses
L1 : **cette fois-ci** il y avait ouais ouais CEFC

⁷⁵ « Il s'agit de l'introduction d'un nouveau référent dans la mémoire immédiate [...], mémoire commune au locuteur et à l'interlocuteur au moment de l'énonciation du SN démonstratif. Ce référent nouveau ne l'est pas totalement, mais est supposé (ou supposé pouvoir être) déjà disponible dans la mémoire longue. Il se trouve réactivé ou réintroduit dans la mémoire immédiate discursive via l'expression d'une proposition portant sur ce référent, proposition supposée également faire partie du stock des propositions déjà acquises dans l'univers de croyance [...] de l'interlocuteur. » (Kleiber et Sock, 2006 : 261)

Ces emplois constituent, par ailleurs, une preuve de plus que l'opposition *-ci* vs *-là* subsiste toujours, dans une certaine mesure, avec les noms à sémantique temporelle, même dans la langue orale informelle. En 54, la particule localisatrice de proximité permet d'insister sur la différence entre l'édition la plus récente de l'exposition et celle de l'année précédente (*cette fois-ci* vs *l'année dernière*). Ce contraste entre plusieurs événements du même genre caractérise, comme le montre Theissen (2008), l'emploi de *cette fois(-ci/-là)*⁷⁶.

Dans les régimes de l'anaphore fidèle et infidèle, la majorité des occurrences avec *-là* laisse clairement comprendre un contraste (explicite ou implicite) du référent à d'autres référents. C'est à ce moment-là que le tchèque emploierait de manière obligatoire TENTO/TENHLE :

55) c'est comme à la braderie je me souviens une fois à la braderie [la rue du temple]_i tu sais c'est [celle vers la vers la piscine]_i [là]_i il y avait eu ouais la la la la la piscine des elle est sur [la rue du temple]_i quand tu descends direction sur [*cette rue-là*]_i CEFC

56) lui il a revendu [un chalet qu'il avait en haut à à Crans]_i pour euh beaucoup plus que ce que ça lui avait coûté il il est tombé sur un qui a euh qui était peut-être multi-millionnaire puis qui voulait à tout prix c- /amorce/ [*ce chalet-là*]_i CEFC

57) j'ai [des shorts]_i qui sont euh je sais pas si euh je crois qu'[ils]_i sont trop petits parce que euh [ils]_i sont trop petits je peux [les]_i enlever sans enlever le bouton et la fermeture et et la la braguette elle s'ouvre toute seule donc j'ai je je je suis tout le temps avec la braguette ouverte
...
quand j'ai [*ces shorts-là*]_i hein pas tout genre pas pas avec tous mes shorts hein CEFC

En 55, la rue dont il est question est implicitement contrastée à d'autres rues (référents ressortissant à la même classe). C'est également ce contraste qui permet de rendre compte de la présence du démonstratif même si le référent a déjà été introduit comme défini, car identifiable indépendamment du discours (le SN défini *la rue du temple* est un défini « sémantique », identifiable au sein de la communauté linguistique de la ville où se déroule le dialogue). La situation est similaire en 56, où le locuteur insiste sur le référent : le multimillionnaire était précisément intéressé par l'achat de ce chalet en particulier et pas par un autre. En 57, le contraste est même très explicite : *quand j'ai ces shorts-là hein pas tout genre pas pas avec tous mes shorts hein*. Que le contraste ne soit cependant pas toujours présent – ou que, du moins, il soit (beaucoup) plus discret –, voilà qui nous semble prouvé par les exemples 58 et 59 :

58) L1 : et là il y a [plein d'arbres]_i autour et l'idée c'est de euh
L2 : ouais on coupe un arbre
L1 : c'est d'aller couper [*ces arbres là*]_i enfin une partie [des arbres]_i de faire du bois CEFC

⁷⁶ Les noms à sémantique temporelle feront l'objet d'une analyse plus détaillée à la fin du Chapitre 5 (Section 5.3).

59) bon donc voilà donc je travaille sur euh sur Coral-Rom là je finis de les les transcriptions quoi j'ai [des des bandes sonores]; j'ai fini les transcriptions de [*ces bandes sonores là*]; CEFC

Aussi dans ces exemples sera-t-il parfaitement naturel de traduire CE en tchèque par TEN, quoique TENTO/TENHLE n'est pas non plus exclu.

4.5.3 Les emplois exophoriques mémoriels

Contrairement au tchèque (parlé informel), présentant une situation équivoque dans la mesure où TEN y reçoit une valeur plus ou moins proche de celle de l'article défini français ou de celle du démonstratif (voir la Section 3.3.4), la situation du français est beaucoup plus transparente : CE reste un démonstratif à part entière, qui, contrairement au défini, procède à une indexation forte du référent sur le contexte :

60) *Cette femme*, elle t'aura tout pris ! (Kleiber, 1986d : 12)

61) il faudrait que je me que je me bouge un peu pour euh pour trouver *ce foutu article de quatre mille mots que je vais devoir traduire en japonais* CEFC

62) oh putain [c']est quoi aussi euh [le film avec]; oh putain euh comment s'appelle [ce film]; ah ah ah ah euh [là ouais où le gars il tombe euh toujours dans les pommes]; CEFC

Comme le fait souvent remarquer la littérature sur le sujet (cf. Himmelmann, 1996 ; Diessel, 1999), étant donné que le référent doit être récupéré par l'interlocuteur via l'activation d'une connaissance/expérience privée, partagée avec le locuteur, l'acte de référence a tendance à s'accompagner d'informations additionnelles susceptibles de faciliter cette récupération :

Furthermore, there is a tendency to incorporate additional anchoring or descriptive information into a recognitional mention to make the intended referent more accessible. Thus, recognitional use often involves relative clauses or other modifiers of similar complexity. (Himmelmann, 1996 : 230)

C'est ce que nous pouvons observer en 61 et en 62. Le SN démonstratif du 61 est assorti de deux expansions, à savoir *de quatre mille mots* et *que je vais devoir traduire en japonais*. En 62, où c'est le locuteur lui-même qui cherche à identifier le référent, le processus de récupération de ce dernier à partir d'un stock de connaissances partagées devient encore plus tangible : en témoignant non seulement la présence de l'information descriptive *là ouais où le gars il tombe euh toujours dans les pommes* mais également les autres commentaires, d'une nature à proprement parler mémorielle (*comment s'appelle*).

Selon Diessel (1999 : 106), les emplois mémoriels du démonstratif se distinguent des autres emplois par le fait qu'ils renvoient à un référent qui est nouveau sur le plan discursif mais « ancien » sur le plan de l'identification : « Recognitional demonstratives are specifically used to mark

information that is discourse new (i.e. unactivated) and hearer old (i.e. pragmatically presupposed)⁷⁷. » À cela s'ajoute le constat très important et déjà mentionné de la nature privée, intime du référent :

More precisely, recognitional demonstratives mark information that is (i) discourse new, (ii) hearer old, and (iii) 'private' (Himmelman uses the term "specific" rather than private). Private information is information that speaker and hearer share due to common experience in the past. It is distinguished from general cultural information shared by all members of the speech community. General cultural information is also hearer old at its first mention, but unlike private hearer old information it is marked by a definite article in English.

(Diessel, 1999 : 106)

Dans la précédente citation, Diessel oppose ainsi la référence mémorielle par le démonstratif aux cas de la référence appartenant à la définitude « sémantique » et marquée par le défini. Dans le cas de ce dernier et contrairement au démonstratif, le référent serait identifiable grâce à une connaissance partagée par l'ensemble des membres d'une communauté linguistique (voir par exemple le SN défini *la rue du temple* en 55 *supra*⁷⁸). Il convient, selon nous, de compléter ce bilan par le constat suivant, qui rend parfois la différence entre le démonstratif et le défini plus ténue et plus difficile à appréhender dans les productions langagières naturelles : certes le défini marque-t-il communément les référents identifiables indépendamment de l'expérience privée des interlocuteurs (contrairement au démonstratif qui, lui, n'accompagne que les référents stockés dans cette dernière), mais il peut également marquer des référents appartenant à cette sphère privée. C'est ce que nous voyons par exemple en 62, énoncé où *ce film* apparaît ensemble avec *le film*. De même le voyons-nous en 63, où l'identification du référent opère via le recours à une connaissance privée :

63) *Le gars* a encore téléphoné aujourd'hui.

Par conséquent, il n'est pas juste de dire que le défini ne soit pas à même d'accompagner une référence mémorielle privée, autrement dit, à introduire un référent « pragmatiquement » défini. Après tout, nous avons bien vu, dans la Section 3.3.4 du Chapitre 3, que le TEN adnominal des emplois mémoriels peut souvent être rendu en français par le défini, fait qui constitue une preuve de sa grammaticalisation, mais aussi d'une certaine zone grise entre les deux unités⁷⁹. Comment expliquer, dans ce cas, la différence, en 62, entre *le film* et *ce film* ? Une fois de plus, elle doit être cherchée,

⁷⁷ Pour une analyse similaire de l'ensemble de la référence mémorielle (qui ne se limite pas aux expressions démonstratives), voir Laury (1997). Himmelman (1996 : 240) écrit que contrairement aux emplois anaphoriques, où il est question de référents activés ou semi-activés, les emplois mémoriels du démonstratif réfèrent à des référents semi-activés ou non-activés. Le lecteur aura remarqué que cette manière de conceptualiser la différence puise dans les approches cognitives de l'indexicalité (voir la Section 1.1.3).

⁷⁸ Hawkins (1978) parle d'« emplois en situation large » ("larger situational use").

⁷⁹ Rappelons la remarque de Trovesi (2002 : 132), citée dans la Section 3.3.4 du chapitre précédent.

selon nous, du côté de l'accessibilité référentielle et de ce que certains auteurs (cf. Chafe, 1994 ; Laury, 1997) appellent le « coût d'activation » du référent. Comme l'explique Laury (1997 : 159) :

The concept of accessibility is closely related to the concept of *activation cost* (Chafe, 1994: 73), roughly, the amount of mental effort required of the addressee in order for him or her to become aware of, in other words, to bring to active consciousness, the idea of the referent mentioned by the speaker.

Il semble ainsi logique d'affirmer que dans une situation où l'identification du référent repose sur des connaissances de la sphère privée, le coût d'activation est moindre lorsque c'est le défini qui est employé, tandis que le démonstratif indique un coût d'activation plus élevé. En disant *le film*, le locuteur suggère que le référent peut être atteint par l'interlocuteur (éventuellement aussi par lui-même, comme c'est le cas ici) avec plus de facilité que lorsqu'il dit *ce film*, le démonstratif s'accompagnant d'un effort d'activation du référent plus conséquent. L'exemple 62 nous semble être une parfaite manifestation de cette tendance : le locuteur commence par utiliser le défini et ne passe au démonstratif qu'au moment où l'identification du référent s'avère une entreprise difficile. De même, *le gars* en 63 supposerait, selon nous, un degré d'activation plus élevé du référent que *ce gars*⁸⁰.

Le français (parlé) semble faire un usage certain d'un cas singulier du déterminant démonstratif mémoriel. Ce cas a fait l'objet de mentions et d'analyses dans la littérature (cf. Carlier et De Mulder, 2006). Il s'agit de la structure partitive *un/une/de ces N + proposition relative*, sachant que la relative est souvent sous-entendue, chose qui aboutit à un fort degré d'expressivité. Notre corpus d'oral informel contient 6 occurrences de ce type. La relative n'est présente avec aucune d'entre elles :

64) c'est une une enflure de première quoi elle me raconte ***de ces trucs*** CEFC

65) il y avait euh il y avait ***un de ces mondes*** CEFC

Remarquons que le nom introduit par le démonstratif construit une référence générique, bien qu'ancrée dans les connaissances privées des interlocuteurs.

En ce qui concerne la présence des particules *-ci/-là*, elle semble rare dans ce régime d'indexicalité. Du moins est-ce ce que montre notre corpus, où *-ci* n'apparaît jamais avec un emploi mémoriel et *-là* n'apparaît que 8 fois (dont 2 fois avec un nom à sémantique temporelle).

Comme nous l'avons déjà écrit à la fin de la Section 3.3.4 du Chapitre 3, certains emplois se présentent à première vue comme permettant d'identifier un référent par la voie mémorielle, alors qu'en vérité, ils ne font que renseigner sur une pure alternative virtuelle. Bien qu'utilisant la forme

⁸⁰ Remarquons au passage que la conception de l'« accessibilité référentielle » que l'on trouve chez Chafe (1994) ne serait pas toute à fait la même que celle que l'on trouve chez Ariel (1988 ; voir la Section 1.1.3 du Chapitre 1), pour qui les référents identifiés par le cadre de la mémoire sont – d'une manière que nous considérons un peu trop sommaire – rangés parmi les référents à accessibilité basse.

définie du démonstratif, la sémantique du SN est indéfinie et le SN ne réfère pas (dans l'approche héritée de Russell (1905), il est question d'un pur « emploi quantificationnel » ; voir également Fodor et Sag, 1982). Notre corpus français d'oral informel contient 9 occurrences de ce type, dont 5 cinq contiennent la particule localisatrice *-là* :

66) L1 : faut pas se laisser aller

...

c'est ça vu que bon comme tu es à la fac tu es libre tu fais un peu euh tu vas à ***ce cours***

L2 : ouais il faut pas se dire à chaque fois qu'il y a une heure de trou bon je vais je vais rien faire

CEFC

67) comme au lycée quand on faisait du ping-pong et il fallait lancer le la

...

la balle à ***cet endroit-là*** parce que sinon c'était pas bon

CEFC

4.5.4 Les emplois « émotionnels »

Nous finirons le présent chapitre avec un mot sur les emplois « émotionnels » du déterminant démonstratif français. Une introduction à ce sujet a déjà été présentée dans la Section 3.5, où nous avons traité des emplois « émotionnels » du démonstratif tchèque, et où une série d'exemples du français figurent déjà. C'est sur cette introduction que nous nous baserons ici. Rappelons notre parti-pris, en vertu duquel il convient de distinguer les emplois « désignatifs » du démonstratif, où celui-ci est porteur d'une subjectivité accrue, et les emplois purement « émotionnels », à savoir ceux où le démonstratif n'entre pas (directement) dans le processus d'identification du référent dénoté par le nom et ces modifieurs. Pour les premiers, nous nous contenterons de reproduire ici l'exemple de Charolles, assorti du commentaire de l'auteur, ainsi que notre traduction d'un exemple cité par Adamec (où le démonstratif apparaît avec la même valeur en tchèque) :

68) Passe-moi ***ce gâteau*** ! (Charolles, 2002 : 121-122)

Si plusieurs gâteaux trônent au milieu de la table, le démonstratif, accompagné d'un geste de monstration sélectif, remplit sa mission sans effet particulier. S'il n'y a, par contre, notoirement qu'un seul gâteau dans l'environnement partagé, l'usage du démonstratif ne paraît motivé que s'il s'accompagne d'une **nuance d'impatience**.

69) Le locuteur s'aperçoit qu'un groupe de personnes est en train de vandaliser sa voiture ; celle-ci n'est pas contrastée à d'autres voitures à l'intérieur de la situation d'énonciation :

« Qu'est-ce que vous faites ? Laissez ***cette voiture*** tranquille ! » (Adamec, 1983 : 159 ; notre traduction)

Rappelons également que souvent, comme l'écrit Bonnard (1972 : 1207), « [...] le démonstratif fait imaginer une qualification, favorable ou défavorable, selon

les cas. » Ensuite, d'autres auteurs (cf. Corblin, 1987 ; Charolles, 2002 ; Danon-Boileau, 1990) mettent en avant le fait que la valeur subjective procède de la présentation du référent à partir de la situation d'énonciation immédiate et des circonstances d'évaluation qui dépendent de cette situation. Pourtant – et c'est cela qui distingue ces emplois du démonstratif de ses emplois prototypiques –, la référence ne se fonde pas sur un contraste interne, le référent étant le seul représentant de sa classe dans la situation donnée.

Concernant les emplois « émotionnels » purs, rappelons qu'ils ont bénéficié d'un traitement théorique chez des auteurs comme Kleiber (1991a), Jonasson (1994 et 1998a) et Gary-Prieur (1994, 2001 et 2011). Dans la Section 3.5., nous avons expliqué pourquoi nous ne souscrivons pas à l'analyse de Gary-Prieur (2011) selon laquelle CE serait ici dépourvu de référence⁸¹. S'agissant de la combinaison *ce* + *nom propre*, Kleiber (1991a) distingue entre deux effets pragmatiques que le locuteur « cherche à produire » en employant le démonstratif : l'effet « de familiarité » / « affectivité » et l'effet « de distanciation ». Ces effets seraient autant d'attitudes du locuteur à l'égard du référent. C'est également Kleiber qui, reprenant à son compte l'analyse de Bonnard (1972), commente la présence au sein de la combinaison susmentionnée d'un terme évaluatif :

[l]'énonciation de *ce cher Pellegrin* ou de *cette pauvre Josette* s'appuie sur une situation antérieure qui justifie la qualification opérée par *cher* et par *pauvre*. Il reste que l'interlocuteur doit aussi avoir accès à une telle situation et, puisqu'il s'agit d'une identification démonstrative, qu'il y ait dans la situation d'énonciation des éléments qui y conduisent. (Kleiber, 1991a : 94)

Nous sommes d'accord avec cette analyse, de même que nous trouvons juste le constat de départ suivant :

À l'hypothèse qu'il ne veut pas utiliser le nom propre non modifié, alors qu'il en a la capacité, ne répond qu'un cas de figure : le locuteur entend marquer quelque chose de plus que la simple énonciation du nom propre. L'emploi du nom propre modifié produit alors un effet de sens particulier, une implicature gricéenne [...]. (Kleiber, 1991a : 85)

Les deux parties de l'analyse, à savoir 1) la nécessité pour l'interlocuteur de pouvoir accéder à la situation qui est à l'origine de l'effet véhiculé par le démonstratif (effet explicité, le cas échéant, par un terme évaluatif) et 2) le fait que l'effet ainsi produit par l'usage « détourné » du démonstratif se laisse expliquer par une implicature pragmatique au sens gricéen du terme, peuvent, selon nous, être étendues à tous les emplois du déterminant démonstratif avec un nom « sémantiquement » défini, y compris un nom en emploi générique :

70) Ils sont fous, *ces Romains* ! (emprunté à Kleiber, 1991a : 95)

⁸¹ Précisons que ce constat n'est explicitement fait par l'auteure qu'au sujet des emplois de CE introduisant un nom propre.

Dans ce fameux énoncé d'*Astérix* bien commenté dans la littérature (Kleiber, 1991a et Gary-Prieur, 2001), l'évaluation au moyen du déterminant démonstratif est conditionnée par la présence, dans la situation d'énonciation, de représentants de la classe des Romains qui viennent d'accomplir un acte jugé fou par le locuteur⁸². Il s'agit ainsi, comme le fait remarquer Gary-Prieur (2001), d'une référence générique construite par le discours. Ensuite, la même auteure relève un autre fait crucial, à savoir qu'en français, seuls les noms dotés d'une extension réduite peuvent se prêter à ce type d'emploi :

[...] si la classe des Romains peut être visée dans sa totalité [...] c'est parce qu'il s'agit d'une sous-catégorie de la classe 'homme', qui s'oppose en l'occurrence à celle des 'Gaulois', dont fait partie le locuteur. Une contrainte pèse donc sur le choix du nom pour que le GN démonstratif permette de viser la classe correspondante. (Gary-Prieur, 2001 : 226)

Cette contrainte relative au choix du nom oppose clairement le français au tchèque, où, comme nous l'avons vu dans la Section 3.5.3, tous les noms peuvent être employés, sans égard pour la largeur de leur extension (tel les noms *lidé* (*gens*) et *peníze* (*argent*)).

Un autre phénomène important qui oppose les deux langues est qu'en français, les emplois « émotionnels » ne peuvent, en l'occurrence, apparaître qu'avec une valeur de subjectivité saillante, valeur qui se prête bien, de manière générale, à une analyse en termes de familiarité/affectivité ou de prise de distance. Il n'en va pas ainsi pour le tchèque, où la présence de TEN peut tout simplement indiquer que le référent a déjà fait l'objet d'une mention antérieure (la « fonction de rappel » d'Adamec, 1983). Cette contrainte supplémentaire concernant le français semble expliquer pourquoi le démonstratif à valeur « émotionnelle » tend davantage à y apparaître dans un énoncé visiblement modalisé, que ce soit sur le plan de la modalité d'énonciation – la nature exclamative de la phrase (cf. l'exemple 70) – ou de la modalité d'énoncé – la présence de termes subjectifs (Kerbrat-Orecchioni, 2014 [1999]), tels les adjectifs évaluatifs (*cher*, *pauvre*) et les noms de qualité (*sacripant*, *salaud* ; Milner, 1978). Il convient ici de redonner la parole à Gary-Prieur, écrivant au sujet de l'exemple 70 :

Il en découle [...] que la généralité énoncée l'est sous la seule responsabilité du locuteur : c'est lui qui assume l'application du prédicat 'être fou' à la classe 'Romains'. On a donc affaire à une généralité subjective. C'est pourquoi n'importe quel prédicat ne convient pas à ce type de généralisation. On comparera par exemple :

Les Romains ont étendu leur empire sur la totalité de la Gaule.

? **Ces Romains** ont étendu leur empire sur la totalité de la Gaule. (2001 : 224)

⁸² « La référence à la classe générique peut être démonstrative, parce qu'elle passe par l'intermédiaire d'éléments saillants dans la situation, à savoir les occurrences de Romains, qui sont en relation d'appartenance avec le référent visé, la classe des Romains. Il y a ainsi un processus de généralisation "déictique" : à partir de quelques Romains qui ont un comportement insensé, on décrète que la classe entière des Romains est folle. »

Et l’auteure de poursuivre ensuite que le second énoncé accueillera beaucoup plus facilement le démonstratif une fois placé sous le signe d’une modalité explicite, exprimée, par exemple, par le modalisateur *tout de même* :

Ces Romains, tout de même, ils ont étendu leur empire sur la totalité de la Gaule. (2001 : 225)

Les deux contraintes prises ensemble font également en sorte que les emplois « émotionnels » purs sont beaucoup plus rares en français qu’en tchèque, la différence étant hors de toute proportion. Cette affirmation est largement corroborée par nos données de corpus : tandis que dans le corpus ORTOFON v1, nous avons attribué cette étiquette à 331 occurrences sur 1 000⁸³, dans le corpus constitué à partir du CEFC, ce chiffre est seulement de 11⁸⁴. Voici 7 de ces occurrences (pour une bonne analyse de l’emploi, la citation d’un contexte plus large s’avère indispensable) :

71) L1 : ouais mais bon pour le lire c'est vrai que c'est dur avec euh maîtriser mille cinq cents

L2 : plus même

L1 : plus de m- plus de mille cinq cents ouais c'est mille cinq cents kanjis pour lire un journal

L2 : minimum dans les mille euh plus je crois plus

L1 : hm comment ils font *ces japonais ils sont fous* CEFC

72) L1 : c'est un département français donc tu es en France quand tu es à la Réunion même si tu es pas sur le même continent sauf que pour nous ben c'est difficile quoi on est un peu douze mille kilomètres je crois voilà quoi on n'est pas en France et donc euh j'ai quand je suis venue ici je fais oui oui c'est la première fois que enfin que je viens en France il fait ah bon tu étais où avant à la Réunion oui mais la Réunion c'est français oui donc tu étais en France je fais oui enfin métropole

L2 : si vous voulez jouer sur les mots

L1 : ils sont pointilleux ces français hein CEFC

73) L1 : je me permettrais pas de lui demander mais euh je me demande si euh s'il m'appelle parce que je parce qu'il a envie de m'entendre ou de savoir comment je vais ou si il m'a il m'appelle juste pour apprendre des nouvelles juste pour savoir vraiment euh si ça va tu vois mais bon après il me disait que il y avait pas forcément tout qui allait bien je sais pas alors on verra mais euh parce qu'il me disait ah si tu sais que tu peux tu peux m'appeler si tu veux et tout ben oui je sais que je peux l'appeler

L2 : ah ces mecs CEFC

74) L1 : mais par contre demain si on y va j'espère que il y aura pas NNAAMMEE euh au deuxième

L2 : ah ouais tu as vraiment pas envie de le voir *ce NNAAMMEE NNAAMMEE* CEFC

⁸³ Dans la suite du travail (la Section 5.1.4 du Chapitre 5 et le Chapitre 7), nous verrons que parfois, l’attribution de cette valeur n’allait pas de soi : il faudrait donc prendre ce chiffre avec une certaine circonspection. Il n’empêche qu’il reste très élevé.

⁸⁴ Dans le sous-corpus composé des pièces de V. Havel, ces occurrences sont au nombre de 10 en français. Il s’agit de traductions d’emplois « émotionnels » de TEN introduisant dans le discours un nom propre de personne, présents dans l’original tchèque. Mais même ici, la proportion est très déséquilibrée, car ces emplois « émotionnels » français ne rendent que 10 des 50 emplois « émotionnels » du texte tchèque. Une analyse plus minutieuse du corpus des pièces fera l’objet des Chapitres 5, 6 et 7.

75) d'ailleurs la dernière fois il passe devant moi il me fait il me dit même pas bonjour alors que tous les profs me disent bonjour [...] lui il passe pff hm sympa la politesse non moi je lui dis bonjour même si je l'aime pas tu vois hm c'est je l'ai jamais porté dans mon cœur *ce cher monsieur NNAAMMEE NNAAMMEE* CEFC

76) L1 : on a regardé *ce Stéphane* euh
 L2 : /suites de syllabes/ Lamb- /amorce/ euh Yann
 L1 : - /amorce/ Yann Lambiel
 L2 : il était tellement drôle
 L1 : mais moi j'ai trouvé que c'était très bien CEFC

77) bon alors qu'est-ce qu'on fait avec *ce NNAAMMEE* quatre cinq ans c'est normal d'après les renseignements pris CEFC

Dans les trois premières (71, 72 et 73), CE introduit un nom employé de manière générique. L'énoncé de l'exemple 71 est fort proche de la phrase d'*Astérix* (70), contenant le même prédicat évaluatif (*être fou*), mais à deux différences près : 1) en 70, l'évaluation s'appuie sur des faits présents dans la situation, contrairement à 71, où elle s'appuie sur l'information qui précède dans le discours (l'obligation pour les Japonais de maîtriser un grand nombre de kanjis pour pouvoir lire un journal) ; 2) 71 ne relève pas d'une généralité discursive, car l'évaluation ne se base pas sur le comportement d'un ou de plusieurs individus appartenant à la classe des Japonais, cette classe étant, dès le début, saisie dans son ensemble. Le même constat peut être fait à propos de 72, énoncé à forte charge ironique (remarquons à cet égard la présence du modalisateur *hein*). Pour 73, en revanche, nous pouvons parler de généralité discursive, car l'énoncé exclamatif présente une évaluation de l'ensemble de la gente masculine à partir du comportement d'un individu masculin particulier, dont le comportement a fait l'objet du discours antérieur⁸⁵. Le reste des emplois contiennent un nom propre. En 75, soulignons la présence de l'adjectif évaluatif *cher*, chargé, au vu de ce qui précède, de fortes connotations ironiques. Un ton ironique accompagne également l'énoncé de 77, où un père réfère à la mauvaise conduite de son fils, mentionnée auparavant. À la différence de 75, l'ironie est ici portée par la seule voix, ne trouvant de support ni dans le lexique ni dans la modalité phrastique. Un certain ton ironique est également présent en 74. En 76, en revanche, la phrase est prononcée sur un ton d'appréciation. Ensuite, nous constatons que si la modalité tend fortement à se manifester sur les plans énonciatif (phrastique) et lexical, des exceptions semblent exister : les énoncés 71, 72 et 73 sont à la modalité phrastique exclamative ; 71 et 72 contiennent, en plus de cela, un prédicat évaluatif et aussi bien 72, 73 et 74 contiennent une interjection (*hein* et *ah*). L'énoncé de 75 contient un adjectif évaluatif (*cher*). Le cas de 76 est un peu particulier dans la mesure où les prédicats

⁸⁵ Cet emploi semble également donner raison à la contrainte relative à l'extension du nom formulée par Gary-Prieur (2001) : le nom *mecs* peut figurer dans un emploi « émotionnel », car cette classe est (implicitement) comparée à celle des femmes, de même qu'en 70, la classe des Romains et comparée à celle des Gaulois.

évaluatifs (*il était tellement drôle, c'était très bien*) n'apparaissent pas dans le même énoncé que l'occurrence du démonstratif, mais dans les deux énoncés qui suivent juste après. En revanche, 77 ne contient rien de tout cela, mais fait tout de même apparaître un ton ironique⁸⁶. Ce dernier se manifeste exclusivement au niveau prosodique, et ne serait donc pas présent à l'écrit même si la transcription des conversations se conformait aux règles orthographiques communes⁸⁷.

Pour résumer au sujet des 11 occurrences et des marques de modalisation qui les accompagnent : 3 occurrences s'accompagnent d'un prédicat évaluatif, 7 s'accompagnent d'une marque lexicale évaluative (adjectif, particule ou interjection), 5 apparaissent dans le cadre de la modalité phrastique exclamative (2, dont l'exemple 73, sont des phrases atypiques sans prédicat exprimé) et l'écrasante majorité d'entre elles (10) font également apparaître un ton expressif, le plus souvent ironique. Une seule ne fait que cela, sans contenir de marques de modalisation au niveau phrastique ou lexical (l'exemple 77). Ajoutons qu'aucune de ces 11 occurrences ne contient une particule localisatrice.

Tous les noms sont des noms « sémantiquement » définis. Dix d'entre eux font apparaître des « concepts individuels » ; un seul, un « concept fonctionnel ». Dans la théorie de Löbner (2011), les « noms fonctionnels » sont des noms à sémantique relationnelle rendus définis par la seule présence d'un argument, le plus souvent de nature possessive⁸⁸ :

Functional nouns are unary function terms [...]. Their meanings are functional concepts, involving one argument: the possessor. The value for a given argument, in a given context of utterance, constitutes the uniquely determined referent of the noun. (Löbner 2011: 282).

Voici l'exemple en question :

78) c'était quoi *ce nom de cette prof* là

CEFC

Le substantif *nom* est saturé par la présence de *cette prof*, argument qui enclenche à lui seul son unicité. Celle-ci procède ainsi du seul contenu sémantique de l'ensemble du SN, raison pour laquelle c'est l'article défini qui est censé apparaître et pas le démonstratif (qui, lui, marque toujours une définitude contextuelle, et donc « pragmatique »). Contrairement au tchèque, ces emplois du démonstratif semblent très rares en français, au point que nous pouvons nous demander si nous n'avons pas plutôt affaire à un emploi non-fonctionnel de *nom*. Mais si tel était le cas, il faudrait qu'une pause sépare l'énonciation du SN *ce nom* et celle de *cette prof*, ce qui ne se produit pas. Nous serions ainsi en droit de conclure qu'il s'agit d'un emploi du démonstratif à valeur « émotionnelle ».

⁸⁶ À moins de ne considérer *bon alors* aussi comme une marque d'expressivité.

⁸⁷ Ce qui n'est pas le cas (la modalité phrastique exclamative n'est pas marquée par le point d'exclamation, etc.).

⁸⁸ Voir la Section 6.1.3.4 pour un meilleur éclaircissement de ces concepts.

Cet emploi serait, de surcroît, très proche des emplois « de rappel » que nous avons vu pour le tchèque, car l'ensemble de l'énoncé s'inscrit dans un registre mémoriel très flagrant⁸⁹.

Quant à la nature du contexte justifiant la réévaluation faite par le démonstratif, et auquel l'interlocuteur doit avoir accès (Kleiber, 1991a : 94), nous constatons que celle-ci est discursive pour la majorité des occurrences (9) et mémorielle pour 2 occurrences⁹⁰. Ce n'est donc jamais la situation d'énonciation elle-même qui permet cette réévaluation mise en place par le démonstratif. Cette réalité est sans doute due à deux facteurs : le nombre assez réduit d'occurrences « émotionnelles » et le profil du corpus lui-même, où tous les emplois renvoyant, de quelque façon que ce soit, au cadre de l'énonciation sont, de manière générale, nettement moins représentés que les emplois renvoyant aux cadres discursif et mémoriel (voir à cet égard la Section 5.2.2.1). N'empêche que ce scénario peut se présenter en français, comme nous le voyons dans l'exemple 70, de même qu'il peut se présenter en tchèque :

79) Co	to	prosím vás	děl-á-te?	Okamžitě	nech-te
que.ACC	MOD		faire-PRS-2PL	immédiatement	laisser-IMP.2PL
být		<i>to</i>	<i>aut-o !</i>		
être.INF	ten.N.ACC.SG	voiture-ACC.SG			

*Qu'est-ce que vous faites ? Laissez **cette voiture** tranquille !* (Adamec, 1983 : 159)

4.6 Résumé

Ce chapitre a été dédié au système du démonstratif en français. Dans le premier volet, nous avons traité de la situation en synchronie, en commençant par les propriétés morphosyntaxiques et sémantiques du paradigme. Nous avons insisté sur la séparation stricte entre les formes employées comme déterminants et celles employées comme pronoms. Les principales raisons pour lesquelles le démonstratif français en position adnominale ressortit à la classe des déterminants ont été présentées et commentées en continuité avec la Section 3.1.3 du chapitre précédent. Nous avons prêté une attention particulière à la question de la distance en rapport avec le démonstratif en français contemporain, tâchant de rendre patente, à l'aide des résultats de notre Questionnaire n°2, la tendance du système à abandonner les formes en *-ci* au profit de celles sans particule ou de celles en *-là*. À cet

⁸⁹ Ainsi, le français connaîtrait lui aussi, bien que dans une bien moindre mesure, ce type d'emploi, au sujet duquel nous étions convaincu qu'il n'existait qu'en tchèque. En même temps, non seulement l'emploi 78 est l'unique de ce type dans tout le corpus d'oral informel, mais, en plus, il s'accompagne d'un ton très expressif (émotif), ce qui est loin d'être systématiquement le cas en tchèque.

⁹⁰ Une certaine hésitation nous semble permise pour l'exemple 74 ; là, l'évaluation pourrait éventuellement s'appuyer à la fois sur la réaction de l'interlocutrice qui précède immédiatement, ainsi que sur une connaissance de plus longue date (le locuteur sait que l'interlocutrice n'aime pas la personne en question, et ce que cette dernière vient de dire ne fait que confirmer cela).

égard, nous avons montré que les formes en *-là* sont également susceptibles de perdre leur faculté de marquer la distance du référent, la présence de la particule aboutissant ainsi à un pur renforcement de la composante sémantique indexicale des formes. Pour clore ce premier volet, nous avons procédé à une comparaison des fréquences du démonstratif adnominal CE dans quatre types de textes différents, à savoir : les textes scientifiques, de presse, de fiction et les conversations. À notre surprise, nous avons constaté que parmi tous ces types de textes, ce sont nos conversations du CEFC qui accueillent le moins d'occurrences de CE, la présence de celui-ci allant croissant, semble-t-il, avec la complexité de la structure des textes (le plus d'occurrences apparaissent dans les textes scientifiques de la plateforme ScienQuest).

Le second volet s'est ouvert sur une brève synthèse de l'émergence du système en diachronie, où nous avons tenu à rappeler les moments clés de l'évolution des paradigmes du latin au français contemporain, évolution qui s'apparente à une transformation radicale. Cette transformation fut marquée, avant tout, par un appauvrissement sémantique (soit une neutralisation de la distinction entre les sphères de distance) et un enrichissement (une spécialisation) syntaxique, et elle fut enclenchée essentiellement par le changement de modèle d'accentuation. De même que dans le cas du tchèque, la première étape de cette évolution fut l'abandon d'un système orienté par rapport à la personne au profit d'un système orienté par rapport à la distance, bien que la personne du locuteur semble avoir continué à jouer un rôle important dans la distribution des formes CIST et CIL en ancien français (cf. Marchello-Nizia, 2003).

Après avoir essayé de soumettre à une étude critique la thèse voyant dans le paradigme CE de la langue moderne une forme forte d'article défini, nous avons consacré une section à la restitution des différentes phases de l'évolution du SN français, tentant de dresser des parallèles entre le latin et le tchèque contemporain, deux systèmes typologiquement proches de par leur souplesse syntaxique et leur nature non-configurationnelle (cf. Carlier et Combettes, 2015).

Le dernier volet a été entièrement consacré, à l'instar de la Section 3.3, aux types d'emploi du déterminant démonstratif. Nous avons insisté sur la frontière assez nette qui, contrairement à ce que nous avons pu observer pour le tchèque, sépare en français le démonstratif du défini dans les régimes situationnel et endophorique, tout en relevant à la fois que cette frontière serait davantage estompée au sein du régime mémoriel. Une comparaison a été tentée entre les possibilités du système de l'anaphore dans les deux langues. Nous avons eu l'occasion, en outre, d'attirer l'attention sur la différence entre les emplois mémoriels – pouvant, selon les cas, recourir autant au démonstratif qu'au défini et sollicitant une connaissance d'ordre privé – et la référence au moyen des noms « sémantiquement » définis, qui, elle, a tendance à s'appuyer sur des connaissances plus généralement partagées au sein d'une communauté linguistique. Nous avons finalement abordé le phénomène des

emplois « émotionnels », afin de montrer que a) ceux-ci sont bien moins nombreux qu'en tchèque et b) qu'ils tendent à s'accompagner aussi bien d'effets pragmatiques que d'indices sémantiques et modaux plus explicites qu'en tchèque.

La première partie de notre travail s'arrête ici. Dans la partie suivante, accueillant le Chapitre 5, nous entreprendrons une série d'analyses quantitatives de nos données, où certaines observations faites dans les quatre chapitres précédents pourront davantage être développées.

Partie II

Analyses quantitatives

5 Le démonstratif adnominal : bilan quantitatif

Dans ce chapitre, nous présenterons un bilan quantitatif contrastif concernant la présence des complexes adnominaux démonstratifs dans notre matériel empirique. Dans un premier temps, nous donnerons les fréquences générale et relative du démonstratif adnominal dans les trois sources empiriques que sont le corpus de conversations défini à partir du CEFC, le corpus ORTOFON v1 et le sous-corpus composé des pièces de V. Havel. Nous entreprendrons une comparaison de la fréquence de TEN adnominal dans ORTOFON v1 avec celles de l'article défini dans les différents types de productions françaises à notre disposition. Nous nous pencherons ensuite sur chacune des trois sources pour étudier la présence et la fréquence d'une série de phénomènes, tels que les types et sous-types d'emploi, la position du complexe nominal par rapport au verbe, la grammaticalisation présumée de TEN, les moyens de transposition des démonstratifs adnominaux des textes du sous-corpus de V. Havel dans leur traduction française, les particules *-ci* et *-là* accompagnant le déterminant démonstratif français, etc. Certains des phénomènes que nous venons de mentionner seront également étudiés ensemble, afin que soient relevées des affinités et des corrélations possibles entre eux. Rappelons que ces phénomènes sont traduits dans les catégories d'annotation définies en amont, dans le Chapitre 2. Mais, au préalable, nous nous arrêterons un instant sur les principales difficultés rencontrées au cours de l'annotation. Ces dernières ont déjà en partie été abordées dans les chapitres précédents. C'est également là que figure une esquisse du bilan quantitatif (voir notamment les Sections 3.3 et 4.5). Aussi le présent chapitre se donne-t-il pour objectif, entre autres, de compléter ce bilan.

5.1 Les difficultés de l'annotation

Parmi les principales raisons pour lesquelles l'annotation, notamment celle en type et sous-type d'emploi, s'est avérée être une entreprise particulièrement délicate figurent le non-accès au contexte situationnel tout comme l'absence d'un contexte discursif plus large, même si les métadonnées dont nous disposons (voir la Section 2.1.2) nous ont parfois aidé à y voir plus clair. Par conséquent, pour une bonne partie des occurrences, il était difficile, sinon impossible, de déterminer le type et/ou le sous-type d'emploi. Faute d'accès à la situation d'énonciation (la configuration adoptée par les locuteurs dans l'espace, leur façon d'interagir, les objets faisant partie de la situation) et à un contexte plus large, il est parfois impossible de déterminer (avec certitude) si le référent d'une expression démonstrative a déjà fait l'objet d'une mention antérieure (emploi anaphorique), s'il figure

directement dans la situation d'énonciation (emploi exophorique situationnel) ou s'il se recrute parmi les connaissances partagées entre les participants à la conversation (emploi exophorique mémoriel). Il faut cependant préciser que même dans un corpus idéal, qui garantirait un bien meilleur accès à toutes ces informations, une partie des occurrences resterait tout de même ambiguë, dans la mesure où l'ambiguïté est une propriété inhérente au langage naturel, contrairement à ce qui se passe avec des exemples forgés et artificiels. Elle est étroitement liée à l'imprécision de la définition du contexte, notion problématique s'il en est. Comme l'écrit Löbner (2011 : 287) en se référant à Stanley et Szabó (2000) :

As to the notion of context: '[n]othing short of an extremely comprehensive theory of linguistic communication could provide a general solution to the problem of context dependence', state Stanley & Szabó (2000: 220). So far, such a theory does not exist.

En même temps, nous aurons également l'occasion de voir, dans le reste de ce travail, que certains emplois semblent résister à l'analyse non pas à cause d'un accès insuffisant au contexte, mais tout simplement parce qu'ils constituent eux-mêmes la source de l'équivoque¹. Dans ce qui suit, nous essaierons de passer en revue les cas d'ambiguïté les plus flagrants. Il convient de souligner que deux cas fondamentaux se sont présentés. Ou bien l'occurrence présentait une certaine ambiguïté, mais des indices permettaient de prendre une décision en faveur d'un type/sous-type d'emploi ; dans de tels cas, nous nous sommes autorisé à trancher au profit d'un type/sous-type, même si cette décision se basait sur des critères heuristiques loin d'être absolus. Ou bien l'ambiguïté était telle qu'elle empêchait le classement de l'occurrence dans un type/sous-type concret : aussi l'occurrence a-t-elle été marquée comme ambiguë (AMB) et écartées des analyses ultérieures.

5.1.1 Endophore, exophore situationnelle ou exophore mémorielle ?

Hormis les cas où il était clairement impossible d'établir si le référent était identifié grâce à sa présence dans la situation, sa mention antérieure ou sa présence dans la mémoire commune des interlocuteurs (typiquement, cela arrivait lorsque le référent apparaissait au tout début de la conversation²), il existe des occurrences dans nos corpus où la prise d'une décision – heuristique et accompagnée d'hésitations – nous semblait tout de même légitime. Typiquement, il s'agit des cas où le référent a fait l'objet d'une mention dans le discours, mais se trouve également dans la situation ou figure dans la mémoire. Dans l'exemple suivant, la locutrice parle du film *C'est arrivé près de chez vous* avec Benoît Poelvoorde comme acteur principal. Quoique le film soit identifié par le recours à la mémoire longue, il continue, par la suite, à figurer dans le texte :

¹ Plus précisément : ils sont porteurs d'une ambiguïté que les locuteurs eux-mêmes, qui les utilisent dans une situation bien déterminée, ne sauraient lever. Nous verrons cela notamment dans les deux chapitres suivants.

² Bien que nous disposions, rappelons-le, de l'ensemble de la transcription de chaque conversation.

- 1) oh toi tu tu tu [l']as regardé euh comment [il]; s'appelle [ce film euh avec Benoît Poelvoorde son premier film]; ah merde euh c'est [C'est Arrivé Près De Chez Vous]; regarde-[le]; si tu aimes l'humour noir regarde-[le]; parce que c'est ça pendant [tout le film]; en fait il joue le rôle euh d'un serial killer et il y a une équipe de télévision de télévision qui le filme pendant son son boulot tu vois et donc lui il se confie et tout mais c'est un le mec c'est un c'est un gros raciste et il dit des trucs dégueulasses et tout et il tue les gens il les viole et tout et eux ils filment pendant tout le truc tu vois il fait des remarques mais [c']est un truc de taré tu vois et [c']est vraiment il est il est il est horrible

...

[*ce film*]; tu as vu enfin tout ce qu'il dit tu vois et en fait euh il est passé euh au festival de Cannes
CEFC

Le référent est d'abord identifié (non sans difficultés) par la locutrice indépendamment du discours (et, bien entendu, indépendamment de la situation d'énonciation) : *ce film euh avec Benoît Poelvoorde, son premier film* et *C'est Arrivé Près de Chez Vous* sont trois SN référant de façon mémorielle. Mais ce référent est aussitôt repris par une série d'expressions anaphoriques : *le, tout le film* et éventuellement aussi *ce* (qui pourrait également référer, de manière résomptive, aux actions de l'antihéros qui viennent d'être exposées). Ainsi, comment analyser la seconde occurrence de *ce film*³ ? S'agit-il toujours d'un emploi mémoriel ou bien sommes-nous passés dans le domaine de l'anaphore ? C'est la seconde hypothèse qui nous semble plus plausible, et ce pour deux raisons : 1) depuis son introduction dans le discours, le référent mémoriel a déjà fait l'objet d'une série de reprises anaphoriques et 2) la seconde occurrence de *ce film* paraît enrichie de tout le contenu descriptif également apparu dans le discours antérieur (le synopsis, le profil de l'antihéros, le type d'humour présent dans le film, etc.).

La situation est plus évidente dès lors qu'un référent mémoriel est soumis à une recatégorisation par le discours :

- 2) L1 : donc du coup ils vous ont sorti [Bienvenue Chez Les Ch'tis]; ils vous [l']ont diffusé là-bas
L2 : j'étais j'étais pas en France à ce moment-là
L1 : si il est il il a deux ans [*ce film*]; CEFC

- 3) počk-ej firm-a Klasik a to jsou . to
attendre-IMP.2SG société-NOM.SG Klasik.NOM.SG et cela.NOM être.PRS.3PL cela.NOM

je ta co doda-l-a to
être.PRS.3SG celui-F.NOM.SG REL livrer-PST-3SG.F ten.N.ACC.SG

Max-ovo céděčk-o .. to dývidýčk-o
de NP.ADJpos-N.ACC.SG disque-ACC.SG ten.N.ACC.SG DVD-ACC.SG

[tu tu tu Traviat-u]; ?
ten.F.ACC.SG Traviata-ACC.SG

...

³ La même question semble se poser, en vérité, pour le SN *tout le film*.

a že tam byl-o napsán-o .. když to volně přelož-í-m ..
 et que là-bas être.PST-3SG.N écrit-3SG.N si cela.ACC librement traduire-PRS-1SG

že takhle by mě-l-a být [ta oper-a]_i ..
 que ainsi AUX.COND.3SG avoir-PST-3SG.F être.INF ten.F.NOM.SG opéra-NOM.SG

inscenován-a
 mis en scène-3SG.F

attends, la société Klasik, ce sont... c'est celle qui a fourni son disque... son DVD de [Traviata]_i à Max

...
et qu'il y avait marqué... si je traduis librement... que c'est comme ça que [l'opéra]_i devrait être mis en scène ORTOFON v1

Là, nous avons considéré automatiquement l'emploi comme anaphorique, même si la recatégorisation est très modeste (entre les deux occurrences, nous pouvons voir un rapport hyponymie-hypéronymie).

Dans l'exemple qui suit, il s'agit d'une conversation au sujet de Malifaux, un jeu d'escarmouche avec des miniatures (figurines). Le locuteur parle de figurines représentant des personnages appelés Ortega dans l'univers du jeu :

4) a to co třeba teď'ka m-á-š tam jsou takov-ý
 et cela.NOM REL MOD maintenant avoir-PRS-2SG là-bas être.PRS.3PL tel-F.NOM.PL

skupin-y .. to jsou Ortegovc-i .. to je
 groupe-NOM.PL ce être.PRS.3PL Ortega-NOM.PL ce être.PRS.3SG

[tahle rodin-a]_i ..
 tenhle.F.NOM.SG famille-NOM.SG

...
 já ne-v-í-m . pŭ-jd-eš si zahrát někam
 moi.NOM NEG-savoir-PRS-1SG FUT-aller-2SG REFL jouer.INF quelque part

a vezm-e-š si [ty Ortegovc-e]_i [těch šest lid-í]_i ..
 et prendre-FUT-2SG REFL ten.M.ACC.PL Ortega-ACC.PL ten.M.ACC.PL six gens-GEN

et ce que tu tiens maintenant, là, il y a des espèces de groupes... ce sont les Ortegas... c'est cette [famille-là]_i...

...
je sais pas, tu vas faire une partie quelque part et tu emmènes [ces Ortegas]_i, [ces six personnes]_i...
 ORTOFON v1

Résumons : dans un premier temps, le locuteur désigne la famille des Ortegas présente dans la situation. Un moment après, il reparle de cette famille. Cette dernière a donc déjà été mentionnée dans le contexte amont, mais elle reste également présente dans la situation. Dans les cas comme celui-ci, il nous semblait plus logique de donner la priorité à la deixis situationnelle, car c'est la situation qui joue le rôle principal dans l'identification du référent : c'est grâce à elle que le référent

a été identifié pour la première fois et il resterait toujours identifiable grâce à elle, quand bien même il n'aurait pas auparavant figuré dans le discours.

Cette préséance de la référence exophorique sur la référence endophorique se manifeste aussi dans les cas déjà décrits dans la Section 4.5.1, à savoir ceux qui relèvent du discours rapporté direct, et où le référent entretient à la fois une relation anaphorique avec un élément de l'entourage discursif et une valeur déictique situationnelle du point de vue du locuteur des paroles rapportées :

- 5) il me dit écoute moi ce que je te propose la pièce moi je sais ce que c'est par contre c'est une pièce qui coûte cinq cents balles mais euh moi je t'emmène on va l'acheter je te la mets puis tu vois là j'étais c'était maintenant quoi c'était maintenant ou jamais puis je fais attends ce gars-là et tout quand même je le connais pas CEFC

Certaines occurrences se sont cependant présentées à nous où il nous semblait impossible de trancher au profit d'un emploi situationnel (ou autre), car, même si le démonstratif apparaît dans un discours rapporté direct et que son référent est identifié par les protagonistes du récit, la voie de cette identification nous reste cachée :

- 6) je me rappelle d'un extrait où je sais plus dans quel film c'est peut-être dans euh La Folie Des Grandeurs où il après il cherche [un gars]_i et il a un de ses un de ses sous-fifres Louis De Funès hein il a un de ses sous-fifres il est borgne et euh mm et il lui dit ouais euh bon allez euh vous allez chercher [*ce gars*]_i et tout vous vous remu~~ remuez toute la ville CEFC

Nous partons de l'apriori que le référent de *ce gars* est identifiable par les protagonistes du film. Mais ce que nous ne pouvons pas savoir, c'est le mécanisme exact qui permet cette identification : est-ce qu'il vient d'être question de lui dans le dialogue entre les gendarmes ? Ou bien ceux-ci sont-ils en train de regarder sa photo ? C'est à cause de cette incertitude que la référence de *ce gars* a été classé comme ambigu (AMB)⁴.

Pour conclure, nous pouvons dire que globalement, dans ce type d'emplois plus ou moins ambigus eu égard au régime d'indexicalité en jeu, l'exophore situationnelle avait, pour les raisons indiquées ci-dessus, la priorité pour nous sur l'endophore, laquelle avait la priorité sur l'exophore mémorielle.

5.1.2 Deixis textuelle, deixis situationnelle ou emploi indéfini

Nous aborderons à présent quelques problèmes que nous avons rencontrés lors de l'annotation d'une série d'occurrences où nous avons hésité entre la deixis textuelle, la deixis situationnelle et un emploi indéfini du démonstratif. Voici d'abord les deux exemples français suivants :

⁴ Rappelons que pour le récepteur du message, l'identification du référent opère par le contexte.

- 7) enfin Dieu sait la conclusion que qu'elle [= la personne qui enregistre la conversation en cours] va qu'elle va faire de *cet enregistrement* CEFC
- 8) mais en fait il manquait quelque chose euh pendant *cet* euh *enregistrement* c'était de la bière CEFC

Le nom *enregistrement* tout comme d'autres noms plus ou moins semblables (*conversation*) pose problème, car ces termes font hésiter l'annotateur entre plusieurs interprétations possibles. Aussi estimons-nous que les emplois 7 et 8 sont différents. En 7, il est possible d'hésiter, d'après nous, entre la deixis textuelle et l'anaphore résomptive : s'agit-il de résumer les propos énoncés jusque-là pour instaurer un nouveau référent comme cela se produit dans l'anaphore résomptive ou bien le SN démonstratif renvoie-t-il à l'ensemble du discours dans sa matérialité (c'est-à-dire non pas en tant que somme de contenus référentiels mais en tant qu'acte de communication) ? C'est cette dernière lecture que nous tenons pour plus juste : celui qui tentera de déchiffrer l'enregistrement ne s'intéressera pas uniquement aux propos communiqués, mais aussi à la façon dont ils ont été communiqués, au matériel linguistique qui les véhicule, etc. En 8, en revanche, la lecture en anaphore résomptive est totalement écartée, et l'interprétation se jouerait entre la deixis textuelle et la deixis situationnelle. Ici, nous penchons vers la seconde : ce n'est pas tant au discours dans sa matérialité que renvoie le locuteur (et encore moins à un contenu référentiel synthétisé), mais au cadre situationnel dans lequel l'enregistrement a eu lieu. Regardons à présent les exemples suivants :

- 9) Le locuteur raconte ses péripéties avec la CAF :

deux jours après je reçois un courrier pour me dire voilà on vous a versé *cette somme-là* mais en fait vous n'y avez pas droit CEFC

- 10) quand je travaillais à l'hôtellerie tu avais les clients jamais contents il y a pas assez de serviettes c'est pas du bon savon faudrait et pourquoi il y a pas *cette chaîne là* CEFC

- 11) La locutrice parle de sa méfiance vis-à-vis de la classe politique et de son népotisme tout en reproduisant le langage supposé des représentants de cette classe :

když vy	ustoup-í-te	nám	ted'ka .	tak my	potom
si vous.NOM	céder-FUT-2PL	nous.DAT	maintenant	alors nous.NOM	après
ustoup-í-me	vám	tam	v	tom	vaš-em
céder-FUT-1PL	vous.DAT	là-bas	dans	ten.M.LOC.SG	vôtre-M.LOC.SG
					point-LOC.SG

si vous nous faites maintenant une concession, alors nous aussi, on va vous céder sur votre point à vous... ORTOFON v1

Ce que partagent ces trois emplois, c'est le fait qu'ils semblent faire partie d'une séquence en discours rapporté direct. Mais à y regarder de plus près, nous découvrons facilement que dans les trois cas, le complexe nominal peut difficilement figurer tel quel dans un discours rapporté. En réalité,

il s'agit d'emplois indéfinis du démonstratif, dont le locuteur se sert, dans le cas présent, pour combler un vide informationnel (« je ne sais plus quelle était la somme en question ») ou bien parce que l'information exacte est sans pertinence réelle pour le discours (« une somme, peu importe laquelle »), éventuellement aussi parce qu'il ne souhaite pas communiquer cette information.

5.1.3 Les différents types d'anaphore

Ici, nous traiterons de plusieurs difficultés rencontrées dans l'analyse des occurrences anaphoriques⁵. Les bases sur lesquelles repose cette analyse ont été introduites dans la Section 4.2.3. Néanmoins, le choix s'est parfois avéré délicat, notamment dans les données françaises, entre l'anaphore coréférentielle (fidèle ou infidèle), indirecte, résomptive et temporelle. Typiquement, des hésitations étaient présentes au sujet de noms à sémantique temporelle :

- 12) pendant les vacances j'ai strictement rien fait en fin de compte bah j'ai fait mon déménagement là de chez mon père et euh du coup après tu as eu euh sa copine qui est descendue alors euh je devais aller chez ma mère à la base pendant *ce temps là* mais vu que elle allait être toute seule à la maison
CEFC

- 13) L1 : pendant [deux semaines]_i tu vas faire ce que tu veux là tu es libre
L2 : ben ouais moi je suis libre et bon ma mère elle travaille et moi pendant [*ce temps*]_i ben j'essaye de de m'occuper
CEFC

Si nous comparons 12 et 13, nous constatons que *ce temps-(là)* reçoit tantôt une référence anaphorique résomptive, car il semble résumer le contenu du passage *après tu as eu euh sa copine qui est descendue* en instituant ainsi un nouveau référent, tantôt il reprend directement le référent de *deux semaines*, et il semble donc constituer une anaphore coréférentielle infidèle. Cette analyse pourrait être contestée dans la mesure où, faisant partie d'un complément circonstanciel de temps, le SN *deux semaines* paraît doté d'une référence très faible, sinon inexistante. Ce cas de figure est mentionné par exemple par Himmelmann (1997 : 118), par Laury (1997 : 25) et par Zíková (2017 : 167). Pour Laury, il s'agirait d'emplois non référentiels, car ne faisant que renseigner sur les circonstances de l'action. Zíková trouve cette analyse controversée, mais ne la rejette pas explicitement. C'est Himmelmann qui propose, d'après nous, la meilleure façon de traiter ces expressions en écrivant :

Der diskursbasierte Referenzbegriff erlaubt graduelle Abstufungen in der Referentialität eines Ausdrucks, die damit zusammenhängen, wie klar ein gegebener Ausdruck der Diskursfunktion *tracking* dient. Als operationalisierbares Kriterium für Diskursreferenz schlage ich vor, die Möglichkeit demonstrativer Bezugnahme zu verwenden. (1997 : 118)

⁵ Nous mettons à part l'anaphore associative, sur laquelle nous dirons un mot plus loin.

Ainsi, dès lors que l'expression peut donner suite à une reprise anaphorique directe (au moyen d'un démonstratif), nous sommes face à un emploi référentiel, quitte à considérer que cette nature référentielle n'est établie qu'a posteriori, au moyen, justement, de la reprise⁶.

À ces deux occurrences de l'expression *ce temps(-là)*, nous pouvons en ajouter une troisième, où l'emploi ne relèverait ni de l'anaphore résomptive ni de l'anaphore infidèle, mais de l'anaphore temporelle, car il ne reprend, en l'occurrence, aucun contenu référentiel du discours précédent :

- 14) voilà si vous l'avez retrouvé hein et puis après je vous dirais les montants exacts je vais les chercher pendant **ce temps** et on fera le le on fera la mise à jour avec les nouveaux montants CEFC

Ces trois exemples nous montrent bien qu'un seul et même nom à sémantique temporelle peut se prêter à plusieurs types d'anaphore différents. Il en va de façon similaire de *ce moment-là*, qui apparaît le plus souvent, dans notre corpus d'oral informel français, avec une valeur d'anaphore temporelle (l'exemple 15), mais qui peut également apparaître avec une valeur d'anaphore coréférentielle infidèle comme en 16, où il reprend la référence du circonstant temporel *tout à l'heure*, voire avec une valeur anaphorique résomptive, servant de connecteur argumentatif (plus que temporel ; l'exemple 17) :

- 15) ben ça me rappelle aussi euh au lycée à Mendès
...
il y avait des jumeaux à **ce moment-là** aussi CEFC

- 16) L1 : c'est parti O K bon on parlait de quoi [tout à l'heure]; on peut on peut reprendre
L2 : à [**ce moment-là**]; CEFC

- 17) et puis si tu dis ça à partir de **ce moment-là** tu enfin les gens se
...
les gens font encore plus attention à leurs expressions CEFC

Dans le cas d'une seule expression démonstrative française contenant un nom à sémantique temporelle, nous avons conclu à la présence d'une anaphore indirecte :

- 18) j'ai vu sortie [le six juillet]; donc euh on verra bien donc ce qu'on pourrait faire parce que normalement la Japan Expo c'est dans ces dans [**ces dates-là**];+? CEFC

⁶ De la même façon qu'un « référent » non-spécifique peut être établi comme proprement référentiel sur le plan discursif via une reprise anaphorique, comme nous l'avons vu dans la Section 1.2.1, entre autres grâce à un exemple de Prince (1981).

Le SN *ces dates-là* est partiellement coréférentiel avec *le six juillet* (*ces dates-là* = le six juillet + d'autres jours proches de cette date). En effet, cet emploi ne s'éloigne guère de l'emploi tchèque suivant :

- 19) [von]_i je takov-*ej* blond'ák víš ? voni
 lui.NOM être.PRS.3SG tel-M.NOM.SG blond.NOM.SG MOD eux.NOM
- [tydlety lid-i]_i +? . [tydle blond'at-ý lid-i]_i +? .. jsou
 tenhleten.M.NOM.PL gens-NOM tenhle.M.NOM.PL blond-M.NOM.PL gens-NOM être.PRS.3PL
- [lui]_i [c']_i est un espèce de blond, tu sais ? [ces personnes-là]_i [ces blonds-là]_i ..., [ils]_i sont
 ORTOFON v1

Ici, les expressions *tydlety lidi* (*ces personnes-là*) et *tydle blond'atý lidi* (*ces personnes blondes-là*) reprennent le référent introduit par ON (LUI), mais elles renvoient aussi à toute la classe des personnes blondes dont celui-ci fait partie.

Outre les noms à sémantique temporelle, c'est l'expression *ce genre de* + *N* qui représentait un défi d'annotation considérable. Comparons les occurrences suivantes :

- 20) L1 : oh j'ai lu un truc qui m'a choquée quoi dans un livre c'est un roman c'est euh Anna Karénine Tolstoï tu vois c'est russe ça se situe au dix-neuvième siècle et euh c'est euh euh un jeune un jeune homme là riche qui monte euh qui aime faire des courses à cheval
 L2 : ouais
 L1 : donc euh là il est en train de monter et de faire une course d'obstacles et il est mal retombé sur son cheval et il lui a brisé les reins
 L2 : oh
 L1 : non ça arrive ce vraiment *ce genre de choses* j- CEFC
- 21) L1 : c'est [le Dumbo]_i c'est tu te mets dans un une nacelle en forme de Dumbo et c'est ça tourne en rond comme ça oh ouais c'est cool hein
 L2 : trop cool non moi j'aime pas [*ce genre de truc*]_i CEFC
- 22) L1 : ben ça reste [Titanic]_i quoi c'est
 L2 : je sais pas moi j'aime ça me plaît pas trop [*ce genre de films*]_i +? si je peux éviter j'évite CEFC
- 23) L1 : à mon frère je lui ai pris un couteau corse parce qu'ils font [des couteaux en bois d'olivier]_i
 ...
 L1 : bien affutés par exemple mais bon lui il coupe pas le saucisson avec il s'en il s'en sert co-~ euh comme objet d'art on va dire en expo
 L2 : oui d'accord oui il y en a qui sont des grands amateurs de [*ce genre d'objet*]_i +? CEFC

Même si la différence entre ces emplois de l'expression est parfois ténue et difficile à décrire en des termes précis, nous trouvons qu'en 20, *ce genre de choses* réfère de manière résomptive en ce qu'il reprend le contenu de l'ensemble de l'histoire du jeune aristocrate qui a cassé les reins de son

cheval⁷. En 21, nous trouvons plus juste de voir en *ce genre de truc* une anaphore coréférentielle infidèle, reprenant le référent du SN *le Dumbo*. Force est cependant de constater que cette décision a été accompagnée d'hésitations, car il serait également possible de voir en cet emploi de l'expression un résumé de la description du principe du manège – *tu te mets dans un une nacelle en forme de Dumbo et c'est ça tourne en rond comme ça* –, et donc une anaphore résomptive. En 22 et en 23, les SN *ce genre de films* et *ce genre d'objet* embrasseraient une référence anaphorique indirecte, car, en plus de référer au film *Titanic* et aux couteaux en bois d'olivier, qui viennent d'être mentionnés, ils renvoient à l'ensemble des films du même genre que *Titanic* (22) et à l'ensemble des objets d'art décoratif, dont font partie les couteaux en bois d'olivier (23). C'est pour cela que nous trouvons qu'un -s manque dans la transcription de 23, où *objet* devrait figurer au pluriel.

En tchèque, 20 trouve une contrepartie en 24, déjà cité dans les Sections 2.2.3 et 3.3.2 :

24) mně to přijd-e úplně divn-ý no nebo to
moi.DAT cela.NOM sembler-PRS.3SG complètement bizarre-N.NOM.SG MOD ou cela.NOM

jak zpíva-l-a Basikov-á v tý latin-ě víš
REL chanter-PST-3SG.F NP-F.NOM.SG en ten.F.LOC.SG latin-LOC.SG MOD

[...] .. *tenhle typ modern-ích pís-n-í*
tenhle.M.NOM.SG genre.NOM.SG moderne-F.GEN.PL chanson-GEN.PL

moi, je trouve ça complètement bizarre ; ou le truc que Basiková a chanté en latin, tu vois... ce genre de chansons modernes
ORTOFON v1

5.1.4 Emploi anaphorique/mémoriel ou emploi « émotionnel » ?

Dans notre échantillon du corpus ORTOFON v1, une question se posait de manière récurrente, à savoir celle de la distinction, pour une partie non négligeable des occurrences, entre un emploi anaphorique ou mémoriel d'une part et un emploi « émotionnel » d'autre part. Aussi allons-nous expliquer à présent comment nous avons procédé dans ces cas équivoques :

25) oni děl-a-jí strašn-ou spoust-u meeting-ů
eux.NOM faire-PRS-3PL horrible-F.ACC.SG masse-ACC.SG meeting-GEN.PL

v [*centr-u měst-a*]_i ..
dans centre-LOC.SG ville-GEN.SG

...
tak . právě hlási-l-i že [*to centr-um*]_i . bud-ou
alors justement annoncer-PST-3PL.M que ten.N.NOM.SG centre-NOM.SG AUX.FUT-3PL

⁷ Certes, il renvoie aussi à l'ensemble des événements du même genre que celui qui vient d'être raconté. Cependant, c'est l'opération mentale d'institution d'un nouveau référent par synthèse qui nous paraît ici la plus importante.

mít nějak-ý ty meeting-y .. a jakmile se
avoir.INF quelque-M.ACC.PL ten.M.ACC.PL meeting-ACC.PL et dès que REFL

zacp-e [centr-um měst-a]_i
boucher-FUT.3SG centre-NOM.SG ville-GEN.SG

ils organisent énormément de meetings [au centre-ville]_i... ; du coup, ils disaient justement que [le centre-ville]_i, qu'il y aura des meetings... et dès que [le centre-ville]_i sera bouché

ORTOFON v1

26) Il s'agit de la description d'un bal :

tak to byl-o .. až do úplně až k [pódi-u]_i .. potom
alors cela.NOM être.PST-3SG.N jusque dans complètement jusque à podium-DAT.SG après

byl-y jakože .. stol-y až k [pódi-u]_i no
être.PST-3PL.M MOD table-NOM.PL jusque à podium-DAT.SG MOD

...
to byl-o k tomu . až k [tomu pódi-u]_i a
ce être.PST-3SG.N à ten.N.DAT.SG jusque à ten.N.DAT.SG podium-DAT.SG et

[pódi-um]_i byl-o voln-ý [...] jakože .. stol-y
podium.NOM.SG être.PST-3SG.N libre-N.NOM.SG MOD table-NOM.PL

až k [pódi-u]_i
jusque à podium-DAT.SG

du coup, c'était... jusque dans jusqu'[au podium]_i ... après, il y avait genre... des tables jusqu'[au podium]_i

c'était jusqu'au jusqu'[au podium]_i et [le podium]_i était libre [...] genre... des tables jusqu'[au podium]_i

ORTOFON v1

27) a bud-ete řešit nějak do toho sklep-a . jak to
et AUX.FUT-2PL régler.INF d'une façon dans ten.M.GEN.SG cave-GEN.SG REL cela.NOM

tam . vlh-l-o ?
là-bas devenir humide-PST-3SG.N

et est-ce que vous allez [essayer de] régler [le truc] dans votre cave, là où ça devenait humide ?

ORTOFON v1

Si nous regardons 25 et 26, nous constatons qu'un référent défini y apparaît d'abord sans TEN avant d'être repris par TEN. Suivant notre premier instinct, nous pourrions en déduire que cette reprise est nécessairement anaphorique (de même que le déterminant démonstratif français peut reprendre des référents déjà établis dans le discours afin de les réactiver ou pour les contraster à d'autres référents ; voir la Section 2.5.2). Or ce n'est pas ici le cas, car les référents sont de nature « sémantiquement » définie (le centre de la ville où la locutrice réside et le podium de la salle des fêtes où se tient le bal). D'où il s'ensuit que le référent est identifiable sans une introduction préalable dans le discours. Mais comme il s'agit d'un défini « sémantique », la présence de TEN avec les mentions ultérieures ne peut non plus être justifiée par un emploi mémoriel. D'ailleurs, si cela en était

vraiment un, comment expliquer que TEN n'introduit pas le nom dès la première mention ? En réalité, le référent du complexe nominal *(to) centrum (města)* est un concept fonctionnel (Löbner, 2011), et le complexe nominal *(to) pódium* semble représenter un cas d'anaphore associative de type situationnel (Czardybon, 2017), car son référent semble identifié via le lien conventionnel existant entre lui et l'ensemble de la situation, à savoir la tenue d'un bal dans une salle des fêtes. La seule explication permettant ainsi de rendre compte de la présence de TEN est qu'il s'agit d'un emploi « émotionnel » du type « de rappel » (Adamec, 1983), signalant (insistant sur le fait) qu'il a déjà été question du référent auparavant. Cela explique également pourquoi, après cette réintroduction par TEN, le référent continue, par la suite, à apparaître sans TEN, exactement comme au début : c'est précisément parce que le locuteur a cessé d'insister sur le fait de la mention antérieure.

Avec l'exemple 27, nous cherchons juste à montrer que ce type d'emploi « émotionnel » apparaît également avec des référents « sémantiquement » définis en première mention : même si aucune mention du référent de la cave – référent au sujet duquel il est logique de conclure qu'il est « sémantiquement » défini, car unique et bien établi dans l'univers du discours de l'interlocuteur (une personne ne possédant d'habitude qu'une seule cave) – ne précède l'emploi de *ten sklep*, le contexte fait clairement la preuve qu'il s'agit d'un rappel d'une « vieille histoire de cave »⁸.

Il y a aussi des cas, dans nos données d'oral informel pour le tchèque, où nous avons hésité entre un emploi « émotionnel » et une anaphore associative :

28) tak to bud-u jako porovnávat ale je to
alors cela.ACC AUX.FUT-1SG MOD comparer.INF mais être.PRS.3SG cela.NOM

hrozně těžk-ý víš ? jako že prostě . pf já vûbec ne-m-á-m
terriblement difficile-N.NOM.SG MOD moi.NOM du tout NEG-avoir-PRS-1SG

ten *teoretick-ej** *základ ..*
ten.M.ACC.SG théorique-M.ACC.SG base.ACC.SG

du coup genre je vais comparer ça, mais c'est trop difficile, tu sais ? genre j'ai pas du tout les bases théoriques ... ORTOFON v1

29) ona ho mě-l-a taky za dv-ě st-ě osmdesát devět
elle.NOM.SG lui.ACC avoir-PST-3SG.F aussi pour deux-ACC cent-ACC.PL quatre-vingt neuf

korun .. jenom bez deset-i eur .. *tu* *doprav-u*
couronne.GEN.PL seulement sans dix-GEN euro.GEN.PL ten.F.ACC.SG transport-ACC.SG
...

⁸ C'est pour cela que *ten sklep* ne réfère pas uniquement à la cave comme telle, mais à la totalité d'une situation la concernant, à savoir, précisément, au problème que l'interlocuteur avait avec la cave et dont il a auparavant fait part au locuteur.

deset eur je včetně dan-ě **ta** **doprav-a**
 dix euro.GEN.PL être.PRS.3SG y compris taxe-GEN.SG ten.F.NOM.SG transport-NOM.SG

elle aussi, elle l'a eu pour deux cents quatre-vingt-neuf couronnes... ; sans compter les dix euros de frais de transport

...
les frais de transport, c'est dix euros, taxe comprise ORTOFON v1

30) vona je těhotn-á ne ?
 elle.NOM.SG être.PRS.3SG enceinte-F.NOM.SG MOD

...
 tak furt po mně tah-á nějak-ý rozum-y
 alors tout le temps après moi.LOC tirer-PRS.3SG quelque-ACC.PL raison-ACC.PL

...
 ať to jako ignoruj-e že **ten** **porod**
 que cela.ACC MOD ignore-PRS.3SG que ten.M.ACC.SG accouchement.ACC.SG

zvládn-e ..
 maîtriser-PRS.3SG

elle est enceinte, tu vois ?

...
du coup, elle essaie tout le temps de me tirer les vers du nez

...
*je lui dit de laisser tomber, que **l'accouchement** va bien se passer..* ORTOFON v1

En 28, la locutrice parle de son mémoire, où elle est obligée de faire un travail de lexicographe consistant à comparer plusieurs dictionnaires entre eux. Elle se plaint de ne pas posséder les bases théoriques en la matière. En 29, il est question des frais de transport associés à la livraison d'un produit. Enfin, en 30, le complexe nominal *ten porod* (*l'accouchement*) s'inscrit dans la situation de la grossesse d'une femme connue des interlocuteurs et des appréhensions que cela génère pour elle. Dans les trois cas, nous avons privilégié une analyse en anaphore associative (du type relationnel en 28 et en 29 et du type situationnel en 30 plutôt qu'une lecture « émotionnelle », laquelle ne peut, pourtant, être complètement écartée non plus. Concrètement, en 28, l'on pourrait conclure à la présence d'un emploi « émotionnel » de type « emphatique » (Adamec, 1983), car l'occurrence apparaît dans un énoncé assez modalisé. L'emploi en 29 serait à la limite analysable comme un emploi « émotionnel de rappel » (*les frais de transport, dont on a déjà parlé*). Quant à 30, l'occurrence pourrait aussi s'interpréter comme un emploi « émotionnel », soit « de rappel » soit même « emphatique » (*cet accouchement, dont elle n'arrête pas de parler*). La raison pour laquelle nous avons opté pour une analyse en anaphore associative est que cette analyse se base davantage sur des critères objectifs, ne nécessitant pas de recours à des jugements subjectifs et, par définition, plus vagues. Ces critères sont les suivants : 1) il existe la possibilité d'établir un lien entre l'occurrence et un cadre mentionné par le discours, lien qui justifie la sémantique définie de l'occurrence en première

mention et 2) contrairement à 26, où l'emploi anaphorique associatif est d'abord introduit sans TEN et seulement après par TEN, ici, il est introduit par TEN dès la première mention.

5.1.5 Autres difficultés d'annotation

Parmi les autres difficultés d'annotation rencontrées figurent les occurrences où le démonstratif apparaît seul, sans nom, que ce soit pour une raison d'aposiopèse ou d'interruption venant de l'extérieur. Dans ces situations, l'occurrence a été soit marquée comme ambiguë (AMB) soit, si le contexte antérieur fournissait un indice permettant de faire une hypothèse suffisamment plausible, un type d'emploi présumé a été établi :

- 31) L1 : c'est comme euh [les cartes aussi que tu achètes pour appeler à l'étranger]_i
L2 : [les prépayées]_i là
L1 : [les] ouais c'est euh [c']_aest pour appeler directement avec le fixe mais [ça]_i te pré~ euh [ça]_i
te prend pas sur ton abonnement fixe en fait
L2 : d'accord non [ça]_i j'utilise pas [*ces*]_i CEFC
- 32) L1 : c'est entre Alicante et Séville
L2 : ouais
L1 : c'est c'est vraiment dans le sud c'est c'est tout petit il y a personne qui connaît euh ah ça ça existe ça ben oui ça y existe je sais j'ai pris le truc le plus le moins connu du monde mais euh
L2 : ah ben ouais mais c'est bien enfin d'un côté d'avoir un truc qui n'est pas immense
L1 : moi je voulais pas m en fait si tu veux moi j
L2 : Madrid c'est pas la même chose
L1 : moi j'ai dit que je voulais pas Madrid hein
L2 : faut voir ce qu'on préfère dans *ces* CEFC

Tandis qu'en 31, le nom absent peut être facilement restitué au vu de ce qui précède (il s'agit des cartes téléphoniques prépayées dont il a été jusque-là question), en 32, le nom ne peut être restitué avec certitude, même si, compte-tenu de ce qui précède, nous pouvons présumer qu'il s'agit probablement d'un nom comme *situation*. Par conséquent, nous avons analysé l'occurrence en 31 comme un emploi anaphorique fidèle, contrairement à 32, qui a été annoté comme ambigu (AMB).

Ensuite, c'est le statut de la particule *-là* qui s'avérait parfois problématique. En général, l'analyse de ce lexème français n'est pas sans poser des difficultés (cf. Kleiber, 1995a et 1995b ; Corblin et Asic, 2016). En ce qui nous concerne, nous laissons de côté toutes les questions du type de savoir si, dans ses emplois standard, *là* est un déictique ou pas et, si oui, à quel point il est comparable aux adverbes déictiques de lieu *ici* et *là-bas*. En revanche, ce qui nous importe, c'est la différence entre les emplois où *-là* doit être analysé comme faisant partie du SN démonstratif, et donc comme étant porteur d'un sens locatif (quoique considérablement neutralisé, comme nous l'avons vu dans la Section 4.1.3) et ceux où il n'en fait pas partie, se comportant ainsi comme ce que nous croyons

juste d'appeler une particule modalisatrice. Si la question de la différence se posait parfois à nous, c'est que certaines transcriptions de la base CEFC ne respectent pas le statut de la particule :

- 33) d'abord est-ce qu'il y a quelqu'un qui est pas d'accord avec euh *ces cadres là* CEFC
- 34) c'était déjà euh *cette ambiance là* CEFC
- 35) ouais donc voilà donc en *ce moment là* puis je te dis non c'est vrai non mais les samedis j'ai remarqué hein CEFC
- 36) ça commence doucement euh parce qu'on n'a eu que quatre jours de cours avant *cette semaine là* CEFC

Alors qu'en 33 et 34, la particule fait partie du SN démonstratif, elle n'en fait pas partie en 35 et en 36. Pourtant, les transcriptions laissent croire qu'elle n'en fait partie dans aucun des quatre énoncés, car dans aucun de ces derniers, le trait d'union n'apparaît. Ainsi, comme dans bien d'autres situations, ici aussi, seul l'enregistrement permettait de trancher.

Mentionnons ici un dernier problème, déjà effleuré dans la Section 4.5.2. Ici, il s'agit de la reprise anaphorique d'un nom ayant occupé la position attribut d'une proposition apparue plus tôt dans le discours :

- 37) L1 : ben [c']_iest pas normal [c']_iest un truc
...
L2 : [ça]_i t'appartient jamais
...
L1 : en fait c'est toi qui [le]_i paies hein [*ce truc-là*]_i CEFC

Bien que le SN *un truc* ne réfère pas (il est question d'une prédication caractérisante), nous avons décidé – certes un peu arbitrairement – de compter le SN *ce truc-là* parmi les occurrences anaphoriques fidèles, partant du principe que la présence du nom *truc* dans la reprise se base, justement, sur la recatégorisation du référent opérée par la prédication qui précède.

D'autres problèmes posés par l'annotation, tels que la position du complexe démonstratif par rapport au verbe, ont été abordés dans la Section 2.2.3 du Chapitre 2.

5.2 Bilan quantitatif

Dans la présente section, nous procéderons d'abord à la présentation des résultats auxquels ont abouti les requêtes formulées pour nos trois principales sources de données. Puis nous effectuerons une comparaison de la fréquence de TEN adnominal dans ORTOFON v1 avec celles de l'article défini LE dans différents types de textes français pour évaluer l'hypothèse de la grammaticalisation de TEN sur le plan quantitatif. Ensuite, nous entreprendrons une analyse quantitative plus détaillée a) des

résultats de notre corpus de conversations créé à partir du CEFC, b) des résultats de l'échantillon du corpus de conversations d'ORTOFON v1 et, enfin, c) des résultats du sous-corpus des pièces de V. Havel.

5.2.1 Constats généraux

Les requêtes appliquées à nos deux corpus d'oral informel, qui ont fait l'objet d'une description détaillée dans les Sections 2.2.1 et 2.2.2, ont abouti dans un tout premier temps aux nombres d'occurrences suivants : 24 096 pour **ORTOFON v1** et 2 015 pour le corpus de conversations défini à partir du **CEFC**. En français, le nombre d'occurrences bruitées, qui ont été éliminées manuellement, était de 144, soit 7,15 %, de toutes les occurrences obtenues. En tchèque, ce nombre était de 377, soit à peu près 15,71 %, dans l'échantillon réduit de 2 409 occurrences, obtenu par le maintien, à partir de la liste initiale générée par la requête, d'une occurrence sur dix. Grâce à l'application du module « 1 feature – many samples » (la loi de Student), nous avons constaté que la fréquence absolue moyenne du bruit dans l'ensemble de la population ($n = 24\,096$ occurrences) serait de 3 785,08 occurrences (voir la Section 2.2.4)⁹.

Si nous enlevons à présent le bruit, nous obtenons, pour nos deux corpus de conversations, les résultats suivants : 1 871 occurrences pour le corpus du CEFC et, pour ORTOFON v1, une moyenne de 20 310,92 occurrences, soit 20 311 après son arrondissement à l'unité. Notons que pour le tchèque, le chiffre est approximatif, car compris à l'intérieur de l'intervalle de confiance calculé, soit 19 917-20 705 occurrences¹⁰.

Qu'en est-il de **la présence**, dans ORTOFON v1, **des lemmes du démonstratif adnominal autres que TEN** ? Dans deux échantillons de 1 000 occurrences constitués tous les deux via le prélèvement d'une occurrence sur dix à partir de la liste des concordances initiale, les lemmes autres que TEN comptent, respectivement, 33 et 34 occurrences. Si nous utilisons l'interface « 1 feature – many samples » (la loi de Student), où nous saisissons les deux valeurs – 33 et 34 –, nous arrivons à une moyenne de 680,42 occurrences du phénomène dans l'ensemble de la population, avec un intervalle de confiance qui s'étend entre 551,4 et 809,5 occurrences, soit entre 2,71 % et 3,99 % de l'ensemble des occurrences du démonstratif adnominal¹¹. Quelque approximatifs qu'ils soient, ces chiffres attestent une faible présence, dans ORTOFON v1, des lemmes du démonstratif autres que

⁹ C'est dans la Section 2.2.4 que nous avons expliqué que vu la taille de la liste des concordances initiale, il aurait été extrêmement laborieux et chronophage d'établir la part du bruit manuellement, raison pour laquelle nous l'avons établie sur une base de 24 échantillons de 100 occurrences, obtenus à partir de l'échantillon de 2 409 occurrences, lui-même généré par le maintien d'une occurrence sur dix à partir de la liste initiale. Nous y avons également expliqué pourquoi nous avons été obligé de nous contenter de l'écart-type assez large de 3,873.

¹⁰ Une fois les chiffres arrondis à l'unité. La valeur 20 310,92 est la moyenne des deux bornes de l'intervalle.

¹¹ L'écart-type = 0,707 ; $\alpha = 0,05$.

TEN en emploi adnominal, la fréquence absolue de ce dernier étant comprise entre l'intervalle 19 366-19 896 occurrences.

Vu les chiffres fournis *supra* ainsi que le nombre de positions qui nous est connu pour les deux sources de données (934 102 pour le CEFC et 1 236 508 pour ORTOFON v1), il nous est désormais possible de calculer **les fréquences relatives** (nombre d'occurrences relativement à un million de positions ; ipm) **du démonstratif adnominal dans les deux corpus**. Cette fréquence s'élève à 16 107,4-16 744,7 (soit 1,30-1,35 % de toutes les positions du corpus) pour ORTOFON v1¹² et à 2 003 (soit 0,21 % de toutes les positions du corpus) pour le corpus du CEFC. S'agissant des lemmes du démonstratif autres que TEN dans ORTOFON v1, leur fréquence relative est comprise entre 445,9 et 654,6 occurrences. Par conséquent, pour le lemme TEN, la fréquence relative est comprise entre 15 661,5 et 16 090,1 occurrences (soit 1,27-1,30 % de toutes les positions du corpus).

Remarque importante : tous ces chiffres doivent être approchés avec circonspection, notamment parce que 1) les annotateurs des deux corpus emploient, comme nous l'avons vu dans le Chapitre 2, des systèmes d'annotation différents pour ce qui est de la transcription de la ponctuation, des marques d'hésitation et des interjections et 2) que les requêtes elles-mêmes que nous avons formulées pour les deux corpus, si elles s'avèrent les meilleures possibles, ne parviennent tout de même pas à isoler la totalité des occurrences visées (voir la Section 2.2.1 pour les requêtes, leur analyse et la justification de leur choix). Bien que nous ayons conscience de ces différences, le clivage n'en est pas moins significatif, témoignant d'une situation radicalement disproportionnée quant à l'emploi du démonstratif adnominal dans les registres parlés informels des deux langues, la fréquence relative dans le corpus tchèque s'élevant à plus de huit fois celle du corpus français.

Dans **notre sous-corpus des pièces de V. Havel**, nous sommes arrivé aux chiffres suivants (rappelons que le relevé des occurrences a en partie été effectué manuellement ; la requête est également décrite dans la Section 2.2.2) : 241 occurrences pour la fréquence absolue¹³ – dont 63 appartenant aux lemmes autres de TEN – et 4 962 occurrences pour la fréquence relative, (soit 0,5 % de toutes les positions du corpus). Du fait de la taille très limitée du sous-corpus (48 568 positions), la fréquence relative doit cependant être considérée avec une grande prudence¹⁴. Si elle se situe au-dessus de celle du CEFC, elle est clairement au-dessous de celle du corpus ORTOFON v1. Une

¹² Cvrček *et al.* (2015 [2010] : 175), étudiant la fréquence de TEN adnominal dans un ensemble de conversations provenant du corpus de tchèque parlé informel ORAL2006 (à peu près 1 000 000 de mots), sont arrivés à un résultat très proche du nôtre (ils ne donnent pas la valeur exacte, mais la fréquence, représentée sur un diagramme, atteint à peu près le chiffre 16 000).

¹³ Ce chiffre a été obtenu après l'élimination du bruit et l'ajout manuel des occurrences qui avaient échappé à la requête.

¹⁴ S'ajoute à cela les différences dans la composition des deux corpus, les corpus de langue orale contenant des signes d'annotation que les corpus de langue écrite ne contiennent pas (voir la Section 2.1.2).

conclusion préliminaire qu’il serait possible de tirer de ce fait est que même dans un tchèque écrit contenant un taux significatif de représentation de la langue orale informelle, l’usage de TEN adnominal est loin d’atteindre les mêmes proportions que dans le tchèque oral informel authentique¹⁵.

Le Tableau 5.1 apporte une synthèse de l’ensemble des observations¹⁶ :

Source	Nombre d’occurrences (fréquence absolue) avec bruit ¹⁷	Nombre d’occurrences (fréquence absolue) sans bruit	Fréquence relative pour un million de positions	Fréquence relative en %
CEFC (Orféo ; 934 102 positions)	2 015	1 871	2 003	0,21 %
ORTOFON v1 (CNT ; 1 236 508 de positions)	24 096	19 917-20 705	16 107,4-16 744,7	1,30-1,35 %
		Dont TEN		
		19 336-19 896	15 661,5-16 090,1	1,27-1,30 %
V. Havel (InterCorp, CNT ; 48 568 positions)	310	241	4 962	0,50 %

Tableau 5.1 : Fréquence absolue du démonstratif en emploi adnominal avec et sans bruit et fréquence relative dans les conversations du CEFC (Orféo), ORTOFON v1 (CNT) et les pièces de V. Havel (InterCorp v10, CNT)

Appliquons à présent la fonction de vraisemblance à la différence de fréquences entre les trois sources : à celle entre les fréquences absolues du corpus du CEFC et d’ORTOFON v1, d’abord, puis à celle entre les fréquences absolues du corpus d’ORTOFON v1 et des pièces de V. Havel¹⁸. Pour ORTOFON v1, nous employons la fréquence absolue moyenne, calculée par l’interface « 1 feature – many samples », soit 20 311 occurrences (voir la Section 2.2.4). Pour la première différence, le score est égal à 13 292,8236 ; pour la seconde, il est égal à 527,9946. Les deux différences sont statistiquement significatives :

Corpus	Score de la fonction de vraisemblance
ORTOFON v1 (CNT) vs le CEFC (Orféo)	13 292,8236
ORTOFON v1 (CNT) vs les pièces de V. Havel (InterCorp v10, CNT)	527,9946

¹⁵ Malheureusement, nous ne pouvons pas non plus exclure un certain poids du facteur diachronique, les pièces ayant été écrites il y a environ 50 ans (voir la Section 2.1.3).

¹⁶ Pour ORTOFON v1, les valeurs inférieure et supérieure de l’intervalle de confiance sont arrondies à l’unité pour le nombre d’occurrences sans bruit (fréquence absolue), au dixième pour la fréquence relative et au centième pour les valeurs exprimées en %. Dans l’ensemble de ce chapitre, les valeurs en % citées dans les tableaux sont toutes arrondies au centième.

¹⁷ Soit le nombre d’occurrences total obtenu par la requête.

¹⁸ Dans les deux cas, les fréquences sont rapportées aux tailles des corpus exprimées en nombre de positions. La valeur critique = 3,841 (degré de liberté = 1 et le seuil de significativité (α) = 0,05).

Tableau 5.2 : Scores de la fonction de vraisemblance appliquée aux différences entre les fréquences absolues du démonstratif en emploi adnominal dans a) ORTOFON v1 (CNT) et le corpus de conversations du CEFC (Orféo) et b) ORTOFON v1 (CNT) et le corpus des pièces de V. Havel (InterCorp v10, CNT) ; valeur critique = 3,841

Rappelons que dans les Chapitres 3 et 4 (Sections 3.1.2 et 4.1.4), nous avons effectué une comparaison des fréquences du démonstratif dans différents types de textes des deux langues. Alors qu'en tchèque, c'est dans notre corpus de conversations que le démonstratif (emplois adnominaux et pronominaux confondus) atteint la plus haute fréquence par comparaison aux trois parties du corpus de référence pour la langue écrite contemporaine SYN2015 et même par comparaison à la partie du corpus PMK sur laquelle se base *Frekvenční slovník mluvené češtiny* (*Dictionnaire des fréquences du tchèque parlé* ; 2007), en français, c'est bien le contraire que nous avons pu observer : notre corpus de conversations affiche la fréquence la plus basse par comparaison aux autres types de textes, appartenant tous à la langue écrite. À cet égard, nous avons proposé ceci comme explication : alors que le tchèque emploie le plus de démonstratifs dans son registre oral informel, où la forme TEN est le plus utilisée à la fois pour son statut d'article en devenir (voir le Chapitre 6) et pour ses valeurs pragmatiques « émotionnelles » (voir le Chapitre 7), en français, le recours à CE – qui garde son statut de démonstratif – va croissant avec la complexité des textes et il est, de manière générale, plus fréquent à l'écrit qu'à l'oral, dans la mesure où ses emplois anaphoriques sont valorisés dans la cohésion des textes à caractère explicatif et argumentatif (d'où le fait que sa fréquence soit le plus élevée dans le domaine des textes scientifiques ; voir la Section 4.1.4).

Rappelons également que dans *Frekvenční slovník mluvené češtiny* (2007), se basant exclusivement sur des productions langagières orales et globalement informelles (voir la Section 3.1.2), TEN apparaît comme le lemme le plus fréquent. Malgré l'indistinction entre les deux types d'emploi, ce constat accorde de la crédibilité à la thèse d'une grammaticalisation en cours de cette forme en article défini. Maintenant que nous disposons du chiffre approximatif de la fréquence de son emploi adnominal dans ORTOFON v1, nous croyons judicieux de comparer cette fréquence avec **la fréquence de l'article défini en français**. En effet, il est une chose bien connue que dans les langues qui disposent d'un article défini, celui-ci atteint une très haute fréquence. Ainsi, Leech, Rayson et Wilson (2001), se basant sur les données du British national corpus, constatent que *the* constitue le lexème le plus fréquent en anglais, et ce à la fois dans les productions orales et écrites. S'agissant du français, vu les difficultés relatives à la lemmatisation et la désambiguïsation des formes grammaticales, il n'est pas aisé d'identifier l'ensemble des occurrences de l'article défini LE au sein d'un même corpus. Ce qui pose problème, c'est l'homographie entre les formes de ce dernier et les formes COD du pronom IL, mais, plus encore, la présence de l'article dans les amalgames *au*, *aux*, *du* et *des* (notamment dans les deux derniers, où aucun critère formel ne permet, à notre connaissance,

de distinguer la valeur du défini précédé de la préposition *de* de celle de l'article partitif)¹⁹. Nous avons d'abord regardé la place du défini dans le *Dictionnaire des fréquences*, qui donne le rang des formes *le*, *la* *l'* et *les*. Il ne fait cependant pas la répartition des homographes entre les formes de l'article et les formes COD du pronom IL et ne comptabilise pas les formes apparaissant dans *au*, *aux*, *du* et *des*. Pour la tranche chronologique 1946-1964, les rangs et les fréquences relatives – rapportées à un million de mots et arrondies au dixième – sont les suivants :

Forme	Rang ²⁰	Fréquence relative en mots
<i>La</i>	2	15 408,0
<i>L'</i>	5	12 358,0
<i>Le</i>	7	11 586,7
<i>Les</i>	10	8 687,7
Total		48 040,4

Tableau 5.3 : Rang et fréquence relative des formes du lemme LE pour la période 1946-1964 dans le *Dictionnaire des fréquences : vocabulaire littéraire des XIXe et XXe siècles* (1971 ; entre 4 et 6 millions de mots)

Dans la tranche en question, le premier rang est détenu par le lexème *de*, dont la fréquence relative atteint le chiffre de 31 105,6 occurrences pour un million de mots. Cependant, lorsque nous additionnons les fréquences relatives de toutes les formes de l'article défini, nous parvenons à la somme de 39 352,7 occurrences pour un million de mots, somme qui indique que ce serait en réalité le lemme LE qui serait le plus fréquent²¹.

Quant aux corpus Écrits scientifiques en français, l'*Est républicain* et le corpus défini par nous dans Frantext à partir de textes de fiction publiés entre 1961-2014, il est possible d'établir la fréquence absolue et relative de LE en dehors des formes *au*, *aux*, *du* et *des* tout en écartant, qui plus est, les formes COD de IL, mais sans pouvoir toutefois accéder aux rangs. Pour le corpus Écrits scientifiques, la fréquence relative de LE s'élève à 67 159,6 occurrences pour un million de mots ; pour le corpus de l'*Est républicain*, elle est de 70 597,3 occurrences pour un million de mots, et, pour notre corpus défini dans Frantext, elle atteint 69 704,9 occurrences pour un million de mots. Une fréquence relative approximative peut aussi être obtenue pour notre corpus de conversation du CEFC, où il est possible

¹⁹ Il en va évidemment de même pour *de la*.

²⁰ Il s'agit de la position dans un index hiérarchique qui liste les formes dans l'ordre décroissant en fonction de la fréquence absolue qu'elles atteignent au sein du corpus représentant la tranche chronologique donnée.

²¹ L'on pourrait certes nous objecter que le chiffre réel pourrait être plus bas pour LE du fait de la non dissociation des emplois articulaire et pronominal, et que le chiffre réel pour *de* est certainement plus élevé du fait de la non inclusion des amalgames *des* et *du*. À cela, il est facile de rétorquer qu'en vérité, il est plus probable que le vrai chiffre pour LE soit même plus élevé que celui indiqué, du fait de la non inclusion des amalgames *du*, *des*, *au* et *aux*. De plus, étant donné que les amalgames *de* et *des* regroupent à la fois *de* et LE, leur présence dans les textes fait autant augmenter la fréquence de *de* que celle de LE.

de chercher le lemme LE tout en essayant autant que faire se peut d'éliminer les formes COD de IL²². Le Tableau 5.4 apporte une synthèse des valeurs obtenues²³ :

Corpus	Fréquence absolue	Fréquence relative pour un million de mots	Fréquence relative en %
<i>Dictionnaire des fréquences</i> (1971) : période 1946-1964 ; entre 4 et 6 millions de mots		48 040,4	4,80 %
Écrits scientifiques en français (ScienQuest ; 5 596 492 de mots)	375 858	67 159,6	6,72 %
<i>l'Est républicain</i> (ScienQuest ; 87 341 867 de mots)	6 166 100	70 597,3	7,06 %
Frantext – échantillon de fiction (1961-2014 ; 2 008 222 de mots)	139 983	69 704,9	6,97 %
		Fréquence relative en positions	
Frantext – échantillon de fiction (1961-2014 ; 2 363 566 de positions)	139 983	59 234	5,92 %
Conversations du CEFC (Orféo ; 934 102 positions)	38 856	41 597,2	4,16 %

Tableau 5.4 : Fréquence absolue et relative de l'article défini LE dans la tranche 1946-1964 du corpus du *Dictionnaire des fréquences* (1971), le corpus Écrits scientifiques en français (ScienQuest), le corpus de *l'Est républicain* (ScienQuest), l'échantillon de textes de fiction publiés entre 1961 et 2014 (Frantext) et le corpus de conversations du CEFC (Orféo)

Hormis le corpus du *Dictionnaire des fréquences* (1971) et le corpus de conversation du CEFC, où la fréquence serait plus basse, les trois autres ensembles de données affichent des valeurs assez proches : il semble donc qu'il n'y ait pas de clivage important entre textes de presse, scientifiques et littéraires pour ce qui est de l'emploi de l'article défini dans le français écrit contemporain. Le cas de l'oral est plus délicat, car nous ne pouvons pas obtenir la fréquence relative en mots.

Afin que nous puissions à présent comparer ces chiffres avec la fréquence relative de TEN adnominal en tchèque, il nous faut connaître celle-ci en nombre de mots : 19 054,3-19 606,1, soit

²² [lemme="le"][!pos="VER"].

²³ Les chiffres des fréquences relatives en mots et en positions sont arrondis au dixième ; ceux des fréquences relatives exprimées en % le sont au centième.

1,91-1,96 % de l'ensemble. Bien qu'il faille toujours être conscient que les chiffres dont nous disposons ne sont qu'approximatifs dans la mesure où, ni pour l'une ni pour l'autre langue, il n'est possible d'identifier toutes les occurrences du phénomène recherché²⁴, ils nous donnent tout de même un tableau général de la situation. En effet, par comparaison avec celle du défini dans les textes français, la fréquence de TEN adnominal établie par nous pour ORTOFON v1 s'avère plutôt basse. Le Tableau 5.5 reprend les valeurs des fréquences relatives en % :

Corpus	Fréquence relative en mots (en %)
<i>l'Est républicain</i> (ScienQuest ; 87 341 867 de mots)	7,06 %
Frantext – échantillon de fiction (1961-2014 ; 2 008 222 de mots)	6,97 %
Écrits scientifiques en français (ScienQuest ; 5 596 492 de mots)	6,72 %
<i>Dictionnaire des fréquences</i> (1971) : période 1946-1964 ; entre 4 et 6 millions de mots	4,80 %
ORTOFON v1 (CNT ; 1 014 786 de mots)	1,91 %-1,96 %
	Fréquence relative en positions (en %)
Frantext – échantillon de fiction (1961-2014 ; 2 363 566 de positions)	5,92 %
Conversations du CEFC (934 102 positions)	4,16 %
ORTOFON v1 (CNT ; 1 236 508 de positions)	1,27 %-1,30 %

Tableau 5.5 : Fréquence relative de l'article défini LE dans l'Est républicain (ScienQuest), l'échantillon de textes de fiction publiés entre 1961 et 2014 (Frantext), le corpus Écrits scientifiques en français (ScienQuest), la tranche 1946-1964 du corpus du *Dictionnaire des fréquences* (1971) et le corpus de conversations du CEFC (Orféo) ; fréquence relative de TEN adnominal dans le corpus de conversations d'ORTOFON v1 (CNT)

Nous voyons ainsi que la fréquence relative de TEN adnominal dans ORTOFON v1 reste bien en-dessous des fréquences relatives de l'article défini LE dans les différents types de textes français en synchronie. Même si nous prenons la valeur supérieure de l'intervalle de la fréquence relative de TEN adnominal dans ORTOFON v1, celle-ci ne fait qu'un tiers de la fréquence relative du défini dans le corpus de conversations. Lorsque nous appliquons la fonction de vraisemblance pour mesurer la significativité statistique de l'écart entre la valeur supérieure de l'intervalle de la fréquence absolue de TEN adnominal dans ORTOFON v1 et celle du défini dans le corpus des conversations, le score est

²⁴ Ou bien, comme c'est le cas du *Dictionnaire des fréquences* (1971), il est impossible d'éliminer la partie d'occurrences indésirables. S'ajoute à cela le fait que les fréquences relatives en tchèque et en français sont influencées par les différences typologiques entre les deux langues : en français (langue analytique), la fréquence relative est toujours moins élevée qu'en tchèque, et ce même si la fréquence absolue dans les textes est identique (cf. Nádvorníková, 2020).

égal à 12 894,4718²⁵. Aussi approximatif qu'il soit, ce bilan incite à la prudence : l'on voit que même au sein de l'oral informel, TEN adnominal est loin d'avoir atteint la même place que l'article défini dans les productions françaises, quelle que soit leur nature.

Dans les trois sections qui suivent, nous examinerons plus en détail les échantillons de 1 000 occurrences prélevés sur les résultats obtenus pour le français et le tchèque oraux informels tout comme les 241 occurrences du sous-corpus des pièces de V. Havel.

5.2.2 Le Corpus d'Étude pour le français contemporain

5.2.2.1 Premiers constats

Le Tableau 5.6, résume les fréquences des types et sous-types d'emploi dans l'échantillon de 1 000 occurrences du déterminant démonstratif français CE²⁶. Se pose, entre autres, la question de savoir dans quelle mesure ce bilan peut être considéré comme représentatif de l'ensemble de la population obtenue ($n = 1871$ occurrences). Afin d'avoir une réponse à cette question, nous avons recouru à l'interface « many features – 1 sample » (voir la Section 2.1.5), où nous avons saisi les nombres d'occurrences pour chacun des types et sous-types d'emploi identifiés au sein dudit échantillon de 1 000 occurrences. Comme le Tableau 5.6 affiche aussi les intervalles de confiance, il permet de voir si le type ou sous-type en question : 1) est susceptible d'apparaître dans un autre échantillon de 1 000 occurrences prélevé sur la population totale²⁷ et, si oui, 2) quelle y sera sa fréquence approximative. Cela nous donne ainsi une idée brute de la représentation des différents types et sous-types d'emploi dans l'ensemble de la population totale. Le graphique de la Figure 5.1 représente les proportions prises, au sein de l'échantillon, par les différents (sous-)types d'emploi²⁸ :

Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des 1 000 occurrences (%)	Intervalle de confiance	
			Borne inférieure	Borne supérieure
ENDOPHORE	489	48,90 %	458	520
Anaphore résomptive	170	17,00 %	147	193
Anaphore fidèle	142	14,20 %	121	164
Anaphore infidèle	112	11,20 %	93	132
Anaphore temporelle	44	4,40 %	32	57
Anaphore indirecte	19	1,90 %	11	28

²⁵ La valeur critique est de 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

²⁶ Nous invitons le lecteur à consulter la Section 4.5 du Chapitre 4 pour des exemples de ces différents (sous-)types d'emploi.

²⁷ À condition, bien entendu, de tenir à l'écart le bruit et dans la situation virtuelle où cette population serait suffisamment grande (rappelons qu'elle ne compte que 1871 occurrences au total).

²⁸ Le graphique reprend les types d'emploi là où ceux-ci ne sont pas divisés en sous-types ; dans le cas contraire, il reprend les sous-types d'emploi.

Cataphore	2	0,20 %	0	5
EXOPHORE A-MEMORIELLE	389	38,90 %	359	419
Deixis temporelle	282	28,20 %	254	310
Deixis situationnelle	93	9,30 %	75	111
Deixis textuelle	14	1,40 %	7	22
EXOPHORE MEMORIELLE	62	6,20 %	47	77
OCCURRENCES AMBIGUES	40	4,00%	28	53
EMPLOIS « EMOTIONNELS »	11	1,10 %	5	18
DEMONSTRATIFS INDEFINIS	9	0,90 %	4	15
TOTAL	1 000	100,00 %		

Tableau 5.6 : Types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences) avec les intervalles de confiance anticipant leur fréquence dans un autre échantillon de la même taille

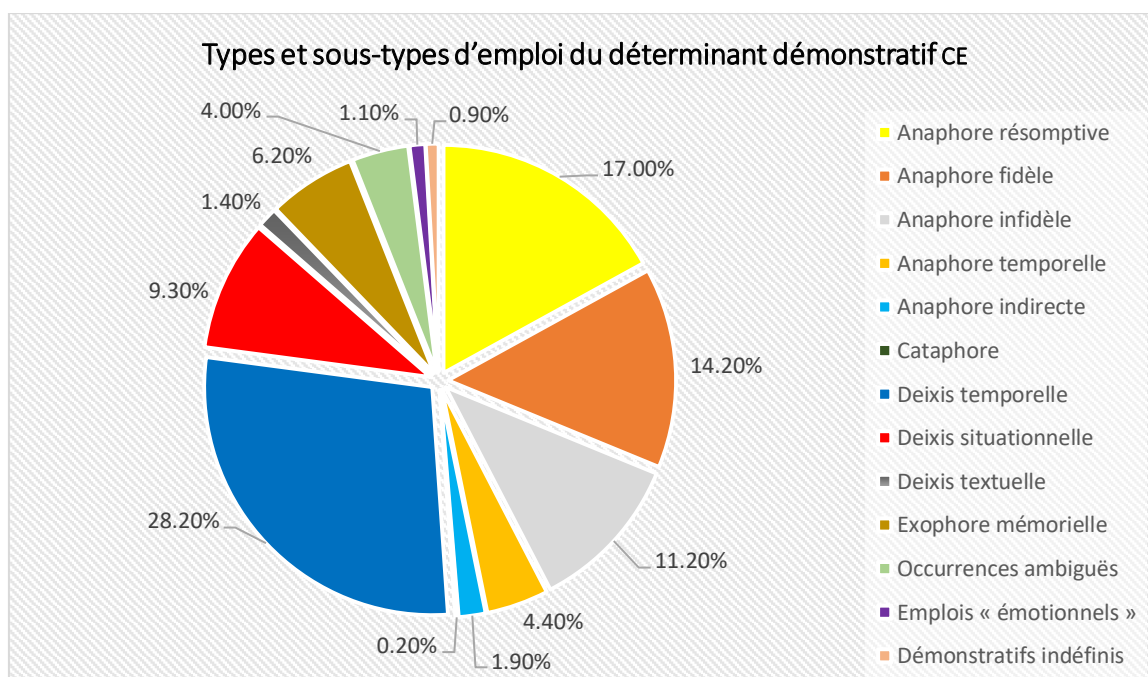


Figure 5.1 : Types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)

Une première série de constats s'impose : la majorité des occurrences de l'échantillon annoté apparaît en emploi endophorique (489, soit 48,9 %), lequel est suivi de l'emploi exophorique a-mémoriel (389, soit 38,9 %). L'échantillon contient relativement peu d'emplois exophoriques mémoriels (62/6,2 %) et très peu d'emplois « émotionnels » (11/1,1 %)²⁹. Au sein des emplois endophoriques, c'est l'anaphore résomptive qui prime, avec 170 occurrences (17,0 % de l'ensemble et 34,8 % des occurrences endophoriques), suivie des anaphores fidèle (142 occurrences, soit 14,2 % de l'ensemble et 29 % des occurrences endophoriques) et infidèle (112 occurrences, soit 11,2 % de l'ensemble et 22,9 % des occurrences endophoriques). Le nombre d'occurrences anaphoriques

²⁹ Dans les Section 5.2.3 et 5.2.4, ces chiffres seront comparés avec les chiffres obtenus pour l'échantillon de conversations tchèques et ceux obtenus pour le sous-corpus des pièces de V. Havel.

temporelles est loin d'être négligeable (44, soit 4,4 % de l'ensemble et 9 % des occurrences endophoriques), surtout si nous tenons compte des spécificités de cet emploi du déterminant démonstratif.

S'agissant des fréquences avec lesquelles les différents types et sous-types sont susceptibles de se reproduire dans un autre échantillon, nous voyons que plus le type d'emploi est représenté dans l'échantillon, plus l'intervalle de confiance est précis : ainsi, pour l'ensemble des emplois endophoriques, il se situe autour de ± 6 %, alors qu'il est très large (et donc inexact) pour l'emploi déictique textuel (+36,36/ -50 %). Pour l'emploi cataphorique, la fréquence est même si basse qu'elle se situe en-dessous du seuil de significativité³⁰ : il est donc impossible de savoir si cet emploi réapparaîtra dans un autre échantillon (la borne inférieure de l'intervalle de confiance – marqué en rouge – atteint zéro).

Regardons à présent le rapport entre les types et sous-types d'emploi et **l'apparition des particules localisatrices -ci et -là**³¹. Le Tableau 5.7 croise l'apparition des particules avec le critère du type et du sous-type d'emploi (l'analyse part donc des (sous-)types d'emploi). Le graphique de la Figure 5.2 représente la part occupée par ces sous(-types) d'emploi au sein des occurrences pourvues de la particule -là³² (l'analyse prend pour point de départ la présence de la particule -là) :

Type et sous-type d'emploi	Nombre d'occurrences		Par rapport à l'ensemble des occurrences du (sous-)type en question (%)	
	-CI	-LA	-CI	-LA
EXOPHORE A-MEMORIELLE	9	23	2,31 %	5,91 %
Deixis situationnelle	5	20	5,38 %	21,51 %
Deixis temporelle	4	2	1,42 %	0,71 %
Deixis textuelle		1		7,14 %
EXOPHORE MEMORIELLE		8		12,90 %
ENDOPHORE	4	146	0,82 %	29,86 %
Anaphore fidèle		31		21,83 %
Anaphore infidèle		25		22,32 %
Anaphore indirecte		4		21,05 %
Anaphore résomptive		60		35,29 %
Anaphore temporelle	4	25	9,09 %	56,82 %
Cataphore		1		50,00 %
DEMONSTRATIFS INDEFINIS		5		55,56 %
OCCURRENCES AMBIGUES	1	6	2,50 %	15,00 %
TOTAL	14	188	1,40 %	18,80 %

³⁰ $\alpha = 0,05$.

³¹ Seules les valeurs avec lesquelles la particule est présente sont représentées. Les % sont, dans l'ensemble de ce chapitre, arrondis au centième.

³² Nous n'avons pas fait de même pour les occurrences pourvues de la particule -ci, vu leur très faible présence.

Tableau 5.7 : Présence de *-ci* et *-là* avec les différents types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)

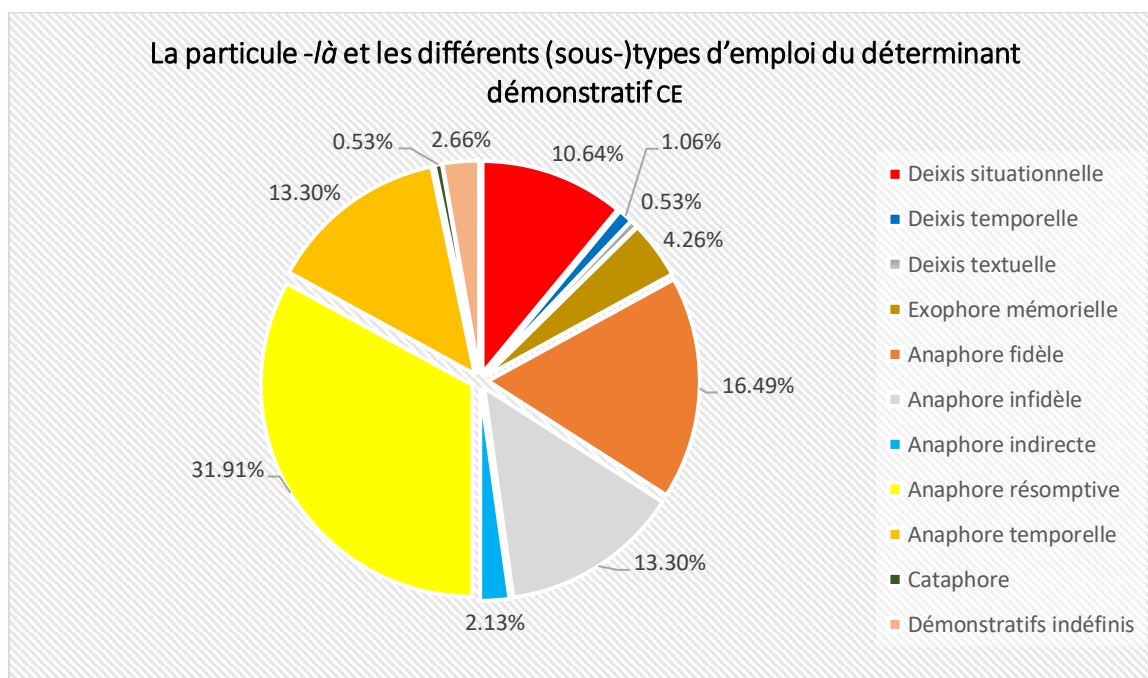


Figure 5.2 : Présence des différents (sous-)types d'emploi du déterminant démonstratif CE au sein des occurrences contenant *-là* dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)

Le tableau et le graphique complètent, sous forme de synthèses, les constats formulés dans les Sections 4.1.4 et 4.5 : très peu d'occurrences de *-ci* (14, soit 1,4 % de l'ensemble), dont 9 ressortissant à l'exophore a-mémorielle et 4 ressortissant au régime de l'anaphore temporelle ; 2) beaucoup plus d'occurrences de *-là* (188, soit 18,8 % de l'ensemble), dont la majorité (146/77,66 %) apparaît avec l'endophore. Ainsi, nous découvrons que les occurrences avec *-là* constituent 29,86 % de toutes les occurrences endophoriques, soit 21,83 % des occurrences anaphoriques fidèles, 22,32 % des anaphoriques infidèles, 35,29 % des anaphoriques résomptives et 56,82 % des anaphoriques temporelles. S'agissant plus particulièrement de ces dernières, au sein de 29 des 44 occurrences, le nom est pourvu d'une particule localisatrice, soit dans 65,9 % des cas.

Les noms qui apparaissent le plus fréquemment avec les anaphores résomptive et temporelle (avec ou sans particule localisatrice) sont les suivants (le nombre d'occurrences est donné entre parenthèses pour chaque nom) :

Anaphore résomptive : *cas* (19), *genre* (19), *moment* (11), *point* (10), *temps* (6), *histoire* (6), *côté* (6), *trucs* (5), *niveau* (4).

Anaphore temporelle : *fois* (16), *moment* (14), *jour* (5), *époque* (4), *soir* (3)³³

³³ Comme seuls les noms les plus fréquents sont représentés, le nombre total n'est pas égal au nombre de toutes les occurrences anaphoriques en question présentes dans l'échantillon.

Le graphique de la Figure 5.2 permet ensuite de répondre à la question de savoir **quel est le (sous-)type d'emploi le plus représenté parmi les occurrences pourvues de -là**. C'est l'anaphore résomptive qui prime (31,91 % de la totalité des occurrences avec -là), suivie des anaphores fidèle (16,49 % de l'ensemble), infidèle et temporelle (toutes les deux 13,30 % de l'ensemble). La part occupée par la deixis situationnelle est de 10,64 %.

De prime abord, il semble ainsi que **des affinités** soient présentes, dans notre échantillon, **entre les occurrences en -là et l'endophore** de manière générale et, plus précisément, **entre les occurrences en -là et les anaphores résomptive et temporelle**. Afin de confirmer cette hypothèse, nous avons recouru au calcul de spécificités. En effet, comme le montrent les graphiques des Figures 5.3 et 5.4, affichant les résultats du calcul, -là (barres vertes) est surreprésenté au sein de l'anaphore résomptive (première colonne) et de l'anaphore temporelle (cinquième colonne) ; de manière plus générale, -là présente une affinité avec l'endophore (première colonne) tout en évitant de s'associer avec l'exophore (a-mémorielle et mémorielle : deuxième et troisième colonnes respectivement), laquelle ne s'associe le plus souvent avec aucune particule (barres rouges)³⁴ :

³⁴ « Par convention, la représentation de la sous-spécificité (ou sous-représentation) se distingue de celle de la sur-spécificité (ou sur-représentation) par un signe moins (-) situé devant l'indice. On s'intéressera alors aux faibles probabilités (donc aux valeurs de $\log_{10}$ importantes) qui rendent compte : soit d'un nombre d'apparitions plus faible que prévu si l'observation est inférieure au mode de la distribution théorique (c'est-à-dire si le nombre d'apparitions de l'événement dans la partie est inférieur au maximum de vraisemblance estimé par notre modélisation hypergéométrique de la distribution [...]) On parlera alors de sous-spécificité ou spécificité négative ; soit d'un nombre d'apparition plus important que prévu si l'observation est supérieure au mode de la distribution théorique. On parlera alors de sur-spécificité ou spécificité positive. » Dans le graphique, « [...] deux lignes rouges délimitent la bande de banalité autour de l'axe d'indice 0 (les barres qui n'en sortent pas sont à considérer comme banales). » <http://textometrie.ens-lyon.fr/html/doc/manual/0.7.9/fr/manual43.xhtml#:~:text=Le%20calcul%20exact%20de%20l,suivant%20la%20loi%20hyperg%C3%A9om%C3%A9trique%20d%C3%A9crite>. Ainsi, une barre comprise dans la bande de banalité correspond à une valeur anticipée par le mode de la distribution théorique. Dans nos calculs, la bande de banalité est fixée à $|2|$ ou à $|3|$, soit 10^{-2} ou 10^{-3} .

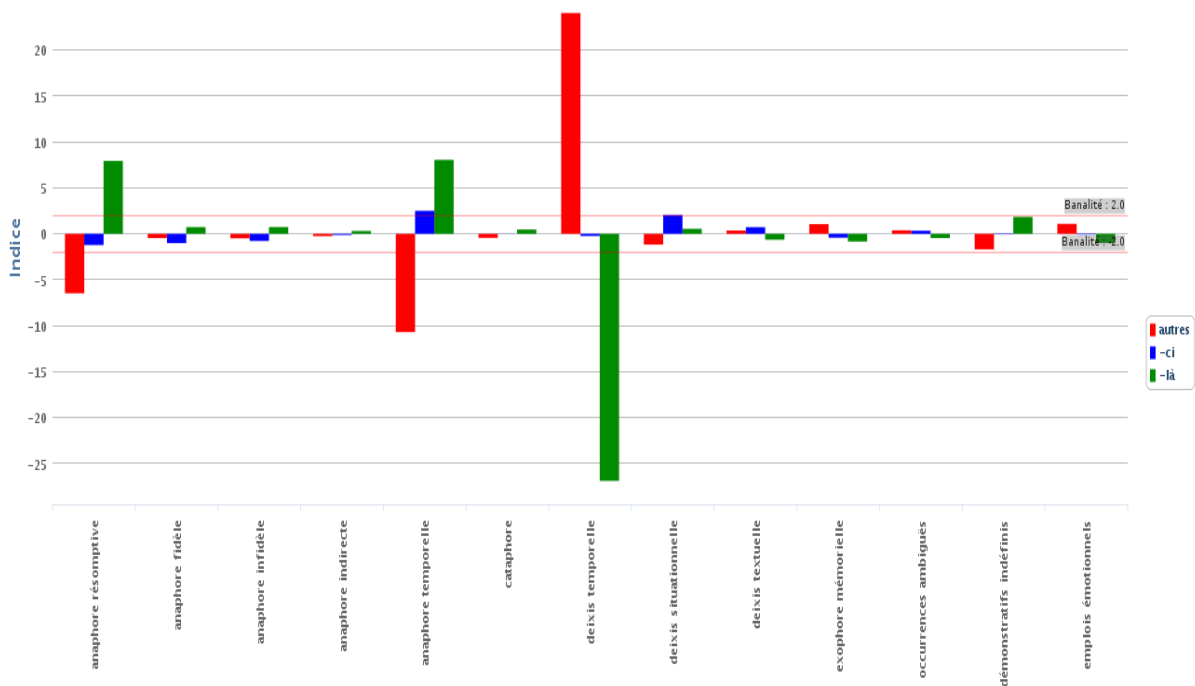


Figure 5.3 : Calcul de spécificités appliqué au lien entre les différents (sous-)types d’emploi du déterminant démonstratif CE et la présence de particules localisatrices dans l’échantillon de conversations du CEFC (Orféo ; 1 000 occurrences) ; rouge = sans particule, vert = *-là*, bleu = *-ci*

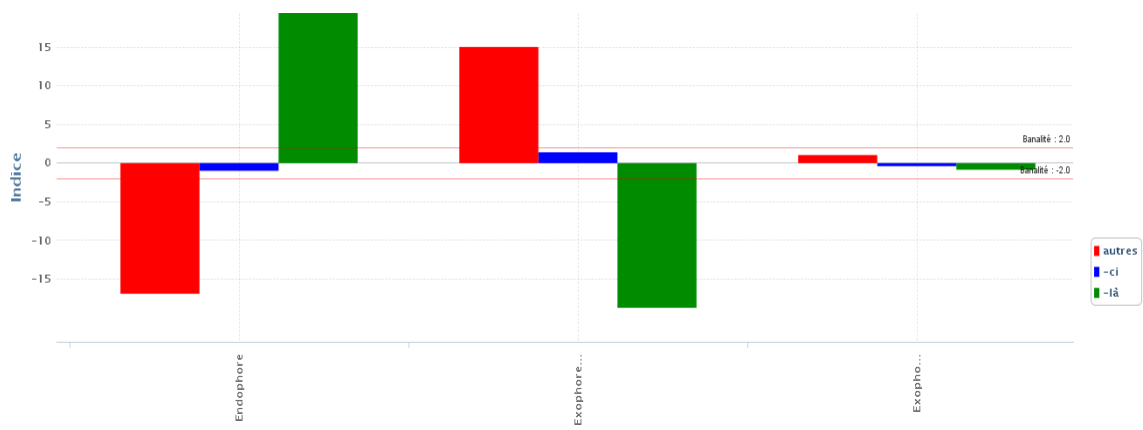


Figure 5.4 : Calcul de spécificités appliqué au lien entre les types d’emploi endophorique, exophorique a-mémoriel et exophorique mémoriel du déterminant démonstratif CE et la présence de particules localisatrices dans l’échantillon de conversations du CEFC (Orféo ; 1 000 occurrences) ; rouge = sans particule, vert = *-là*, bleu = *-ci*

C’est à la fin du présent chapitre, dans la Section 5.3, que nous entreprendrons une analyse plus poussée de tous les noms à sémantique temporelle dans les deux échantillons de conversations, à la fois dans le domaine de l’endophore et de la deixis temporelle.

5.2.2.2 Position de l'occurrence de CE par rapport au verbe de la proposition

Le Tableau 5.8 et le graphique de la Figure 5.5 *infra* apportent une synthèse des fréquences en fonction des positions de l'occurrence du SN démonstratif par rapport au verbe de la proposition telles que celles-ci ont été définies dans la Section 2.2.3³⁵. Toutes les fréquences sont complétées d'un intervalle de confiance, calculé par l'interface « many features – 1 sample » du CNT et permettant, comme nous venons de le voir *supra* pour les (sous-)types d'emploi, d'établir la fréquence approximative que les différentes positions atteignent au sein de l'ensemble de la population totale (plus précisément : dans un autre échantillon de la même taille)³⁶ :

Position	Nombre d'oc- currences	Par rapport à l'ensemble des occurrences de l'échantillon (%)	Intervalle de confiance	
			Borne inférieure	Borne supérieure
POSITION POSTVERBALE	359	35,90 %	330	389
COMPLEMENT DE PHRASE	356	35,60 %	326	386
DISLOCATION A DROITE	144	14,40 %	123	166
POSITION INCONNUE	70	7,00 %	55	86
DISLOCATION A GAUCHE	48	4,80 %	35	62
POSITION PREVERBALE	19	1,90 %	11	28
CLIVAGE	4	0,40 %	1	8
TOTAL	1 000	100,00 %		

Tableau 5.8 : Position de l'occurrence du déterminant démonstratif CE par rapport au verbe de la proposition dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)

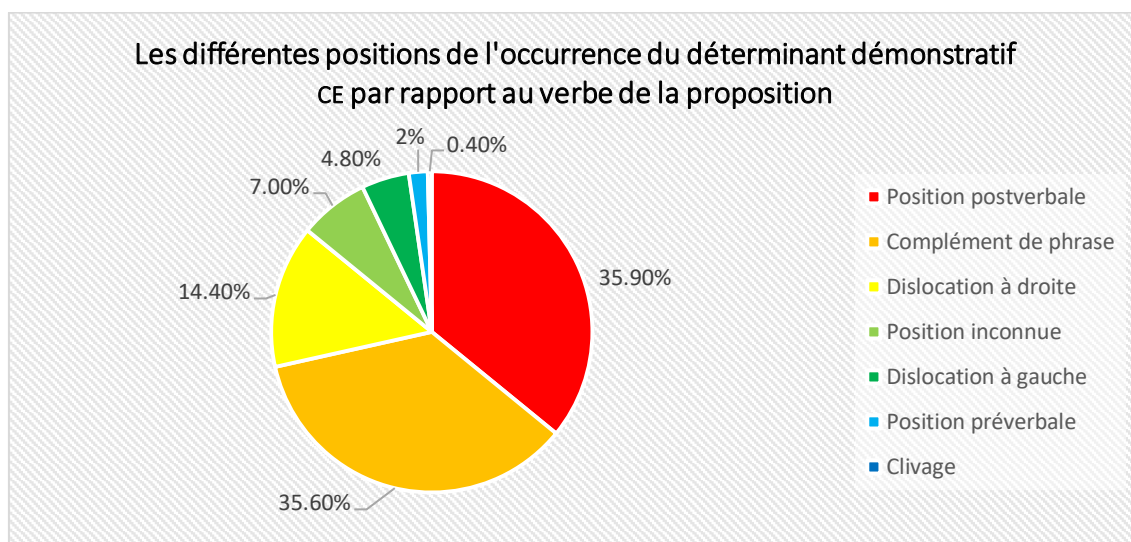


Figure 5.5 : Représentation en % des différentes positions de l'occurrence du déterminant démonstratif CE par rapport au verbe de la proposition dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)

³⁵ Le lecteur trouvera également dans cette section des exemples des différentes positions pour les deux échantillons de conversations.

³⁶ Tout en écartant les occurrences bruitées ; $\alpha = 0,05$.

Cet ensemble de données nous permet de formuler une autre série de constats : notre échantillon contient le plus d'occurrences en positions postverbale (359) et de complément de phrase (356). Un nombre significatif d'occurrences apparaît également en dislocation à droite (144).

Croisons à présent les données du Tableau 5.8 avec celles du Tableau 5.6 pour voir s'il existe **des affinités entre**, d'une part, **l'occurrence par rapport au verbe et**, d'autre part, **les différents types et sous-types d'emploi**. Les chiffres sont présentés dans une série de tableaux dans l'ordre décroissant, à commencer par la position par rapport au verbe la plus répandue³⁷. Ces tableaux sont complétés par des graphiques qui affichent les proportions des différents (sous-)types d'emploi au sein de chaque type de position par rapport au verbe de la proposition³⁸ : alors que les tableaux adoptent le point de vue des (sous-)types d'emploi, les graphiques adoptent celui de la position par rapport au verbe³⁹. Seule ne sera pas traitée la position du clivage, très peu présente dans l'échantillon (2 occurrences endophoriques et 2 occurrences exophoriques a-mémorielles). Pour ce qui est de la position inconnue, vu la difficulté d'analyse qu'elle présente, nous nous limiterons à la catégorie des types d'emploi.

POSITION POSTVERBALE		
Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
ENDOPHORE	185	37,83 %
Anaphore résomptive	69	40,59 %
Anaphore fidèle	60	42,25 %
Anaphore infidèle	42	37,50 %
Anaphore indirecte	9	47,37 %
Anaphore temporelle	5	11,36 %
EXOPHORE A-MEMORIELLE	100	25,71 %
Deixis temporelle	60	21,28 %
Deixis situationnelle	30	32,26 %
Deixis textuelle	10	71,43 %
EXOPHORE MEMORIELLE	43	69,35 %
OCCURRENCES AMBIGUES	20	50,00 %
DEMONSTRATIFS INDEFINIS	8	88,89 %
EMPLOIS « EMOTIONNELS »	3	27,27 %
TOTAL	359	35,90 %

Tableau 5.9 : Rapport entre la position postverbale et les différents types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 359 occurrences)

³⁷ Seuls sont représentés les types et sous-types d'emploi apparaissant dans la position en question.

³⁸ Les graphiques représentent les types d'emploi là où ceux-ci ne sont pas divisés en sous-types ; s'ils le sont, ce sont les sous-types qui sont représentés.

³⁹ Enfin, le calcul de spécificités permet de faire une synthèse de ces deux points de départ de l'analyse (voir la Section 2.1.6).

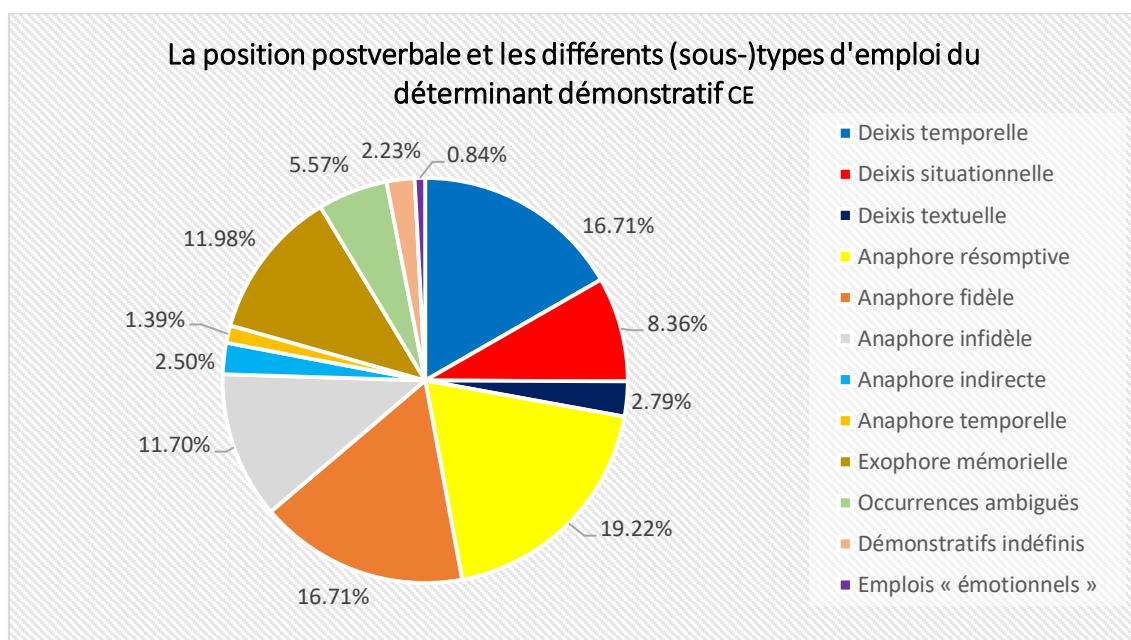


Figure 5.6 : Représentation en % des différents (sous-)types d'emploi du déterminant démonstratif CE au sein de la position postverbale dans l'échantillon de conversations du CEFC (Orfeo ; 359 occurrences)

Concernant les SN démonstratifs de **la position postverbale**, si la plupart d'entre eux apparaissent avec une valeur endophorique (185/51,53 %), beaucoup d'entre eux apparaissent également avec une valeur exophorique a-mémorielle (100/27,86%). Au sein de l'endophore, ce sont les anaphores résomptive (69/19,22 %) et fidèle (60/16,71%) qui priment. Au sein de l'exophore a-mémorielle, nous trouvons notamment la deixis temporelle (60/16,71 %). Si l'exophore mémorielle ne représente que 11,98 % de toutes les occurrences apparaissant en position postverbale, la majorité des occurrences exophoriques mémorielles de notre échantillon, soit 43 des 62 (69,35 %), occupe cette position. Cela prouve **une affinité certaine entre l'exophore mémorielle et la position postverbale**, représentée par le graphique de la Figure 5.12 *infra*. La part des emplois endophoriques apparaissant en position postverbale est de 37,83 %, et celle des emplois exophoriques a-mémoriels est de 25,71 %.

COMPLEMENT DE PHRASE		
Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
EXOPHORE A-MEMORIELLE	221	56,81 %
Deixis temporelle	207	73,40 %
Deixis situationnelle	14	15,05 %
ENDOPHORE	122	24,95 %
Anaphore résomptive	48	28,24 %
Anaphore temporelle	36	81,82 %
Anaphore fidèle	22	15,49 %
Anaphore infidèle	12	9,84 %
Anaphore indirecte	2	10,53 %

Cataphore	2	100,00 %
EXOPHORE MEMORIALE	7	11,29 %
OCCURRENCES AMBIGUES	6	15,00 %
TOTAL	356	35,60 %

Tableau 5.10 : Rapport entre la position de complément de phrase et les différents types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orfeo ; 356 occurrences)

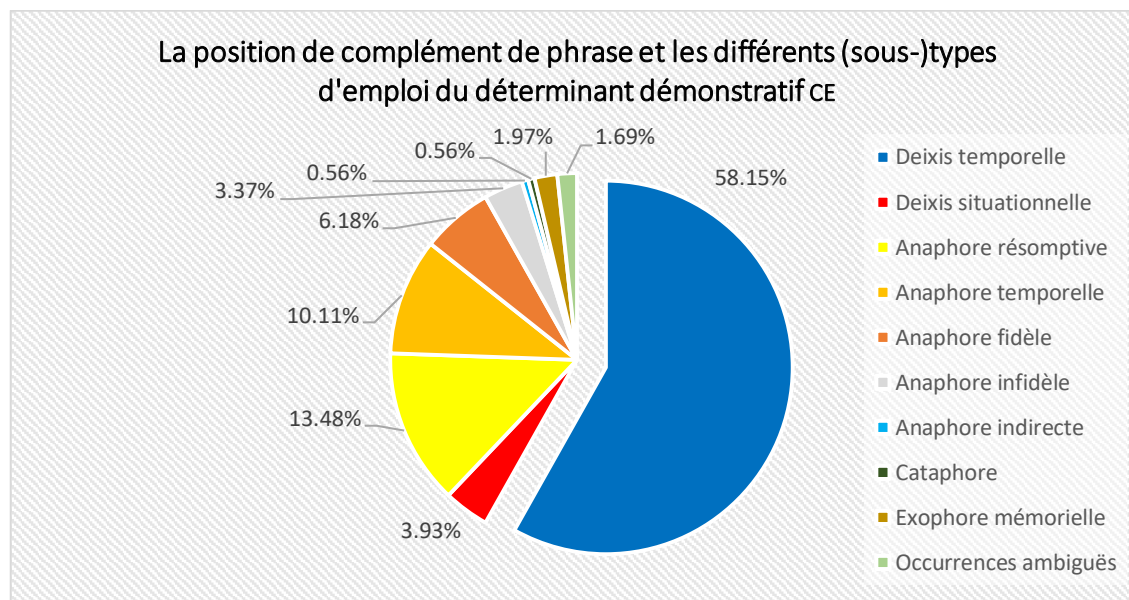


Figure 5.7 : Représentation en % des différents (sous-)types d'emploi du déterminant démonstratif CE au sein de la position de complément de phrase dans l'échantillon de conversations du CEFC (Orfeo ; 356 occurrences)

Parmi les occurrences placées en position de complément de phrase, la majorité appartient à la deixis temporelle (207/58,14 % ; en revanche, très peu relèvent de la deixis situationnelle). Il n'est guère étonnant que la majorité des occurrences endophoriques placées en cette position ressortissent aux anaphores résomptive et temporelle prises ensemble (84 occurrences, soit 23,59 % de la totalité des occurrences occupant cette position). En effet, ce sont ces deux types d'anaphore qui semblent les plus à même d'établir un cadre pour l'action décrite dans la proposition. Du côté de l'anaphore résomptive, cela se fait notamment grâce à la présence de noms fonctionnant comme des connecteurs argumentatifs ou temporels : *dans ce cas-là*, *à ce moment-là*, *pendant ce temps*, etc. Il est ainsi symptomatique que des 48 occurrences anaphoriques résomptives en position de complément de phrase, le nombre de celles contenant l'un des trois noms *cas*, *moment* et *temps* s'élève à 33 (18 pour *cas*, 9 pour *moment* et 6 pour *temps*⁴⁰). Du côté de l'anaphore temporelle, nous constatons que 14 occurrences sur 36 placées en position de complément de phrase contiennent le

⁴⁰ Ainsi, comparés aux fréquences de ces trois noms dans le régime de l'anaphore résomptive, ces chiffres révèlent que dans la totalité ou dans l'écrasante majorité des cas, lorsque l'un des trois apparaît en anaphore résomptive, l'occurrence est placée en position de complément de phrase (18 occurrences sur 19 pour *cas*, 9 occurrences sur 11 pour *moment* et 6 occurrences sur 6 pour *temps*).

nom *fois* et 10 contiennent le nom *moment*⁴¹. Nous attirons également l'attention sur le fait que 36 des 44 occurrences anaphoriques temporelles de l'échantillon apparaissent en position de complément de phrase (81,82 %). En conséquence, nous pouvons dire que notre corpus révèle **une affinité entre l'anaphore temporelle** d'une part **et la position syntaxique de complément de phrase** d'autre part. Une affinité encore plus grande peut être observée **entre la position de complément de phrase et la deixis temporelle**, où 207 des 282 occurrences, soit 73,4 %, occupent la place d'un complément de phrase⁴². Ce chiffre dépasse de loin toutes les autres positions, y compris la deuxième position la plus présente avec la deixis temporelle, à savoir la position postverbale (60 des 282 occurrences, soit 21,28 %, voir *supra*). Rappelons que la part de la deixis temporelle au sein la totalité des occurrences apparaissant en position de complément de phrase est également très significative : 207 des 356, soit 58,14 %. De la même façon que les anaphores résomptive et temporelle, la deixis temporelle sert à la mise en place d'un cadre (temporel) de l'action ou du procès exprimés par le verbe. Les affinités entre les emplois anaphorique et déictique temporels et la position de complément de phrase sont observables grâce au graphique de la Figure 5.12 *infra*, affichant les résultats du calcul de spécificités pour les différentes positions et les différents (sous-)types d'emploi

DISLOCATION A DROITE		
Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
ENDOPHORE	104	21,27 %
Anaphore infidèle	38	31,15 %
Anaphore fidèle	37	26,06 %
Anaphore résomptive	26	15,29 %
Anaphore indirecte	3	15,79 %
EXOPHORE A-MEMORIELLE	24	6,17 %
Deixis situationnelle	22	23,66 %
Deixis textuelle	2	14,29 %
EXOPHORE MEMORIELLE	6	9,68 %
EMPLOIS « EMOTIONNELS »	5	45,45 %
OCCURRENCES AMBIGUËS	5	12,50 %
TOTAL	144	14,40 %

Tableau 5.11 : Rapport entre la position de la dislocation à droite et les différents types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orfeo ; 144 occurrences)

⁴¹ Les nombres totaux respectifs d'occurrences anaphoriques temporelles contenant les noms *fois* et *moment* sont de 16 et de 14.

⁴² Bien entendu, ces affinités, et notamment celle qui concerne l'anaphore temporelle, doivent être relativisées au vu de la taille limitée de l'échantillon.

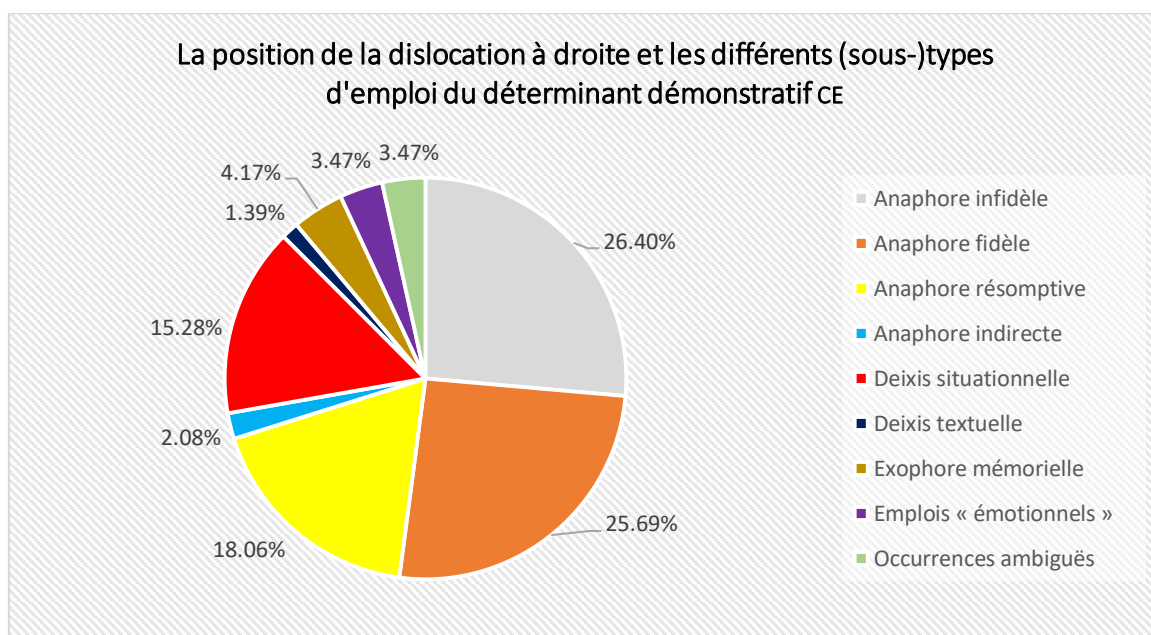


Figure 5.8 : Représentation en % des différents (sous-)types d'emploi du déterminant démonstratif CE au sein de la position de la dislocation à droite dans l'échantillon de conversations du CEFC (Orféo ; 144 occurrences)

POSITION INCONNUE

Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
ENDOPHORE	37	7,56 %
EXOPHORE A-MEMORIELLE	19	4,88 %
OCCURRENCES AMBIGUES	7	17,50 %
EXOPHORE MEMORIELLE	4	6,45 %
EMPLOIS « EMOTIONNELS »	2	18,18 %
DEMONSTRATIFS INDEFINIS	1	11,11 %
TOTAL	70	7,00 %

Tableau 5.12 : Rapport entre la position inconnue et les différents types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 70 occurrences)

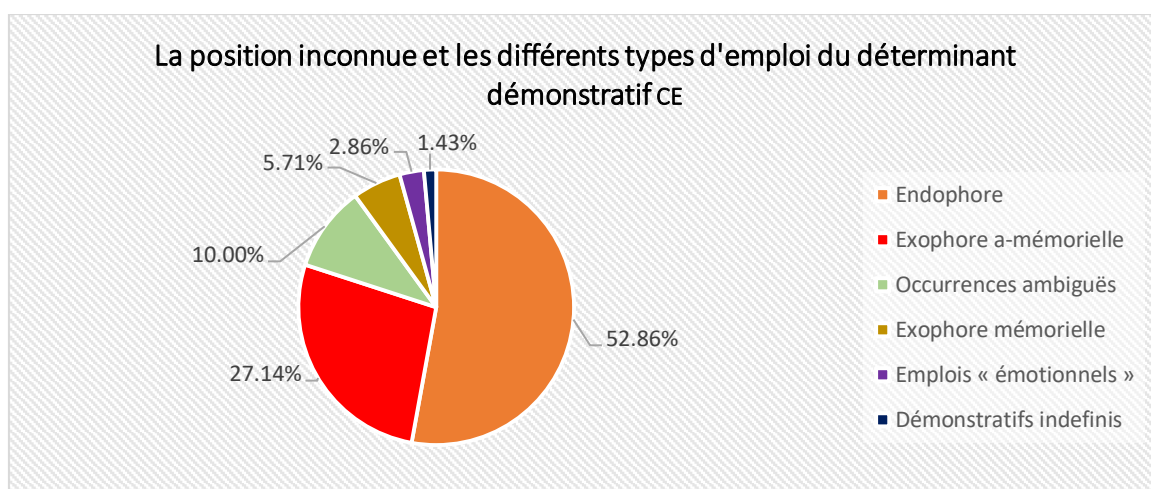


Figure 5.9 : Représentation en % des différents types d'emploi du déterminant démonstratif CE au sein de la position inconnue dans l'échantillon de conversations du CEFC (Orféo ; 70 occurrences)

DISLOCATION A GAUCHE

Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
ENDOPHORE	31	6,34 %
Anaphore infidèle	11	9,02 %
Anaphore fidèle	11	7,75 %
Anaphore résomptive	8	4,71 %
Anaphore indirecte	1	5,26 %
EXOPHORE A-MEMORIELLE	13	3,34 %
Deixis situationnelle	11	11,83 %
Deixis temporelle	1	0,35 %
Deixis textuelle	1	7,14 %
EXOPHORE MEMORIELLE	2	3,23 %
EMPLOIS « EMOTIONNELS »	1	9,10 %
OCCURRENCES AMBIGUES	1	2,50 %
TOTAL	48	4,80 %

Tableau 5.13 : Rapport entre la position de la dislocation à gauche et les différents types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orfeo ; 48 occurrences)

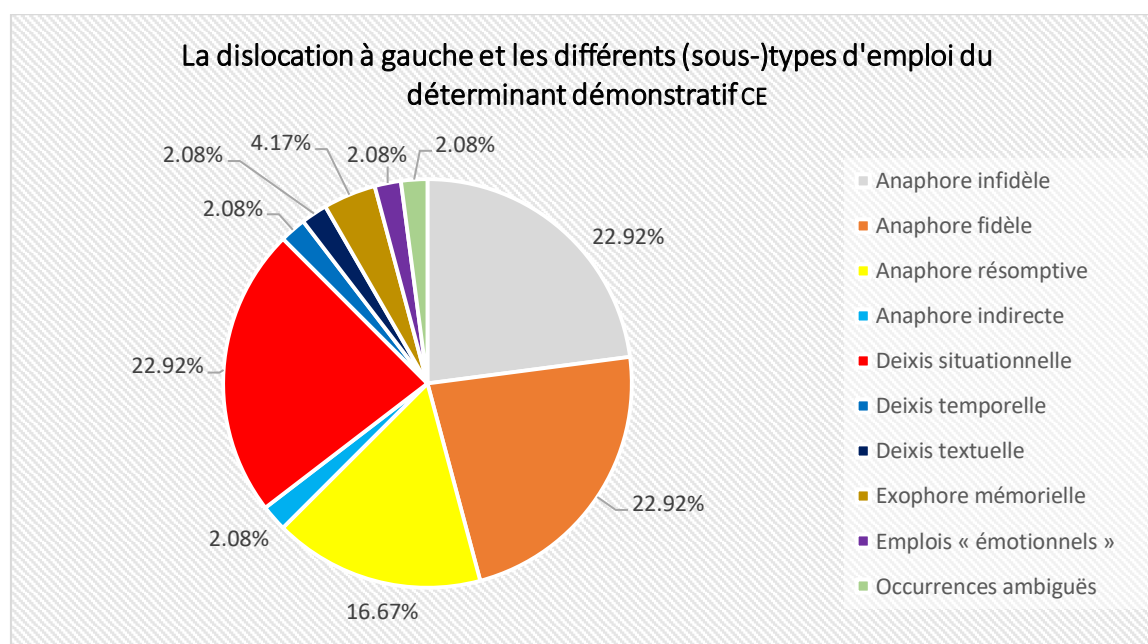


Figure 5.10 : Représentation en % des différents (sous-)types d'emploi du déterminant démonstratif CE au sein de la position de la dislocation à gauche dans l'échantillon de conversations du CEFC (Orfeo ; 48 occurrences)

Examinons maintenant de plus près la situation concernant **les deux types de dislocation**. D'abord, **une affinité** est présente dans notre échantillon **entre la dislocation à droite et l'endophere** (voir le graphique de la Figure 5.12 pour les résultats du calcul de spécificités) : 104 des 144 occurrences placées en cette position sont anaphoriques (soit 72,22 %), contre seulement 24 qui sont déictiques (16,67 % ; l'écrasante majorité, soit 22, relève de la deixis situationnelle). Quant à **la dislocation à gauche**, qui est bien moins représentée dans le corpus (48 occurrences), 31 occurrences

de celle-ci accueillent un emploi anaphorique (64,59 %) et 13 un emploi déictique (27,08 % ; dont 11 cas de déixis situationnelle). Si nous partons ici de la synthèse de la différence entre les deux types de dislocation faite par Ziková (2017 : 237) et rappelée dans la Section 2.2.3, la présence plus élevée de la dislocation à droite avec l'anaphore pourrait sembler quelque peu surprenante, dans la mesure où les référents subissant la dislocation à droite sont censés se recruter davantage parmi ceux qui sont identifiables via la situation d'énonciation, contrairement aux référents faisant l'objet de la dislocation à gauche. Mais cette situation pourrait aussi se justifier, d'après nous, par la prise en considération de deux facteurs. D'abord, il y a le fait que bon nombre des occurrences anaphoriques (infidèles et fidèles pour la plupart) en dislocation à droite de notre échantillon semblent reprendre un référent déjà bien établi dans le discours et saillant (apparaissant souvent, avant cette reprise, en position thématique ; cf. Ariel, 1988) afin de maintenir sa continuité. C'est bien ce qui se passe en 38 et 39⁴³ :

38) L1 : je lui ai raconté l'histoire [du chat]_i [du petit chat qu'on a failli avoir]_i

L2 : l'histoire [du chat]_i ah

L1 : hm [que finalement nous n'avons pas eu]_i euh de quoi on a parlé

L2 : et d'ailleurs [il]_i devenu quoi [*ce petit chat*]_i CEFC

39) L1 : oh j'ai lu un truc qui m'a choquée quoi dans un livre c'est un roman c'est euh Anna Karénine Tolstoï tu vois c'est russe ça se situe au dix-neuvième siècle et euh c'est euh euh [un jeune un jeune homme là riche qui monte euh qui aime faire des courses à cheval]_i

L2 : ouais

L1 : donc euh là [il]_i est en train de monter et de faire une course d'obstacles et [il]_i est mal retombé sur son cheval et [il]_i lui a brisé les reins CEFC

...

L2 : mais [il]_i pesait combien [*ce cet homme*]_i CEFC

En 38, le référent anaphorique est identifié par la mémoire commune des interlocuteurs et, au moment de sa reprise par le SN anaphorique *ce petit chat*, il occupe déjà une position importante dans le discours. En 39, le référent est rendu identifiable par le discours. Dans les deux cas, il bénéficie d'un statut important dans le discours, où il a déjà été mentionné plusieurs fois avant d'être repris par le SN anaphorique disloqué. Après avoir soigneusement examiné, une après une, toutes les occurrences anaphoriques en dislocation à droite avec le contexte large dans lequel elles apparaissent, nous constatons qu'au sein de ces 104 occurrences, le référent de 49 d'entre elles (47,12 %) est identifié par le discours⁴⁴, contrairement à 55 occurrences (52,88 %), où son identification a sa source en dehors du discours, à savoir dans la mémoire commune des interlocuteurs (il est donc question de reprises anaphoriques de référents dont l'identification n'a pas, pour autant, ses origines dans le

⁴³ Ainsi, c'est l'information apportée par la prédication où figure le constituant disloqué qui est davantage mise en valeur, car présentée comme nouvelle (rhématique) ; voir la Section 2.2.3.

⁴⁴ Font évidemment partie de ce sous-ensemble tous les emplois anaphoriques résomptifs et indirects, mais aussi quelques reprises de référents anaphoriques associatifs.

discours même, soit d'un référent dit « nouveau » par Laury, 1997). Ce facteur pourrait également contribuer à l'explication de la haute fréquence d'occurrences anaphoriques dans cette position syntaxique : bien que nous soyons face à des emplois anaphoriques, le référent possède un degré d'accessibilité et de proéminence plus élevé, car il ne se recrute pas à partir des informations discursives⁴⁵

- 40) L1 : [NNAAMMEE]_i va partir une année en Erasmus
 je sais pas si tu vois qui [c']_iest [NNAAMMEE]_i
 L2 : très belle très intelligente
 L1 : oh là là [elle]_i me désespère [*cette fille*]_i

Quant aux référents exophoriques (a-mémoriels et mémoriels confondus) faisant l'objet d'une dislocation, ceux-ci sont deux fois plus nombreux à apparaître dans une dislocation à droite (30 occurrences) que dans une dislocation à gauche (15 occurrences).

Ensuite, concernant la dislocation à gauche, 31 des 48 référents (soit 64,58 %) se prêtent à un emploi anaphorique. Parfois, il s'avère évident que lorsqu'il apparaît dans cette position, le référent est traité par le locuteur comme moins accessible et moins proéminent⁴⁶ que s'il apparaissait en dislocation à droite, et que cette position permet justement d'en renforcer la proéminence, tout en pouvant, éventuellement, faire démarrer une nouvelle chaîne référentielle (voir la Section 2.2.3). Cela est en général vrai pour les occurrences exophoriques et aussi pour une partie des occurrences anaphoriques :

- 41) mais [*cette carte-ci*]_i moi je [la]_i trouve pas très attends ici CEFC
- 42) L1 : oh la la [*ce fromage à la fraise*]_i là [c']_iétait immonde
 L2 : une horreur Ficello à la fraise
 L1 : quoi ah mais [c']_iétait même pas [c']_iétait des petites boules
 L2 : ouais mais [ça]_i avait la même texture que les Ficello
 L1 : ouais comme la Vache Qui Rit le Ficello c'est comme c'est comme du Babybel CEFC
- 43) L1 : il nous disait hier que par rapport au pointeur aux collections et aux fichiers qu'il y a chaque fois [des pavés de cours]_i mais que [*ces pavés*]_i il nous [les]_i donne pas
 L2 : oui ouais CEFC
- 44) L1 : et là il y a [une une jeune flic une jeune policière qui euh qui veut faire ses preuves]_i tout ça machin et euh il y a une affaire d'un tueur en série qui euh qui qui dépèce ses victimes en fait qui leur qui leur enlève la peau

⁴⁵ Cette explication improvisée semble pourtant contredire l'approche d'Ariel (1988), exposée dans la Section 1.1.3, car selon cette dernière, les référents identifiés par la mémoire se distingueraient justement par une accessibilité plutôt basse, dans la mesure où ils se recrutent à partir d'informations stockées dans la mémoire longue. Peut-être est-ce leur reprise anaphorique qui fait ici toute la différence.

⁴⁶ Comme nous le verrons dans la Section 6.2.1.2, la proéminence est un concept complexe, déterminé par plusieurs facteurs, dont le statut ontologique du référent, sa fréquence d'apparition dans le discours, la position syntaxique dans laquelle le référent apparaît et son statut thématique-rhématique.

L2 : ah quelle horreur

L1 : et euh et donc en fait [*cette jeune cette jeune policière*]_i là [*elle*]_i va se rendre compte que Hannibal Lecter il sait qui c'est

L2 : d'accord

L1 : alors [*elle*]_i va aller le voir et ils vont résoudre l'affaire tous les deux en fait bon lui dans sa cellule et [*elle*]_i à l'extérieur et hum et en fait il accepte de [*lui*]_i donner des renseignements sur euh sur ce sur ce tueur en série si [*elle*]_i en échange [*elle*]_i répond à des questions que lui va [*lui*]_i poser sur sa vie privée à [*elle*]_i CEFC

L'exemple 41 ressortit à la deixis situationnelle, l'exemple 42 appartient à l'exophore mémorielle et les exemples 43 et 44, eux, relèvent de l'anaphore. Nous constatons que dans les quatre cas, le référent voit sa prééminence augmenter, pouvant jusqu'à se trouver investi du rôle de référent prééminent⁴⁷ d'une chaîne de référence importante, comme nous le voyons en 42 et 44.

Pourtant, nous ne sommes ici que face à une tendance et non pas à une règle absolue, car apparaissent également dans cette position quelques référents bien établis et prééminents, et dont la prééminence est instaurée (bien) avant la reprise en dislocation à gauche :

45) L1 : [Berlusconi]_i s'il te plaît tu sais [Berlusconi]_i je demandais à ma à ma colocataire euh parce que ça m'étonne quand même qu'[il]_i soit réélu tu vois et ma colocataire italienne je lui demandais mais mais quand même enfin [Berlusconi]_i euh

L2 : hein non mais

L1 : disons les choses comme [c']_iest [*ce mec*]_i [c']_iest quand même une blague

L2 : ah oui c'est pas possible d'avoir un président comme ça

L1 : et elle elle me regarde elle fait oui je sais bien d'ailleurs tout le monde que je connais le pense mais

L2 : mais [il]_i est réélu

L1 : [il]_i est réélu et je comprends pas pourquoi [il]_i est réélu parce que tout le monde [le]_i déteste CEFC

46) oh évidemment [la voiture]_i je [l']_iai encore amenée chez le garagiste

...

il va falloir que je [l']_iamène aussi faire des révisions parce que là [*elle*]_i [*elle*]_i fait un drôle de clac

...

moi [la mienne]_i [*elle*]_i est passée sur le marbre hein

...

et [*elle*]_i est passée sur le marbre parce que [*cette voiture*]_i vraiment [*elle*]_i est super

CEFC

En 45, la dislocation à gauche d'un référent déjà assez prééminent (*Berlusconi*) pourrait s'expliquer par la recatégorisation de ce dernier via un autre nom (*mec*) et par sa qualification via le prédicat *être une blague*. Cette explication tient aussi en partie pour 46, où le référent disloqué est qualifié de *super*. Ce référent est très prééminent dans le récit fait par la locutrice, sachant que parfois, des pans de discours assez larges séparent ses nombreuses reprises, le plus souvent réalisées au moyen de pronoms personnels, mais aussi par un SN possessif (*la mienne*).

⁴⁷ De par sa position thématique, entre autres.

En somme, les référents de nos occurrences anaphoriques en dislocation à gauche sont un peu plus souvent identifiables par le discours (19 cas) que via une connaissance extra-discursive, mémorielle (12 cas). Pour conclure au sujet de la dislocation, si des tendances conformes aux constats du Chapitre 2 (Section 2.2.3) sont présentes dans notre échantillon, il reste pourtant difficile d'appréhender la différence entre les deux types de dislocation en des termes catégoriques et autrement qu'à travers la notion de continuum.

POSITION PREVERBALE		
Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
EXOPHORE A-MEMORIELLE	10	2,57 %
Deixis situationnelle	10	10,75 %
ENDOPHORE	8	1,64 %
Anaphore fidèle	5	3,52 %
Anaphore résomptive	2	1,18 %
Anaphore infidèle	1	0,82 %
OCCURRENCES AMBIGUES	1	2,50 %
TOTAL	19	1,90 %

Tableau 5.14 : Rapport entre la position préverbale et les différents types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 19 occurrences)

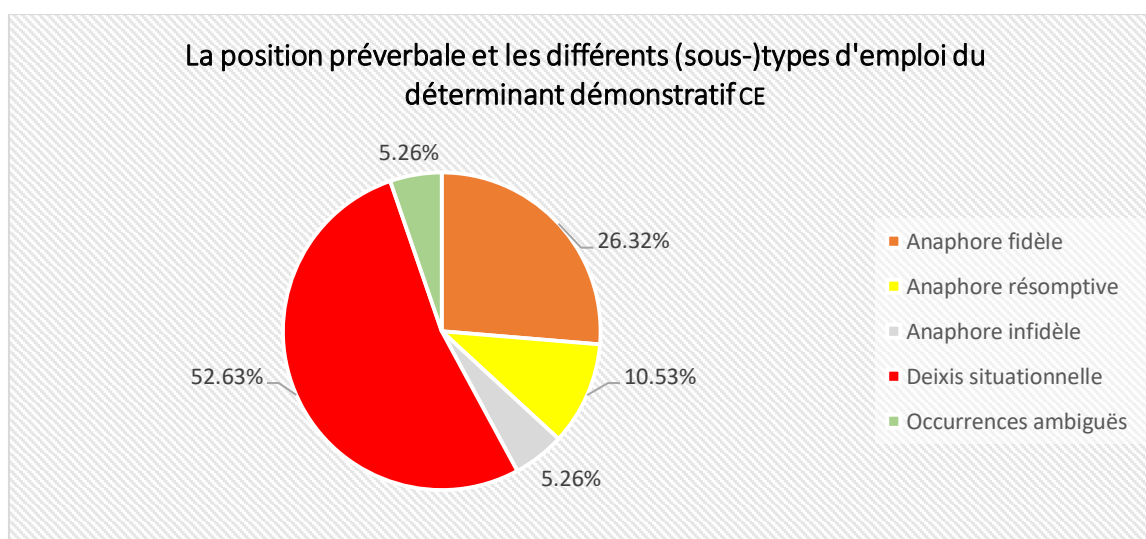


Figure 5.11 : Représentation en % des différents (sous-)types d'emploi du déterminant démonstratif CE au sein de la position préverbale dans l'échantillon de conversations du CEFC (Orféo ; 19 occurrences)

La faible représentation de **la position préverbale** est une chose plutôt étonnante⁴⁸. Cette position ne concerne que 19 occurrences, dont 10 sont déictiques situationnelles et 8 sont endophoriques. Le fait que les occurrences soient partagées entre ces deux emplois pourraient partiellement s'expliquer à travers le constat d'Ariel (1988) vu dans la Section 1.1.3, à savoir que les

⁴⁸ Une fois de plus, il ne faut jamais perdre de vue la taille limitée de l'échantillon.

référents se recrutant au sein de ces deux cadres tendent à afficher un degré d'accessibilité plus élevé que les référents mémoriels, ce qui pourrait favoriser leur apparition en position préverbale, donc thématique.

Concernant les emplois indéfinis et « émotionnels », leur faible représentation au sein de toutes les positions rend impossible toute généralisation à leur sujet. Quant aux occurrences ambiguës, la majorité d'entre elles (50 %) adoptent la position postverbale⁴⁹.

Nous allons clore cette section par la Figure 5.12, dont le graphique affiche notamment les **affinités** susmentionnées, à savoir : 1) celle **entre la position postverbale et l'exophore mémorielle**, 2) celle **entre la position de complément de phrase et les emplois déictique et anaphorique temporelles**, 3) celle **entre la position en dislocation à droite et l'endophore** (représentée notamment par les anaphores infidèle et fidèle) et, enfin, 4) celle entre la position préverbale et la deixis situationnelle (cette dernière affinité doit cependant être approchée avec circonspection, du fait de la très faible représentation de la position préverbale). Pour une meilleure lisibilité du graphique, certains (sous-)types d'emploi qui ne se démarquant par aucun indice de spécificité saillant ne sont pas représentés :

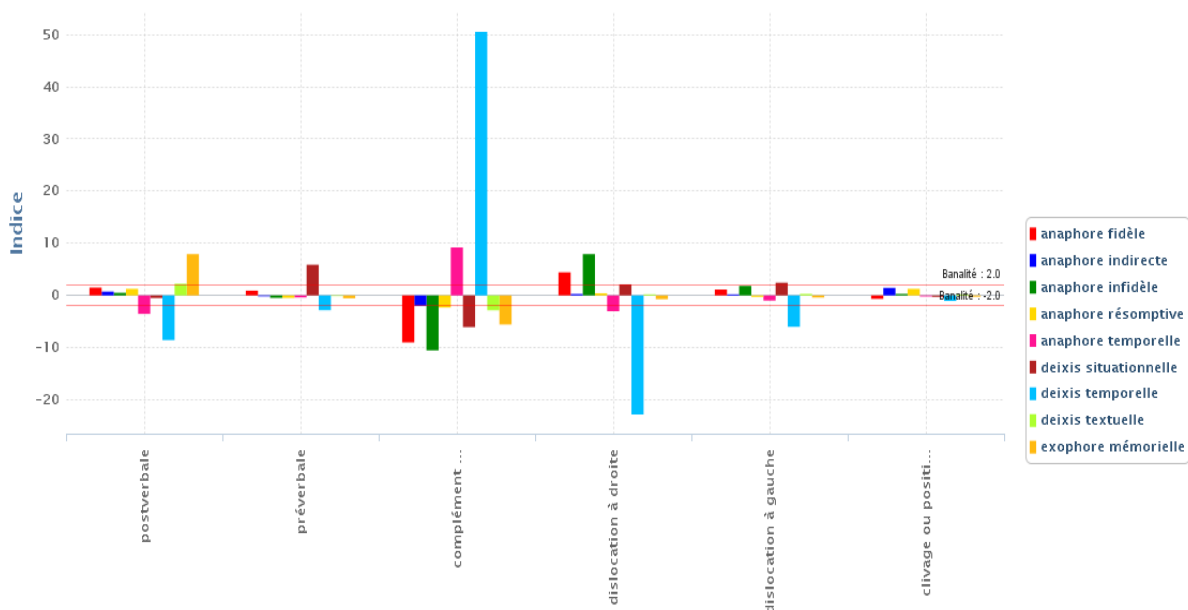


Figure 5.12 : Calcul de spécificités appliqué au lien entre les différentes positions de l'occurrence du déterminant démonstratif CE par rapport au verbe de la proposition et les différents (sous-)types d'emploi (hormis les emplois ambigus, « émotionnels », cataphoriques et indéfinis) dans l'échantillon de conversations du CEFC (Orféo ; 938 occurrences)

⁴⁹ Le statut de ces occurrences nous empêche de les analyser plus avant.

5.2.2.3 Les disfluences dans les occurrences de CE

Les Tableaux 5.15 et 5.16 renseignant sur la présence de disfluences (amorces inachevées, répétitions, marques d'hésitation, pauses entre le démonstratif et le nom, interruptions ou les combinaisons de ces phénomènes) avec les différents types et sous-types d'emploi ainsi qu'avec les différentes positions que le SN démonstratif occupe par rapport au verbe de la proposition (le point de vue adopté est celui des (sous-)types d'emploi et des positions par rapport au verbe). Les graphiques des Figures 5.13 et 5.14 représentent, quant à eux, les parts que les (sous-)types d'emploi et les différentes positions par rapport au verbe occupent au sein de l'ensemble des occurrences contenant des disfluences (ces dernières étant ainsi le point de départ de l'analyse) :

Type et sous-type d'emploi	Nombre d'occurrences avec des disfluences	Par rapport à l'ensemble des occurrences du (sous-) type en question (%)
ENDOPHORE	38	7,79 %
Anaphore résomptive	15	8,82 %
Anaphore fidèle	12	8,45 %
Anaphore infidèle	7	6,25 %
Anaphore temporelle	2	4,55 %
Anaphore indirecte	2	10,53 %
EXOPHORE A-MEMORIELLE	11	2,83 %
Deixis temporelle	6	2,13 %
Deixis situationnelle	5	5,38 %
OCCURRENCES AMBIGUES	11	27,50 %
EXOPHORE MEMORIELLE	7	11,29 %
EMPLOIS « EMOTIONNELS »	2	18,18 %
TOTAL	69	6,90 %

Tableau 5.15 : Proportion des occurrences du déterminant démonstratif CE affichant des disfluences pour chaque type et sous-type d'emploi dans l'échantillon de conversations du CEFC (Orfeo ; 69 occurrences)

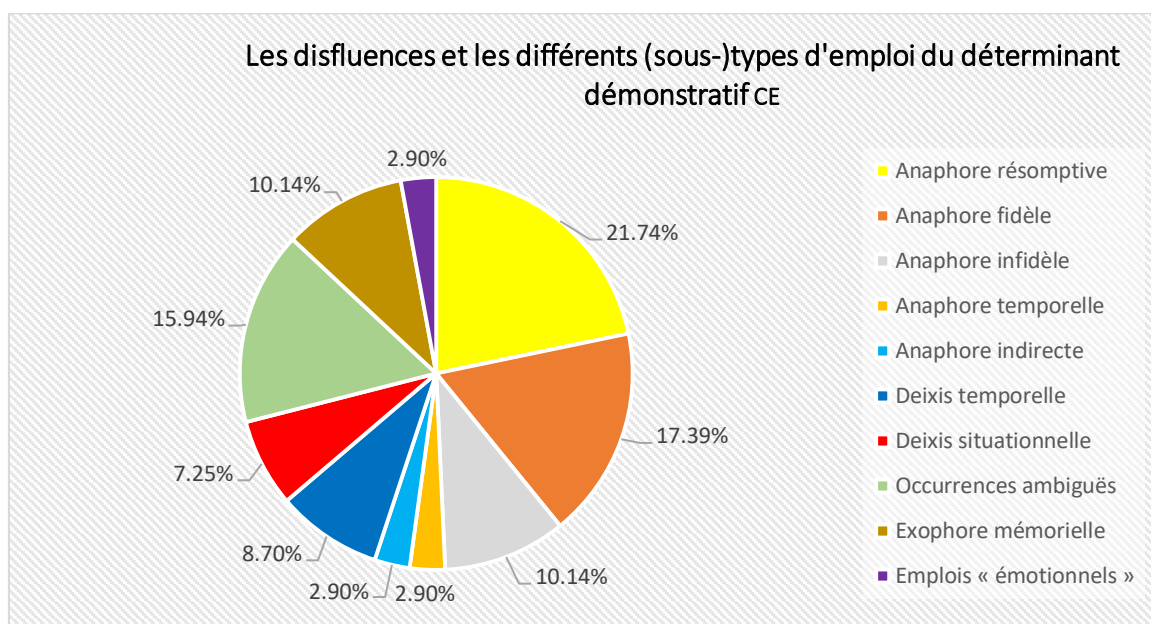


Figure 5.13 : Représentation en % des différents (sous-)types d'emploi du déterminant démonstratif CE parmi les occurrences affichant des disfluences dans l'échantillon de conversations du CEFC (Orféo ; 69 occurrences)

Position	Nombre d'occurrences avec des disfluences	Par rapport à l'ensemble des occurrences de la position en question (%)
POSITION POSTVERBALE	28	7,20 %
COMPLEMENT DE PHRASE	14	3,93 %
POSITION INCONNUE	10	14,29 %
DISLOCATION A DROITE	8	5,56 %
DISLOCATION A GAUCHE	7	14,58 %
POSITION PREVERBALE	2	10,52 %
TOTAL	69	6,90 %

Tableau 5.16 : Proportion des occurrences du déterminant démonstratif CE affichant des disfluences pour chaque position de l'occurrence par rapport au verbe dans l'échantillon de conversations du CEFC (Orféo ; 69 occurrences)

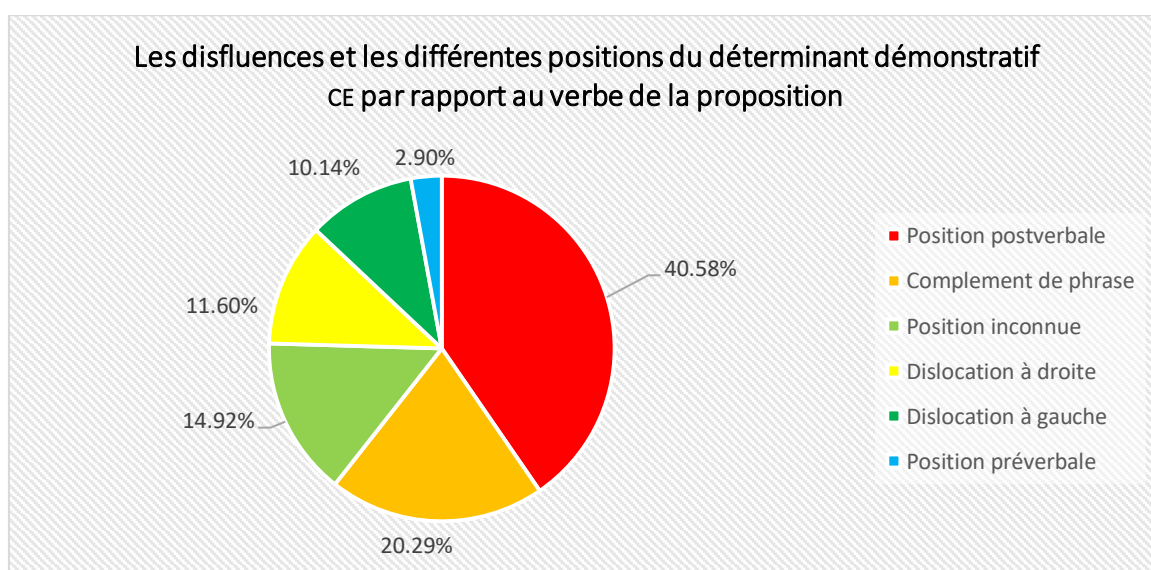


Figure 5.14 : Représentation en % des différentes positions du déterminant démonstratif CE par rapport au verbe parmi les occurrences affichant des disfluences dans l'échantillon de conversations du CEFC (Orféo ; 69 occurrences)

Nous relevons, avant tout, les faits suivants : 1) 6,9 % de toutes les occurrences présentent des disfluences. Ces occurrences sont le plus souvent de nature anaphorique résomptive (21,74 % de l'ensemble) et, de manière générale, bien plus souvent de nature anaphorique (55,07 % de l'ensemble) que de nature exophorique (26,09 % de l'ensemble)⁵⁰. 15,94 % de toutes les occurrences contenant des disfluences relèvent de l'emploi ambigu. Concernant l'emploi anaphorique, la présence de disfluences peut en partie s'expliquer par une certaine hésitation lors de la saisie du référent, hésitation pouvant être entraînée, entre autres, soit par la recatégorisation (le cas de l'anaphore infidèle – voir l'exemple 46 *infra*) ou l'institution dans le discours de ce dernier (le cas de l'anaphore résomptive – voir l'exemple 47 *infra*). Si 2 emplois « émotionnels » sont également accompagnés de disfluences, ce nombre, trop bas, rend toute généralisation non pertinente. Comme nous pouvons le constater grâce au graphique de la Figure 5.15, le calcul de spécificités montre **une affinité entre la présence de disfluences et les occurrences ambiguës** (quatrième colonne) ; en revanche, les occurrences exophoriques a-mémorielles (deuxième colonne) « évitent » les disfluences, la spécificité étant ici très basse :

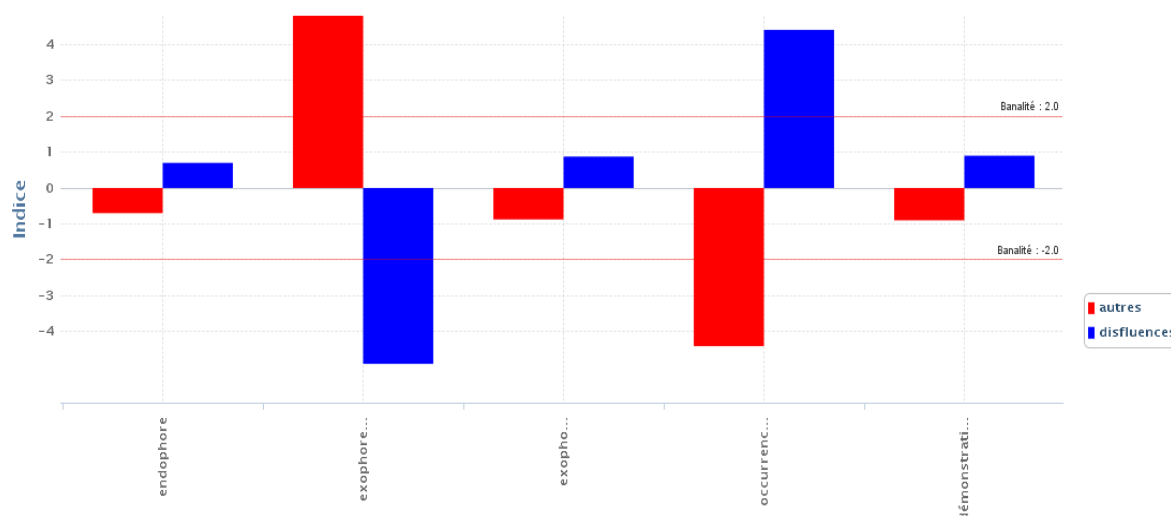


Figure 5.15 : Calcul de spécificités appliqué au lien entre les types d'emploi du déterminant démonstratif CE et la présence de disfluences dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)

Pour ce qui est de **la position par rapport au verbe**, ce sont la position postverbale et celle du complément de phrase qui accueillent le plus d'occurrences contenant des disfluences, soit, respectivement, 40,58 % et 20,29 % de l'ensemble. Cette fréquence n'est pas non plus négligeable pour la position syntaxique inconnue (14,92 % de l'ensemble), ce qui ne se présente pas comme

⁵⁰ Cette différence est statistiquement significative, car si nous appliquons la fonction de vraisemblance (fréquence absolue des emplois endophoriques contenant des disfluences = 38 ; fréquence absolue des emplois exophoriques contenant des disfluences = 18 ; la taille de la population (n) = 69), le score est égal à 17,0209 (la valeur critique = 3,841 (degré de liberté = 1 et α = 0,05)).

quelque chose de surprenant. En effet, il n'est nullement contre-intuitif que le statut syntaxique flou de ce type d'occurrences s'accompagne de marques de rupture aux niveaux phonétique, lexical et syntaxique⁵¹. Quant à la dislocation à gauche (10,14 % de l'ensemble apparaissent avec des disfluences), les disfluences pourraient trouver une explication dans le statut encore fragile, pas suffisamment établi, du référent dans le discours⁵². Dans ce qui suit, nous entreprendrons une étude plus détaillée des disfluences.

À regarder de plus près les occurrences apparaissant avec les disfluences, nous pouvons discerner plusieurs cas de figure grossièrement délimités, que nous croyons être au nombre de 4. Le cas de figure de loin le plus représenté (56 occurrences, soit 81,16 % de la totalité) est celui où les disfluences apparaissent dans **un contexte d'hésitation et/ou de légères reformulations du propos** :

47) ils se sont cachés dans [les montagnes qu'il y a vers euh Santiago de Cuba à l'Est de Santiago]; je sais plus comment [elles]; s'appellent putain mais [elles]; sont magnifiques c'est [là]; où ils ont fait les premiers plants de café c'est le café c'est les français qui ont été chassés de Haïti tu sais quand Haïti a acquis son indépendance mh mh attends c' est pas son indépendance enfin il y a eu un moment où ils ont été chassés par euh par un un nationaliste haïtien tous les colonialistes français et ils ont fui ils ont atterri à Cuba à l'Est donc vers Santiago de Cuba et euh c'est eux qui ont fait les premières plantations de café dans [*ces dans ces massifs de montagnes*]; CEFC

48) à quoi ça te sert de raisonner les gens juste euh juste te dire pourquoi tu pourquoi s~ tu en es arrivé enfin pourquoi il y a cette situation et puis euh et puis en finir avec *cette euh cette atmosphère un peu malsaine* quoi qui règne CEFC

49) mais c'est vrai que *le gâ~ ce petit gâteau au chocolat qu'ils ont là* CEFC

En 47, 48 et 49, nous voyons que les locuteurs cherchent à formuler leur propos de la meilleure façon possible, ce qui s'accompagne d'une hésitation et d'une répétition du démonstratif en 47 et 48 et d'une amorce inachevée en 49. Dans 7 de ces occurrences, la recherche du meilleur moyen d'expression a débouché sur une aposiopèse :

50) jusqu'au bord de la plage oh je te jure c'était trop impressionnant il y avait le bateau il faisait *de ces euh* CEFC

Dans un cas, l'hésitation semble explicitement traduire non seulement une quête de moyens d'expression appropriés, mais aussi la quête du bon référent⁵³:

⁵¹ Notons que cette affinité fonctionne également dans l'autre sens, à savoir en partant de la position : 14,29 % de toutes les occurrences en position inconnue présentent des disfluences.

⁵² Cependant, l'application du calcul de spécificités ne démontre aucune affinité entre une position en particulier et la présence de disfluences.

⁵³ Nous ne relevons là qu'un cas où cette quête du bon référent nous semble très explicite. Mais cela ne veut pas dire qu'elle soit absente des autres emplois accompagnés de disfluences, car souvent, recherche de moyens d'expression appropriés et du bon référent semblent aller de pair, au point qu'il est impossible de les séparer.

- 51) alors comment ils vont *euh ces ces ces jeunes euh NNAAMMEE NNAAMMEE* ça va mh mh il est toujours à Paris CEFC

Les autres cas de figure sont clairement minoritaires en comparaison avec celui qui vient d'être décrit. Dans le cas de 8 occurrences (soit 11,59 %), le locuteur, après s'être embarqué dans la formulation de son propos, fait marche arrière et en vient à le reformuler de manière plus radicale :

- 52) L1 : alors bon l'école mistralienne et l'école traditionnelle
L2: ben ils écrivent pas de la même façon en fait *ces deux gra~* ça correspond à deux graphies différentes pour dire la même chose CEFC

- 53) comment tu peux ne pas aimer le fromage moi je ne comprends pas *cette euh* je peux je peux pas comprendre le fromage c'est ma vie à moi CEFC

Trois occurrences (4,35 %) sont celles où la disfluence est due au fait que le locuteur tente de prendre la parole au moment où parle son interlocuteur, mais, ne réussissant pas à le faire, il finit par se taire :

- 54) L1 : puis à côté à côté on aura aussi tout ce qui est euh ben alcool pour la soirée donc faut aussi voir le vin quoi et ça à la limite parce que moi le plus simple c'est que eux ils ramènent l'alcool pour la soirée moi je vois le reste le vin et voilà
L2 : *ben dans ce cas-là dans ce cas-là* CEFC

En 54, le locuteur n°2 essaie d'imposer sa parole au moment où le locuteur n°1 parle, mais n'y arrive pas⁵⁴.

Dans deux cas (2,90 %), c'est le locuteur lui-même qui se fait interrompre par son interlocuteur :

- 55) L1 : faut voir ce qu'on préfère dans *ces*
L2 : moi j'ai dit je voulais pas je sais pas si tu as été regarder par curiosité CEFC

Toutes les occurrences du déterminant démonstratif présentant des disfluences et placées soit en dislocation à gauche soit en dislocation à droite relèvent du premier cas de figure (contexte d'hésitation et/ou de légères reformulations du propos) :

- 56) [*cette cette querelle*]_i [*elle*]_i [*elle*]_i fait plus de mal que CEFC

- 57) tu sais [*cette cette proximité entre ce mec qui a qui a deux deux facettes comme ça*]_i la folie et l'intelligence rien que [*ça*]_i déjà [*c'*]_iest angoissant CEFC

À la fois en 56 et en 57, les répétitions semblent traduire une incertitude de la locutrice dans la saisie d'un référent (de nature anaphorique résomptive) encore relativement instable car non

⁵⁴ En réalité, la prise de parole du locuteur n°2 est simultanée à celle du locuteur n°1, contrairement à ce que laisse croire la transcription linéaire de la conversation.

suffisamment établi dans le discours⁵⁵. Pour 5 des 7 occurrences en dislocation à gauche accompagnées de disfluences, l'identification du référent repose sur le discours : il s'agit ou bien de cas d'anaphore résomptive ou bien de cas d'anaphore fidèle ou infidèle dont le référent est identifié via un antécédent indéfini situé non loin de l'anaphorique⁵⁶.

Parmi les 8 occurrences placées en dislocation à droite qui sont accompagnées de disfluences, 5 contiennent un référent identifiable via le discours (cf. 39 et 58) et 3 contiennent un référent identifiable par la mémoire des interlocuteurs (cf. 51 et 59) :

- 58) la plage elle était nickel ah il y avait pratiquement personne la plage elle était super belle [l'eau];
 bon il y avait des algues hein dans [l'eau]; mais
 ...
 c' est pour ça moi j'aimais pas trop euh j'aimais pas trop [y]_i aller [*dans cette euh eau*]_i
 CEFC
- 59) bon puis # alors on va [y]_i arriver [*à ce ce cher*]_i CEFC

Néanmoins, contrairement à la dislocation à gauche, l'argument du statut insuffisamment établi du référent est ici plus difficilement acceptable, vu que la dislocation à droite implique, de manière générale, un statut plus thématique (saillant) et plus discursivement établi de ce dernier (voir la Section 5.2.2.2 *supra* et la Section 2.2.3). D'autres explications pourraient éventuellement être proposées au cas par cas. En 39, par exemple, la dislocation apparaît dans la séquence de la locutrice n°2, alors que toutes les reprises du référent sont, jusque-là, apparues chez la locutrice n°1 : ainsi, l'hésitation de la locutrice n°2 pourrait se justifier par le fait que pour elle, le référent serait bien moins établi dans le discours que pour sa collègue. En 59, l'hésitation semble refléter l'attitude perplexe du locuteur parlant de son fils en bas âge, dont il trouve problématique le comportement.

Regardons maintenant comment la situation se présente dans notre échantillon de conversations tchèques d'ORTOFON v1 pour poursuivre notre comparaison du comportement du démonstratif adnominal dans l'oral informel des deux langues.

⁵⁵ Voir aussi l'exemple 44 *supra*.

⁵⁶ Dans un cas, le référent est identifié par la situation, et dans un autre, la voie d'identification du référent est ambiguë.

5.2.3 ORTOFON v1

5.2.3.1 Premiers constats

Commençons avec un tableau synthétique des types et sous-types d'emploi de l'échantillon de 1 000 occurrences du démonstratif adnominal tchèque⁵⁷. Précision importante : contrairement au français, qui ne dispose que d'un seul lemme – CE –, le tchèque, comme nous l'avons vu dans le Chapitre 3, dispose de plusieurs lemmes : TEN, TENTO, TENHLE, TENHLETEN, TADYTEN, TAMTEN, ONEN, etc⁵⁸. Par conséquent, tous ces lemmes étaient visés par la requête et ils sont ou peuvent ainsi être inclus dans l'échantillon et le Tableau 5.17. De même que pour le corpus de conversations françaises, nous donnons également l'intervalle de confiance, anticipant la fréquence du (sous-)type d'emploi en question dans un autre échantillon de 1 000 occurrences prélevées sur la liste des concordances originelle⁵⁹, et renseignant ainsi sur sa place dans l'ensemble de cette liste. Comme taille de référence de cette liste (soit la population totale), nous avons utilisé la fréquence moyenne du démonstratif adnominal telle qu'elle a été calculée par l'interface « 1 feature – many samples », soit 20 311 occurrences (voir la Section 5.2.1 *supra*). Le graphique de la Figure 5.16 représente les proportions occupées, au sein de l'échantillon, par les différents (sous-)types d'emploi⁶⁰ :

Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des 1 000 occurrences (%)	Intervalle de confiance	
			Borne inférieure	Borne supérieure
EMPLOIS « EMOTIONNELS »	331	33,10 %	302	361
EXOPHORE MEMORIELLE	285	28,50 %	257	313
ENDOPHORE	216	21,60 %	191	242
Cataphore	63	6,30 %	48	78
Anaphore fidèle	58	5,80 %	44	73
Anaphore associative	54	5,40 %	41	68
Anaphore infidèle	15	1,50 %	8	23
Anaphore indirecte	13	1,30 %	6	20
Anaphore résomptive	12	1,20 %	6	19
Anaphore temporelle	1	0,10 %	0	3
EXOPHORE A-MEMORIELLE	69	6,90 %	54	85
Deixis situationnelle	62	6,20 %	47	77
Deixis temporelle	6	0,60 %	2	11
Deixis textuelle	1	0,10 %	0	3
OCCURRENCES AMBIGUES	81	8,10 %	65	98

⁵⁷ Nous invitons le lecteur à consulter les Sections 3.4, 3.4 et 3.5 du Chapitre 3 pour des exemples de ces différents (sous-)types d'emploi.

⁵⁸ Rappelons que si seul TEN peut être touché par la grammaticalisation, c'est l'ensemble des formes qui ont été étudiées, excepté dans la Section 5.2.3.2.

⁵⁹ En écartant le bruit ; $\alpha = 0,05$.

⁶⁰ Il reprend les types d'emploi là où ceux-ci ne sont pas divisés en sous-types ; dans le cas contraire, il reprend les sous-types d'emploi.

EMPLOIS « METALINGUISTIQUES »	14	1,40 %	7	22
DEMONSTRATIFS INDEFINIS	4	0,40 %	1	8
TOTAL	1 000	100 %		

Tableau 5.17 : Types et sous-types d'emploi du démonstratif adnominal tchèque (tous lemmes confondus) dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences) avec les intervalles de confiance anticipant leur fréquence dans un autre échantillon de la même taille

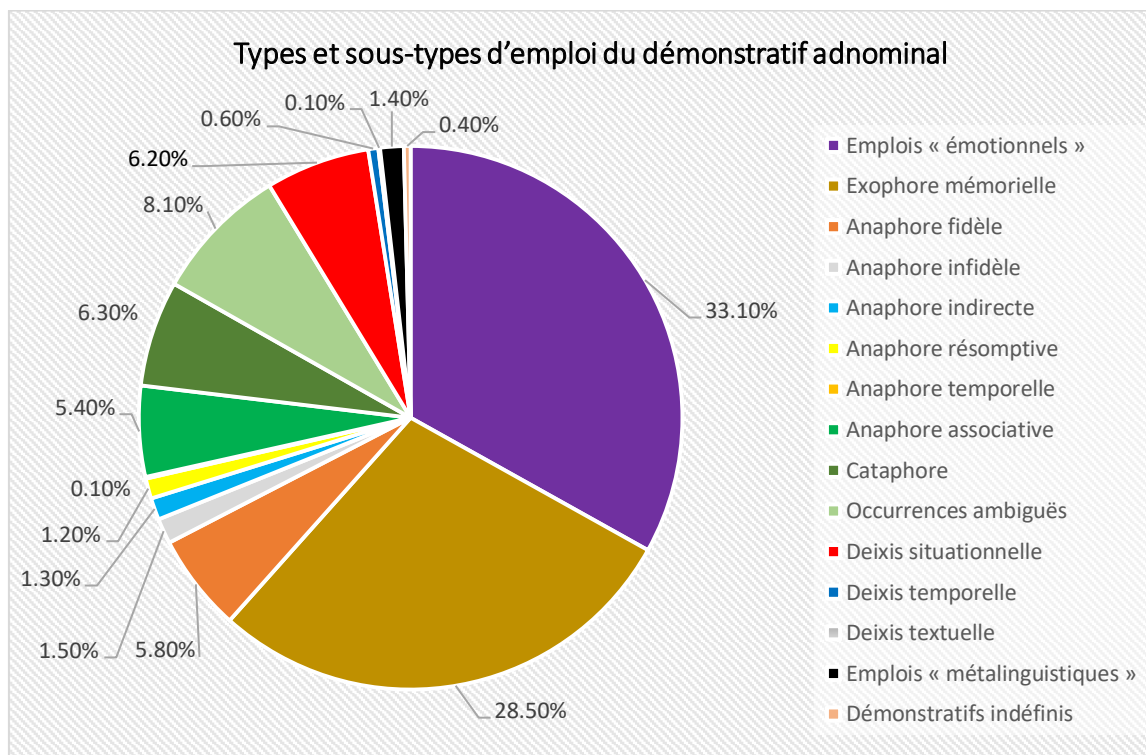


Figure 5.16 : Types et sous-types d'emploi du démonstratif adnominal (tous lemmes confondus) dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

En ce qui concerne les fréquences des différents types et sous-types d'emploi, la situation est fort différente de celle du français. Contrairement au corpus de conversations françaises, où ce sont les emplois endophoriques (489 occurrences, soit 48,9 % de l'ensemble) et exophoriques a-mémoriels (389 occurrences, soit 38,9 % de l'ensemble) qui priment (voir la Section 5.2.2.1), l'échantillon tchèque est essentiellement composé d'emplois « émotionnels » (331 occurrences, soit 33,1 % de l'ensemble) et mémoriels (285 occurrences, soit 28,5 % de l'ensemble), les emplois endophoriques n'arrivant qu'en troisième position (216 occurrences, soit 21,6 % de l'ensemble). Cette dernière catégorie n'affiche pas la même composition en tchèque, où les occurrences cataphoriques (63, soit 29,17 % de toutes les occurrences endophoriques) et anaphoriques associatives (54, soit 25 % de toutes les occurrences endophoriques) contribuent considérablement au chiffre final. Tandis que les secondes ne se trouvent pas du tout en français, les premières n'y sont qu'extrêmement peu représentées (2 occurrences). En tchèque, le chiffre élevé d'occurrences cataphoriques est presque exclusivement dû à la présence de la cataphore intra-syntagmatique (62 des 63 occurrences, soit 98,41%). D'autre part, l'échantillon tchèque nous semble contenir une seule occurrence d'anaphore

temporelle, contrairement au français qui en comporte 44. Le français contient également beaucoup plus d'anaphores résomptives (170 contre seulement 12 en tchèque, soit 5,56 % de toutes les occurrences endophoriques) et d'anaphores infidèles (112 contre 15 en tchèque, soit 6,94 % de toutes les occurrences endophoriques). De manière générale, nous constatons qu'à l'exception de l'anaphore associative et de la cataphore, tous les sous-types d'emploi endophorique sont davantage représentés en français⁶¹. L'essentiel de ce qui vient d'être dit atteste, sur le plan quantitatif, les deux tendances relatives au registre parlé informel du tchèque mentionnées dans l'Introduction générale, à savoir : 1) la grammaticalisation de TEN adnominal (se manifestant, avant tout, à travers sa haute fréquence en emplois mémoriel, cataphorique intra-syntagmatique et dans l'anaphore associative ; voir le Chapitre 6) et 2) la fréquence abondante de valeurs dites « émotionnelles » avec TEN adnominal (voir le Chapitre 7).

Il nous semble évident grâce à notre analyse des données que la cause de ces divergences est avant tout de nature structurelle, à savoir qu'elle procède des différences entre les deux systèmes linguistiques. Nous notons, tout d'abord, une forte disproportion dans les fréquences des occurrences « émotionnelles » – 331 sur 1 000 en tchèque contre seulement 11 sur 1 000 en français – et mémorielles – 285 sur 1 000 en tchèque contre seulement 62 sur 1 000 en français⁶². En effet, l'emploi « émotionnel » du démonstratif est à la fois beaucoup plus répandu et connaît une variété plus importante de valeurs en tchèque parlé informel (voir à ce sujet les Sections 3.5 et 4.5.4 des Chapitres 3 et 4, mais également l'ensemble du Chapitre 7 à venir). Toujours en tchèque parlé informel, la plus forte présence du démonstratif dans les emplois mémoriels est due au fait que celui-ci y apparaît souvent, avec cette valeur, là où le français utilise le défini (voir à cet égard les Section 3.3.4 et 6.2.1.3, mais aussi la Section 5.2.4.4 *infra*). En revanche, l'exophore a-mémorielle est, avec ses 69 occurrences sur 1 000, assez faiblement représentée dans nos données tchèques (dans nos données françaises, ce type d'emploi compte 389 occurrences sur 1 000)⁶³. Comme cela a déjà été relevé, dans les données françaises, ce chiffre est avant tout dû à une présence abondante de la deixis temporelle : 282 occurrences contre uniquement 6 en tchèque. Ce clivage s'explique, quant à lui, par le fait que le tchèque utilise beaucoup moins le démonstratif avec les expressions déictiques temporelles,

⁶¹ La représentation des emplois anaphorique temporel et déictique textuel est si faible qu'il est impossible de confirmer leur présence dans le reste des données. Rappelons que pour ce qui est de l'anaphore temporelle, sa présence basse en tchèque s'expliquerait en partie par le fait que contrairement au français, le tchèque emploie souvent, dans ces cas, les adverbes déictiques (*pro*)*tentokrát* pour *cette fois*-(*ci*/là) et *tehdy* ou *tenkrát* pour à *cette époque*(-là) / *en ce temps*-là.

⁶² Les deux différences sont significatives sur le plan statistique : si nous appliquons la fonction de vraisemblance (fréquences absolues pour le tchèque = 331 et 285 ; fréquences absolues pour le français = 11 et 62 ; la taille des deux populations (*n*) = 1 000), nous obtenons le score de 439,0086 pour la première et le score de 185,4723 pour la seconde ; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

⁶³ La différence est significative sur le plan statistique : si nous appliquons la fonction de vraisemblance (fréquence absolue pour le tchèque = 69 ; fréquence absolue pour le français = 389 ; la taille des deux populations (*n*) = 1 000), nous obtenons le score de 313,5744 ; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

préférant, dans ce contexte, recourir à des adverbes (voir la Section 3.3.1 ; cette préférence serait en phase avec la nature davantage synthétique du tchèque). Une explication semblable permet de rendre partiellement compte de la disparité dans le domaine de l'anaphore résomptive, à laquelle le français (oral informel) semble recourir davantage que le tchèque pour structurer le discours sur les plans argumentatif et temporel⁶⁴. Toujours en ce qui concerne l'endophore, les emplois anaphoriques associatif et cataphorique intra-syntagmatique sont ceux où le démonstratif est proscrit en français (voir la Section 4.5.2).

Par ailleurs, nous remarquons aussi que le tchèque ne contient qu'une seule occurrence déictique textuelle (sur 1 000), contre 14 en français (sur 1 000). Nous notons également, dans l'échantillon tchèque, la présence de 14 emplois « métalinguistiques » (type d'emploi qui n'existe pas en français) et de 4 démonstratifs indéfinis (contre 9 en français). Enfin, il faut aussi dire que les occurrences du tchèque se sont avérées, de manière générale, plus difficiles à analyser (pour une bonne part d'entre elles, nous avons hésité entre un emploi « émotionnel » et un emploi anaphorique/mémoriel, voir la Section 5.1.4), raison pour laquelle l'échantillon tchèque contient plus de deux fois plus d'occurrences ambiguës que l'échantillon français⁶⁵. Il semblerait par ailleurs que cette plus grande indétermination puisse être le reflet d'un système en mouvement, où les différentes valeurs seraient moins figées car en pleine évolution.

Regardons également **la présence**, au sein de l'échantillon de 1 000 occurrences, **des occurrences du démonstratif adnominal appartenant à d'autres lemmes que le lemme TEN**. Le nombre de ces occurrences s'élève uniquement à 33 (soit 3,3 % de l'ensemble de l'échantillon). Les voici avec, pour chacune, le (sous-)type d'emploi auquel elle se prête :

⁶⁴ Ce n'est sans doute pas la seule raison : la surreprésentation de l'anaphore temporelle, de la deixis temporelle et de l'anaphore résomptive dans l'ensemble de l'échantillon français à l'opposé de l'échantillon tchèque est également due au fait que par comparaison à nos données françaises, la population tchèque de base contient beaucoup plus d'occurrences appartenant aux autres (sous-)types d'emploi. Par conséquent, dans l'échantillon que nous en avons retenu, les trois sous-types en question se font aussi plus rares (car plus « dilués »). La différence entre les fréquences des emplois endophoriques en général ainsi que celle entre les fréquences de la seule anaphore résomptive sont statistiquement significatives : la fonction de vraisemblance affiche le score de 166,5148 pour la première et de 177,6053 pour la seconde (fréquences absolues pour le tchèque = 216 et 12 ; fréquences absolues pour le français = 489 et 170 ; la taille des deux populations (n) = 1 000 ; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$)).

⁶⁵ L'application de la fonction de vraisemblance montre que la différence est statistiquement significative pour les emplois déictiques textuels (13,5317), métalinguistiques (19,5068) et ambigus (15,0661), mais ne l'est pas pour les emplois indéfinis du démonstratif (1,9861) ; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

Lemme	Nombre d'occurrences
TENHLETEN	13
Deixis situationnelle	6
Emplois « émotionnels »	2
Anaphore fidèle	2
Anaphore infidèle	1
Anaphore indirecte	1
Anaphore résomptive	1
TENHLE	11
Deixis temporelle	4
Deixis situationnelle	3
Anaphore infidèle	1
Anaphore résomptive	1
Exophore mémorielle	1
Occurrences ambiguës	1
TENTO	3
Anaphore résomptive	2
Deixis temporelle	1
TADYTEN	2
Deixis situationnelle	2
TADYHLETEN	1
Deixis situationnelle	1
TAMTEN	1
Démonstratif indéfini	1
TUHLETEN	1
Anaphore résomptive	1
TUTEN	1
Deixis temporelle	1
TOTAL	33

Tableau 5.18 : Présence des lemmes du démonstratif tchèque autres que TEN en emploi adnominal dans l'échantillon de 1 000 occurrences d'ORTOFON v1 (CNT) avec leurs sous-types d'emploi

Nous constatons ainsi que le lemme TEN, avec ses 967 occurrences, est de loin le plus représenté dans notre échantillon. Au sein de la minorité que représentent les autres lemmes, ce sont TENHLETEN et TENHLE qui priment⁶⁶. En revanche, d'autres lemmes ne sont qu'extrêmement peu, voire pas du tout présents (dont ONEN). TENHLETEN compte 6 occurrences déictiques situationnelles et 5 occurrences endophoriques. C'est également le seul lemme à part TEN à connaître des emplois « émotionnels », quoique ces derniers ne soient qu'au nombre de 2. TENHLE compte 4 occurrences déictiques temporelles, 3 occurrences déictiques situationnelles, 2 occurrences endophoriques, une

⁶⁶ Le premier se présente deux fois sous la forme *tendleten*, et le second se présente deux fois sous la forme *tendle*.

occurrence exophorique mémorielle et une occurrence ambiguë⁶⁷. La majorité écrasante constituée par TEN est un autre indice quantitatif de la grammaticalisation de ce lexème.

5.2.3.2 La grammaticalisation de TEN adnominal

À présent, regardons précisément ce qu'il en est des occurrences de TEN qui ont été évaluées comme engagées dans un processus de grammaticalisation. Rappelons que le nombre de ces occurrences (451) est comparé au nombre total d'occurrences de TEN, soit 967. Rappelons aussi que les critères exacts sur lesquels a reposé cette évaluation seront présentés de manière détaillée dans le Chapitre 6, où une multitude d'exemples seront également cités et commentés. Le Tableau 5.19 croise le critère de la grammaticalisation et les différents types et sous-types d'emploi en présentant les chiffres dans l'ordre décroissant. Il compare ainsi le nombre d'occurrences évaluées comme étant en voie de grammaticalisation avec le nombre total d'occurrences pour chaque type et sous-type en donnant le pourcentage des occurrences en voie de grammaticalisation par rapport à cette totalité. Le graphique de la Figure 5.17 affiche la part que les différents (sous-)types occupent au sein du groupe des occurrences en voie de grammaticalisation :

GRAMMATICALISATION DE TEN ADNOMINAL			
Type et sous-type d'emploi	Occurrences grammaticalisées	Nombre total d'occurrences du (sous-) type en question	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
EXOPHORE MEMORIELLE	257	284	90,49 %
ENDOPHORE	150	206	72,82 %
Cataphore	63	63	100,00 %
Anaphore associative	54	54	100,00 %
Anaphore fidèle	25	56	44,64 %
Anaphore indirecte	5	12	41,67 %
Anaphore infidèle	3	13	23,08 %
Anaphore résomptive	0	7	0,00 %
Anaphore temporelle	0	1	0,00 %
EXOPHORE A-MÉMORIELLE	44	51	86,27 %
Deixis situationnelle	44	50	88,00 %
Deixis textuelle	0	1	0,00 %
EMPLOIS « EMOTIONNELS »	0	329	0,00 %
DEMONSTRATIFS INDEFINIS	0	3	0,00 %
EMPLOIS « METALINGUISTIQUES »	0	14	0,00 %
OCCURRENCES AMBIGUËS	0	80	0,00 %
TOTAL	451	967	46,64 %

⁶⁷ Vu leur faible présence dans l'échantillon, il n'est malheureusement possible de faire aucune observation fiable au sujet de l'emploi de ces autres lemmes du démonstratif.

Tableau 5.19 : Proportion des occurrences de TEN adnominal évaluées comme engagées dans le processus de grammaticalisation pour chaque type et sous-type d'emploi dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

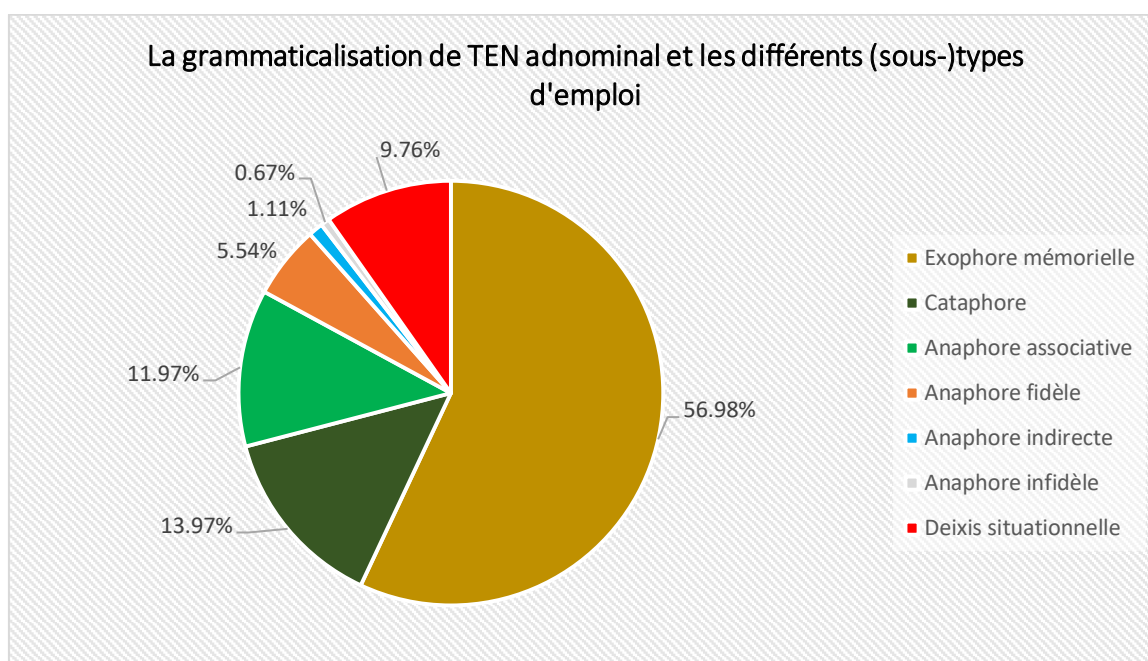


Figure 5.17 : Représentation en % des différents (sous-)types d'emploi parmi les occurrences de TEN adnominal encourageant le processus de grammaticalisation dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Nous constatons que la grammaticalisation est le plus présente avec les occurrences anaphoriques associatives (100 %) et cataphoriques (dont toutes sauf une seule relèvent de la cataphore intra-syntagmatique, soit 98,41 %) ⁶⁸ et mémorielles (90,49 %). Ce sont notamment les deux premières qui font qu'elle est, de manière générale, élevée dans le domaine des emplois endophoriques (72,82 % de toutes les occurrences endophoriques). Suivent tout de suite après les emplois exophoriques a-mémoriels (86,27 % ; grâce à la seule deixis situationnelle, où la proportion atteint 88,00 % de l'ensemble). Par ailleurs, la grammaticalisation est totalement absente des emplois « émotionnels », car dans ce type d'emploi, TEN garde son fonctionnement plein de démonstratif, chose qui rend justement possible l'effet émotionnel produit (voir la Section 3.5). Elle est également incompatible avec les emplois « métalinguistiques » et indéfinis. S'agissant des occurrences ambiguës, nous les avons exclues du calcul faute d'indices suffisants qui nous permettraient de les évaluer relativement à ce critère. Le graphique de la Figure 5.18, affichant les résultats du calcul de spécificités, illustre **les forts liens** qui existent **entre la grammaticalisation de TEN et l'anaphore associative** (cinquième colonne), **la cataphore** (sixième colonne), **les emplois mémoriels** (dixième colonne) et, dans une moindre mesure, les emplois déictiques situationnels (huitième colonne) :

⁶⁸ Une seule occurrence cataphorique nous semble relever d'un emploi associatif :
*[...] studuješ češtinu tak musíš vidět **ten** rozdíl . kyselý . a navinulý [...]*

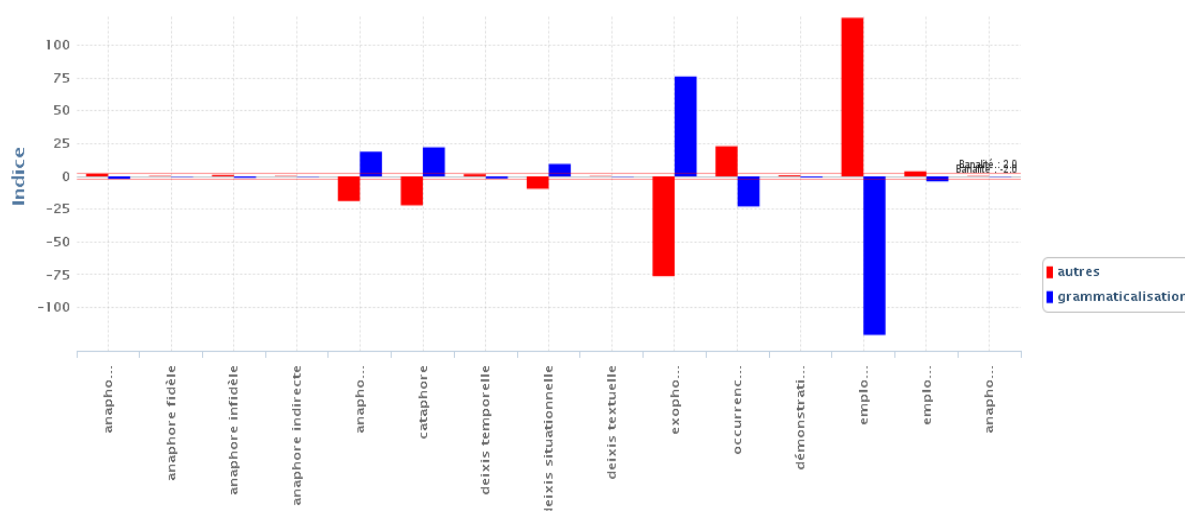


Figure 5.18 : Calcul de spécificités appliqué au lien entre les différents (sous-)types d'emploi et les occurrences de TEN adnominal encourageant le processus de grammaticalisation dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Au total, 46,54 % de toutes les occurrences de TEN adnominal seraient, selon nous, engagées dans un processus de grammaticalisation.

5.2.3.3 Les propriétés combinatoires du démonstratif adnominal avec d'autres mots fonctionnels

Regardons à présent la proportion des occurrences où le démonstratif se combine avec un autre élément fonctionnel (pour la définition du terme, voir notamment la Section 3.1.3). Une précision importante s'impose : nous n'avons pas tenu compte ici des quantificateurs universels VŠECHEN/VEŠKERÝ (*tous/toutes* pour la totalité externe des noms dénombrables pluriels et *tout/toute* pour la totalité externe des noms indénombrables) et CELÝ (*tout/toute* ; totalité interne), car leurs équivalents se combinent, en français, avec les déterminants dits « centraux » (ceux qui occupent la « central determiner slot »), pouvant ainsi fonctionner comme des pré-déterminants (ils se placent à gauche de ces derniers)⁶⁹. De même, nous n'avons pas non plus tenu compte des numéraux (cardinaux et ordinaux) qui suivent le démonstratif⁷⁰ et des équivalents tchèques des lexèmes français MEME (SAMÝ; *ta samá otázka* = *la même question*), AUTRE (DALŠÍ ; *ten další účastník* = *l'autre participant*) et SEUL (JEDINÝ ; *tenhle jediný účastník* = *ce seul participant*), car tous ces éléments se combinent

[...] *tu fais des études de tchèque, du coup, tu dois sentir la différence : acide et acidulé* [...] ORTOFON v1
Soit : la différence entre [les termes d'] « acide » et « acidulé ». Nous reviendrons à cet exemple dans la Section 6.2.2.2 du Chapitre 6.

⁶⁹ En tchèque, OBA, l'équivalent de l'anglais *both*, peut également se comporter de la sorte : *oba (ti) mládenci* = *both young men*. Il n'en va pas de même pour le quantificateur universel de la totalité distributive KAŽDÝ, traduit, selon les circonstances, par *tout/toute* ou *chaque* (*toute excuse, chaque participant*). Un fait intéressant relatif à VŠECHEN/VEŠKERÝ est que ces lexèmes peuvent aussi suivre le démonstratif : *ti všichni lidé / všichni ti lidé* (*tous ces/les gens*) ; dans le cas où le quantificateur suit le démonstratif, nous sommes d'avis qu'il serait possible de l'analyser comme un quantificateur flottant (cf. Dowty et Brodie, 1984 ; voir la Section 7.2.3).

⁷⁰ En revanche, nous en avons tenu compte s'ils précédaient le démonstratif (combinaison impossible en français). Nous avons également tenu compte de la combinaison *ten jeden kluk* (**ce un garçon*), également impossible en français.

également en français avec un déterminant central, se plaçant à droite de ce dernier et étant, pour cela, parfois appelés « post-déterminants »⁷¹. En résumé, seuls nous intéressaient les éléments fonctionnels qui, dans les langues connaissant la catégorie des déterminants, agissent comme des déterminants centraux, et s'excluent donc mutuellement avec le démonstratif⁷², car c'est là que réside la particularité du tchèque en comparaison avec ces langues. Le Tableau 5.20 croise la présence de cumuls avec les différents types et sous-types d'emploi en présentant les chiffres dans l'ordre décroissant. Il compare le nombre d'occurrences comportant des cumuls avec le nombre total d'occurrences pour chaque type et sous-type. Le graphique de la Figure 5.19 affiche la part que les différents (sous-)types occupent au sein du groupe des occurrences comportant des cumuls :

CUMULS		
Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
EXOPHORE MEMORIELLE	32	11,23 %
EMPLOIS « EMOTIONNELS »	20	6,04 %
OCCURRENCES AMBIGUËS	7	8,64 %
ENDOPHORE	6	2,78 %
Anaphore fidèle	2	3,45 %
Anaphore résomptive	2	16,67 %
Anaphore indirecte	1	7,69 %
Cataphore	1	1,59 %
Anaphore temporelle	0	0,00 %
Anaphore associative	0	0,00 %
EMPLOIS « METALINGUISTIQUES »	4	28,57 %
DEMONSTRATIFS INDEFINIS	1	25,00 %
EXOPHORE A-MEMORIELLE	0	0,00 %
Deixis situationnelle	0	0,00 %
Deixis temporelle	0	0,00 %
Deixis textuelle	0	0,00 %
TOTAL	70	7,00 %

Tableau 5.20 : Proportion des occurrences du démonstratif adnominal se cumulant avec un autre mot fonctionnel pour chaque type et sous-type d'emploi dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 70 occurrences)

⁷¹ Pour le français, ce parti est pris par exemple par Riegel, Pellat et Rioul (2016 [1994] : 305), qui parlent de « postdéterminants » pour les numéraux, *seul*, *propre*, etc., et d'« identificateurs » pour *autre* et *même*.

⁷² La seule exception est représentée par NĚJAKÝ (QUELQUE/UN CERTAIN), dont la variante française au pluriel *quelques* peut fonctionner comme postdéterminant : *ces quelques remarques*. La raison de ce choix est qu'en tchèque, la forme peut se combiner avec le démonstratif même au singulier, sachant qu'elle peut, en plus, le précéder (*nějaký ten*) ou le suivre (*ten nějaký*).

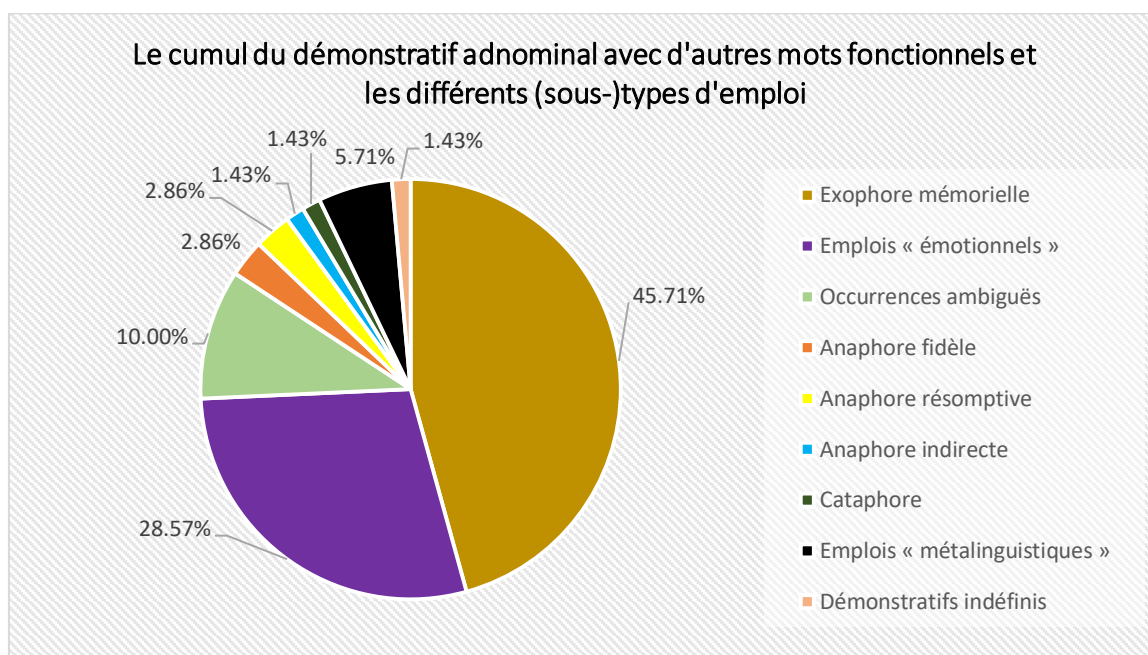


Figure 5.19 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal parmi les occurrences où le démonstratif se cumule avec un autre mot fonctionnel dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 70 occurrences)

Le Tableau 5.20 nous permet de voir que les cumuls apparaissent seulement avec 7 % des occurrences. Il convient tout de suite de relever que toutes les occurrences concernées sauf une seule sont des occurrences de TEN (celle qui fait exception emploie une forme de TENTO). Le cumul est totalement absent des emplois exophoriques a-mémoriels. Il est le plus élevé au sein des emplois « métalinguistiques » (28,57 % ; mais le nombre total de ces occurrences est très faible⁷³), mémorielles (11,23 %) et ambiguës (8,64 %). Du point de vue des pourcentages rapportés aux différents (sous-)types (le graphique de la Figure 5.19), le plus d'occurrences avec les cumuls apparaissent en emploi mémoriel (45,71 %), puis en emploi « émotionnel » (28,57 %)⁷⁴.

À présent, nous allons voir plus concrètement quels sont **les éléments fonctionnels se combinant avec le démonstratif au sein d'un même complexe nominal**. Le Tableau 5.21 reprend les différentes combinaisons représentées dans notre corpus par fréquence décroissante avec, chaque fois, les chiffres relatifs aux types et/ou sous-types d'emploi :

⁷³ Raison pour laquelle nous ne mentionnons même pas l'emploi indéfini.

⁷⁴ C'est dans les Chapitres 6 et 7 que les différentes combinaisons du démonstratif avec un autre élément fonctionnel seront analysées relativement aux emplois auxquels elles se prêtent.

Cumul	Type ou sous-type d'emploi	Nombre d'occurrences
TAKOVÝ TEN (TAKÝ TEN, TAKOVÝHLE TEN)	Exophore mémorielle	20
	Anaphore résomptive	1
	Anaphore indirecte	1
	Emplois « émotionnels »	1
	Total	23
TEN + possessif	Emplois « émotionnels »	12
	Exophore mémorielle	3
	Occurrences ambiguës	2
	Anaphore fidèle	1
	Démonstratifs indéfinis	1
	Total	19
NĚJAKÝ TEN	Emplois « émotionnels »	7
	Exophore mémorielle	4
	Occurrences ambiguës	3
	Emplois « métalinguistiques »	3
	Cataphore	1
	Total	18
TEN + le numéral JEDEN	Exophore mémorielle	2
	Anaphore fidèle	1
	Total	3
Numéral + TEN	Exophore mémorielle	2
Possessif + TEN	Emplois « métalinguistiques »	1
NĚKTERÝ TEN	Occurrences ambiguës	1
KAŽDÝ TEN	Occurrences ambiguës	1
Kolik těch	Exophore mémorielle	1
TAKOVÝ TENTO	Anaphore résomptive	1
TOTAL		70

Tableau 5.21 : Différents types de cumuls du démonstratif adnominal avec un autre mot fonctionnel et différents types et sous-types d'emploi au sein desquels ils apparaissent dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Le bilan est le suivant : le cumul le plus présent dans notre corpus est celui où **TEN** (et une fois **TENTO**) **se combine au démonstratif indéfini TAKOVÝ** (TEL ; avec, dans deux cas, ses variantes **TAKOVÝHLE** et **TAKÝ**⁷⁵). Cette combinaison, le plus directement traduisible en français par la combinaison *CE + genre/type/espèce(s) de*, est surtout utilisée avec l'exophore mémorielle (20 occurrences sur 24) :

⁷⁵ La première, contenant *hle*, est davantage présente sur le territoire de la Bohême, tandis que **TAKÝ** est une variété dialectale morave et silésienne : dans ORTOFON v1, **TAKOVÝHLE** atteint une fréquence relative de 354,1 occurrences pour un million de positions parmi les locuteurs originaires de Bohême et 67,2 occurrences parmi ceux originaires de Moravie et Silésie ; pour **TAKÝ**, la fréquence relative est de 170,1 occurrences dans la partie Moravie-Silésie et seulement de 5,4 occurrences dans la partie Bohême. Par la présence de *hle*, **TAKOVÝHLE** est également plus fortement déictique, d'où l'impossibilité pour cette forme d'apparaître en exophore mémorielle et le fait qu'elle est limitée aux emplois anaphorique et exophorique situationnel. La fréquence relative de la combinaison **TAKOVÝ TEN** ([lemma="takový"] [lemma="ten"]) atteint 559,6 occurrences pour un million de positions dans l'ensemble d'ORTOFON v1.

60) L1 : já ne-snáš-í-m
moi NEG-supporter-PRS-1SG

L2 : *takov-ý ty kumbálek-y*
tel-M.ACC.PL ten.M.ACC.PL cagibi-ACC.PL

L1 : jo .. jak je to v těch panelák-ách
oui comme être.PRS.3SG cela.NOM dans ten.M.ABL.PL tour-LOC.PL

L1 : je déteste

L2 : *ces espèces de cagibis*

L1 : ouais... genre comme on en voit dans les tours (des cités HLM) ORTOFON v1

Il y a ici un contraste intéressant avec l'échantillon de conversations françaises, où nous n'avons recensé que 5 occurrences au total de *cette espèce de* (3), *ces espèces de* (1) et *ce genre de* (1) se prêtant à l'emploi mémoriel. En revanche, l'échantillon français contient 30 occurrences de *ce genre de* en emploi anaphorique, notamment résomptif (18 occurrences sur la totalité de 170, soit 10,59 %) et indirect (9 occurrences sur la totalité de 19, soit 47,37 %). Une série d'exemples de ce type figure dans la Section 5.1.3, où nous avons également essayé de décrire les critères sur la base desquels *ce genre de* a été classé dans les différents types d'anaphore. Nous voudrions montrer ici que dans les rares cas où il est employé anaphoriquement, TAKOVÝ TEN présente les mêmes défis d'analyse :

61) když jsou tam ty Tchajwanc-i a .. a Číňan-i
quand être.PRS.3PL là-bas ten.M.NOM.PL Taïwanais-NOM.PL et Chinois-NOM.PL

a *takovýdl-e ty národnost-i*
et tel-F.NOM.PL ten.F.NOM.PL nationalité-NOM.PL

quand il y a des Taïwanais et des Chinois et *ce genre de nationalités* ORTOFON v1

62) že se to otevírá-l-o když mě-l-i [určit-ý
que REFL cela.NOM ouvrir-PST-3SG.N quand avoir-PST-3PL.M certain-M.ACC.SG

počet opakován-í]i a [*takov-ých těch věc-t*]i+? ..
nombre.ACC.SG répétition-GEN.PL et tel-F.GEN.PL ten.F.GEN.PL chose-GEN.PL

que ça s'ouvrait (devenait accessible) quand ils étaient arrivés à un certain nombre de répétitions et [*ce genre de choses*]i+? ORTOFON v1

63) L1 : oh j'ai lu un truc qui m'a choquée quoi dans un livre c'est un roman c'est euh Anna Karénine Tolstoï tu vois c'est russe ça se situe au dix-neuvième siècle et euh c'est euh euh un jeune un jeune homme là riche qui monte euh qui aime faire des courses à cheval

L2 : ouais

L1 : donc euh là il est en train de monter et de faire une course d'obstacles et il est mal retombé sur son cheval et il lui a brisé les reins

L2 : oh

L1 : non ça arrive ce vraiment *ce genre de choses* j- CEFC

64) L1 : ben ça reste [Titanic]i quoi c'est

L2 : je sais pas moi j'aime ça me plaît pas trop [*ce genre de films*]_{i+?} si je peux éviter j'évite
CEFC

L'exemple 61 relève, selon nous, de l'anaphore résomptive, son référent étant le fruit d'une opération de synthèse des deux référents précédemment introduits (*ty Tchajwanci a Čiňani*). *Takovýdle ty národnosti* (*ce genre de nationalités*) inclut, bien entendu, d'autres nationalités que les Taïwanais et les Chinois. Ainsi, cet exemple présente une certaine analogie avec l'exemple français 63, qui lui aussi relève de l'anaphore résomptive. En revanche, l'exemple 61 est plus proche de 64, les deux appartenant, selon nous, à l'anaphore indirecte. Contrairement à ce qui se produit dans ses emplois mémoriels, *ce genre de* en emploi anaphorique ne peut pas toujours être traduit en tchèque par TAKOVÝ TEN (en 64, par exemple, l'on serait obligé de le rendre par *takovýhle filmy*, *tyhlety filmy* ou *tenhle typ filmů* (littéralement *ce genre de films*⁷⁶). Il n'en reste pas moins que même en emploi anaphorique, les deux expressions peuvent parfois fonctionner comme des équivalents de traduction (cf. 61, 62).

La deuxième combinaison la plus répandue est celle de TEN suivi d'un possessif (19 occurrences dont 12 en emploi « émotionnel » ; voir le Tableau 5.21⁷⁷) :

65) <i>ten</i>	<i>náš</i>	<i>slavn-ěj</i>	..	<i>Jan</i>	<i>Evangelist-a</i>
ten.M.NOM.SG	nôtre.M.NOM.SG	célèbre-M.NOM.SG		NP.M.NOM.SG	NP-M.NOM.SG
 <i>Purkyn-ě ..</i>					
NP-M.NOM.SG					
 <i>notre célèbre... Jan Evangelista Purkinje...</i>					
ORTOFON v1					

En 65, l'emploi de la combinaison nous rappelle fortement l'exemple cité par Goyens (1994 : 110) et mentionné dans la Section 4.4 :

66) <i>nostr-um</i>	<i>ill-um</i>	<i>Caton-em</i>
nôtre-M.ACC.SG	ce-M.ACC.SG	NP-M.ACC.SG
<i>notre célèbre Caton</i>		

Dans les deux cas, ni le démonstratif ni le possessif ne modifie l'extension du nom tête et ne contribue à l'identification du référent. En conséquence, nous sommes ici face à un emploi « émotionnel » à valeur que l'on pourrait qualifier d'affective, le locuteur exprimant son attitude admirative vis-à-vis d'un référent en raison de sa notoriété. Néanmoins, cette combinaison connaît également des emplois en tchèque où l'analyse s'avère plus délicate, comme nous le verrons dans les

⁷⁶ Nous laissons ici de côté tous les problèmes d'analyse de cette expression en français : *genre* fonctionne-t-il encore comme un nom ou *ce genre de* est-il utilisé comme un déterminant complexe introduisant une autre tête nominale (cf. Wilmet, 1986) ?

⁷⁷ Dans l'ensemble d'ORTOFON v1, la fréquence relative de la combinaison TEN + possessif ([lemma= "ten"])[tag= "PS.....-----.*"]) atteint 391,4 occurrences pour un million de positions.

Chapitres 6 et 7, consacrés à la grammaticalisation de TEN et aux emplois « émotionnels » du démonstratif tchèque.

Une seule occurrence de notre échantillon présente l'ordre renversé des deux éléments fonctionnels. Comme nous l'avons déjà écrit, cet ordre est normalement impossible. Il est ici présent parce qu'il s'agit d'un emploi « métalinguistique » du démonstratif (la locutrice est en recherche du terme approprié pour la position d'attribut, recherche qui s'accompagne d'une hésitation et d'une aposiopèse)⁷⁸. Tout semble indiquer que ce contexte très particulier est le seul où cette combinaison puisse apparaître :

67) to je	Veronik-a .	to je	<i>naš-e</i>	<i>ta</i>
ce être.PRS.3SG	NP-F.NOM.SG	ce être.PRS.3SG	nôtre-F.NOM.SG	ten.F.NOM.SG
ça, c'est Veronika,	c'est notre	notre...		ORTOFON v1

La combinaison qui suit immédiatement, avec un nombre d'occurrences s'élevant à 18, est celle de NĚJAKÝ TEN (voir l'exemple 1 du Chapitre 3 *nějaké to pivo*). Cette combinaison se prête à une variabilité d'emplois bien plus grande, comme nous le montre le Tableau 5.21. Ces emplois seront également étudiés dans les Chapitres 6 et 7.

Parmi les combinaisons restantes, nous relevons notamment celle de TEN avec un numéral (dans les deux ordres), ainsi que trois autres occurrences, combinant NĚKTERÝ ((UN) CERTAIN), KAŽDÝ (*chaque*) et kolik (*combien de*) avec TEN⁷⁹.

Il reste à mentionner 4 cas qui n'ont pas été inclus dans les statistiques et où l'autre élément fonctionnel n'est pas antéposé (comme TEN), mais postposé au nom. Ces éléments sont : *svůj* (*son*), *nějakej* (*quelque/un certain*), *jedna* (*une*) et *jakási* (*une certaine*). Cet ordre ne produit pas le même effet marqué archaïsant que celui qui a été décrit dans les Sections 2.2.1 et 3.1.3. Dans les 4 cas, l'enregistrement indique que l'élément fonctionnel postposé est comme ajouté après coup par le locuteur. Il s'agit donc, au contraire, d'un trait qui relève de l'oralité :

68) jak	mě-l-y	ty	ps-y .	ty	tam
REL	avoir-PST-3PL.F	ten.M.ACC.PL	chien-ACC.PL	celui-ci.F.NOM.PL	là-bas
mě-l-y	<i>ten</i>	<i>ps-í</i>	<i>sjezd</i>	<i>nějak-ej</i>	
avoir-PST-3PL.F	ten.M.ACC.SG	canin-M.ACC.SG	congrès.ACC.SG	quelque-M.ACC.SG	

⁷⁸ Il ne s'agit donc pas d'une postposition stricto sensu du démonstratif. L'ensemble d'ORTOFON v1 contient, au total, 10 occurrences de ce type ([tag="PS.....-.*"] [lemma="ten"]).

⁷⁹ Dans l'ensemble d'ORTOFON v1, la fréquence relative de la combinaison NĚJAKÝ TEN ([lemma="nějaký"] [lemma="ten"]) atteint 279 occurrences pour un million de positions. La combinaison TEN NĚJAKÝ, elle, est beaucoup moins présente dans le corpus. Malheureusement, elle est difficile à viser par la requête [lemma="ten"] [lemma="nějaký"], celle-ci aboutissant à une majorité d'occurrences où TEN est un pronom au singulier neutre.

5.2.3.4 Position de l'occurrence du démonstratif adnominal par rapport au verbe de la proposition

Comme nous l'avons fait pour notre échantillon français, nous allons maintenant examiner les positions des occurrences par rapport au verbe de la proposition. Le Tableau 5.22 et le graphique de la Figure 5.20 *infra* apportent une synthèse des fréquences en fonction des positions de l'occurrence du démonstratif adnominal par rapport au verbe de la proposition. Toutes les fréquences sont complétées d'un intervalle de confiance, calculé par l'interface « many features – 1 sample » du CNT (voir la Section 5.2.2.2)⁸⁰:

Position	Nombre d'oc- currences	Par rapport à l'ensemble des occurrences de l'échantillon (%)	Intervalle de confiance	
			Borne inférieure	Borne supérieure
POSITION POSTVERBALE	494	49,40 %	463	525
POSITION PREVERBALE	187	18,70 %	163	212
POSITION INCONNUE	104	10,40 %	85	123
COMPLEMENT DE PHRASE	100	10,00 %	82	119
DISLOCATION A DROITE	76	7,60 %	60	93
DISLOCATION A GAUCHE	39	3,90 %	27	51
TOTAL	1 000	100,00 %		

Tableau 5.22 : Position de l'occurrence du démonstratif adnominal par rapport au verbe de la proposition dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

⁸⁰ Nous employons la fréquence absolue moyenne du démonstratif adnominal, soit le chiffre 20 311 (voir la Section 5.2.1) ; $\alpha = 0,05$.

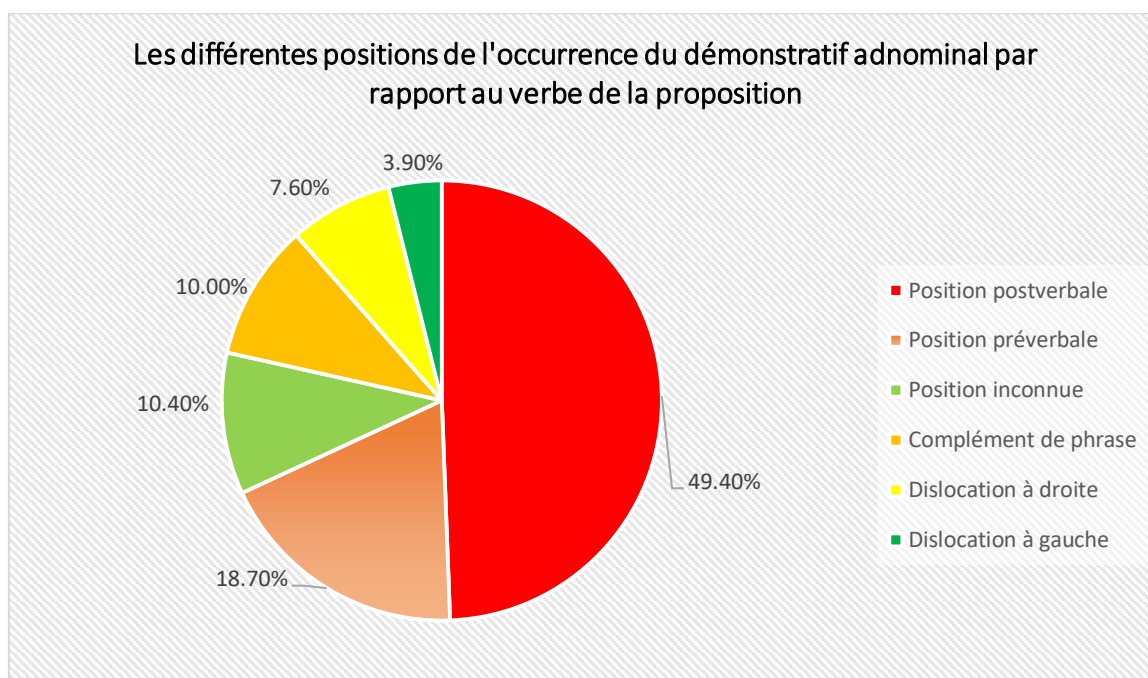


Figure 5.20 : Représentation en % des différentes positions de l'occurrence du démonstratif adnominal par rapport au verbe de la proposition dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Le Tableau 5.22 permet de faire les constats suivants : la position dominante est la position postverbale (494 occurrences, soit 49,4 % de l'ensemble). Cette situation n'est pas sans rappeler celle de l'échantillon français, où ce chiffre n'atteint cependant que 359 occurrences (35,9 % de l'ensemble)⁸¹. Pourtant, pour le reste, les données des deux langues divergent, car la deuxième position la plus représentée en tchèque est la position préverbale (187 occurrences, soit 18,7 % de l'ensemble), alors que celle-ci n'est que très faiblement présente dans l'échantillon français (19 occurrences, soit 1,9 % de l'ensemble). Vu qu'en tchèque 34,22 % des occurrences apparaissant dans cette position relèvent des emplois « émotionnels » et 18,72 % relèvent de l'exophore mémorielle (voir à cet égard le graphique de la Figure 5.21 – *infra*), il semblerait que la faible représentation de cette position dans l'échantillon français soit en partie due à la faible représentation de ces deux types d'emploi (11 occurrences « émotionnelles » et 62 occurrences mémorielles contre 331 et 285 dans l'échantillon d'ORTOFON v1).

Nous constatons également une disproportion dans les compléments de phrase, avec 100 occurrences en tchèque (10 % de l'ensemble) contre 356 occurrences en français. Le corpus tchèque contient moins de dislocations à droite : 76 (7,6 % de l'ensemble) contre 144 en français. En revanche, les valeurs sont plus proches pour la dislocation à gauche : 39 occurrences en tchèque (3,9 % de l'ensemble) et 48 occurrences en français. Le nombre de cas où la position n'a pas pu être déterminée est de 104 (10,4 de l'ensemble ; 70 en français).

⁸¹ La différence est statistiquement significative, le score de la fonction de vraisemblance étant de 37,382 (la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$)).

Qu'en est-il des **affinités entre la position de l'occurrence par rapport au verbe et les différents types et sous-types d'emploi** ? Comme dans la Section 5.2.2.2, la situation est présentée à travers une série de tableaux (à commencer par la position par rapport au verbe la plus répandue⁸²), complétée d'une série de graphiques qui affichent les proportions des différents (sous-)types d'emploi au sein de chaque type de position⁸³. De même que pour l'échantillon de conversations françaises, seuls les types d'emploi seront renseignés pour la position inconnue.

POSITION POSTVERBALE		
Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
EXOPHORE MEMORIELLE	175	61,40 %
EMPLOIS « EMOTIONNELS »	157	47,43 %
ENDOPHORE	88	40,74 %
Cataphore	31	49,20 %
Anaphore fidèle	25	43,10 %
Anaphore associative	20	46,30 %
Anaphore infidèle	5	33,33 %
Anaphore résomptive	4	33,33 %
Anaphore indirecte	3	23,08 %
EXOPHORE A-MEMORIELLE	25	36,23 %
Deixis situationnelle	23	37,10 %
Deixis temporelle	1	16,67 %
Deixis textuelle	1	100,00 %
OCCURRENCES AMBIGUES	37	45,68 %
EMPLOIS « MÉTALINGUISTIQUES »	9	64,29 %
DEMONSTRATIFS INDEFINIS	3	75,00 %
TOTAL	494	49,40 %

Tableau 5.23 : Rapport entre la position postverbale et les différents types et sous-types d'emploi du démonstratif adnominal dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

⁸² Seuls sont représentés les types et sous-types d'emploi apparaissant dans la position en question.

⁸³ Les graphiques représentent les types d'emploi là où ceux-ci ne sont pas divisés en sous-types ; s'ils le sont, ce sont les sous-types qui sont représentés. De même que dans la Section 5.2.2.2, les tableaux adoptent le point de vue des (sous-)types d'emploi, et les graphiques adoptent celui de la position par rapport au verbe.

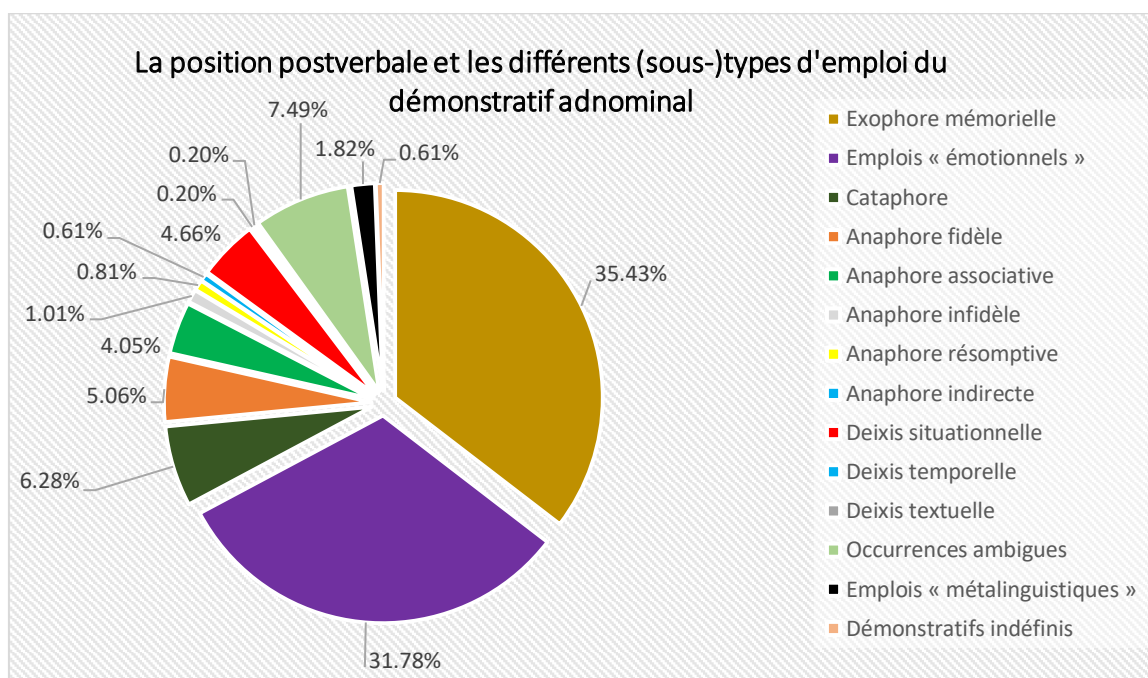


Figure 5.21 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal au sein de la position postverbale dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Si la **position postverbale** est la plus représentée dans les échantillons des deux langues, nous voyons que contrairement à l'échantillon de conversations françaises, où la majeure partie des occurrences figurant dans la position postverbale étaient de nature endophorique (51,53 % de l'ensemble), ici, elles sont avant tout de nature mémorielle (175 occurrences, soit 35,43 %) et « émotionnelle » (155 occurrences, soit 31,78 %), les emplois endophoriques n'occupant que la troisième place (88 occurrences, soit 17,81 %). Nous nous expliquons cette disparité, avant tout, par le simple fait que l'emploi endophorique est plus représenté en français qu'en tchèque (489 occurrences vs 216 occurrences), alors que c'est tout l'inverse pour l'emploi mémoriel (62 occurrences vs 285 occurrences) et « émotionnel » (11 occurrences vs 331 occurrences)⁸⁴. L'emploi exophorique a-mémoriel – beaucoup moins représenté en tchèque de manière générale – n'atteint, dans cette position, que le chiffre de 25 (soit 5,11 % contre 100 en français). Une fois de plus, tout cela a partie liée avec les différences dans la valeur linguistique (Saussure, Bally, Sechehaye, 1960 [1916]) des deux unités CE et TEN : tandis que CE est un démonstratif « fort » en ce sens qu'il n'a rien perdu de son fonctionnement indexical, la valeur de TEN se situe à la charnière entre celle de CE et celle de LE, tantôt plus proche de l'une et tantôt de l'autre (voir à cet égard la Section 3.3). Cela expliquerait en partie pourquoi TEN est moins présent en emplois situationnel et anaphorique (où l'on lui préfère parfois TENTO/TENHLE) et plus présent en emploi mémoriel (où, au contraire, il sera souvent employé là où le français emploie LE).

⁸⁴ Voir les Tableaux 5.6 et 5.17 et les graphiques des Figures 5.1 et 5.16 *supra*.

De même que pour nos données françaises, le recours au calcul de spécificités révèle l'existence d'une **affinité certaine entre la position postverbale et l'emploi mémoriel**, comme nous pouvons l'observer sur le graphique de la Figure 5.22 (première colonne) :

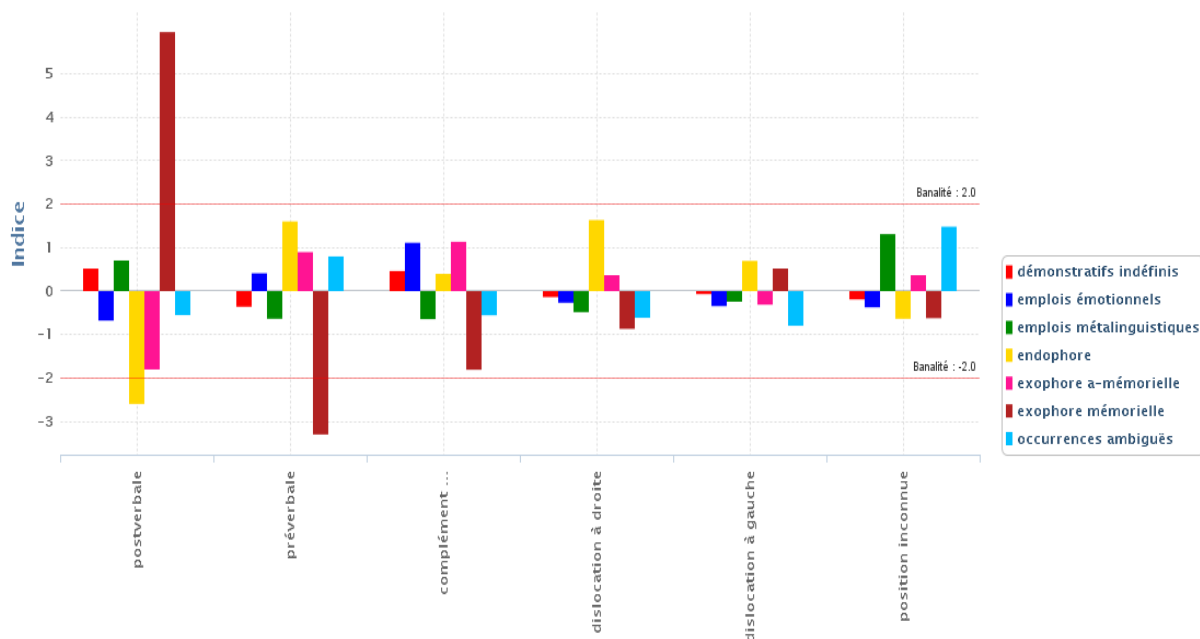


Figure 5.22 : Calcul de spécificités appliqué au lien entre les différentes positions de l'occurrence du démonstratif adnominal par rapport au verbe de la proposition et les différents types d'emploi dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Nous nous expliquons cette affinité – présente dans les deux ensembles de données – par le profil des référents mémoriels : ces référents, bien qu'identifiables, sont nouveaux du point de vue discursif (ils ne sont pas « activés » ; pour cette notion, voir la Section 6.2.1.2 du Chapitre 6), d'où leur tendance à apparaître en position postverbale, qui est la position rhématique.

Certaines analogies rapprochent les deux échantillons quant à la proportion occupée par les différents (sous-)types d'emploi à l'intérieur de cette position syntaxique : la position postverbale accueille 40,74 % de tous les démonstratifs endophoriques tchèques et 37,83 % de tous les démonstratifs endophoriques français, 45,68 % de l'ensemble des occurrences ambiguës tchèques et 50 % de l'ensemble des occurrences ambiguës françaises, tout comme 36,23 % des démonstratifs exophoriques a-mémoriels tchèques et 25,71 % des démonstratifs exophoriques a-mémoriels français. Dans l'échantillon tchèque, cette position accueille également 61,4 % de tous les emplois mémoriels et 47,43 % de tous les emplois « émotionnels ».

POSITION PREVERBALE

Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
EMPLOIS « EMOTIONNELS »	64	19,34 %
ENDOPHORE	51	23,61 %
Anaphore associative	19	35,19 %
Anaphore fidèle	15	25,87 %
Cataphore	7	11,11 %
Anaphore indirecte	4	30,77 %
Anaphore infidèle	4	26,67 %
Anaphore résomptive	2	16,67 %
EXOPHORE MEMORIALE	35	12,28 %
EXOPHORE A-MÉMORIALE	17	24,64 %
Deixis situationnelle	17	27,42 %
OCCURRENCES AMBIGUES	19	23,46 %
EMPLOIS « MÉTALINGUISTIQUES »	1	7,14 %
TOTAL	187	18,70 %

Tableau 5.24 : Rapport entre la position préverbale et les différents types et sous-types d'emploi du démonstratif adnominal dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

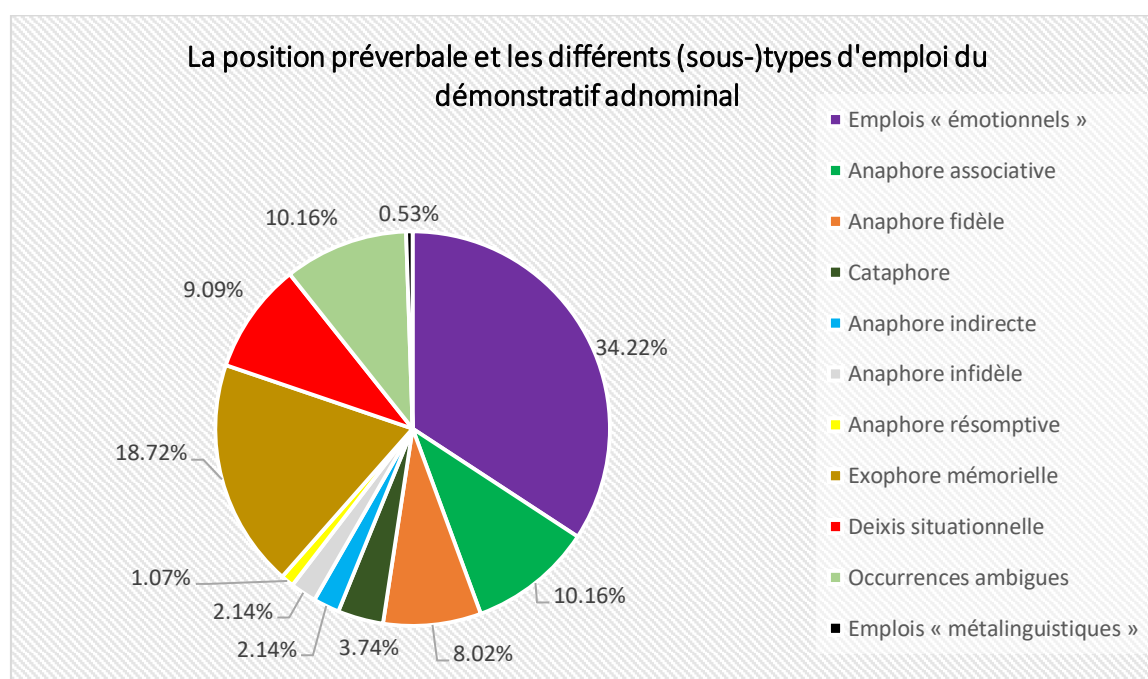


Figure 5.23 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal au sein de la position préverbale dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Dans la **position préverbale**, ce sont les emplois « émotionnels » qui priment (34,22 % de l'ensemble), suivis des emplois endophoriques (27,27 % de l'ensemble ; 10,16 % relèvent de l'anaphore associative, 8,02 % de l'anaphore fidèle et 3,74 % de la cataphore) et mémoriels (18,72 % de l'ensemble). La deixis situationnelle représente 9,09 % des occurrences en position préverbale.

POSITION INCONNUE		
Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
EMPLOIS « EMOTIONNELS »	33	9,97 %
EXOPHORE MEMORIALE	26	9,12 %
ENDOPHORE	19	8,80 %
OCCURRENCES AMBIGUES	14	17,28 %
EXOPHORE A-MEMORIALE	8	11,59 %
EMPLOIS « MÉTALINGUISTIQUES »	4	28,57 %
TOTAL	104	10,40 %

Tableau 5.25 : Rapport entre la position inconnue et les différents types et sous-types d'emploi du démonstratif adnominal dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

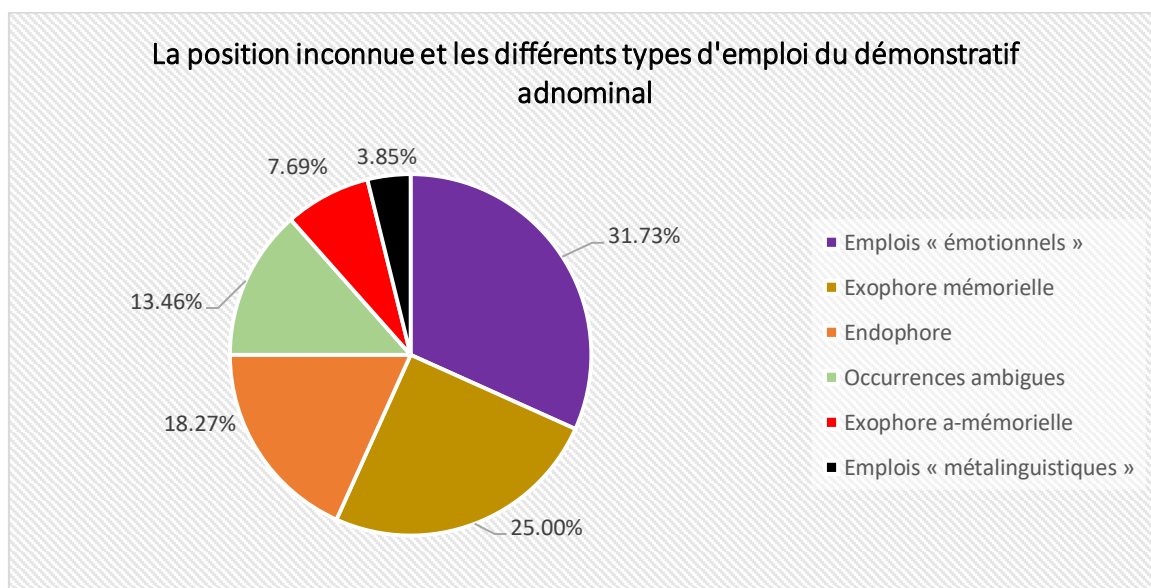


Figure 5.24 : Représentation en % des différents types d'emploi du démonstratif adnominal au sein de la position inconnue dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

COMPLEMENT DE PHRASE

Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
EMPLOIS « EMOTIONNELS »	40	12,08 %
ENDOPHORE	23	10,65 %
Cataphore	10	15,87 %
Anaphore résomptive	4	33,33 %
Anaphore infidèle	4	26,67 %
Anaphore fidèle	2	3,45 %
Anaphore associative	1	1,85 %
Anaphore indirecte	1	7,69 %
Anaphore temporelle	1	100,00 %
EXOPHORE MEMORIALE	19	6,67 %
EXOPHORE A-MÉMORIALE	11	15,94 %
Deixis situationnelle	6	9,68 %
Deixis temporelle	5	83,33 %
OCCURRENCES AMBIGUES	6	7,41 %
DEMONSTRATIFS INDEFINIS	1	25,00 %
TOTAL	100	10,00 %

Tableau 5.26 : Rapport entre la position de complément de phrase et les différents types et sous-types d'emploi du démonstratif adnominal dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

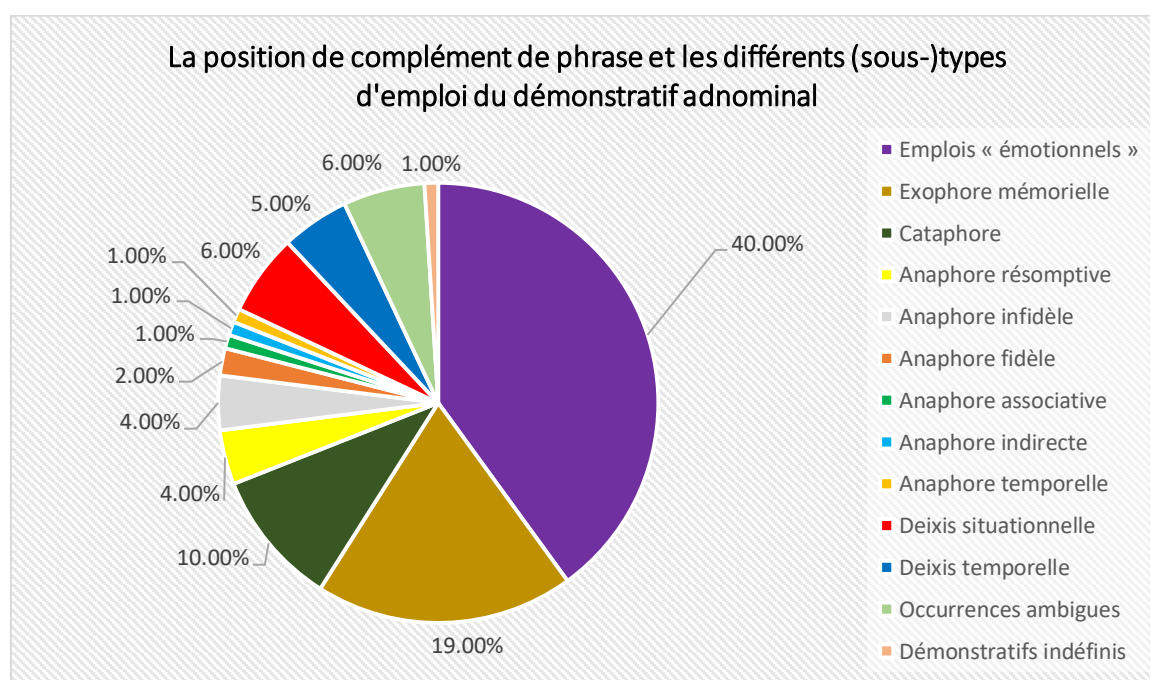


Figure 5.25 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal au sein de la position de complément de phrase dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Quant à la **position de complément de phrase**, la plupart des occurrences sont des emplois « émotionnels » (40 %), suivis des emplois endophoriques (23 % ; 10 % de cataphores) et mémoriels (19 %). Les emplois exophoriques a-mémoriels constituent 11 % de l'ensemble. Cette situation est incomparable avec celle du français, où, rappelons-le, 207 occurrences déictiques temporelles (58,15 %), 48 occurrences anaphoriques résomptives (13,48 %) et 36 occurrences anaphoriques temporelles (10,11 %) figurent dans cette position et font office de cadre accueillant la prédication de la proposition. Cette disproportion entre les deux langues n'a rien de surprenant si nous tenons compte de l'affinité déjà soulignée entre le complément de phrase et les emplois déictique et anaphorique temporels et l'emploi anaphorique résomptif : comme ces trois sous-types d'emploi ne sont que très peu représentés dans le corpus tchèque, c'est également cette position syntaxique qui connaît un taux de représentation bien plus faible⁸⁵.

DISLOCATION A DROITE		
Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
EMPLOIS « EMOTIONNELS »	25	7,55 %
ENDOPHORE	24	11,11 %
Anaphore associative	10	18,52 %
Anaphore fidèle	9	15,52 %
Anaphore indirecte	2	15,38 %
Anaphore résomptive	1	8,33 %
Anaphore infidèle	1	6,67 %
Cataphore	1	1,59 %
EXOPHORE MEMORIELLE	17	5,96 %
EXOPHORE A-MEMORIELLE	6	8,70 %
Deixis situationnelle	6	9,68 %
OCCURRENCES AMBIGUËS	4	4,94 %
TOTAL	76	7,60 %

Tableau 5.27 : Rapport entre la position de la dislocation à droite et les différents types et sous-types d'emploi du démonstratif adnominal dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

⁸⁵ Il a été démontré *supra* que la différence est statistiquement significative pour les emplois déictiques temporels et anaphoriques résomptifs. Elle l'est également pour les emplois anaphoriques temporels, le score de la fonction de vraisemblance étant de 53,7382 ; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

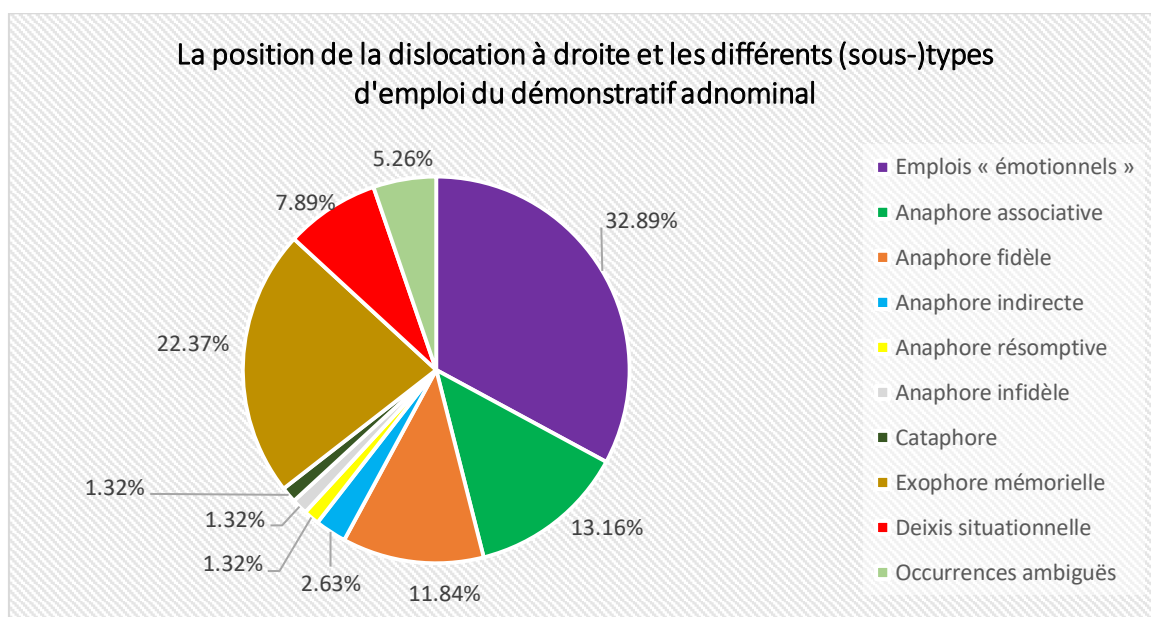


Figure 5.26 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal au sein de la position de la dislocation à droite dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

DISLOCATION A GAUCHE

Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
EXOPHORE MEMORIELLE	13	4,56 %
EMPLOIS « EMOTIONNELS »	12	3,63 %
ENDOPHORE	11	5,09 %
Cataphore	6	9,52 %
Anaphore indirecte	2	15,38 %
Anaphore fidèle	2	3,45 %
Anaphore associative	1	1,85 %
EXOPHORE A-MEMORIELLE	2	2,90 %
Deixis situationnelle	2	3,23 %
OCCURRENCES AMBIGUËS	1	1,23 %
TOTAL	39	7,60 %

Tableau 5.28 : Rapport entre la position de la dislocation à gauche et les différents types et sous-types d'emploi du démonstratif adnominal dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

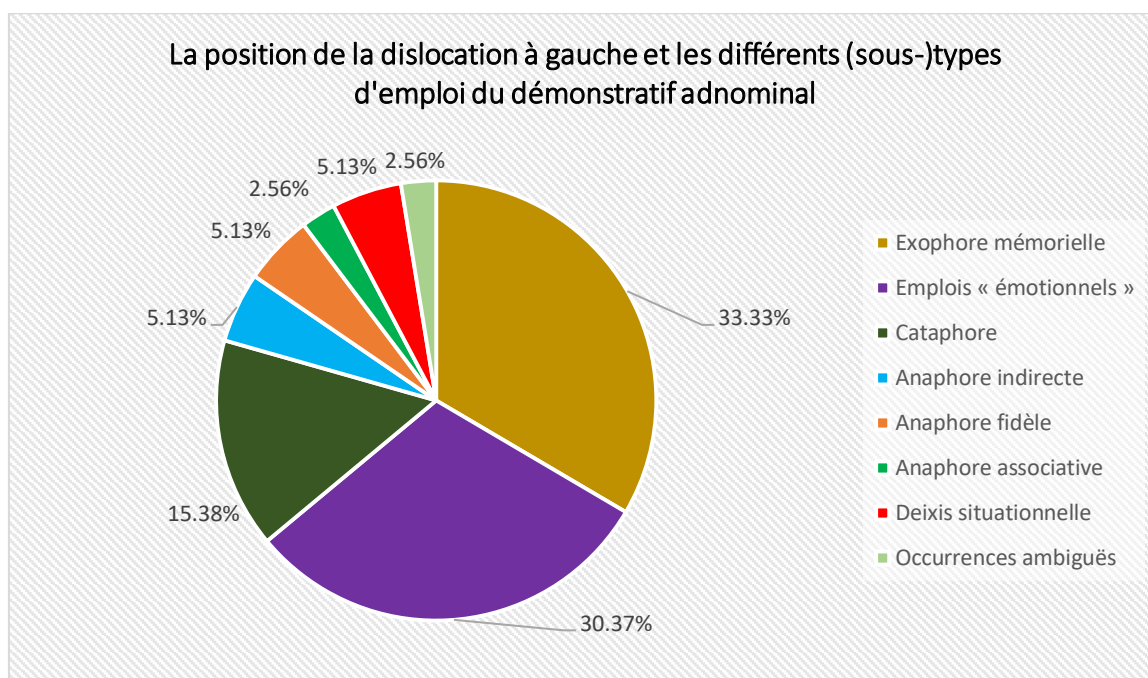


Figure 5.27 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal au sein de la position de la dislocation à gauche dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Pour ce qui est des occurrences apparaissant en **dislocation à droite**, 32,89 % d'entre elles sont de nature « émotionnelle », 31,58 % de nature endophorique⁸⁶, 22,37 % de nature mémorielle et 7,89 % de nature déictique situationnelle. Rappelons que dans l'échantillon de conversations françaises, la majorité des occurrences occupant cette position, soit 72,2 %, était de nature anaphorique. Au sein de **la dislocation à gauche**, 33,33 % des occurrences sont mémorielles, 30,37 % d'entre elles sont « émotionnelles », et 28,21 % d'entre elles sont endophoriques (15,38 % de cataphores).

Une différence nous frappe d'emblée : tandis que la dislocation à droite contient 10 anaphores associatives, la dislocation à gauche n'en contient qu'une seule⁸⁷. Pour rendre compte de cette disproportion, nous proposons l'explication suivante : par rapport aux référents des premières occurrences des autres expressions anaphoriques⁸⁸, ceux des anaphores associatives sont déjà établis dans le discours et traités davantage comme une information connue et présupposée, car leur identification procède d'un lien implicite tissé avec le discours précédent. C'est également la raison

⁸⁶ Parmi les 24 cas d'endophore, 10 ressortissent à l'anaphore associative et 9 à l'anaphore fidèle.

⁸⁷ La différence semble être statistiquement significative selon la fonction de vraisemblance, affichant le score de 4,0608 (la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$)) ; néanmoins, force est de dire que la p -valeur (0,0439) se rapproche ici de α et que les valeurs sont basses pour que l'on puisse en tirer une conclusion fiable.

⁸⁸ C'est-à-dire les premières reprises.

pour laquelle, à la différence de la majorité des autres anaphores (mise à part l'anaphore résomptive⁸⁹), leur référent est introduit comme défini dès la première mention :

- 69) *když fot-í-š . . . z výšk-y* .. tak [to]_i vypad-á
 quand photographe-PRS-2SG de hauteur-GEN.SG alors cela.NOM avoir l'air-PRS.3SG
- pak [ta fotk-a]_i jak ..*
 après ten.F.NOM.SG photo-NOM.SG comme
- si tu photographies en hauteur, [ça ressemble]_i, [la photo]_i à...* ORTOFON v1

S'agissant de la nature des référents des autres types d'anaphore occupant cette position (14 occurrences au total), nous constatons que dans presque tous les cas (sauf peut-être un seul), ceux-ci sont identifiés via leur introduction antérieure dans le discours (souvenons-nous qu'en français, ils étaient dans plus de la moitié des cas identifiés grâce à une connaissance extra-discursive)⁹⁰. Mais nous constatons en même temps que dans la plupart des cas, le référent a déjà acquis un certain degré de proéminence et de notoriété au moment où il est repris par l'expression anaphorique, à force d'être apparu plusieurs fois dans le discours (dans 7 cas, nous considérons TEN comme « grammaticalisé » ; voir la Section 5.2.3.2) :

- 70) *venku byl-a .. [žensk-á]_i .. jsem pochopi-l-a že [to]_i*
 dehors être.PST-3SG.F femme-NOM.SG AUX.PST.1SG comprendre-PST-1SG.F que cela.NOM
- bud-e asi [ta .. prodavačk-a]_i .. a sedě-l-a tam*
 AUX.FUT-3SG MOD ten.F.NOM.SG vendeuse-NOM.SG et être assis-PST-3SG.F là-bas
- [čern-á kočk-a]_{ii} .. i na dálk-u byl-o vidět jak*
 noir-F.NOM.SG chatte-NOM.SG même sur distance-ACC.SG être.PST-3SG.N voir.INF comme
- m-á nádhern-ý zelen-ý oč-i .. a [vona]_i .. [ju]_{ii}*
 avoir-PRS.3SG splendide-N.ACC.PL vert-N.ACC.PL œil-ACC.PL et elle.NOM.SG elle.ACC.SG
- tak jako [vypada-l-a]_{ii} moc dobře .. tak já jsem*
 MOD avoir l'air-PST-3SG.F trop bien alors moi.NOM AUX.PST.1SG
- tak jako chvíl-u vodpočív-a-l-a*
 MOD moment-ACC.SG se reposer-PTCP-SG.F
- a potom jsem tam šl-a .. da-l-a jsem*
 et après AUX.PST.1SG là-bas aller.PTCP-SG.F donner-PTCP-SG.F AUX.PST.1SG

⁸⁹ Dans ce cas, l'établissement du référent dans le discours requiert une opération d'inférence plus explicite et, semble-t-il, encore plus coûteuse en effort mental que l'établissement du référent d'une anaphore « classique », car il n'y a pas reprise mais création d'un référent nouveau.

⁹⁰ Dans les deux échantillons, la dislocation à droite contient plus d'occurrences du démonstratif adnominal dont le référent est identifiable grâce à des informations extra-discursives (dans cette classe, nous rangeons également les occurrences « émotionnelles ») que d'occurrences dont le référent est identifiable grâce au discours. Pour les deux échantillons, la différence est statistiquement significative, la fonction de vraisemblance affichant le score de 23,704 pour le français et 18,2098 pour le tchèque ; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

s [nim-a]_{i+ii} řeč . s [kočk-ou]_{ii} i s [prodavačk-ou]_i ..
avec elle.INS.PL conversation.ACC.SG avec chatte-INS.SG même avec vendeuse-INS.SG

[mě-l-a]_{ii} fakt nádhern-ý ty zelen-ý voč-i
avoir-PST-3SG.F vraiment splendide-N.ACC.PL ten.N.ACC.PL vert-N.ACC.PL œil-ACC.PL

...
a [ba* baba]_i mně řek-l-a že .. [vona]_{ii} k [ní]_i
et bonne femme-NOM.SG moi.DAT dire-PST-3SG.F que elle.NOM.SG chez elle.DAT.SG

chod-i na svačink-u . [ta kočk-a]_{ii}
aller-PRS.3SG à goûter-ACC.SG ten.F.NOM.SG chatte-NOM.SG

dehors, il y avait [une bonne femme]_i... j'ai compris que [ça]_i serait sans doute [la... vendeuse...]_i et il y avait [une chatte noire]_{ii} qui était assise là... ; même à distance, on voyait qu'[elle]_i avait des yeux verts splendides..., et [l'autre]_i [l']_{ii} avait... [elle]_{ii} avait vraiment l'air joli... ; du coup, moi, je me suis posée un peu et puis je suis rentrée... causer avec [elles]_{i+ii}, avec [la chatte]_{ii} et [la vendeuse]_i, aussi... ; [elle]_{ii} avait vraiment des yeux verts splendides (comme je disais) ; et [la bonne femme]_i, [elle]_i m'a dit que... qu'[elle]_{ii} venait chez [elle]_i à l'heure du goûter, [la chatte]_{ii}
ORTOFON v1

En 70, on assiste à l'alternance des deux référents les plus proéminents du récit, à savoir la vendeuse et la chatte. Par conséquent, lorsqu'apparaît la dislocation à droite de *ta kočka (la chatte)*, le référent de l'expression est déjà très bien établi dans le discours.

Le nombre d'occurrences exophoriques (mémoires et a-mémoires confondues) est, comme en français, moins élevé en dislocation à gauche qu'en dislocation à droite (15 vs 23 ; en français : 15 vs 30)⁹¹. Cependant, la composition de ces occurrences est différente dans les deux langues, car en français, nous trouvons avant tout des démonstratifs exophoriques a-mémoriels, tandis qu'en tchèque, ces derniers sont minoritaires par rapport aux démonstratifs exophoriques mémoriels (uniquement 6 occurrences déictiques situationnelles en dislocation à droite et 2 en dislocation à gauche contre 17 et 13 occurrences mémorielles en dislocation à droite et en dislocation à gauche)⁹².

Pour ce qui concerne les 11 emplois endophoriques en dislocation à gauche, 6 d'entre eux sont de nature cataphorique intra-syntagmatique :

71) a ted'ka .. [v tom barák-u co bydl-i]_i [...] že jo
et maintenant dans ten.M.LOC.SG maison-LOC.SG REL habiter-PRS.3SG MOD
[tam]_i jí zemře-l nějak-ej star-ej dědek
là-bas elle.DAT.SG mourir-PST.3SG.M quelque-M.NOM.SG vieux-M.NOM.SG croûton.NOM.SG

et là... [dans la maison où elle habite]_i, tu vois, il y a un espèce de vieux croûton qui est mort

⁹¹ Malheureusement, cette différence n'est significative sur le plan statistique ni pour l'un ni pour l'autre échantillon : le score de la fonction de vraisemblance étant de 0,7743 pour le français et de 2,086 pour le tchèque ; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

⁹² Cette différence serait, quant à elle, significative pour l'exophore mémorielle en dislocation à droite : $L = 16,7052$, pour l'exophore mémorielle en dislocation à gauche : $L = 13,7108$ et pour l'exophore a-mémorielle en dislocation à gauche : $L = 8,137$, mais ne le serait pas pour l'exophore a-mémorielle en dislocation à droite : $L = 3,5125$; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

En dislocation à droite, seule une occurrence cataphorique intra-syntagmatique apparaît. Peut-être serait-il possible par conséquent de faire l'hypothèse que les référents des cataphores intra-syntagmatiques sont moins établis dans le discours que ceux des anaphores associatives (bien qu'ici aussi, le référent soit directement établi comme défini par le complément restrictif). Malheureusement, ces hypothèses restent au stade de la spéculation vu le nombre limité des occurrences anaphoriques associatives et cataphoriques intra-syntagmatiques disloquées.

Comme nous l'avons constaté pour le français, les référents des dislocations à gauche se présentent souvent comme moins établis dans le discours (moins proéminents), le locuteur pouvant recourir à cette position dans le but, précisément, d'augmenter leur proéminence. Cet effet semble présent dans le cas des emplois endophoriques, mémoriels, ainsi qu'« émotionnels ». Pour ce qui est des premiers, nous invitons le lecteur à comparer les exemples 69 et 70 avec l'exemple 71. Pour ce qui est des deuxièmes, voici les trois exemples qui suivent :

- 72) L1 : m-á-m tam prosím tě . [tu vest-u]_i .. [tu zimn-í]_i
avoir-PRS-1SG là-bas s'il te plaît ten.F.ACC.SG gilet-ACC.SG ten.F.ACC.SG hivernal-F.ACC.SG

[ta]_i by se moh-l-a vyprat ..
celui-ci.F.NOM.SG AUX.COND.3SG REFL pouvoir-PTCP-SG.F laver-INF

...
L2 : no já už jsem [ji]_i přece pra-l-a jednou .
MOD moi.NOM déjà AUX.PST.1SG elle.ACC.SG pourtant laver-PTCP-SG.F une fois

[tu vest-u]_i
ten.F.ACC.SG gilet-ACC.SG

L1 : est-ce qu'il y a [ma veste]_i ... [celle pour l'hiver]_i, s'il te plaît ? il faudrait peut-être [la]_i laver

...
L2 : mais je [l']_i ai déjà lavée une fois, [la veste]_i ORTOFON v1

- 73) L1 : já m-á-m totiž pocit jestli si vzpomín-á-š [tam
moi.NOM avoir-PRS-1SG MOD sensation.ACC.SG si REFL rappeler-PRS-2SG là-bas

někde když se jed-e jakoby z Rybář do Drahotuš]_i tak [tam]_i ..
quelque part quand REFL aller-3SG MOD de Rybáře.GEN.PL à Drahotuše.GEN.PL alors là-bas

L2 : stáv-á vod-a [na tom pol-í]_i .
se tenir-PRS.3SG eau-NOM.SG sur ten.N.LOC.SG champ-LOC.SG

L1 : parce que moi, j'ai l'impression... si tu te rappelles, [là quand on va de Rybáře à Drahotuše]_i, alors [là-bas]_i

L2 : il y a souvent de l'eau, dans [le champ]_i ORTOFON v1

- 74) L1 : v zim-ě jezd-í buďto s ní .. a . toho ..
en hiver-LOC.SG rouler-PRS.3SG soit avec elle.INS.SG et ten.M.ACC.SG

nissan-a	nebo	[<i>tu</i>	<i>toyot-u</i>] _i	[tu] _i
nissan-ACC.SG	ou	ten.F.ACC.SG	toyota-ACC.SG	celui-ci.F.ACC.SG
necha-l ..	tej ..			
laisser-PST.3SG.M	ten.F.DAT.SG			
L2 : manželec-e	ne	bejval-ej		
épouse-DAT.SG	MOD	ancien-F.DAT.SG		

L1 : *en hiver, c'est soit avec elle qu'il roule... et la... nissan ou [la toyota]_i, il [l']_i a laissée*
L2 : *à son ex-femme, non ?* ORTOFON v1

Tandis qu'en 72 et 73, les référents (le gilet et le champ) se présentent comme une information plus solidement établie dans le discours (en 72, le fait que le référent a été mentionné plus tôt contribue également à cet effet), celui de 74 (la toyota) se présente comme moins établi, moins thématique (saillant) et plus difficilement identifiable pour l'interlocuteur. La dislocation à gauche permet ainsi d'asseoir le référent au sein de l'univers du discours des participants à la communication.

Ce même effet nous semble tangible même lorsque le référent est de nature « émotionnelle », et que le démonstratif n'intervient donc pas (directement) dans son identification. Nous invitons le lecteur à comparer 75 (dislocation à droite) avec 76 (dislocation à gauche). En 75, il s'agit de la neuvième occurrence du référent, contrairement à 76, où il s'agit de sa première occurrence⁹³ :

75) já	mus-í-m	taky začít	hučet	do tý
moi.NOM	devoir-PRS-1SG	aussi commencer.INF	rebattre les oreilles.INF	dans ten.F.GEN.SG
Markétk-y	aby	je-l-a	taky	[tam] _i
NP-F.GEN.SG	que.CONJ+AUX.COND.3SG	aller-PST-3SG.F	aussi	là-bas
[<i>do tý</i>	<i>Franci-e</i>] _i			
en ten.F.GEN.SG	France-GEN.SG			
<i>moi aussi, je dois convaincre Markétka pour qu'elle aille aussi [là-bas]_i, [en France]_i</i>				
ORTOFON v1				

76) [<i>v</i>	<i>té</i>	<i>chemi-i</i>] _i	[tam] _i	jd-ou	velk-ý	peníz-e ..
dans ten.F.LOC.SG	chimie-LOC.SG	là-bas	aller-PRS.3PL	grand-M.NOM.PL	argent-NOM.PL	
<i>[dans l'industrie chimique]_i, il y a des grosses sommes qui vont [là-bas]_i ...</i>						
ORTOFON v1						

5.2.3.5 Les disfluences dans les occurrences du démonstratif adnominal

De même que nous l'avons fait pour le français dans la Section 5.2.2.3, nous apportons ici une synthèse de l'association de disfluences (amorces inachevées, répétitions, marques d'hésitation, pauses entre le démonstratif et le nom, interruptions ou les combinaisons de ces phénomènes) avec

⁹³ Dans les deux exemples, il semble être question d'un emploi « émotionnel » à valeur « de rappel » ; voir la Section 3.5.3 pour la présentation de cette notion d'Adamec (1983).

les types et sous-types d'emploi et avec les positions du complexe démonstratif par rapport au verbe de la proposition où il figure. Les Tableaux 5.29 et 5.30 informent de la présence de disfluences avec les différents types et sous-types d'emploi, tout comme avec les différentes positions du démonstratif adnominal par rapport au verbe. Les graphiques des Figures 5.28 et 5.29 affichent les parts que les différents (sous-)types d'emploi et positions par rapport au verbe occupent au sein de l'ensemble des occurrences contenant des disfluences :

Type et sous-type d'emploi	Nombre d'occurrences avec des disfluences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
EXOPHORE MÉMORIELLE	27	9,47 %
EMPLOIS « EMOTIONNELS »	16	4,83 %
EMPLOIS « MÉTALINGUISTIQUES »	14	100 %
ENDOPHORE	10	4,63 %
Cataphore	5	7,94%
Anaphore fidèle	3	5,17 %
Anaphore associative	2	3,70 %
OCCURRENCES AMBIGUËS	9	11,11 %
EXOPHORE A-MÉMORIELLE	7	10,14 %
Deixis situationnelle	7	11,29 %
TOTAL	83	8,30 %

Tableau 5.29 : Proportion des occurrences du démonstratif adnominal affichant des disfluences pour chaque type et sous-type d'emploi dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

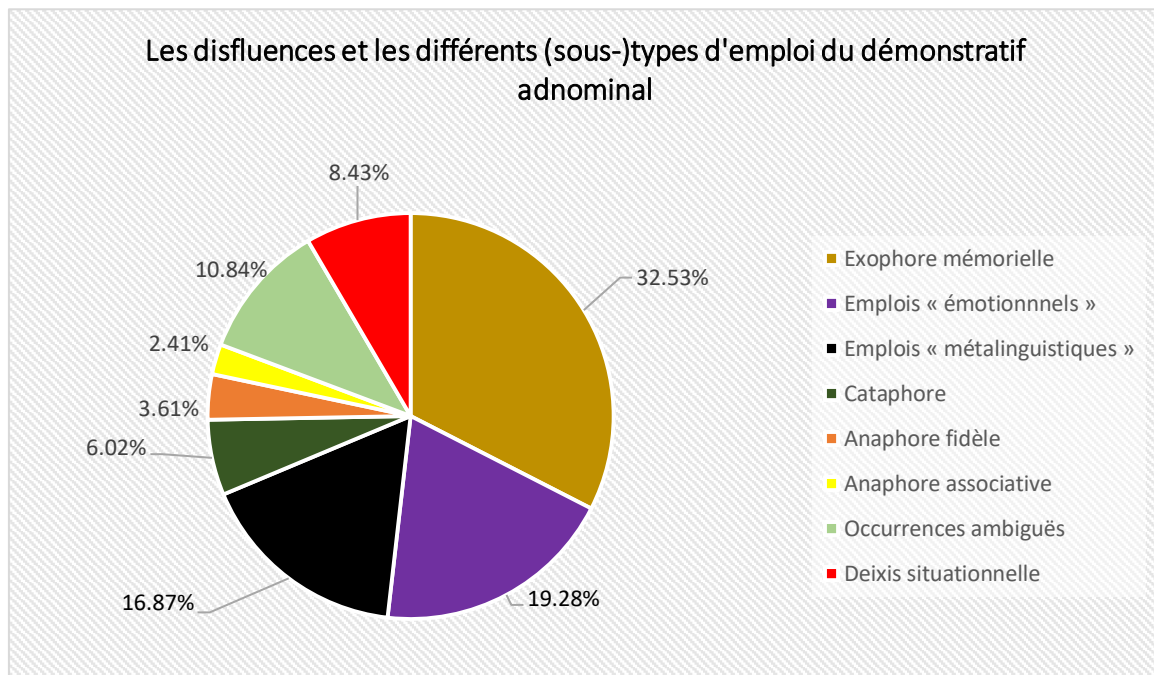


Figure 5.28 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal parmi les occurrences affichant des disfluences dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Position	Nombre d'occurrences avec des disfluences	Par rapport à l'ensemble des occurrences de la position en question (%)
POSITION POSTVERBALE	49	9,92 %
POSITION PREVERBALE	12	6,42 %
POSITION INCONNUE	12	11,54 %
DISLOCATION A DROITE	4	5,26 %
COMPLEMENT DE PHRASE	3	3,00 %
DISLOCATION A GAUCHE	2	5,13 %
TOTAL	83	8,30 %

Tableau 5.30 : Proportion des occurrences du démonstratif adnominal affichant des disfluences pour chaque position de l'occurrence par rapport au verbe dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

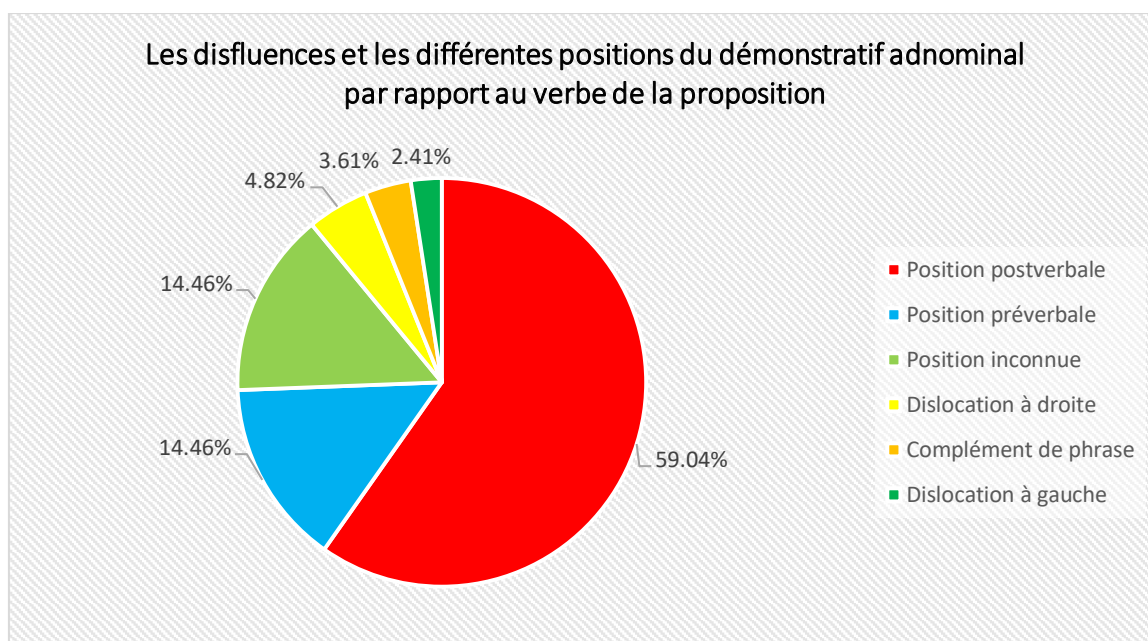


Figure 5.29 : Représentation en % des différentes positions du démonstratif adnominal par rapport au verbe parmi les occurrences affichant des disfluences dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Le Tableau 5.29 nous montre que 8,3 % de toutes les occurrences présentent des disfluences (soit un peu plus qu'en français, où le chiffre atteint 6,9 % ; voir la Section 5.2.2.3)⁹⁴. Ce sont, avant tout, les emplois « métalinguistiques » qui en contiennent : dans cette catégorie, toutes les occurrences sont marquées par des disfluences, constat qui est parfaitement en phase avec la nature de ces emplois, décrite dans la Section 3.3.3. **L'affinité entre les emplois « métalinguistiques » et la présence de disfluences** est illustrée au travers des résultats du calcul de spécificités, présentés par le graphique de la Figure 5.30 (les emplois « métalinguistiques » sont représentés par la septième colonne) :

⁹⁴ La différence n'est cependant pas significative sur le plan statistique : $L = 1,3974$; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

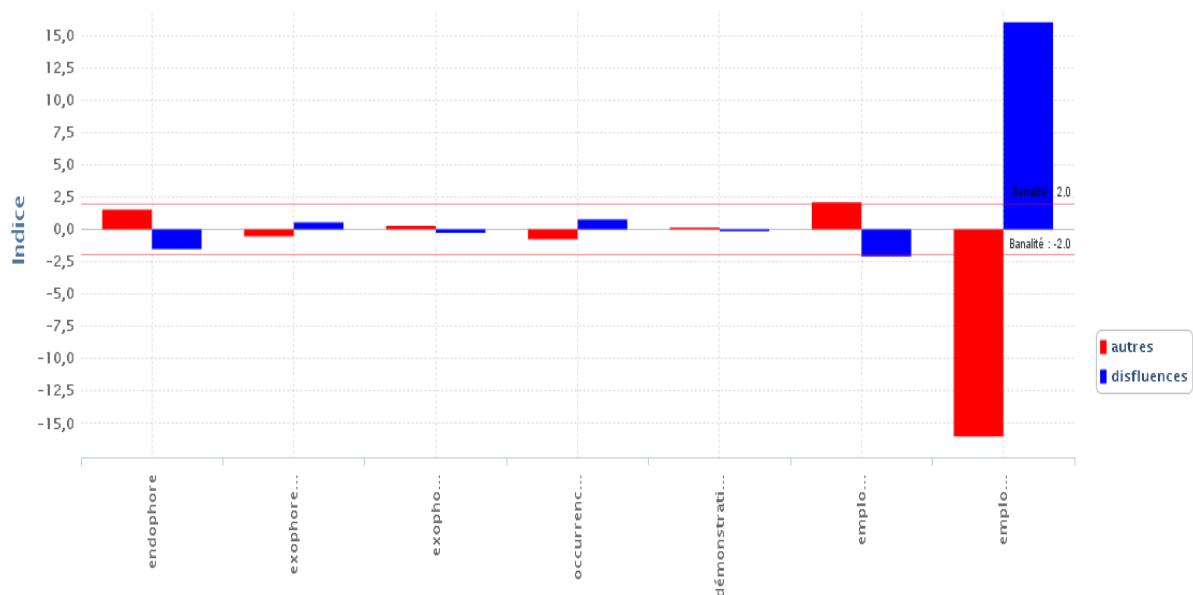


Figure 5.30 : Calcul de spécificités appliqué au lien entre les types d'emploi du déterminant démonstratif CE et la présence de disfluences dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Nous constatons ensuite que des disfluences accompagnent à peu près deux fois plus les emplois exophoriques (soit 9,6 % de la totalité de ces emplois, mémoriels et a-mémoriels confondus) et ambigus (11,11 % de la totalité de ces emplois) que les emplois « émotionnels » et endophoriques (soit, respectivement, 4,83 % et 4,63 % des ensemble de ces emplois)⁹⁵. Vu depuis l'ensemble des occurrences affichant des disfluences, l'exophore représente 40,96 % de ces dernières, suivie des emplois « émotionnels » (19,28 %) et « métalinguistiques » (16,87 %), les emplois endophoriques ne représentant que 12,05 %⁹⁶.

En ce qui concerne **la position par rapport au verbe**, ce sont la position postverbale, préverbale et inconnue qui reçoivent le plus d'occurrences contenant des disfluences (59,04 % pour la première et 14,46 % pour les deux autres)⁹⁷.

Nous allons maintenant procéder à une analyse plus détaillée des disfluences. Dans l'écrasante majorité des cas (79 occurrences, soit 91,18 % de l'ensemble), elles sont dues – de même qu'en français – à **des hésitations et/ou de légères reformulations du propos**. Font partie de cette sous-catégorie tous les emplois « métalinguistiques ». Dix-neuf occurrences relevant de cette catégorie (dont 4 emplois « métalinguistiques ») combinent disfluences et cumul d'éléments fonctionnels (cf. l'exemple 78) :

⁹⁵ S'agissant de l'exophore et de l'endophore, c'est le contraire qui se produit dans l'échantillon français, où des disfluences sont présentes avec 3,99 % des emplois exophoriques et 7,79 % des emplois endophoriques (voir la Section 5.2.2.3).

⁹⁶ Dans l'échantillon français, nous avons vu que les occurrences concernées étaient plus souvent de nature anaphorique (55,07 % de l'ensemble) que de nature exophorique (26,09 % de l'ensemble).

⁹⁷ Dans l'échantillon français, ce sont, avant tout, la position postverbale et celle du complément de phrase (voir la Section 5.2.2.3). Comme pour ce dernier échantillon, ici non plus, l'application du calcul de spécificités ne démontre aucune affinité entre une position en particulier et la présence de disfluences.

77) ale	jako	jak	vypad-a-j	ty .. ty	sluchátk-a
mais	MOD	comment	avoir l'air-PRS-3PL	ten.N.NOM.PL	écouteur-NOM.PL
jako	jak	to	vypad-á	to	vůbec
MOD	comment	cela.NOM	avoir l'air-PRS.3SG	cela.ACC	du tout

mais par contre à quoi ils ressemblent, les... les écouteurs ; genre à quoi ça ressemble, ça, je sais pas du tout
ORTOFON v1

78) to	je	takov-á	ta ta	misk-a	potravinov-á
ce être.PRS.3SG	tel.F.NOM.SG	ten.F.NOM.SG	bol-NOM.SG	alimentaire-F.NOM.SG	
<i>c'est cet espèce de bol alimentaire</i>					ORTOFON v1

79) no	a	Páj-e	kvet-e	ibišek	úplně	ve velk-ým ..
MOD	NP-F.DAT.SG	fleurir-PRS.3SG	hibiscus.NOM.SG	complètement	en grand-N.LOC.SG	
a	ten	můj	by	asi	potřebova-l	
et	ten.M.NOM.SG	mien.M.NOM.SG	AUX.COND.3SG	MOD	avoir besoin-PTCP.SG.M	
ostříhat	všechn-y	ty ..	všechn-y	větv-e		
couper.INF	tout-F.ACC.PL	ten.F.ACC.PL	tout-F.ACC.PL	branche-ACC.PL		

Pája, son hibiscus, il est tout en fleurs... ; et le mien, il faudrait lui couper toutes les... toutes les branches
ORTOFON v1

En 77, le SN est anaphorique : les locutrices s'entretiennent d'une paire d'écouteurs dont l'une s'apprête à faire cadeau à l'autre, et qui a déjà fait l'objet de plusieurs mentions antérieures. Le bol de 78 est un référent identifiable par la mémoire longue (« cet espèce de bol, tu vois ? »). Remarquons que le profil hésitant de l'occurrence se manifeste également par la postposition de l'adjectif, laquelle est marquée en tchèque. En 79, nous sommes face à un emploi « métalinguistique », la locutrice étant à la recherche du bon terme lexical, à savoir *větvě* (*branches*).

C'est sur les emplois « métalinguistiques » qu'il convient ici de s'arrêter. Comme nous l'avons écrit dans les Sections 2.2.3 et 3.3.3, ce type d'emploi fait le plus souvent apparaître une pause entre le démonstratif et le nom qu'il introduit, ainsi qu'une accentuation de la forme du démonstratif. Dans la Section 3.3.3, nous avons également signalé que le nom ainsi introduit n'est pas nécessairement de nature définie. Après avoir soigneusement étudié nos emplois « métalinguistiques », nous pouvons constater les faits suivants : 1) tous sauf un contiennent une pause entre TEN et l'élément qui suit ; 2) dans tous sauf un, TEN porte l'accent⁹⁸ ; 3) dans 9 cas, nous pouvons dire avec certitude que l'accent porté par TEN est un accent à intonation montante, créant un effet de suspense et d'attente (pour 2 occurrences, il semble difficile d'établir la courbe intonative, trop discrète, et, enfin, 3 occurrences

⁹⁸ Il s'agit ici de la même occurrence qui ne comporte pas de pause entre TEN et l'élément qui suit et où, semble-t-il, TEN n'est pas accentué.

n'ont pas d'accent montant⁹⁹) ; 4) la majorité des noms ainsi introduits est de nature indéfinie, sachant que deux d'entre eux font l'objet d'une quantification universelle avec VŠECHEN ; dans deux cas, le nom est employé de façon générique ; dans un cas, le nom est de nature définie non générique et dans un autre, une aposiopèse nous empêche d'en déterminer la nature. Ainsi, la différence principale avec les autres démonstratifs adnominaux accompagnés de disfluences – dont toutes les occurrences françaises – est qu'il s'agit ici avant tout de la recherche du terme lexical approprié. Ce terme peut accessoirement être un nom à sémantique (référentielle) définie, mais c'est loin d'être systématiquement le cas : en réalité, il semble être le plus souvent indéfini, et il peut même être en emploi non référentiel, comme en 80 :

80) vona	jí	ta	učitelk-a	řek-l-a	že
elle.NOM.SG	elle.DAT.SG	ten.F.NOM.SG	enseignante-NOM.SG	dire-PST-3SG.F	que
třeba jako	vypíš-e	<i>nějak-ý</i>	<i>ty .. ty</i>	<i>názv-y</i>	
MOD	proposer-FUT.3SG	quelque-M.ACC.PL	ten.M.ACC.PL	nom-ACC.PL	
<i>těch</i>	<i>bakalářek</i>				
ten.GEN.PL	mémoire de licence.GEN.PL				

genre la prof lui a dit qu'elle proposerait des... des sujets de mémoires de licence ORTOFON v1

En français, en revanche, lorsque le SN démonstratif s'accompagne de disfluences, il est toujours employé en emploi référentiel défini, que la quête du bon terme soit présente ou non (voir la Section 5.2.2.3).

Parmi tous ces énoncés, nous avons par ailleurs identifié 4 cas où le locuteur voit sa parole interrompue par son interlocuteur :

81) L1 : furt	ne-m-á-me	ještě [...]	vylišovan-ý	<i>ty ..</i>
toujours	NEG-avoir-PRS-1PL	encore	pressé-N.ACC.PL	ten.N.ACC.PL
L2 : čtrnáct	dnů			
quatorze	jour.GEN.PL			
L2 : sedmnáct	dnů ..	tak to	je	pohod-a
dix-sept	jour.GEN.PL	alors ce	être.PRS.3SG	sérénité-NOM.SG
L1 : <i>ty</i>	<i>originál-í</i>	<i>cédéčk-a ..</i>	to	ještě ne-m-á-me
ten.N.ACC.PL	original-N.ACC.PL	disque-ACC.PL	cela.ACC	encore NEG-avoir-PRS-1PL

L1 : on a toujours pas pressé les...

L2 : quatorze jours

L2 : dix-sept jours... du coup, on est tranquilles

L1 : les disques originaux, on les a toujours pas

ORTOFON v1

⁹⁹ Parmi celles-ci se trouve celle qui manque de pause et d'accentuation tout court.

Pour ce qui concerne les complexes nominaux démonstratifs accompagnés de disfluences et placés en dislocation, nous en recensons 4 en dislocation à droite et 2 en dislocation à gauche¹⁰⁰. Ces 6 emplois relèvent du premier cas de figure (disfluences dues à la présence d'hésitations et/ou de légères reformulations). L'identification du référent ne dépend pas du discours, sauf dans un cas (en dislocation à droite) :

- 82) [*Ten* .. *kostel*]_i [to]_i byl-a taky prác-e .. takov-á ..
 ten.M.NOM.SG église.NOM.SG ce être.PST-3SG.F aussi travail-NOM.SG tel-F.NOM.SG
- ne blb-á . takov-á .. hezk-á .. ale .. holt jsme
 non bête-F.NOM.SG tel-F.NOM.SG joli-F.NOM.SG mais MOD AUX.PST.1PL
- se .. háda-l-i
 REFL disputer-PTCP-PL.M

[cette... église]_i, [ça]_i aussi, [c']_i était un espèce de travail... genre... pas nul, plutôt chouette, au fond, mais on se disputait beaucoup à l'époque... ORTOFON v1

- 83) vono .. se [to]_i trošku měn-í [*to p* to počas-í* ..]_i
 cela.NOM REFL cela.NOM un peu change-PRS.3SG ten.N.NOM.SG temps-NOM.SG
- [ça]_i change un peu, [le t* le temps]_i* ORTOFON v1

En 82, nous sommes face à un cas d'identification situationnelle du référent (l'église est présente dans la situation d'énonciation). En 83, le démonstratif se prête à un emploi « émotionnel de rappel » (le référent a déjà été mentionné auparavant ; voir la Section 3.5.3) : il est question du temps au sens météorologique en général, sur l'ensemble de la planète.

À présent, nous compléterons ce tableau en abordant la situation du sous-corpus des quatre pièces de V. Havel.

¹⁰⁰ Vu la représentation trop faible de ces occurrences, il n'est pas possible d'en tirer une conclusion digne de ce nom.

5.2.4 Le sous-corpus des pièces de V. Havel (InterCorp v10)

Rappelons d’abord les propriétés de cet ensemble. Il contient 38 120 mots et 48 568 positions en tchèque et 52 626 mots et 64 540 positions en français. Rappelons également que le sous-corpus, ainsi que la motivation de son choix, sont présentés dans l’Introduction générale et dans la Section 2.1.3 du Chapitre 2. Au vu de ce qui a été constaté dans la Section 5.2.1 *supra* – une fréquence relative bien plus faible de TEN adnominal (soit 0,5 % de toutes les positions du corpus, contrairement à 1,30-1,35 % dans ORTOFON v1) – nous pouvons nous attendre à ce que la langue de cet ensemble soit plus éloignée de la langue parlée informelle.

5.2.4.1 Premiers constats

Le Tableau 5.31 recense les types et sous-types d’emploi du démonstratif adnominal dans le sous-corpus des pièces de V. Havel (241 occurrences au total) avec, pour chaque type et sous-type, l’intervalle de confiance qui anticipe sa fréquence dans un autre échantillon hypothétique de 241 occurrences¹⁰¹. Le graphique de la Figure 5.31 représente les proportions occupées, au sein de l’échantillon, par les différents (sous-)types d’emploi¹⁰².

Type et sous-type d’emploi	Nombre d’occurrences	Par rapport à l’ensemble des 241 occurrences (%)	Intervalle de confiance	
			Borne inférieure	Borne supérieure
ENDOPHORE	102	42,32 %	87	117
Anaphore résomptive	47	19,50 %	35	59
Anaphore fidèle	28	11,62 %	19	38
Cataphore	12	4,98 %	6	19
Anaphore infidèle	11	4,56 %	5	18
Anaphore indirecte	2	0,83 %	0	5
Anaphore associative	1	0,41 %	0	3
Anaphore temporelle	1	0,41 %	0	3
EMPLOIS « EMOTIONNELS »	50	20,75 %	38	63
EXOPHORE MEMORIELLE	42	17,43 %	31	54
EXOPHORE A-MEMORIELLE	36	14,94 %	26	47
Deixis situationnelle	27	11,20 %	18	37
Deixis temporelle	8	3,32 %	3	14
Deixis textuelle	1	0,41 %	0	3
OCCURRENCES AMBIGUES	8	3,32 %	3	14
EMPLOIS « METALINGUISTIQUES »	2	0,83 %	0	5
DEMONSTRATIFS INDEFINIS	1	0,41 %	0	3

¹⁰¹ En écartant le bruit. La formulation d’« échantillon hypothétique » se justifie évidemment par le fait que notre échantillon présent englobe l’ensemble de la liste des concordances valides.

¹⁰² Il reprend les types d’emploi là où ceux-ci ne sont pas divisés en sous-types ; dans le cas contraire, il reprend les sous-types d’emploi.

TOTAL	241	100,00 %

Tableau 5.31 : Types et sous-types d'emploi du démonstratif adnominal tchèque (tous lemmes confondus) dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences) avec les intervalles de confiance anticipant leur fréquence dans un autre échantillon de la même taille

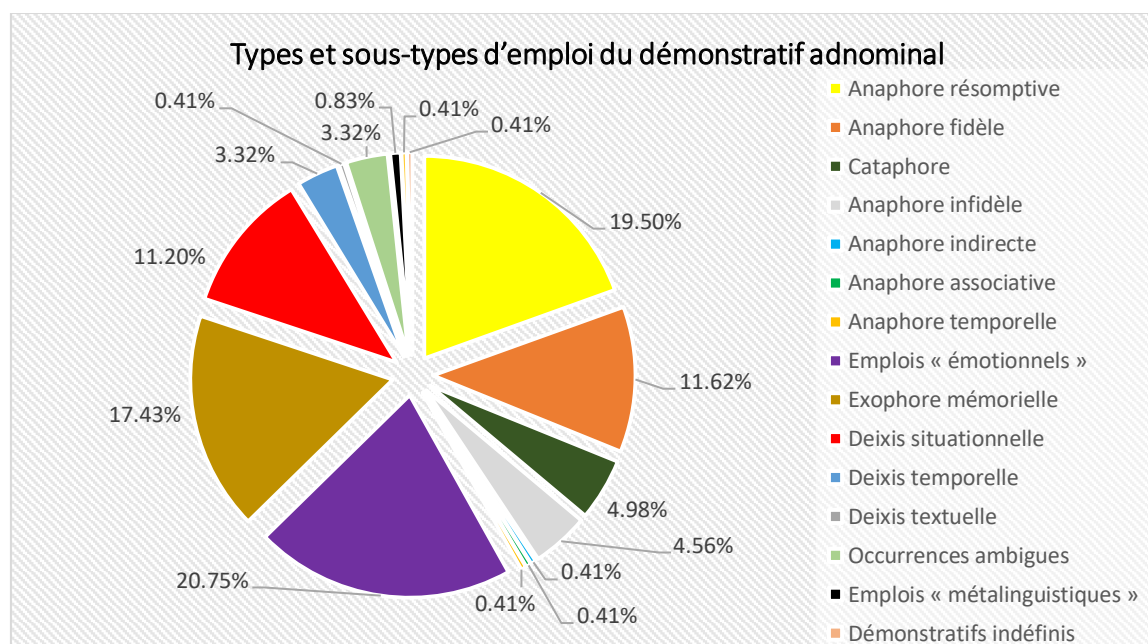


Figure 5.31 : Types et sous-types d'emploi du démonstratif adnominal (tous lemmes confondus) dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Le Tableau 5.31 montre que la majorité des occurrences du démonstratif adnominal se prête à l'emploi endophorique (102, soit 42,32 % de la totalité). Dans cet ensemble, c'est l'anaphore résomptive qui prime, avec 47 occurrences (19,52 % du tout). Les emplois « émotionnels » suivent (50 occurrences, soit 20,75 % du tout), puis les emplois mémoriels (42 occurrences, soit 17,43 % du tout). L'exophore a-mémorielle est représentée par 36 occurrences (14,94 % du tout), dont 27 sont situationnelles (11,20 % du tout). Si nous comparons ces données avec celles obtenues pour ORTOFON v1, plusieurs différences nous frappent aussitôt : le sous-corpus des pièces est plus riche en démonstratifs endophoriques (42,32 % de l'ensemble contre 28,5 % de l'ensemble pour ORTOFON v1) et exophoriques a-mémoriels (14,94 % de l'ensemble contre seulement 6,9 % de l'ensemble pour ORTOFON v1). En revanche, il est plus pauvre en emplois « émotionnels » (20,75 % de l'ensemble contre 33,1 % de l'ensemble pour ORTOFON v1) et mémoriels (17,43 % de l'ensemble contre 28,5 % de l'ensemble). La composition des emplois endophoriques des deux sources est également assez différente : au sein d'ORTOFON v1, 29,17 % d'entre eux sont de nature cataphorique, contre seulement 11,76 % dans le sous-corpus des pièces de V. Havel. Les pièces ne contiennent qu'une seule anaphore associative (soit 0,41 % de l'ensemble), alors qu'ORTOFON v1 en contient 54 (soit 5,4 % de l'ensemble). En revanche, 46,08 % de tous les emplois endophoriques

de ce sous-corpus sont résomptifs, contre seulement 5,56 % pour ORTOFON v1¹⁰³ ! Vu la taille assez limitée de l'échantillon (moins d'un quart de celle des échantillons de conversations), la définition de l'intervalle de confiance s'avère difficile, et bien des (sous-)types d'emploi affichent la valeur 0.

Il semble aller de soi qu'une partie de ces divergences est directement liée aux différences de nature entre les deux types de textes. Ainsi, la présence plus haute de la référence exophorique a-mémorielle et la présence plus basse de la référence exophorique mémorielle dans notre sous-corpus des pièces suggèrent que ces dernières recourent davantage, pour identifier les référents, au cadre spatio-temporel de l'action en cours. La motivation de cela est, en vérité, tout à fait compréhensible : une inégalité d'accès à ces référents entre, d'une part, les personnages et, d'autre part, le lecteur/spectateur. Dans les conversations, c'est l'inverse qui se produit, car les référents sont également accessibles à toutes les parties, les informations menant à leur identification pouvant donc davantage rester implicites. De plus, cette identification est aussi plus redevable à des connaissances et souvenirs intimes, cachées à une tierce personne. La plus faible présence d'emplois « émotionnels » semble conforme à l'attente formulée *supra*, soit un profil moins informel de la langue des pièces : à l'opposé d'ORTOFON v1, contenant d'authentiques productions orales informelles, les dialogues des pièces ne font qu'imiter ce registre. La proportion de l'anaphore résomptive dans le sous-corpus est plus proche de celle constatée pour l'échantillon du CEFC (soit 34,76 % de toutes les occurrences endophoriques ; voir la Section 5.2.2.1). Deux explications majeures nous semblent ici pertinentes : 1) une parole davantage structurée des pièces, due à leur caractère plus formel et 2) le facteur relatif à la place occupée par certains autres (sous-)types d'emploi dans ORTOFON v1 (notamment les types mémoriel et « émotionnel »), déjà mentionné *supra*, et qui pourrait faire en sorte que l'emploi anaphorique résomptif est sous-représenté dans notre échantillon.

Le Tableau 5.32 représente **la fréquence des démonstratifs autres que TEN** par sous-types d'emploi :

¹⁰³ L'application de la fonction de vraisemblance montre la significativité statistique de la différence de fréquences dans le domaine de l'exophore a-mémorielle ($L = 14,1802$), de l'exophore mémorielle ($L = 13,1274$), de l'endophore ($L = 40,4967$), des emplois « émotionnels » ($L = 14,7485$), de l'anaphore résomptive ($L = 106,7604$) et de l'anaphore associative ($L = 17,1002$) ; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

Lemme	Nombre d'occurrences
TENTO	41
Anaphore résomptive	19
Deixis temporelle	8
Deixis situationnelle	7
Anaphore fidèle	5
Deixis textuelle	1
Occurrences ambiguës	1
TENHLE	20
Anaphore résomptive	11
Deixis situationnelle	8
Anaphore indirecte	1
ONEN	1
Démonstratif indéfini	1
TÁMHLE TEN	1
Deixis situationnelle	1
TOTAL	63

Tableau 5.32 : Présence des lemmes du démonstratif tchèque en emploi adnominal autres que TEN dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Le tableau indique que la fréquence des démonstratifs autres que TEN atteint 63 occurrences, soit 26,14 % du total. Ce chiffre est beaucoup plus élevé que celui de l'échantillon des 1 000 occurrences d'ORTOFON v1 (33 occurrences, soit 3,3 % de l'ensemble). Cette disproportion remet partiellement en cause la comparabilité des deux sources, ces formes – notamment TENTO – étant largement minoritaires dans l'échantillon de tchèque oral informel (voir la Section 5.2.3.1). D'autre part, ce chiffre est en grande partie dû à la pièce *La grande roue*, dont la langue est, de manière générale, bien plus formelle que celle des autres pièces. Ainsi, le texte de cette pièce contient à lui seul 30 des 41 occurrences de TENTO (soit 73,17 %). C'est la raison pour laquelle cette forme est, après TEN, la plus représentée dans le sous-corpus. Elle est suivie de TENHLE, avec 20 occurrences. Le sous-corpus contient également une occurrence de ONEN, en emploi indéfini. Quant à TENTO, il compte 24 occurrences anaphoriques (essentiellement résomptives) et 16 occurrences déictiques (dont 7 situationnelles et 8 temporelles). TENHLE compte 12 occurrences anaphoriques (essentiellement résomptives) et 8 occurrences déictiques situationnelles. TENTO et TENHLE sont donc avant tout valorisés dans deux cadres, à savoir celui de l'anaphore résomptive, où il est question de créer un nouveau référent via la synthèse d'un contenu discursif antérieur (30 occurrences au total, soit 49,18 % de toutes les occurrences de ces deux formes) et dans celui de l'exophore a-mémorielle, référant à des entités de la situation d'énonciation (15 occurrences, soit 24,59 % de l'ensemble) ou à des entités temporelles (8 occurrences, soit 13,11 % de l'ensemble). Le sous-corpus contient également une occurrence de *támhle ten*, dont les deux mots, bien que graphiés séparément, fonctionnent sans doute comme la forme fusionnée du lemme TAMHLETEN.

Le recours à l'interface « many features – 1 sample » génère les intervalles de confiance suivants pour TEN et l'ensemble des autres lemmes¹⁰⁴ :

Lemme	Nombre d'occurrences	Intervalle de confiance	
		Borne inférieure	Borne supérieure
TEN	178	164	191
Les autres lemmes	63	50	77

Tableau 5.33 : Fréquence absolue du lemme TEN et celle des autres lemmes du démonstratif en emploi adnominal dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

À titre de comparaison, le lemme TENTO (emplois adnominaux et pronominaux confondus) atteint la fréquence relative de 1 091,3 occurrences pour un million de positions dans notre sous-corpus de V. Havel, 399,5 occurrences dans ORTOFON v1 et 430,3 occurrences dans notre sous-corpus de théâtre de référence. Pour TENHLE, les fréquences relatives, présentées dans le même ordre, sont les suivantes : 947,1, 999,6 et 1 735,8. En se basant uniquement sur ces différences – dont 3 sont significatives sur le plan statistique lorsque nous appliquons la fonction de vraisemblance aux fréquences absolues des deux lemmes¹⁰⁵ –, il est possible de conclure à un ton plus formel des pièces de V. Havel, et ce non seulement par comparaison avec l'oral informel, mais aussi par comparaison avec le genre des pièces de théâtre dans son intégralité¹⁰⁶.

5.2.4.2 La grammaticalisation de TEN adnominal

De même que nous l'avons fait pour l'échantillon d'ORTOFON v1, nous allons présenter le bilan des occurrences de TEN adnominal qui ont été évaluées comme étant engagées dans un processus de grammaticalisation¹⁰⁷. Le nombre d'occurrences considérées comme en voie de grammaticalisation est comparé au nombre total d'occurrences de TEN, à savoir 178. Le Tableau 5.34 tient compte des différents types et sous-types d'emploi et renseigne sur le pourcentage des occurrences considérées comme en voie de grammaticalisation au sein de l'ensemble des occurrences

¹⁰⁴ $\alpha = 0,05$. Ces intervalles anticipent les fréquences du lemme TEN et des autres lemmes dans un autre échantillon hypothétique de la même taille.

¹⁰⁵ Pour la comparaison entre le sous-corpus de V. Havel et le sous-corpus de théâtre de référence, $L = 32,7209$ pour TENTO et 20,2734 pour TENHLE ; pour la comparaison entre le sous-corpus de V. Havel et ORTOFON v1, $L = 37,1902$ pour TENTO et 0,1312 pour TENHLE ; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

¹⁰⁶ Ajoutons cependant qu'en partant toujours de ce même critère, il serait moins formel que celui des parties « non-fiction » (fréquence relative de TENTO/TENHLE : 4 559,9 / 191,1) et « presse » (fréquence relative de TENTO/TENHLE : 2 091,3 / 413,8) de SYN2015.

¹⁰⁷ Rappelons que les critères de cette évaluation seront détaillés dans le Chapitre 6 pour chacun des (sous-)type d'emploi concernés.

appartenant à chacun de ces derniers. Le graphique de la Figure 5.32 affiche la part que les différents (sous-)types occupent au sein du groupe des occurrences en voie de grammaticalisation :

GRAMMATICALISATION DE TEN ADNOMINAL

Type et sous-type d'emploi	Occurrences grammaticalisées	Nombre total d'occurrences du (sous-) type en question	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
EXOPHORE MEMORIELLE	32	42	76,19 %
ENDOPHORE	27	66	40,91 %
Cataphore	12	12	100,00 %
Anaphore fidèle	11	23	47,83 %
Anaphore résomptive	1	17	5,88 %
Anaphore infidèle	1	11	9,09 %
Anaphore associative	1	1	100,00 %
Anaphore indirecte	1	1	100,00 %
Anaphore temporelle	0	1	0,00 %
EXOPHORE A-MÉMORIELLE	7	11	63,64 %
Deixis situationnelle	7	11	63,64 %
EMPLOIS « EMOTIONNELS »	0	50	0,00 %
EMPLOIS METALINGUISTIQUES	0	2	0,00 %
OCCURRENCES AMBIGUES	0	7	0,00 %
TOTAL	66	178	37,08 %

Tableau 5.34 : Proportion des occurrences de TEN adnominal évaluées comme engagées dans le processus de grammaticalisation pour chaque type et sous-type d'emploi dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

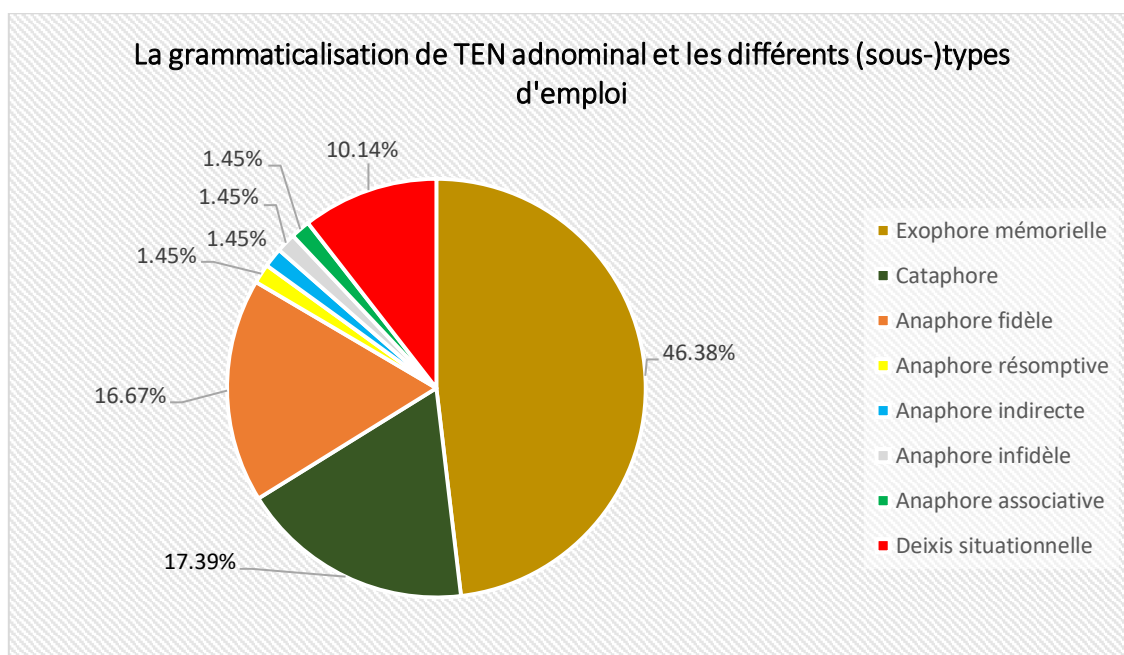


Figure 5.32 : Représentation en % des différents (sous-)types d'emploi parmi les occurrences de TEN adnominal encourageant le processus de grammaticalisation dans les pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Le Tableau 5.34 suggère que la grammaticalisation touche surtout les occurrences exophoriques mémorielles (76,19 % de l'ensemble). Viennent ensuite l'exophore a-mémorielle (63,64 % de l'ensemble) et l'endophore (40,91 % de l'ensemble). Dans ce dernier cas, la grammaticalisation est le plus présente dans le domaine de la cataphore (100 % ; toutes les occurrences relèvent de la cataphore intra-syntagmatique). La grammaticalisation accompagne également une partie considérable des anaphores fidèles (47,83 %). Dans le domaine de l'exophore a-mémorielle, seule la deixis situationnelle est concernée, bilan qui reflète celui d'ORTOFON v1 (TEN n'apparaît pas avec la deixis temporelle). Nous avons déjà vu pourquoi la grammaticalisation ne peut se combiner avec les emplois « émotionnels » et « métalinguistiques » et pourquoi son évaluation s'avère très problématique pour les occurrences ambiguës. Somme toute, 37,08 % de toutes les occurrences de TEN adnominal seraient concernées par la grammaticalisation dans le sous-corpus, sachant que 46,38 % d'entre elles ressortissent à l'exophore mémorielle et 40,91 % ressortissent à l'endophore.

Le chiffre total est plus bas que celui obtenu pour ORTOFON v1 (46,54 % de toutes les occurrences de TEN)¹⁰⁸. Nous remarquons également qu'au sein des deux principaux types d'emploi concernés par la grammaticalisation (exophorique mémoriel et endophorique), la proportion des occurrences de TEN en voie de grammaticalisation est plus basse dans ce sous-corpus que dans l'échantillon d'ORTOFON v1 (76,19 % contre 86,27 % de toutes les occurrences mémorielles et 40,91 % contre 72,82 % de toutes les occurrences endophoriques). Vu que les deux différences sont

¹⁰⁸ Malheureusement, la différence n'est pas statistiquement significative : $L = 3,0853$; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

statistiquement significatives – $L = 6,132$ pour la première et $20,8885$ pour la seconde¹⁰⁹ – il se peut que cela soit (en partie) dû au caractère davantage informel des conversations, caractère qui se manifesterait, entre autres¹¹⁰, par un taux de grammaticalisation plus élevé de TEN adnominal. Quant à l'endophrase, la différence est avant tout due à la présence élevée de l'anaphore associative dans les conversations. Or comme nous le verrons dans le Chapitre 6, l'emploi de TEN avec certains types de cette dernière semble propre, justement, au tchèque parlé informel.

Le graphique de la Figure 5.33, visualisant les résultats du calcul de spécificités, met en valeur **les liens entre la grammaticalisation de TEN et la cataphore** (sixième colonne), **ainsi que la grammaticalisation de TEN et les emplois mémoriels** (huitième colonne) :

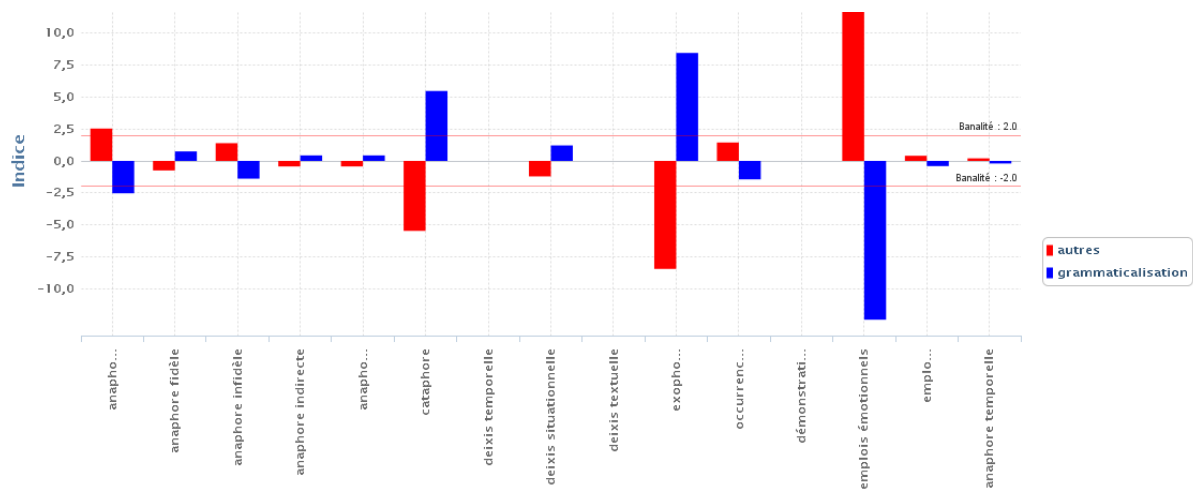


Figure 5.33 : Calcul de spécificités appliqué au lien entre les différents (sous-)types d'emploi et les occurrences de TEN adnominal encourageant le processus de grammaticalisation dans les pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

5.2.4.3 Les propriétés combinatoires du démonstratif adnominal avec d'autres mots fonctionnels

Le Tableau 5.35 fait le bilan des cumuls du démonstratif adnominal avec d'autres éléments fonctionnels (voir la Section 5.2.3.3 pour les éléments fonctionnels qui ont été inclus dans le calcul et ceux qui en ont été exclus) tout en croisant la présence de ces cumuls avec les différents types et sous-types d'emploi. Le graphique de la Figure 5.34 affiche la part que les différents (sous-)types occupent au sein du groupe des occurrences comportant des cumuls :

¹⁰⁹ La valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

¹¹⁰ Cf. aussi l'emploi plus élevé de TENTO et l'occurrence de ONEN ainsi qu'une présence plus basse d'emplois « émotionnels » (voir la Section 5.2.4.1).

CUMULS

Type et sous-type d'emploi	Nombre d'occurrences	Par rapport à l'ensemble des occurrences du (sous-)type en question (%)
EMPLOIS « EMOTIONNELS »	10	20,00 %
EXOPHORE MEMORIELLE	5	11,90 %
ENDOPHORE	5	7,69 %
Anaphore fidèle	3	13,04 %
Anaphore résomptive	2	11,76 %
OCCURRENCES AMBIGUES	2	28,57 %
TOTAL	22	9,13%

Tableau 5.35 : Proportion des occurrences du démonstratif adnominal se cumulant avec un autre mot fonctionnel pour chaque type et sous-type d'emploi dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

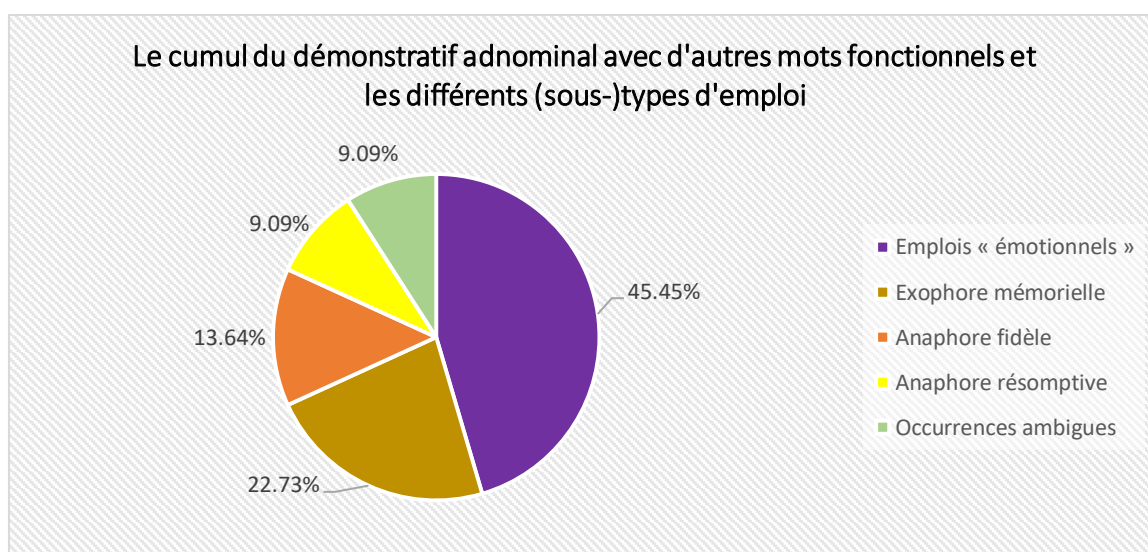


Figure 5.34 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal parmi les occurrences où le démonstratif se cumule avec un autre mot fonctionnel dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Quels sont les éléments fonctionnels se combinant avec le démonstratif au sein d'un même complexe nominal ? La réponse à cette question est apportée par le Tableau 5.36 :

Cumul	Type ou sous-type d'emploi	Nombre d'occurrences
TEN + possessif	Emplois « émotionnels »	4
	Anaphore fidèle	3
	Exophore mémorielle	2
	Occurrences ambiguës	2
	Anaphore résomptive	1
	Total	12
NĚJAKÝ TEN	Emplois « émotionnels »	5
<i>Těch několik</i>¹¹¹	Exophore mémorielle	2

¹¹¹ NĚKOLIK = plusieurs.

TAKOVÝ TEN	Exophore mémorielle	1
TENTO + Possessif	Anaphore résomptive	1
Trochu¹¹² toho	Emplois « émotionnels »	1
TOTAL		22

Tableau 5.36 : Différents types de cumuls du démonstratif adnominal avec un autre mot fonctionnel et les différents types et sous-types d'emploi avec lesquels ils apparaissent dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Dans le sous-corpus des pièces de V. Havel, le cumul le plus représenté est celui de **TEN se combinant avec un élément fonctionnel possessif** (12 occurrences, soit 54,54 % de l'ensemble). Il est suivi de la combinaison **NĚJAKÝ TEN**, dont les 5 occurrences (22,72 % de l'ensemble) se prêtent à des emplois « émotionnels ». Au total, le nombre de cumuls atteint, dans le sous-corpus, le chiffre de 22, soit 9,13 % de toutes les occurrences du démonstratif adnominal. Dans l'échantillon d'ORTOFON v1, il s'agit de 7 %, sachant que la combinaison qui y figure le plus est **TAKOVÝ TEN** (23 occurrences, soit 32,86 % de l'ensemble, contrairement à une seule occurrence dans le sous-corpus des pièces), suivie de **TEN + possessif** (27,14 % de l'ensemble) et de **NĚJAKÝ TEN** (18 occurrences, soit 25,71 % de l'ensemble). Nous notons également que les deux cumuls ne sont pas employés tout à fait de la même façon dans les deux ensembles de données : dans les conversations, le premier affiche 12 occurrences « émotionnelles » (soit 63,16 % de la totalité) et le second, 7 occurrences « émotionnelles » (soit 38,89 % de la totalité), tandis que dans le sous-corpus, le premier n'affiche que 4 emplois « émotionnels » (33,33 % de l'ensemble) et le second, 5 emplois émotionnels, soit la totalité.

À titre de comparaison, nous avons regardé **la fréquence relative des trois cumuls suivants** dans le sous-corpus de V. Havel, le corpus de théâtre de référence, ORTOFON v1 et la partie « fiction » du corpus SYN2015 : 1) **TAKOVÝ TEN**, 2) démonstratif (hormis **TAKOVÝ**) + possessif et 3) **NĚJAKÝ TEN**¹¹³. Ces fréquences sont reprises par le Tableau 5.37 :

	Le sous-corpus de V. Havel	Le corpus de théâtre de référence	ORTOFON v1	SYN2015 – fiction
TAKOVÝ (TAKÝ/ TAKOVÝHLE) TEN	20,6	32,3	568,5	22,5
démonstratif (sauf TAKOVÝ) + possessif	350	635,5	397,1	327,1
NĚJAKÝ TEN	103	29	279	18,6

¹¹² TROCHU = *un peu de*.

¹¹³ [lemma= "TAKOVÝ|TAKÝ|TAKOVÝHLE"] [lemma= "ten"], [tag= "PD.....-----.*" & !lemma= "takový"] [tag= "PS.....-----.*"] et [lemma= "nějaký"] [lemma= "ten"]. Les chiffres obtenus sont arrondis au dixième.

Tableau 5.37 : Fréquence relative (nombre d’occurrences pour un million de positions) des cumuls TAKOVÝ TEN, « démonstratif (sauf TAKOVÝ) + possessif » et NĚJAKÝ TEN dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 48 568 positions), le corpus de théâtre de référence (CNT ; 1 301 447 de positions), ORTOFON v1 (CNT ; 1 236 508 de positions) et la partie « fiction » de SYN2015 (CNT ; 41 619 587 de positions)

En partant des analyses d’Uhlířová (1992) ainsi que de notre propre sentiment de la langue, lequel nous dit que l’emploi de cumuls contenant le démonstratif va croissant, en tchèque, avec l’oralité, ce petit relevé conforte globalement le constat déjà formulé dans les Sections 5.2.4.1 et 5.2.4.2 *supra*, à savoir que le ton des pièces de V. Havel concernées est globalement plus formel que ceux d’ORTOFON v1 et du corpus de théâtre de référence (seule la combinaison NĚJAKÝ TEN est plus fréquente dans le sous-corpus des pièces que dans le corpus de théâtre de référence)¹¹⁴.

5.2.4.4 La transposition du démonstratif adnominal dans les traductions françaises

Nous ferons ici le bilan des différentes traductions du démonstratif adnominal tchèque¹¹⁵. Nous renvoyons le lecteur au Tableau 2.9, situé à la fin de la Section 2.2.3, pour les types de transposition utilisés. Voici, tout d’abord, un bilan très général, qui fera, par la suite, l’objet d’analyses plus ciblées :

Type de transposition	Nombre d’occurrences	Par rapport à l’ensemble des 241 occurrences (%)
Déterminant démonstratif	83	34,44 %
Article défini	53	21,99 %
Absence de transposition (traduction libre)	42	17,43 %
Autre transposition (déterminant quantificateur, adverbe, <i>de</i> d’extensité nulle, déterminant exclamatif, absence de déterminant, pronom personnel, etc.)	33	13,69 %
Déterminant possessif	25	10,37 %
Article indéfini	5	2,07 %
TOTAL	241	100,00 %

Tableau 5.38 : Différents types de transposition du démonstratif adnominal tchèque dans les traductions françaises du sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT)

¹¹⁴ La différence en termes de fréquences absolues est statistiquement significative pour TAKOVÝ TEN chez V. Havel et dans ORTOFON v1 ($L = 44,805$), pour « démonstratif (sauf TAKOVÝ) + possessif » chez V. Havel et dans le corpus de théâtre de référence ($L = 7,2283$) ainsi que chez V. Havel et dans la partie « fiction » de SYN2015 ($L = 7,2283$), pour NĚJAKÝ TEN chez V. Havel et dans ORTOFON v1 ($L = 6,9277$), chez V. Havel et dans le corpus de théâtre de référence ($L = 5,1216$) et, enfin, chez V. Havel et la partie « fiction » de SYN2015 ($L = 5,1216$) ; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

¹¹⁵ Afin de dissiper tout malentendu, toutes les occurrences ont été incluses dans le bilan, même celles où le démonstratif se combine avec un autre élément fonctionnel.

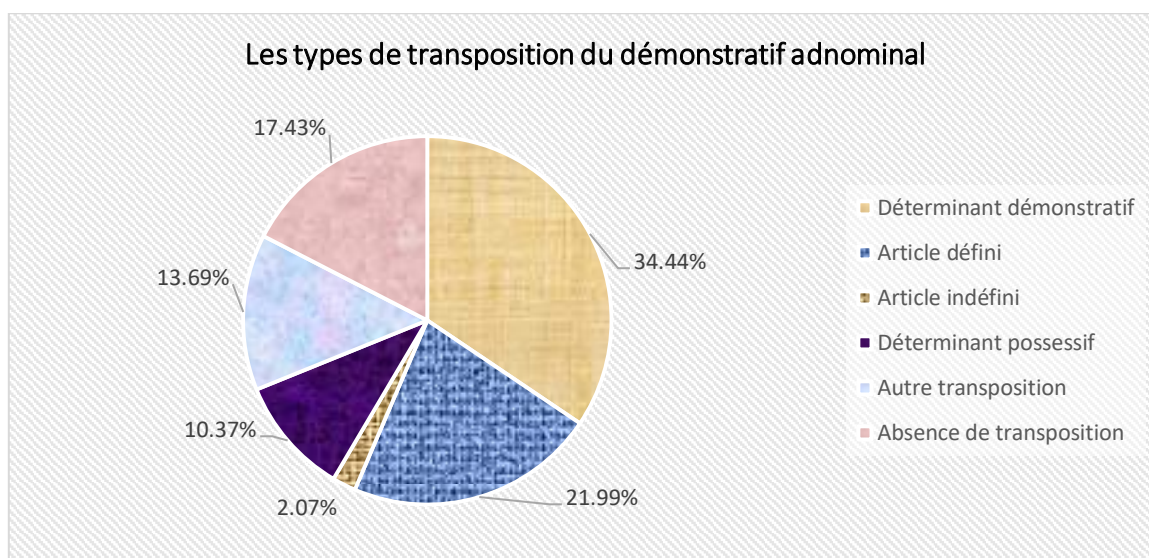


Figure 5.35 : Différents types de transposition du démonstratif adnominal tchèque dans les traductions françaises du sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT)

Nous voyons ainsi que dans la majorité des cas, le démonstratif adnominal est rendu en français par le déterminant démonstratif (34,44 %) ¹¹⁶. Ce dernier est suivi de l'article défini (21,99 %). Dans l'ensemble de cette analyse, il nous semble pertinent de partir du présupposé que c'est précisément à celui-ci que les traducteurs recourront le plus spontanément pour rendre les occurrences de TEN jugées grammaticalisées (voir à cet égard notamment la Section 3.3 du Chapitre 3). Dans 10,37 % des cas, les traducteurs ont opté pour un déterminant possessif. L'article indéfini, quant à lui, n'est que très faiblement représenté (2,07 %). Un nombre considérable des occurrences du démonstratif tchèque (13,69 %) a été transposé par un autre moyen que ceux qui viennent d'être mentionnés. Il reste une partie encore plus importante des occurrences qui n'a pas été transposée du tout, le traducteur ayant opté pour une traduction libre (17,43 %). Voici un exemple pour chacun de ces cas de figure, présentés dans l'ordre où ils figurent dans le Tableau 5.38 (nous citons d'abord l'original, puis sa traduction) :

- 84) Podchytit i **tyhle věci** patří přece, hergot, tak nějak k práci s člověkem!
Ça fait partie, nom d'un chien, du travail, voyons, du travail sur l'homme, d'encourager ces choses-là ! V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10
- 85) Poslyš, ale ted' upřímně: to myslíš vážně s **tím pivovarem**?
Ecoute, parlons franchement et sans détour : ton job à la brasserie, ce n'est pas sérieux ! ¹¹⁷
V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

¹¹⁶ Dans l'ensemble des traductions des pièces, la fréquence relative globale du déterminant démonstratif CE est à peu près deux fois supérieure à celle du déterminant démonstratif CE traduisant un démonstratif adnominal tchèque. Il résulte de cela qu'à peu près la moitié de toutes les occurrences de CE ne sont pas issues de la transposition du démonstratif adnominal tchèque.

¹¹⁷ Nous sommes ici face à un procédé de traduction appelé « explication » (cf. Vinay et Darbelnet, 1960) : *la brasserie* → *le job à la brasserie*.

86) Nikdy jsem tomu nevěřil, ale teď vidím, jak **to dítě** dává člověku úplně novou perspektivu [...] ¹¹⁸
Je l'ignorais autrefois, mais maintenant j'en suis sûr. Un enfant nous donne une perspective plus juste [...] V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

87) Nevím jak ty, Willy, ale já dobře věděl – když jsem se na **tohle řemeslo** dával – že mě čeká jen nevděk!

Toi, je ne sais pas, Tony, mais moi, quand j'ai commencé dans mon métier, je savais ce qui m'attendait : ... l'ingratitude générale !

V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

88) No bóže, **trochu toho deště** na zpáteční cestě!

Oh, mon Dieu, un peu de pluie au retour ! ¹¹⁹

V. Havel, *La fête en plein air*, 1963, trad. par F. Kérel, 1964, InterCorp v10

89) A ty myslíš, že teď – když je Macheath zase tady – neexistuje už opravdu žádná cesta, jak z něho **ty informace** dostat?

/ littéralement : est-ce que vraiment il n'existe plus aucun moyen de lui soutirer ces informations ?

Tu ne crois pas qu'il y a un moyen de faire parler Maxence maintenant qu'il est repris ?

V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

Croisons à présent les types de transposition du démonstratif adnominal avec le critère des différents types et sous-types d'emploi. Dans les cas où la différence existe, nous ferons également la part entre les occurrences en voie de grammaticalisation et celles qui ne le sont pas :

DETERMINANT DEMONSTRATIF

Type et sous-type d'emploi	Nombre d'occurrences	Dont les occurrences grammaticalisées	La proportion des occurrences grammaticalisées (%)
EXOPHORE A-MEMORIELLE	18	3	16,67 %
Deixis situationnelle	17	3	17,65 %
Deixis textuelle	1	0	0,00 %
EXOPHORE MEMORIELLE	16	12	75,00 %
ENDOPHORE	36	2	5,56 %
Anaphore résomptive	22	0	0,00 %
Anaphore fidèle	6	0	0,00 %
Anaphore infidèle	6	0	0,00 %
Cataphore	2	2	100,00 %
EMPLOIS « ÉMOTIONNELS »	11	0	0,00 %
OCCURRENCES AMBIGUES	2	0	0,00 %
TOTAL	83	17	20,48 %

Tableau 5.39 : Différents (sous-)types d'emploi des occurrences du démonstratif adnominal tchèque dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT) rendues, en français, par le déterminant démonstratif (InterCorp v1, CNT ; 241 occurrences)

¹¹⁸ *To dítě* est ici employé de manière générique. Le traducteur maintient cette généricité, mais il l'exprime au moyen de l'indéfini.

¹¹⁹ De la combinaison de TROCHU (*un peu de*) avec TEN, où ce dernier est doté d'une fonction « émotionnelle », seul *un peu de* est gardé en français.

ARTICLE DEFINI

Type et sous-type d'emploi	Nombre d'occurrences	Dont les occurrences grammaticalisées	La proportion des occurrences grammaticalisées (%)
EXOPHORE A-MEMORIELLE	5	4	80,00 %
Deixis situationnelle	4	4	100,00 %
EXOPHORE MEMORIELLE	14	14	100,00 %
ENDOPHORE	18	18	100,00 %
Anaphore fidèle	7	7	100,00 %
Cataphore	7	7	100,00 %
Anaphore résomptive	1	1	100,00 %
Anaphore associative	1	1	100,00 %
Anaphore infidèle	1	1	100,00 %
Anaphore indirecte	1	1	100,00 %
EMPLOIS « ÉMOTIONNELS »	16	0	0,00 %
TOTAL	53	36	67,92 %

Tableau 5.40 : Différents (sous-)types d'emploi des occurrences du démonstratif adnominal tchèque dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT) rendues, en français, par l'article défini (InterCorp v1, CNT ; 241 occurrences)

ARTICLE INDEFINI

Type et sous-type d'emploi	Nombre d'occurrences	Dont les occurrences grammaticalisées	La proportion des occurrences grammaticalisées (%)
EMPLOIS « ÉMOTIONNELS »	3	0	0,00 %
ENDOPHORE	2	1	50,00 %
Anaphore fidèle	2	1	50,00 %
TOTAL	5	1	20,00 %

Tableau 5.41 : Différents (sous-)types d'emploi des occurrences du démonstratif adnominal tchèque dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, le CNT) rendues, en français, par l'article indéfini (InterCorp v1, CNT ; 241 occurrences)

DETERMINANT POSSESSIF

Type et sous-type d'emploi	Nombre d'occurrences	Dont les occurrences grammaticalisées	La proportion des occurrences grammaticalisées (%)
EXOPHORE A-MEMORIELLE	2	0	0,00 %
Deixis situationnelle	2	0	0,00 %
EXOPHORE MEMORIELLE	3	1	33,33 %
ENDOPHORE	11	3	27,27 %
Anaphore fidèle	5	3	60,00 %
Anaphore résomptive	5	0	0,00 %
Anaphore infidèle	1	0	0,00 %
EMPLOIS « ÉMOTIONNELS »	6	0	0,00 %
OCCURRENCES AMBIGUES	3	0	0,00 %
TOTAL	25	4	16,00 %

Tableau 5.42 : Différents (sous-)types d'emploi des occurrences du démonstratif adnominal tchègue dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT) rendues, en français, par le déterminant possessif (InterCorp v1, CNT ; 241 occurrences)

AUTRE TRANSPOSITION			
Type et sous-type d'emploi	Nombre d'occurrences	Dont les occurrences grammaticalisées	La proportion des occurrences grammaticalisées (%)
EXOPHORE A-MEMORIELLE	7	0	0,00 %
Deixis temporelle	6	0	0,00 %
Deixis situationnelle	1	0	0,00 %
EXOPHORE MEMORIELLE	3	2	66,67 %
ENDOPHORE	11	0	00,00 %
Anaphore résomptive	4	0	0,00 %
Anaphore fidèle	4	0	0,00 %
Cataphore	2	2	100,00 %
Anaphore infidèle	1	0	0,00 %
EMPLOIS « ÉMOTIONNELS »	10	0	0,00 %
OCCURRENCES AMBIGUES	2	0	0,00 %
TOTAL	33	4	12,12 %

Tableau 5.43 : Différents (sous-)types d'emploi des occurrences du démonstratif adnominal tchègue dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT) rendues, en français, par un autre moyen (InterCorp v1, le CNT ; 241 occurrences)

ABSENCE DE TRANSPOSITION			
Type et sous-type d'emploi	Nombre d'occurrences	Dont les occurrences grammaticalisées	La proportion des occurrences grammaticalisées (%)
EXOPHORE A-MEMORIELLE	4	0	0,00 %
Deixis temporelle	2	0	0,00 %
Deixis situationnelle	2	0	0,00 %
EXOPHORE MEMORIELLE	6	3	50,00 %
ENDOPHORE	24	1	4,17 %
Anaphore résomptive	15	0	0,00 %
Anaphore fidèle	4	0	0,00 %
Anaphore infidèle	2	0	0,00 %
Cataphore	1	1	100,00 %
Anaphore indirecte	1	0	0,00 %
Anaphore temporelle	1	0	0,00 %
EMPLOIS « ÉMOTIONNELS »	4	0	0,00 %
EMPLOIS « METALINGUISTIQUES »	2	0	0,00 %
DÉMONSTRATIF INDÉFINI	1	0	0,00 %
OCCURRENCES AMBIGUES	1	0	0,00 %
TOTAL	42	4	9,52 %

Tableau 5.44 : Différents (sous-)types d'emploi des occurrences du démonstratif adnominal tchègue dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT) non traduites en français (InterCorp v1, CNT ; 241 occurrences)

En nous appuyant sur la série de tableaux qui précèdent, nous pouvons faire une série de constats. D'abord, seules 20,48 % des occurrences rendues en français par le démonstratif contiendraient un TEN en cours de grammaticalisation. Ce nombre est très élevé au sein de l'emploi mémoriel : 12 occurrences sur 16, mais très bas ou complètement absent au sein des autres emplois. Ce chiffre est plutôt conforme à nos attentes, dans la mesure où le choix entre le défini et le démonstratif semble justement le moins évident lors de la traduction en français d'un emploi mémoriel de TEN (contrairement à ses emplois endophoriques et encore plus à ses emplois situationnels, où la frontière entre valeur démonstrative et valeur articulaire est plus nette ; voir à cet égard la Section 3.3 et le Chapitre 6).

La plupart des occurrences rendues en français par le défini, soit 67,92 %, sont celles où TEN est en cours de grammaticalisation, exception faite notamment des emplois « émotionnels » (16 occurrences). Ce dernier fait n'est pas non plus surprenant, car il arrive assez souvent que l'effet « émotionnel » produit par TEN ou bien ne soit pas rendu en français, ou bien qu'il y soit rendu par un autre moyen, plus en harmonie avec les possibilités systémiques de cette langue, dont le défini (voir le Chapitre 7). La transposition via l'article indéfini français est très minoritaire, ne permettant ainsi pas de tirer une quelconque conclusion. Parmi les occurrences transposées en français par le déterminant possessif, des occurrences grammaticalisées de TEN constitueraient 16 % (soit 4 des 26 transpositions par le possessif ; 3 appartiennent à l'anaphore fidèle). Pour les autres moyens de transposition, ce chiffre atteint 12,12 %. Enfin, l'absence de transposition, touchant avant tout l'emploi endophorique (24 des 42 occurrences qui ne sont pas transposées sont de nature endophorique), affiche également un pourcentage faible d'occurrences de TEN en voie de grammaticalisation : 9,52 %.

Si nous envisageons la situation **du point de vue des (sous-)types d'emploi**, nous notons dans un premier temps qu'en ce qui concerne **les emplois mémoriels jugés grammaticalisés** (32 occurrences au total), 43,75 % d'entre eux sont transposés en français par l'article défini :

90) Notak ze začátku byl **ten spor** zajímavý, pak se ale trochu zvrhl, že?

*Bon alors, au début **la discussion** était intéressante, mais ensuite elle a plutôt dégénéré, pas vrai ?*

V. Havel, *La fête en plein air*, 1963, trad. par F. Kérel, 1964, InterCorp v10

Néanmoins, la seconde transposition la plus courante s'effectue par le déterminant démonstratif (37,5 % ; voir *supra* pour l'explication proposée) :

91) Uvědom si, že odpovídáš za **ta děvčata** přímo jejich matkám ;

*Tu es responsable de **ces petites** directement devant leurs mères ;*

Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

Les emplois mémoriels qui ne sont pas, selon nous, engagés dans le processus de **grammaticalisation** ne sont qu'au nombre de 10. Cependant, malgré ce petit nombre, un fait n'est pas à négliger, à savoir l'absence totale de transposition par l'article défini (ce constat est conforme à nos attentes : le défini est davantage susceptible d'accompagner un emploi grammaticalisé, voir *supra*) :

- 92) Třeba ti to přijde malicherné, ale já myslím, že i tyhle věci představují takový jakýsi tmel, který drží rodin dohromady a pomáhá vytvářet **takový ten správný pocit, že máš doma skutečné zázemí**.
Ça peut te paraître négligeable, mais je suis convaincu que ces choses-là servent à cimenter les liens de famille. C'est en y veillant qu'on arrive à ce merveilleux sentiment rassurant d'avoir à la maison, chez soi, une véritable base, un fondement en quelque sorte.

V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

Dans le domaine de l'exophore **a-mémorielle**, 7 occurrences ont été évaluées comme **grammaticalisées**, dont toutes appartiennent à la deixis situationnelle. Comme dans le cas des emplois mémoriels grammaticalisés, c'est l'article défini qui prévaut dans leur transposition en français (57,14 % ; cf. l'exemple 93), mais ce dernier est à nouveau suivi d'assez près du déterminant démonstratif (42,86 % ; cf. l'exemple 94)¹²⁰ :

- 93) Uvědomuješ si, jaké dramatické napětí vzniká mezi ní a **tím kinžálem**?
As-tu remarqué le champ de tension dramatique qui se forme entre la statue et le sabre turc ?

V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

- 94) Potřeboval jsem totiž, aby šla přesně do **toho výklenku** [...]
il m'en fallait absolument une à la dimension de cette niche [...]

V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

Si nous regardons la **partie non-grammaticalisée des occurrences exophoriques a-mémorielles** (soit 29 occurrences), nous constatons que parmi celles-ci, 20 relèvent de la deixis situationnelle, 8 de la deixis temporelle et une seule de la deixis textuelle. Au sein de la première catégorie, c'est la transposition par le déterminant démonstratif qui prime (70 % de la totalité), alors que l'article défini n'apparaît qu'une fois. Au sein de la deuxième catégorie, le démonstratif adnominal a soit été rendu par un autre moyen (75 % de la totalité) soit il n'a pas été rendu du tout. Là où il a été rendu par un autre moyen, le traducteur a recouru 4 fois à un adverbe ou à une locution adverbiale temporels :

- 95) Tak Mary, a **od tohoto okamžiku** má naše organizace pod kontrolou prakticky celé podsvětí!
/ littéralement : à partir de ce moment-ci
C'est chose faite, Maria, à dater d'aujourd'hui notre organisation a presque tout ce monde-là sous son contrôle !

V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

¹²⁰ La pertinence de ces constats doit évidemment être relativisée au vu du nombre total très faible (12 occurrences).

Nous pouvons ainsi constater que cette dernière tendance est en quelque sorte inverse à celle que nous avons relevée pour nos deux corpus d'oral informel, où le français emploie beaucoup plus de démonstratifs que le tchèque avec la deixis temporelle (voir les Sections 3.3.1 et 4.5.1). L'unique emploi déictique textuel du sous-corpus a été rendu par un déterminant démonstratif.

Regardons à présent **les emplois endophoriques**. Parmi ceux-ci, 27 ont été **analysés comme engagés dans le processus de grammaticalisation**, dont notamment 12 emplois cataphoriques intra-syntagmatiques et 11 emplois anaphoriques fidèles. Au sein des seconds, l'article défini a été utilisé par les traducteurs dans la majorité des cas (soit 63,64 % ; cf. l'exemple 96), suivi du déterminant possessif (soit 27,27 % ; cf. l'exemple 97)¹²¹ :

96) Il vient d'être question d'un petit magot :

To jmění už skoro mám.

Le petit magot, je l'ai presque ! V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

97) Il vient d'être question d'un brocart :

Nezapomenete mi na ten brokát, pane Peachume, vid'te?

Vous n'oubliez pas mon brocart, hein, mon cher Volland ?

V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

Dans un cas, c'est l'article indéfini qui a été utilisé. Au sein de la cataphore intra-syntagmatique, l'article défini a été utilisé dans 58,33 % des cas. Le déterminant démonstratif a tout de même été choisi dans 2 cas (soit 16,67 %), de même qu'un autre type de transposition. Pour les 4 occurrences endophoriques restantes (anaphores résomptive, associative, infidèle et indirecte, respectivement), c'est l'article défini qui a été choisi à chaque fois.

La majorité des **occurrences endophoriques**, soit 75, ne serait cependant **pas touchée par le processus de grammaticalisation**. Au sein de cet ensemble, c'est l'anaphore résomptive qui est la plus fréquente, avec 46 occurrences. Parmi celles-ci, 47,83 % sont transposées en français par le démonstratif (cf. l'exemple 84 *supra*, où *tyhle věci (ces choses-là)* synthétise ce qui vient d'être dit). Si l'article défini n'apparaît jamais, 32,61 % des occurrences ne sont pas transposées du tout. Le possessif apparaît dans 5 cas (10,87 % de l'ensemble) et une autre transposition dans 4 cas (8,7 % de l'ensemble). Les anaphores fidèle et infidèle comptent, respectivement, 17 et 10 occurrences. Au sein de la première, 6 occurrences ont été rendues par le démonstratif, 4 ont été rendues par un autre moyen et 4 n'ont pas été transposées. Au sein de la seconde, 6 occurrences ont été transposées par le démonstratif, les 4 occurrences restantes étant réparties entre le déterminant possessif (une occurrence), une autre transposition (une occurrence) et l'absence de transposition (2 occurrences).

¹²¹ Le français emploie, de manière générale, plus de constructions avec le possessif que le tchèque, en partie parce que le tchèque exprime plus souvent la possession de façon externe, au moyen d'un datif. C'est le cas en 97, où *mi* est un datif d'intérêt.

Dans les deux cas, nous notons l'absence totale de l'article défini. Restent 2 occurrences, dont l'une relève de l'anaphore indirecte et l'autre de l'anaphore temporelle. Ni l'une ni l'autre n'est transposée en français.

Du côté des **emplois « émotionnels »** (50 occurrences au total), l'article défini est choisi par les traducteurs dans 32 % des cas (cf. l'exemple 98, où TEN possède une valeur « de rappel » – le locuteur rappelle la carrière de dramaturge de l'interlocuteur, laquelle a déjà fait l'objet du dialogue), le déterminant démonstratif, dans 22 % des cas (cf. l'exemple 99, où le démonstratif marque une prise de distance) et une autre transposition dans 20 % des cas (cf. l'exemple 88 *supra*, où seul le quantificateur *trochu* est retenu en français : *trochu toho deště* → *un peu de pluie*) :

98) A to jste znal taky různý herečky, když jste psal pro **to divadlo**?

Toutes ces belles actrices là, vous les connaissiez aussi, quand vous étiez écrivain pour le théâtre ?

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

99) Kdo je to vlastně **ten Kohout**?

Pendant que j'y pense, qui est ce Kohout ?

V. Havel, *Audience*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

Nous relevons encore la présence du déterminant possessif dans 6 cas, soit 12 %. Nous ne nous attarderons pas ici sur les occurrences ambiguës. En revanche, les 2 emplois « métalinguistiques » ne sont pas transposés en français, ce qui n'est guère surprenant, car ce type d'emploi n'y possède pas vraiment d'équivalent (voir les Sections 5.2.2 et 5.2.3).

Pour résumer, nous retenons principalement les deux faits suivants : 36 occurrences de TEN adnominal évaluées comme étant en voie de grammaticalisation sont transposées par l'article défini en français, soit 54,55 % ; 17 de ces emplois le sont par le déterminant démonstratif, soit 25,76 %¹²². Alors que la première donnée est conforme à nos attentes, la seconde les met quelque peu à l'épreuve. Dans le Chapitre 6, traitant de la grammaticalisation de TEN, nous essaierons de proposer des explications à ce phénomène, explications allant au-delà d'une simple interférence possible entre l'original et la traduction. Dans le Chapitre 7, traitant des emplois « émotionnels » du démonstratif adnominal tchèque, nous nous pencherons de plus près sur les choix de traduction de ce type d'emploi. Davantage d'exemples seront présentés et commentés dans ces deux chapitres.

¹²² Cette différence est statistiquement significative : $L = 11,58$; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

5.3 Les noms à sémantique temporelle : comparaison des deux langues

Nous terminerons le présent chapitre par un bilan contrastif portant sur la présence, dans nos deux ensembles de données de conversations, de noms à sémantique temporelle introduits par un démonstratif adnominal. Nous développerons ainsi les pistes mises au jour dans les sections précédentes. Comme nous l'avons souligné à plusieurs reprises (voir les Sections 5.1.3, 5.2.2.1, 5.2.3.1, 5.2.4.1), nos données pour le français contiennent beaucoup plus de noms porteurs d'un sens temporel, et ce à la fois en emploi endophorique et en emploi exophorique temporel¹²³.

5.3.1 Les noms à sémantique temporelle en emploi endophorique

Nous commencerons par les noms temporels en emploi endophorique. Dans les données de l'échantillon d'ORTOFON v1, nous recensons uniquement 14 occurrences de ce type. Le Tableau 5.45 les présente avec leurs sous-types d'emploi ainsi que leurs positions par rapport au verbe de la proposition. Celle-ci est indiquée au moyen des abréviations utilisées lors de l'annotation et expliquées dans le Chapitre 2 (Section 2.2.3).

Nom et sous-type(s) d'emploi	Nombre d'occurrences	Position de l'occurrence par rapport au verbe
DOBA (EPOQUE/TEMPS)	4	
Anaphore indirecte	2	CP
Anaphore temporelle	1	CP
Cataphore intra-syntagmatique	1	CP
DEN (JOUR)	4	
Cataphore intra-syntagmatique	3	CP
Anaphore résomptive	1	CP
MOMENT (MOMENT)	1	
Cataphore intra-syntagmatique	1	CP
HODINA (HEURE)	1	
Cataphore intra-syntagmatique	1	CP
DATUM (DATE)	1	
Anaphore associative	1	POST
TÝDEN (SEMAINE)	1	
Cataphore intra-syntagmatique	1	CP
VÍKEND (WEEK-END)	1	
Cataphore intra-syntagmatique	1	LD
ROK (ANNEE)	1	
Cataphore intra-syntagmatique	1	CP
TOTAL	14	

¹²³ Dans les deux langues, une poignée d'occurrences de noms temporels sont également employées en exophore mémorielle. En français, c'est notamment le cas du nom *vacances* (cf. **ces vacances** *tu as fait quoi ?* CEFC).

Tableau 5.45 : Noms à sémantique temporelle en emploi endophorique (fréquence, sous-type(s) d'emploi et position(s) par rapport au verbe) dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Le tableau montre que c'est en emploi cataphorique intra-syntagmatique qu'apparaissent la plupart des occurrences (9). L'anaphore indirecte est représentée par 2 occurrences du nom DOBA. Les trois autres sous-types d'emploi, n'apparaissant tous qu'une seule fois, sont l'anaphore temporelle (l'exemple 100), résomptive et associative.

100)	její	tatínek	byl	taky jakoby	středoškolsk-ej
	son.F.NOM.SG	papa.NOM.SG	être.PST.3SG.M aussi MOD		de lycée-M.NOM.SG
	profesor	ale to v tý	dob-ě	byl-o	hodně
	professeur.NOM.SG	mais ce à ten.F.LOC.SG	époque-LOC.SG	être.PST.3SG-N	beaucoup
	<i>son papa était lui aussi prof au lycée ; mais à l'époque, ça voulait dire beaucoup, ça</i>				
	CEFC				

Regardons maintenant comment la situation se présente **dans l'échantillon de conversations françaises**, où nous enregistrons 81 occurrences endophoriques avec un nom temporel. Seuls sont représentés les sept noms les plus fréquents, occupant, au total, 74 de ces 81 occurrences (soit 91,36 % ; l'ensemble des 81 occurrences contenant un nom temporel constitue 16,56 % de toutes les occurrences endophoriques de l'échantillon) :

Nom et sous-type(s) d'emploi	Nombre d'occurrences par nom et par sous-type d'emploi	Position de l'occurrence par rapport au verbe	Nombre d'occurrences par position par rapport au verbe
MOMENT	29		
Anaphore temporelle	14	CP	10
		CLIVAGE	2
		POST	2
Anaphore résomptive	11	CP	11
Anaphore infidèle	2	POST	2
Anaphore fidèle	1	CP	1
Cataphore	1	CP	1
FOIS	16		
Anaphore temporelle	16	CP	14
		POST	1
		?	1
TEMPS	7		
Anaphore résomptive	6	CP	6
Anaphore temporelle	1	CP	1
JOUR	6		
Anaphore résomptive	5	CP	5
Anaphore infidèle	1	CP	1
ÉPOQUE	6		
Anaphore temporelle	4	CP	4
Anaphore infidèle	1	CP	1

Anaphore fidèle	1	CP	1
SOIR	5		
Anaphore temporelle	3	CP	2
		POST	1
Anaphore infidèle	1	CP	1
Anaphore fidèle	1	CP	1
ANNÉE	5		
Anaphore fidèle	4	CP	4
Anaphore temporelle	1	CP	1
TOTAL	74		

Tableau 5.46 : 7 noms à sémantique temporelle en emploi endophrasique les plus fréquents (fréquence, sous-type(s) d'emploi et position(s) par rapport au verbe) dans l'échantillon de conversations du CEFC (Orfeo ; 1 000 occurrences)

Grâce au Tableau 5.46, nous pouvons voir que c'est le nom temporel MOMENT qui connaît la plus haute fréquence dans l'endophrase : 29 occurrences au total, dont 14 en anaphore temporelle et 11 en anaphore résomptive. Nous renvoyons ici le lecteur aux exemples 15, 16 et 17, cités dans la Section 5.1.3 *supra*. Suit le nom FOIS (16 occurrences), se prêtant exclusivement à l'anaphore temporelle :

- 101) L1 : oh évidemment ma voiture je l'ai encore amenée chez le garagiste
L2 : pourquoi *cette fois-ci*
L1 : un jour je vais pour démarrer j'enclenche la première et la roue arrière elle était bloquée
CEFC

Le nom TEMPS connaît 7 occurrences, dont 6 en anaphore résomptive (cf. les exemples 12, 13 et 14 de la Section 5.1.3). Avec leurs 6 occurrences, les noms JOUR et EPOQUE se partagent la position du quatrième nom temporel le plus présent dans l'endophrase. JOUR est avant tout employé en anaphore résomptive (5 occurrences) alors que EPOQUE l'est surtout en anaphore temporelle (4 occurrences). SOIR et ANNEE se partagent la cinquième position, avec 5 occurrences chacun, sachant que SOIR connaît 3 occurrences anaphoriques temporelles et ANNEE 4 occurrences anaphoriques fidèles. Les autres noms – non représentés dans le tableau – ressortissent aux anaphores fidèle et infidèle. La plupart des occurrences des noms (64, soit 86,49 %) occupent la position d'un complément de phrase. Le fait que 54 des occurrences répertoriées par le Tableau 5.46 (soit 72,97 %) apparaissent à la fois en emplois anaphoriques résomptif ou temporel et en position de complément de phrase rejoint ce qui a été dit dans la Section 5.2.2.2, à savoir que ces deux types d'anaphore semblent les plus susceptibles d'établir un cadre pour l'action décrite par le contenu de la proposition.

Voici maintenant le Tableau 5.47, où nous nous intéressons à **la présence de particules localisatrices** avec les noms temporels français susmentionnés :

Nom	Nombre d'occurrences	Sous-type d'emploi	Nombre d'occurrences par sous-type d'emploi
-CI			
FOIS	5	Anaphore temporelle	5
-LA			
MOMENT	29	Anaphore temporelle	14
		Anaphore résomptive	11
		Anaphore infidèle	2
		Anaphore fidèle	1
		Cataphore	1
JOUR	6	Anaphore résomptive	5
		Anaphore infidèle	1
EPOQUE	5	Anaphore temporelle	3
		Anaphore fidèle	1
		Anaphore infidèle	1
SOIR	4	Anaphore temporelle	3
		Anaphore infidèle	1
TEMPS	2	Anaphore résomptive	2
ANNEE	1	Anaphore temporelle	1
TOTAL	47		
ABSENCE DE PARTICULE			
FOIS	11	Anaphore temporelle	11
TEMPS	5	Anaphore résomptive	4
		Anaphore temporelle	1
ANNEE	4	Anaphore fidèle	4
EPOQUE	1	Anaphore temporelle	1
SOIR	1	Anaphore fidèle	1
TOTAL	22		

Tableau 5.47 : Présence de particules localisatrices avec les 7 noms à sémantique temporelle les plus fréquents en emploi endophorique (fréquence, sous-type(s) d'emploi) dans l'échantillon de conversations du CEFC (Orfêo ; 1 000 occurrences)

Ce tableau montre qu'au total, 47 des 74 occurrences de noms temporels concernés dans l'endophore contiennent la particule *-là* (soit 63,51 %), 22 ne contiennent aucune particule (29,73 %) et 5 contiennent la particule *-ci* (6,76 %). La particule *-là* est présente avec toutes les occurrences des noms MOMENT et JOUR. Quant à *-ci*, il n'apparaît qu'avec 5 occurrences du nom FOIS, employées, comme toutes les occurrences de ce nom introduites par le démonstratif, en anaphore temporelle. Contrairement à ce qui pourrait sembler intuitivement, il n'est pas vrai que tous les emplois endophoriques des noms temporels exigent la présence de *-là*, quoique *-là* soit tout de même présent, dans nos données, avec 58,02 % de ces noms :

102) il y avait encore Thatcher ou pas à *cette époque-là* CEFC

103) si tu veux à *cette époque* ma tante c'était sa patronne ma mère CEFC

Il pourrait sembler étonnant que, d'une part, *-là* accompagne systématiquement le nom MOMENT, alors que tout risque d'ambiguïté entre une interprétation endophorique et exophorique est écarté pour ce nom (c'est la présence de la préposition qui tranche : à *ce moment(-là)* vs *en ce moment(-ci)*), et que, d'autre part, un nom comme ANNEE semble apparaître sans problème en emploi endophorique sans *-là*, sachant que *cette année* tout seul est davantage susceptible d'une interprétation déictique.

Ajoutons encore qu'au total, l'endophore constitue 28,57 % de toutes les occurrences avec *-ci* et 77,66 % de toutes les occurrences avec *-là*.

5.3.2 Les noms à sémantique temporelle en emploi déictique temporel

Dans les données d'ORTOFON v1, nous constatons la présence de seulement 3 noms temporels en emploi déictique temporel : TÝDEN (SEMAINE), DOBA (EPOQUE/TEMPS) et SEMESTR (SEMESTRE). Le premier y connaît 4 occurrences, dont 3 en position de complément de phrase et une seule en position postverbale. Les deux autres n'y connaissent qu'une seule occurrence chacune, placée en position de complément de phrase¹²⁴.

Les données françaises, en revanche, contiennent 282 occurrences de noms temporels employés en deixis temporelle, soit 28,2 % de toutes les occurrences (voir la Section 5.2.2.1). Les 7 noms les plus représentés – constituant, au total, 241 de ces 282 occurrences (85,46 %) – figurent dans le Tableau 5.48 :

Nom	N° d'occurrences par nom	Position de l'occurrence par rapport au verbe	N° d'occurrences par position par rapport au verbe	Particule localisatrice ¹²⁵
ANNEE	53	CP	46	
		POST	6	
		?	1	
MATIN	46	CP	39	
		POST	6	
		?	1	
SOIR	42	CP	29	
		POST	11	
		?	2	

¹²⁴ Dans les textes des pièces, nous retrouvons 5 fois le nom CHVÍLE (MOMENT/INSTANT), 2 fois le nom OKAMŽIK (INSTANT/MOMENT) et une fois le nom SOBOTA (SAMEDI).

¹²⁵ En cas d'absence de particule localisatrice, la case reste vide.

MOMENT	41	CP	35	
		POST	3	
		?	3	
SEMAINE	31	POST	14	1 occurrence de <i>là</i>
		CP	12	1 occurrence de <i>là</i>
		?	3	
		CLIVAGE	1	
		LD	1	
WEEK-END	16	CP	13	
		POST	3	
APRES-MIDI	12	CP	11	
		POST	1	
TOTAL	241			

Tableau 5.48 : 7 noms à sémantique temporelle les plus fréquents en emploi déictique temporel (fréquence, position par rapport au verbe, présence de particules localisatrices) dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)

Comme nous l'avons déjà montré dans la Section 5.2.2.2, la plupart des noms temporels en emploi déictique temporel occupent la place d'un complément de phrase (plus précisément, 73,40 % ; voir le Tableau 5.10 *supra*). La part des occurrences à particule localisatrice étant extrêmement faible, seules 2 occurrences du nom SEMAINE contiennent *-là* (exemple 104 et 105) :

- 104) L1 : dès la première semaine on a un devoir à rendre pour monsieur NNAAMMEE une analyse poussée de La Princesse Mononoke de Hayao Miyazaki
L2 : pff oui mais il laisse il laisse plus qu'une semaine je crois
L1 : si il laisse une semaine entière
L2 : qu'une semaine
L1 : ouais pour faire euh un petit exposé euh
L2 : *ce~ cette semaine-là* pour la prochaine
- 105) L1 : et euh sinon tu as fait quoi toi le week-end euh tout euh ces vacances-là
L2 : mh qu'est-ce que j'ai fait j'ai été j'ai été deux ou trois fois sur Aix le soir
L1 : ouais ah au Sunset
L2 : non j'ai pas été au Sunset j'ai été dans un bar à chicha parce que chez moi j'en ai une L1 : ah ouais ben j'en ai fait pendant euh pendant euh *cette semaine-là*

Ce qui est avant tout étonnant, c'est que nous ne retrouvons pas une seule occurrence avec *-ci*. Il s'avère ainsi que non seulement la présence de ce dernier ne serait absolument pas nécessaire en deixis temporelle, mais que les locuteurs tendraient même à l'éviter¹²⁶.

Partant d'une intuition se basant sur les données du Tableau 5.47, nous avons appliqué le calcul de spécificités pour déterminer si **une affinité** est présente **entre**, d'une part, **les SN démonstratifs contenant le nom *fois* et *-ci*** et, d'autre part, **les SN démonstratif contenant le nom *moment* et *-là***.

¹²⁶ Rappelons qu'au total, seules 1,42 % des occurrences déictiques temporelles de l'échantillon contiennent *-ci* et 0,71 % contiennent *-là*.

Pour ce faire, nous avons entré dans le calcul les données présentées par le Tableau 5.49 *infra*. Afin d'augmenter la significativité du calcul, nous avons tenu compte de l'ensemble des 1 871 occurrences valides du SN démonstratif générées par la requête décrite dans la Section 2.2.1 :

	CE + <i>-ci</i>	CE + <i>-là</i>	CE sans <i>-ci/-là</i>	Total
<i>fois</i>	6	0	16	22
<i>moment</i>	0	60	79	139
autres noms	21	253	1436	1710
Total	27	313	1531	1871

Tableau 5.49 : Données utilisées pour le calcul de spécificités appliqué au lien entre CE + *fois* + *-ci* et CE + *moment* + *-là* dans l'ensemble des 1 871 occurrences générées par la requête au sein des conversations du CEFC (Orféo)

Les résultats du calcul, affichés via le graphique de la Figure 5.36, appuient cette intuition, notamment pour CE + *moment* + *-là* :

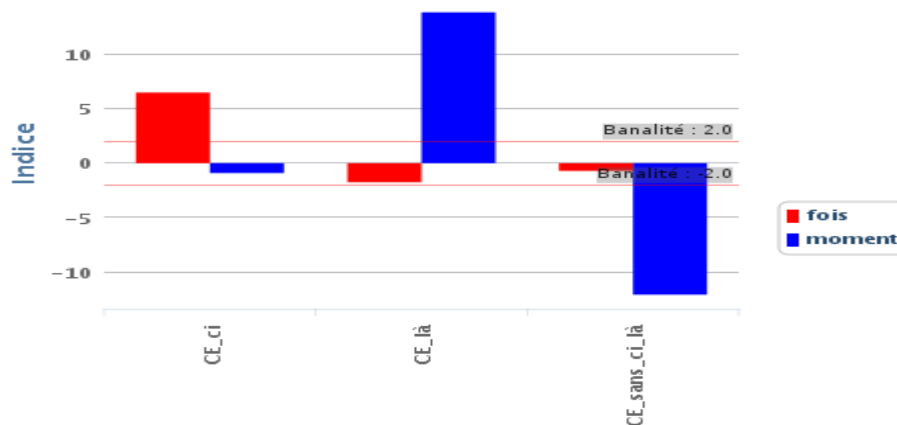


Figure 5.36 : Calcul de spécificités appliqué à CE + *fois* + *-ci* et CE + *moment* + *-là* dans l'ensemble des 1 871 occurrences générées par la requête au sein des conversations du CEFC (Orféo)

Enfin, vu le nombre très bas de noms à sémantique temporelle introduits par un démonstratif défini que nous avons constaté pour l'échantillon de conversations tchèques, il nous semble opportun de **rapporter ce bilan à un ensemble plus vaste de données**. C'est pour cela que nous avons regardé la fréquence relative approximative de cette même combinaison affichant l'ensemble des noms tchèques à sémantique temporelle mentionnés dans le Tableau 5.45 (DOBA, DEN, MOMENT, HODINA, DATUM, TÝDEN, VÍKEND et ROK), auxquels nous avons encore ajouté les noms SEMESTR (SEMESTRE), CHVÍLE (MOMENT, INSTANT¹²⁷), ODPOLEDNE (APRÈS-MIDI), VEČER (SOIR) et RÁNO (MATIN), dans la totalité d'ORTOFON v1¹²⁸. La fréquence est approximative, car nous n'avons cherché que les cas où le nom est directement précédé d'un démonstratif adnominal défini, et elle s'élève à 388,2

¹²⁷ Ce mot est plus usité en tchèque que son synonyme MOMENT.

¹²⁸ [tag= "PD.....*"]

& !lemma="takový"] [lemma="odpoledne|večer|ráno|doba|den|moment|hodina|datum|týden|víkend|rok|semestr|chvíle"].

occurrences pour un million de positions. En français, la fréquence relative des noms à sémantique temporelle mentionnés dans les Tableaux 5.47 et 5.48 (MOMENT, FOIS, TEMPS, JOUR, EPOQUE, SOIR, ANNEE, MATIN, SEMAINE, WEEK-END et APRES-MIDI), auxquels nous avons ajouté HEURE, DATE et SEMESTRE, s'élève, dans l'ensemble de notre corpus de conversations¹²⁹, à 602,2 occurrences pour un million de positions. Aussi approximatifs que soient ces deux résultats (passage à la trappe des occurrences au sein desquelles le nom est modifié par un ou plusieurs adjectifs ou post-déterminants, nombre limité de noms recherchés), la différence est statistiquement significative après l'application de la fonction de vraisemblance aux fréquences absolues : $L = 20,977^{130}$. Il semble ainsi légitime d'avancer que l'oral informel français emploie davantage de noms à sémantique temporelle précédés d'un démonstratif défini, quoique le clivage soit bien moindre que ne le laisse croire l'analyse des seuls échantillons.

5.4 Résumé et conclusion

Le but de ce chapitre était d'étudier notre matériel empirique dans une optique essentiellement quantitative. Dans un tout premier temps (Section 5.1), nous avons passé en revue les principaux écueils et dilemmes liés à notre entreprise d'annotation, en complétant les éléments présentés notamment dans le Chapitre 2, et nous avons tenté de justifier nos partis-pris.

Nous avons ouvert le deuxième volet (Section 5.2) en introduisant les chiffres des fréquences absolues et relatives du démonstratif adnominal dans les trois sources de données. Pour ce qui concerne l'oral informel des deux langues, les données correspondent tout à fait à nos attentes, à savoir que le tchèque fait un usage beaucoup plus important du démonstratif adnominal que le français. L'écrasante majorité des occurrences pourvues du démonstratif affiche la forme TEN (soit à peu près 96,47 %). Le sous-corpus des pièces de V. Havel fait un usage plus élevé du démonstratif adnominal que le corpus de conversations françaises, mais bien moindre que celui de conversations tchèques. Force est cependant de constater que la pertinence de cette comparaison est en partie remise en question par plusieurs facteurs, dont notamment les tailles divergentes des sources (et, avant tout, la taille « minuscule » du sous-corpus des pièces de V. Havel), l'impossibilité d'identifier absolument toutes les occurrences recherchées, les chiffres approximatifs des fréquences dont nous disposons pour le corpus de conversations tchèques, mais également le ton bien plus formel de ces pièces¹³¹.

¹²⁹ [pos = "DET.DEM"] [lemme="moment|temps|jour|époque|soir|année|matin|semaine|week-end|après-midi|heure|date|semestre"].

¹³⁰ La différence est également significative si nous comparons les nombres de toutes les occurrences des noms en question dans les deux corpus (et non seulement leurs occurrences introduites par un démonstratif défini) : $L = 450,1677$. La valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

¹³¹ Ce sont les analyses réalisées dans le présent chapitre (Section 5.2.4) qui nous ont fait prendre la pleine mesure de cette formalité.

Afin d'avoir une idée purement quantitative de l'avancée de la grammaticalisation de TEN adnominal en tchèque parlé informel, nous avons comparé la fréquence relative de ce dernier dans ORTOFON v1 avec la fréquence relative atteinte par l'article défini dans différents types de textes français. Cette comparaison a montré que, de par sa fréquence, TEN reste tout de même loin du statut d'article défini, a fortiori si l'on tient compte du fait qu'une partie considérable de ses occurrences sont dotées d'une valeur « émotionnelle », emploi où tout soupçon de grammaticalisation doit être écarté¹³².

Nous avons ensuite procédé à l'étude de chacune des trois sources de données en nous intéressant à plusieurs phénomènes. Là aussi, des différences considérables ont été constatées, à commencer par la fréquence des différents types et sous-types d'emploi du démonstratif adnominal au sein de l'échantillon¹³³. Si nous comparons les données de l'oral informel, il est frappant de constater l'importance de la place occupée en tchèque par les emplois « émotionnels » et mémoriels. Ce constat est tout à fait en phase avec **notre hypothèse principale**, selon laquelle l'abondance de TEN adnominal au sein de ce registre procède à la fois 1) de sa capacité à exprimer une série de valeurs pragmatiques en contextes où il ne participe pas directement à l'identification d'un référent et 2) de l'extension de ses emplois en direction d'un futur article défini (voir à cet égard l'Introduction générale). L'oral informel français, en revanche, est marqué avant tout par les emplois endophoriques et déictiques temporels. De même, le profil des emplois endophoriques diffère de manière significative dans les deux échantillons avec, en français, un usage abondant de l'anaphore résomptive ainsi qu'une présence importante de l'anaphore temporelle. De façon très différente, le tchèque use beaucoup de la cataphore (intra-syntagmatique) et de l'anaphore associative. Comme nous le verrons plus concrètement dans le chapitre suivant, ces deux derniers emplois sont, eux aussi, tout à fait cruciaux en rapport avec le taux d'occurrences considérées comme engagées dans le processus de grammaticalisation.

De grandes divergences existent également eu égard aux positions des occurrences par rapport au verbe de la proposition, à savoir, notamment, la quasi-absence de la position préverbale en français et sa présence fort élevée en tchèque, tout comme une fréquence bien plus élevée des compléments de phrase en français. Nous avons insisté sur une série d'affinités entre type d'emploi et position par rapport au verbe, à commencer, en français, par celle entre le complément de phrase et les emplois

¹³² Même ici, la fiabilité du résultat obtenu doit être relativisée, comme nous l'avons écrit plus haut, et ce notamment parce que nous sommes dans l'incapacité d'identifier toutes les occurrences des deux lemmes (surtout pour ce qui est du défini français). En même temps, cette insuffisance pourrait aussi être perçue comme encourageante pour notre conclusion, car si le nombre d'occurrences du défini français est en vérité plus élevé que celui constaté, il s'ensuit que le clivage de fréquence entre ce dernier et TEN adnominal est d'autant plus grand.

¹³³ Rappelons ici que l'exactitude de l'intervalle de confiance anticipant la présence d'un (sous-)type d'emploi dans un autre échantillon hypothétique de la même taille va croissant avec le taux de représentation de ce (sous-)type d'emploi au sein de l'échantillon étudié. Par conséquent, tout pronostic concernant les (sous-)type d'emploi peu représentés s'avère délicat.

déictique temporel et anaphorique temporel ou entre la dislocation à droite et l'endopore, affinités étayées par le calcul de spécificités. Dans les deux échantillons de conversations, la position postverbale et l'exopore mémorielle sont fortement liées, comme l'a également montré l'application du calcul de spécificités. Les dislocations à droite et à gauche ont fait l'objet d'un intérêt tout particulier : après avoir étudié la nature des référents apparaissant dans ces deux positions, il nous semble possible de donner en partie raison aux auteurs qui ont théorisé les différences entre les deux types de dislocation (voir pour cela la Section 2.2.3), bien que le bilan s'appuyant sur nos données reste en partie mitigé.

En tchèque, la grammaticalisation de TEN en emploi adnominal et la combinaison du démonstratif avec un autre élément fonctionnel ont fait l'objet d'une étude liminaire. Comme le présent chapitre n'a présenté qu'une ébauche des deux phénomènes, des aspects importants restent évidemment à creuser dans les deux chapitres suivants. En ce qui concerne la grammaticalisation, celle-ci semble avant tout associée aux emplois mémoriel, cataphorique, anaphorique associatif et, dans une moindre mesure, situationnel. Cette tendance fera l'objet d'une analyse qualitative approfondie dans le Chapitre 6. Pour ce qui est des cumuls, ce sont les combinaisons TAKOVÝ + TEN, TEN + possessif et NĚJAKÝ + TEN qui priment. Les disfluences affectant l'expression démonstrative constituent un autre point intéressant. Ici, nous avons insisté sur la spécificité des emplois dits « métalinguistiques », apparaissant, eux aussi, exclusivement en tchèque.

Concernant le sous-corpus des quatre pièces de théâtre, nous avons prêté une attention particulière à la transposition du démonstratif tchèque dans les traductions françaises. Cette étude doit également être approfondie, et ce notamment pour deux raisons : 1) afin de rendre compte du fait qu'une occurrence de TEN qu'un nombre suffisant d'indices permet pourtant de rapprocher d'un article défini soit tout de même rendue, en français, par le démonstratif et 2) pour explorer les différentes manières de rendre en français l'emploi « émotionnel » du démonstratif : l'effet « émotionnel » est-il maintenu ou supprimé ? S'il est maintenu, par quel moyen ? Dans le reste de ce travail, nous essaierons d'apporter des réponses à toutes ces questions et à d'autres encore, à commencer par le grand sujet qu'est la grammaticalisation de TEN.

Le dernier volet (Section 5.3) a été consacré aux noms à sémantique temporelle introduits par un démonstratif adnominal dans les deux échantillons de conversations. Il s'avère que ce phénomène apparaît plus souvent dans les conversations françaises que tchèques, la principale cause de cette disparité semblant être de nature structurelle : la préférence du tchèque pour les expressions adverbiales afin de référer au temps (à la fois en deixis temporelle et en endopore). Elle s'explique aussi, en partie, par la forte présence, dans le registre parlé informel du français, des connecteurs argumentatifs contenant un nom temporel, dont notamment *à ce moment-là*. D'autres constats

intéressants concernant le français ont pu être mis au jour, tels que la quasi-absence de la particule *-ci*, y compris en emploi déictique temporel et la présence systématique de la particule *-là* avec le substantif *moment* lorsque celui-ci apparaît en endophore. De manière générale, l'application du calcul de spécificités à nos données françaises dans la Section 5.2.2.1 suggère par ailleurs qu'une affinité serait présente entre la particule *-là* et les emplois anaphoriques résomptif et temporel.

En guise de résumé, les faits que nous retenons avant tout pour la dernière partie du présent travail sont les suivants : 1) un clivage abyssal en matière de fréquence du démonstratif adnominal entre l'oral informel tchèque et français (sa fréquence dans l'échantillon tchèque fait 8,2 fois celle dans l'échantillon français) ; 2) une présence élevée d'occurrences de TEN où celui-ci a été évalué comme engagé dans un processus de grammaticalisation en article défini, et ce notamment en emplois mémoriel, cataphorique intra-syntagmatique et anaphorique associatif ; 3) une présence élevées d'occurrences de TEN où nous avons évalué celui-ci comme porteur d'un effet pragmatique « émotionnel ». Les phénomènes 2) et 3) sont directement liés au constat 1) et ils seront respectivement abordés dans les Chapitres 6 et 7, constituant la troisième partie du travail.

Partie III

Analyses qualitatives

6 La grammaticalisation du démonstratif TEN en article défini

Le présent chapitre a pour objectif de traiter du phénomène de la grammaticalisation de TEN en article défini. Par conséquent, il se limitera presque exclusivement à la partie tchèque de nos données. Si l'échantillon créé à partir du corpus ORTOFON v1 sera ici notre principale source, nous tiendrons également compte des données du sous-corpus des pièces de V. Havel. Quant aux données de l'oral informel français, elles nous aideront, le cas échéant, à mieux illustrer et étayer nos explications. Voici le plan que nous avons adopté pour cette partie.

Dans le cadre du premier volet, théorique, nous proposerons tout d'abord une introduction à la notion de grammaticalisation en nous appuyant essentiellement sur la monographie de Marchello-Nizia (2006c), mais aussi sur les travaux de Heine et Kuteva (2002), Givón (1971 et 2015), Hagège (1993), Heine (2002), Traugott et Hopper (2003 [1993]), Kuryłowicz (1965), Langacker (1985), Lehmann (2015 [1982]), Meillet (1982 [1912]), Peyraube (2002), Traugott (1980, 2003, 2005 [1995] et 2010), etc. Nous nous focaliserons ensuite plus particulièrement sur la problématique de l'émergence des articles définis à partir de démonstratifs adnominaux en enchaînant sur ce qui a été dit dans la Section 1.3.4. du Chapitre 1 et en nous fondant notamment sur les travaux de Carlier et De Mulder (2006, 2007, 2010 et 2011), Diessel (1999), Greenberg (1978a), Heine et Kuteva (2002), Himmelmann (1997 et 2001), Laury 1997 et C. Lyons (1999). Dans ce cadre, nous examinerons la distinction entre les définitudes « sémantique » et « pragmatique », distinction à laquelle Löbner (1985 et 2011) a donné une assise conceptuelle, et qui s'avère très porteuse, selon nous, pour l'étude de la progression de la grammaticalisation des démonstratifs en articles définis. La thèse de doctorat de Czardybon (2017) ainsi que l'étude d'Ortmann (2014) constituent des sources particulièrement précieuses pour notre propre étude du tchèque, car leurs auteurs ont appliqué l'approche de Löbner à la grammaticalisation du démonstratif en défini dans les langues slaves occidentales. Une introduction à la situation du tchèque sera proposée à la suite, où nous ferons un bref état de l'art de la problématique tout en rappelant son traitement par une série de spécialistes de cette langue (notamment Adamec, 1983 ; Berger, 1993 ; Mathesius, 1926 ; Vey, 1996 et Ziková, 2017 et 2018).

S'ensuivra un second volet, consacré à l'analyse de nos données. Nous partirons pour cela notamment des exemples mentionnés et commentés dans la Section 3.3. Ces exemples nous serviront de tremplin pour aller vers les différents types et sous-types d'emploi de TEN, que nous nous donnerons pour but d'examiner plus avant en rapport avec la grammaticalisation. Si un travail constant avec les données statistiques portant sur notre matériel empirique tchèque, figurant dans le Chapitre 5, est nécessaire pour étayer nos analyses, la distinction löbnérienne entre définitudes

« sémantique » et « pragmatique » constituera le principal axe théorique autour duquel ces analyses seront articulées. Ainsi, seront d'abord étudiés les contextes appartenant à la définitude « pragmatique », puis les contextes se trouvant à la charnière entre les deux types de définitude. Nous verrons que c'est le second type de contextes qui est sans conteste le plus révélateur au sujet de la progression de la grammaticalisation. Les travaux de Laury (1997) et de Zíková (2017 et 2018), s'intéressant à la naissance et la propagation des articles définis dans le domaine de l'anaphore textuelle de deux langues vivantes (le finnois pour la première et le tchèque pour la seconde), nous ont été d'une utilité inestimable lors de la définition des critères d'évaluation du degré de grammaticalisation de TEN. Nous tâcherons de fournir un résumé de ces critères à la fin du volet théorique du présent chapitre (Section 6.1.4), avant d'aborder l'analyse des données. Ce résumé s'accompagnera d'une série d'hypothèses relatives à la situation de TEN adnominal dans le tchèque (parlé) informel de nos jours, hypothèses que le second volet cherchera à confirmer ou à infirmer.

Le traitement de chaque type d'emploi inclura un bref bilan concernant les données correspondantes du sous-corpus des pièces de V. Havel. Dans le cadre de ce bilan, nous nous demanderons dans quelle mesure les choix des traducteurs reflètent la situation de l'original et tâcherons de trouver les principales explications pour les cas où la traduction française semble s'en écarter.

6.1 Grammaticalisation et émergence des articles définis : présentation générale

La notion linguistique de grammaticalisation a fait l'objet de nombreuses études. Malgré cela, elle continue à faire débat pour sa grande complexité. Nous dirons d'abord un mot au sujet de la grammaticalisation en général pour ensuite nous centrer sur l'apparition des articles définis.

6.1.1 La grammaticalisation

Dans un de ses travaux consacrés à ce sujet, Traugott (2005 [1995] : 32), s'inscrivant, comme nombre de théoriciens de la grammaticalisation, dans un paradigme cognitiviste de la linguistique, définit le processus comme suit :

'Grammaticalisation' is the process whereby lexical items or phrases come through frequent use in certain highly constrained local contexts to be reanalyzed as having syntactic and morphological functions, and, once grammaticalized, continue to develop new grammatical functions [...]

Marchello-Nizia (2006c : 24) écrit que

[l]a grammaticalisation est un processus complexe [...] qui touche tous les niveaux de l'analyse linguistique, mais il semble qu'au départ il s'agisse presque toujours d'un phénomène pragmatico-sémantique.

Un peu plus loin dans sa monographie, la même auteure (2006c : 31) résume à propos de ce qui est souvent considéré comme étant les trois traits définitoires de la grammaticalisation :

Trois traits fondamentaux définissent le processus de ‘grammaticalisation’ : il est progressif et unidirectionnel, et aboutit à la création d’une unité dont l’emploi peut être obligatoire lorsqu’il s’agit d’exprimer la notion pour laquelle il avait été à l’origine optionnellement choisi.

Regardons donc de plus près ces trois traits¹. Quant à la progressivité, il a été observé que la grammaticalisation se distingue par un processus d’affaiblissement graduel touchant l’unité qui s’y est engagée. C’est d’abord Meillet (1982 [1912] : 139), l’auteur même du terme de « grammaticalisation », qui a relevé ce fait :

La constitution des formes grammaticales par dégradation progressive de mots jadis autonomes est rendue possible par les procédés qu’on vient de décrire sommairement, et qui consistent, on le voit, en un affaiblissement de la prononciation, de la signification concrète des mots et de la valeur expressive des mots et des groupes de mots.

Comme il se dégage de cette citation, l’affaiblissement en question est généralisé, se manifestant (ou pouvant se manifester) sur plusieurs niveaux à la fois. Parmi ceux-ci, Meillet mentionne les niveaux phonétique (« la prononciation »), sémantique (« la signification ») et expressif (« la valeur expressive »). À cela, il faudrait ajouter le niveau morphologique, les unités tendant vers une érosion du signifiant (voir *infra*).

La citation de Meillet fait également apparaître le terme de « mot autonome ». Ce terme est à prendre au sens de « mot lexical », par opposition à « mot grammatical ». Une autre terminologie rendant compte de cette dichotomie est utilisée par Marchello-Nizia (2006c), qui distingue entre les « catégories majeures » et les « catégories mineures (secondaires) » du lexique. De façon prototypique, la grammaticalisation consisterait donc à produire des unités linguistiques grammaticales à partir d’unités linguistiques lexicales. Pourtant, comme nous l’avons déjà souligné dans la Section 1.3 en nous référant à Diessel (1999 : 152), ce schéma n’est pas validé par la grammaticalisation du démonstratif, ce dernier étant lui-même un mot grammatical « primitif » (à savoir non issu d’un mot lexical) :

[...] demonstratives are not ordinary grammatical markers, and hence it would make sense if they do not derive from lexical items as all other grammatical markers. [...] If demonstratives are not derived from lexical items, there would be two different sources from which grammatical markers may emerge: lexical expressions and demonstratives.

D’où la rectification de la définition par Kuryłowicz (1965 : 69), qui précise que :

¹ Nous ne les décrirons ici que dans leurs grands traits. Ce n’est que dans la section suivante que nous verrons davantage ce que cela implique pour les démonstratifs adnominaux en cours de grammaticalisation.

Grammaticalization consists in the increase of the range of a morpheme advancing from a lexical to a grammatical or **from a less grammatical to a more grammatical status**² [...]

Sur le plan sémantique, la nature progressive de la grammaticalisation fait, comme le remarque Hagège (1993 : 22), apparaître des stades intermédiaires :

Un modèle tel que celui de la grammaticalisation pose qu'un phénomène se déroule de façon progressive, et donc que dans le passage du stade initial où la nouvelle unité n'existait pas, au stade final où elle s'est intégrée pleinement au système grammatical, il existe toujours un stade intermédiaire (ou des stades) où les deux systèmes coexistent.

Les étapes clés de cette progression peuvent être exprimés par le schéma suivant, inspiré par Heine (2002) :

A, Ab, Ba, B

Ainsi, la première phase (A) correspond à l'état dit « initial »³, à savoir celui où l'unité conserve pleinement son sens d'origine. Lors de la deuxième phase (Ab), un nouveau sens est dérivé par inférence au sein d'un (ensemble de) contexte(s) particulier(s) (le rôle de celui-ci est donc essentiel). Cette phase est appelée par Heine (2002) « contexte de transition » (“bridging context”). Vient ensuite la phase désignée par Heine comme « contexte de passage » (“switch context” ; aB), où

[l]e sens initial est relégué à l'arrière-plan (mais toujours accessible), et le nouveau sens permet au mot d'apparaître dans des contextes tout à fait nouveaux, et pas seulement dans les contextes ambigus qui avaient permis son apparition. (Marchello-Nizia, 2006c : 23)

Enfin, le tout s'achève dans la phase dite « de conventionalisation » (B), où seul le sens dérivé est maintenu au sein du/des contexte(s) en question, au détriment du sens d'origine, cependant que ce dernier continue parfois à être présent dans le système. C'est ce que Hagège (1993, 200-202) appelle la « preuve par anachronie » (“the proof by Anachrony principle”) : les deux sens co-existent indépendamment et peuvent se combiner, comme nous le constatons dans l'exemple suivant, où le verbe ALLER est d'abord utilisé en sa qualité de semi-auxiliaire et ensuite en sa qualité de verbe de mouvement⁴ :

1) Je **vais aller** au cinéma. (Marchello-Nizia, 2006c : 43)

² Nous soulignons.

³ Nous reprenons ici les traductions françaises des termes de Heine qui ont été proposées par Marchello-Nizia (2006c : 23).

⁴ De même Traugott : “[...] grammaticalized forms may coexist alongside non-grammaticalised ones, sometimes for many centuries [...]” (2005 [1995] : 32).

Pour ce qui concerne la nature unidirectionnelle du processus, celle-ci doit d'abord être comprise au sens de l'irréversibilité : il semble globalement admis que le processus ne saurait retourner en arrière⁵. Marchello-Nizia (2006c : 32) remarque que ce principe agit à trois niveaux différents :

1. au niveau *formel* : dans le cours d'un processus de grammaticalisation, l'évolution se fait soit sans changement de forme, soit en allant vers une forme plus réduite, mais jamais vers une forme plus étoffée que la forme de départ ;
2. au niveau *catégoriel* : on va toujours d'une catégorie majeure (nom, verbe, adjectif) vers une catégorie mineure ;
3. au niveau *sémantique* : le sens lexical évolue vers un sens grammatical plus général et plus abstrait.

À ceci près que, eu égard à ce qui a été souligné *supra* à propos du démonstratif, les points 2. et 3. doivent être relativisés (le démonstratif adnominal constitue déjà une catégorie mineure et n'est pas doté de sens lexical). S'agissant des changements encourus au niveau sémantique, ils sont souvent désignés par le terme de « semantic bleaching » (cf. Givón, 1971 ; Langacker, 1985 ; Lehmann, 2015 [1982] ; Peyraube, 2002 ; ce dernier propose l'équivalent français de « javélisation »⁶). Pour sa part, Givón (2015) remarque que l'unidirectionnalité est notamment à mettre sur le compte de l'érosion du signifiant (1^{er} niveau), auquel il est impossible, par la suite, de faire porter un sens plus étoffé.

Cette érosion a fait l'objet de nombreuses descriptions (cf. Meillet, 1982 [1912] ; Lehmann, 2015 [1982]). Un exemple parlant est la réduction du démonstratif latin *illam*, en position adnominale, à la forme de l'article défini français *la*. Cependant, ce phénomène n'est pas universel : en français, la succession adjectif + nom *beau coup* a donné l'adverbe et le déterminant quantificateur *beaucoup*, en tchèque, la forme du démonstratif adnominal TEN reste identique même lorsqu'elle est employée avec une valeur proche de l'article défini, etc. Outre l'érosion au sens propre, les auteurs relèvent également la perte de l'autonomie prosodique de l'unité (« l'affaiblissement prosodique et accentuel », Marchello-Nizia, 2006c : 40). Celle-ci peut être exprimée moyennant le schéma suivant :

⁵ Ou seulement très difficilement. Ajoutons cependant que la nature unidirectionnelle du processus a également fait l'objet de débats (cf. Andersen, 2001).

⁶ Un grand débat existe cependant autour de la question de savoir s'il est vraiment légitime de parler d'appauvrissement sémantique plutôt que d'une redistribution du sens, entre autres parce que ce dernier migre en partie du domaine de la sémantique vers celui de la pragmatique (cf. Traugott, 2005 [1995] et Traugott et Hopper, 2003 [1993]) et que le sens d'arrivée est parfois plus complexe que celui de départ (cf. Marchello-Nizia, 2006c). Ainsi, dans le processus de grammaticalisation d'un démonstratif en article défini, celui-ci voit d'abord son sens pragmatique renforcé (cf. Carlier et De Mulder, 2007 et 2010). Mais même une fois le processus abouti, il semble contestable de considérer l'article défini résultant comme sémantiquement plus pauvre que le démonstratif de départ : certes, l'ancien démonstratif a perdu son sens indexical (éventuellement aussi sa capacité à marquer la distance) ; en revanche, il a acquis la capacité à marquer l'unicité (maximalité) dénotative, qui lui faisait auparavant défaut. Il convient également de souligner le caractère foncièrement cognitiviste de la notion de « bleaching ».

unité lexicale → morphème → clitique → affixe → marque zéro (Marchello-Nizia, 2006c : 40)

La fameuse phrase de Givón (1971 : 413) selon laquelle « la morphologie d’aujourd’hui est la syntaxe d’hier » résume bien ce cheminement vers de moins en moins d’autonomie prosodique, mais aussi morpho-syntaxique et sémantique.

Pour ce qui est du caractère obligatoire de l’emploi de la nouvelle unité, il est décrit par Marchello-Nizia (2006c : 53) comme suit : « [...] lorsque pour exprimer une notion grammaticale, il y a obligation à employer le nouveau morphème, et seulement celui-ci [...] ». Il nous semble que c’est le cas, par exemple, des démonstratifs CIST et CIL en ancien français, apparus à l’issue de la grammaticalisation de la combinaison des anciens démonstratifs ISTE et ILLE avec la particule *ecce* et dont l’usage est devenu obligatoire pour marquer la possibilité d’identifier le référent moyennant son indexation forte sur le contexte. Ce serait également le cas d’un article défini, apparu suite à la grammaticalisation d’un démonstratif, dont l’emploi sera désormais obligatoire pour marquer l’unicité dénotative d’un nom employé en discours (voir *infra*)⁷.

Le caractère obligatoire de l’unité va inévitablement de pair avec un accroissement de la fréquence d’apparition de celle-ci. Il est cependant à préciser qu’une fréquence d’emploi accrue semble accompagner la grammaticalisation de manière générale⁸ :

Lorsqu’une forme est en cours de grammaticalisation, les nouveaux emplois ainsi générés font que sa fréquence d’apparition dans les énoncés augmente : cette croissance du nombre des occurrences peut être l’un des signes d’une grammaticalisation à ses débuts. Ensuite, les emplois en tant que morphèmes croîtront selon un schéma que A. Kroch (1989 : 223) a nommé ‘courbe en S’ (‘S-curve’) : les emplois augmentent lentement d’abord, gagnant un contexte puis un autre ; dans un second temps la fréquence augmente rapidement et également en tous contextes, avant de ralentir, formant enfin une sorte de palier. (Marchello-Nizia, 2006c : 42)

6.1.2 Grammaticalisation, expressivité et subjectivité

Pour Meillet (1982 [1912]), l’expressivité se trouve à l’origine de tout processus de grammaticalisation. C’est le besoin de se démarquer (et de se faire remarquer) du locuteur, sa recherche d’intensité dans le langage, qui constituerait le facteur déclencheur de la grammaticalisation. Voilà donc une approche centrée sur la personne du locuteur. Avec le développement de la pragmatique, l’accent est peu à peu déplacé sur le rapport entre le locuteur et l’interlocuteur :

⁷ Dans les deux cas, l’unité qui devient plus tard obligatoire n’est, pendant un temps, qu’une option (CIL / (*ecce*) ILLE ; ILLE → LE / Ø).

⁸ Y compris donc lorsque celle-ci ne débouche pas sur une unité dont l’emploi devient obligatoire dans certains contextes.

Mais le développement de la pragmatique a infléchi la problématique posée par Meillet, la déplaçant du locuteur vers la relation entre le locuteur et l'allocutaire. C'est sur l'effet produit, et sur la volonté de le produire, qu'on insiste désormais. (Marchello-Nizia, 2006c : 25)

Dès lors, à la motivation de se démarquer succède celle d'insister sur la pertinence du message :

The reason for the ubiquity of subjectification presumably lies in the speaker's attempts to communicate the relevance of what is said to the communicative event, which includes hearers as well as speakers, but which ultimately depends for its occurrence on the speaker. (Traugott, 2005 [1995] : 46)

Traugott est ainsi la première à faire un usage systématique du terme de « subjectification », bien que ce terme ait traversé, dans ses travaux, une évolution conceptuelle. Dans le cadre de son acception la plus récente, Traugott met en avant le rôle des intentions du locuteur. Ce sont elles qui expliquent le recours à des emplois « subjectifs » d'unités linguistiques qui, au départ, étaient parfaitement neutres. Dans cette conjoncture, c'est donc le point de vue du locuteur qui apparaît comme un facteur crucial. C'est ainsi qu'un verbe de mouvement comme ALLER en est venu à acquérir sa valeur de semi-auxiliaire : au départ, le locuteur aurait eu recours à la métaphore du déplacement physique réalisé dans le but d'accomplir quelque chose ; plus tard, « l'expression d'un mouvement réel » aurait cédé la place à « l'idée d'un mouvement abstrait » (Marchello-Nizia, 2006c : 28).

À en croire Traugott, la « subjectification » serait intrinsèque à tous les processus de grammaticalisation (cf. 2005 [1995] : 46), contrairement à d'autres principes et mécanismes qui peuvent ou non être en jeu, tels que la « métaphorisation ». Cette dernière serait uniquement présente, selon Traugott, dans certains cas de grammaticalisation, et davantage comme résultat que comme principe motivant le processus. Ainsi, il est possible de parler de métaphorisation pour le cas de l'auxiliarisation du verbe ALLER, dans la mesure où l'on peut considérer que la trajectoire dans le temps (sens après la grammaticalisation) est calquée sur celle dans l'espace (sens avant la grammaticalisation). Toujours est-il que pour Traugott et d'autres (cf. Traugott et König, 1991 ; Traugott et Dasher, 2002 ; Traugott et Hopper, 2003 [1993]), cette explication est marginalisée pour ces prétendues limites au profit du rôle de la métonymie (« association in linguistic context » ; Traugott, 2005 [1995] : 36) et des inférences pragmatiques. Comme le fait remarquer Marchello-Nizia (2006c : 37) :

Cette position correspond également à celle de la pragmatique discursive (cf. Grice : une implicature dans le discours peut se figer, se conventionnaliser, se 'grammaticaliser' ; ex. : *dès lors que* : valeur temporelle ? valeur logique causale (selon le principe '*post quem, ergo propter quem*', c'est-à-dire 'après X = à cause de X').

Dans cette perspective, une unité linguistique employée dans un type de contexte donne lieu à des inférences (« implicatures conversationnelles » ; Grice, 1975⁹) qui auront tendance – dans le contexte en question – à se conventionnaliser, produisant de nouvelles valeurs sémantiques désormais codées par le système.

Un autre principe important en jeu dans la grammaticalisation porte le nom d'« extension analogique » (“analogical extension”, cf. Traugott, 2005 [1995] : 36). Ce principe explique comment le nouveau sens dérivé par inférence réussit à se répandre dans d'autres contextes, avec lesquels il était incompatible au départ (voir la phase appelée « contexte de passage » *supra*). Ainsi, les principes de métonymie-inférences et d'extension analogique seraient complémentaires en ce sens qu'ils interviendraient à des moments différents du processus. Dans le cas de l'auxiliarisation du verbe ALLER, il y a d'abord l'inférence pragmatique portant la valeur du futur, dérivée à partir de l'association, dans un contexte, du verbe de mouvement avec celui dénotant une autre activité : lorsque je signale à mon interlocuteur que je compte me déplacer physiquement afin d'accomplir une action, celui-ci aura tous les droits d'en déduire que j'accomplirai cette action dans le futur¹⁰. Une fois que l'inférence en question se sera figée, elle pourra s'étendre à des contextes dont elle était proscrite au début, à savoir ceux où le sens de mouvement du verbe ALLER est contradictoire avec le sens statique du verbe à l'infinitif (cf. *il va pleuvoir*).

Toujours est-il qu'aucun des deux mécanismes ne parvient à rendre compte de la motivation première du changement, laquelle réside dans la subjectivité du locuteur, sa volonté d'insister, devant l'interlocuteur, sur son intention d'accomplir l'action (il impose ainsi à l'interlocuteur son point de vue). Pour simplifier donc, nous pouvons dire qu'alors que le principe de métonymie-inférences et celui d'« extension analogique » répondent à la question du *comment*, la subjectivité (« subjectification » chez Traugott) répondrait à la question du *pourquoi*.

D'autres mécanismes peuvent accompagner le processus de grammaticalisation, tels que la « réanalyse ». Bien que certains (cf. Peyraube, 2002) voient dans la « réanalyse » un phénomène à extension très large et comprenant aussi la grammaticalisation, il est aujourd'hui plus communément admis que la « réanalyse » et la « grammaticalisation » doivent être dissociées, et que la première peut, dans certains cas, apparaître comme un moment de la seconde. L'évolution de la négation en français est un exemple parlant de « réanalyse » : dans *il ne marche pas*, *pas* – interprété comme un nom (du latin *passus*) complément d'objet direct du verbe *marcher* et utilisé uniquement pour renforcer le sens négatif de *ne* – est réinterprété comme morphème forclusif de la négation, ayant le

⁹ Voir la Section 7.1.3.1 pour une présentation plus détaillée de ce terme.

¹⁰ « En effet, en annonçant qu'il va quelque part pour faire quelque chose (*I am going to X to do Y*), le locuteur dit en même temps son intention de le faire. Il s'agit d'une inférence : du mouvement vers un but on infère la recherche d'un but, une intention, donc un futur. » (Marchello-Nizia, 2006c : 36)

même degré d'importance pour celle-ci que le morphème *ne*, voire, plus tard dans le registre familier, comme le pivot même de la négation après la chute de *ne*. Une fois que la réanalyse s'est produite, *pas* a vu son usage s'étendre aux emplois incompatibles avec son sens premier (ceux ne contenant pas un verbe de mouvement). Une réanalyse peut également être postulée pour la grammaticalisation du verbe ALLER en semi-auxiliaire du futur proche : d'une structure où celui-ci constitue le pivot de la prédication, l'on passe à une structure où il ne fait qu'accompagner le verbe à l'infinitif (ancien infinitif de progrédience) devenu, lui, ce pivot.

En guise de transition vers la section suivante, où nous traiterons plus spécifiquement de la grammaticalisation des démonstratifs adnominaux en articles définis, nous allons ici insister sur le fait que parmi les produits possibles de la grammaticalisation, l'on compte, entre autres, ce que Marchello-Nizia (2006c : 58) appelle l'« apparition dans la grammaire d'une nouvelle catégorie qui en était jusque-là absente ». L'auteure (2006c : 58) précise que :

Cela peut se faire soit par la création d'une nouvelle forme, unique à l'origine, soit d'un nouveau paradigme. Ce type d'innovation appartient à ce qu'E. Benveniste nommait les « transformations innovantes », « produites par la disparition ou l'apparition de classes formelles, modifiant ainsi l'effectif des catégories vivantes » (1968 : 126-7). **Ainsi par exemple, l'apparition des articles défini et indéfini dans les langues romanes pour coder la définitude, catégorie sémantique qui en latin n'était pas marquée grammaticalement – de même dans les langues germaniques, et plus récemment en finnois et en estonien** (Laury 1997)¹¹.

Ainsi, dans le cas de l'émergence des articles, la notion logico-sémantique de définitude devient également catégorie grammaticale, dans la mesure où les locuteurs se trouvent dans l'obligation de l'exprimer au moyen de ces nouvelles unités.

6.1.3 Le processus de grammaticalisation des démonstratifs en articles définis

Dans cette section, nous rappellerons d'abord certains constats généraux à propos de l'émergence des articles définis à partir de démonstratifs adnominaux. Ensuite, nous examinerons le processus du point de vue de sa motivation pragmatique, en soulignant le rôle de la subjectivité et de l'intersubjectivité, et nous finirons en abordant trois approches concurrentes du phénomène, à savoir celle qui met l'accent sur le rôle de l'emploi anaphorique, celle qui insiste sur la présence de l'ancien démonstratif avec les noms en emploi générique et, enfin, celle qui est centrée sur le rôle de l'emploi mémoriel dans le processus. Nous expliquerons pourquoi, à l'instar d'une série d'auteurs, nous nous inscrivons dans la dernière de ces approches.

¹¹ Nous soulignons.

6.1.3.1 Constats généraux

Tout d'abord, force est de constater que la catégorie des articles est plutôt rare dans les langues du monde :

From a typological viewpoint, the grammatical category of the articles is rather uncommon. According to Dryer (1989), articles would be attested in only one third of the languages of the world. Only 8 per cent would have both a definite and an indefinite article. Moreover the spread of this phenomenon is geographically very unequal, with a high incidence in (western) European languages (for an overview, see Himmelmann 1997: 195-207; Bauer 2007; Dryer 2008). From a historical viewpoint, it is established with respect to the European languages, for which written records enable us to observe long-term evolutions, that, apart from Greek, the grammatical category of articles is a recent phenomenon. Hence, the existence of this category is not a feature inherited from Indo-European. (Carlier et De Mulder, 2011 : 522)

Heine et Kuteva (2006) soulignent l'influence des langues germaniques et romanes, dotées d'articles, sur les langues se trouvant à leur contact tout en remarquant que cette influence est plus prononcée dans la partie sud de l'Europe (d'où la présence d'articles définis pleinement constitués en bulgare et en macédonien, contrairement aux autres langues slaves). Cependant, cette influence (de l'allemand, en l'occurrence) est également présente au sein du groupe des langues slaves d'Europe centrale¹², à savoir les langues slaves occidentales et le slovène, contenant toutes un démonstratif dont la grammaticalisation en article défini est entamée et plus ou moins avancée (cf. Berger, 1993 et 2008, Czardybon, 2017 et Trovesi, 2002) :

One interesting observation is that the determiners of the West Slavic languages have developed much further than the investigated non-West Slavic ones with Bulgarian as an exception. A possible reason for this could be the influence of German, which is attested for these languages. (Czardybon, 2017 : 95)

Depuis le Moyen-Âge, l'évolution culturelle et historique de ces nations (jadis les tribus des Slaves occidentaux et alpins) est en contact étroit avec des populations germanophones. Ainsi, une longue et vaste aire traverse l'ensemble du continent européen, où Slaves et Allemands ont cohabité dans des conditions sociales et politiques semblables. La situation linguistique qui a émergé dans ces régions abritant des populations mixtes a, dans une certaine mesure, favorisé la propagation du bilinguisme. Bien qu'il ne soit pas possible d'établir aujourd'hui avec précision l'ampleur et l'intensité de ce processus, il n'en demeure pas moins incontestable que les populations slaves sont, dans leur totalité, entrées en contact avec l'allemand. Parmi les classes sociales supérieures et dans les milieux urbains, l'allemand était très en vogue, sachant qu'à Prague, de même qu'à Bautzen et à Ljubljana, il était souvent préféré pour des raisons de prestige¹³. (Trovesi, 2002 : 129)

¹² Trovesi (2002 : 129) emploie le terme d'« aire géographique et culturelle d'Europe centrale » („středoevropský geografický a kulturní areál“).

¹³ Nous traduisons : „Kulturní a historický vývoj těchto národů je již od středověku (slovanské kmeny západních a alpských Slovanů) v těsném styku s německy mluvícím obyvatelstvem a přes celý evropský kontinent se takto vytvořilo dlouhé a široké pásmo, na němž za podobných sociálních a politických podmínek žili Slované a Němci vedle sebe. Jazyková situace, jež v těchto oblastech se smíšeným obyvatelstvem vznikla, podporovala do určitého stupně šíření bilingvismu. Ačkoli dnes není možné přesně stanovit do jaké míry a s jakou intenzitou tento jev probíhal, zůstává nepopíratelným faktem, že veškeré slovanské obyvatelstvo přišlo do styku s němčinou: ve vyšších společenských vrstvách a ve městech se němčina těšila značné oblibě a v Praze stejně jako v Budyšině a v Lublani byla často z prestižních důvodů upřednostňována.“

Comme nous l'avons écrit dans la Section 1.3.4. du Chapitre 1, dans la mesure où il tend à s'engager dans un processus de « polygrammaticalisation » (Diessel, 1999 : 115), le démonstratif constitue la source de plusieurs unités de la grammaire à travers les langues du monde, parmi lesquelles se trouve l'article défini. Rappelons ici également la citation de Carlier et De Mulder (2011 : 523) à ce sujet :

Even though the grammatical category of articles as such is far from being universal, the grammaticalization process that leads to its development exhibits cross-linguistic regularities: in the majority of cases, the definite article originates from a weakened demonstrative [...]

Même si d'autres sources à l'origine des articles définis sont attestées dans les langues du monde – telles que les possessifs et les pronoms personnels, voire certains lexèmes nominaux ou verbaux, dont typiquement des marqueurs de cohésion textuelle comme LEDIT (cf. Himmelmann, 2001 : 837-840 et Carlier et De Mulder, 2011 : 523) –, ces autres sources jouent clairement un rôle marginal par comparaison avec les démonstratifs. En ce qui concerne la grammaticalisation des démonstratifs en articles définis, bien que le processus soit loin d'être entièrement compris, les chercheurs semblent être d'accord sur toute une série de constats, à commencer par : 1) le fait que c'est la forme distale du démonstratif (ou la forme associée à la 3^{ème} personne) ou la forme neutre par rapport à l'expression de la distance qui constitue habituellement le candidat idéal¹⁴ et 2) le fait qu'une fois que la forme s'est engagée dans le processus, elle traverse une phase d'affaiblissement morpho-phonétique et sémantique. Sur le plan morpho-phonétique, la forme peut perdre une partie de sa substance (cf. *illum* → *le* pour le français, *þæt* → *the* pour l'anglais). De manière moins radicale, elle peut juste subir une réduction de la voyelle, voire uniquement perdre sa capacité à être accentuée, c'est-à-dire se cliticiser avec l'élément nominal qui la suit ou la précède ; cf. '*der* → *der* pour l'allemand, '*ten* → *ten* pour le polonais et certaines variétés dialectales moravo-silésiennes du tchèque (voir la Section 3.1.1.). La perte d'autonomie morpho-phonétique peut aller jusqu'à l'affixation de la forme au nom (cf. le cas du macédonien et du suédois)¹⁵. Sur le plan sémantique, nous pouvons ici parler de « bleaching », dans la mesure où le démonstratif perd notamment sa sémantique indexicale, lui permettant d'identifier un référent via son ancrage dans le contexte immédiat. En outre, il peut également perdre sa capacité à coder la distance du référent par rapport à un centre déictique, à condition que cette capacité ait été présente au départ. Tous ces phénomènes

¹⁴ Parmi les différents types de démonstratifs eu égard à l'expression de la distance, Greenberg (1978a : 61) conclut que ce sont les formes distales ou de la 3^{ème} personne qui sont les plus propices à produire des articles définis. Ce constat doit cependant être relativisé au vu, justement, des langues slaves occidentales, où la forme engagée dans un processus de grammaticalisation est toujours celle qui ne marque explicitement ni la proximité ni la distance.

¹⁵ "Schroeder (2006) shows that there is geographical continuum: in western European languages, the definite article is formally distinct from the demonstrative (e.g. English, Spanish, French); in central Europe the definite article has the same form as the demonstrative pronoun (e.g. German) [...]" (Carlier et De Mulder, 2011 : 523)

ont été bien décrits par de nombreux auteurs (cf. Carlier et De Mulder, 2006, 2010 et 2011 ; Diessel, 1999 ; Greenberg, 1978a ; Heine et Kuteva, 2002 ; Himmelmann, 1997 et 2001 ; C. Lyons, 1999).

Rappelons ici le cycle de l'article défini proposé par Greenberg (1978a ; emprunté à Carlier et De Mulder, 2006 : 97) :

Stade 0	Stade I	Stade II	Stade III
DÉMONSTRATIF > ARTICLE DÉFINI > ARTICLE NON GÉNÉRIQUE > MARQUEUR NOMINAL			

Ce schéma synthétise l'évolution du démonstratif adnominal débouchant, dans un premier temps, sur l'article défini, puis sur un marqueur de spécificité, voire, dans l'ultime stade de la grammaticalisation, sur un pur marqueur de morphologie nominale. Il faut cependant dire que le schéma souffre d'une certaine imprécision et qu'il a été remis en question depuis. Ainsi, Himmelmann (1997) reproche à Greenberg de s'être trop limité à des données provenant de langues nigéro-congolaises et dit ne pas avoir identifié, dans ses propres données, un seul exemple de transition entre les stades I et II¹⁶.

Du point de vue formel (syntaxique), deux critères de reconnaissance de la classe des articles existent :

Two essential, though not sufficient formal criteria are as follows. First, articles are grammatical elements which occur only in nominal expressions. Second, their position within such expressions is fixed (i.e. they occur either to the left or to the right of a noun but not alternatively on either side). (Himmelmann, 2001 : 832)

Il découle de cela que seuls les démonstratifs en position adnominale peuvent se grammaticaliser en articles définis. Par ailleurs, nous renvoyons le lecteur à la Section 3.1.3. pour une description plus détaillée de ce qui se produit précisément au niveau syntaxique, à savoir l'émergence d'un SN à proprement parler à partir d'un simple « complexe nominal ».

6.1.3.2 Le rôle de la subjectivité et de l'intersubjectivité

Dans leur article de 2011, consacré à la naissance des articles définis, Carlier et De Mulder écrivent :

¹⁶ En revanche, il met en avant l'existence de langues contenant des articles spécifiques mais ne contenant aucun article défini : „Sollte es sich herausstellen, dass, wie oben angedeutet, außerhalb der europäischen und afroasiatischen Sprachen keine ‚echten‘ Definitartikel zu finden sind, dann stellt sich die Frage, ob die von Greenberg postulierte Entwicklungsabfolge (Demonstrativum – Definitartikel – spezifischer Artikel) tatsächlich universal ist, oder ob nicht der Definitartikel eine europäisch-afroasiatische Idiosynkrasie darstellt. Denn wenn in Familien, in denen spezifische Artikel gut und häufig belegt sind [...] keine Spuren von ‚echten‘ Definitartikeln erkennbar sind, dann ist es nicht unwahrscheinlich, dass die Entwicklung vom Demonstrativum zum spezifischen Artikel auch ohne das Zwischenstadium Definitartikel erfolgen kann.“ (1997 : 207)

Although it is true that the grammaticalization of the definite article involves at some point a weakening of the demonstrative force of its source, more attention should be paid to the pragmatic factors that set off the grammaticalization process to understand the evolutionary pathway from distal demonstrative to definite article. (Carlier et De Mulder, 2011 : 528)

Comme il a été suggéré plus haut, la grammaticalisation serait essentiellement motivée par la volonté de celui qui parle d'imposer son point de vue afin d'insister sur une partie du message (cf. Traugott, 2005 [1995]). Dans leur monographie, Traugott et Hopper (2003 [1993]) défendent la thèse selon laquelle la subjectivité (l'expressivité) interviendrait notamment dans les premiers stades de la grammaticalisation. C'est à ce moment-là que se produit ce que les auteurs désignent par le terme de « renforcement pragmatique » (“pragmatic strengthening”). Ce terme, Traugott le définit comme le renforcement de l'attitude subjective du locuteur : « I have argued here that what is strengthened is specifically the subjective stance of the speaker. » (Traugott, 2005 [1995] : 48) Concrètement, dans le cas de la grammaticalisation du démonstratif en article défini, cela veut dire que le locuteur recourt au démonstratif pour insister sur le statut identifiable d'un référent via son indexation forte sur le contexte, même si la possibilité pour l'interlocuteur d'identifier ce référent va de soi :

[...] in a language stage where zero determination is the rule for a certain type of nominal reference, in this case definite reference, the use of a demonstrative or another linguistic expression in an article-like role is a conscious strategy by which the speaker intends to externalize for the hearer the discourse status of a referent. (Carlier et De Mulder, 2010 : 253)

C'est ce que les deux auteurs mettent en avant dans leurs études consacrées, respectivement, à l'apparition de l'article défini en latin tardif (2010) et aux premiers stades de son évolution en ancien français (2007)¹⁷. Ces études s'appuient sur *La Chronique de Frédégaire* et ses *Continuations*, écrites en Gaule entre les 7^{ème} et 8^{ème} siècles et marquées par un fort degré d'oralité. Le latin tardif se démarque par rapport au latin classique par un emploi extensif des lexèmes ILLE et IPSE en position adnominale. Comme le démontrent les textes de Frédégaire, les deux lexèmes avaient pour fonction d'insister sur la nature identifiable d'un référent discursif. Plus précisément, les auteurs relèvent qu'alors qu'IPSE avait la fonction d'insister sur l'identité du référent d'une mention à l'autre (fonction tout à fait compréhensible vu son sens originel), ILLE entraînait quant à lui l'obligation pour le lecteur d'une nouvelle identification du référent :

In Late Latin, definite reference to the participants in the narrative is normally marked by zero determination. However, when the writer wants to attract the reader's attention to the referential act, either because the precise and exact identification of the intended referent is not guaranteed, or because the referent has topical or focal

¹⁷ Les travaux d'Epstein (1993, 1994 et 1995) représentent une autre référence importante parmi ceux qui soulignent le rôle de la subjectivité dans le processus d'émergence de l'article défini en français. Or pour Epstein, les premiers emplois du défini français seraient caractérisés par la substitution totale d'une valeur « expressive » à la valeur référentielle originelle du démonstratif, postulat que nous n'acceptons pas. Nous souscrivons ainsi à la perspective de Carlier et De Mulder, selon qui ces deux valeurs, loin de s'exclure, sont étroitement liées.

status, he develops explicit markers for this purpose. He makes an extended use of *ille* and, even more markedly, of *ipse* by exploiting their original meanings: by using the identity marker *ipse*, he gives the instruction to resume a previously mentioned referent identically; by means of the demonstrative *ille*, he asks the reader to make a new identification of the referent. Since the use of these expressions is marked with respect to zero determination, it can have as such the effect of attracting attention to their referent and of conveying its topical or focal value. (Carlier et De Mulder, 2010 : 262)

Il s'ensuit de cela tout d'abord qu'IPSE serait notamment employé pour marquer la continuité d'un référent discursif, alors qu'ILLE interviendrait davantage à des moments de discontinuité. Ensuite, l'on voit que l'usage des deux proto-déterminants n'est pas uniquement affaire d'identifiabilité du référent, mais aussi d'expression de son statut thématique-rhématique et de sa prééminence discursive¹⁸. À ce premier stade de l'évolution de l'article défini, marqué par la subjectivité, se substituerait un autre stade, davantage marqué, quant à lui, par l'intersubjectivité :

However, if the use of a demonstrative *ille* signals that the writer attracts the reader's attention to the reference act – amongst other things because he is not sure that the hearer will be able to identify the referent –, this implies that the writer has a representation of the reader's text model and takes into account this representation of the reader's text model when choosing the referential expression. In as far as the demonstrative is an expression of the writer's awareness of the beliefs of the reader, it is also an expression of intersubjectivity.

(Carlier et De Mulder, 2010 : 263)

Le passage de la subjectivité (« subjectification ») à l'intersubjectivité (« intersubjectification ») représenterait un changement de perspective dans la communication du message : le centre est déplacé de la perspective du locuteur/auteur sur celle de l'interlocuteur/lecteur (Traugott, 2003 : 129). Comme le soulignent les auteurs de l'étude, tandis qu'à la première situation correspond davantage l'emploi du démonstratif (une unité de langue éminemment subjective, comme nous l'avons vu notamment dans la Section 3.5.1), la seconde situation appelle plus naturellement l'article défini.

6.1.3.3 La grammaticalisation du démonstratif en article défini : trois approches concurrentes

Si l'origine des articles définis est en général bien connue (voir *supra*), il n'en va pas de même du processus de grammaticalisation lui-même. En effet, il s'avère plutôt malaisé de décrire à la fois la trajectoire exacte qu'empruntent, dans les langues, les articles définis naissants, ainsi que les critères auxquels un ancien démonstratif adnominal doit se conformer pour pouvoir être considéré comme un article défini pleinement constitué. Aussi présenterons-nous dans cette section trois approches majeures de la transformation du démonstratif en article défini et nous expliquerons pourquoi nous souscrivons à la dernière de ces approches. De même que nous l'avons fait dans la

¹⁸ L'une des raisons pour lesquelles IPSE a finalement été abandonné dans sa fonction de (proto-)article est que contrairement à ILLE, son usage est confiné au domaine de l'endophore.

section précédente, nous suivrons ici en grande partie l’argumentation de Carlier et De Mulder (2010 et 2011).

La première approche du phénomène met l’accent principal sur le rôle de l’anaphore dans l’ensemble du processus. Ainsi, les partisans de cette approche – cf. Berger (1993), Christophersen (1939), Czardybon (1978), Givón (2001), Greenberg (1978a), Diessel (1999), C. Lyons (1999) et, dans une certaine mesure, Laury (1997) – soutiennent que l’emploi du démonstratif doit devenir systématique en contexte anaphorique, d’où il s’étendrait aux autres contextes d’emploi. C. Lyons (1999) souligne également l’importance des emplois situationnels. Par opposition à cette thèse « minimaliste », la thèse de Krámský (1972) suppose que l’usage de l’ancien démonstratif doit devenir systématique avec les emplois génériques du nom. En l’occurrence, cette thèse est beaucoup trop stricte, dans la mesure où nombre de langues considérées comme ayant des articles définis (comme l’anglais) ne les emploient pas avec la plupart des noms génériques au pluriel. Par conséquent, nous pouvons d’emblée rejeter cette thèse. Enfin, d’autres auteurs encore – cf. Himmelmann (1997 et 2001), Carlier et De Mulder (2006, 2010 et 2011) et Löbner (1985 et 2011) – soulignent l’importance des emplois qui se situent entre les définitudes « pragmatique » et « sémantique » :

Neither Lyons’ hypothesis nor the hypothesis deriving the definite article from an anaphoric demonstrative referring to a previously mentioned discourse participant accounts for this meaning shift from direct reference to indirect reference, which is crucial in the development of the definite article.

(Carlier et De Mulder, 2011 : 527)

En effet, l’hypothèse des emplois situationnel et anaphorique peut, il est vrai, rendre compte de la grammaticalisation du démonstratif en contextes de définitude « pragmatique », contextes qui peuvent accueillir à la fois le démonstratif et le défini :

2) a) Je voudrais **le** gâteau.

b) Je voudrais **ce** gâteau. (exemple inspiré par Gary-Prieur, 2011 : 66)

Nous avons vu dans la Section 1.2.1 que la seule différence entre 2a et 2b est qu’en 2b, le locuteur attire l’attention de l’interlocuteur sur un gâteau parmi plusieurs gâteaux présents dans la situation, contrairement à 2a, où le gâteau est le seul représentant de la classe des gâteaux. *Idem* pour les emplois anaphoriques :

3) a) [Une voiture]_i était rangée devant la porte. C’est [**cette voiture**]_i que j’ai prise et non **la tienne**.

b) [Une voiture]_i était rangée devant la porte. C’est [**la voiture**]_i que j’ai prise et non **la moto**.
(Corblin, 1987 : 221)

Dans les deux cas, l'emploi de l'article défini répond au principe de « contraste externe » (cf. Corblin, 1987 : 217 ; voir la Section 1.2.2), mais dans les deux cas, le référent est identifié de manière directe, à savoir grâce au contexte (situationnel ou discursif) immédiat.

En revanche, dans les exemples 4 et 5, l'identification du référent opère différemment :

- 4) Après être allés au cinéma, ils ont encore discuté *du film*.

(Czardybon, 2017 : 90 ; notre traduction)

- 5) Est-ce que tu peux descendre dans *la cave* chercher des pommes de terre ?

En 4, le référent du film est identifié grâce au lien associatif stéréotypé existant, dans une communauté de locuteurs, entre l'activité d'aller au cinéma et le film visionné dans le cadre de cet événement. En 5, le nom *cave* est doté d'un référent unique dans l'univers du discours du locuteur et de l'interlocuteur (*notre cave, la cave de notre maison*).

Himmelmann (1997) fut le premier, à notre connaissance, à souligner l'importance, pour le processus de grammaticalisation du démonstratif en article défini, de son emploi mémoriel :

Dagegen scheint es mit weitaus einfacher, einen Zusammenhang und damit einen möglichen Entwicklungübergang zwischen dem anamnestischen Gebrauch der Demonstrativa und den semantisch-definiten Gebrauchskontexten für Definitartikel zu sehen. In beiden Fällen wird auf einen im Redeuniversum ‚vorhandenen‘ Referenten verweisen, Vorerwähtheit spielt keine Rolle. **Was sich im Grammatikalierungsprozeß ändert ist, dass die für anamnestischen Gebrauch definitorische Bezugnahme auf spezifisches Wissen aufgegeben und durch den Bezug auf allgemeines Weltwissen ersetzt wird (allgemeines Weltwissen ist sowohl für assoziativ-anaphorischen wie abstrakt-situativen Gebrauch von zentraler Bedeutung). Dabei handelt es sich um einen Generalisierungs- bzw. Abstraktionsprozess, der als typisch für Grammatikalierungsprozesse gelten darf**¹⁹.

(Himmelmann, 1997 : 95)

Comme il apparaît de cette citation, l'argumentation de Himmelmann repose sur l'existence d'une contiguïté (voire une certaine zone grise) incontestable entre l'emploi mémoriel („anamnestischer Gebrauch“) d'une part et les emplois anaphoriques associatifs et sémantiquement uniques („abstrakt-situativer Gebrauch“) d'autre part. Les deux phénomènes partagent la caractéristique suivante : le référent n'est ni présent physiquement dans la situation ni mentionné dans le discours antérieur immédiat. En conséquence, son identification opère via une connaissance partagée. Or ce qui change, c'est la nature de cette connaissance, spécifique/privée dans le cas de l'emploi mémoriel²⁰ et générale dans le cas des emplois anaphoriques associatifs et définis

¹⁹ Nous soulignons.

²⁰ Voilà la raison pour laquelle ce type d'emploi du démonstratif s'accompagne fréquemment d'une relative précisant l'identité du référent et/ou de propositions incidentes du type « tu sais » (cf. Carlier et De Mulder, 2011 : 261), le locuteur s'efforçant de faciliter à l'interlocuteur la tâche qu'est l'identification du référent.

sémantiques (voir aussi la Section 4.5.3). Carlier et De Mulder (2010 : 258) adoptent la position de Himmelmann en étudiant l'émergence de l'article défini en latin tardif à partir du démonstratif ILLE :

In its recognitional use, the demonstrative *ille* is similar to the definite article to the extent that it does not require the referent to be available in the context or the immediate situation. However, there also is an important difference: in the recognitional use of the demonstrative, the referent is identified on the basis of specific knowledge that is shared by speaker and hearer; **in characteristic uses of the definite article, those where it cannot be replaced by a demonstrative determiner, the identification of the referent is based on stereotypical knowledge, or knowledge that is presumably shared by all members of the relevant speech community**²¹.

Nous tenons pour particulièrement importante la partie de la citation marquée en gras, car les auteurs y impliquent que le défini partage avec le démonstratif une sphère d'emploi (il peut ainsi, comme nous l'avons mentionné dans la Section 1.1.2 et comme nous l'avons vu tout au long du présent travail, connaître des emplois indexicaux). Mais parallèlement à cela, ils insistent sur la position selon laquelle les emplois prototypiques du défini sont justement ceux qui n'admettent pas le démonstratif.

Pour sa part, Himmelmann (1997 : 95) développe cette hypothèse en écrivant :

[...] ist dieser Gebrauchskontext [= celui de l'emploi mémoriel] m.E. für die Theorie der Artikelgenese von zentraler Bedeutung, weil hier die Grenze zwischen pragmatisch-definiten und semantisch-definiten Gebrauchsweisen überschritten wird. **Empirisch zeigt sich die Relevanz dieses Gebrauchskontexts darin, dass typischerweise nur diejenigen Demonstrativa, die in diesem Kontext konsistent gebraucht werden, zu Definitartikel grammatikalisiert werden können**²².

L'argument empirique en question représente, selon nous, un vrai argument de poids. En effet, dans la plupart des langues romanes, l'article défini a évolué à partir de ILLE, démonstratif qui était le seul à connaître des emplois mémoriels. De même en tchèque, TEN, qui évolue en article défini, représente de loin la forme la plus employée en emploi mémoriel (voir les Sections 5.2.3.1 et 5.2.4.1 du Chapitre 5).

Le principal avantage que la thèse de Himmelmann détient par rapport à la thèse concurrente « anaphorico-déictique » est donc la possibilité de rendre compte de la transition entre les définitudes « pragmatique » et « sémantique » :

As pointed out by Himmelmann (1997; 2001), the recognitional use of the distal demonstrative allows us to understand the transition between the distal demonstrative and the definite article: the distal demonstrative becomes a definite article when the presumably shared knowledge necessary to identify the referent is no longer presented as specific to the speaker and the hearer, but as consisting of information presumably shared by all members of the speech community. (Carlier et De Mulder, 2010 : 258)

²¹ Nous soulignons.

²² Nous soulignons.

Mais il y a un autre avantage que cette thèse détient par rapport à la thèse « anaphorico-déictique ». Certains partisans de cette dernière (cf. Harris, 1978 et C. Lyons, 1999) relèvent que le démonstratif en cours de grammaticalisation perd, dans les contextes anaphoriques et déictiques, sa capacité à coder la distance du référent, laquelle distance n'est plus pertinente (nous avons vu dans la Section 1.3.2 que l'expression de distance n'est qu'exceptionnelle dans le cas du défini). Cependant, si l'article défini était vraiment issu des démonstratifs apparaissant en emplois anaphorique et situationnel, plusieurs formes d'articles définis à valeur interchangeable existeraient dans la plupart des langues du monde, dans la mesure où la plupart des langues connaissent, dans ces deux types d'emploi, une alternance de formes proximales et distales ou de formes associées aux participants à l'interlocution. Ainsi, les langues romanes, par exemple, contiendraient probablement plusieurs formes d'article défini, issues des trois démonstratifs latins, ce qui n'est pas le cas²³.

Pour conclure cette section, nous pouvons dire qu'en ce qui nous concerne, nous souscrivons pleinement à la thèse de Himmelmann. Il faut néanmoins préciser trois choses. D'abord, la thèse concurrente ne saurait être complètement discréditée, dans la mesure où une présence accrue²⁴ du démonstratif adnominal dans l'anaphore constitue clairement le signe d'une grammaticalisation en cours vers l'article défini, comme ont su le démontrer dans leurs travaux des auteurs tels que Czardybon (2017), Diessel (1999), Laury (1997) et Zíková (2017 et 2018). Diessel (1999 : 128-129) résume à ce sujet :

The use of anaphoric demonstratives is usually confined to non-topical antecedents that tend to be somewhat unexpected, contrastive or emphatic [...] When anaphoric demonstratives develop into definite articles their use is gradually extended from non-topical antecedents to all kinds of referents in the preceding discourse. In the course of this development, demonstratives lose their deictic function and turn into formal markers of definiteness.

Ensuite, comme nous le verrons dans la section suivante, il existe des langues et des dialectes disposant d'un paradigme d'articles définis dits « forts », dont l'emploi se limite, justement, aux contextes de la seule définitude « pragmatique ». Compte-tenu de ce fait, il s'avère plutôt difficile de tracer une ligne de démarcation claire entre ce type d'articles définis et les démonstratifs dont ils ont évolué²⁵. De fait, cette réalité remet partiellement en question ce qui est impliqué par la citation de Carlier et De Mulder (2010 : 258) *supra*, à savoir qu'il serait seulement légitime de parler d'article

²³ “As pointed out by Greenberg (1978:61) and Himmelmann (1997:96-98) from a typological perspective, the most common source of the definite article is the distal demonstrative or the third person demonstrative. If there is free alternation between proximal and distal demonstratives in the contexts of use where the shift from demonstrative is initiated, how can we account for this widespread tendency?” (Carlier et De Mulder, 2010 : 247)

²⁴ Mais pas forcément obligatoire, voir *infra*.

²⁵ Voir également Diessel (1999 : 119) : “The division between grammatical items and endophoric demonstratives is strictly speaking an idealization. There is no clear-cut borderline that separates endophoric demonstratives from grammatical markers.”

défini une fois que l'ancien démonstratif a pénétré dans le domaine de la définitude « sémantique » pour y acquérir une place solide²⁶.

Et, enfin, comme l'écrit encore Himmelmann (1997 : 96) :

Falsch wäre es, daraus zu folgern, dass der Gebrauch in diesem Kontext [= le contexte mémoriel] die Ursache für die Grammatikalisierung zum Definitartikel ist. Er ist nur ein Indiz dafür, dass eine weitere Grammatikalisierung erfolgen kann.

Ainsi, l'emploi mémoriel n'est pas à l'origine de la grammaticalisation du démonstratif en article défini, celle-ci étant plus probablement due, d'une part, aux facteurs pragmatico-subjectifs mentionnés dans la Section 6.1.2 *supra* (l'insistance sur la nature identifiable du référent, son statut particulier dans le discours, etc.), éventuellement aussi à des facteurs plus directement linguistiques, tels que la perte de la flexion nominale à l'époque du latin tardif, du roman et de l'ancien français²⁷. En effet, il est probable que ce dernier phénomène a à son tour encouragé l'émergence et la présence de l'article, et ce via deux facteurs se succédant dans le temps. D'abord, comme le relèvent Carlier et Combettes :

[...] dans la perspective d'une cohérence quant au mode d'expression de la grammaire, l'affaiblissement progressif de la morphologie suffixale conduit à grammaticaliser davantage les marqueurs grammaticaux antéposés à la tête, à leur faire perdre le poids pragmatique. (2015 : 20)

Ainsi, comme la perte de la flexion débouche sur une linéarité de l'ordre des constituants de même que sur une plus grande hiérarchie syntaxique entre eux, le démonstratif en cours de grammaticalisation, situé dans la contiguïté du nom, en vient à perdre une bonne part de son autonomie syntaxique, sémantique et pragmatique pour former avec lui une certaine unité. Cette subordination explicite de la forme au nom qu'elle actualise a sans doute encouragé et accéléré la poursuite de sa grammaticalisation déjà commencée.

Ensuite (plus tard), le nom n'étant plus du tout en mesure de marquer, moyennant un système casuel sujet à l'érosion et au syncrétisme, les différentes fonctions qu'il remplit au sein de la phrase, ce rôle est désormais en partie relégué aux actualisateurs de ce dernier, où la déclinaison se maintient plus longtemps, avant de disparaître à son tour²⁸. Par conséquent, la langue aurait gagné à continuer

²⁶ Himmelmann (1997 et 2001) adhère à la même position.

²⁷ Pour l'évolution de la morphologie nominale du latin au français ainsi que l'impact que cette évolution a eu sur le plan de la syntaxe, voir par exemple Harris (1978) ou Carlier et Combettes (2015)

²⁸ L'allemand moderne serait, à notre avis, un bon exemple de cette évolution : bien qu'appartenant aux quelques langues germaniques où l'expression du cas a survécu, celle-ci est en vérité extrêmement rudimentaire sur le nom et considérablement affaiblie sur l'adjectif (grand degré d'isomorphisme), alors qu'elle reste plus riche sur l'article et certains autres déterminants. C'est cela même qui a permis à cette langue de garder un ordre des mots davantage régi par des contraintes pragmatiques que celui du français (cf. la possibilité de thématiser un constituant en le plaçant en début de phrase, indépendamment de son statut syntaxique). Pour le français, Schøsler (1984 et 2018) constate effectivement que la déclinaison s'est maintenue plus longtemps sur l'article que sur le nom et l'adjectif. Cependant, il reste la question de savoir dans quelle mesure cette réalité a pu avoir un impact sur l'expansion du domaine des emplois de l'article défini,

d'« investir » dans le développement de l'article, qui, vue sa présence particulièrement élevée, aurait joué un rôle important dans le marquage du cas.

Ces deux phénomènes pourraient être représentés de la manière suivante :

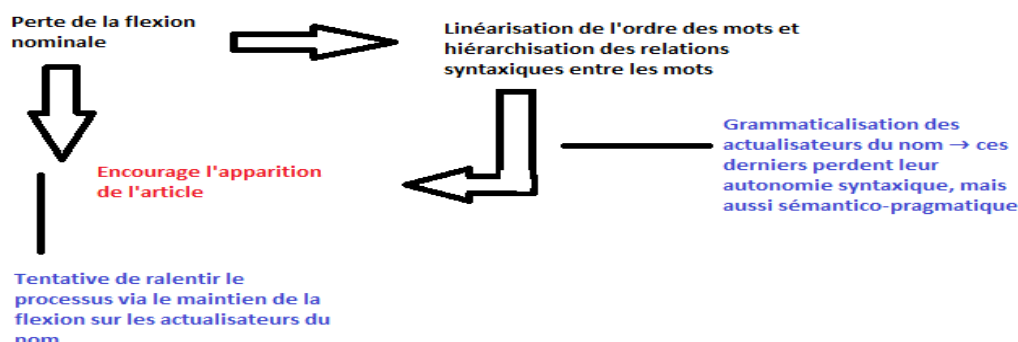


Figure 6.1 : Comment l'érosion de la flexion nominale a pu encourager le développement de l'article défini

Néanmoins, si la perte de la flexion nominale et la plus grande configurationnalité du système qu'elle entraîne (cf. Carlier et Combettes, 2015) semble encourager le développement de l'article, elle ne saurait être considérée comme sa cause première et directe. Car il faut admettre en toute honnêteté et toujours avec Himmelmann (1997 : 228) que la question de la motivation du processus reste, jusqu'à présent, sans réponse satisfaisante :

Vielmehr ist m.E. die Grammatikalisierung von artikelähnlichen Elementen [...] ähnlich zu konzipieren wie Lautwandelerscheinungen: Über die Frage, warum Sprechende zu einer bestimmten Zeit in einem bestimmten Kontext ein Demonstrativum ‚außer‘ der sonst üblichen Gebrauchskontexte gebraucht, kann praktisch nur spekuliert werden (ebenso wie über die Ursachen, die Sprechende dazu veranlassen ein /b/ stärker als üblich zu aspirieren).

6.1.3.4 Définitude « pragmatique » et définitude « sémantique »

Dans ses deux articles publiés dans le *Journal of semantics* (1985 et 2011), Löbner expose son approche de la définitude intitulée « Concept Types and Determination Theory » (CTD). Au sein de son travail se trouve la distinction entre les deux types de définitude susmentionnés²⁹. L'approche de Löbner est articulée sur une base logico-sémantique, la définitude étant entendue au sens de l'unicité. Il s'agit, selon lui, d'une propriété sémantique que certains noms, à savoir les noms dits « individuels » (“individual”) et « fonctionnels » (“functional”), possèderaient par défaut. Ces deux types de noms débouchent sur autant de concepts nominaux. Selon Löbner, ils sont a priori prédestinés à la

déjà présent dans le système au moment concerné. De plus, comme l'écrivent Carlier et Simonenko : “However, this system of case marking on determiners is already defective in Old French, since it only appears on determiners agreeing with masculine nouns, as feminine *l*-forms no longer exhibit this opposition.” (2020 : 398)

²⁹ Cette distinction s'inspire, entre autres, du travail de Hawkins (1978) et de celui d'Ebert (1971).

référence unique. Parallèlement à ces deux catégories de noms, deux autres catégories de noms existent, a priori prédestinées, elles, à la référence non-unique³⁰ : les noms « sortaux » (“sortal”) et « relationnels » (“relational”).

Il importe ici de préciser que chez Löbner, le terme d’unicité ne revêt pas tout à fait le même sens premier que chez Russell (1905) et ses successeurs. En effet, c’est davantage au sens de « non ambiguïté » qu’il faut d’abord l’entendre :

I am not happy with the term *unique(ness)*, but there does not seem to be a better choice. In German, I would use the term *eindeutig*, which means something similar to *unambiguous*. A description is ‘eindeutig’ if it can apply to only one case. A description is ‘unique’ if there is just a single case it applies to, but there may be more such cases possible. In a given context, *girl with a red hood* may amount to a unique description because it may happen that there is exactly one such case in that context; but the description would not be *eindeutig*. *Pope* is *eindeutig* because the concept ‘pope’ is such that there can be only one such person. [...] Thus, ‘Eindeutigkeit’ entails uniqueness but not conversely.” (Löbner, 2011 : 281)

Ainsi, c’est cette unicité inhérente aux noms individuels et fonctionnels qui constitue le point explicatif de départ pour le traitement des autres expressions définies. D’où également la critique faite par Löbner de l’unicité telle qu’elle est théorisée par Russell (1905), dans la mesure où celle-ci ne peut rendre compte que de la définitude des expressions accueillant des concepts sortaux et relationnels, lesquelles expressions sont uniques, mais ne sont pas *eindeutig*³¹. De même, la formule russellienne $\exists x (N(x) \ \& \ VP(x)) \ \& \ \forall y (N(y)) \rightarrow y = x$ ne peut s’appliquer aux expressions accueillant des concepts fonctionnels et individuels. En effet, comme il n’existe, dans ce cas, qu’un seul x , la partie $\forall y (N(y)) \rightarrow y = x$ de la formule est tout simplement redondante.

Dans la terminologie de Löbner (2011 : 281), les noms individuels sont définis comme suit : « Their meanings are individual concepts that assign a unique referent to every appropriate context of utterance. » Des exemples de cette catégorie seraient des noms tels que *France*, *pape* ou *hôtel de ville*. Alors que les deux premiers ressortissent à la catégorie des *unica*, le dernier ne réfère de manière unique que pour les habitants d’une ville particulière.

Quant aux noms fonctionnels :

Their meanings are functional concepts, involving one argument: the possessor. Functional concepts are restricted to a domain of appropriate possessors and of appropriate contexts of utterance. The value for a given

³⁰ Pour nous, malgré son ingéniosité, l’approche de Löbner comporte deux grands défauts. D’abord, elle ne se préoccupe guère des emplois non référentiels des SN définis (pour l’importance de la dissociation de l’unicité et de la référence, voir la Section 1.2.1, où nous avons fait nôtre la position de Donnellan (1966)). Ensuite, le terme lui-même d’unicité s’avère problématique pour les démonstratifs, ainsi que nous l’avons également vu dans la Section 1.2.1.

³¹ Quant aux expressions démonstratives comme *ce livre*, non seulement elles ne sont pas *eindeutig*, mais elles ne semblent pas non plus être uniques, car leur emploi dans une situation implique nécessairement que celle-ci accueille plus d’un cas de figure auquel l’expression peut s’appliquer.

argument, in a given context of utterance, constitutes the uniquely determined referent of the noun.
(Löbner, 2011 : 282)

Les termes de parenté appelant un seul argument possessif – tels les noms *mère* et *père* – constituent des exemples prototypiques de cette catégorie. Cependant, l'argument en question n'est pas nécessairement de nature possessive au sens strict (ce terme est donc davantage une simplification) : dans les SN *le train pour Paris* et *la différence entre toi et moi*, les noms *train* et *différence* sont tout autant des concepts fonctionnels, car saturés par des arguments syntaxiques enclenchant l'unicité de l'ensemble du SN.

Une précision d'une grande importance s'impose au sujet des noms individuels et fonctionnels, à savoir que leur référent n'est pas non plus complètement indépendant du contexte d'énonciation. Comme le précise Löbner (2011 : 284) au sujet des noms individuels³² :

The referent of an individual noun depends on the context of utterance, including a time index, a location index and a constellation of facts (i.e. a "world" in the sense of possible world semantics). The range of indices for which an individual noun refers to a particular entity may vary widely. Some have almost globally uniform reference (*US president*, *Ghana*), while others like *mosque*, *station*, *exit* and so on have a particular referent only in smaller domains. For such individual nouns, the index for fixing the unique referent has to be determined more precisely. Proper names, in particular given names, have a small range of uniform reference. This does not contradict the assumption of semantic uniqueness of proper names since uniqueness is relative to a given context of utterance³³.

Pour Löbner, les noms propres font ainsi également partie des concepts individuels, de même que les pronoms personnels anaphoriques et déictiques³⁴.

Ensuite, la théorie CTD reconnaît l'existence des noms « sortaux » :

Sortal nouns are unary predicate terms [...]. They characterize their potential referents in terms of properties. Their meanings are sortal concepts. In a given context of utterance, there may be zero, one or more entities the noun denotes. (Löbner, 2011 : 280)

Les noms *individu* et *chien* sont de bons exemples de cette catégorie.

Et, enfin, il existe une catégorie de noms appelés « noms relationnels » : « Relational nouns are binary predicate terms [...]. Their meanings are binary relational concepts, involving a further argument in addition to the referential argument » (Löbner, 2011 : 281). Font partie de ces noms les

³² Dans son étude de 1985, Löbner considérait encore les noms individuels comme une sous-catégorie des noms fonctionnels. Ce n'est que dans son étude de 2011 qu'il en a fait une catégorie à part.

³³ Dans le système de Hawkins (1978), cette distinction correspond à peu près à celle entre ce qu'il appelle « larger situational use based on general knowledge » et « larger situational use based on specific knowledge ».

³⁴ L'association des pronoms personnels avec la définitude « sémantique » pourrait paraître quelque peu incongrue et contre-intuitive vu qu'il s'agit d'unités linguistiques dont le mode référentiel dépend étroitement du contexte. Cependant, dans la théorie CTD, c'est l'unicité référentielle qui compte avant tout. Or les pronoms personnels *je*, *tu* et *il* sont toujours dotés d'une référence unique.

termes de parenté *frère, sœur, oncle, tante, cousin*, etc. Dans le cas des noms fonctionnels – qui, il faut bien le reconnaître, constituent une sous-catégorie singulière des noms relationnels –, l’argument possessif seul suffit pour déclencher l’unicité. En revanche, dans le cas des noms relationnels propres, afin que l’unicité puisse être établie, la présence d’une information empruntée au contexte immédiat est elle aussi indispensable (cf. *mon ami = l’ami à moi dont il a été question / que voici*).

Les noms individuels et fonctionnels constituent le domaine de la définitude « sémantique ». Le terme de « définitude pragmatique », quant à lui, regroupe tous les cas où un nom qui, de base, n’est pas prédisposé à la référence unique – donc un nom sortal ou relationnel – en vient à référer de manière unique grâce à l’intervention d’un contexte immédiat situationnel ou discursif³⁵. Nous sommes ici face aux emplois exophoriques et endophoriques.

Il est un aspect crucial de la théorie CTD, à savoir le fait que de nombreux noms possèdent une nature ambiguë par défaut, dans la mesure où ils peuvent donner lieu à plusieurs concepts à la fois. À titre d’exemple, le substantif *maison* débouche sur un concept fonctionnel dans le SN *ma maison*. Ici, la présence seule de l’argument possessif fait en sorte, dans le contexte d’emploi en question, que le nom réfère de façon unique. Mais le même substantif peut également aboutir à un concept sortal dans un emploi comme le suivant :

6) À Lyon, il y a *beaucoup de jolies maisons*.

Ainsi, les deux concepts coexistent de manière latente au sein de la sémantique du nom *maison*, constituant ce que Löbner (2011 : 310) désigne par le terme de « variants sémantiques » (“meaning variants”). Le choix de l’une ou de l’autre interprétation dépendra de l’usage particulier du nom, la théorie CTD rendant compte de cette fluctuation moyennant ce qu’elle appelle les « changements (glissements) sémantiques de niveau zéro » (“level zero semantic shifts”) :

The type distinction is relativized by polysemy: a noun represents a certain type only with respect to a given lexical reading. For example, *child* has a relational reading ‘direct descendant of’ and a sortal reading ‘non-adult’; *moon* has an individual reading ‘moon of the earth’ and a relational reading ‘natural satellite of’; *subject* can be a grammatical term (functional), it can mean ‘member of a state’ (relational) or ‘individual, person’ (sortal). (Löbner, 2011 : 282)³⁶

Deux autres types de changements sémantiques sont reconnus par Löbner. D’abord, les « changements de 1^{er} niveau » (“level 1 semantic shifts”), processus lors duquel des interprétations

³⁵ Nous pouvons également ajouter « mémoriel ».

³⁶ Pour Löbner, les différentes interprétations conceptuelles auxquelles un nom peut se prêter sont envisagées comme autant d’acceptions sémantiques à part entière, et le passage entre elles relève de la polysémie. Pour nous, ce parti-pris ne va pas de soi, dans la mesure où ces interprétations ne figurent pas toujours dans les dictionnaires en tant qu’acceptions autonomes du même lexème.

sémantiques sont dérivées à partir d'interprétations déjà lexicalisées, sans pour autant qu'elles soient elles-mêmes lexicalisées. Typiquement, ce phénomène comprend des noms tels que *voiture*, *appartement* ou *brosse à dent*, qui tendent à n'avoir qu'un seul référent au sein d'une situation d'emploi où ils appellent un argument possessif (*ma voiture*, etc.). Pourtant, ces interprétations fonctionnelles resteraient volatiles et non codées dans le lexique. Enfin, il y a les « changements de 2nd niveau » (“level 2 semantic shifts”), puisant dans les informations apportées par le contexte entendu au sens immédiat du terme ([...] « application of shifts that draw on particular contextual information » [...] ; Löbner, 2011 : 310). Ainsi, un nom sortal tel que *homme* aboutit à une interprétation individuelle moyennant une mention dans le discours précédent :

7) [Un homme]_i entra. [**Cet homme**]_i m'avait l'air familier.

Les changements de 2nd niveau relèvent de la définitude « pragmatique ».

Résumons ce qui vient d'être dit jusque-là : dans la théorie CTD, les noms sont répartis en quatre catégories en fonction de deux propriétés, à savoir l'unicité inhérente et la sémantique relationnelle. Par conséquent, les quatre types conceptuels résultent de la combinaison de ces deux propriétés : les noms sortaux sont, de manière inhérente, non relationnels et non uniques, les noms relationnels (propres) sont relationnels et non uniques, les noms fonctionnels sont relationnels et uniques et, enfin, les noms individuels sont uniques et non relationnels :

	Sémantique relationnelle inhérente [+R]	Sémantique relationnelle non inhérente [-R]
Unicité inhérente [+U]	noms fonctionnels (<i>mère, père, tête, cause, âge</i>)	noms individuels (<i>soleil, France, hôtel de ville, président, cuisine</i>)
Unicité non inhérente [-U]	noms relationnels (<i>frère, sœur, ami, jambe, doigt</i>)	noms sortaux (<i>chien, individu, feuille, pierre, arbre</i>)

Tableau 6.1 : Les quatre types de noms dans la théorie CTD (Löbner, 2011)

Deux autres termes sont cruciaux pour la théorie CTD : la détermination dite « congruent » et celle dite « incongruent »³⁷. Le premier terme désigne les situations où un nom est employé conformément à son (principal) type conceptuel. Dans les faits, cela veut dire qu'un nom (a priori) individuel ou fonctionnel se prête à un emploi unique, tandis qu'un nom (a priori) sortal ou relationnel se prête à un emploi non unique. Lorsque ce n'est pas le cas, nous sommes face à la détermination

³⁷ Les termes de Löbner étant en anglais, leur traduction en français n'est pas toujours aisée : « détermination concordante » versus « non concordante » ?

« incongruent ». Les SN tels que *le soleil, ma mère, l'âge de mes parents, un ami à moi et un individu étrange* sont des exemples de la détermination « congruent », contrairement à *un soleil, une mère tardive, mon ami et le/cet individu*, qui relèvent de la détermination « incongruent ».

Les deux types de définitude peuvent être représentés sur une échelle. Cette dernière contient une ligne de démarcation (marquée en jaune), une sorte de « zone tampon » constituée de cas davantage ambigus, situés à la charnière entre les définitudes « sémantique » et « pragmatique ». Il s'agit, avant tout, des noms sortaux ou relationnels rendus uniques par la présence d'une proposition relative « établissante » (cf. Hawkins, 1978 : 131), des noms sortaux (éventuellement relationnels) modifiés par les adjectifs superlatifs et les ordinaux, ainsi que des différents cas relevant de l'anaphore associative. Löbner (2011 : 320) propose l'échelle suivante :

pragmatic ↓	deictic with sortal CNPs ³⁸	
uniqueness ↓	< anaphoric with sortal CNPs	
	< sortal CNPs with establishing relative clauses	
	< functional CNPs with explicit definite possessor	
	< definite associative anaphors (DAA)	
	< individual CNPs	↑
	< proper names	↑
	< 3 personal pronouns	↑ semantic
	< 1, 2 personal pronouns	uniqueness

Cependant, cette échelle a l'inconvénient d'être à la fois trop approximative (elle n'opère pas de distinction entre les différents types d'anaphore associative, ne mentionne pas le cas des noms avec les superlatifs et les ordinaux, etc.) et trop inexacte pour le tchèque. Czardybon (2017 : 26) modifie l'échelle en l'adaptant à la situation des langues slaves occidentales. Nous trouvons le résultat plus juste pour le tchèque (parlé informel)³⁹ :

³⁸ CNP = « common noun phrase »

³⁹ Une échelle presque identique est proposée par Ortmann (2014 : 314).

pragmatic ↓	deictic sortal nouns	
uniqueness ↓	< anaphoric sortal nouns	
	< sortal nouns with complements establishing uniqueness	
	< relational DAAs	
	< part-whole DAAs	
	< complex individual concepts	
	< lexical individual and functional nouns	↑
	< proper names	↑ semantic
	< personal pronouns	uniqueness

Plusieurs choses diffèrent ici en comparaison avec l'échelle précédente. D'abord, la troisième ligne du haut ne se limite pas uniquement au cas de figure du nom modifié par une relative « établissante » (exemple 8), mais indique qu'il peut également s'agir d'autres types de compléments, comme une subordonnée conjonctive pure (exemple 9) ou certaines constructions prépositionnelles/adverbiales (exemple 10)⁴⁰ :

8) Jacques vit avec la *femme* qu'il a épousée.

9) L'*éventualité* que Marie réussira à l'examen

10) La *voisine* de dessus est passée ce matin.

Les différents types d'anaphore associative ont été présentés dans la Section 2.2.3. Il est à noter que l'échelle de Czardybon ne tient pas compte de l'anaphore associative situationnelle. En ce qui concerne les « concepts individuels complexes », cette catégorie inclut les noms modifiés par les superlatifs, les ordinaux et les cas similaires (la modification du nom par des adjectifs du type *prochain, suivant, précédent, dernier*, etc.).

Dans Dvořák (2020 : 299), nous avons proposé notre propre modification de l'échelle de l'unicité en partant de celle de Czardybon (2017) et en tenant compte de nos propres analyses des données du corpus ORTOFON v1⁴¹. Plus tard dans ce chapitre, soit dans la Section 6.2.2, nous verrons

⁴⁰ Pour tous ces cas de figure, Ortmann (2014) emploie le terme d'« autophoric reference ».

⁴¹ Dans cette étude publiée dans la revue *Naše řeč*, nous avons également employé des données du corpus ORTOFON v1, obtenues par la procédure d'échantillonnage et d'annotation exposée dans le Chapitre 2. Cependant, comme le principal sujet de l'étude était la grammaticalisation de TEN, l'échantillon contenait 1 000 occurrences de ce dernier, excluant tous les autres lemmes. Outre cela, nous tenons pour important de préciser que depuis la publication de l'étude, notre regard de chercheur a en partie évolué. Cette évolution se caractérise notamment par davantage de scepticisme à l'égard du rattachement des occurrences aux différents (sous-)types d'emploi : en revoyant nos choix d'annotation, là où le doute demeurerait patent, nous avons préféré recatégoriser comme ambiguës certaines occurrences. Aussi la catégorie des occurrences ambiguës a-t-elle gagné en taille. Ensuite, notre regard a en partie changé pour ce qui est des emplois « émotionnel » et mémoriel : alors que nous trouvons désormais que le premier est plus représenté parmi nos occurrences, le second, au contraire, l'est moins qu'il ne nous paraissait au début. Enfin, les critères d'évaluation de la grammaticalisation de TEN ont eux aussi évolué en allant vers plus de réticence et de scepticisme. Le point majeur est à cet égard le suivant : contrairement à l'étude publiée dans *Naše řeč*, nous ne considérons plus ici comme grammaticalisées

les réalités concrètes qui, à nos yeux, plaident pour le changement de l'ordre des éléments à l'intérieur de la zone en jaune. Chaque cas de figure est accompagné d'un exemple :

deictic sortal nouns (*Can you see that car?*) < anaphoric sortal nouns (*A man came in. This man looked familiar to me*) < sortal nouns with complements establishing uniqueness (*the woman whom I married*) < situational DAAs (*Yesterday, I went to the cinema but the film was a flop*) < relational DAAs (*This novel is great. I know the author personally*) < complex individual concepts (*the biggest rainforest in the world*) < part-whole DAAs (*I have bought a new car recently. The engine is already broken*) < lexical individual and functional nouns (*the Prime Minister, my mother*) < proper names (*Jack*) < personal pronouns (*I, me*)

De même que dans les deux schémas précédents, la zone charnière est marquée en jaune. Les cas de figure situés à gauche de celle-ci relèvent de la définitude « pragmatique » pure, alors que ceux situés à droite ressortissent à la définitude « sémantique » pure.

L'échelle a pour but de hiérarchiser les différents phénomènes qui y sont représentés eu égard, à la fois, à la définitude (la capacité de ces derniers à former des SN définis) et à la fréquence/probabilité d'apparition de l'article défini. Il s'agit, bien évidemment, d'une tentative heuristique et, de ce fait, nécessairement imparfaite. Comme l'écrit Czardybon (2017 : 25) :

Ortmann in particular provides typological evidence for the structuring of the scale. Still, the question remains whether the scale is universal and can capture the distribution of the definite articles cross-linguistically.

Bien que l'échelle soit susceptible de changer en fonction des langues étudiées (ainsi, nous verrons par exemple qu'en tchèque, contrairement à ce qu'indique l'échelle de Löbner (2011), les noms fonctionnels saturés par un argument possessif représentent un cas de la pure définitude « sémantique »), elle possède une double vertu. D'abord, elle permet de représenter les tendances relatives à l'apparition des articles définis en synchronie, dans les langues qui en disposent. À cet égard, l'article défini peut apparaître avec la définitude « pragmatique », domaine qu'il partage avec le démonstratif. Son apparition devient systématique, en règle générale, avec les noms modifiés par les compléments instaurant l'unicité tout comme avec les anaphores associatives, les « concepts individuels complexes », les noms fonctionnels⁴² et les noms individuels, bien que ce dernier domaine affiche déjà des irrégularités (*la Grande Bretagne* vs *Ø Great Britain* ; *le lac de Genève* vs *Ø Lake*

les occurrences déictiques situationnelles et mémorielles où TEN, en plus d'exprimer la possibilité d'identifier un référent unique de sa classe au sein du contexte, est porteur d'une valeur subjective accrue (voir les Sections 3.5.2, 6.2.1.1 et 6.2.1.3). Compte-tenu de cela, la proportion finale des occurrences évaluées comme grammaticalisées que nous donnons ici est un peu plus basse : 46,64 % vs 53,60 % de l'ensemble des occurrences de TEN.

⁴² Pour ce qui est des noms fonctionnels, l'apparition de l'article défini ou son absence dépendent en partie de la sémantique des possessifs dans la langue en question. Ainsi, l'article n'apparaît pas lorsque le possessif est lui-même porteur du présupposé d'unicité (le français moderne et l'anglais vs l'ancien français et l'italien).

Geneva). À partir des noms propres de personnes, la présence du défini se fait rare dans les langues du monde, bien que des exceptions existent. C'est parce que la définitude est à ce point inhérente à ceux-ci (le nom propre sert, avant tout, à identifier un particulier précis au sein d'un univers du discours) que son marquage en devient redondant. Ce marquage est ressenti comme encore plus redondant avec les pronoms personnels :

Explicit definiteness marking usually cuts off at some point on the scale. In standard written German and in English, proper names are used without a definite article. In colloquial German or Modern Greek, proper names do take definite articles. Very rarely, even personal pronouns receive definiteness marking. Maori personal pronouns in nontopical use and proper names receive a definite article *a* [...] (Löbner, 2011 : 320)

La seconde raison de l'utilité de l'échelle réside dans sa capacité à évaluer l'avancée du processus d'émergence et de propagation de l'article défini en diachronie. En effet, comme le résument Hofherr et Zribi-Hertz, (2014 : 9) :

From a diachronic viewpoint, definite determiners indeed spread along the scale [...] from deictic and anaphoric uses to semantically unique uses like associative anaphora, individual nouns (*the sun*) and generic NPs (*the Panda*)⁴³.

Ortmann (2014 : 298) fait remarquer dans son étude appliquant la théorie CTD à une certaine quantité de langues du monde⁴⁴ :

If a language employs definiteness markers in only some contexts but not in others, these contexts will always comprise those of pragmatic uniqueness, in which the functional load of the article is obvious for it denotes a shift from SN/RN to IC/FC⁴⁵.

Par conséquent, afin d'évaluer le degré de grammaticalisation d'un démonstratif, il est fort utile de voir quel niveau il atteint sur l'échelle de l'unicité et si oui ou non sa présence est systématique avec l'emploi qui y correspond. Hormis une hausse significative de la fréquence du démonstratif en emplois exophorique et endophorique, voire sa présence obligatoire dans ces contextes, la meilleure preuve d'une grammaticalisation en cours est l'apparition du démonstratif dans des contextes se rapprochant de la définitude « sémantique » et admettant mal, pour cela, les démonstratifs qui conservent l'intégralité de leur fonctionnement indexical.

Un autre argument empirique fort parle en faveur de la distinction entre les deux types de définitude, à savoir l'existence, dans certaines langues et dialectes, de deux, voire trois paradigmes d'article défini. Des auteurs travaillant sur les langues germaniques et leurs divers dialectes se sont

⁴³ La théorie CTD ne rend cependant pas compte du fonctionnement des noms en emploi générique.

⁴⁴ Des langues germaniques, des langues slaves occidentales, l'ancien géorgien et le maori.

⁴⁵ SN = sortal nouns, RN = relational nouns, FC = functional nouns, IC = individual nouns.

aperçu qu’un nombre considérable d’entre eux contiennent deux formes d’article défini, dont une plus réduite sur les plans phonétique et morphologique. Cette forme « faible » manifeste une grande affinité pour les contextes de définitude « sémantique », alors que la forme plus étoffée, dite parfois « forte », s’emploie notamment dans les contextes de définitude « pragmatique » :

In various German dialects, as well as other languages, the difference between semantic definite reference and pragmatic definite reference is reflected in the use of two different article paradigms. While a *strong* (i.e. morphologically full) article is used with pragmatic definites, a *weak* (i.e. morphologically reduced) article is used with semantic definites. (Studler, 2014 : 146)

Pour la première fois, ce phénomène est relevé et décrit avec méticulosité par Ebert (1971) pour le frison, l’un des dialectes du fering. Dans ce dialecte, deux paradigmes d’article défini existent : l’article-D et l’article-A. Le premier est utilisé avec la définitude « pragmatique » pure, mais aussi dans les cas où le nom est rendu unique par la présence d’une relative « établissante », tandis que le second est utilisé avec la définitude « sémantique » pure et l’anaphore associative. Les dialectes de l’allemand parlés en Suisse contiennent également un double paradigme. Grâce au Tableau 6.2, emprunté à Studler (2014 :144), nous pouvons ainsi observer le continuum qui existe entre ces deux formes et le démonstratif dont elles sont toutes les deux issues⁴⁶ :

	Masculin	Féminin	Neutre
Article défini « faible »	<i>de</i> Maa (l’homme)	<i>d</i> Frou (la femme)	<i>s</i> Chend (l’enfant)
Article défini « fort »	<i>dä</i> Maa	<i>di</i> Frou	<i>das</i> Chend
Démonstratif	<i>dää</i> Maa	<i>die</i> Frou	<i>daas</i> Chend

Tableau 6.2 : Démonstratif et Double paradigme de l’article défini dans les dialectes allemands de Suisse (Studler, 2014 : 144)

Le maori, une langue polynésienne appartenant à la famille des langues austronésiennes, contient même trois formes d’article défini, dont la plus réduite est employée avec les noms propres et les pronoms personnels (cf. notamment Bauer, 1993). C’est de nouveau Ortmann (2014) qui, se basant sur son échantillon de langues étudiées, élabore la notion de scission quant au marquage de la définitude par l’article. Cette notion s’applique aussi bien à des langues disposant d’un article défini pleinement constitué qu’à des langues où un article est seulement en phase d’émergence. L’auteur

⁴⁶ En allemand standard, la forme « faible » se retrouve avec les amalgames associant l’article défini avec certaines prépositions ; cf. : *zu + der* → *zur*, *in + dem* → *im* (cf. Schwarz, 2009).

explique que cette scission peut être de deux types : la « scission du type I » et la « scission du type II »⁴⁷ :

Split I: Pragmatic uniqueness is marked by the definite article, whereas semantic uniqueness is unmarked. The use of demonstrative pronouns as definite articles in some West Slavic languages essentially covers the contexts of pragmatic uniqueness.

[...]

Split II: Pragmatic and semantic uniqueness are distinguished by different forms, either lexically or phonologically. The contrast of strong and weak definite articles in dialects of German is shown to reflect this distinction [...]. (Ortmann, 2014 : 294)

En ce qui concerne le tchèque (parlé) informel, nous nous efforcerons de démontrer dans le second volet du présent chapitre (Section 6.2) qu'il se rapproche le plus d'une « scission du type I » et nous nous appliquerons à décrire la situation relative à la grammaticalisation du démonstratif TEN à l'appui de nos données empiriques.

6.1.4 Le statut du lexème TEN en tchèque : démonstratif ou article ?

Nombre d'auteurs (cf. Adamec, 1983 ; Berger, 1993 ; Krámský, 1972 ; Machač et Zíková, 2014 ; Mathesius, 1926 ; Meyerstein, 1972 ; Orlandini, 1981 ; Souček, 2011 ; Trovesi, 2002 ; Vey, 1946 et Zíková, 2017 et 2018) ont jusqu'à présent remarqué que le tchèque, qui continue globalement à être considéré comme une langue dépourvue d'articles (cf. Dryer et Haspelmath, 2013 ; Haspelmath, 2001), serait ou pourrait être en train d'acquérir un article défini. Ainsi, Machač et Zíková (2014 : 99) écrivent :

[...] it seems reasonable to argue that the Czech demonstrative *ten* is on its way to becoming a definite article – indeed, on a way which stretches over hundreds of years.

Citons également Trovesi (2002 : 135), qui écrit ceci au sujet de TEN adnominal en tchèque et du TÓN adnominal en haut-sorabe :

[L]es pronoms *ten* et *tón* apparaissent avec une fréquence élevée. Il est donc possible d'affirmer que ces pronoms occupent, à l'échelle évolutive, la place entre la catégorie des pronoms démonstratifs au sens étroit du terme et l'article défini. En même temps, il n'est pas évident de déterminer si ce processus de grammaticalisation continue ou s'il s'est figé dans les emplois décrits *supra*⁴⁸.

À cet égard, il a été relevé que le démonstratif TEN – réalisé, dans la plupart des cas, de manière non accentuée et neutre au regard de la distinction spatiale – apparaît dans des contextes n'admettant

⁴⁷ Ne sont évidemment pas concernées, par cette scission, les langues où le démonstratif n'a pas du tout entamé sa grammaticalisation en défini (cf. le russe) et celles où il n'existe qu'une forme de défini, employée avec les deux types de définitude (cf. le français).

⁴⁸ Nous traduisons : „[V]ýskyt zájmen *ten* a *tón* má vysokou frekvenci, lze tedy tvrdit, že jsou tato zájmena na vývojovém stupni mezi kategorií ukazovacích zájmen v užším slova smyslu a určitým členem. Přitom však není zřejmé, zda tento gramatikalizační proces bude pokračovat anebo zda ustrnul ve výše popsaném užití.“

pas ou plus difficilement d'authentiques démonstratifs. Ces contextes ressortissent, comme nous l'avons vu dans le Chapitre 3, aux régimes de l'exophore situationnelle, de l'endophore et de l'exophore mémorielle. Son rapprochement avec l'article défini semble le plus évident en contexte situationnel. Rappelons ici les constats de Mathesius (1926 : 39) et de Berger (1993 : 463) :

Tout d'abord, force est de souligner que dans le tchèque parlé, le pronom épithète *ten* n'a plus la capacité à pointer de manière autonome dans une situation extérieure. Dans cette fonction, il a été complètement évincé par les formes composées (*tenhle*, *tenhle ten*, *tuhle ten*, *tadyhle ten*, y compris, dans les dialectes, *tu ten*, *tam ten*, *tamhle ten*)⁴⁹.

Wenn wir *ten* außer Acht lassen, das, wie ich gezeigt habe, nie wirklich deiktische Funktion hat [...] ⁵⁰

Cette difficulté de TEN à fonctionner comme un démonstratif en contexte situationnel (à l'opposé des autres contextes) est sans doute une conséquence du fait que c'est précisément en contexte situationnel que la différence en termes de distance et, de manière plus générale, le marquage du contraste entre plusieurs référents appartenant à la même catégorie, est la plus importante : TEN n'est plus employé comme démonstratif, car sa neutralité en matière d'expression de la distance est, dans ce type de contexte, perçue comme une déficience. Zíková (2017 : 132) conclut d'une manière assez semblable à l'égard de l'emploi situationnel de TEN :

En tchèque, la référence aux objets se trouvant dans la sphère immédiate du locuteur et étant identifiables de façon univoque (car le locuteur est en contact direct, sensoriel, avec eux) doit emprunter l'expression *ten*⁵¹.

Comparant la sphère d'emploi du démonstratif tchèque à celle du démonstratif russe, Adamec (1983 : 158-159) parvient à la conclusion suivante :

S'il est tout à fait évident, dans la situation, de savoir quel est l'objet visé par le locuteur (par exemple, lorsqu'il y figure comme le seul objet de son espèce), le pronom peut ne pas être utilisé du tout. Et c'est précisément sur ce point qu'existe une certaine différence quantitative entre les deux langues étudiées : en russe, nous trouvons un nombre beaucoup plus élevé de non-emplois du pronom au sein de la situation décrite qu'en tchèque [...] Il est possible d'en déduire qu'en tchèque, le pronom *ten* est souvent employé avec une fonction déictique même là où ce n'est pas absolument indispensable du point de vue des besoins de la communication et, par cela même, qu'il y acquiert, en quelque sorte, le statut d'article défini⁵².

⁴⁹ Nous traduisons : „Předně je důležité vytknouti, že v hovorové češtině přívlastkové zájmeno *ten* nemá už schopnosti samostatně ukazovati ve vnější situaci. V tom úkolu je zcela zatlačily tvary složené (*tenhle*, *tenhle ten*, *tuhle ten*, *tadyhle ten*, dial. i *tu ten*, *tam ten*, *tamhle ten*).“

⁵⁰ Rappelons que cette conclusion doit être relativisée par le constat empirique de Jodas (2010), à savoir que TEN peut toujours fonctionner comme un vrai démonstratif en emploi situationnel chez certains locuteurs moraves et silésiens.

⁵¹ Nous traduisons : „K předmětům, které se vyskytují v bezprostředním dosahu mluvčího a jsou jednoznačně identifikovatelné (neboť je s nimi mluvčí v přímém, smyslovém kontaktu), musí být v češtině referováno výrazem *ten*.“

⁵² Nous traduisons : „Jestliže je z aktuální konsituace zcela zřejmé, o který předmět mluvčímu jde (figuruje-li v ní např. jediný předmět svého druhu), nemusí být zájmena užito vůbec. A právě v tomto bodu existuje jistý kvantitativní rozdíl mezi oběma sledovanými jazyky: v ruštině nacházíme daleko větší počet neužití zájmena v popsané situaci nežli v češtině [...] Z toho lze vyvodit, že v češtině se užívá zájmena *ten* v deiktické funkci často i tam, kde to z hlediska sdělovacích potřeb není zcela nezbytné, čímž toto zájmeno v češtině poněkud nabývá charakteru určitého členu.“

L’auteur démontre que cette différence entre les deux langues ne se limite pas à la seule sphère situationnelle. L’exemple suivant, emprunté à la pièce de K. Čapek *La maladie blanche*, met en scène un emploi anaphorique de TEN. Il y est question d’un procès en justice, dont l’identification ne pose aucun problème, car il a été mentionné plusieurs fois en amont :

11) Co	potřebu-je-te,	abyste	<i>ten</i>
que.ACC	avoir besoin-PRS-2PL	que.CONJ+AUX.COND.2PL	ten.M.ACC.SG
<i>proces</i>	vyhrá-l?		
procès.ACC.SG	gagner-PTCP.SG.M		

Que vous faut-il pour gagner le procès ? (K. Čapek, emprunté à Adamec, 1983 : 160)

Contrairement au tchèque, où la présence de TEN est tout à fait naturelle, le russe recourt beaucoup plus spontanément, dans ce type de situations, au substantif nu.

De manière générale, la présence élevée de TEN en emploi anaphorique a été démontrée par plusieurs études (cf. Adamec, 1983 ; Berger, 1993 ; Czardybon, 2017 ; Křížková 1971 ; Nědolužko, 2004 ; Štícha, 1999 ; Souček, 2011 ; Trovesi, 2002 ; Zíková, 2017 et 2018⁵³). Toutefois, malgré cette présence accrue, TEN n’est pas obligatoire dans l’anaphore, et sa présence ou absence au sein de ce régime d’indexicalité reste en partie régie par des contraintes diverses et pas nécessairement liées à la perspective de la grammaticalisation (voir la Section 6.2.1.2 *infra*).

Pour ce qui est de l’emploi mémoriel, c’est là que TEN occuperait une place particulièrement importante, voire – à en croire certains auteurs (cf. Adamec, 1983 ; Nědolužko, 2004) – bien plus importante que dans les autres langues slaves, y compris occidentales⁵⁴. Or un taux de présence élevé du démonstratif dans ce contexte semble être un trait qui caractérise les langues où ce dernier a entamé sa voie vers l’article défini (cf. Laury, 1997 pour le finnois). Plus loin (Section 6.2.1.3), nous verrons qu’il est des énoncés où TEN semble quasi-obligatoire dans ce type d’emploi.

En plus de ces trois types de contextes, ressortissant tous à la définitude « pragmatique », TEN tend à proliférer dans les contextes marquant la transition, sur l’échelle de l’unicité, entre les définitudes « pragmatique » et « sémantique ». Ces contextes sont : 1) ceux où le nom est rendu unique par la présence d’une subordonnée ou certains tours prépositionnels/adverbiaux (les « compléments établissant l’unicité » dans la terminologie de Czardybon (2017 : 26)), 2) ceux où le nom est rendu unique par la présence d’un superlatif, un ordinal ou un terme semblable (les « concepts individuels complexes » ; Czardybon, 2017 : 26) et, enfin, 3) certains contextes

⁵³ Zíková (2018 : 119) écrit que dans son matériel empirique, composé uniquement de narratifs informels sur un thème imposé (quinze participants devaient raconter trois histoires courtes après les avoir visionnées), TEN apparaît avec plus de 60 % des complexes nominaux référant à une entité qui a été mentionnée au moins une fois dans le discours antérieur.

⁵⁴ Nos données empiriques étayaient cette conclusion : 284 des 967 occurrences de complexes nominaux introduits par TEN ont été évaluées comme relevant de l’emploi mémoriel (soit 29,4 %).

d'anaphore associative⁵⁵. La présence de TEN avec les noms modifiés par des superlatifs a déjà été relevée par Zubatý dans ses brèves études de 1916 et 1920. Selon l'auteur – en partie influencé par la longue tradition du purisme tchèque⁵⁶ –, l'usage de TEN serait uniquement correct dans cette position lorsque le complexe nominal est employé en emploi référentiel. Dans ce dernier cas, TEN serait un (pronom) démonstratif. En revanche, lorsque le complexe nominal ne réfère pas (ne permet pas l'identification d'une entité précise), l'emploi de TEN est jugé comme étant un germanisme et serait fautif :

La présence du pronom *ten* avec le superlatif des adjectifs est peut-être le seul cas de figure où nous employions ce pronom par une imitation servile de l'article défini allemand ; [partout ailleurs] où ce germanisme est pointé, il en va sans doute toujours ainsi à cause d'une compréhension imparfaite du *pronom* (et non pas de l'*article*), qui est légitime en tchèque, et où c'est plutôt son omission qui est fautive [...]. Avec le superlatif de l'adjectif, c'est avec ténacité que le mot *ten* subsiste dans le rôle d'article défini, utilisé de la sorte non seulement par nos commerçants, habitués à réfléchir à l'allemande au sujet des affaires (c'est la meilleure marchandise, la chose la plus chère est souvent celle qui coûte le moins, etc.), et par des personnes semblables, mais aussi par certains de nos écrivains et journalistes⁵⁷. (Zubatý, 1920 : 74)

Le critère qui permettrait, selon l'auteur, de déterminer que TEN est employé correctement avec le superlatif serait la possibilité de le remplacer par les formes plus fortes sur le plan indexical, à savoir TENTO et ONEN. Pour notre part, nous devons dire que cette position ne nous convainc guère. En effet, que le complexe nominal au superlatif rende possible l'identification d'une entité précise ou pas, la présence de TEN – ou celle de l'article défini allemand ou français – est appelée justement par le seul sens restrictif du superlatif, ce qui n'est pas le cas lorsqu'un « vrai » démonstratif est employé. Il suffit de comparer les exemples 12a et 12b pour s'apercevoir que les deux situations ne sont pas interchangeables (sans même s'interroger sur le degré d'acceptabilité de 12b) :

12) a)	to	ne-jlepš-í	zbož-í
	ten.N.NOM.SG	SUP-meilleur-N.NOM.SG	marchandise-NOM.SG
	la meilleure marchandise		
b)	(?) toto	ne-jlepš-í	zbož-í
	tento.N.NOM.SG	SUP-meilleur-N.NOM.SG	marchandise-NOM.SG
	(?) cette meilleure marchandise		

⁵⁵ Pour des exemples des deux premiers phénomènes, voir la Section 3.3.3. Pour un exemple du troisième phénomène, voir la Section 3.3.2.

⁵⁶ Pour cette notion, voir par exemple Berger (2008) ou Garvin (1983).

⁵⁷ Nous traduisons : „Zájmeno *ten* u 3. stupně přídavných jmen je snad jediný doklad, kde toto zájmeno klademe otrockým napodobením německého určitého členu; kde tento germanismus jinde hledávají, děje se tak snad všude nedokonalým porozuměním *zájmenu* (nikoli *členu*) v češtině oprávněnému, a chybou bývá spíše jeho vypouštění [...]. U 3. stupně přídavného jména se slovo *ten* v úloze určitého členu drží velmi houževnatě, užívají ho tak nejen na př. naši obchodníci, zvyklí o obchodních věcech mysliti po německu (to je to nejlepší zboží, ta nejdražší věc bývá ta nejlacinější atp.), a jiní lidé podobní, ale i naši někteří spisovatelé a novináři.“

Quelques années plus tard, la position de Mathesius (1926) paraît bien plus souple, car ce dernier reconnaît le phénomène en question comme un emploi légitime de TEN, de la même manière que lorsque TEN anticipe une relative « établissante » :

Ten serait également doté d'une fonction distinctive lorsqu'il est employé avec le superlatif ou une expression superlative, que ce dernier apparaisse de façon autonome, avec un sens de substantif, par ex. : *C'est le meilleur (le pire) ; c'est toujours le premier (la première chose) ; c'est déjà le dernier (la dernière chose)*, ou avec une proposition relative, par exemple : *Il m'a percuté sur le côté, et ça, c'est le pire qui puisse t'arriver ; c'est la meilleure qualité qu'on fabrique*⁵⁸. (Mathesius, 1926 : 39-40)

Ce à quoi il ajoute aussitôt (1926 : 40) que de telles constructions sont fortement marquées sur le plan émotionnel :

Néanmoins, de telles phrases, comme l'implique la nature du superlatif elle-même, font également apparaître une forte coloration émotionnelle⁵⁹.

Nous tâcherons de mieux explorer cette hypothèse dans la Section 6.2.2.3 *infra*, au moment où nous traiterons de la présence de ce cas de figure parmi nos données.

Avec la définitude « sémantique » pure, en revanche (les « concepts individuels et fonctionnels lexicaux »), TEN ne sert pas à marquer l'unicité, mais un engagement émotionnel du locuteur avec le référent, comme nous l'avons vu dans la Section 3.5.3. C'est cela même que constate Czardybon (2017 : 75) au sujet du fonctionnement de TEN dans ces contextes en polonais :

The first reason for the occurrence of *ten* in cases of semantic uniqueness is when the speaker is emotionally affected. This can be a positive or negative emotion such as happiness or anger. In such an emotional context, *ten* can occur regardless of the concept type since *ten* functions to indicate emotional involvement and not unique reference⁶⁰.

Se pose ainsi la question de savoir quel est vraiment le statut de TEN en tchèque. Pour certains, TEN serait déjà un article défini pleinement constitué. Vey (1946 : 42) est d'avis que « *[t]en* est employé couramment en valeur d'article dans la langue parlée et dans tous les dialectes. » Un point de vue similaire est adopté par Hlavsa (1975), selon qui TEN adnominal serait un équivalent de l'article défini anglais. Or une telle conclusion semble inadmissible au vu de ce qui vient d'être dit, à savoir l'impossibilité pour TEN de marquer l'unicité dénotative d'un concept « sémantiquement »

⁵⁸ Nous traduisons : „Rozlišovací funkci má asi zájmeno *ten* také u superlativu nebo u výrazu superlativního, ať kladeného samostatně ve významu substantiva, na př.: *To je při tom to nejlepší (nejhorší); to je vždycky to první; to už je to poslední*, ať ve spojení s větou relativní, na př.: *Vrazil do mne s boku, a to je to nejhorší, co se ti může stát; to je ta nejlepší jakost, která se vyrábí.*“

⁵⁹ Nous traduisons : „Ovšem v takovýchto větách, jak to s sebou nese sama povaha superlativu, ozývá se silně i zabarvení emocionální.“

⁶⁰ La situation est la même lorsque TEN est employé avec un nom à valeur générique.

unique. Les emplois « émotionnels » mis à part, les « concepts individuels et fonctionnels lexicaux » apparaissent sans TEN en tchèque, à la différence des langues comme le français et l'anglais :

13) ENG : <i>the sun</i>	<i>the president of the French Republic</i>
FR : <i>le soleil</i>	<i>le président de la République française</i>
CZ : <i>slunce</i>	<i>prezident Francouzské republiky</i>

Les auteurs souscrivant à une position plus nuancée sont toutefois plus nombreux, à commencer par Mathesius, dont l'étude susmentionnée se termine par le constat suivant, fort pertinent à notre avis :

L'emploi fréquent du *ten*, *ta*, *to* adnominal dans le tchèque parlé peut mener à la question de savoir si ce dernier n'acquiert pas déjà, dans certains cas, le caractère de **l'article défini**. Que l'affaiblissement du sens déictique ait parfois avancé très loin est évident du fait que lors de la traduction dans une langue étrangère telle que l'anglais, il suffit souvent d'utiliser l'article défini à la place du pronom *ten*. Il en va ainsi, avant tout, du pronom distinctif *ten* avec les expressions comparatives, avec le superlatif et avant les propositions relatives déterminatives. Dans ces trois cas de figure, *ten* adnominal s'approche réellement, dans le tchèque parlé, de l'article défini. Mais il ne fait que s'en approcher pour l'instant, car son emploi peut toujours s'expliquer par le contexte phrastique ou situationnel. Il est seulement possible de parler d'un vrai article lorsque son usage découle nécessairement du sens du substantif lui-même, qu'il soit question de détermination singulière (article individualisant) ou générale (article générique). Or rien de tel ne peut être constaté même dans le tchèque parlé⁶¹. (Mathesius, 1926 : 41)

Nous trouvons une conclusion similaire dans l'étude de Trovesi (2002 : 135), consacrée à la grammaticalisation du démonstratif en tchèque, en haut-sorabe et en slovène, tout comme dans celle de Souček (2011 : 71), se basant sur le corpus parallèle InterCorp du CNT et comparant la fréquence et la distribution des démonstratifs adnominaux TEN et SE en tchèque et en finnois à celles du défini anglais *the* :

Le caractère non-obligatoire et expressif de *ten* en tchèque et de *tón* en sorabe ne permet pas de considérer ceux-ci comme des articles définis grammaticalisés, comme c'est le cas des langues d'Europe occidentale⁶².

⁶¹ Nous traduisons : „Hojně užívání přívlastkového *ten*, *ta*, *to* v hovorové češtině může vésti k otázce, nenabývá-li již v některých případech povahy **určitého členu**. Že oslabení deiktického významu tu postoupilo leckdy hodně daleko, je patrné z toho, že při překladu do cizí řeči, na př. do angličtiny, mnohdy stačí za české zájmeno *ten* užití určitého členu. Je to zejména při rozlišujícím zájmeně *ten*, u výrazů komparativních, u superlativů a před diferencujícími větami relativními. V těchto třech případech opravdu se přívlastkové *ten* v hovorové češtině blíží určitému členu. Ale teprve se mu blíží, neboť vždy ještě se dá jeho užití vysvětliti zdůrazněním větné a situační souvislosti. O skutečném členu se může mluvit teprve tehdy, když jeho užívání nutně vyplývá už z významu samotného substantiva, ať jde o určitost singulární (člen individualisující) či obecnou (člen generický). Nic takového však se ani v hovorové češtině nedá zjistiti.“

⁶² Nous traduisons : „Neobligatornost a expresivita *ten* v češtině a *tón* v lužické srbštině nedovolují, aby tato zájmena byla považována za gramatikalizované určité členy, jak je tomu v západních evropských jazycích.“

L'analyse des fonctions choisies des pronoms démonstratifs et de l'article défini en tchèque, en finnois et en anglais n'a pas montré que les pronoms démonstratifs des deux premières langues remplissent, de façon systématique, la fonction de l'article défini⁶³.

D'autres auteurs adhérant à la même position (cf. Berger, 1993 ; Czardybon, 2017 et en partie aussi Zíková, 2017 et 2018) s'appuient tout particulièrement sur l'« argument anaphorique » (voir la Section 6.1.3.3) : TEN ne serait pas un article défini dans la mesure où son emploi avec l'anaphore n'est pas systématique.

S'agissant du discours grammatical (des grammaires de référence), force est de remarquer que même si la tendance générale est toujours celle de traiter TEN avec les démonstratifs, il existe tout de même des voix qui commencent à tenir compte de l'ambiguïté démonstratif-article. Ainsi, Cvrček *et al.*, proposant une grammaire du tchèque non normative basée sur des données du Corpus national tchèque (le corpus de tchèque écrit SYN2005 et le corpus de tchèque parlé informel ORAL2006), écrivent :

Dans le tchèque parlé, le déterminant *ten* possède une fréquence et des fonctions similaires, en position adnominale, à l'article défini *the* en anglais, même si dans le tchèque écrit, il n'est que rarement utilisé dans la détermination nominale. Les deux formes du tchèque s'influencent mutuellement, d'où probablement le fait que la détermination au moyen du pronom *ten* n'est pas figée, dans le tchèque parlé, de la même façon qu'elle l'est dans les langues à articles définis [...] ⁶⁴ (2015 [2010] : 176)

Le pronom *ten* est **souvent** utilisé, dans la langue parlée, comme un déterminant des substantifs, exprimant ainsi la définitude, similairement à l'article *the* en anglais [...] ⁶⁵. (2015 [2010] : 262)

C'est, le lecteur l'aura compris, le mot *souvent* (mis en gras par nous) qui fait toute la différence.

Dans la suite du présent chapitre, nous tâcherons d'apporter notre propre contribution à ce débat complexe. Dès l'abord, que l'on nous permette néanmoins de formuler ici **une série de critères et d'hypothèses relatifs à l'évaluation du statut de TEN vis-à-vis de la grammaticalisation**, qui nous ont servi de fil conducteur dans l'ensemble de nos analyses. Les critères reflètent une certaine idiosyncrasie des différents types et sous-types d'emploi, déjà relevée dans la Section 3.3 du Chapitre 3. Dans le domaine de l'emploi situationnel, nous étions parti du critère de la présence obligatoire de TEN ; par conséquent, nous le considérons comme grammaticalisé dès lors que son absence aurait fait basculer l'emploi soit dans la catégorie des emplois « sémantiquement » définis

⁶³ Idem : „Analýza vybraných funkcí ukazovacích zájmen a určitého členu v češtině, finštině a angličtině neukázala, že by v prvních dvou jazycích demonstrativní zájmena plnila systematicky funkci určitého členu.“

⁶⁴ Nous traduisons : „V mluvené češtině má determinátor *ten* v pozici před substantivem podobnou frekvenci i podobné funkce jako určitý člen *the* v angličtině, v psané češtině se ale k určení substantiv užívá jen zřídka. Obě formy češtiny se ovlivňují, a patrně proto není determinace zájmenem *ten* v mluvené češtině ustálena jako v jazycích s určitým členem [...]“

⁶⁵ Nous traduisons : „Zájmena *ten* se užívá v mluveném jazyce často jako determinátoru substantiv, vyjadřuje tedy určitost podobně jako v angličtině člen *the* [...]“

soit dans celle des emplois indéfinis (voir la Section 6.2.1.1). Comme nous allons le voir, deux cas font exception à cette règle : a) l'emploi situationnel de TEN exprime une subjectivité accrue du locuteur et b) la forme est accentuée, fonctionnant ainsi comme les démonstratifs TENTO/TENHLE (voir la Section 3.1.1). En ce qui concerne l'emploi anaphorique, les critères sont plus complexes. D'abord, il convient de distinguer entre l'anaphore fidèle et les autres types d'anaphore, sachant que la grammaticalisation a nettement moins de chance de se produire au sein de ces derniers, où TEN joue souvent un rôle important dans le maintien de la coréférence. Au sein de l'anaphore fidèle, deux cas de figure excluent également une lecture de TEN en article défini : a) lorsque TEN est employé pour reprendre un référent venant tout juste d'être introduit dans le discours afin de l'établir dans son accessibilité référentielle (d'où notre décision de parler ici de démonstratif « établissant ») et b) lorsque TEN reprend un référent qui, bien que déjà établi dans le discours, a depuis été abandonné par ce dernier et y revient après quelque temps – le recours au démonstratif a ainsi pour rôle de réactiver le référent (d'où notre décision de parler ici de démonstratif « réactivant »). De même que dans le cadre de l'emploi situationnel, doivent aussi être exclues de toute considération en termes de grammaticalisation toutes les occurrences où TEN est le marqueur d'une subjectivité accrue (comme nous allons le voir, ce cas correspond aux autres types que l'anaphore fidèle). Les occurrences où TEN anaphorique se combine avec d'autres actualisateurs du nom doivent être considérées comme un cas à part. Cela ne veut pas dire que la grammaticalisation ne puisse les atteindre ; seulement, il faut les analyser au cas par cas et avec une vigilance particulière. S'agissant de l'emploi mémoriel, l'exclusion de la grammaticalisation des cas où TEN est doté d'une valeur subjective accrue vaut également pour eux. En outre, il importe, ici aussi, de réserver un traitement à part aux occurrences contenant des cumuls. Excepté ces deux cas, nous partons du présupposé que TEN en emploi mémoriel, se rapproche du défini français et peut être considéré comme encourageant la grammaticalisation.

Dans le domaine des emplois à la charnière entre les deux types de définitude, la situation est plus transparente : dès lors que TEN apparaît, il est susceptible d'une interprétation grammaticalisée. Les choses semblent juste un peu plus ambiguës dans le domaine de l'anaphore associative, où la présence de TEN avec certains noms peut être due à un effet pragmatique « émotionnel ». Pour savoir si tel est le cas, il convient de transposer l'énoncé en question dans un contexte émotionnellement neutre et d'observer si l'occurrence sera maintenue ou pas. Le jugement de nos dix informateurs s'est également révélé très utile ici. Pour ce qui est des « concepts individuels complexes », leur singularité réside dans ce que la présence de TEN grammaticalisé y est souvent encouragée par celle d'un effet « émotionnel » : voilà donc un type d'emploi où ces deux valeurs semblent inextricablement liées

(voir la citation de Mathesius *supra*). Elle semble aussi encouragée par l'absence de nom, entraînant une substantivation de l'adjectif (voir la Section 3.4).

Quant aux hypothèses, les voici. D'abord, en faisant écho au constat de Himmelmann (2001 : 832), nous nous sentons obligé de rappeler que le démonstratif occupe, dans les complexes nominaux de nos données tchèques, une position assez stable en ce sens qu'il précède toujours le nom (cf. Veselovská, 2014), bien qu'il puisse s'y combiner avec d'autres éléments fonctionnels (cf. Uhlířová, 1992). Ainsi, cette hypothèse a déjà été validée à travers les analyses précédentes. Les autres hypothèses sont celles-ci (pour chacune d'entre elles, nous indiquons la section où elle sera examinée) :

- 1) La présence de TEN sera ressentie comme obligatoire en emploi situationnel, hormis les cas où ce dernier est porteur d'un effet de subjectivité accrue ou ceux où il est accentué (Section 6.2.1.1).
- 2) Sa présence sera ressentie comme préférable dans ceux des contextes anaphoriques fidèles où il ne remplit ni sa fonction « établissante » ni sa fonction « réactivante », soit les contextes où il ne remplit pas les fonctions d'un démonstratif (Section 6.2.1.2).
- 3) Elle sera ressentie comme fortement préférable (sinon obligatoire) dans les contextes mémoriels où TEN ne se combine pas avec un autre élément fonctionnel et où il n'est pas porteur d'un effet de subjectivité accrue (Section 6.2.1.3).
- 4) La présence de TEN sera ressentie comme préférable à son absence avec les noms sortaux ou relationnels pourvus de compléments qui établissent leur unicité dénotative (Section 6.2.2.1).
- 5) Elle le sera également avec certaines anaphores associatives, du moins celles du type situationnel (Section 6.2.2.2).
- 6) Avec les « concepts individuels complexes », la présence de TEN sera susceptible d'être encouragée par celle d'un effet pragmatique « émotionnel » ainsi que par l'absence de nom, soit en contexte d'adjectif substantivé (Section 6.2.2.3).
- 7) Enfin, vu la place occupée par TEN dans les contextes de définitude « pragmatique » et les contextes se situant entre les définitudes « pragmatique » et « sémantique », le tchèque (parlé) informel de nos jours tend vers une « scission du type 1 » (Ortmann ; 2014 : 294)

Dans l'examen de toutes ces hypothèses, le recours au sentiment de la langue de nos informateurs tchèques (et, accessoirement, au nôtre propre) a joué un rôle incontournable. En effet, l'une des principales faiblesses de ce travail est l'impossibilité pour nous de comparer les nombres d'occurrences introduites par TEN avec l'ensemble des occurrences du (sous-)type concerné⁶⁶, y compris celles où TEN est absent, pour voir quelle est la proportion exacte de la variante avec TEN. Via le questionnement de locuteurs natifs, nous avons tenté de pallier ce défaut.

⁶⁶ Les « concepts individuels complexes » représentent une certaine exception à cela (voir la Section 6.2.2.3).

6.2 L'étude des données

Dans le présent volet, nous examinerons, sur la base de nos données annotées, les différents types et sous-types d'emploi de TEN quant à sa grammaticalisation en article défini. Le plan adopté reflètera globalement celui de la Section 3.3 du Chapitre 3, à la seule différence près que la structure principale reposera sur la distinction entre les deux types de définitude décrits dans la Section 6.1.3.4. Comme TEN n'apparaît pas avec une fonction d'article en contextes de définitude « sémantique » pure (voir *supra*), ces contextes ne seront pas traités ici en détail (ils le seront, en revanche, dans l'ensemble du Chapitre 7).

6.2.1 Les contextes de définitude « pragmatique » pure

Seront, dans un premier temps, étudiés les emplois appartenant à la définitude « pragmatique » pure. Ces emplois sont : l'emploi déictique situationnel, l'emploi anaphorique et l'emploi mémoriel. Au total, notre échantillon de conversations contient 334 occurrences de TEN ressortissant à l'ensemble de ces trois types d'emploi et considérés par nous comme grammaticalisés (voir le Tableau 5.19 de la Section 5.2.3.2), soit 74,06 % de toutes les occurrences évaluées comme grammaticalisées (voir les critères d'évaluation de la grammaticalisation énumérés à la fin de la Section 6.1.4 *supra*). Dans le sous-corpus des pièces de Havel, ce chiffre s'élève à 53 (voir le Tableau 5.34 de la Section 5.2.4.2), soit 80,30 % de toutes les occurrences que nous avons évaluées comme grammaticalisées.

6.2.1.1 L'emploi exophorique situationnel

Dans notre échantillon d'ORTOFON v1, nous avons analysé 50 occurrences de TEN comme appartenant à ce type d'emploi (nous laissons ici de côté la seule occurrence que nous avons analysée comme étant un emploi déictique textuel et dont le statut quant à la grammaticalisation nous semble difficile à évaluer). De ces 50 occurrences, 44 seraient, d'après nous, engagées dans une grammaticalisation vers l'article défini. Nous commencerons par reproduire ici l'exemple du Chapitre 3 :

14) L1 : no jo já to m-á-m proti východ-u
MOD moi.NOM cela.ACC avoir-PRS-1SG contre Est-DAT.SG

jenže když je tam na *tom* *stolečk-u* u
mais quand être.PRS.3G là-bas sur ten.M.LOC.SG petite table-LOC.SG auprès de

toho *radiátor-u* ..
ten.M.GEN.SG radiateur- GEN.SG

L2 : tak oni ty orchidej-e ne-m-a-jí takov-ý
 MOD eux.NOM ten.F.NOM.PL orchidée-NOM.PL NEG-avoir-PRS-3PL tel-N.ACC.SG

ne ne-m-a-jí rád-y takov-ej .. žár .. víš ?
 non NEG-avoir-PRS-3PLcontent-F.NOM.PL tel-M.ACC.SG chaleur.ACC.SG MOD

L1 : bon, je l'ai placée contre l'Est, mais quand elle est là, sur la petite table près du radiateur...

L2 : bon, les orchidées n'aiment pas... je veux dire elles n'aiment pas autant de chaleur, tu vois ?

ORTOFON v1

Tout porte à croire que la pièce où se trouvent le locuteur et l'interlocuteur ne contient qu'une seule petite table et qu'un seul radiateur. Par conséquent, TEN ne remplit pas son rôle de démonstratif, consistant à choisir un référent parmi d'autres référents du même type moyennant son indexation forte sur le contexte. Les référents sont identifiés à partir du cadre situationnel comme étant chacun l'unique représentant de sa catégorie. La forme est non-accentuée et, si la présence d'un geste ostensif n'est pas exclue, elle n'est pas nécessaire. La situation est analogue en 15, où les personnes roulent ensemble en voiture (l'enregistrement laisse entendre le bruit du moteur) :

15) ty .. mysl-í-š že to je rozumn-é? .
 toi.NOM penser-PRS-2SG que cela.NOM être.PRS.3SG raisonnable-N.NOM.SG

před **tou** **zatáčk-ou** jo?
 devant ten.F.INS.SG virage-INS.SG MOD

tu crois que c'est raisonnable ? [de faire ça] devant le virage ?

Il est possible de déduire de l'ensemble de la situation que le copilote émet un doute sur la nature d'un acte que le conducteur réalise juste avant que la voiture ne s'engage dans un virage, en l'occurrence le seul virage (jugé pertinent) qui soit présent à ce moment-là.

Il peut arriver que le référent ne soit pas également perceptible, dans la situation, par tous les participants au dialogue :

16) L1 : já ti tu vynd-á-m **to** **céděčk-o**
 moi.NOM toi.DAT là sortir-FUT-1SG ten.N.ACC.SG CD-ACC.SG

L2 : já jsem tam mě-l-a céděčk-o?
 moi.NOM AUX.PST.1SG là-bas avoir-PTCP-SG.F CD-ACC.SG

L1 : je vais te sortir le CD

L2 : il y avait un CD à moi ?

En 16, le locuteur semble partir du présupposé que son interlocutrice est au courant de la présence d'un CD au sein de la situation, alors que sa réaction prouve que c'est faux (y contribue sans doute le fait que le CD en question se trouve à l'intérieur de l'ordinateur, et n'est donc pas directement

perceptible par elle : contrairement au locuteur, elle n'est pas directement en contact avec l'ordinateur)⁶⁷.

Nous avons soumis à l'évaluation de nos informateurs l'exemple suivant, où le contexte général est plus facile à expliquer que dans les trois exemples précédents. Les personnes avaient pour tâche de choisir entre les énoncés a), b) et c) :

17) La personne à gauche s'adresse à la personne à droite pour lui demander de lui passer le ballon :



a)	<i>D-ej</i> donner-IMP.2SG <i>Donne-moi la balle !</i>	<i>mi</i> moi.DAT	míč! ballon.ACC.SG
b)	<i>D-ej</i> donner-IMP.2SG <i>Donne-moi cette balle !</i>	<i>mi</i> moi.DAT	tenhle tenhle.M.ACC.SG
c)	<i>D-ej</i> donner-IMP.2SG <i>Donne-moi la balle !</i>	<i>mi</i> moi.DAT	ten ten.M.ACC.SG
			míč! ballon.ACC.SG

Tous les informateurs sans exception ont choisi l'énoncé c), percevant ainsi la présence de TEN comme nécessaire, apportant ainsi une confirmation à l'hypothèse n°1, formulée *supra*⁶⁸. TENHLE impliquerait que plusieurs ballons sont présents dans la situation.

Il est pourtant un cas de figure essentiel où nous n'avons pas considéré TEN comme étant en voie de grammaticalisation. Nous en avons déjà relevé l'existence dans la Section 3.5.2 ainsi que parmi les critères énumérés à la fin de la Section 6.1.4 *supra* :

18) Co	to prosím vás	děl-á-te?	Okamžitě	nech-te
que.ACC	MOD	faire-PRS-2PL	immédiatement	laisser-IMP.2PL

⁶⁷ Il ne serait pas non plus faux, à notre avis, de voir dans cet emploi un coup de force présuppositionnel éventuel, et ce lorsque le locuteur serait bien conscient du fait que l'interlocutrice n'est pas en mesure d'identifier le référent. Dans ce cas, l'emploi serait proche de l'exemple suivant, cité par Epstein (2002 : 354) : « M. "Did you hear about *the fight*?" A. "What fight?" M. "Between Bob and Grandpa ..." ». Comme le relève Ziková (2017 : 89), de tels emplois ne permettent pas à l'interlocuteur d'identifier un référent grâce à un cadre établi ; c'est le défini lui-même qui met en place un tel cadre et, en ce faisant, il permet d'introduire un nouveau thème dans le discours. Il nous semble par ailleurs que le coup de force est souvent moindre lorsque le complexe nominal défini contient un concept fonctionnel. Ainsi, une question comme « Avez-vous la carte du magasin ? », qu'une caissière pose à un client au moment de son passage à la caisse serait un coup de force moins « violent », car la présence du complément possessif suffit à établir l'existence du référent dans l'univers du discours de l'interlocuteur. Ce dernier sera moins enclin à poser la question « Quelle carte ? ».

⁶⁸ Une comparaison avec le slovaque et le polonais est ici intéressante. Alors qu'en slovaque, la situation est la même qu'en tchèque (TEN est obligatoire), en polonais, la situation neutre est celle où TEN est absent, car sa présence semble automatiquement entraîner un effet de subjectivité accrue – le locuteur exprime sa colère, son irritation, son impatience, etc. (cf. les exemples 18 et 19).

být *to* *aut-o* !
être.INF ten.N.ACC.SG voiture-ACC.SG

*Qu'est-ce que vous faites ? Laissez **cette voiture** tranquille !* (Adamec, 1983 : 159)

Nous avons écrit au sujet de ce type d'emploi que bien que le référent soit le seul représentant de sa classe dans la situation, le démonstratif semble être porteur d'une valeur affective accrue (font pencher vers cette lecture la nature du contexte, où les personnes interpellées sont en train de vandaliser la voiture du locuteur, ainsi que la modalité phrastique exclamative). De ce fait, TEN ne serait pas simplement équivalent au défini, dans la mesure où il resterait doté d'un certain sémantisme déictique, sémantisme qui serait à l'origine de la valeur affective produite. La preuve est que même en français et en anglais⁶⁹, le démonstratif sera préféré au défini. L'exemple suivant, emprunté à nos données, est analogue (la locutrice reproduit la réaction de l'entourage d'une femme qui, après avoir agressé des policiers dans un espace public, vient d'être menottée par eux) :

19) ted' všichn-i lid-i na ni kouka-l-i
maintenant tout.M.NOM.PL gens-NOM sur elle.ACC.SG regarder-PST-3PL.M

říka-l-i páni *ta* *žensk-á*
dire-PST-3PL.M MOD ten.F.NOM.SG femme-NOM.SG

*et là tout le monde la regardait en disant mais alors **cette femme*** ORTOFON v1

Un autre fait singulier concernant ces exemples est que TEN y est à la fois difficilement suppressible et impossible à remplacer par une forme forte sans que le sens de l'énoncé en soit profondément altéré. Comme l'écrit Berger (1993 : 187), si son absence n'est pas habituelle (et, pour nous, elle est même clairement agrammaticale en 18 et 19), son remplacement par une forme forte comme TENHLE débouchera sur un contraste (*cette voiture-ci et pas une autre*) qui se marie mal avec l'expressivité recherchée. Nous osons même conclure que si TEN subsiste encore dans ce type d'emploi avec un sens indexical (bien que sans doute affaibli), il en est ainsi justement parce qu'aucun contraste entre plusieurs référents du même type n'est présent (voir *supra*). Notre échantillon de conversations contient 2 occurrences de ce type, et notre sous-corpus de V. Havel en contient 4.

Rappelons que notre échantillon de conversations contient une occurrence où TEN est accentué en emploi situationnel. Cet emploi n'est évidemment pas touché par la grammaticalisation (il peut être remplacé par une forme forte comme TENHLE ; voir les critères cités à la fin de la Section 6.1.4)⁷⁰ :

⁶⁹ Voir les exemples de Fillmore (1982 : 42-44) dits « acknowledging », présentés dans la Section 3.5.2.

⁷⁰ Cependant, contrairement à ce que laisse supposer l'étude de Jodas (2010), la locutrice n'est pas originaire de Moravie de l'Est ou de Silésie, mais de Lubnice, un village situé entre la région de Moravie du Sud et celle de Vysočina. Cet emploi de TEN semble d'autant plus étrange chez elle qu'elle utilise couramment, dans le reste de la conversation, des traits

- 20) L1 : tady kolik holek jd-e jako pryč?
 ici combien fille.GEN.PL aller-PRS.3SG MOD ailleurs
 ...
 L2 : vod nás . tady z tý stran-y
 de nous.GEN ici de ten.F.GEN.SG côté-GEN.SG
 ...
 L1 : z tý st* z **té** **stran-y** . jasně
 de ten.F.GEN.SG de ten.F.GEN.SG côté-GEN.SG bien sûr

L1 : ici, combien de filles s'en vont, genre ?

L2 : de chez nous, ici, de ce côté

*L1 : de ce... de **ce côté**, bien sûr*

ORTOFON v1

Et, enfin, il est un troisième cas qui n'a pas été considéré comme relevant de la grammaticalisation, et que nous n'avons pas relevé dans le cadre des critères de grammaticalisation à la fin de la Section 6.1.4 :

- 21) jo .. takže si muž-e-š vybrat **tady z těch** . **šilenc-ů**
 ouais du coup REFL pouvoir-PRS-2SG choisir.INF ici de ten.M.GEN.PL fou-GEN.PL
ouais, du coup, tu peux choisir parmi ces fous-là ORTOFON v1

En 21, il est question d'une liste électorale établie à l'occasion des élections régionales. La construction est particulière en ce sens que l'adverbe déictique *tady* (*ici*) et TEN se rapprochent, par leur contiguïté phono-syntaxique, de la forme amalgamée du démonstratif nominal TADYTEN, bien qu'ils soient séparés par la préposition *z* (*de*). Dans la Section 3.1.1, nous avons brièvement relevé le problème relatif à la graphie (*tadyten* vs *tady ten*) tout en considérant que cette graphie ne reflète pas vraiment – en tout cas, pas dans ORTOFON v1 – la différence entre les deux analyses sémantiques possibles. Ce qui nous fait dire que la même ambiguïté se présente lorsqu'une préposition s'intercale entre les deux éléments est le fait que dans certaines occurrences du type de 21, ces éléments ne sont pas séparés par une pause. Dans ce dernier cas, la situation semble ainsi analogue à celle où la préposition se retrouverait devant le premier des éléments, et où *tady ten* ne peuvent pas faire l'objet d'une analyse dissociée :

- 22) jo .. takže si můžeš vybrat **z tady těch** . **šilenců**

Trois occurrences de ce type ont été détectées dans l'échantillon de conversations et une seule a été détectée dans le sous-corpus de V. Havel.

En plus des scénarii qui viennent d'être mentionnés, il pourrait en exister, selon nous, un quatrième, non représenté dans notre corpus, et où non seulement TEN devrait être considéré comme

du « tchèque commun » (y compris les formes du paradigme TENHLE), comme c'est de plus en plus le cas, de nos jours, pour les locuteurs de cette partie du territoire de la République tchèque (voir également Jodas, 2010) et la Section 3.1.1.

encourant une grammaticalisation, mais où cette grammaticalisation serait même plus avancée, car l'emploi se situerait entre les définitudes « pragmatique » et « sémantique ». Il s'agit de l'emploi « **déictique situationnel associatif** »⁷¹, équivalent de l'anaphore associative dans le domaine situationnel :

23) Une personne entre dans un supermarché et pose à un membre du personnel la question suivante :

Où se trouve **le rayon des surgelés**, s'il vous plaît ?

Ici, malgré l'impossibilité, pour le locuteur, de percevoir et d'identifier un référent concret (le locuteur ne voit pas le rayon, sinon sa question ne serait pas pertinente), le SN est défini grâce au seul lien conventionnel qui existe entre les supermarchés et leurs différents rayons (*dans tout supermarché, il y a d'habitude un rayon des surgelés*), autrement dit, entre un cadre et un de ses éléments. Il est cependant curieux qu'à l'heure actuelle, le tchèque parlé informel semble ne pas (encore) employer TEN dans ce contexte : en 23, le nom y sera employé nu⁷².

S'agissant des pièces de V. Havel, nous considérons que 7 des 11 occurrences de TEN en emploi situationnel reflètent sa grammaticalisation en cours (voir la Section 5.2.4.2). C'est le cas, d'après nous, de l'occurrence suivante, où la madone est le seul objet de son espèce présent au sein du cadre situationnel :

24) MICHAL Ale stejně – je to verk! BEDŘICH Co? MICHAL: **Ta madona** – BEDŘICH Hm – MICHAL Uvědomuješ si, jaké dramatické napětí vzniká mezi ní a tím kinžálem?

MICHAEL : Tu peux dire ce que tu veux, mais c'est un chef d'œuvre ! FERDINAND : Quoi ? MICHAEL : Cette vierge gothique. FERDINAND : Mmh. MICHAEL : As-tu remarqué le champ de tension dramatique qui se forme entre la statue et le sabre turc ?

V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

Il peut donc sembler étonnant que le traducteur français ait opté pour le démonstratif. Nous estimons que l'explication peut être attribuée à la volonté du traducteur de faire marquer à Michal son attitude émotionnelle vis-à-vis de l'objet, son attachement à celui-ci (voir *supra*), mais aussi à la volonté de lui faire mettre en place un contraste implicite entre la madone concernée et d'autres madones (seule la forme CE peut marquer ce contraste, contrairement à TEN ; voir *supra*). Cette dernière explication vaudrait également pour l'emploi de TENHLE (plutôt que TEN) et de CE (plutôt que LE) dans l'exemple suivant :

⁷¹ Aucun travail de ceux que nous avons consultés ne relève ce cas de figure, d'où la présence des guillemets.

⁷² Sauf en emploi « émotionnel ».

- 25) MICHAL Viš, z čeho mám ovšem největší radost? BEDŘICH Z čeho? MICHAL Z *téhle gotické madony*! Potřeboval jsem totiž, aby šla přesně do *toho výklenku*, a všude měli jako naschvál buď větší, nebo menší – BEDŘICH A to nešlo *ten výklenek* zvětšit?

MICHAEL. *Mais, sais-tu ce qui était ma plus grande satisfaction ?* FERDINAND. *Quoi ?* MICHAEL. *Cette vierge gothique, tu vois, il m'en fallait absolument une à la dimension de cette niche. Et je n'en voyais que de trop grandes ou de trop petites, c'était comme un mauvais sort !* FERDINAND. *Ne pouvait-on agrandir la niche ?*

V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

En 25, le recours à *TENHLE* pour désigner la madone établirait un contraste (implicite) entre la madone en question et d'autres madones qui, bien qu'absentes de la situation, n'en existeraient pas moins dans l'esprit des participants à la conversation (soit toutes les autres madones auxquelles ces derniers ont eu affaire jusque-là). Ce contraste serait maintenu en français par l'usage de *CE*. Le traducteur continuerait ensuite à produire le même effet en employant *CE* avec la première occurrence du nom *niche*, même si l'auteur a opté, dans ce cas, pour *TEN* (la seconde mention du référent *niche* est cependant introduite, en français, par le défini).

Par ailleurs, ce que notamment les exemples 18 et 19 *supra* nous permettraient de voir, c'est que contrairement aux formes fortes tchèques, qui semblent toujours mettre en place un contraste (fût-il implicite) entre le référent visé et d'autres référents de la même catégorie, le démonstratif français *CE* peut – comme *TEN* – aussi s'employer dans une situation où un seul représentant de la catégorie référentielle est présent, sans pour autant que cela déclenche automatiquement le même effet de contraste⁷³.

Au total, des 7 occurrences déictiques situationnelles grammaticalisées, 4 ont été rendues par le défini (cf. la seconde occurrence de *výklenek* en 25) et 3 ont été rendues par le démonstratif (cf. en 24 et la première occurrence de *výklenek* en 25)⁷⁴. S'agissant des 4 occurrences non-grammaticalisées selon nous, elles affichent le même profil que les exemples 18 et 19 cités *supra*, ce qui veut dire que *TEN* y est porteur d'une valeur affective accrue :

- 26) Čert mi byl *toho chlapa* dlužen!
J'avais bien besoin de ce pauvre petit con !

V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

En 26, bien que seul représentant de sa classe, le référent apparaît dans un contexte très marqué émotionnellement : en témoignent la modalité phrastique exclamative tout comme le nom familier *chlap* (*gars, mec*). Dans la traduction française, cette charge modale est encore amplifiée par la

⁷³ En revanche, cet effet apparaîtra inévitablement lorsque *CE* sera complété par *-ci* ou *-là*.

⁷⁴ Pour ce qui est des occurrences déictiques situationnelles introduites par une forme forte du démonstratif (16 cas), c'est le démonstratif qui a été utilisé la plupart du temps en français (soit dans 11 cas) ; le possessif a été employé 2 fois, le défini, une reformulation et l'emploi d'un autre moyen grammatical une fois chacun (voir la Section 5.2.4.4).

présence des adjectifs évaluatifs *pauvre* et *petit* et par le « nom de qualité » (Milner, 1976) *con*. Compte-tenu de cette modalité, il n'est nullement surprenant que dans 3 des 4 cas, TEN soit traduit par CE (une fois, le traducteur opte pour une transposition verbale, donc une reformulation⁷⁵).

6.2.1.2 L'emploi anaphorique

Comme nous l'avons déjà indiqué dans la Section 3.3.2, l'évaluation de la possible grammaticalisation de TEN en emploi anaphorique s'avère une tâche difficile. Cette difficulté est due à plusieurs facteurs. D'abord, contrairement à l'emploi situationnel – où la présence de TEN reflète une situation entièrement différente de la présence des formes fortes du démonstratif et où l'absence de TEN tend vers l'agrammaticalité (cf. les exemples 17, 18, 19 et 26) –, en contexte anaphorique, la sphère d'emploi de TEN recouvre partiellement, d'une part, celle des formes fortes TENTO/TENHLE, ainsi que, d'autre part, celle de la détermination zéro. Ensuite, l'usage de TEN en anaphore obéit à un ensemble de contraintes plus ou moins rigides ou au contraire dynamiques, contraintes qui relèvent de la structure informationnelle de l'énoncé ou qui sont d'ordre pragmatique (la possibilité du locuteur de diriger l'attention de l'interlocuteur sur un élément de l'énoncé), mais qui tiennent également compte du degré d'ancrage et d'activation du référent au sein du discours.

Le premier phénomène (le chevauchement de la valeur de TEN avec celle des formes fortes proximales et avec celle de l'anaphore nue) ayant été abordé dans la Section 3.3.2, nous n'y reviendrons plus dans cette section. En revanche, c'est le second et le troisième phénomène que nous aborderons ici. Nous commencerons par **les contraintes d'ordre thématique-rhématique** en reproduisant ici les constats de Štícha (1999), dont l'analyse est, à notre connaissance, la plus complète. Pour rappel, TEN est obligatoire lorsqu'il apparaît avec un rhème contextuellement ancré (la reprise d'un référent en position rhématique ; voir la Section 3.3.2) :

- 27) Když jste odeš-l-i s Old-ou do toho kin-a,
 quand AUX.PST.2PL partir-PTCP-PL.M avec NP-M.INS.SG à ten.N.GEN.SG cinéma-GEN.SG
- uděla-l-i jsme si s Luc-y **ta** **játr-a.**
 faire-PTCP-PL.M AUX.PST.1PL REFL avec NP-F.INS.SG ten.N.ACC.PL foie-ACC.PL

Quand vous êtes partis au cinéma avec Olda, Lucy et moi, on s'est fait le foie à manger.
 (V. Havel, emprunté à Štícha, 1999 : 128)

Dans notre échantillon de conversations, nous enregistrons la présence de 31 occurrences appartenant, respectivement, à l'anaphore fidèle (24), infidèle (4) et indirecte (3) introduites par TEN

⁷⁵ *Nech toho bulení, viš, že to nesnáším!* (littéralement : Arrête **ce chialement** !)
Arrête de chialer, s'il te plaît ! Tu sais que j'ai horreur de ça !

V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

et occupant la position postverbale (voir la Section 5.2.3)⁷⁶. Dans notre sous-corpus de Havel, le nombre de ce type d'occurrences s'élève à 8. Au vu de ce qui vient d'être dit, nous constatons que l'emploi de TEN est obligatoire dans cette position, quel que soit par ailleurs son statut vis-à-vis de la grammaticalisation :

28) já	jsem	si	koupi-l	[ledov-ý	kaf-e] _i
moi.NOM	AUX.PST.1SG	REFL	acheter-PTCP.SG.M	glacé-N.ACC.SG	café-ACC.SG
...					
já	jsem	pi-l	[to	ledov-ý	kaf-e] _i
moi.NOM	AUX.PST.1SG	boire-PTCP.SG.M	ten.N.ACC.SG	glacé-N.ACC.SG	café-ACC.SG
<i>j'ai acheté [un café glacé]_i</i>					
...					
<i>moi, j'étais en train de boire [mon café glacé]_i</i>					

ORTOFON v1

En effet, l'absence de TEN dans la reprise du référent *café glacé* mettrait en péril le lien de coréférence, le complexe nominal *ledový kafe* pouvant être facilement interprété comme indéfini.

Ensuite, l'emploi de TEN semble obligatoire chaque fois que son absence avec le nom de la reprise risquerait d'entraîner une interprétation générique, et donc non-coréférente avec l'antécédent :

29) Byl	tam	[klid] _i	a [ten	klid] _i
être.PST.3SG.M	là-bas	calme.NOM.SG	et ten.M.NOM.SG	calme.NOM.SG
ji	utěšova-l.			
elle.ACC.SG	consoler-PST.3SG.M			
<i>Il y avait [du calme]_i et [ce calme]_i la consolait.</i>				

(M. Kundera, emprunté à Štícha, 1999 : 128)

En absence de TEN, le nom *klid* (*calme*) pourrait, dans sa seconde occurrence, être interprété comme générique. Remarquons qu'en français, la même ambiguïté se présente et doit être levée par le démonstratif. Dans notre échantillon de conversations, il nous semble que cette ambiguïté risque de se produire, en cas d'omission de TEN, 8 fois, dont une fois avec l'anaphore fidèle, 3 fois avec l'anaphore infidèle et 4 fois avec l'anaphore indirecte. Regardons l'exemple 30, qui relève de l'anaphore fidèle, et où le complexe nominal démonstratif figure en dislocation à droite :

30) voni	lid-i	se	stěh-uj-ou	a tak	ty
eux.NOM	gens-NOM	REFL	déménager-PRS-3PL	et alors	ten.M.ACC.PL

⁷⁶ Nous nous intéressons ici aux seuls emplois endophoriques de TEN adnominal relevant de la définitude « pragmatique » pure et au sein desquels la grammaticalisation a certaines chances de se produire. En nous basant sur la réalité de nos données, nous nous sentons ainsi contraint d'exclure l'emploi anaphorique temporel (deux occurrences seulement dans l'ensemble de nos données, toutes deux non-grammaticalisées) et l'emploi anaphorique résomptif (une seule occurrence grammaticalisée sur une totalité de 24 dans l'ensemble de nos données ; voir les Sections 5.2.3.2 et 5.2.4.2). Ensuite, il faut écarter tous les emplois endophoriques se situant à la charnière entre les définitudes « pragmatique » et « sémantique », car ceux-ci feront l'objet d'une analyse à part dans la Section 6.2.2 *infra*.

star-ý vieux-M.ACC.PL	cajk-y bric-à-brac-ACC.PL	nos-ej porter-PRS.3PL	do popelnic-e à poubelle-GEN.SG
a [voni] _i et eux.NOM REFL	si to zase cela.ACC MOD	ber-ou prendre-PRS.3PL	[ty cik* cikán-ī] _i ten.M.NOM.PL gitan-NOM.PL

les gens, ils déménagent, et, du coup, ils jettent le vieux bric-à-brac à la poubelle ; et après, [ils]_i le récupèrent, [ces gitans-là]_i
ORTOFON v1

Ici, la reprise cible un groupe particulier de gitans que la locutrice a mentionné auparavant. L'absence du démonstratif pourrait présenter un certain risque, pour le complexe nominal, de faire l'objet d'une interprétation générique (*les gitans en général*). Force est pourtant de reconnaître que contrairement à l'exemple précédent, cette interprétation serait ici davantage écartée par le contexte. Dans notre sous-corpus de Havel, ce type d'occurrence est attesté dans 10 cas, dont 4 relevant de l'anaphore fidèle et 6 de l'anaphore infidèle⁷⁷.

Le dernier contexte où TEN est exigé est celui où il apparaît dans la reprise anaphorique d'un référent présenté comme thématique mais rejeté en fin d'énoncé, comme en 31, où *ty šaty* (*cette robe*) est la première reprise anaphorique d'un référent qui vient d'être introduit dans le discours (il est question d'une fille qui vient d'acquérir une nouvelle robe) :

- 31) Mě-l-a krásn-ou **FI**gur-u v *těch* *šat-ech*.
avoir-PST-3SG.F belle-F.ACC.SG taille-ACC.SG dans ten.M.LOC.PL robe-LOC.PL tantum
Elle avait une belle taille, dans cette/la robe. (J. Škvorecký, emprunté à Štícha, 1999 : 129)

Le complément prépositionnel *v těch šatech* (*dans la robe*) est un circonstant extra-prédicatif (qu'il serait également possible de placer en tête de phrase), l'accent principal étant ainsi placé sur le mot *figuru* (*taille*) et une pause étant marquée avant le circonstant (d'où l'insertion de la virgule en français). Sur le plan de la présentation de l'information, ce cas n'est pas loin de la dislocation à droite, dans la mesure où, rappelons-le, le circonstant est porteur de l'information thématique. Mais la situation est tout de même différente : si dans le cadre de la dislocation à droite, le constituant extrait apparaît aussi comme un thème situé en fin d'énoncé, il est représenté au sein de la structure matrice par un pronom (ou un adverbe) qui l'anticipe⁷⁸. Notons que si TEN était absent en 31, *v šatech* serait susceptible d'une lecture générique (de même que *en robe* en français : *de manière générale, elle avait une belle taille en robe / chaque fois qu'elle portait une robe*).

Dans notre échantillon de conversations, ce scénario nous semble se produire deux fois :

⁷⁷ En ce qui concerne les anaphores infidèle et indirecte, cela ne nous surprend guère, car dans nos données, les noms apparaissant dans ces emplois sont, de manière générale, davantage susceptibles d'une lecture générique, et ce notamment pour deux raisons : 1) le fait de posséder un sens tendant vers une généralité par défaut (*člověk* (*humain*)) et 2) le fait d'apparaître au pluriel (*holky/děvčata* (*filles*)). Comme dans le cas précédent, seules sont ici considérées les occurrences anaphoriques fidèles, infidèles et indirectes de TEN.

⁷⁸ Fût-ce en anaphore zéro, comme cela peut se produire en tchèque, lorsque le constituant occupe la fonction sujet (voir la Section 2.2.3).

32) právě [Runtour]_i začína-l v Plzn-i jestli se ne-mýl-i-m
justement Runtour commencer-PST.3SG.M à Pilsen.LOC.SG si REFL NEG-tromper-PRS-1SG
...
a co je možn-é zhlédnout vlastně při [tom běh-u]_i
et que.ACC être.PRS.3SG possible-N.NOM.SG voir.INF MOD pendant ten.M.LOC.SG course-LOC.SG
c'est justement [Roundtour]_i qui commençait à Pilsen, si je ne me trompe pas
...
et qu'est-ce qu'on peut voir; justement, pendant [la course]_i ? ORTOFON v1

En 32, en absence de TEN, le substantif *běh* (*course*) présenterait le risque d'une lecture générique et non-coréférente avec la première mention (*Runtour*), renvoyant, elle, à une course concrète.

Pour ce qui concerne **les contraintes pragmatiques**, c'est elles qui semblent les plus difficiles à résumer. De manière générale, il s'agit de la possibilité, pour le locuteur, de recourir à l'usage de TEN dans le but de placer un référent au premier plan. Ce faisant, le locuteur signale que le référent est davantage digne d'attention que les autres référents, car il joue un rôle plus important dans le discours. Cette fonction a notamment été étudiée par Epstein (1993, 1994, 1995) en rapport avec les premiers stades de l'évolution de l'article défini en ancien français. Epstein assimile cette fonction à ce qu'il appelle la « fonction expressive » (cf. 1995) de l'article défini, qu'il tient, par ailleurs, à séparer de sa fonction référentielle :

Nous avons montré que l'article défini en AF possède une fonction expressive, à savoir qu'il peut indiquer la prééminence d'un référent. Il peut mettre en valeur un référent dont la prééminence est contextuelle (importance syntagmatique) ou un référent dont la prééminence découle de son statut à l'intérieur d'une catégorie (importance paradigmaticque). (1995 : 66)

Pour l'auteur, la fonction expressive est avant tout basée sur la subjectivité de la personne du locuteur⁷⁹, ce qui expliquerait, selon lui, pourquoi le défini apparaît en ancien français même là où son apparition semble enfreindre les règles sémantico-référentielles qui régissent son emploi (en contextes génériques, non référentiels, lorsque le référent ne peut être identifié par l'interlocuteur, etc.). Nous n'allons pas ici entrer dans les détails de la théorie d'Epstein, retenant uniquement ce qui nous semble pertinent pour l'étude de TEN anaphorique en tchèque, à savoir la notion d'emphase. Un rapprochement entre cette notion, telle qu'elle a été développée par Epstein pour l'ancien français, et certains emplois de TEN en tchèque a déjà été fait par Kresin (2008), Trovesi (2002) et Zíková (2017). La première, utilisant également des corpus de tchèque parlé informel, met notamment en avant les contextes où TEN introduit un nom « sémantiquement » unique, éventuellement générique :

⁷⁹ Nous retrouvons ici la notion de subjectivité, abordée dans la Section 6.1.3.2 *supra*.

33) to je ce être.PRS.3SG	problém problème.NOM.SG	tý ten.F.GEN.SG	nevyrovnan-ý inégal-F.GEN.SG	společnost-i, société-GEN.SG
<i>ta</i> ten.F.NOM.SG	<i>Prah-a</i> Prague-NOM.SG	proud-í se dynamiser-PRS.3SG	rychlejc, plus vite	
než se que REFL	rozvíj-í développer-PRS.3SG	<i>ten</i> ten.M.NOM.SG	<i>venkov...</i> campagne.NOM.SG	

c'est le problème de la société inégale, Prague se dynamise beaucoup plus vite que se développe la campagne...
(emprunté à Kresin, 2008 : 246 ; modifié)

En 33, le locuteur aurait décidé de marquer les référents de *Praha* (Prague) et *venkov* (la campagne) par TEN afin d'accentuer leur proéminence ainsi que le clivage qui les sépare dans les faits⁸⁰. S'inscrivant dans l'approche d'Epstein, Kresin (2008) et, dans une certaine mesure, Trovesi (2002) considèrent même ce type d'emploi (la « highlighting function » chez Kresin) comme crucial pour l'ensemble de la grammaticalisation de TEN en défini. Trovesi relève qu'en tchèque parlé informel, TEN apparaît autant avec des noms propres de personnes et de lieux qu'avec des noms à sémantique générique⁸¹. Il conclut son étude par le constat suivant :

[S]ur le plan sémantique et fonctionnel, ils [= les démonstratif adnominaux *ten* et *tón* en tchèque et en haut-sorabe] couvrent plus ou moins l'ensemble du spectre d'emploi de l'article défini, à commencer par les cas de référence anaphorique jusqu'aux cas des groupes nominaux dotés d'une référence générique ou renvoyant à un référent défini de manière inhérente [...]⁸²

En effet, ce parti-pris semble en phase avec l'importance de l'expressivité et de la subjectivité dans le processus de grammaticalisation, relevée par nombre d'auteurs (voir la Section 6.1.2). Il faut toutefois dire qu'il s'avère peu plausible – du moins dans la conjoncture actuelle – au vu de la valeur démonstrative « émotionnelle » très saillante que TEN possède dans ces contextes, où le nom concerné (propre ou générique) est loin de la sphère de la définitude « pragmatique »⁸³.

S'agissant finalement du **degré d'ancrage et d'activation du référent dans le discours**, il importe d'abord de rappeler ici quelques notions. L'activation du référent dépend directement de la distance qui sépare sa mention de la mention précédente la plus proche. Cette distance est dite

⁸⁰ “In this example, *ten* highlights the focal points of the contrast: Prague and the countryside. Word order also reinforces the focus on *Praha*: as in this example, it is usually mentioned either clause-initially, as the point of departure, or clause-finally, as the most communicatively dynamic part of the clause.” (Kresin, 2008 : 246)

⁸¹ Au sein de son corpus, contenant des productions tchèques, haut-sorabes et slovènes, ce dernier emploi apparaît uniquement dans les données tchèques.

⁸² Nous traduisons : „[V]ýznamově a funkčně pokrývají víceméně celé spektrum použití určitého členu – to jest od anaforického odkazování po případy nominálních skupin s generickou referencí anebo s inherentně již daným určitým referentem [...]“

⁸³ Vu le profil logico-sémantique des noms *Praha* et *venkov*, nous ne sommes plus ici, évidemment, dans le domaine de l'anaphore *stricto sensu*, car la référence de ces noms ne dépend pas du contexte. À l'instar de Šimík (2015), nous pouvons parler d'« anaphore discursive pragmatique ».

« distance anaphorique » (cf. Ariel, 1988 ; Givón, 1983 ; voir la Section 1.1.3). Quant à l’ancrage du référent, nous entendons par là que le référent est établi dans le discours. Un référent est établi du moment qu’il a été repris en deuxième mention par une expression définie anaphorique après avoir été introduit via une expression indéfinie⁸⁴. Il est une chose connue que l’une des fonctions typiquement assumées par le démonstratif en contexte anaphorique est justement cet ancrage d’un nouveau référent dont l’identification dépend du discours (cf. Christophersen, 1939 ; Corblin, 1983 ; Diessel, 1999 ; Himmelmann, 1996 ; Kleiber, 1986c). Il va sans dire que les deux notions (activation et ancrage) doivent être dissociées, car un référent activé n’est pas forcément ancré (établi) et vice versa.

D’autres concepts s’avèrent pertinents pour l’analyse des référents anaphoriques, telle la « persistance » (“decay” ; Givón, 1983), exprimant le nombre de fois où le référent sera repris ultérieurement (voir également la Section 1.1.3), et la « proéminence », qui est un concept complexe et dépendant d’une série de facteurs. Comme l’écrit Zíková (2017 : cf. 188-189), à qui nous empruntons globalement sa définition, la proéminence est d’abord déterminée par le statut ontologique du référent : sont plus proéminents les référents animés et notamment humains. Ensuite, elle est déterminée par la persistance et la fonction syntaxique dans laquelle le référent apparaît⁸⁵. Comme le relève également Zíková (2017 : 184), les référents proéminents sont souvent caractérisés par un niveau d’activation élevé⁸⁶. Outre son rôle consistant à établir un référent juste après son introduction dans le discours (voir *supra* ; nous proposons de parler de « démonstratif établissant »), le démonstratif est également utilisé de manière régulière lorsqu’il s’agit de réactiver un référent proéminent et déjà établi, mais disparu du discours depuis un certain temps (cf. Laury, 1997 ; Zíková, 2017 et 2018 ; nous parlerons de « démonstratif réactivant »)⁸⁷. Enfin, sera dit « nouveau », à l’instar de Laury (1997), tout référent en première mention, qu’il soit défini ou indéfini (que son identification dépende ou non du discours).

Les deux contextes qui viennent d’être présentés sont ainsi ceux où les démonstratifs semblent être le plus utilisés en emploi anaphorique dans les langues du monde, que ces langues disposent ou pas d’articles définis. Aussi TEN ne peut-il – a priori – être considéré comme un article en devenir lorsqu’il apparaît dans ces deux contextes (voir les critères de la fin de la Section 6.1.4). Nous disons

⁸⁴ Ainsi délimitée, la notion ne peut être appliquée qu’aux référents rendus définis par le discours. Cependant, rien n’empêche à notre avis de considérer par défaut comme établis tous les référents dont l’identification a sa source en dehors du discours (elle s’appuie sur un autre « cadre » ; Ariel, 1988).

⁸⁵ Les référents en fonction de sujet sont, de manière générale, plus proéminents que ceux en fonction de COD, de COI et de circonstant (cf. Brown, 1983).

⁸⁶ Elle maintient cependant que les deux concepts sont indépendants l’un de l’autre. Quant à nous, il nous semble logique de considérer que l’activation contribue directement à la proéminence, dans la mesure où cette dernière est directement influencée par le nombre de mentions à la fois passées et à venir, ainsi que par la distance qui sépare ces mentions.

⁸⁷ Laury (1997) réserve le terme de « référent accessible » à ce cas de figure.

a priori, car l'exception la plus patente, c'est lorsque plusieurs référents viennent d'être introduits dans le discours à la fois, référents entre lesquels un contraste existe. À ce moment-là, l'emploi de TEN correspond naturellement à celui du défini français (voir les Sections 1.2.2 et 6.1.3.3). Zíková cite plusieurs exemples de ce cas de figure, présents dans son corpus de textes narratifs :

34) ehm uvid-í pod kol-em zaparkovan-ýho [aut-a]_i [...] [ňák-ou
MOD voir-PRS.3SG sous roue-INS.SG garé-N.INS.SG voiture-GEN.SG quelque-F.ACC.SG

bankovku]_{ii}, [kter-ou]_{ii} chc-e sebrat, ale ne-jd-e to,
billet-ACC.SG qui-F.ACC.SG vouloir-PRS.3SG ramasser.INF mais NEG-aller-PRS.3SG

protože [to aut-o]_i na [ní]_{ii} stoj-í
parce que ten.N.NOM.SG voiture-NOM.SG sur elle.LOC.SG être garé-PRS.3SG

ben sous la roue d'[une voiture garée]_i, il voit [un billet]_{ii}, et il veut [le]_{ii} ramasser, mais il y arrive pas parce que [la voiture]_i est garée dessus (sur [lui]_{ii}) (Zíková, 2017 : 295)

Ici, la présence de TEN est plus proche de LE, car le référent voiture entre en « contraste externe » (cf. Corblin, 1987) avec le référent billet.

Un démonstratif accompagnant les référents qui sont à la fois établis dans le discours (donc à partir de leur troisième mention) et activés se rapprocherait, quant à lui, du défini (**l'hypothèse n°2 supra**). C'est encore Laury (1997 : 224) qui enregistre dans son étude portant sur l'émergence du défini en finnois une hausse majeure, en l'espace de moins de cent ans seulement, du marquage par le démonstratif SE des référents activés : tandis que dans les textes datant de l'époque 1880-1930, seuls 29 % des référents activés sont introduits par SE, ce chiffre atteint 88 % au sein des textes datant de l'époque 1960-1990. Il en résulte que le démonstratif finnois est entretemps devenu un marqueur bien plus conséquent des référents identifiables même lorsque ceux-ci apparaissent au-delà des deuxièmes mentions, et peuvent donc être considérés comme ancrés dans le discours. Zíková (2018 : 110) résume ainsi avec justesse à propos de TEN anaphorique :

Il existe ainsi deux possibilités testables : si *ten* se comporte (encore) comme un démonstratif typique, son taux d'apparition devrait être limité dans les mentions répétées, la référence aux référents en question étant, dans ce cas, garantie par les syntagmes lexicaux correspondants sans *ten*. En revanche, si le lexème *ten* n'est pas (n'est plus) un démonstratif typique, il figurera plus souvent dans les mentions répétées et tendra à s'y imposer comme moyen explicite de marquage de la définitude, et ce au détriment des syntagmes lexicaux correspondants sans *ten*⁸⁸.

⁸⁸ Nous traduisons : „Existují tedy dvě testovatelné možnosti: pokud se výraz *ten* chová (stále ještě) jako typické demonstrativum, v opakovaných zmínkách by měl být jeho výskyt omezený, odkazování k příslušným referentům v takovém případě zajišťují korespondující lexikální fráze bez *ten*. Naopak pokud výraz *ten* typickým demonstrativem (už) není, bude v opakovaných zmínkách vystupovat častěji a prosazovat se zde jako explicitní prostředek značení určitosti, a to na úkor korespondujících lexikálních frází bez *ten*.“

Aussi était-il primordial pour nous, lors de l'évaluation de la grammaticalisation de TEN dans notre propre échantillon, d'examiner le statut du référent qu'il introduisait, l'occurrence ayant une forte chance d'être évaluée comme grammaticalisée dès lors qu'il s'agissait d'une troisième mention et au-delà d'un référent activé⁸⁹. Nous reproduisons ici la séquence où il est question de papiers peints, figurant déjà dans l'Introduction générale :

- 35) L1 : ti soused-i jak jsem ti vo nich vyprávě-l-a
ten.M.NOM.PL voisin-NOM.PL REL AUX.PST.1SG toi.DAT de eux.LOC raconter-PTCP-SG.F
- jak tam bydle-l-i u mě . ty NP jak . voni
REL là-bas habiter-PST-3PL.M chez moi.ACC ten.M.NOM.PL NP REL eux.NOM
- mě-l-i [tapety]_i tam padesát let
avoir-PST-3PL.M papier peint-NOM.PL là-bas cinquante an.GEN.PL
- L2 : no tak to já je ne*
MOD moi.NOM eux.ACC NEG
- L1 : a ona ještě byl-a kuřák . a vona rád-a
et elle.NOM.SG en plus être.PST-3SG.F fumeur.NOM.SG et elle.NOM.SG content-F.SG
- děla-l-a tydlet-y panenky víš
fabriquer-PST-3SG.F tenhleten-F.ACC.PL poupée-ACC.PL MOD
- L2 : já v-í-m
moi.NOM savoir-PRS-1SG
- L1 : ona skoro byl-a už na smrt nemocná .
elle.NOM.SG presque être.PST-3SG.F déjà à mort.ACC.SG malade-F.NOM.SG
- L1 : já jsem tam přiš-l-a ona mi
moi.NOM AUX.PST.1SG là-bas arriver-PTCP-SG.F elle.NOM.SG moi.DAT
- ukazova-l-a panenky .
montrer-PST-3SG.F poupée-ACC.PL
- já řík-á-m no Dáš-o ale je to na úkor .
moi.NOM dire-PRS-1SG MOD NP-F.SG.VOC mais être.PRS.3SG cela.NOM à détriment.ACC.SG
- se urazi-l-a . no poněvač . a ted' tam přiše-l .
REFL vexer-PST-3SG.F MOD parce que et maintenant là-bas arriver-PST.3SG.M
- nov-ej majitel .
nouveau-M.NOM.SG propriétaire.NOM.SG
- a ty představ si [ty tapet-y]_i muse-l
et ten.F.ACC.PL imaginer.IMP.2SG REFL ten.F.ACC.PL papier peint-ACC.PL devoir-PST.3SG.M
- strhat . [to]_i byl-y ještě takov-ý ze západn-ího
arracher.INF ce être.PST-3PL.F encore tel-F.NOM.PL de occidental-N.GEN.SG

⁸⁹ Lorsque nous parlons de mentions répétées d'un référent, nous entendons par là, évidemment, toutes les mentions répétées et non seulement celles réalisées par TEN adnominal.

Německ-a rozumíš .
 Allemagne-GEN.SG MOD

proto mi tam děla-l-y sbíječ-k-y
 c'est pour ça que moi.DAT là-bas faire-PST-3PL.F marteau-piqueur-NOM.PL

[to vše-cko]_i šl-o dolů to byl-o příšern-ý
 cela.NOM tout-N.NOM.SG aller.PST-3SG.N en bas ce être.PST-3SG.N horrible-N.NOM.SG

tak dlouho [**ty** **tapet-y**]_i tam byl-y že jo
 si longtemps ten.F.NOM.PL papier peint-NOM.PL là-bas être.PST-3PL.F MOD

L1 : ces voisins que je t'ai parlé d'eux, ceux qui vivaient chez moi, ces NP, ceux qui avaient chez eux [des papiers peints]_i depuis cinquante ans

L2 : bon, moi, je les connais pas

L1 : et en plus de ça, elle fumait et elle aimait faire ces espèces de poupées, tu vois

L2 : je vois

L1 : elle était presque mortellement malade, j'y suis arrivée, elle m'a montré des poupées, je lui dis mais Dáša, ça, c'est au détriment de... elle s'est vexée parce que bon... et maintenant, il y a un nouveau propriétaire qui est arrivé, et imagine qu'il a été obligé d'arracher [ces papiers peints]_i, en plus [c']était des papiers peints d'Allemagne de l'Ouest, tu vois, c'est pour ça que j'avais des marteaux-piqueurs à côté de chez moi, il fallait bien enlever [tout ça]_i, c'était horrible, ça faisait si longtemps que [les papiers peints]_i étaient là, tu vois

ORTOFON v1

Rappelons notre analyse de la séquence. La première reprise du référent *tapety* (*papiers peints*) est réalisée par TEN adnominal, dont le rôle est d'établir le référent dans le discours (à ce titre, il n'est pas surprenant, d'ailleurs, que cette reprise occupe la position thématique de l'énoncé). Les deux reprises subséquentes emploient le démonstratif pronominal. La quatrième reprise, de nouveau réalisée par un complexe nominal avec TEN et où le référent est de nouveau thématisé, intervient peu de temps après, le référent étant donc activé et discursivement établi. Cet emploi de TEN peut, à notre avis, être considéré comme proche d'un article défini⁹⁰. Voici un autre exemple, où la première locutrice parle de son achat prévu de chaussures de danse :

36) L1: když slyše-l že moj-e příšt-í výplat-a
 quand entendre-PST.3SG.M que mien-F.NOM.SG prochain-F.NOM.SG salaire-NOM.SG

pů-jd-e na [tanečn-í bot-y kter-é stoj-í
 FUT-aller-3SG à de danse-F.ACC.PL chaussure-ACC.PL qui-F.NOM.PL coûter-PRS.3PL

tř-i tisíc-e]_i tak se zača-l strašně chechtat
 trois-ACC mille-ACC.PL alors REFL commencer-PST.3SG.M terriblement rigoler-INF

...

L2: a tak co bot-y to chc-e jako já jsem to co
 MOD chaussure-ACC.PL cela.NOM vouloir-PRS.3SG MOD moi.NOM AUX.PST.1SG MOD

⁹⁰ Neuf parmi nos dix informateurs ont choisi la version où le nom est introduit, dans cette occurrence, par TEN. Un seul a choisi la version avec le nom nu.

- jsem si kupova-l-a třeba do tanečn-ích
AUX.PST.1SG REFL acheter-PTCP-SG.F par exemple pour cours de danse-GEN.PL
- ...
- L2: tak buď to třeba hrozně klouza-l-o nebo to mě-l-o
alors soit cela.NOM par exemple terriblement glisser-PST-3SG.N ou cela.NOM avoir-PST-3SG.N
- podrážk-u takov-ou že se ani ne-votoč-í-š
semelle-ACC.SG tel-F.ACC.SG que REFL même NEG-tourner-PRS-2SG
- a to potom se ne-d-á
et cela.NOM après REFL NEG-donner-PRS.3SG
- L1: hmm no a na tohle na to tang-o tak to
MOD et pour tenhle.N.ACC.SG pour ten.N.ACC.SG tango-ACC.SG alors cela.NOM
- chc-e vyloženě ty tanečn-í bot-y
vouloir-PRS.3SG vraiment ten.F.ACC.PL de danse-F.ACC.PL chaussure-ACC.PL
- L2: hmm to mus-í bejt hrozně super mít
MOD cela.NOM devoir-PRS.3SG être-INF terriblement super.N.NOM.SG avoir-INF
- tanečn-í bot-y
de danse-F.ACC.PL chaussure-ACC.PL
- ...
- L1: jo mus-í a já se na [ně]_a neskonale těš-í-m
ouais devoir-PRS.3SG et moi.NOM REFL à elle-ACC.PL infiniment se réjouir-PRS-1SG
- a poříd-í-m si [je]_i až po Vánoc-ích
et se procurer-FUT-1SG REFL elle.ACC.PL seulement après Noël-LOC.PLtantum
- ...
- L2: uděl-á-š si takov-ýho poježíšk-a
faire-FUT-2SG REFL tel-M.ACC.SG après père Noël-ACC.SG
- L1: jo protože k Ježíšk-ovi dostan-u prach-y
ouais parce que à Noël-DAT.SG recevoir-FUT.1SG pognon-ACC.PLtantum
- a prach-y pů-jd-ou do [botiček]_i
et pognon-NOM.PLtantum FUT-aller-3PL dans petite chaussure-GEN.PL
- L2: jo takhle . to m-á-š dobře naplánovan-ý
MOD cela.ACC avoir-PRS-2SG bien planifié-N.ACC.SG
- L1: hlavně mus-í-m dostat tř-i tisíc-e ..
surtout devoir-PRS-1SG recevoir-INF trois-ACC mille-ACC.PL
- ...
- L1: tři a půl nej-lépe proto* .. nej-lépe čtyři
trois-ACC et demi-ACC.SG SUP-mieux parce* SUP-mieux quatre.ACC
- abych mě-l-a ještě na cest-u .
que.CONJ+AUX.COND.1SG avoir-PTCP-SG.F encore pour route-ACC.SG
- do Prahy a zpátky protože
à Prague-GEN.SG et de retour parce que

[ty	bot-y] _i	bud-u .	nakupovat	v Praze
ten.F.ACC.PL	chaussure-ACC.PL	AUX.FUT-1SG	acheter.INF	à Prague-LOC.SG

L1 : quand il a entendu que mon prochain salaire serait englouti par [des chaussures de danse qui coûtent trois mille couronnes]_i, il a éclaté de rire

L2 : et alors ! des chaussures, il en faut bien ! moi, ce que j'ai acheté par exemple pour mes cours de danses... alors soit ça glissait comme pas possible soit ça avait une semelle qui t'empêchait même de tourner sur toi ; et là, c'est vraiment pas possible

L1 : ouais, et pour faire du... du tango, il faut absolument les chaussures de danse

L2 : ouais, ça doit être trop cool d'avoir des chaussures de danse

L1 : ça, c'est sûr; et moi, j'ai trop hâte de [les]_i avoir; et je me [les]_i achète seulement après Noël

L2 : tu te fais un père Noël tardif

L1 : ouais, parce qu'à Noël, je vais recevoir du pognon, et le pognon, lui, il va servir à acheter [les petites chaussures]_i

L2 : ça alors, t'as bien planifié les choses

L1 : surtout, il faut que je reçoive trois mille... trois mille cinq cents dans l'idéal parce quatre mille dans l'idéal, pour pouvoir me payer un aller-retour à Prague parce que [les chaussures]_i, je vais [les]_i acheter à Prague*

ORTOFON v1

Le référent est introduit dans le discours par un complexe nominal nu. S'ensuit un échange au sujet des chaussures de danse en général. C'est le pronom personnel *ně* (elle.ACC.PL) qui reprend le référent en deuxième mention afin de l'établir dans le discours. Une troisième et une quatrième reprise suivent peu après, réalisées, respectivement, par un pronom personnel et un nom nu. Ainsi, six propositions plus tard, au moment de sa dernière reprise via TEN en position thématique, le référent reste bien activé dans l'esprit des locutrices, raison pour laquelle nous considérons que TEN remplit ici une fonction semblable à celle du défini français. Par ailleurs, nous pouvons nous interroger sur l'absence de TEN dans l'avant-dernière reprise, d'autant que selon Štícha, TEN devrait apparaître obligatoirement pour assurer la lecture coréférente d'un référent rhématisé (voir l'exemple 27 *supra*). Pourtant, cette coréférence ne nous semble nullement remise en question par l'absence de TEN : une tierce personne entendant le dialogue comprendra sans peine – grâce à l'ensemble du cadre évoqué par le discours – qu'il s'agit des chaussures que la locutrice envisage d'acheter. Nous proposons deux pistes d'explication partielle. D'abord, la volonté de la locutrice de recourir à une focalisation forte du référent afin de marquer son affection pour l'objet convoité ; cette hypothèse est étayée à la fois par la position rhématique du nom et par l'usage du diminutif affectif *botičky* (*petites chaussures*). Mais aussi la présence dans le contexte antérieur de deux référents entre lesquels s'établit un contraste externe (Corblin, 1987) : *botičky* (*les petites chaussures*) et *prachy* (*le pognon*).

Après avoir soigneusement examiné toutes les **reprises anaphoriques fidèles** avec TEN dans l'échantillon de conversations et dans le sous-corpus de Havel, nous estimons que 25 d'entre elles relèvent de la grammaticalisation dans l'échantillon oral, soit 44,64 % de l'ensemble de 56 occurrences (voir la Section 5.2.3.2), et que 11 d'entre elles relèvent de la grammaticalisation dans le

sous-corpus de Havel (soit 47,83 % de l'ensemble de 23 occurrences ; voir la Section 5.2.4.2). Voici l'exemple d'une occurrence où nous ne considérons pas TEN comme grammaticalisé :

- 37) ta tet-a ona tam m-á [...] právě [nějak-ou
ten.F.NOM.SG tante.NOM.SG elle.NOM.SG là-bas avoir-PRS.3SG MOD quelque-F.ACC.SG
- kamarádk-u]_i a vona byd-l-i ve Švýcarsk-u a m-á
copine-ACC.SG et elle.NOM.SG habiter-PRS-3SG en Suisse-LOC.SG et avoir-PRS.3SG
- jakoby ten .. její přítel je Švýcar a
MOD ten.M.NOM.SG sien.M.NOM.SG copain.NOM.SG être.PRS.3SG Suisse.NOM.SG et
- [*ta* *kamarádk-a*]_i normálně mluv-í i česky jo
ten.F.NOM.SG copine-NOM.SG MOD parler-PRS.3SG même tchèque MOD
- což je jako fajn
ce qui.NOM être.PRS.3SG MOD cool.N.NOM.SG
- cette tante, elle a [...] [une copine]_i là-bas et elle habite en Suisse et elle a [...] genre son copain, il est Suisse, et [cette copine-là]_i, elle parle tchèque et tout, ce qui est cool* ORTOFON v1

Le complexe nominal *ta kamarádka* constitue la première reprise d'un référent qui vient d'être introduit par le complexe nominal (*nějakou kamarádku*)⁹¹, raison pour laquelle TEN a ici été considéré comme étant plus proche de CE.

En anaphore infidèle, en revanche, TEN est bien moins susceptible de se rapprocher du défini français, ce qui se comprend notamment par la nécessité de maintenir le lien de coréférence en situation de recatégorisation du référent (voir la Section 3.3.2 et les critères à la fin de la Section 6.1.4). Une situation « extrême » de ce cas de figure est celle où la recatégorisation se fait moyennant un « nom de qualité », dont typiquement un vulgarisme. C'est cela même qui se produit en 38, où le locuteur parle de personnes qui lui doivent de l'argent (l'on voit que le français emploie ici le démonstratif) :

- 38) oni mně dluž-i-jú *ty* *kurv-y*
eux.NOM moi.DAT devoir-PRS-3PL ten.M.NOM.PL salopard.NOM.PL
- padesát čtyři-i tisíc
cinquante quatre-NOM mille.GEN.PL
- ces salopards, ils me doivent cinquante-quatre milles (couronnes)* ORTOFON v1

⁹¹ Parmi nos informateurs, à qui nous avons demandé de choisir entre *ta kamarádka*, *tahle kamarádka* et *kamarádka* (nom nu), sept ont opté pour la première, une pour la deuxième et deux pour la troisième possibilité. Ce dernier choix défie notre conclusion selon laquelle TEN serait ici un vrai démonstratif. Une piste d'explication possible : l'établissement d'un contraste externe entre les référents copine et copain.

En somme, seules 3 occurrences anaphoriques infidèles de TEN sur 13 (soit 23,08 %) ont été analysées par nous comme relevant de la grammaticalisation dans l'échantillon de conversations et une seule sur 11 dans le sous-corpus de Havel (voir les Sections 5.2.3.2 et 5.2.4.2). Seules ont été considérées comme grammaticalisées les occurrences où la recatégorisation est modeste au point que même le français peut employer le défini sans que la coréférence s'en trouve menacée. L'exemple 39 est une reprise de 32, cité *supra* :

- 39) právě [Runtour]_i začína-l v Plzn-i jestli se ne-mýl-í-m
justement Runtour commencer-PST.3SG.M à Pilsen.LOC.SG si REFL NEG-tromper-PRS-1SG
...
a co je možn-é zhlédnout vlastně při [tom běh-u]_i
et que.ACC être.PRS.3SG possible-N.NOM.SG voir.INF MOD pendant ten.M.LOC.SG course-LOC.SG
c'est justement [Roundtour]_i qui commençait à Pilsen, si je ne me trompe pas
...
et qu'est-ce qu'on peut voir, justement, pendant [la course]_i ? ORTOFON v1

En 39 (reprise de 32), il est question d'une course locale dont le nom est Runtour. Après avoir été introduit par son nom propre, le référent est successivement repris par les noms communs *běh* et *závod* (tous deux traduits par *course*)⁹² dans la suite du discours. Or il est une évidence pour les participants à la conversation que Runtour est une course. En conséquence, la recatégorisation du référent n'apporte, en réalité, aucune information nouvelle et ne risque pas de mettre en péril le lien de coréférence. La situation est similaire en 40 (l'auteur écrit, littéralement, *l'humain (ten člověk)*) :

- 40) Víš, [děcko]_i člověka hodně změní – najednou začneš úplně nově, jinak, hlouběji – chápat život – přírodu – lidi – a tvůj život dostává najednou – ať chceš nebo ne – tak nějak novou dimenzi jiný rytmus, jiný obsah, jiný řád – že je to tak, Věro? VĚRA Přesně tak! Vždyť si vem jenom tu odpovědnost, kterou najednou máš: na tobě záleží, jaký [ten člověk]_i bude [...]

[Un enfant]_i, [ça]_i te transforme un homme – c'est comme ça, Ferdinand. Tout à coup, on a une perception plus profonde de la nature, des autres, de la vie tout court. On n'y peut rien, l'existence gagne une autre dimension, un nouveau rythme, un contenu autre et un ordre plus rigoureux. C'est ça, Vera, n'est-ce pas ? VERA. Exactement. Imagine de quelle responsabilité tu es soudain investi. Le devenir [du petit être]_i dépend de toi seul.

V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

Considérons à présent le cas de l'**anaphore indirecte**. Celle-ci est globalement peu représentée dans nos corpus : 15 occurrences au total, dont 13 avec TEN. Ici aussi, la grammaticalisation a moins de chance d'apparaître, chose qui s'explique par la nature partielle de la coréférence⁹³. Dans nos

⁹² Il s'agit d'une reprise réalisée par un hyperonyme.

⁹³ Dans la Section 2.2.3, nous avons présenté cette notion plutôt complexe en expliquant que deux scénarii pouvaient se produire eu égard à la coréférence, celle-ci étant soit partielle soit absente. Avec un démonstratif adnominal, la coréférence semble toujours être présente et partielle.

corpus, il existe deux cas de figure : soit le référent de la reprise inclut celui de l'antécédent (exemple 41), soit c'est l'inverse qui se produit (exemple 42) :

41) L1 : on je rád protože pak vždycky dostan-e
 lui.NOM être.PRS.3SG content.M.NOM.SG parce que après toujours recevoir-PRS.3SG
 [sušenku]_i
 cookie-ACC.SG

L2 : tak kvůli [těm sušenk-ám]_{i+?} by si
 alors à cause de ten.F.DAT.PL cookie-DAT.PL AUX.COND.3SG REFL

L1 : to vytrp-í
 cela.ACC endure-PRS.3SG

L1 : il est content parce qu'à chaque fois, il reçoit [un cookie]_i

L2 : du coup, pour [ces cookies]_{i+?}, il

L1 : il tient bon ORTOFON v1

42) Katk-a [...] mě-l-a tady na návštěv-ě [dv-a Francouz-e]_{i+ii}
 NP-F.Nom.SG avoir-PST-3SG.F ici sur visite-LOC.SG deux-M.ACC Français-ACC.PL

jak byl-a ve Franci-i tak tam chodi-l-i
 comme être.PST-3SG.F en France-LOC.SG alors là-bas aller-PST-3PL.M

lézt na nějak-ou stěn-u
 escalader.INF sur quelque-F.ACC.SG mur-ACC.SG

...
 [ten jeden]_i je student
 ten.M.NOM.SG un.M.NOM.SG être.PRS.3SG student.NOM.SG

a [jeden]_{ii} už prac-uj-e
 et un.M.NOM.SG déjà travailler-PRS-3SG

Katka [...] a accueilli chez elle [deux Français]_{i+ii} ; au moment où elle était en France, ils allaient faire des escalades sur un mur d'escalade [...] [l'un]_i (d'entre eux) est étudiant et [l'autre]_{ii} travaille déjà. ORTOFON v1

Nous estimons qu'il y a grammaticalisation uniquement dans le second des cas de figure. Notons qu'en 42, le français semble recourir spontanément au défini, et que le tchèque peut recourir au nom nu sans que la coréférence soit mise en danger. Il en va d'une manière semblable en 43, emprunté à notre sous-corpus de Havel :

43) JIM Zajímala by mě ještě jedna věc: dosud jsme mluvili převážně o tom, jak nejlépe [ženské]_{i+ii} získávat. Neméně důležitou se mi však zdá být otázka, jak se [jich]_{i+ii} potom zbavovat – MACHEATH Myslíš [těch neodbytných]_{ii}?

JULES Il y a encore quelque chose qui m'intéresse. Jusqu'à maintenant on a surtout parlé de la méthode pour conquérir [les petites femmes]_{i+ii}. Comment s'[en]_{i+ii} débarrasser après, ça m'a l'air aussi une question très importante ... MAX Tu veux parler [des pots de colle]_{ii}, comme on les appelle, n'est-ce pas ? V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

La différence d'avec 42 est que l'anaphorique et son antécédent apparaissent ici en emploi générique⁹⁴. Rappelons qu'au total, le nombre d'occurrences de TEN apparaissant au sein de ce régime et que nous jugeons grammaticalisées est de 5 sur 12 (41,67 %) dans l'échantillon de conversations et de une sur une dans le sous-corpus des pièces (voir les Sections 5.2.3.2 et 5.2.4.2).

Enfin, bien que TEN semble mal se prêter à la grammaticalisation lorsqu'il apparaît en **anaphore résomptive** (voir *supra*), nos données contiennent tout de même une occurrence analysée comme telle et où TEN est certainement grammaticalisé :

44) MACHEATH Chcete tím říct, že jsem v pasti? LOCKIT Mám **ten dojem** –

MAX Ça signifie que je suis bel et bien pris au piège ? TOURMAKOFF J'en ai **l'impression**...

V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

Relevons ici deux phénomènes. D'abord, nous voyons bien que le français use d'une formulation légèrement différente de celle du tchèque : tandis qu'en tchèque, c'est le complexe nominal *ten dojem* qui institue un nouveau référent en résumant le contenu de l'énoncé *že jsem v pasti* (*que je suis bel et bien pris au piège*), dans la traduction française, c'est le pronom adverbial *en* qui est responsable de cette opération, le nom *impression* étant ainsi employé comme un concept fonctionnel saturé par l'argument possessif *en* (*j'ai l'impression de cela*). Ensuite, le nom *dojem* (*impression*) est un nom abstrait apparaissant dans un tour bien figé et où la présence de TEN est quasi-obligatoire. Dans les Sections 2.2.3 et 3.3.2, nous avons commenté le tour *tím pádem* (*du coup, par conséquent*), fonctionnant comme un connecteur logique, et où TEN est indispensable. Nous pouvons encore ajouter à ces deux tours celui de *v tom (tomto/tomhle) případě* (*dans ce cas(-là)*), apparaissant 4 fois dans nos données (dont 3 fois dans le sous-corpus de Havel). Ces trois tours figés (nous pourrions en trouver d'autres) ont tous en commun de contenir un nom abstrait, d'exiger la présence de TEN (éventuellement d'un autre démonstratif) et d'être valorisés dans la structuration du discours⁹⁵.

Intéressons-nous maintenant aux **occurrences anaphoriques de TEN éligibles à la grammaticalisation et où celui-ci se combine avec un autre élément fonctionnel** (voir les Sections 5.2.3.3 et 5.2.4.3 du chapitre précédent). Il convient d'abord de dire que la grammaticalisation ne peut être envisagée pour les combinaisons qui rendent l'ensemble du complexe nominal indéfini sur le

⁹⁴ Dans ce cas précis, nous nous interrogeons si, malgré la position postverbale (rhématique) du référent, la lecture anaphorique requiert absolument la présence de TEN. La raison de cette hésitation est l'ellipse du nom : comme celui-ci doit, pour la compréhension de l'énoncé, être restitué à partir du contexte antérieur, l'interprétation du complexe nominal ne peut être qu'anaphorique. En même temps, la présence de TEN est d'autant plus naturelle qu'il s'agit, justement, d'un complexe nominal avec ellipse du nom (voir la Section 3.4).

⁹⁵ Compte-tenu de sa remplaçabilité par TENTO/TENHLE, TEN ne saurait être considéré comme grammaticalisé dans *v tom případě*. Rappelons que dans notre corpus de conversations françaises, ce rôle est avant tout assuré par des tours anaphoriques résomptifs figés contenant les noms *cas* et *moment* (voir les Sections 2.2.3, 4.5.2 et 5.2.2.1).

plan dénotatif, car cela serait en contradiction directe avec le sémantisme du défini. Malheureusement, il n'est pas toujours chose aisée de trancher entre la lecture définie et indéfinie. Ainsi, dans la combinaison TAKOVÝ TEN, vu que le premier élément est indéfini, il serait logique de considérer l'ensemble du complexe nominal comme indéfini (voir la citation d'Uhlířová, 1992 : 254 *infra*). D'autre part, comme nous l'avons vu dans la Section 5.2.3.3, la combinaison permet l'identification d'une entité, bien que TAKOVÝ ait justement pour fonction de marquer une certaine difficulté de cette identification, difficulté due à la nature imprécise et/ou insuffisamment familière du référent dans l'esprit du locuteur (voir la Section 5.2.3.3, où nous avons relevé une certaine proximité de cette construction avec CE + *genre/type/espèce(s) de*).

La combinaison NĚJAKÝ (QUELQUE/UN CERTAIN) TEN pose globalement le même problème d'analyse dans la mesure où le premier élément est certes indéfini, mais le tout peut également être employé pour saisir un référent présentant des aspects inconnus, voire suscitant distanciation et méfiance chez le locuteur⁹⁶. Et pour rendre les choses encore plus complexes, il faut rappeler que la combinaison NĚJAKÝ TEN peut aussi être investie d'un sens partitif, auquel cas le premier élément constitue un complexe nominal à part entière :

45) <i>nějak-é</i>	<i>to</i>	<i>piv-o</i>
quelque-N.NOM.SG	ten.N.NOM.SG	bière-NOM.SG
<i>un peu de (cette/la) bière</i>		

La même lecture partitive peut également avoir lieu avec les combinaisons KAŽDÝ TEN, NĚKTERÝ TEN, *kolik těch, trochu toho*⁹⁷ et numéral + TEN. En revanche, lorsque c'est le démonstratif qui occupe la première position, l'ensemble jouit toujours d'une interprétation définie. Uhlířová (1992 : 254) résume :

Il est ainsi possible de conclure : en principe, il est vrai que si plusieurs déterminants à la fois sont présents au sein d'un SN, c'est celui qui se trouve le plus à gauche qui décide de la définitude / indéfinitude du SN. Cela veut dire que si des déterminants de types différents, c'est-à-dire un déterminant défini et un déterminant indéfini, se retrouvent ensemble dans un SN, leur ordre est distinctif sur le plan sémantique. En coprésence de deux, éventuellement de plus de déterminants indéfinis, des variations dans l'ordre des mots sont possibles, qui permettent, dans certains cas, d'exprimer diverses nuances sémantiques accessoires⁹⁸.

⁹⁶ Voir à cet égard Uhlířová (1992 : 251) : « La combinaison d'un déterminant indéfini avec le déterminant défini *ten* la plus souvent réalisée est celle de *takový ten*, où le sens de *takový* est plus ou moins synonyme de celui de *nějaký* avec, en plus, la nuance possible de "je ne sais pas exactement lequel, en fait, je ne sais pas lequel, je cherche le terme approprié, je ne trouve pas l'expression exacte", etc. De même que dans la combinaison avec le déterminant *nějaký*, *ten* réfère aux sphères contextuelles les plus larges. » Nous traduisons : „Nejčastěji realizovanou kombinací neurčitěho determinátoru s určitým determinátorem *ten* je hovorové a obecně *takový ten*, kde *takový* má význam víceméně synonymní s *nějaký*, navíc s možným odstínem „nevím přesně jaký, vlastně nevím jaký, hledám vhodný atribut, nenapadá mě přesný výraz“ atd. *Ten*, podobně jako v kombinaci s determinátorem *nějaký*, referuje k nejširším kontextovým sféram.“

⁹⁷ KAŽDÝ = *chaque*, NĚKTERÝ = (UN) CERTAIN, *kolik* = *combien de*, *trochu* = *un peu de*.

⁹⁸ Nous traduisons : „Lze tedy uzavřít: V zásadě platí, že je-li v NP přítomno více determinátorů současně, rozhoduje o příznaku určenosti / neurčenosti NP *ten*, který stojí nejvíce vlevo. To znamená, že pokud jsou v NP přítomny

De manière générale, vu son profil équivoque, la combinaison TAKOVÝ TEN a été écartée de la grammaticalisation dans ce travail. Quant aux combinaisons NĚJAKÝ TEN, KAŽDÝ TEN, NĚKTERÝ TEN, *kolik těch*, *trochu toho* et numéral + TEN, ainsi que toutes celles où TEN occupe la place la plus à gauche du complexe nominal, elles ont été évaluées pour la grammaticalisation lorsqu'elles apparaissaient dans un emploi a priori compatible avec celle-ci, soit un emploi endophorique fidèle, infidèle ou indirect ou un emploi mémoriel. Le domaine de la deixis situationnelle, quant à lui, n'accueille aucun cumul dans nos données (voir les Sections 5.2.3.3 et 5.2.4.3).

Dans l'échantillon de conversations, ces occurrences ne figurent qu'au nombre de 2 – TEN JEDEN (TEN + UN) et TEN SVŮJ (TEN + SON⁹⁹) –, dont aucune n'a été évaluée comme contenant un TEN grammaticalisé. Dans le sous-corpus de Havel, elles sont au nombre de 3 et toutes contiennent TEN suivi d'un possessif : TEN VÁŠ (2 fois) et TEN SVŮJ (une fois). Ces trois occurrences n'ont pas non plus été évaluées comme relevant de la grammaticalisation. Tâchons de justifier notre parti-pris à l'aide des deux exemples suivants :

- 46) m-á-š na hlav-ě takov-ou kamer-u teda pardon m-á-š tam
avoir-PRS-2SG sur tête-LOC.SG tel-F.ACC.SG caméra-ACC.SG MOD pardon avoir-PRS-2SG là-bas
- [projektor]_i jo pak tam m-á-š ještě jakoby senzor jo
projecteur.ACC.SG MOD après là-bas avoir-PRS-2SG encore MOD capteur.ACC.SG MOD
- takže [...] ten senzor ti vlastně přijím-á to co ty
du coup ten.M.NOM.SG capteur.NOM.SG toi.DAT MOD capter-PRS.3SG ce.qui.ACC.SG toi.NOM
- projekt-uj-eš skrze [**ten** **svůj** **dataprotektor**]_i
projecter-PRS-2SG à travers ten.M.ACC.SG ton.REFL.M.ACC.SG projecteur de données-ACC.SG
- sur la tête, t'as une espèce de caméra, pardon, je veux dire t'as [un projecteur]_i, tu vois ; puis t'as aussi un genre de capteur ; du coup, ce capteur capte ce que tu projettes avec [ton projecteur de données]_i*
ORTOFON v1

- 47) SLÁDEK Co vy jste to vlastně psal, jestli se můžu zeptat? VANĚK [Divadelní hry]_i - SLÁDEK [Divadelní hry]_i? A [to]_i se hrálo někde na divadle? VANĚK Ano - SLÁDEK Hm. No jo, no jo. Tak [divadelní hry]_i, říkáte.
- ...
- SLÁDEK O čem byly [**ty vaše hry**]_i?

SLÁDEK. Sans indiscretion – c'était quoi que vous écriviez ? VANEK. [Des pièces de théâtre]_i. SLÁDEK. [Des pièces de théâtre]_i... et alors quelque part, dans un théâtre, on [les]_i jouait ?

determinátory různého druhu, tj. determinátor určitý a determinátor neurčitý, je jejich pořadí sémanticky distinktivní. Při souvškytu dvou, ev. více determinátorů neurčitosti jsou možné slovosledné variace, jimiž lze v některých případech navíc vyjadřovat různé druhotné sémantické odstíny.“

⁹⁹ En tchèque, le paradigme SVŮJ est un possessif réfléchi, employé pour exprimer la possession relativement à toutes les personnes grammaticales, mais uniquement lorsque le possesseur est le même que le sujet de la proposition : *Dejte mi své slovo* = *Donnez-moi votre parole* vs *Dali mi vaše slovo* = *Ils m'ont donné votre parole*.

VANEK. Oui. SLADEK. Hum... [des pièces de théâtre]_i... Ouais, ouais, ouais, ouais, ouais, ouais, dites voir...

...
SLADEK. Et vous alors, [vos pièces]_i parlaient de quoi ?

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

Dans les deux cas, nous avons TEN suivi d'un possessif. Remarquons d'abord que dans les deux cas, la traduction française (dont la première est de nous), recourt au seul possessif. En effet, comme nous l'avons rappelé dans la Section 4.1.2, le français fait partie des langues où le possessif, du fait de son alignement syntactico-sémantique sur le défini, porte le présupposé d'unicité (« de totalité/maximalité » pour les noms pluriels et indénombrables ; cf. Hawkins, 1978) : le SN qu'il introduit est, par conséquent, quasiment toujours interprété comme défini¹⁰⁰. Or en tchèque, le possessif est neutre sur le plan de la définitude, pouvant introduire à la fois des complexes nominaux définis et indéfinis. Si dans le cas des complexes nominaux *vaše hry* et *svůj dataprojektor*, la pragmatique fait clairement pencher vers une interprétation définie (une personne possède un seul projecteur de données dans le cadre établi par le discours, et l'on pourrait difficilement poser la question portant sur le contenu d'un ensemble de pièces dans une situation où celui-ci ne serait pas présupposé dans sa totalité), la présence de TEN remplit au moins une double fonction : 1) elle entérine cette interprétation définie, de sorte qu'aucun doute n'est permis à son sujet et 2) elle rend explicite le lien de coréférence entre le complexe nominal en question et son antécédent. Contrairement à la situation en français, cette coréférence irait beaucoup moins de soi si TEN faisait défaut, a fortiori lorsque, comme en 47, une distance considérable entre les deux expressions (soit 33 propositions en 47) fait perdre au référent une bonne part de son activation discursive.

Ensuite, force est de constater que les deux noms (*dataprojektor*, *hry*) sont ici employés comme des concepts relationnels, à savoir comme des concepts appelant la présence d'un argument possessif, cependant que cet argument ne suffit pas à lui seul à établir l'unicité dénotative (voir la Section 6.1.3.4). Cette dernière est en réalité le résultat de la synergie entre l'argument possessif et le contexte, où la première mention du référent a été introduite. Ainsi, TEN sert ici à indexer le référent sur ce contexte, exactement comme lorsqu'il apparaît tout seul. Et vu qu'il s'agit tantôt de la toute première reprise du référent (en 46), tantôt d'une reprise du référent suite à une période de désactivation de ce dernier (en 47), TEN ne fonctionne pas comme un défini, mais comme un démonstratif (il s'agit, respectivement, d'un TEN « établissant » et d'un TEN « réactivant », voir la fin de la Section 6.1.4).

¹⁰⁰ Exception faite du cas de son apparition en fonction syntaxique d'attribut caractérisant (non référentiel) où il introduit un nom relationnel pur (« non-fonctionnel » ; cf. Löbner, 2011) : *C'est mon ami/frère*. Bien sûr, il serait également possible d'opter, dans la traduction française, pour le démonstratif (*ce projecteur de données, ces pièces*).

Somme toute, nous constatons que parmi toutes **les occurrences anaphoriques** fidèles de TEN **du sous-corpus de Havel** analysées comme grammaticalisées, 7 sont **transposées en français** par le défini, 3 par le possessif et une par l'article indéfini (voir la Section 5.2.4.4). Quant aux occurrences anaphoriques fidèles non-grammaticalisées, 5 d'entre elles sont transposées par le démonstratif, 2 par un autre moyen, 2 par le possessif (soit 2 des 3 cumuls TEN + possessif mentionnés *supra*, dont l'exemple 46), une par l'article indéfini et 2 ne sont pas transposées (le propos fait l'objet d'une reformulation en français). Dans le domaine de l'anaphore infidèle, la seule occurrence grammaticalisée – soit l'exemple 39 *supra* – est transposée par le défini. Parmi les occurrences non-grammaticalisées, 6 sont rendues par le démonstratif, une par un autre moyen, une par le possessif et 2 ne sont pas transposées pour cause de reformulation. Enfin, une seule occurrence de TEN – soit l'exemple 42 *supra* – se prête à l'emploi anaphorique indirect. Elle est considérée comme grammaticalisée et transposée par le défini.

Nous relevons par ailleurs que les transpositions effectuées par l'indéfini et celles effectuées par un autre moyen grammatical relèvent elles aussi d'une légère reformulation du propos, d'un point de vue légèrement différent que le traducteur adopte :

48) Co to neslo, **ty hry?**

*Combien ça rapporte, **une pièce de théâtre** ?*

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

49) FILCH No přece proto, že mám své zásady! Jednou z nich totiž je, že nikdy neudám [kolegu]_i, i kdybych proti [němu]_i měl sebevětší výhrady – LOCKIT Což platí zřejmě i pro případ, kdy vás [**ten kolega**]_i sám udal –

*FAUCHE Au nom de mes principes, justement ! Et il y en a un auquel je tiens plus particulièrement : ne jamais dénoncer [un compagnon de lutte]_i – même si nous nous opposons...
TOURMAKOFF Et même, si je comprends bien, dans le cas où [**celui-ci**]_i vous dénonce ?*

V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

En 48, les traducteurs ont opté pour un indéfini en emploi générique (*une pièce de théâtre en général*), alors que l'auteur renvoie à un ensemble spécifique de pièces, mentionné en amont. En 49, le traducteur utilise un pronom démonstratif. Ce pronom reprend – de même que le complexe nominal démonstratif de l'original – la référence d'un SN indéfini – *un compagnon de lutte* (*un collègue* en tchèque). Quoi qu'il en soit, une telle reformulation s'écarte moins de l'original que le cas que nous analysons comme une absence de transposition (cf. l'exemple 50) :

50) JOHN Pan Macheath tu má [důležitou návštěvu]_i –

...

JENNY Dobrá, **počkám tedy venku, až [ta návštěva]_i skončí**

GEORGES Il y a une visite très importante pour Monsieur Maxence...

...

ESTELLE Bon, *j'attendrai dehors...* / vs dans l'original : *j'attendrai dehors jusqu'à ce que la visite soit terminée*
V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

Le fait de rendre certains emplois anaphoriques de TEN (y compris ceux qui ne se combinent pas avec le possessif en tchèque) par le déterminant possessif reflète le penchant structurel du français, déjà mentionné dans la Section 5.2.4.4, à davantage recourir à ce moyen. Rappelons qu'au total, 11 occurrences anaphoriques du sous-corpus de Havel (éligibles ou pas à la grammaticalisation) sont rendues par le possessif (voir la Section 5.2.4.4) :

51) Komu myslíš, že jsme chtěli *ty desky* přehrát [...]?

Qui, crois-tu, devait avoir la primeur de nos disques nouveaux ?

V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

En 51, il s'agit de la seconde reprise d'un référent précédemment introduit, à savoir de disques vinyles que l'un des protagonistes avait ramenés de son voyage en Suisse. Comme en 46 et 47, le possessif français combine la présence d'un argument possessif avec l'indexation du référent sur le contexte discursif (*les nouveaux disques à nous dont il a été question*). Dans l'approche löbnérienne, le nom constitue ainsi un concept relationnel. Contrairement à 46 et 47, TEN est ici considéré comme grammaticalisé, car il apparaît en troisième mention du référent, et que cette deuxième reprise se trouve à proximité de la première reprise, au moyen de laquelle le référent a été à la fois établi dans le discours et activé (la fonction « établissante » du démonstratif, voir *supra*).

Enfin, il est un dernier facteur sur lequel il pourrait s'avérer intéressant de se pencher ici, à savoir **la fonction syntaxique du complexe nominal anaphorique**, déjà mentionnée *supra*. Ce facteur fait l'objet d'une analyse à la fois chez Laury (1997) et chez Zíková (2017). La première observe que la fonction syntaxique influe sur le taux de marquage du nom par le démonstratif adnominal SE. Dans son corpus, ce taux obéit à une tendance croissante suivant le schéma :

complément du nom < sujet¹⁰¹ < complément d'objet indirect < complément d'objet direct

Les analyses faites par Zíková sur son propre corpus étayent largement cette tendance, la seule différence étant que les noms en fonction de sujet apparaissent, chez elle, encore moins souvent avec TEN que les noms en fonction de complément du nom. L'explication du phénomène qu'elle livre au lecteur de son travail repose sur l'argument selon lequel les référents définis introduits en fonctions de COD et de COI requièrent, de manière générale, un marquage plus explicite de leur définitude :

¹⁰¹ Pour la fonction sujet, l'auteure opère, en plus, la distinction entre les sujets des verbes non-transitifs et ceux des verbes transitifs, sachant que parmi ses propres données, les sujets nominaux des premiers sont encore moins marqués par SE que les noms des seconds.

L'explication est en vérité simple : la position de COD et celle de COI ont ceci de particulier qu'elles introduisent dans le discours des référents nouveaux, non-identifiables et indéfinis. Conscient qu'il est de cette convention qui détermine les attentes de l'interlocuteur, le locuteur accompagne sa violation d'un signal explicite de la nature identifiable du référent. Se trouve, derrière cela, le principe pragmatique général suivant : moins l'information est prédictible, plus elle est communiquée explicitement ; plus l'information est prédictible, plus elle est communiquée implicitement. Et comme le sujet est, de manière générale, occupé par des référents identifiables, nulle nécessité de les marquer explicitement comme tels¹⁰². (2017 : 113)

De plus, l'auteure s'intéresse également aux occurrences de TEN adnominal au sein de la fonction de complément circonstanciel, arrivant au constat que c'est le complément circonstanciel de lieu qui affiche le plus grand nombre d'occurrences de TEN adnominal dans ses données (la position de celui-ci eu égard à ce critère se situe entre le complément du nom et le COI).

Nous avons ainsi analysé la fonction syntaxique de nos propres occurrences de TEN adnominal apparaissant en emplois anaphoriques fidèle, infidèle et indirect, autant d'emplois endophoriques qui ressortissent à la définitude « pragmatique » et où la grammaticalisation peut se produire. Les Tableaux 6.3 et 6.4 présentent les nombres d'occurrences appartenant aux fonctions : sujet, COD, COI, complément du nom et complément circonstanciel¹⁰³ :

	Sujet	COD	COI	Complément du nom	Complément circonstanciel
Anaphore fidèle	22	12	3	4	13
Anaphore infidèle	6	1	1	1	3
Anaphore indirecte	5	1	0	0	4
TOTAL	33	14	4	5	20

Tableau 6.3 : Nombre d'occurrences de TEN adnominal apparaissant en emplois anaphoriques fidèle, infidèle et indirect en fonctions de sujet, COD, COI, complément du nom et complément circonstanciel dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

	Sujet	COD	COI	Complément du nom	Complément circonstanciel
Anaphore fidèle	6	10	0	1	2
Anaphore infidèle	2	4	3	1	3
Anaphore indirecte	0	0	1	0	0
TOTAL	8	14	4	2	5

Tableau 6.4 : Nombre d'occurrences de TEN adnominal apparaissant en emplois anaphoriques fidèle, infidèle et indirect en fonctions de sujet, COD, COI, complément du nom et complément circonstanciel dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

¹⁰² Nous traduisons : „Vysvětlení je ve skutečnosti jednoduché: pozice objektu a oblique jsou charakteristické tím, že do diskursu uvádějí nové, tedy neidentifikovatelné, neurčité referenty. Mluví si je této konvenční vazby, která určuje posluchačovo očekávání, vědom, a proto její narušení provází explicitním signálem identifikovatelnosti referentu. Obecný pragmatický princip v pozadí: čím méně predikovatelná informace, tím explicitněji je komunikována; čím predikovatelnější informace, tím implicitněji formulována. A protože subjekty jsou typicky obsazovány identifikovatelnými referenty, není třeba je jako identifikovatelné explicitně značit.“

¹⁰³ Nous ne distinguons pas entre les différents types de compléments circonstanciels. Une poignée de cas, n'appartenant à aucune de ces fonctions, n'a pas été comptée.

Les résultats de nos données se montrent « décevants » au vu des attentes issues du bilan de Laury (1997) et de Ziková (2017) : dans notre échantillon de conversations, la position de sujet est celle qui accueille le plus d'occurrences de TEN, celle de COD n'arrivant, quant à elle, qu'en 3^{ème} position. Dans le sous-corpus de Havel, où le nombre d'occurrences concernées est cependant assez faible, il s'agit de la 2^{ème} position la plus représentée, après celle de COD. La fonction de vraisemblance, appliquée aux fréquences absolues de TEN adnominal dans les différentes fonctions syntaxiques de l'échantillon de conversations, affiche un score statistiquement significatif pour la différence entre la fonction de sujet et toutes les autres fonctions¹⁰⁴.

Il faut cependant préciser que contrairement aux deux auteures susmentionnées, qui ont étudié la proportion des occurrences introduites par le démonstratif adnominal au sein de toutes les expressions référentielles apparaissant dans chaque fonction syntaxique, nos propres analyses se limitent aux seules occurrences introduites par le démonstratif. Par conséquent, si les résultats repris dans les Tableaux 6.3 et 6.4 nous permettent de dire que la fonction sujet n'est pas sous-représentée, dans notre échantillon, parmi les occurrences de TEN, ces résultats ne nous disent strictement rien sur l'inverse, à savoir si les occurrences de TEN sont sous-représentées au sein de l'ensemble des sujets syntaxiques. Afin de répondre à cette question, il nous faudrait avoir analysé toutes les occurrences anaphoriques de notre corpus, tâche qui dépasserait de loin les limites et les objectifs de notre travail¹⁰⁵.

Mais ce n'est pas tout. En étudiant différentes sources de référence relatives à la grammaticalisation du démonstratif en défini, nous nous sommes aperçu que celles-ci contiennent deux thèses (en partie) antagonistes. Si la première – celle qui vient d'être présentée – se base sur des arguments sémantico-pragmatiques pour justifier une présence plus faible du démonstratif au sein de la fonction de sujet, la seconde, elle, mobilise d'autres arguments sémantico-pragmatiques pour avancer, au contraire, que c'est dans cette fonction que le démonstratif serait le plus fréquent. Telle est, par exemple, la position de Selig (1992) et de Trovesi (2002) :

¹⁰⁴ $L = 4,7756$ pour la différence entre le sujet et le complément circonstanciel, $13,9806$ pour celle entre le sujet et le COD, $32,7345$ pour celle entre le sujet et le COI et, enfin, $29,464$ pour celle entre le sujet et le complément du nom ; la valeur critique = $3,841$ (degré de liberté = 1 et $\alpha = 0,05$). Les fréquences absolues ont été comparées à l'ensemble de la population (n), soit toutes les occurrences anaphoriques fidèles, infidèles et indirectes introduites par TEN.

¹⁰⁵ Aussi inspirants et déterminants qu'ils soient pour nos recherches, les travaux de ces deux auteures ont un objectif clairement différent du nôtre : étudier la grammaticalisation du démonstratif en article défini à l'intérieur de l'emploi anaphorique. Certes, ce choix les amène inévitablement à analyser les autres types d'emploi dans la mesure où ceux-ci apparaissent dans leur corpus. Mais en lui-même, ce corpus est choisi/conçu de manière à maximiser le nombre d'occurrences anaphoriques. De plus, sa taille limitée – 1 968 expressions référentielles annotées par Laury et 3 472 expressions référentielles annotées par Ziková – permet une exploration beaucoup plus approfondie de l'ensemble et non seulement des expressions introduites par le démonstratif adnominal, et ce eu égard à toute une batterie de variables (nature animée ou non du référent, position de l'expression vis-à-vis du verbe, l'ordre de la mention du référent, etc.). Ajoutons encore qu'en ce qui concerne la position vis-à-vis du verbe, Ziková (2017 : 263) constate que parmi ses données, les occurrences introduites par TEN ne manifestent d'affinité avec aucune position en particulier.

[...] fokalisiert werden die Satzfunktionen, die bedingt durch die Rollenverteilung in der semantisch-propositionalen Ebene bereits Affinitäten zu kommunikativ relevanten Referenten haben [...] (Selig, 1992 : 203)

La preuve qu'en tchèque et en sorabe, l'expression de la définitude d'un groupe nominal au moyen de *ten* et *tón* est étroitement liée à (et peut-être même directement dépendante de) l'accentuation emphatique et expressive du référent en question, c'est la fréquence élevée de ces pronoms démonstratifs en fonction d'article au nominatif et à l'accusatif : 41,51 % au nominatif et 27,86 % à l'accusatif en tchèque et 62,59 % au nominatif et 24,55 % à l'accusatif en sorabe. L'objet [direct] et le sujet sont les éléments les plus importants dans la phrase et, du point de vue pragmatique, ils se trouvent déjà au centre de l'attention du locuteur et de l'interlocuteur. L'emploi de *ten* en tchèque et de *tón* en sorabe permet de continuer sa focalisation au sein de l'énoncé¹⁰⁶. (Trovesi, 2002 : 134)

Précisons que de la même manière que nous, Trovesi ne prend en compte que les occurrences introduites par le démonstratif, mais, contrairement à nous, il s'intéresse à tous les types d'emploi et non seulement à l'anaphore. Cette différence mise à part, nos propres résultats relatifs au domaine anaphorique ne s'éloignent pas radicalement des siens : dans l'échantillon de conversations, 43,42 % des occurrences avec TEN apparaissent en fonction de sujet et 18,42 % en fonction de COD ; dans le sous-corpus de Havel, 32 % apparaissent en fonction de sujet et 36 % en fonction de COD.

Nous constatons ainsi que les deux positions exposées *supra* aboutissent à une dissonance cognitive, dissonance que nous n'avons ici ni les moyens ni l'intention de réduire. Disons juste qu'une voie de sortie possible de cette aporie pourrait être de considérer que cette tendance antagoniste reflète deux fonctions différentes de TEN, sachant que, selon le contexte, c'est tantôt l'une ou l'autre qui est susceptible d'être valorisée : la fonction qui consiste à identifier le référent, particulièrement pertinente lorsque celui-ci occupe la position d'objet, et la fonction pragmatique, consistant à marquer la proéminence du référent, qui, elle, touche avant tout les référent en position de sujet.

6.2.1.3 L'emploi exophorique mémoriel

À plusieurs reprises dans ce travail (les Sections 3.3.4, 4.5.3 et 5.2.4.4), nous avons relevé les spécificités de l'emploi mémoriel, dont la ligne de démarcation en partie floue entre l'emploi du défini (ou, dans les langues qui n'en disposent pas, du nom nu) et celui du démonstratif. Répétons que ce choix dépend, en premier lieu, du coût d'activation du référent dans les esprits du locuteur et de l'interlocuteur (voir la Section 4.5.3). Mais comment mesurer ce coût d'activation ? Cette difficulté fait que c'est précisément dans le domaine de l'emploi mémoriel que l'évaluation de la

¹⁰⁶ Nous traduisons : „Důkazem toho, že v češtině a lužické srbštině je vyjádření určenosti nominální skupiny prostřednictvím *ten* a *tón* úzce spojeno (a možná i přímo na něm závislé) s emfatickým a expresivním zdůrazněním míněného referentu, je vysoký výskyt těchto ukazovacích zájmen ve funkci členu v nominativu a akuzativu: v češtině v nominativu 41,51 %, v akuzativu 27,86 %, v lužické srbštině v nominativu 62,59 % a v akuzativu 24,55 %. Předmět a podměť jsou nejdůležitějšími elementy ve větě a z pragmatického hlediska jsou již ve středu pozornosti mluvčího a posluchače; užití českého *ten* a lužického *tón* umožňuje jeho další zdůraznění v promluvě.“

grammaticalisation s'est avérée la plus délicate¹⁰⁷. Il serait tentant de dire qu'une occurrence de TEN adnominal en emploi mémoriel pourrait être considérée comme engagée dans le processus de grammaticalisation dès lors que sa traduction en français requiert la présence de l'article défini. Pourtant, ce critère est d'une efficacité très limitée, car, comme nous venons justement de le rappeler, le locuteur d'une langue à articles a souvent le choix entre le défini et le démonstratif pour renvoyer à un référent dont l'identification se base sur des connaissances partagées entre lui et l'interlocuteur. Nous reproduisons ici l'exemple déjà commenté dans la Section 3.3.4 :

- 52) L1 : tady jsou **ty** **promenád-y** ne ? kolem
ici être.PRS.3PL ten.F.NOM.PL promenade-NOM.PL MOD autour
- ty** **chodník-y**
ten.M.NOM.PL trottoir-NOM.PL
- L2 : jo tady-hlenc je **ta** **promenád-a**
MOD ici-SUFFIXE DEICTIQUE être.PRS.3SG ten.F.NOM.SG promenade-NOM.SG
- L1 : ici, il y a **les/ces promenades** et **les/ces trottoirs** tout autour, non ?
L2 : ouais, là, il y a **la/cette promenade**

De plus, le test de la traduction pose plusieurs problèmes à la fois : s'il est question d'une traduction proposée par nous, celle-ci risque forcément d'être biaisée. Si nous prenons une traduction publiée (ce que nous allons en partie faire, mais rappelons que cette option n'est possible que pour les occurrences du sous-corpus de Havel), non seulement elle risque aussi d'être biaisée, mais, en plus, de représenter un cas de traduction libre qui ne se soucie guère de la fidélité à la formulation de l'original (voir notamment la Section 5.2.4.4).

En même temps, une question importante qui se pose pour les énoncés comme 52 est la suivante : TEN est-il ressenti comme obligatoire par un locuteur tchèque ? Les référents des complexes nominaux *ty promenády* et *ty chodníky* sont identifiables bien que nouveaux sur le plan discursif (cf. Laury, 1997). En ce qui nous concerne, nous avons tendance à considérer ici TEN comme obligatoire et proche du défini français, donc comme grammaticalisé. S'il était absent, les noms seraient interprétés ou bien comme référant à des définis « sémantiques » (des « permanently established individual concepts » de Gerland et Horn, 2010 ; voir la Section 3.5.3), ou bien – vu surtout leur position rhématique au sein de l'énoncé – comme des indéfinis (voir la fin de la Section 6.1.4). Cette dernière interprétation est appuyée par nos informateurs, à qui nous avons demandé de choisir entre *ty promenády/ty chodníky/ta promenáda* et *promenády/chodníky/promenáda*. Les personnes ayant

¹⁰⁷ Si elle n'est pas non plus évidente en contexte anaphorique, ce dernier domaine permet, comme nous l'avons vu dans la section précédente, de se baser sur une série de critères objectifs, notamment le degré d'activation du référent, mesuré par la distance anaphorique entre sa mention actuelle et précédente, et le degré de son ancrage dans le discours, mesuré par l'ordre de la mention.

choisi la version avec les noms nus sont au nombre de 4. C'est une personne en particulier qui nous a fait savoir que pour elle, la variante avec TEN implique clairement qu'il s'agit d'un référent défini identifié via une mémoire commune : les deux personnes regardent des photos prises lors de vacances passées ensemble ; à l'inverse, la variante où les noms sont nus impliquerait automatiquement une lecture indéfinie : le locuteur regarde des photos prises lors de vacances où l'interlocuteur est parti seul. Dans nos consignes, nous n'avions pas précisé qu'il était question du premier scénario (voir la Section 3.3.4), omission qui a fini par s'avérer bénéfique pour la question traitée ! Un autre exemple a été présenté à nos informateurs :

53) Pamatuj-eš si na (*toho*) *divn-ýho* *chlápka*,
rappeler-PST-2SG REFL à ten.M.ACC.SG étrange-M.ACC.SG type-ACC.SG

co jsme ho potka-l-i v Řím-ě?
REL AUX.PST.1PL lui.ACC rencontrer-PTCP-PL.M à Rome-LOC.SG

Tu te rappelles le/ce type étrange qu'on a croisé à Rome ?

En 53, la valeur de TEN est équivoque : il est impossible de trancher entre la lecture en défini et celle en démonstratif. Dans les deux cas, la relative fait clairement, pour nous, l'objet d'une interprétation restrictive, contribuant ainsi directement à l'identification du référent (sans elle, l'énoncé *Tu te rappelles le/ce type ?* est extrêmement vague). Ici, la variante avec TEN a été choisie par 9 informateurs sur les 10, bilan qui fait clairement pencher vers l'hypothèse de grammaticalisation¹⁰⁸.

Comme nous l'écrivions *supra* (Section 6.1.4), une hausse de la fréquence du démonstratif en emploi mémoriel constitue souvent un signe de sa grammaticalisation en défini. Ainsi, Laury a démontré dans son travail consacré à la grammaticalisation du démonstratif SE en finnois parlé que la présence de celui-ci va croissant, dans les occurrences mémorielles, à mesure que sa grammaticalisation progresse :

[...] in the earlier narratives¹⁰⁹, no noun phrases with new but identifiable referents had been *se*-marked, but in the data from the 1930s and 40s, a full 60% of such noun phrases are *se*-marked. (1997 : 189)

Au vu de ces faits, il nous a semblé le plus opportun de considérer par défaut les référents mémoriels introduits par TEN comme participant de la grammaticalisation, exception faite notamment des cas où TEN se cumule avec TAKOVÝ, ambigus pour la présence de ce dernier (voir la section

¹⁰⁸ De nouveau, il est intéressant de voir que la situation semble être la même en slovaque, où une forte préférence pour TEN existe à la fois lorsque le nom n'est pas modifié par une relative (comme en 52) et lorsqu'il l'est (comme en 53). Le polonais (parlé informel) semble, quant à lui, plus réticent à utiliser TEN, notamment lorsque le nom n'est pas modifié par une relative.

¹⁰⁹ Ces textes narratifs datent de la période 1880-1930.

précédente) et ceux où il se cumule avec NĚJAKÝ (soit l'hypothèse n°3 *supra*). Dans ce dernier cas, la grammaticalisation peut se produire, selon nous, uniquement lorsque la construction fait l'objet d'une lecture partitive (voir *infra*). Un troisième cas d'exception s'ajoute à cela, déjà relevé à plusieurs reprises (cf. la Section 3.5.2 et la Section 6.2.1.1 *supra*), à savoir celui où TEN, bien que participant à l'identification du référent, véhicule une valeur affective accrue :

- 54) Il est question de faire reconnaître la paternité d'une fille à l'ancien compagnon d'une femme que les locuteurs connaissent :

a ted' mě zajím-á . když je .. dejme tomu prostě že od tý
et maintenant moi.ACC intéresser-PRS.3SG si être.PRS.3SG MOD de ten.F.GEN.SG

Nikol-y že jo kdyby vona .. dodatečně přida-l-a
NP-F.GEN.SG MOD si.CONJ+AUX.COND.3SG elle.NOM.SG a posteriori ajouter-PTCP-SG.F

otcovstv-í hmm k tomu hajzl-ovi
paternité-ACC.SG MOD à ten.M.DAT.SG salaud-DAT.SG

et maintenant, ce qui m'intéresse... si... admettons que pour Nikola, qu'elle... qu'elle lui fasse reconnaître sa paternité, à cette canaille
ORTOFON v1

- 55) L1 : a .. minul-ej tejd-en jste hrá-l-i ? ..
et dernier-M.ACC.SG semaine-ACC.SG AUX.PST.2PL jouer-PTCP-PL.M

L2 : v sobot-u no ..
en samedi-LOC.SG MOD

L1 : a vyhrá-l-i jste ? ..
et gagner-PTCP-PL.M AUX.PST.2PL

L2 : dva jedna na tom hrozn-ým sněh-u ..
deux un sur ten.M.LOC.SG horrible-M.LOC.SG neige-LOC.SG

L1 : et la semaine dernière, vous avez joué ?

L2 : samedi, ouais...

L1 : et vous avez gagné ?

L2 : 2 à 1, sur cette neige horrible
ORTOFON v1

Aussi bien en 54 qu'en 55, l'emploi de TEN – transposé par nous, dans les deux cas, par CE en français –, représente l'équivalent mémoriel de celui en 18, 19 et 26 (occurrences déictiques situationnelles) et de celui en 38 (occurrence anaphorique). De même qu'en 38, en 54, TEN introduit un « nom de qualité » à fort sémantisme subjectif péjoratif. En 55, le nom introduit par TEN est modifié par l'adjectif *hrozný* (*horrible, affreux*), lui aussi doté d'un fort sémantisme subjectif (cf. Kerbrat-Orecchioni, 2014 : 79-82).

Comme nous l'avons écrit dans la Section 5.2.3.2, 90,49 % des occurrences mémorielles de TEN de l'échantillon de conversations tchèques sont, d'après nous, touchées par la grammaticalisation.

Dans le sous-corpus de Havel, il s'agit de 76,19 % (Section 5.2.4.2). Ces chiffres ont été obtenus suite à l'application des critères susmentionnés, déjà énumérés à la fin de la Section 6.1.4. Les exemples 56 et 57, tirés du sous-corpus de Havel, affichent la même valeur de TEN que 54 et 55, raison pour laquelle ils n'ont pas été considérés comme grammaticalisés :

56) Přirozeně to bylo celý docela jinak, jenomže [*ta sketa*]_i – [ten Mlynárik ze spilky]_i – víte, kterej [to]_i je?

Bien sûr, ça c'était pas du tout passé comme ça, mais [ce petit salaud]_i, [ce Mlynarik]_i de la mise en bouteilles, vous [le]_i voyez ?

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

57) A *ty lesy*!

Et ces forêts ! V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

Remarquons que dans les deux exemples, les traducteurs ont opté pour le démonstratif, choix qui corrobore notre analyse (voir également la Sections 6.2.1.1). En 56, les traducteurs renforcent la valeur subjective du démonstratif via l'emploi péjoratif de l'adjectif *petit* ; en 57, notons la présence de la modalité phrastique exclamative (cf. aussi les exemples 17 et 25 *supra*). Dans l'échantillon de conversations, les occurrences mémorielles de TEN considérées comme non-grammaticalisées sont au nombre de 27, dont 2 relèvent du cas de figure illustré en 54-57 (soit les exemples 54 et 55). Quant au reste (soit 25 occurrences), il s'agit de cumuls de TEN adnominal avec d'autres éléments fonctionnels. Dans le sous-corpus de Havel, le nombre total d'occurrences mémorielles considérées comme non-grammaticalisées s'élève à 10, dont 9 relèvent du cas de figure illustré en 54-57¹¹⁰ et une seule affiche le cumul TAKOVÝ + TEN.

Nous allons à présent aborder **les occurrences mémorielles où TEN se cumule avec un autre élément fonctionnel** afin de tâcher de justifier ce qui nous a amené à les qualifier tantôt de grammaticalisées, tantôt de non-grammaticalisées. Parmi **les occurrences contenant des cumuls et qui n'ont pas été évaluées comme grammaticalisées**, 21 présentent le cumul TAKOVÝ + TEN (dont 20 apparaissent dans notre échantillon de conversations). En ce qui concerne les 5 autres, 4 d'entre elles affichent le cumul NĚJAKÝ + TEN, et une fait apparaître le cumul de TEN avec le numéral JEDEN (UN). Cependant, à y regarder de plus près, nous voyons que deux autres éléments fonctionnels suivent, au sein de cette dernière occurrence, TEN JEDEN après une courte pause, à savoir TAKOVÝ TEN :

58) La locutrice parle du prix d'un demi-litre de myrtilles chez un commerçant :

¹¹⁰ Quatre de ces occurrences apparaissent dans un énoncé à modalité phrastique exclamative (cf. l'exemple 57), 2 contiennent un substantif à sémantique subjective (cf. l'exemple 56), une affiche un adjectif à sémantique subjective et 4 font apparaître à la fois un substantif et un adjectif à sémantique subjective.

pětapadesát kor-un **ten** **jeden .** **takov-ěj** **ten**
cinquante-cinq couronne-GEN.PL ten.M.NOM.SG un.M.NOM.SG tel-M.NOM.SG ten.M.NOM.SG

co v* **půllitr** asi to je ..
? demi-litre.NOM.SG MOD ce être.PRS.3SG

cinquante-cinq couronens [il coûtait], ce... cet espèce de chope de demi-litre ORTOFON v1

À vrai dire, nous avons hésité à analyser cet emploi de TEN comme un emploi « métalinguistique », vu notamment la courte pause entre *jeden* et *takovej*. En effet, il semble qu'outre accomplir un acte de référence, la locutrice soit en quête du bon terme lexical.

De manière générale, la combinaison NEJAKÝ TEN n'est pas, quant à elle, facile à analyser. Nous avons écrit *supra* que le seul cas où TEN puisse y être considéré comme grammaticalisé est celui où la combinaison est susceptible d'une interprétation partitive. Or ce n'est pas ce qui se produit dans les 4 occurrences mémorielles de notre échantillon de conversations :

59) ty mus-í-š vytvořit bakalářsk-ou prác-i kter-á se
toi.NOM devoir-PRS-2SG créer.INF de licence-F.ACC.SG travail-ACC.SG qui.F.NOM.SG REFL

ne-sm-í nebo by oni m-a-jí **nějak-ý**
NEG-pouvoir-PRS-3SG ou AUX.COND.3SG eux.NOM avoir-PRS-3PL quelque-M.ACC.SG

ten **klíč** že oni pozn-a-jí jestli #s to
ten.M.ACC.SG clé-ACC.SG que eux.NOM reconnaître-PRS-3PL si AUX.PST.2SG cela.ACC

stáh-l-a z internet-u nebo okopči-l-a
télécharger-PTCT-SG.F de Internet-GEN.SG ou copier-PTCT-SG.F

*tu dois faire un mémoire de licence qui doit pas... ou alors... ils ont **cette espèce de clé** qui fait qu'ils savent si tu l'as téléchargé depuis Internet ou si tu l'as copié* ORTOFON v1

L'interprétation la plus naturelle de la combinaison dans ces 4 emplois serait, selon nous, celle-ci : *ce X, je ne sais pas trop lequel/dont je n'ai qu'une idée vague* (voir la citation d'Uhlířová, 1992 : 251 *supra*). Aussi croyons-nous que la traduction la plus appropriée de la combinaison est, ici aussi, CE + *genre/type/espèce(s) de*. En revanche, par comparaison avec TAKOVÝ + TEN, la présence de NĚJAKÝ nous semble introduire encore davantage d'approximation et de tâtonnements dans la saisie du référent.

Venons-en maintenant aux **occurrences mémorielles contenant des cumuls et ayant été analysées comme grammaticalisées**. Dans l'échantillon de conversations, ces occurrences sont au nombre de 7. Nous avons d'abord la combinaison TEN + possessif (3 occurrences ; l'exemple 60), suivie de la combinaison numéral + TEN (2 occurrences ; l'exemple 61). Nous identifions ensuite les deux combinaisons suivantes, représentées chacune par une occurrence : TEN + numéral et *kolik* (*combien de*) + *těch* (les exemples 62 et 63) :

- 60) Les locuteurs s'entretiennent d'une connaissance commune dont ils n'apprécient pas le comportement en général, et qui s'est récemment signalée par d'autres excès :

ale vždyť mu to tam lid-i říká-l-i to byl-a
MOD lui.DAT cela.ACC là-bas gens-NOM dire-PST-3PL.M ce être.PST-3SG.F

vod-a na jeho mlejn .. zača-l mlet
eau.NOM.SG sur son.M.ACC.SG moulin.ACC.SG commencer-PST.3SG.M bredouiller.INF

ty **svo-je** **nesmysl-y**
ten.M.ACC.PL son.REFL-M.ACC.PL nonsens-ACC.PL

mais des gens là-bas le lui disaient, c'était comme apporter de l'eau à son moulin... il a juste commencé à débiter ses sornettes (des siennes) ORTOFON v1

- 61) Le locuteur raconte le drame d'un plongeur pratiquant la plongée en apnée (le plongeur, en manque d'oxygène, a de la peine à remonter à la surface) :

v **jedn-é** **té** **soutěž-i** ho táh-l-i
dans un-F.LOC.SG ten.F.LOC.SG compétition-LOC.SG lui.ACC tirer-PST-3PL.M

tam ho taha-l-i dokonce ručně
là-bas lui.ACC tirer-PST-3PL.M même manuellement

dans l'une des compétitions, ils l'ont même... ils ont même dû le sortir à la main ORTOFON v1

- 62) jak **tu** **jedn-u** **noc** jsme tam [...] nějak
comment ten.F.ACC.SG un-F.ACC.SG nuit.ACC.SG AUX.PST.2PL là-bas d'une façon

povídá-l-i já ne-v-í-m dokdy
bavarder-PTCP-PL.M moi.NOM NEG-savoir-PRS-1SG jusque quand

tu vois, la nuit, là, où on a bavardé, je sais plus jusqu'à quelle heure... ORTOFON v1

- 63) teď ne-v-í-m přesně kolik ne-v-í-m přesně **kolik**
maintenant NEG-savoir-PRS-1SG exactement combien NEG-savoir-PRS-1SG exactement combien

těch **řemeslník-ů** už zaplati-l-o
ten.M.GEN.PL artisan-GEN.PL déjà payer-PST-3SG.N

maintenant, je sais pas combien... je sais pas exactement combien des artisans ont déjà réglé ORTOFON v1

Regardons d'abord 60. Nous estimons que l'identification du référent (les sornettes) s'effectue à la fois par la présence de l'argument possessif exprimé par le possessif réflexif SVŮJ¹¹¹ et par la mobilisation d'une connaissance mémorielle partagée par les interlocuteurs : *les sornettes qu'il a tendance à raconter, tu sais / des siennes*. En français, ces deux voies d'identification du référent sont codées, au niveau du signifiant, par le seul possessif *ses*. Nous partons de la position suivante : si TEN était, en tchèque, plus proche d'un démonstratif, il serait plus naturel de rendre l'ensemble de la

¹¹¹ Nous sommes face à un concept relationnel.

combinaison par *ces sornettes à lui*, ce qui ne nous semble pas être ici le cas. Les exemples 61 et 63 ont ceci en commun que la combinaison des deux éléments fonctionnels s’y prête à une lecture partitive : *l’une parmi les compétitions* et *combien parmi les artisans*. Les deux ensembles en question – celui des compétitions et celui des artisans – sont présentés, dans la conversation, comme identifiables dans leur totalité via la mémoire commune des interlocuteurs, d’où l’emploi du défini en français. En 62, TEN JEDEN permet lui aussi d’identifier le référent par la voie mémorielle : *la nuit, tu sais laquelle*¹¹². Remarquons que 62 et 63 contiennent (*já*) *nevím* (*je ne sais pas*), tour dont la modalité épistémique portant sur un événement du passé semble également refléter la nature mémorielle de l’ensemble de l’énoncé. S’y ajoute, en 62, le mot *jak*, dont la présence est forte dans le contexte mémoriel du tchèque parlé informel. Nous explorerons davantage les singularités de ce mot dans le Chapitre 7 (Section 7.2.3).

Dans le corpus de Havel, il est question des 2 occurrences suivantes de TEN + NĚKOLIK (*plusieurs, quelques*) :

- 64) *Žádám vás o to jménem těch několika hezkých chvil, které jsme spolu prožili, a těch několika bláznivých plánů, které jsme spolu už nestačili uskutečnit!*

Je vous le demande au nom de ces instants inoubliables que nous avons vécus ensemble, en souvenir des rêves les plus fous que seul le manque de temps nous a empêchés de réaliser ensemble ! V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

Dans les deux cas, le traducteur a évité de reproduire le cumul en français, bien qu’il eût pu employer la combinaison *les/ces quelques instants/rêves*. Remarquons aussi que le traducteur a employé tantôt CE, tantôt LE, ce qui, une fois de plus, montre la ténuité de la frontière entre les deux déterminants en français¹¹³. Si nous avons décidé de considérer ces occurrences de TEN comme grammaticalisées, c’est que la présence de celui-ci nous semble ici nécessaire : *několika hezkých chvil* et *několika bláznivých plánů* seraient interprétés comme des indéfinis.

Si nous nous focalisons à présent sur **les transpositions, dans les traductions françaises des pièces de Havel, des occurrences mémorielles de TEN adnominal** considérées comme grammaticalisées et de celles qui ne le sont pas, nous constatons qu’au sein de la seconde catégorie (10 occurrences au total), les traducteurs ont recouru au démonstratif dans 4 cas, au possessif dans 2 cas, à un autre moyen grammatical dans un cas et à une reformulation du propos dans 3 cas. Au sein de la première catégorie (32 occurrences au total), nous enregistrons la présence de l’article défini

¹¹² LE/CE + UN + N n’est pas possible en français, contrairement à LE/CE combinés aux autres numéraux cardinaux (LES/CES + deux + N, etc.). *Cette nuit-là* serait également acceptable comme traduction.

¹¹³ Bien qu’il faille reconnaître que dans la seconde occurrence, la présence du défini est favorisée par celle du superlatif, lequel n’apparaît pas en tchèque.

dans 14 cas, celle du démonstratif dans 12 cas, une reformulation dans 3 cas, un autre moyen dans 2 cas et, enfin, le possessif dans un seul cas¹¹⁴:

- 65) Uvědom si, že odpovídáš za **ta děvčata** přímo jejich matkám ;
*Tu es responsable de **ces petites** directement devant leurs mères ;*
Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

En 65 (exemple déjà cité dans la Section 5.2.2.4), les filles en question ne sont identifiables ni à travers la situation d'énonciation ni à travers une mention dans le discours antérieur, mais grâce à une connaissance partagée. Il serait tout à fait possible d'imaginer, en français :

- 66) *Tu es responsable **des petites** directement devant leurs mères ;*

De même, il serait envisageable de remplacer le défini dans la traduction de l'exemple 67 (exemple également cité dans la Section 5.2.2.4) par le démonstratif (exemple 68) :

- 67) Il s'agit d'une discussion qui vient d'avoir lieu en la présence des interlocuteurs, qui en gardent ainsi une mémoire fraîche¹¹⁵ :

Notak ze začátku byl **ten spor** zajímavý, pak se ale trochu zvrhl, že?
*Bon alors, au début **la discussion** était intéressante, mais ensuite elle a plutôt dégénéré, pas vrai ?*
V. Havel, *La fête en plein air*, 1963, trad. par F. Kérel, 1964, InterCorp v10

- 68) *Bon alors, au début **cette discussion** était intéressante, mais ensuite elle a plutôt dégénéré, pas vrai ?*

Présentons ici deux exemples de reformulation du propos par les traducteurs. Nous remarquerons notamment le fait que dans les deux cas, le complexe nominal tchèque est rendu par un tour au verbe conjugué :

- 69) Jste tichej, pracovitej člověk, chodíte každy den, nemáte kolem všeho blbý řeči jako ostatní dělníci, vejplata vám stačí – a **při tom nedostatku lidí**, že jo? / littéralement : *vu le/ce manque de personnes*

*Vous êtes calme et travailleur, vous venez tous les jours, vous ne râlez pas tout le temps comme tous ces autres imbéciles, votre paye vous suffit – alors **vu qu'on manque de main-d'œuvre**, n'est-ce pas ?* V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

- 70) Vybojuju ti **to místo ve skladě** [...] / littéralement : *le/ce poste au magasin*
*Je me battraí jusqu'au bout **pour que tu deviennes magasinier** [...]*

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

¹¹⁴ Ce paragraphe développe les données relatives à la transposition de TEN mémoriel présentées dans la Section 5.2.4.4.

¹¹⁵ Peut-être la fraîcheur de ce souvenir est-elle la raison première du choix du traducteur. Ainsi, nous rejoignons ici ce qui a été dit dans la Section 4.5.3, à savoir qu'en français, le choix du démonstratif reflète, au sein du régime de l'exophore mémorielle, une identification du référent plus coûteuse en effort mental (cf. Chafé, 1994 et Laury, 1997).

Voici enfin deux exemples d'emploi d'un autre moyen grammatical par les traducteurs. Vu la subjectivité produite par le choix des termes lexicaux et par le recours à la modalité exclamative, nous considérons ici la présence de TEN comme non-grammaticalisée, car participant à l'expression de cette subjectivité :

71) *Ta děvka!*

Quelle salope ! Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

72) *Tý srandy*, co tady bylo!

Mais alors, quelle rigolade.

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

Dans les deux cas, TEN, apparaissant dans un énoncé à modalité exclamative, est transposé par le déterminant exclamatif QUEL. En 71, nous sommes face à un énoncé monorhématique (une prédication condensée du type *To je (ale) děvka!* (*C'est une salope !*)), de même qu'en 57 *supra*. Mais c'est notamment la construction en 72, où le complexe nominal *tý srandy* apparaît au génitif, qui mérite une analyse. Cette construction a été relevée et commentée, entre autres, par Trávníček (1951 : 633 et 1210) et par Berger (1993 : 95-96). Le premier parle de « génitif émotionnel » („citový genitiv“), le second, de « propositions averbales avec un TEN adjectival au génitif partitif » („verblose Sätze mit adjektivischem *ten* und dem partitiven Genitiv“). Toutefois, nous sommes d'avis que le terme « partitif » – utilisé par les deux auteurs – n'est pas vraiment pertinent, dans la mesure où il ne s'agit pas de prélever une quantité sur un ensemble prédéfini. En résumant la position de Trávníček (1951 : 633 et 1210), Berger (1993 : 95) écrit ceci :

Das Pronomen *ten* wird als Zeichen für die emotionale Haltung des Sprechers interpretiert, der Genitiv steht laut Trávníček für die Verwunderung über eine große Menge [...], während der ebenfalls mögliche Nominativ nur noch die emotionale Bedeutung habe.

Or s'il est vrai que le génitif ajoute à l'attitude émotionnelle du locuteur le sens d'étonnement face à une quantité élevée du référent, il en va ainsi, d'après nous, parce que de telles constructions feraient apparaître en structure profonde un quantificateur ellipsé, lequel quantificateur appelle le génitif :

73) *Tolik* *srand-y!*

tellement rigolade-GEN.SG

Tellement de rigolade !

Il est également intéressant que ce quantificateur puisse apparaître en même temps que TEN, même si le recours à une proposition à verbe conjugué semble nécessaire pour que cela puisse se produire :

74) Tam ti byl-o *tolik* *tý* *srand-y!*
 là-bas toi.DAT être.PST-3SG.N tellement ten.F.GEN.SG rigolade-GEN.SG
 Mais, alors, il y en avait, de la rigolade !

Cette possibilité de co-occurrence de TEN avec le quantificateur pourrait s'expliquer, selon nous, par le fait que la valeur de TEN se rapprocherait ici de sa fonction « émotionnelle de rappel » telle que celle-ci a été définie par Adamec (1983 ; voir la Section 3.5.3). De sorte qu'il n'interviendrait pas dans le marquage de l'extensité du nom, mais servirait uniquement à signaler que l'information a déjà été mentionnée : *Comme je disais, il y en avait de la rigolade !*

Nous aborderons à présent l'autre grand volet de la seconde partie du présent chapitre, à savoir l'étude de la grammaticalisation de TEN adnominal dans les contextes occupant une position intermédiaire entre les définitives « pragmatique » et « sémantique ».

6.2.2 Les contextes à la charnière entre les définitives « pragmatique » et « sémantique »

Dans cette section, nous nous intéresserons à la présence de TEN adnominal grammaticalisé au sein des contextes constituant la transition entre les définitives « pragmatique » et « sémantique ». Rappelons ici, en guise d'introduction, la signification toute particulière que nous attachons à cet ensemble de contextes, à l'instar des auteurs tels que Carlier et De Mulder (2010), Czardybon (2017), Himmelmann (1997), Hofherr et Zribi-Hertz (2014), Löbner (2011) et Ortmann (2014). Comme nous avons pu le constater dans les Sections 6.1.3.3 et 6.1.3.4, ces chercheurs mettent tour à tour en avant le caractère exclusif de ces contextes à l'égard de l'usage du défini – ils n'admettent pas la présence des démonstratifs forts (non atteints par la perte de leur sémantique déictique) – et l'importance que représente pour un défini en devenir la pénétration au sein de ces contextes, d'où il pourra par la suite s'étendre aux contextes de définitive « sémantique » pure. Il s'agit donc, dans la diachronie de la constitution de l'article défini, d'une étape charnière dont le marquage serait indispensable à la poursuite du processus de grammaticalisation.

Dans la Section 6.1.3.4, nous avons reproduit trois propositions de représentation de ce type de contextes sur une échelle de l'unicité, schéma de la définitive organisé de façon ascendante à partir des contextes de définitive « pragmatique » pure et proposé pour la première fois par Löbner (2011). Nous avons nous-mêmes proposé une de ces trois représentations dans notre étude de 2020, où nous avons apporté quelques modifications à l'échelle proposée par Czardybon (2017) tout en tenant compte des tendances observées lors de l'analyse de nos propres données d'ORTOFON v1. Comme nous estimons que ce schéma a fait ses preuves, nous le prendrons aussi comme point de départ dans la présente section :

deictic sortal nouns (*Can you see that car?*) < anaphoric sortal nouns (*A man came in. This man looked familiar to me*) < sortal nouns with complements establishing uniqueness (*the woman whom I married*) < situational DAAs (*Yesterday, I went to the cinema but the film was a flop*) < relational DAAs (*This novel is great. I know the author personally*) < complex individual concepts (*the biggest rainforest in the world*) < part-whole DAAs (*I have bought a new car recently. The engine is already broken*) < lexical individual and functional nouns (*the Prime Minister, my mother*) < proper names (*Jack*) < personal pronouns (*I, me*) (Dvořák, 2020 : 299)

En suivant globalement ce schéma, nous traiterons donc d'abord le cas des noms sortaux (éventuellement relationnels) rendus uniques par la présence d'une proposition subordonnée relative « établissante », une proposition subordonnée conjonctive pure ou un complément de nature prépositionnelle/adverbiale, puis l'anaphore associative et, enfin, les « concepts individuels complexes ». Au total, notre échantillon de conversations contient 117 occurrences de TEN appartenant à l'ensemble de ces cas de figure (voir le Tableau 5.19 de la Section 5.2.3.2), soit 11,7 % de toutes les occurrences et 25,94 % de toutes les occurrences évaluées comme grammaticalisées. Dans le sous-corpus des pièces de Havel, il s'agit de 13 occurrences (voir le Tableau 5.34 de la Section 5.2.4.2), soit 5,39 % de toutes les occurrences et 19,7 % de toutes les occurrences évaluées comme grammaticalisées.

6.2.2.1 Les noms sortaux ou relationnels avec des compléments établissant l'unicité dénotative

Avec les « concepts individuels complexes », cette catégorie relève du phénomène de la cataphore intra-syntagmatique (cf. la Section 1.1.2). Dans le paradigme de la CTD, les noms – sortaux ou relationnels – constituant la tête de ces complexes nominaux subissent, du fait de la présence du complément rabattant leur extension sur leur extensité, un « changement de niveau zéro » pour produire un concept individuel (cf. Löbner, 2011 : 310).

Nous avons relevé 29 occurrences de ce type dans notre échantillon de conversations et 8 occurrences dans notre sous-corpus de Havel. Le Tableau 6.5 indique les nombres d'occurrences en fonction des différents types de compléments qui établissent l'unicité du nom :

Type de contexte	Nombre d'occurrences	
	L'échantillon de conversations	Le sous-corpus des pièces de Havel
Proposition subordonnée relative « établissante »	23	8
Proposition subordonnée conjonctive pure	4	0
Complément prépositionnel/adverbial	2	0

Tableau 6.5 : Nombre d'occurrences où l'unicité dénotative du nom est instaurée par a) une subordonnée relative « établissante », b) une subordonnée conjonctive pure et c) un complément prépositionnel/adverbial dans l'échantillon de 1 000 occurrences d'ORTOFON v1 (CNT) et dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Nous voyons que c'est au moyen d'une **relative « établissante »** que la plupart des noms concernés sont établis dans leur unicité dénotative : 82,14 % de toutes les occurrences de l'échantillon de conversations et la totalité des occurrences du sous-corpus de Havel. Rappelons que nous empruntons le terme de « relative établissante » à Hawkins (1978 : 131), qui parle de « referent-establishing relative clause ». Selon l'auteur, une relative de ce type « établit un référent défini » (1978 : 140). En souscrivant à la position définie par Donnellan (1966 ; voir notamment la Section 1.2.1), nous préférons remplacer le terme de « référent » par celui de « dénoté », les complexes nominaux au sein desquels le nom est rendu unique sur le plan dénotatif par la présence d'un complément ne donnant pas nécessairement lieu à un acte référentiel, du moins pas au sens où nous l'entendons. Toutes les subordonnées relatives restrictives (déterminatives) ne sont évidemment pas des relatives « établissantes ». Il est à noter que ce statut est dépendant non seulement du sens du nom tête, mais aussi du sens des éléments constituant la subordonnée et, enfin, de facteurs conjoncturels d'ordre extralinguistique. S'agissant de ces derniers, nous voyons, par exemple, que la relative restrictive de l'énoncé en 75 sera uniquement interprétée comme « établissante » au sein d'un contexte social monogame :

75) J'ai rencontré *la femme* que Jean a épousée.

En soi, le nom *femme* est susceptible en français de deux lectures conceptuelles, découlant directement de sa polysémie : 1) être humain du sexe féminin (concept sortal) et 2) épouse (concept fonctionnel). En 75, c'est le premier sens qui est présent (contrairement à *ma femme*). Dans un cadre social monogame, la relative sera naturellement interprétée comme « établissante »¹¹⁶, mais pas dans un cadre social polygame, où un contexte supplémentaire sera indispensable à l'établissement de l'unicité dénotative (*La femme qu'il a récemment épousée, comme tu le sais / celle dont on a parlé l'autre jour en disant qu'ils allaient se marier, etc.*).

¹¹⁶ Il est possible, bien entendu, que Jean ait divorcé pour épouser une autre femme en secondes noces. Mais dans ce cas, il sera automatiquement question de sa femme du moment de l'énonciation, ce référent étant ainsi fixé par l'« index temporel » (Löbner, 2011 : 284 ; voir la Section 6.1.3.4 *supra*).

Pour ce qui est du sens des éléments de la relative, c'est l'exemple suivant, emprunté à Löbner, qui nous permet d'illustrer ce propos :

76) What's wrong with Bill?

- a) *The woman* he went out with last night was nasty to him.
- b) *The woman* was nasty to him.
- c) *The woman* who was from the South was nasty to him.
- d) *The woman* who some sailor dated last night spread nasty rumours about Bill. (Löbner, 1985 : 314 ; modifié)

Löbner relève que seule la variante a) contient une relative « établissante », et ce grâce à l'ancrage situationnel via l'anaphorique *he*, renvoyant à *Bill*, la nature de l'événement « sortir avec quelqu'un », l'ancrage temporel via le circonstant *last night* (aboutissant à la singularité de l'événement, dans la mesure où, d'habitude, l'on ne sort qu'une seule fois par nuit avec quelqu'un) et, enfin, l'objet de l'action « sortir avec quelqu'un », à savoir, ici, une femme :

Now to every event of the sort go-out-with-somebody there is one and only one object figuring in the object role (note that the object may be complex). In our case, the object is said to be a woman, hence a single person. (Löbner, 1985 : 315)

Si l'exemple de notre échantillon d'ORTOFON v1 mentionné dans la Section 3.3.4 et reproduit *infra* n'a pas été considéré comme contenant une relative « établissante », c'est que la relative restrictive qu'il contient ne nous semble pas apporter suffisamment d'éléments d'ancrage contextuel qui permettraient de conclure, de façon convaincante et univoque, à l'unicité du référent, en conséquence de quoi TEN a été ici classé parmi les occurrences mémorielles¹¹⁷ :

77) a *ten* *olivov-ej* *olej ..* *co* *jsme* *kupova-l-i ..*
et ten.M.NOM.SG d'olive-M.NOM.SG huile.NOM.SG REL AUX.PST.1PL acheter-PTCP-PL.M

tak ten je taky už podražen-ej
alors celui-ci.M.NOM.SG être.PRS.3SG aussi déjà plus cher-M.NOM.SG

et *cette/l'huile d'olive... qu'on achetait...*, alors elle aussi, elle coûte déjà plus cher
ORTOFON v1

En effet, l'action d'acheter, exprimée par le verbe au prétérit, n'est ni restreinte à un ou plusieurs événements ponctuels, ni délimitée dans sa durée. S'ajoute à cela la nature du référent (un type/une marque d'huile d'olive), laquelle nature ne tend pas forcément à la singularité, car l'on peut acheter plusieurs types/marques d'huile d'olive lors d'une période de temps. Il serait, par conséquent,

¹¹⁷ Plus exactement et comme nous l'avons écrit dans la Section 3.3.4, son fonctionnement est à la fois mémoriel (renvoi à un souvenir partagé) et cataphorique (renvoi à la subordonnée relative).

impossible d'exclure l'interprétation selon laquelle l'huile en question ne serait qu'une parmi plusieurs¹¹⁸. En revanche, l'exemple suivant, également introduit dans la Section 3.3.4, lui, contient une relative « établissante » :

78) už to	není	ta	<i>Srí Lanka-a</i>			
dějà ce	être.PRS.3SG.NEG	ten.F.NOM.SG	Sri Lanka-NOM.SG			
co	to	byl-a	třeba	před pět-i	rok-ama	jo
REL	ce	être.PST-3SG.F	par exemple	avant cinq-INS	an-INS.PL	MOD

ce n'est plus le Sri Lanka que c'était il y a par exemple cinq ans, tu vois ORTOFON v1

Nous sommes ici face à la réalité appelée par Kleiber (1991a : 97) le « fractionnement du porteur du nom propre » : le référent du nom propre est considéré uniquement sous un aspect de son existence (aspect temporel ou autre ; d'autres exemples : *le Paris de la Libération*, *le Camus de La Peste*, etc.).

Dans la Section 3.3.3, nous avons abordé l'existence de l'emploi dit « **déterminatif** » (cf. Diessel, 1999 : 108) de TEN adnominal tchèque. Nous avons illustré ce phénomène en recourant à des exemples empruntés à Ziková et Volín (2012), dont l'un est ici reproduit :

79) Obslouž-í	jen	ty	<i>zákazník-y</i> ,	<i>kter-í</i>	<i>potřeb-uj-í</i>
servir-PRS.3SG	seulement	ten.M.ACC.PL	client-ACC.PL	qui-M.ACC.PL	avoir besoin-PRS-3PL
<i>lék-y</i>		<i>bez</i>	<i>předpis-u.</i>		
médicament-ACC.PL		sans	ordonnance-GEN.SG		

Il sert uniquement les clients qui viennent chercher des médicaments sans ordonnance.
(Ziková et Volín, 2012 : 36)

Nous avons écrit au même endroit que ces constructions ont ceci de particulier que le rôle de la relative s'y limite à l'expression de l'appartenance du dénoté à une certaine classe, d'où le terme de « relative classificatoire », employé par les deux auteurs. Clairement, en 79, le nom n'est pas employé dans un emploi référentiel, mais « attributif » (Donnellan, 1966) : *les clients qui viennent chercher des médicaments sans ordonnance, peu importe qui ils sont précisément*. Il peut ainsi être question d'emplois génériques comme en 79 et en 80, mais pas nécessairement (cf. l'exemple 81) :

80) A nejvýznamnějšími epochami v dějinách jsou v důsledku toho <i>ty epochy</i> ,	<i>které se dovedou opřít</i>
<i>o střední vrstvy a položit do jejich rukou své ideály [...]</i>	

¹¹⁸ L'ambiguïté serait levée, par exemple, si l'on ajoutait un circonstant temporel : *L'huile qu'on a achetée jeudi dernier...* Admettons cependant qu'il n'est pas toujours aisé de trancher entre l'interprétation « établissante » ou non d'une relative restrictive, difficulté qui découle des difficultés que l'on rencontre lorsqu'on tente de déterminer la quantité de contexte nécessaire, d'un cas à l'autre, à l'instauration de l'unicité sémantique d'un nom (cf. Stanley et Szabó, 2000 : 220)..

*En conséquence de quoi les plus grandes époques de l'histoire sont **celles** qui savent s'appuyer sur les classes moyennes et leur remettre leurs idéaux en main propre [...] / littéralement : les époques qui...* V. Havel, *La fête en plein air*, 1963, trad. par F. Kérel, 1964, InterCorp v10 (FOR)

81) tam	v	těch	zem-ích	<u>kde</u>	<u>už</u>	<u>je</u>	<u>eur-o</u>
là-bas	dans	ten.F.LOC.PL	pays-LOC.PL	où	déjà	être.PRS.3SG	euro-NOM.SG
<i>là-bas, dans les pays où ils ont déjà l'euro...</i>				ORTOFON v1 (INF)			

En 81, la relative permet de circonscrire une catégorie particulière de pays, à savoir de ceux qui utilisent l'euro comme devise. Compte-tenu du fait qu'il s'agit d'un ensemble a priori bien délimité et fermé (bien qu'il ne soit pas figé une fois pour toutes, mais susceptible d'évoluer dans le temps), une interprétation générique n'est pas ici opportune. Pour autant, est-ce que le locuteur accomplit un acte de référence au moyen de cet énoncé ? Rien n'est moins sûr, car si l'énoncé peut être mobilisé pour référer aux états membres de la zone euro au moment de l'énonciation (soit en novembre 2012), l'interprétation pragmatiquement beaucoup plus plausible est celle en vertu de laquelle le locuteur ne se soucie guère de l'identité de ces pays (il serait probablement incapable d'en faire la liste exhaustive) et ne cherche en vérité qu'à opérer une délimitation catégorielle : *tous les pays utilisant l'euro, quels qu'ils soient / peu importe lesquels*. Au total, nous sommes d'avis que 10 parmi les relatives « établissantes » sont interprétables comme « classificatoires » dans notre échantillon d'ORTOFON v1 et 2 dans notre sous-corpus de Havel. Celle en 81 en fait très probablement partie.

Au sein des relatives « établissantes », Ortmann (2014 : 311) opère la distinction entre celles qu'il dit « dépendantes du contexte » et celles qu'il dit « indépendantes du contexte ». Alors que les premières s'ancrent dans le contexte d'énonciation et en appellent à des connaissances d'ordre privé, dans les secondes, « [...] the NPs [...] exhibit unique reference independently of the speech situation [...] ». En 76a *supra*, la relative est dépendante du contexte, car elle nécessite la capacité à identifier l'individu Bill ainsi qu'un ancrage situationnel permettant de déterminer la référence du circonstant *last night*. Il en va similairement en 78, où le circonstant temporel *il y a cinq ans*, déictique, s'interprète par rapport à l'énonciation. À l'inverse, les relatives dans les exemples 79, 80 et 81 *supra* sont indépendantes du contexte. Nous reviendrons à cette différence, qui semble peser sur la présence ou l'absence de TEN en tchèque¹¹⁹.

¹¹⁹ Précisons que l'(in)dépendance de la relative « établissante » du contexte d'énonciation n'a a priori rien à voir avec sa nature « classificatoire ». Si, dans nos exemples, toutes les relatives « établissantes » contextuellement indépendantes semblent être « classificatoires », contrairement aux relatives « établissantes » contextuellement dépendantes, qui ne le sont pas, une relative « établissante » peut tout aussi bien être dépendante du contexte et « classificatoire » (*L'équipe qui gagnera ce match gagnera la finale*) qu'elle peut être indépendante du contexte et recevoir une lecture (a priori) « non-classificatoire » (*J'ai rencontré l'homme qui est prétendant au trône de France*).

Venons-en à présent aux occurrences où le nom est rendu unique soit par la présence d'**une subordonnée conjonctive pure** (cf. l'exemple 82) soit par celle d'**un complément prépositionnel/adverbial** (cf. l'exemple 83) :

- 82) ale on samozřejmě m-á potom **ten** *psychick-ej*
 mais lui.NOM évidemment avoir-PRS.3SG après ten.M.ACC.SG mental-M.ACC.SG
- problém** že jako se znova přípravu-je kdyby
 problème.ACC.SG that MOD REFL de nouveau préparer-PRS.3SG si.CONJ+AUX.COND.3SG
- byl náhod-ou dalš-í hlad
 être. PTCP.SG.M par hasard autre-M.NOM.SG famine.NOM.SG
- mais après, lui, le problème mental [qu'il a], c'est qu'il continue à se préparer au cas où il y aurait une autre famine*
 ORTOFON v1 (INF)
- 83) no a **ten** *tejden* **předtím** bych jakoby
 MOD et ten.M.ACC.SG semaine.ACC.SG avant AUX.COND.1SG MOD
- já bych je-l-a
 moi.NOM AUX.COND.1SG partir-PTCP-SG.F
- ouais, et la semaine d'avant, moi, je... genre, je partirais*
 ORTOFON v1

L'énoncé en 82 s'inscrit dans un récit traitant d'un personnage des romans de J. London qui a développé un mécanisme de survie consistant à amasser de la nourriture en périodes d'abondance. Le nom *problém* (*problème*) est rendu unique par une subordonnée conjonctive pure introduite par la préposition *že* (*que*). En 83, l'unicité est établie par l'expression adverbiale *předtím* (*d'avant*)¹²⁰.

Qu'en est-il de **la traduction de TEN** dans les 8 occurrences du sous-corpus de Havel contenant une relative « établissante » ? Nous constatons que TEN est rendu par l'article défini 3 fois, par un autre moyen grammatical 2 fois (dont l'exemple 80), par le déterminant démonstratif 2 fois ; une fois, il n'est pas transposé. Le recours au démonstratif est compréhensible au vu de l'ambiguïté potentielle de la valeur de TEN, pouvant parfois être interprété non pas comme un pur support structurel d'une relative « établissante », et donc comme un proto-article, mais comme un véritable démonstratif, investi d'une référence allant au-delà du complexe nominal (voir à cet égard la Section 3.3.3) :

- 84) Ale nejen to: protože jsi to byl ty, kdo ho do tohoto světa uvrhl, kdo mu ho nabídl k užívání a kdo ho v něm orientuje, začínáš najednou pocít'oval daleko větší odpovědnost i za **ten svět**, který toho kluka obklopuje [...]

Oui, mais j'aimerais ajouter autre chose. Vu que c'est toi qui as mis cet enfant au monde, qui, pour ainsi dire, met la terre à sa disposition, en lui donnant un enseignement qui lui permette de

¹²⁰ Étymologiquement, *předtím* est issu de la jonction entre la préposition *před* (*avant*) et la forme neutre du pronom TEN.

s'y orienter, tu te découvres d'un jour à l'autre une responsabilité accrue pour ce monde, qui entoure ton enfant. V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

En 84, le complexe nominal permet une lecture référentielle situationnelle ou anaphorique, lecture qui s'accompagne ici d'une interprétation appositive de la relative (en français, cette dernière est explicitée dans le texte par la présence de la virgule).

Enfin, la question du caractère obligatoire de la présence de TEN dans ces structures demande à être posée. Après avoir étudié tous les exemples concernés, nous nous sentons contraint de conclure que si la variante avec TEN est, d'un cas à l'autre, plus ou moins ressentie comme préférable à celle sans TEN, nous ne saurions aller jusqu'à dire que TEN est devenu obligatoire. Il est d'autant plus difficile de trancher cette question que le jugement doit tenir compte de la dichotomie entre le registre (oral) informel et le registre (écrit) formel¹²¹. Comme l'a constaté Czardybon au sujet de TEN introduisant un nom modifié par une relative « établissante » :

Ten in Czech seems to be optional with NPs modified by context-independent relative clauses [...] With context-dependent relative clauses, *ten* is preferred [...] or there is interspeaker variation between preferred and optional use [...] Furthermore, my informants mention that in those cases in which *ten* is optional, the presence or absence is dependent on the difference between spoken and written Czech. It rather shows up in spoken Czech while it rather does not in written Czech. (2017 : 88)

Nous pouvons en partie appuyer cette conclusion tout en partant à la fois de notre propre sentiment de locuteur natif et des jugements de nos informateurs. Ainsi, en 79, 80 et 81, l'effacement de TEN est ressenti par nous comme plus acceptable qu'en 78 et qu'en 85 *infra*, sachant que cet effacement nous semble d'autant plus facile si l'énoncé s'inscrit dans le registre formel (ce qui est manifestement le cas en 79 et 80). Quant à 78 et 85 *infra*, la suppression de TEN produirait un énoncé moins naturel, y compris au sein du registre formel¹²².

Les réponses de nos informateurs reflètent un ressenti semblable au nôtre. Outre les énoncés 80, 81 et 82, nous avons soumis à leur jugement l'énoncé suivant :

85) Byl	tu	ten	kluk,	<u>co</u> <u>s nim</u>
être.PST.3SG.M	ici	ten.M.NOM.SG	garçon.NOM.SG	REL avec lui.INS
<u>chod-í</u>	<u>naše</u>	<u>Věr-a.</u>		
sortir-PRS.3SG	nôtre.F.NOM.SG	NP-F.NOM.SG		

¹²¹ L'appartenance à l'un (INF) ou à l'autre (FOR) registre a été marquée pour les exemples 80, 81, 82 et 85, soit ceux qui ont été soumis au jugement de nos informateurs.

¹²² L'exemple 84 se situe un peu à l'écart de ces considérations, car, d'abord, il n'est pas sûr que nous soyons ici face à une relative « établissante » (voir *supra*), et, ensuite, si tant est que nous le soyons, la relative peut également s'interpréter comme ayant une fonction « classificatoire ».

*Il y a **le gars** que notre Věra sort avec lui qui est passé.* (INF)

Dans tous ces énoncés, nous leur avons donné pour consigne de choisir entre la formulation avec TEN et la formulation sans TEN. Le Tableau 6.6 renseigne, pour chaque exemple, sur le nombre de personnes ayant choisi la variante avec TEN :

Exemple (n°)	Nombre d'informateurs ayant choisi TEN
80 (« les époques »)	6
81 (« les pays »)	6
82 (« le problème mental »)	3
85 (« le gars »)	10

Tableau 6.6 : Nombre de personnes ayant choisi d'employer TEN avec les exemples 80, 81, 82 et 85 (sur 10 informateurs natifs)

D'abord, la version avec TEN a été choisie à l'unanimité dans le cas de l'exemple 85, contenant une relative « établissante » dépendante du contexte et s'inscrivant ostensiblement dans le registre informel. Une informatrice nous a par ailleurs fait savoir que pour elle, l'usage de TEN devient beaucoup plus naturel en 85 lorsque l'interlocuteur est au courant de l'existence du garçon. Cette remarque tout à fait passionnante prouve combien le sujet est en vérité complexe et mériterait une étude beaucoup plus approfondie. Dans nos futures recherches, nous ne manquerons pas d'exploiter ces pistes. Pour les exemples 80 et 81, où le nom est modifié par une relative « établissante » indépendante du contexte, l'opinion des informateurs est partagée : six personnes préfèrent la variante avec TEN. À l'opposé de ce que constate Czardybon, la différence de registre ne semble donc pas avoir d'incidence sur le choix de TEN (80 est formel, alors que 81 penche plutôt vers le registre informel). Enfin, le score bas de l'exemple 82 nous surprend, dans la mesure où nous nous attendions à ce que la majorité préfère ici la présence de TEN (laquelle présence correspond à la fois à la version de l'énoncé tel qu'il apparaît dans ORTOFON v1 et à notre propre sentiment de la langue).

Pour conclure, l'**hypothèse n°4** se voit en partie confirmée, mais en partie seulement. S'il semble y avoir une forte préférence pour TEN dans les occurrences où l'unicité du nom découle de la présence d'une relative « établissante » dépendante du contexte d'énonciation, cette préférence se fait plus discrète lorsque l'unicité est atteinte par une relative « établissante » indépendante du contexte d'énonciation. Enfin, le bilan incite à la prudence lorsque l'unicité est établie par une conjonctive pure.

6.2.2.2 L'anaphore associative

Rappelons dans un premier temps la définition de l'anaphore associative :

Associative anaphora involves a hidden link or anchor which has to be introduced earlier. The expressions are interpreted as though the link or anchor were explicit and anaphoric. (Löbner, 1998 :1)

C'est donc la récupération du lien contextuel en question qui permet d'interpréter l'ensemble du complexe nominal comme défini sur le plan dénotatif. Dans le paradigme de la théorie CTD, le nom constituant la tête d'un tel complexe est analysé comme un concept fonctionnel (implicitement) saturé par l'argument possessif qu'est le lien contextuel restauré¹²³. Dans la Section 2.2.3, nous avons introduit une typologie faisant la distinction entre trois types d'anaphore associative, soit le type « partie-tout », le type « relationnel » et le type « situationnel ». Rappelons ici que la première distinction – celle entre les types partie-tout et relationnel – fut proposée par Schwarz dans sa thèse de doctorat (2009).

Dans ce travail, où il traite de la définitude et de ses expressions en allemand, l'auteur distingue entre deux types de définitude : la définitude basée sur l'unicité et celle qui est basée sur l'« anaphoricité ». Bien que le cadre théorique de ce travail ne se réfère jamais aux travaux de l'approche de la théorie CTD, il existe des similitudes frappantes entre la distinction de Schwarz et celle de Löbner, dans la mesure où le premier type de définitude (celui basé sur la notion d'unicité) correspond globalement à ce que Löbner appelle la définitude « sémantique »¹²⁴, tandis que le second correspond globalement à la définitude « pragmatique »¹²⁵. Dans la Section 6.1.3.4, nous avons vu que certaines langues disposent de deux paradigmes d'article défini, dont la distribution obéit à cette différence. Aussi Schwarz relève-t-il qu'en allemand standard, seuls les contextes où la définitude procède de l'unicité permettent l'amalgame du défini certaines prépositions (cf. *zu der* → *zur*). L'anaphore associative – appelée, à l'instar de Clark (1975), « bridging » – présente, selon Schwarz, la particularité suivante : lorsque l'entité dénotée par la tête nominale constitue une partie de l'entité dénotée par l'argument possessif récupéré à partir du contexte, c'est la variante amalgamée de l'article qui est préférée (voire la seule possible), contrairement aux cas où l'entité dénotée par le nom entre dans une relation avec l'entité de l'argument, mais n'en constitue pas pour autant une partie. Dans ces derniers cas, c'est la variante non-amalgamée qui est préférée (voir Schwarz, 2009 : 34-40 ; ce constat s'appuie sur un questionnaire soumis à 28 informateurs natifs). Il s'ensuivrait donc

¹²³ Nous sommes face à un « changement de 2^{er} niveau », l'argument possessif étant fourni par le contexte (Löbner, 2011 : 312).

¹²⁴ Pour Schwarz, le terme d'« unicité » est ainsi restreint dans son extension, correspondant à ce que Löbner entend par la notion d'« Eindeutigkeit » (voir la Section 2.1.3.4).

¹²⁵ Le choix du terme d'« anaphoricité » est sans doute dû au fait que Schwarz limite ses analyses au seul domaine endophorique de la définitude « pragmatique ».

logiquement que l'anaphore associative de type partie-tout (cf. l'exemple 86) serait plus proche du pôle de la définitude « sémantique » que celle de type relationnel (cf. l'exemple 87) :

- 86) Der Kühlschrank war so groß, dass der Kürbis problemlos **im Gemüsefach** untergebracht werden konnte. Et non pas : **in dem** Gemüsefach

Le réfrigérateur était si grand que la citrouille pouvait sans aucun problème être rangée dans le compartiment à légumes. (Schwarz, 2009 : 34 ; notre traduction)

- 87) Das Theaterstück missfiel dem Kritiker so sehr, dass er in seiner Besprechung kein gutes Haar **an dem Autor** ließ. Plutôt que : **am** Autor

La pièce a déplu au critique au point que celui-ci a démolit l'auteur dans son entretien. (Schwarz, 2009 : 34 ; notre traduction)

C'est Czardybon (2017 : 66) qui, se rapportant à un échange privé avec Löbner, introduit dans sa thèse de doctorat un troisième type d'anaphore associative, soit le type dit « situationnel ». La particularité de ce type est la suivante : alors que les deux autres types d'anaphore associative contiennent un nom intrinsèquement relationnel, ce type affiche, quant à lui, un nom qui n'est pas intrinsèquement relationnel, mais qui est seulement modifié de telle sorte à dégager un concept à sémantique relationnelle¹²⁶ :

Both types of DAAs have in common that the head of the DAA is represented by an inherently relational noun. According to Löbner (p.c.), there are DAAs with a [+R] noun as head [...] and 'situational DAAs' with a [-R] noun as head.

En voici deux exemples, dont le premier a déjà été en partie cité dans les Sections 2.2.3 et 3.3.2 :

- 88) Une fois arrivé dans la cuisine, j'ai remarqué que **le réfrigérateur** était cassé. (Czardybon, 2017 : 66 ; notre traduction)

- 89) Après être allés au cinéma, ils ont encore discuté **du film**. Ils avaient aimé **les acteurs**. (Czardybon, 2017 : 90 ; notre traduction)

Il nous semble important de relever ici un autre trait par lequel ce type d'anaphore associative diffère des deux autres. Alors que dans les deux autres types, l'argument possessif saturant le nom fonctionnel semble globalement restituable tel quel, au sein de l'anaphore associative situationnelle, il semblerait que ce soit plutôt l'ensemble de la situation décrite qui en vient à constituer cet argument, et ce via une opération de synthèse obéissant au mécanisme de l'anaphore résomptive. Ce mécanisme apparaît clairement dans la première partie de 89 :

¹²⁶ Ainsi, un nom sortait subit un « changement sémantique de 1^{er} niveau » pour déboucher sur un concept fonctionnel (voir la Section 6.1.3.4).

après être allé au cinéma → cette sortie au cinéma → le film = le film de cette sortie au cinéma

Cependant, cette analyse ne fonctionne pas toujours : *les acteurs* en 89 sont sans doute *les acteurs du film*, et, en 88, *le réfrigérateur* est plus probablement *le réfrigérateur de la cuisine* que *le réfrigérateur du passage dans la cuisine*¹²⁷.

Qu'en est-il de la présence de TEN en contexte d'anaphore associative en tchèque (parlé informel) ? Avant de procéder à l'analyse de nos données, nous rapportons d'abord ce que constate Czardybon (2017 : 90) dans sa comparaison des langues slaves :

I observe that with some situational DAAs *ten* is regarded as optional – as in (81)¹²⁸. For (81), it is reported that, especially in spoken colloquial Czech, *ten* can occur while it is avoided in written language. The Czech data shows that situational DAAs can be considered as a separate class of DAAs since only here do we find the possibility of the presence of *ten* in contrast to part-whole and relational DAAs with which *ten* is absent in Czech.

À en croire l'auteur, TEN serait uniquement autorisé avec le type situationnel. Nous tâcherons ici de démontrer que la réalité est plus subtile, car les deux autres types – notamment le type relationnel – semblent également l'admettre (bien que dans une moindre mesure).

En analysant soigneusement toutes nos occurrences anaphoriques associatives, nous sommes arrivé au bilan suivant : notre échantillon de conversations contient 34 occurrences appartenant au type situationnel, 18 occurrences appartenant au type relationnel et 2 occurrences relevant du type partie-tout¹²⁹. La seule occurrence du sous-corpus de Havel relève du type relationnel. Réitérons néanmoins ce que nous avons déjà écrit dans la Section 5.1.4, à savoir le statut équivoque d'une partie des occurrences et les hésitations qu'a suscité leur classement parmi les différents (sous-)types d'emploi. Les résultats doivent ainsi être considérés avec une once de scepticisme.

Parmi nos occurrences anaphoriques associatives introduites par TEN, **le type situationnel** est le plus représenté des trois¹³⁰ :

¹²⁷ Il est également à noter que dans certaines anaphores associatives des deux autres types, il arrive que l'argument ne soit pas explicitement présent dans le contexte, et qu'il faille le « chercher » par inférence, en partant d'une autre information qui, elle, en fait partie. Cf. *Demain, je vais lui demander ce qu'il en pense. Dès que j'ai sa réponse, je te tiens au courant. → la réponse à la question que je vais lui poser au moment de lui demander ce qu'il en pense / de l'action de demander ce qu'il en pense*. Ainsi l'argument à restituer (*la question*) est-il lui-même implicite, et doit être inféré à partir du sens de *demander ce qu'il en pense*. Alors que le lien réponse-question est de nature relationnelle, celui entre l'action de demander et la question est de nature situationnelle.

¹²⁸ Dans notre numérotation, il s'agit de l'exemple 89.

¹²⁹ Nous verrons cependant que le profil de l'une de ces occurrences n'est pas univoque.

¹³⁰ La différence est statistiquement significative : $L = 9,6396$ pour la différence entre les fréquences des types situationnel et relationnel et $L = 49,1902$ pour celle entre les fréquences des types situationnel et parti-tout ; la valeur critique = 3,841 (degré de liberté = 1 ; $\alpha = 0,05$). De nouveau, il faut ici reconnaître les limites de notre approche, liées au fait que le nombre d'occurrences introduites par TEN n'est pas comparé au nombre total d'occurrences de l'emploi en question ; ainsi, nos données nous autorisent à dire que le type situationnel est le plus représenté au sein de l'ensemble des

90) L1: nahoře nad tím byl-y tak obrovitánsk-ý pavučin-y
 en haut au-dessus cela.INS être.PST-3PL.F si géant-F.NOM.PL toile d'araignée-NOM.PL

L2: jako že se tam rok-y ne-vysáva-l-o jo ?
 MOD REFL là-bas an-ACC.PL NEG-passer l'aspirateur-PST-3SG.N MOD

L1: ne on ten týpek říká-l že *ti*
 non lui.NOM ten.M.NOM.SG gars.NOM.SG dire-PST.3SG.M que ten.M.NOM.PL

pavou-ci to uděla-l-i za hodin-u takov-ou
 araignée-NOM.PL cela.ACC faire-PST-3PL.M en heure-ACC.SG tel-F.ACC.SG

obrovsk-ou pavučin-u
 géant-F.ACC.SG toile d'araignée-F.ACC.SG

L1 : là-haut, juste au-dessus, il y avait des toiles tellement géantes...

L2 : genre on y avait pas passé l'aspirateur pendant des années, c'est ça ?

L1 : non, le gars, lui, il a dit que les araignées avaient fait ça en l'espace d'une heure, une toile géante comme ça ORTOFON v1

91) L1: jak . jak vlastně . děl-á-š to aby ses
 comment comment MOD faire-PRS-2SG cela.ACC que.CONJ+AUX.COND REFL+2SG

proces uzdraven-í . jak to vlastně . jak to děl-á-š?
 processus.NOM.SG guérison-GEN.SG comment cela.ACC MOD comment cela.ACC faire-PRS-2SG

L2: no zakukl-i-m se . a ček-á-m až . z
 MOD chrysalider-PRS-1SG REFL et attendre-PRS-1SG que de

té *kukl-y* .. se vylíhn-e zdrav-ý člověk
 ten.F.GEN.SG chrysalide-GEN.SG REFL éclore-FUT.3SG sain-M.NOM.SG homme.NOM.SG

L1: comment... genre comment tu fais, concrètement, pour... processus de guérison... comment tu... comment tu fais, en fait?

L2 : ben, je me chrysalide et j'attends qu'un homme en bonne santé sorte de la chrysalide ORTOFON v1

92) voni tam děl-a-jí mátov-ej čaj ale strašlivě
 eux.NOM là-bas faire-PRS-3PL de menthe-M.ACC.SG thé.ACC.SG mais terriblement

sladk-ej .. vopravdu *ta* *mát-a* je tam hrozně cítit
 fort-M.ACC.SG vraiment ten.F.NOM.SG menthe.NOM.SG être.PRS.3SG là-bas terriblement sentir.INF

eux, ce qu'il préparent là-bas, c'est du thé à la menthe ; mais alors ils le font vachement sucré... vraiment, la menthe, on la sent super fort ORTOFON v1

En 90, il est question d'une chambre remplie de toiles d'araignée et des araignées qui les ont tissées. L'introduction des premières dans le discours permet aussitôt l'introduction des secondes via un complexe nominal défini : ce sont, littéralement, *les araignées des toiles en question*. En 91, il en va de même de *la chrysalide*, grâce au lien situationnel entre l'action de se chrysalider et son produit,

occurrences avec TEN, mais ne nous renseignent pas sur la proportion des occurrences avec TEN parmi toutes les occurrences anaphoriques associatives d'un type donné.

une chrysalide. Enfin, l'analyse de 92, s'inscrivant dans le récit d'un séjour au Maroc, est peut-être plus délicate, notamment en français, où le SN *la menthe* figure déjà dans le contexte antérieur. Quant à la version tchèque, il est évident que *máta* ne peut constituer une reprise anaphorique directe de *mátovej*, car ce dernier est un adjectif relationnel, équivalent littéral du SN prépositionnel français *de menthe*. Mais en français, il est à peine plus réaliste de supposer un lien anaphorique direct entre les deux occurrences du SN *la menthe*, car au sein de *thé à la menthe*, le SN, ne semble pas posséder une référence¹³¹. Dans les deux cas, il s'agirait donc de *la menthe du thé à la menthe*.

Dans notre Questionnaire n°3, nous avons soumis au jugement de nos informateurs la première partie de l'exemple 89, l'exemple 90 ainsi que l'exemple suivant :

- 93) Když jsem veše-l do toho krám-u, tak se mě
 quand AUX.PST.1SG entrer-PTCP-SG.M dans ten.M.GEN.SG magasin-GEN.SG alors REFL moi.ACC
 (ta) prodavačk-a hned pta-l-a, co chc-i.
 ten.F.NOM.SG vendeuse-NOM.SG tout de suite demander-PST-3SG.F que.ACC vouloir-PRS-1SG

Quand je suis entré dans le magasin, la vendeuse m'a tout de suite demandé ce que je cherchais.

Les informateurs ont reçu pour consigne de choisir entre la version où les noms *film*, *pavouci* et *prodavačka* n'étaient pas introduits par TEN et celle où ils l'étaient. Huit informateurs ont choisi la variante avec TEN en 89 et neuf l'ont choisi en 90. En revanche, seule une personne a opté pour la version avec TEN dans l'exemple 93. Cela pourrait s'expliquer par le fait que le nom *prodavačka* (qui n'est pas relationnel de manière inhérente) pourrait ici se prêter à une lecture indéfinie, dans la mesure où il peut facilement y avoir plusieurs vendeurs et vendeuses dans un magasin¹³².

Venons-en maintenant au **type relationnel** :

- 94) já jsem jí říka-l-a že to už m-á
 moi.NOM AUX.PST.1SG elle.DAT.SG dire-PTCP-SG.F que cela.ACC déjà avoir-PRS.3SG
 úplně to sam-é když půjd-e studovat bakalář-e ..
 complètement ten.N.ACC.SG même.N.ACC.SG si aller-FUT.3SG étudier.INF licence-ACC.SG
 a bud-e to mít takov-ou jakoby větš-í váh-u
 and AUX.FUT-3SG cela.NOM avoir.INF tel-F.ACC.SG MOD plus grand-F.ACC.SG poids-ACC.SG

¹³¹ Sémantiquement, l'ensemble à *la menthe* se rapproche d'un adjectif. Une analyse limite consisterait, selon nous, à poser que le SN *la menthe* est ici doté d'une référence générique.

¹³² Nous serions ainsi face à une anaphore associative indéfinie (cf. Czardybon, 2017 : 66) : le nom reste interprété en rapport avec le contexte précédent, mais n'est pas employé avec l'unicité dénotative. Curieusement, cette même interprétation est possible en 90 (*des araignées*), alors que le choix de TEN y a fait la quasi-unanimité. Une personne nous a fait savoir que pour elle, la présence de TEN implique nécessairement une lecture mémorielle : *la vendeuse, tu sais laquelle, celle dont on a déjà parlé*.

než ta	vošk-a	[...]	přítom ta	náročnost
que ten.F.NOM.SG	école professionnelle-NOM.SG		pourtant ten.F.NOM.S	exigence.NOM.SG
je	stejn-á	si	mysl-í-m	
être.PRS.3SG	pareil-F.NOM.SG	REFL	penser-PRS-1.SG	

*du coup, je lui ai dit qu'en fin de compte, ça revient au même si elle fait une licence... genre ça aura plus de poids que l'école professionnelle [...] alors que **les exigences** sont les mêmes, je crois*

ORTOFON v1

95) ale já m-á-m můj .. telefon v chodb-ě ..
 mais moi.NOM avoir-PRS-1SG mien.M.ACC.SG téléphone-ACC.SG dans couloir-LOC.SG

teda pevn-ou	link-u	..	tak na tu	kdybyste
MOD fixe-F.ACC.SG	ligne-ACC.SG		alors sur celui-ci.F.ACC.SG	si.CONJ+AUX.COND.2PL

zavola-l	to	vám	řek-n-u	to	číslo
appeler-PTC.SG.M	celui-ci.N.ACC.SG	vous.DAT	dire-FUT-1SG	ten.N.ACC.SG	numéro-ACC.SG

*mais moi, mon téléphone, il est dans le couloir... c'est-à-dire mon fixe... alors est-ce que c'est là que vous pourriez m'appeler ? **Le numéro**, je vais vous le dire*

ORTOFON v1

96) L1: sam-á modřin-a je [...] tam konečně si vyčt-e .. že
 tout-F.NOM.SG bleu-NOM.SG être.PRS.3SG là-bas finalement REFL lire-FUT.3S que

to	je [...]	z	<u>těch</u>	<u>lék-ů</u>
cela.NOM	être.PRS.3SG	de	ten.M.GEN.PL	médicament-GEN.PL

L2: no to já si vždycky čt-u **ty** **nežádouc-í** **účink-y**
 MOD moi.NOM REFL toujours lire-PRS.1SG ten.M.ACC.PL secondaire-M.ACC.PL effet-ACC.PL

L1 : elle est toute couverte de bleus... là-bas, elle pourra finalement lire que c'est à cause des médocs

*L2 : ben moi, je lis toujours **les effets secondaires***

ORTOFON v1

Commentons. En 94, la locutrice parle de l'une de ses amies, laquelle hésite entre intégrer une formation de licence et une formation dans une école professionnelle. C'est ainsi que la locutrice en vient à comparer les exigences des deux formations. Contrairement aux noms *araignée*, *chrysalide*, *menthe* et *vendeuse* en 90, 91, 92 et 93, *náročnost* (*exigences*) est un nom à nature relationnelle inhérente : l'on parle d'habitude des exigences de quelque chose ; ici, en l'occurrence, *des exigences d'une [formation de] licence*. Le mécanisme est identique en 95 : *le numéro de mon fixe*. Le cas de 96 est un peu plus complexe, car si *effets secondaires* est un terme lexical composé relationnel par nature (*les effets secondaires de quelque chose*), l'argument possessif qui doit ici être restitué n'est pas exactement le complexe nominal *těch léků* (*des médocs*), mentionné dans le texte. En vérité, c'est le circonstant *vždycky* (*toujours*), établissant le cadre de la prédication, qui empêche cette lecture. Dans le paradigme de la grammaire générative, l'on emploie volontiers le terme de « quantifier domain/scope » (cf. Fodor et Sag, 1982 ; Schwarz, 2009 : 43-74). Ici, le circonstant est un quantificateur dont le domaine comprend toutes les situations/scenarii se trouvant sous sa portée, à

Dans les deux cas, la tête de l'anaphore associative est le nom *lidi* (*gens*). En l'occurrence, grâce à son sens général et son extension assez vague, ce nom s'adapte facilement au contexte particulier où il figure. Mais peut-on poser, pour ce nom, l'existence d'un sens lexicalisé relationnel (autrement dit : ce sens fait-il partie des variants sémantiques figés du nom (voir la Section 6.1.3.4)) ? A priori, cela nous paraît à la limite acceptable pour 97, dans la mesure où le nom *gens* semble souvent se prêter à l'interprétation qui est celle d'*habitants* : *lidi tčech rozvojových zemí* (*les gens des pays en voie de développement*). En revanche, une telle analyse est exclue pour 98, où le même nom s'interprète comme *politiques* / *acteurs de la scène politique*. Or ce variant sémantique – entièrement dépendant du discours et ne pouvant survivre en dehors de celui-ci – n'est nullement inscrit au niveau de la langue. C'est pour cela que 98 a été rattaché au type situationnel, contrairement à 97, qui a été rattaché au type relationnel. En dépit de ce type d'hésitations, il n'en reste pas moins que nos données contiennent suffisamment d'occurrences de type relationnel pour nous autoriser à affirmer qu'en tchèque parlé informel, TEN peut s'employer dans ce contexte.

Nos données contiennent une occurrence cataphorique que nous avons également classée, après de longues hésitations, parmi les emplois associatifs de type relationnel :

99) ježiš uč-í-š	studuj-e-š	čestín-u	tak	mus-í-š	vidět
MOD enseigner-PRS-2SG	étudier-PRS-2SG	tchèque-ACC.SG	alors	devoir-PRS-2SG	voir.INF
ten	rozdl .	<u>kysel-ý .</u>		<u>a navinul-ý</u>	
ten.M.ACC.SG	différence.ACC.SG	acide-M.NOM.SG		et acidulé-M.NOM.SG	
<i>ma foi, tu enseignes... tu fais des études de tchèque, du coup, tu dois sentir la différence :</i>					
<u>acide et acidulé</u>					ORTOFON v1

En 99, l'expression endophorique précède l'argument possessif qui sature contextuellement le nom relationnel *rozdl* (*différence*) : *la différence entre [les termes d'] « acide » et « acidulé »*.

L'énoncé soumis à l'évaluation de nos informateurs était le suivant :

100)	Když jsem	posledně	nastoupi-l	v Praz-e	do <u>autobus-u</u> ,
	quand AUX.PST.1SG	la dernière fois	monter-PTCP.SG.M	à Prague-LOC.SG	dans bus-GEN.SG
	tak	mě	<i>(ten)</i>	<i>řidič</i>	ani ne-pozdravi-l.
	alors	moi.ACC	ten.M.NOM.SG	chauffeur	même NEG-saluer-PST.3SG

*La dernière fois que je suis monté dans un bus à Prague, **le chauffeur** m'a même pas dit bonjour.*

Seuls 2 informateurs sur 10 ont choisi ici la variante avec TEN, ce qui témoigne de davantage de réticences quant à l'emploi de TEN avec ce type d'anaphore associative, y compris dans le registre parlé informel. Il n'en reste pas moins que la variante avec TEN est tout à fait possible.

Enfin, les deux occurrences que voici peuvent, selon nous, être analysées comme des cas qui relèvent de l'**anaphore associative partie-tout** :

- 101) L1: nech-á-š to bejt .. to zrcátk-o
 laisser-FUT-2SG cela.ACC être.INF ten.N.ACC.SG rétroviseur-ACC.SG
- ...
 L2: [v] ponděl-í se vyměn-í **ten** **kryt**
 en lundi-LOC.SG REFL changer-FUT.3SG ten.M.NOM.SG coque-NOM.SG
- a bud-e to vyřízen-ý
 et AUX.FUT-3SG ce réglé-N.NOM.SG

L1 : pour le rétroviseur, tu laisses tomber ?

...
 L2 : lundi, on change **la coque**, et ça sera réglé ORTOFON v1

- 102) já jsem se byl-a totiž podívat i v tom
 moi.NOM AUX.PST.1SG REFL être.PTCP-SG.F à savoir voir.INF même dans ten.M.LOC.SG
- byt-ě .. úplně náhodou v tom jejich nov-ým
 appartement-LOC.SG tout à fait par hasard dans ten.M.LOC.SG leur-M.LOC.SG nouveau-M.LOC.SG
- ...
 no [...] m-a-j to jakože .. parádn-í ..
 MOD avoir-PRS-3PL cela.ACC MOD génial-N.ACC.SG
- jakože uzpůsoben-í **těch** **pokoj-ů** a todle
 MOD configuration-ACC.SG ten.M.GEN.PL pièce-GEN.PL et MOD

parce que moi, je suis allée voir l'appart... complètement par hasard, leur nouvel appart

...
 ben, c'est vraiment... nickel...genre la configuration **des pièces** et tout ORTOFON v1

En 101, la tête nominale est *kryt* (*coque/couvercle*). Ce nom ressortit, de façon inhérente, à la sémantique relationnelle : il est naturellement question de la coque/du couvercle de quelque chose. De plus, la coque/le couvercle fait partie de l'objet qu'il couvre, et c'est pour cela que si anaphore associative il y a, il ne peut s'agir que de type partie-tout. Au vu de la conclusion de Czardybon (2017 : 90), nous avons, dans un premier temps, cherché à écarter cette hypothèse tout en envisageant d'autres interprétations possibles. Il serait éventuellement concevable de postuler une lecture en emploi « émotionnel », et ce dans une situation où le référent serait d'abord introduit sans TEN pour ensuite être repris dans une occurrence avec TEN. Il s'agirait ainsi de la reprise d'un référent « sémantiquement » défini, et donc d'un emploi « émotionnel de rappel » (voir à ce sujet les exemples 26 et 27 de la Section 5.1.3). Or ce n'est pas ici le cas, car le référent n'apparaît qu'une seule fois dans la conversation. Il nous semble par ailleurs qu'une analyse en emploi mémoriel ne peut non plus être complètement exclue : *la coque dont on a parlé l'autre jour / ce type de coques qu'il y a sur les rétroviseurs*. Dans ce scénario, le nom *coque* serait un nom sortal, et l'identification du référent

relèverait de la définitude « pragmatique » pure¹³³. Cette analyse pourrait ici être d'autant plus légitime que l'ensemble de l'échange s'inscrit dans le registre mémoriel : le locuteur 1 rappelle au locuteur 2 l'incident qui est arrivé récemment à ce dernier lorsque sa voiture a été vandalisée. Malgré cette interprétation alternative, l'anaphore associative de type partie-tout nous semble être une analyse tout à fait envisageable de l'occurrence. En 102, c'est également une interprétation en termes d'anaphore associative partie-tout qui nous semble la plus plausible. En effet, le nom *místnost* (*pièce/chambre*) compte parmi ses variants sémantiques figés (codés par la langue) des variants relationnels : *une pièce de mon appartement, les pièces de mon appartement*. Les pièces n'ont pas figuré plus tôt dans la conversation, et l'interlocuteur n'est pas en mesure de les identifier par la voie mémorielle, moyennant une connaissance partagée avec la locutrice, car tout son savoir à leur sujet se limite au compte-rendu que la locutrice lui en fait. Toutefois, vu le ton de la totalité de l'énoncé – deux mentions du modalisateur *jakože* et l'adjectif *parádní* (*génial*), la modalité phrastique exclamative –, TEN pourrait aussi relever ici d'un effet « émotionnel » d'affectivité.

C'est l'énoncé suivant qui a été soumis à l'évaluation de nos informateurs :

- 103) Petr si koupi-l nov-ý aut-o.
 NP.M.NOM.SG REFL acheter-PST.3SG.M nouveau-N.ACC.SG voiture-ACC.SG
- Hned druh-ej den se mu rozbi-l
 tout de suite deuxième-M.ACC.SG jour.ACC.SG REFL lui.DAT tomber en panne-PST.3SG.M
- (ten)** **motor.**
 ten.M.NOM.SG moteur.NOM.SG

*Petr s'est acheté une nouvelle voiture. Dès le lendemain, **le moteur** est tombé en panne.*

Personne n'a choisi la variante avec TEN, ce qui était tout à fait conforme à nos attentes, car la présence de celui-ci apparaît ici comme peu naturelle, voire agrammaticale. Comment se fait-il alors que TEN soit acceptable en 102 et non en 103 ? Voilà une question à laquelle nous ne sommes pas, pour l'instant, à même d'apporter une réponse satisfaisante, si ce n'est l'hypothèse mentionnée *supra*, à savoir qu'en 102, TEN ne relève pas, en réalité, d'un emploi mémoriel, mais d'un emploi « émotionnel ». À cet égard, nous tenons à remercier Mme Nádvořníková, qui nous a fait part de son sentiment, qui est le suivant : si l'exemple 103 était transformé de telle sorte qu'il traduirait une émotivité patente (dans le style de 104), TEN pourrait devenir acceptable :

¹³³ Et non pas du lien associatif entre le rétroviseur et sa coque. Nous estimons que si cette lecture est possible, c'est, entre autres, parce que le lien associatif entre les deux dénotés est moins solidement établi dans le discours : l'on pourrait ainsi identifier la coque tout en contournant le lien qui la lie au rétroviseur. Or ce contournement serait plus difficile dans un cas où le lien est plus solide (plus évident), comme par exemple dans celui d'une voiture et de son moteur (l'exemple 103 *infra*). Ces subtilités sont, à notre avis, autant de pistes qui pourraient être exploitées par une future recherche sur l'anaphore associative, dont la nôtre propre.

- 104) Si představ, že Petr si koupi-l nov-ej
 REFL imaginer.IMP.2SG que NP.M.NOM.SG REFL acheter-PST.3SG.M nouveau-N.ACC.SG
- bourák, ale hned druh-ej den se mu
 bagnole-ACC.SG mais tout de suite deuxième-M.ACC.SG jour.ACC.SG REFL lui.DAT
- rozbi-l **ten** **motor!**
 tomber en panne-PST.3SG.M ten.M.NOM.SG moteur.NOM.SG

*Imagine que Petr s'est acheté une nouvelle bagnole, mais dès le lendemain, **le moteur** est tombé en panne !*

Dans aucun des énoncés mentionnés jusqu'ici et ressortissant à l'anaphore associative (de quelque type que ce soit), la présence de TEN n'est obligatoire, y compris dans le registre informel. Si elle est clairement préférée dans certains énoncés appartenant au type situationnel, elle est davantage évitée avec d'autres énoncés de ce même type. Elle est possible avec (certains ?) énoncés de type relationnel, mais c'est son absence qui est ici préférée, y compris dans le registre informel. Enfin, s'agissant du type partie-tout – le type qui, des trois, est le plus proche du domaine de la définitude « sémantique » (cf. Czardybon, 2017 ; Dvořák, 2020) –, la présence de TEN est clairement exclue dans certains (la majorité des ?) cas. Pourtant, d'autres cas semblent admettre TEN dans le registre informel, sans que nous soyons capable, à l'heure actuelle, de donner une explication satisfaisante de cette dissymétrie. TEN adnominal étant le plus répandu et le plus acceptable avec le type situationnel, **l'hypothèse n°5** a été confirmée par nos données. Le Tableau 6.7 reprend les nombres de nos informateurs ayant choisi la variante avec TEN pour chaque exemple :

Exemple (n°)	Nombre d'informateurs ayant choisi TEN
89 (« le cinéma et le film »)	8
90 (« les toiles et les araignées »)	9
93 (« le magasin et la vendeuse »)	1
100 (« le bus et le chauffeur »)	2
103 (« la voiture et le moteur »)	0

Tableau 6.7 : Nombre de personnes ayant choisi d'employer TEN avec les exemples 89, 90, 93, 100 et 103 (sur 10 informateurs natifs)

6.2.2.3 Les « concepts individuels complexes »

Cette dernière section sera consacrée aux « concepts individuels complexes ». Comme déjà mentionné dans la Section 6.1.3.4 *supra*, ce terme recouvre **les noms modifiés par les adjectifs au superlatif, les ordinaux, ainsi que par les compléments analogues** tels que les adjectifs *prochain, suivant, précédent, dernier, aîné, antérieur, postérieur*, etc. De même que dans les cas examinés dans la Section 6.2.2.1, traitant des noms sortaux ou relationnels avec des compléments établissant l'unicité dénotative, le modifieur est ici l'élément qui, à lui seul, enclenche l'unicité dénotative du nom. Mais nous avons également indiqué dans la Section 6.1.4, en nous rapportant aux travaux de Mathesius (1926) et de Trovesi (2002), que lorsqu'il apparaît dans ce contexte, TEN, en plus de marquer l'unicité dénotative, est souvent doté d'une valeur « émotionnelle ». Aussi l'un des enjeux majeurs de la présente section sera-t-il de tâcher d'élucider la manière dont ces deux fonctions de TEN s'articulent.

Commençons par un relevé quantitatif des occurrences apparaissant dans nos données. Au total, nous répertorions 33 occurrences de ce type dans notre échantillon de conversations, contre 4 dans le sous-corpus de Havel :

L'échantillon de conversations		Le sous-corpus des pièces de V. Havel
Adjectif au superlatif	4	3
Numéral ordinal	17	1
Autres compléments	12	
SAMÝ/STEJNÝ (MÊME)	2	
DALŠÍ (SUIVANT)	2	
POSLEDNÍ (DERNIER)	2	
STARÝ (VIEUX/ANCIEN)	2	
OSTATNÍ (AUTRE)	1	
HLAVNÍ (PRINCIPAL)	1	
BUDOUCÍ (FUTUR)	1	
PŮVODNÍ (ORIGINEL)	1	

Tableau 6.8 : Nombre d'occurrences où l'unicité dénotative du nom est instaurée par a) un adjectif au superlatif, b) un numéral ordinal et c) un autre complément analogue dans l'échantillon de 1 000 occurrences d'ORTOFON v1 (CNT) et dans le sous-corpus des quatre pièces des V. Havel (InterCorp v1, CNT ; 241 occurrences)

Nous pouvons voir grâce au Tableau 6.8 que les occurrences contenant un numéral ordinal sont les plus nombreuses. Le superlatif est moins représenté que l'on ne pourrait s'y attendre. Une place considérable est également occupée (dans l'échantillon de conversations) par des compléments autres que le superlatif et le numéral ordinal. Voici une série d'exemples (les exemples 104 et 106 sont des reprises, respectivement, des exemples 27 et 26 de la Section 3.3.3) :

- 105) jo a todleto jsi vidě-l-a ? .. jak v tý na
MOD cela.ACC AUX.PST.2SG voir-PTCP-SG.F comment dans ten.F.LOC.SG sur

tý **nej-rušnějš-í** **ulic-i** v Los Angeles
 ten.F.LOC.SG SUP-plus animé-F.LOC.SG avenue-LOC.SG à Los Angeles.LOC.SG

jsou takov-ý-dle upoutávk-y
 être.PRS.3PL tel-F.NOM.PL-SUFFIXE DEICTIQUE panneau-réclame-NOM.PL

et ça, tu l'as vu ? ces espèces de panneaux-réclames qu'il y a dans l'avenue la plus animée de Los Angeles ?
 ORTOFON v1

- 106) VĚRA Pro Michala je totiž radost vařit! I **ten nejskromnější nápad** dokáže ocenit a pochválit [...] / littéralement : l'idée la plus simple

VERA. Oui, mais il faut que j'avoue que faire la cuisine pour Michael est un régal. C'est une merveille comme il s'entend à remarquer et à louer la moindre trouvaille [...]

V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

- 107) a **ten druh-ej** **film** ten dáva-l-i
 et ten.M.ACC.SG second-M.ACC.SG film celui-ci.M.ACC.SG donner-PST-3PL.M

včera na dvoj-ce..
 hier sur deux-LOC.SG

le second film, lui, il est passé hier sur la deuxième chaîne... ORTOFON v1

- 108) L1: když tam ne-pojed-eš ted'ka tam pojed-eš za rok nebo
 si là-bas NEG-aller-FUT.2SG maintenant là-bas aller-FUT.2SG dans an.ACC.SG ou

za dv-a že jo .. to asi ti ne-uteč-e snad ..
 dans deux.M.ACC MOD cela.NOM MOD toi.DAT NEG-échapper-FUT.3SG MOD

L2: to ne ale **ten dalš-í** **rok** už to bud-e
 MOD mais ten.M.ACC.SG suivant-M.ACC.SG an.ACC.SG déjà ce AUX.FUT-3SG

prostě jin-ý že jo
 simplement différent-N.NOM.SG MOD

L1 : Si tu y vas pas maintenant, tu y vas dans un an ou deux, non ? de toute façon, tu vas pas le rater, quoi...

L2 : ben non, mais l'année suivante, ça sera plus pareil, tu vois ORTOFON v1

- 109) představ-a že jsem v nějak-ý restaurac-i j-í-m
 idée-NOM.SG que être.PRS.1SG dans quelque-F.LOC.SG restaurant-LOC.SG manegr-PRS-1SG

chc-i si užít prostě buď přítomnost prostě **těch**
 vouloir-PRS-1SG REFL profiter.INF MOD soit présence-ACC.SG MOD ten.M.GEN.PL

ostatn-ích **lid-í** co tam jsou se mnou
 autre-M.GEN.PL gens-GEN REL là-bas être.PRS.3PL avec moi.INS

l'idée que je suis dans un resto, je mange, je veux profiter genre soit de la présence genre des autres personnes qui sont là avec moi [...]
 ORTOFON v1

En 105 et 106, c'est le superlatif qui est employé. En 107, l'unicité dénotative du nom est enclenchée par la présence d'un numéral ordinal. Pour ce qui concerne le sous-corpus de Havel, contenant 3 occurrences du superlatif et une occurrence de l'ordinal, c'est l'article défini qui est utilisé

Un premier constat important qui se dégage de nos données est le suivant : à l'intérieur de 13 des 37 occurrences recensées (soit 35,14 %), le complexe introduit par TEN ne contient pas de nom¹³⁴. Il convient, selon nous, de dissocier deux cas de figure au sein de cet ensemble. Soit un nom peut être restitué à partir du contexte, soit ce n'est pas le cas. Voici deux exemples qui illustrent cette différence :

- V. Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

480

tchèque, où l'adjectif apparaît sans nom, aussi bien le superlatif que TEN figurent à la forme neutre du singulier (*le plus malin qu'on puisse faire*)¹³⁶.

En elle-même, la présence d'un superlatif, ordinal ou autre complément semblable est considérée par nous comme une marque de la grammaticalisation en cours de TEN. Or rappelons-nous que dans la Section 3.4, nous avons également relevé le rôle que TEN peut jouer dans la substantivation de l'adjectif¹³⁷. C'est la raison pour laquelle, partout où TEN non seulement introduit un complément de la sorte, mais où, en plus de cela, il apparaît dans un contexte sans nom, sa proximité avec l'article défini n'en devient que plus apparente.

Afin d'avoir une idée des proportions des occurrences avec et sans TEN, nous avons cherché toutes les occurrences du superlatif adjectival dans l'ensemble d'ORTOFON v1. Pour ce faire, nous avons formulé deux requêtes pour trouver : 1) toutes les formes adjectivales au superlatif (qu'elles soient ou non précédées de TEN) et 2) les seules formes adjectivales précédées de TEN. Pour des raisons pratiques, nous nous sommes limité au superlatif de supériorité, le corpus permettant de formuler des requêtes précises isolant cette forme synthétique de l'adjectif¹³⁸. La première requête a débouché sur 529 résultats, soit une fréquence relative (occurrences par million de positions) de 427,82. La seconde a donné 74 résultats, soit une fréquence relative de 59,85 occurrences pour un million de positions. Ainsi, le nombre d'adjectifs superlatifs de supériorité introduits par TEN est bien inférieur au nombre d'adjectifs superlatifs de supériorité apparaissant sans TEN¹³⁹ : les premiers ne constituent que 14 % de toutes les occurrences sélectionnées, 86 % des occurrences de l'adjectif au superlatif de supériorité n'étant pas introduites par TEN.

À lui seul, ce constat constitue une preuve tangible du statut non-obligatoire de TEN avec les « concepts individuels complexes » (vu la ressemblance de fonctionnement syntactico-sémantique entre le superlatif et les autres types de compléments ici concernés, le même constat peut être étendu

¹³⁶ Pour les cas comme 111, où la restitution d'aucun nom n'est possible, une analyse alternative de l'ensemble peut être proposée, où TEN est appréhendé comme un pronom, et l'adjectif, comme son complément. Or ce qui nous rend réticent face à cette analyse, c'est d'abord la symétrie syntaxique avec les cas contenant un nom, mais aussi le fait qu'en français, de tels cas existent aussi : **Le pire**, *c'est qu'il se rends même pas compte !* Dans cet exemple, aucun nom ne peut être directement restitué, et pourtant, nul ne prétendrait que *le* est un pronom. Voir aussi l'espagnol *lo peor...*

¹³⁷ Syntaxiquement parlant, les compléments qui nous intéressent ici fonctionnent tous comme des adjectifs. Au sujet des ordinaux et de certains éléments semblables, il est également possible de parler de « post-déterminants » (cf. la Section 4.1.2).

¹³⁸ Analytique, la forme du superlatif d'infériorité est plus difficile à cibler. Les deux requêtes étaient celles-ci : « [tag="A.....3.*"] » pour tous les adjectifs au superlatif de supériorité et « [tag="P[D].*"] [tag="A.....3.*"] » pour les adjectifs au superlatif de supériorité précédés de TEN (seules les occurrences où le superlatif suit immédiatement TEN ont été ciblées). Dans les deux cas, la liste des concordances contenait une certaine quantité de « bruit » qu'il a fallu éliminer manuellement.

¹³⁹ Cette différence est significative sur le plan statistique : $L = 387,0145$; la valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

à ces compléments)¹⁴⁰. Ensuite, il est loin d'être insignifiant que parmi les 74 complexes nominaux au superlatif de supériorité qui sont introduits par TEN, 50 ne contiennent pas de nom (soit 67,57 %). Vu la taille du corpus, le fait que ces occurrences représentent plus des deux tiers de la totalité des occurrences introduites par TEN montre que si les locuteurs peuvent faire usage de TEN pour introduire un adjectif superlatif de supériorité dans n'importe quel contexte, ils le font de loin le plus souvent là où TEN, en l'absence de nom, est également doté de sa fonction de substantivation (voir la Section 3.4 pour l'introduction à ce phénomène). S'ajoute à cela le fait qu'au sein de l'ensemble de la population des 529 occurrences du superlatif de supériorité (avec ou sans TEN), seules 288 ne contiennent pas de nom (soit 54,44 %). Cette différence est statistiquement significative, l'application de la fonction de vraisemblance donnant le score de 4,6552¹⁴¹. Une preuve de plus qu'en ce qui concerne l'adjectif au superlatif de supériorité, il existe une affinité, dans ORTOFON v1, entre la présence de TEN et l'absence d'un nom.

Se pose à présent la question de savoir quels autres facteurs, outre l'absence de nom, sont favorables à l'apparition de TEN. En s'appuyant sur les travaux de bohémistes existants ainsi que sur les résultats de son propre questionnaire, soumis à des informateurs tchèques, Czardybon (2017 : 92) résume :

In Czech, the determiner can be used optionally with NPs with ordinal numbers and superlatives [...] However, this has the function to indicate emotional affectedness, similarly to standard Polish. **If the referent has not been mentioned before and the sentence is uttered in an emotionally neutral context, *ten* does not normally occur with superlatives and ordinals**¹⁴².

Selon l'auteur, deux facteurs influeraient ainsi sur la présence de TEN : la coloration émotionnelle du contexte où l'expression est employée¹⁴³ et la question de savoir si oui ou non le référent de cette expression a déjà fait l'objet d'une mention antérieure. Nous avons sollicité nos informateurs à propos de la présence de TEN dans l'énoncé de l'exemple 105 *supra* ainsi que dans les énoncés suivants :

112)	U	nás	na krám-ě	m-á-me	jenom
	chez	nous.GEN	sur magasin-LOC.SG	avoir-PRS-1PL uniquement	
	(to)	<u>nej-kvalitnějš-í</u>		<u>zbož-í.</u>	
	ten.N.ACC.SG	SUP-de meilleure qualité-N.ACC.SG		marchandise-ACC.SG	
- Chez nous, dans notre magasin, nous n'avons que la meilleure marchandise.*

¹⁴⁰ Bien qu'il faille souligner que tous les adjectifs au superlatif sans TEN ne sont pas forcément analysables comme étant employés en complexe nominal : en position syntaxique d'attribut, le superlatif sans TEN peut, en absence de nom, être analysé comme un simple adjectif, en même temps qu'il peut être analysé comme faisant partie d'un complexe nominal à adjectif substantivé.

¹⁴¹ La valeur critique = 3,841 (degré de liberté = 1 et $\alpha = 0,05$).

¹⁴² Nous soulignons.

¹⁴³ Voir également les remarques de Mathesius à ce sujet, reproduites dans la Section 6.1.4 *supra*.

(exemple inspiré par Zubatý, 1920 ; voir *supra*)

113)	Wilshire	Boulevard	je	
	Wilshire	Boulevard.NOM.SG	être.PRS.3SG	
	(ta)	<u>nej-rušnějš-í</u>	<u>ulic-e</u>	v Los Angeles.
	ten.F.NOM.SG	SUP-plus animé-F.NOM.SG	avenue-NOM.SG à Los Angeles-LOC.SG	

Wilshire Boulevard est l'avenue la plus animée de Los Angeles.

Nous constatons que le complexe nominal de 105 met en scène un référent caractérisé par les deux faits suivants : 1) l'interlocutrice est censée l'identifier en recourant à sa mémoire (*cette avenue-là, tu vois laquelle ?*) et 2) nous remarquons, chez la locutrice, une attitude de distanciation avec le référent, celle-ci cherchant à « [...] indiquer que la connaissance qu'[elle] a du référent n'est pas parfaite ou totale, que celui-ci présente des aspects inconnus, troubles, mystérieux, etc. » (Kleiber, 1991a : 92). En 112, le complexe nominal apparaît également dans un contexte teinté émotionnellement : le locuteur, recourant au registre parlé informel¹⁴⁴, prétend mettre en avant la qualité de la marchandise qu'il vend. Enfin, seul 113 met en scène un complexe nominal dans un contexte parfaitement neutre, à la fois sur le plan émotionnel et mémoriel. Aussi n'est-t-il guère surprenant que la variante avec TEN ait été choisie par une majorité, voire la totalité de nos informateurs en 105 et 112 (8 pour 105 et 10 pour 112), tandis qu'en 113, personne ne l'a choisie¹⁴⁵ :

Exemple (n°)	Nombre d'informateurs ayant choisi TEN
105 (« l'avenue la plus animée de Los Angeles – t'as vu ça ? »)	8
112 (« la meilleure marchandise »)	10
113 (« l'avenue la plus animée de Los Angeles – neutre »)	0

Tableau 6.9 : Nombre de personnes ayant choisi d'employer TEN avec les exemples 105, 112 et 113 (sur 10 informateurs natifs)

Regardons à présent l'énoncé suivant :

114)	to	jsem	byl-a	ted'ka	v Drážďan-ech	když
	MOD	AUX.PST.1SG	être.PTCP-SG.F	maintenant	à Dresde-LOC.PL	quand
	byl-y	ty	<u>nej-větš-í</u>	<u>slejšák-y</u>		
	être.PST-3PL.M	ten.M.NOM.PL	SUP-plus grand-M.NOM.PL	saucée-NOM.PL		

¹⁴⁴ En témoigne le tour *na krámě*, contenant la préposition *na* là où la langue standard emploie la préposition *v* et le nom familier *krám* (*magasin*) là où la langue standard emploie *obchod*.

¹⁴⁵ À en croire les choix de nos informateurs, la situation semble être la même en slovaque. En polonais, TEN peut apparaître dans les cas du type 112, où la coloration émotionnelle (affectivité ou prise de distance) est saillante, il ne le peut pas dans les énoncés du type 105.

L'emploi de TEN dans l'énoncé 114 se rapprocherait le plus, selon nous, de la valeur « émotionnelle de rappel » : la locutrice en appelle à la mémoire de l'interlocutrice au sujet des averses récentes. Comme le relève Adamec (1983 : 157, 165) à propos de ce type d'emploi, TEN fonctionne ici comme un commentaire mémoriel condensé : *les saucées les plus violentes qu'il y a eu à ce moment-là et que tu te rappelles certainement*. L'absence de TEN produirait, ici aussi, un énoncé bien moins naturel, notamment dans le registre (parlé) informel.

Pour résumer, en 105, 112 et 114, le complexe nominal au superlatif est employé dans un contexte qui n'est pas émotionnellement neutre, contrairement à 113. En 105 et 114, en plus, il est question d'un référent se présentant comme « déjà mentionné »¹⁴⁶, dans la mesure où il est censé être identifié, par l'interlocuteur, à travers sa mémoire. Seul l'exemple 113 admet comme choix prioritaire, dans le registre (parlé) informel, un complexe nominal sans TEN, alors que les autres exemples appellent la présence de TEN. La thèse résumée par Czardybon (2017 : 92) et reproduite *supra* se voit confortée par ce constat. Nous en voulons pour preuve supplémentaire que l'énoncé 113 appellera lui aussi la présence de TEN une fois transposé dans un contexte « émotionnellement » marqué :

- 115) Wilshire Boulevard v Los Angeles je zaručeně
 Wilshire Boulevard.NOM.SG à à Los Angeles-LOC.SG être.PRS.3SG à coup sûr

ta nej-rušnějš-í ulic-e,
 ten.F.NOM.SG SUP-plus animé-F.NOM.SG avenue-NOM.SG

co jsem kdy vidě-l!
 REL AUX.PST.1SG jamais voir-PTCP.SG.M

Wilshire Boulevard à Los Angeles, c'est, à coup sûr, l'avenue la plus animée que j'ai jamais vue !

En 115, l'engagement émotionnel du locuteur est signalé, à l'écrit, par la présence de la modalité d'énonciation exclamative et celle du modalisateur *zaručeně* (*à coup sûr*). D'autres indices signalent la coloration émotionnelle de l'énoncé dans nos exemples, dont le contexte plus large dans lequel celui-ci s'inscrit, la présence d'un lexique à dénotation et/ou connotations positive(s), et, à l'oral, la prosodie (telle qu'une insistance sur la forme du superlatif), éventuellement aussi l'expression non verbale d'émotions comme le rire. L'exemple 116 illustre la synergie de plusieurs facteurs à l'origine du « profil émotionnel » de l'énoncé :

¹⁴⁶ Sans qu'une mention antérieure explicite du référent apparaisse nécessairement dans l'échange verbal en question, l'important étant que le référent soit traité comme pouvant être identifié par le recours à une mémoire commune.

- 116) na přelom-u zář-í říjn-a jezd-í-me na Smrad'avk-u
 à tournant-LOC.SG septembre-GEN.SG octobre-GEN.SG aller-PRS-1PL à Smrad'avka-ACC.SG
- ve čtyř-ech případ-ech z pět-i je tak nádhern-é počas-í ..
 dans quatre-LOC cas-LOC.PL de cinq-GEN être.PRS.3SG si splendide-N.NOM.SG temps-NOM.SG
- to je takov-é to bab-í lét-o
 ce être.PRS.3SG tel-N.NOM.SG ten.N.NOM.SG indien-N.NOM.SG été-NOM.SG
- to je prostě *to* nej-hezč-í .. z toho
 ce être.PRS.3SG MOD ten.N.NOM.SG SUP-plus joli-N.NOM.SG de ten.M.GEN.SG
- podzim-u co muž-e být
 automne-GEN.SG REL pouvoir-PRS.3SG être-INF
- au tournant de septembre et octobre, nous allons à Smrad'avka ; dans quatre cas sur cinq, il fait un de ces temps magnifiques... c'est cet espèce d'été indien ; c'est tout simplement ce qu'il y a de plus beau en automne*
- ORTOFON v1

Hormis l'adjectif *nejhezčí* (le plus beau), l'énoncé contient un autre adjectif doté d'une forte sémantique positive – *nádherné* (splendide, magnifique) –, précédé par l'intensif *tak* (tellement, si). Remarquons également la présence du modalisateur *prostě* (tout simplement) et le fait que la combinaison de TEN avec le nom générique *podzim* (automne) débouche elle aussi sur un effet pragmatique « émotionnel ». Ensuite, le complexe nominal au superlatif lui-même est modifié par la relative *co může být*, voulant dire, littéralement, *qui puisse être*, et exprimant la plus haute valeur positive envisageable. À tout cela, il faut ajouter la prosodie enthousiaste de la locutrice, accompagnée d'un rire à connotations positives au moment de l'énonciation de l'adjectif *nejhezčí*¹⁴⁷.

Nous constatons que dans la majorité de nos occurrences, le complexe nominal apparaît dans un contexte marqué émotionnellement, qu'il s'agisse d'une valeur « de rappel » (cf. les exemples 105 et 114), d'affectivité (cf. les exemples 112, 115 et 116) ou de prise de distance (cf. l'exemple 104). Parmi les 37 occurrences concernées, nous estimons qu'au moins une de ces trois valeurs est présente dans 31 d'entre elles. Certains énoncés, comme 105, semblent cumuler plusieurs de ces trois valeurs¹⁴⁸. En plus de cela, 8 énoncés combinent, d'après nous, l'effet « émotionnel » avec l'absence de nom (la totalité de ces cas étant, rappelons-le, au nombre de 13). Il découle de ce bilan que très peu d'occurrences (plus précisément 3, à notre avis) ne s'inscrivent ni dans un contexte émotionnellement marqué ni en l'absence de nom¹⁴⁹.

À partir de tout ce qui a précédé ainsi que des choix de nos informateurs, il est possible de conclure que la présence de TEN est loin d'être obligatoire, bien qu'elle soit fortement encouragée par l'inscription de l'énoncé dans un contexte émotionnellement marqué, où elle est susceptible de

¹⁴⁷ Nous rappelons que la prosodie peut être analysée grâce aux enregistrements disponibles sur le site du CNT.

¹⁴⁸ Par ailleurs, il faut dire que la nature précise de l'effet pragmatique produit reste parfois difficile à établir.

¹⁴⁹ Dans 2 de ces 3 occurrences, le complexe nominal contient, en plus du complément, une subordonnée relative restrictive (cf. l'exemple 109 *supra*), phénomène qui pourrait également favoriser la présence de TEN.

produire les trois valeurs susmentionnées¹⁵⁰, de la même manière qu'elle est fortement encouragée par l'absence de nom. En conséquence, **l'hypothèse n°6** est confirmée par nos analyses. Soulignons au passage que ces trois valeurs sont exactement les mêmes que celles qui ont été identifiées et circonscrites, dans leurs traits majeurs, dans la Section 3.5.3. Dans cette section, il était question des emplois « émotionnels » purs de TEN, à savoir des occurrences où TEN ne possède pas de fonction grammaticale – il ne marque ni l'unicité dénotative du complexe nominal ni la possibilité d'identifier un référent au moyen de l'ensemble du complexe nominal –, mais uniquement une fonction pragmatique. Ici, en revanche, il arrive le plus souvent que ces deux fonctions se combinent, dans la mesure où TEN 1) marque l'unicité dénotative du complexe nominal¹⁵¹ et 2) véhicule au moins l'un des trois effets pragmatiques. Accessoirement, 3) le complexe nominal peut accomplir un acte de référence, dans lequel cas TEN en vient aussi à coder, dans la structure de la langue, le présupposé d'identification d'un référent. De par ses fonctions 1) et 3), TEN joue le même rôle, dans le système du tchèque parlé informel, que les articles définis des langues qui en disposent. Que l'ensemble du complexe nominal soit ou non employé pour référer semble n'avoir aucun impact sur la présence ou l'absence de la/des valeur(s) pragmatique(s) : en vérité, la plupart des exemples cités dans cette section n'accomplissent pas un acte de référence, à l'exception notamment de 105, 107 et 110¹⁵². En 113, la présence de TEN est de loin préférable à son absence, alors que c'est le contraire qui se produit en 115 ; pourtant, les deux exemples mettent en scène un emploi non-référentiel du complexe nominal placé en position syntaxique d'attribut.

6.3 Résumé et conclusion

Dans ce chapitre, nous nous sommes employé à examiner la question de savoir si un article défini est en train d'émerger dans le registre parlé informel du tchèque contemporain. Une introduction à la notion complexe de grammaticalisation et, plus particulièrement encore, à la grammaticalisation des démonstratifs en articles définis était un prérequis nécessaire à l'analyse concrète de nos données. C'est pour cela que nous y avons consacré la Section 6.1. Ont également été présentés, dans le cadre de cette section théorique, les deux types de définitude distingués par Löbner et l'ensemble de la théorie CTD. C'est en nous fondant sur cette distinction que nous avons mieux pu examiner, par la suite, l'état des choses en tchèque parlé informel. Les partis-pris de

¹⁵⁰ La notion de « fonction émotionnelle de rappel », inspirée directement par Adamec (1983), permet de subsumer les « mentions répétées » (Czardybon, 2017 : 92) sous la catégorie des « contextes émotionnellement marqués ».

¹⁵¹ Ou, exprimé en termes d'extension, il signale l'extensivité extensive du nom (Wilmet, 1986 : 194 ; cf. la Section 1.2.2). Rappelons encore que dans ses emplois de démonstratif pur, TEN participe également à la circonscription de l'extension du nom (cf. Wilmet, 1986 : 477 ; cf. la Section 1.2.2).

¹⁵² En 115, un doute subsiste : s'agit-il d'une prédication identifiante (emploi référentiel) ou d'une prédication qualifiante (emploi non-référentiel) ?

plusieurs auteurs au sujet du statut de TEN adnominal en tchèque ont été exposés à la fin de ce premier volet (Section 6.1.4), ainsi que le fait que ces derniers convergent en majorité : si TEN a bien entamé un parcours vers l'article défini, ce parcours est loin d'être abouti pour le moment. Nous en voulons pour preuves les faits suivants : 1) TEN ne fonctionne pas comme un article défini avec les « concepts individuels et fonctionnels lexicaux », soit au sein de la définitude « sémantique » pure ; 2) dans les autres emplois, la présence de TEN est plus ou moins préférée à son absence, mais rares sont, en vérité, les contextes où son absence est clairement ressentie comme agrammaticale (de fait, le seul contexte de ce type semble être celui de la deixis situationnelle). Nous avons bâti **nos** propres **hypothèses n°1-6** sur la base de ces constats.

Les analyses que nous avons présentées au lecteur dans le second volet (Section 6.2) – en partant des emplois qui relèvent de la définitude « pragmatique » pure pour aborder ensuite ceux qui s'étendent entre les définitudes « pragmatique » et « sémantique » – ont abouti à une série de conclusions plus ou moins attendues (car formulées par nos prédécesseurs) ou, au contraire, plutôt surprenantes. Reprenons ici uniquement quelques découvertes majeures : la grammaticalisation est plus facile à évaluer en deixis situationnelle, où le nombre de facteurs décisifs qui interviennent dans l'évaluation est bien moins élevé qu'en régimes anaphorique et mémoriel (voir les critères de la fin de la Section 6.1.4). En régime anaphorique, les limites du présent travail nous ont forcé à laisser à part certains facteurs pouvant influencer sur le choix entre TEN, un démonstratif fort, le nom nu et d'autres moyens anaphoriques encore. C'est ainsi que nous avons avant tout articulé nos critères de jugement de la grammaticalisation autour des notions d'activation du référent et de son statut (non) établi dans le discours. De plus, il s'avère que l'analyse de la grammaticalisation au sein de l'anaphore textuelle peut parfois déboucher sur des positions en partie contradictoires, comme nous l'avons vu à la fin de la Section 6.2.1.2 au sujet de la fonction syntaxique de l'anaphorique.

Dès lors que TEN est investi d'une charge affective (en plus de marquer la possibilité d'identifier le référent), la grammaticalisation n'a pas lieu, quand bien même les autres conditions seraient réunies (voir **les hypothèses n°1 et 3**). En régime anaphorique, cela va de pair, en plus, avec la nécessité de recourir à un démonstratif pour s'assurer du maintien de la coréférence, mise à mal par le choix d'un « nom de qualité » au sein de la reprise. Les occurrences où TEN se cumule avec d'autres mots grammaticaux ont aussi été étudiées. Nous avons pu voir que souvent, le statut de ces occurrences au regard de la grammaticalisation doit être déterminé via une analyse au cas par cas, car une même combinaison peut être porteuse, dans le discours, de valeurs sémantico-référentielles différentes (voir les critères dans la Section 6.1.4). Dans les traductions des occurrences en français, la valeur de TEN est rendue avec plus ou moins de fidélité, sachant que les traducteurs disposent parfois d'une marge de manœuvre certaine, dans la mesure où l'original lui-même admet plus d'une interprétation (cf. TEN

peut parfois être interprété davantage comme démonstratif ou comme défini, une subordonnée relative peut, à son tour, parfois être ressentie comme restrictive ou non, etc.).

À plusieurs reprises, nous avons été contraint d'admettre que les critères sur lesquels reposaient nos jugements concernant la grammaticalisation comprenaient une dose de subjectivité et d'approximation, notamment dans le cas des emplois anaphorique (hormis l'anaphore associative) et mémoriel. Nous avons essayé de pallier ce défaut 1) en examinant autant que faire se pouvait le contexte large d'emploi de l'occurrence (nombre de mentions antérieures, distance entre ces mentions, degré d'activation du référent au moment de sa reprise via TEN, etc.), 2) en procédant parfois à des tests de reformulations des énoncés accueillant l'occurrence et, enfin, 3) avec les jugements de nos dix informateurs tchèques. Les retours de ces derniers nous ont permis de faire plus d'une découverte, notamment en ce qui concerne la présence davantage à préconiser ou à éviter de TEN dans les contextes à la charnière entre les deux types de définitude, mais aussi les facteurs susceptibles de favoriser ou de décourager cette présence. Dans certains cas, la version avec TEN entraîne même un changement d'interprétation radical.

Nous avons pu avoir la confirmation, grâce à nos informateurs, du fait que TEN est non seulement ressenti comme obligatoire en contexte de deixis situationnelle ne contenant qu'un seul représentant de la classe dénotée par le nom et ses compléments (hormis les situations où il est porteur d'une subjectivité accrue ; voir **l'hypothèse n°1**), mais qu'il constitue également le choix de loin préféré en contexte mémoriel (**l'hypothèse n°2**) et dans une partie du contexte anaphorique, soit, notamment, en anaphore fidèle où TEN n'est ni « établissant » ni « réactivant » pour le référent (**l'hypothèse n°3**). Pour ce qui est des contextes de transition entre les deux types de définitude, nous avons pu constater que TEN y a acquis une place incontestable. Dans une partie de ces contextes, son usage constitue, pour ainsi dire, l'option neutre. Nous songeons ici, avant tout, à certaines occurrences de type situationnel de l'anaphore associative (**l'hypothèse n°5**), aux noms modifiés par une subordonnée relative « établissante » (une partie de **l'hypothèse n°4**) et aux « concepts individuels complexes ». S'agissant de ces derniers, nous croyons avoir apporté des preuves à **l'hypothèse n°6**, selon laquelle la présence de TEN est largement encouragée par la subjectivité/expressivité du locuteur (cf. Mathesius, 1926 et Czardybon, 2017), tout comme par l'absence de nom.

À cet égard, nous avons identifié trois grands effets pragmatiques « émotionnels » produits par l'apparition de TEN au sein de cette catégorie et correspondant à ceux qui se produisent également lorsque TEN accompagne un concept « sémantiquement » défini (voir la Section 3.5.3). Nous pensons avoir démontré que l'analyse de ces effets (à la fois leur présence et leur nature) est un processus complexe où une série de facteurs demandent à être pris en compte, à commencer par la sémantique

du complexe nominal, mais aussi la totalité de l'énoncé où celui-ci apparaît, les éléments phonétiques suprasegmentaux (la prosodie) et les expressions non verbales (le rire). Le fait qu'un effet pragmatique (aussi discret soit-il) puisse, dans nos données, être attribué à la majorité des « concepts individuels complexes » introduits par TEN renvoie à l'existence, au sein de cette catégorie, d'un lien étroit entre le simple marquage de la définitude et la présence des effets pragmatiques étudiés. Ainsi, il semble plausible que l'effet pragmatique en question influe sur le processus de grammaticalisation de TEN tout en l'encourageant et peut-être même en l'accéléralant. Cette hypothèse s'accorde avec les travaux sur l'évolution de l'article défini mettant l'accent sur le rôle de la pragmatique et de la subjectivité dans le processus de grammaticalisation, travaux que nous avons brièvement présentés dans le premier volet de ce chapitre (voir notamment la Section 6.1.2). Rappelons que selon leurs auteurs (cf. Epstein, 1995 ; Traugott et Hopper, 2003 [1993]), le rôle de la subjectivité est particulièrement important dans les premières phases du processus. S'agissant du tchèque (parlé) informel, le processus en serait, en effet, à ses phases initiales (voir *infra*).

En ce qui concerne l'absence de nom à l'intérieur du complexe nominal, les chiffres que nous avons pu obtenir en consacrant une étude plus exhaustive au cas du seul superlatif de supériorité dans l'ensemble d'ORTOFON v1 vont dans ce sens.

Compte-tenu non seulement de l'apparition de TEN en tant qu'article défini avec la définitude « pragmatique » (analysée dans la Section 6.2.1), mais aussi de sa prolifération dans les contextes charnières analysés dans la Section 6.2.2, le tchèque parlé informel se rapproche le plus de ce qu'Ortmann désigne par le terme de « scission du type I » (2014 : 294 ; voir la Section 6.1.3.4). C'est donc l'ensemble de **l'hypothèse n°7** qui se voit appuyée par nos analyses. Néanmoins, il ne fait que s'en rapprocher, car, comme nous avons pu le constater, son emploi est rarement figé au point d'être obligatoire. Se pose, évidemment, la question suivante : le processus se poursuivra-t-il dans l'avenir, de sorte qu'en l'espace d'un ou de plusieurs siècles, non seulement TEN pourrait devenir obligatoire dans tous les registres du tchèque au sein de tous les contextes situés en dehors de la définitude « sémantique » pure, mais qu'il commencera à s'étendre même à ces derniers ?

Il faut dire que la situation du tchèque est particulièrement imprévisible à cet égard. D'abord, nous sommes loin du profil analytique des langues romanes et de certaines langues germaniques contemporaines : le tchèque est une langue où le système flexionnel se porte très bien, à la différence, par exemple, du latin tardif et du roman (cf. Herman, 1985). Que le lecteur sache que nous partons ici – comme à la fin de la Section 6.1.3.3 – de la simple observation typologique suivante : la naissance des articles a souvent accompagné, dans les langues du monde, le changement typologique qu'est le passage d'un système davantage synthétique et non configurationnel vers un système davantage analytique et configurationnel (pour le français, voir par Marchello-Nizia, 1995 ou Carlier

et Combettes, 2015). Nous sommes, par ailleurs, tout à fait conscient de la complexité de la question tout comme de l'importance de ne pas tomber dans le piège d'une explication de cause à effet trop directe, et donc factice :

Il n'y a pas de relation de cause à effet entre l'effritement de la flexion suffixale et la création de nouvelles formes grammaticales antéposées à la tête nominale ou verbale. Le développement des prépositions, des articles, des pronoms sujets, des formes analytiques du verbe, tel que le parfait analytique, est bel et bien antérieur à l'affaiblissement de la morphologie suffixale. Néanmoins, dans la perspective d'une cohérence quant au mode d'expression de la grammaire, l'affaiblissement progressif de la morphologie suffixale conduit à grammaticaliser davantage les marqueurs grammaticaux antéposés à la tête, à leur faire perdre le poids pragmatique. (2015 : 20)

Quoi qu'il en soit, le tchèque demeure une langue penchant nettement vers le pôle de la non-configurationnalité, tendance qui se manifeste non seulement par la richesse de son système flexionnel et la liberté de son ordre des mots (« [...] régi par des facteurs pragmatiques plutôt que par la syntaxe [...] », Carlier et Combettes, 2015 : 17), mais aussi par un degré plus faible de hiérarchisation entre les constituants, que ce soit au sein du complexe verbal ou du complexe nominal (voir les Section 3.1.3 et 4.4).

Au vu de ce qui a été exposé à la fin de la Section 6.1.3.3, à savoir que le processus de l'émergence des articles, une fois entamé, semble être favorisé par l'érosion de la flexion nominale, il y a une certaine légitimité à croire qu'en tchèque aussi, le processus en question – que nous nous sommes employé à décrire dans le second volet de ce chapitre – aurait de fortes chances d'aboutir si un jour cette flexion s'affaiblissait. Or, répétons-le, rien ne semble l'indiquer.

De surcroît, il y a toujours la norme soutenue – bien plus réticente à l'expansion de TEN comme article – dont l'autorité reste de nos jours incontestable (cf. Cvrček *et al.*, 2015 [2010] : 176 *supra*). Il se pourrait ainsi que l'ombre de cette norme « hante » le locuteur au point de l'empêcher d'employer le TEN articulaire dans le registre informel de manière complètement décomplexée. En comparant la situation relative à la grammaticalisation du démonstratif en tchèque et en haut-sorabe, Trovesi (2002 : 134) constate que dans la première des deux langues, même engagé dans ce processus, le démonstratif est toujours davantage teinté d'expressivité :

Il existe, à mon sens, une différence dans le degré d'expressivité entre l'emploi de *ten* adnominal en tchèque et de *tón* adnominal en sorabe. Influencée par l'usage non-marqué de *der* en allemand, le sentiment de la langue des Sorabes, bilingues, voit en *tón* une marque de définitude moins expressive que celui des Tchèques. Dans le cas du locuteur natif tchèque, l'exemple de la langue soutenue est bien plus prégnant, qui évite cet usage excessif¹⁵³.

¹⁵³ Nous traduisons : „Mezi užitím atributivního *ten* v češtině a *tón* v lužické srbštině podle mého mínění existuje rozdíl ve stupni expresivnosti. Jazykové povědomí dvojjazyčných lužických Srbů vnímá pod vlivem nepříznakového užití *der* v němčině *tón* jako méně expresivní prostředek k vyjádření určenosti, než je tomu v češtině v případě zájmena *ten*. Český

Cela expliquerait en partie pourquoi le paradigme TÓN est obligatoire, en haut sorabe, en contexte d'anaphore textuelle (indépendamment de la structure thème-rhème), en contexte où le nom est modifié par une relative « établissante » ainsi que – du moins chez la jeune génération – dans tous les types d'anaphore associative et avec les « concepts individuels complexes » apparaissant en position rhématique (Czardybon, 2017 : 87-92).

En tchèque, un scénario similaire pourrait seulement se produire, à notre avis, si la norme soutenue en venait à perdre l'essentiel de l'autorité qu'elle exerce sur le bon usage de la langue à travers l'école, les médias de service public, les journaux et la littérature¹⁵⁴. Comme nous avons pu le constater notamment dans le Chapitre 3, c'est le dernier de ces trois domaines qui semble le plus ouvert à l'innovation : les textes de la partie fiction du corpus représentatif du tchèque écrit en synchronie SYN2015 sont ceux qui contiennent de loin le plus d'occurrences de TEN et de TENHLE, contrairement aux deux autres parties – non-fiction et presse –, où les auteurs se soucient davantage de la norme soutenue (voir la Section 3.1.2). Cela est tout à fait compréhensible, dans la mesure où la fiction a tout intérêt à s'inscrire dans une alternance et superposition de registres (langue du narrateur vs langue des personnages, etc.). Ainsi, faisant se côtoyer (au moins) deux normes, elle peut refléter assez fidèlement la situation de quasi-diglossie propre au tchèque contemporain (cf. Čermák, 2014). D'autres types de textes propices à cette rencontre sont, par exemple, les entretiens dirigés (voir la Section 3.1.2) ou certains autres types de communication semi-formelle tels que les réunions de travail ou la correspondance (y compris électronique) entre personnes proches¹⁵⁵.

Afin de renouer ici avec le volet théorique de la présente partie, il importe de se poser la question de savoir quel stade de la « courbe en S » de Kroch (1989 : 223 ; voir la Section 6.1.1) a déjà été atteint par le processus de grammaticalisation de TEN en tchèque (parlé) informel. Au vu de ce qui a été montré ici, nous croyons juste de poser qu'il s'agit, pour l'instant, (encore) du début de cette courbe, soit le moment où la fréquence des emplois est en hausse lente, la forme se répandant progressivement de contexte en contexte, avant de connaître une montée en flèche dans tous les contextes à la fois (donc également ceux de définitude « sémantique »). De même convient-il de s'interroger sur la place de TEN adnominal marqueur d'unicité dénotative dans le schéma inspiré par Heine (2002 ; voir la Section 6.1.1) :

rodilý mluvčí má kolem sebe v nesrovnatelně větší míře příklad standardního jazyka, který se tomuto nadměrnému užití vyhýbá.“

¹⁵⁴ Pour comprendre le poids de cette norme, la prise en compte de l'histoire du 18^{ème} et du 19^{ème} siècle de l'actuelle République tchèque est incontournable. Nous aborderons brièvement ce sujet dans la Conclusion générale.

¹⁵⁵ Cette dernière s'inscrit au sein de l'ensemble des « textes écrits de la communication quotidienne » („schriftliche Texte der Alltagskommunikation“ ; Berger, 1993).

Nous estimons que cette place commence à se rapprocher du stade Ba du schéma, soit du « contexte de passage ». En effet, comme l'écrit Marchello-Nizia, à ce stade de l'évolution,

[l]e sens initial est relégué à l'arrière-plan (mais toujours accessible), et le nouveau sens permet au mot d'apparaître dans des contextes tout à fait nouveaux, et pas seulement dans les contextes ambigus qui avaient permis son apparition. (Marchello-Nizia, 2006c : 23)

Nous avons pu observer, dans nos données, que TEN peut apparaître dans des positions où d'authentiques démonstratifs sont difficilement, voire pas du tout admis. Ainsi, non seulement TEN apparaît-il en situation où nul contraste entre plusieurs référents de la même classe n'est en jeu, mais il a même « débordé » du seul domaine des « contextes ambigus qui avaient permis son apparition », à savoir celui de la définitude « pragmatique » pure. Un bémol : comme nous avons également pu le voir, dans ces derniers contextes, TEN continue, en même temps, à pouvoir fonctionner comme un authentique démonstratif¹⁵⁶. En conséquence, il n'est pas juste de dire que son sens initial soit complètement relégué à l'arrière-plan. La place actuelle de TEN adnominal en tchèque (parlé) informel se situerait donc davantage entre les phases Ab et Ba.

Nous avons écrit dans la Section 6.1.1 en paraphrasant Heine (2002) que la phase Ab correspond à la dérivation par inférence d'un nouveau sens au sein d'un (ensemble de) contexte(s) où l'expression est utilisée. Dans notre cas, ce processus s'apparente à l'émergence de la fonction de TEN consistant à marquer l'unicité (maximalité) dénotative d'un nom en régime de définitude « pragmatique » pure. Cette fonction elle-même paraît être une conséquence indirecte de l'insistance sur la possibilité pour l'interlocuteur d'identifier le référent. C'est là, précisément, que l'inférence en question apparaîtrait : s'obstinant à marquer la possibilité d'identifier le référent du nom et ses compléments même lorsque le contexte ne l'exige pas, le locuteur crée un effet pragmatique d'insistance sur l'interlocuteur. Il y a ainsi renforcement pragmatique (Traugott, 2005 [1995])¹⁵⁷. C'est également là que la subjectivité, s'accompagnant de la notion de point de vue – en l'occurrence, celui du locuteur – entre en jeu (voir la Section 6.1.2). Cependant, cet effet ne caractériserait que les débuts du processus. Plus tard, il semble se faire plus discret, voire disparaître complètement, ce qui aboutit à la possibilité pour l'ancien démonstratif d'apparaître également en contextes 1) où la référence cesse d'être tributaire du contexte immédiat, voire 2) là où il n'y a même pas d'acte de référence accompli (voir à cet égard certains des exemples cités et commentés dans la Section 6.2.2). Ces contextes correspondent à la

¹⁵⁶ Notamment dans le contexte anaphorique (avec ses fonctions « établissante » et « réactivante », mais aussi sa fonction de recatégorisation du référent ; voir la Section 6.2.1.2), moins dans le contexte mémoriel et assez peu, finalement, dans le contexte situationnel.

¹⁵⁷ Dans le prochain chapitre, nous essaierons de mieux circonscrire la notion d'inférence pragmatique en rapport avec différents types de fonctionnement du démonstratif.

transition entre les définites « pragmatique » et « sémantique », et, plus tard encore, à ceux relevant de la définitive « sémantique » pure. Dans tous ces contextes, le corollaire de l'inférence (le marquage de l'unicité dénotative) subsiste, alors que l'inférence en tant que telle (l'insistance sur le statut identifiable du référent) n'est plus présente.

Comme nous l'avons aussi relevé au début de ce chapitre, la mutation en défini n'est pas sans conséquences sur le sens du démonstratif qui s'y est engagé : si la forme a perdu son sens indexical d'origine (cf. Kleiber, 1991b), elle développe, en revanche, la faculté de marquer l'unicité dénotative. Ainsi, il y a à la fois perte et gain de sens. Dans le cadre de notre Conclusion générale, nous tâcherons de synthétiser, à l'aide d'un tableau, l'ensemble des étapes que TEN a déjà traversées ou qu'il pourrait encore traverser au cours de sa grammaticalisation en article défini.

7 Les emplois « émotionnels » du démonstratif TEN

Dans ce dernier chapitre, nous nous focaliserons sur les emplois que nous avons pris la décision, dans la Section 3.5 du Chapitre 3, d'appeler « émotionnels ». Nous ne reviendrons plus ici sur ceux parmi les emplois de TEN où ce dernier, bien que potentiellement porteur d'une subjectivité accrue, soit intervient directement dans l'identification du référent, soit exprime l'unicité dénotative du complexe nominal, soit encore les deux. Tous ces cas ont été introduits dans le Chapitre 3, où nous avons utilisé à leur sujet le terme d'« emplois désignatifs ». Ces cas de figure ont été traités en détail dans le Chapitre 6 précédent en rapport avec les différents (sous-)types d'emploi. Ainsi, l'analyse sera ici limitée au seul cas des emplois « émotionnels purs », à savoir ceux où TEN ne contribue pas – du moins pas directement – à l'identification du référent. Rappelons que selon notre hypothèse principale, ces emplois constituent, ensemble avec les emplois où TEN est en voie de grammaticalisation, l'un des deux principaux facteurs de la haute fréquence du démonstratif adnominal en tchèque parlé informel (voir l'Introduction générale). Voici à présent le plan du chapitre.

Dans un premier volet, essentiellement théorique, nous reviendrons brièvement sur le choix du terme « emplois émotionnels » en tâchant d'inscrire ce choix dans un débat plus large, et nous rappellerons également certains constats introduits à leur sujet dans les Section 3.5.3 et 4.5.4. Cela nous servira d'introduction à la problématique. Ensuite, nous approfondirons notre approche de ce type d'emploi en recourant à la théorie CTD et, plus précisément, à la notion de « définitude sémantique ». Enfin, nous tenterons d'en proposer un modèle universel d'explication articulant sémantique référentielle et pragmatique, modèle bâti sur les notions d'« implicature conversationnelle » de Grice (1975) et de « sous-entendu » de Ducrot (1980).

Un second volet sera ensuite consacré aux analyses concrètes des différents cas de figure présents dans nos données. La structure principale de ces analyses obéira à la distinction opérée dans la Section 3.5.3 et réitérée dans la Section 6.2.2.3, soit celle entre les trois effets/valeurs pragmatiques majeurs auxquels aboutit la présence de TEN. Ces valeurs sont : l'affectivité, la prise de distance et le rappel du fait qu'il a déjà été question du référent/de l'information énoncée¹. Outre l'effet produit par TEN et certaines particularités plus ou moins réservées à l'un des trois sous-types d'emploi « émotionnel », nous nous intéresserons à la place du complexe nominal au sein de la proposition, ce qui nous permettra de voir si des tendances récurrentes existent quant à celle-ci. Si nos données de corpus tchèques nous serviront ici de principal réservoir de matériel empirique, d'autres données – y

¹ L'un des enjeux de ce chapitre sera précisément de démontrer que dans le cadre des emplois « émotionnels » (notamment ceux « de rappel »), TEN n'introduit pas nécessairement un nom employé référentiellement.

compris françaises – n'en seront pas moins mobilisés pour appuyer ou compléter notre propos. Pour ce qui concerne les occurrences puisées dans le sous-corpus des pièces de V. Havel, il nous importera également d'examiner leur transposition en français. Exactement comme dans le chapitre précédent, il s'agira de répondre à la question de savoir dans quelle mesure les choix des traducteurs reflètent la situation de l'original et comment la langue cible – où le phénomène connaît une bien moindre ampleur (cf. les Sections 3.5.3, 4.5.4, 5.2.2.1 et 5.2.3.1) – s'en accommode. Des exemples français – puisés ou non dans notre échantillon de conversations – seront avant tout mobilisés dans la Section 7.1.1 *infra*, dans le cadre de l'introduction à la problématique. Compte-tenu de ces éléments, il est possible de dire que nous adopterons ici une approche contrastive, dans la mesure où nous partirons du français pour aussitôt aborder la même thématique en tchèque.

7.1 Les emplois « émotionnels » : pour un modèle unifié de traitement

Dans le présent volet, il convient, dans un premier temps, de faire une brève introduction à l'ensemble des emplois concernés. S'ensuivra une section dans laquelle sera rappelée la place de ces emplois au sein de la théorie CTD. Nous terminerons avec l'exposition d'un modèle d'explication sémantico-pragmatique du fonctionnement des emplois « émotionnels ».

7.1.1 Les emplois « émotionnels » : introduction

C'est dans la Section 3.5.3 que nous avons indiqué que de nombreuses dénominations ont cours dans les deux milieux linguistiques. Rappelons-en ici quelques-unes. Si l'on commence, par la tradition tchèque, il y a d'abord Mathesius qui, dans son article pionnier de 1926, utilise le terme de « funkce emocionální » (littéralement : « fonction émotionnelle »). Dans l'étude d'Adamec parue en 1983 se trouve la distinction entre les fonctions appréhendées par l'auteur comme primaires (concrètement : les « fonctions d'identification » déictique, anaphorique et cataphorique) et celles appréhendées comme périphériques, à savoir les « fonctions secondaires » („sekundární funkce“). C'est ensuite chez Štícha (1999) que nous trouvons la désignation explicite de « fonctions périphériques » :

L'on peut, à mon avis, considérer comme généralement accepté (ou acceptable) le fait qu'il est possible d'attribuer au lexème *ten* trois ensembles de fonctions : 1) les fonctions déictiques (monstratives) dénotatives (d'identification), 2) les fonctions de renvoi contextuel et discursif (anaphoriques et cataphoriques) et 3) les fonctions expressives, dites « additionnelles ». Les deux premiers ensembles se trouvent au centre du système, le troisième, à sa périphérie. (1999 : 124)²

² Nous traduisons : „Lze myslím pokládat za obecně akceptovaný (resp. akceptovatelný) fakt, že lexému *ten* lze přisoudit trojí okruh funkcí: 1. deiktické (ukazovací) funkce denotační (identifikační); 2. kontextově-diskurzivní (anaforické a

Comme nous le lisons dans la citation, ce même auteur recourt également aux termes d'« expressifs » et d'« additionnels », autant d'adjectifs mettant l'accent sur le caractère marginal de ces fonctions. Un autre adjectif – celui d'« affectifs » – apparaît également dans la tradition tchèque (cf. Trávníček, 1951 ; Voráč, 1946). Mais, comme nous avons également pu le constater dans la Section 3.5.2, cette dichotomie entre fonctions primaires et secondaires ne se limite pas à la seule tradition tchèque. À cet égard, nous reproduisons ici la citation de Fillmore :

Elements of demonstrative systems that have primarily the Informing function are the demonstrative adverbs [...] those that have mainly the Identifying function may be thought of as the prototypic uses of demonstratives, the Acknowledging use being derivative and secondary. (Fillmore, 1982 : 45)

Dans la tradition française, c'est sous la plume de Kleiber (1991a) que nous retrouvons le terme d'« affectif », ce dernier distinguant, rappelons-le, entre les emplois « affectifs » à proprement parler (aussi dits « de familiarité ») et les emplois « de distanciation » au sein des cas où le déterminant démonstratif CE introduit dans le discours un nom propre employé dans sa fonction prototypique (à savoir là où il est doté d'un sens instructionnel).

De par notre choix de l'adjectif « émotionnel », nous nous inscrivons dans le chemin tracé par l'étude de Mathesius. Le lecteur aura ensuite remarqué que notre distinction des trois effets pragmatiques susmentionnés se base sur trois sources essentielles, que nous citons ici dans leur intégralité et selon l'ordre chronologique : l'entrée « Les mots démonstratifs » du tome 2 du *Grand Larousse de la langue française* de 1972, rédigée par Bonnard, l'étude d'Adamec « České zájmeno ten a jeho ruské ekvivalenty », publiée en 1983 et, enfin, ledit article de Kleiber « Du nom propre non modifié au nom propre modifié : le cas de la détermination des noms propres par l'adjectif démonstratif » de 1991, où l'auteur fait explicitement sienne une partie des thèses de Bonnard.

Comme nous l'avons déjà écrit dans la Section 3.5.3, Bonnard a proposé, selon nous, la meilleure explication du fonctionnement des emplois qui nous intéressent ici : en présentant le référent sur fond des circonstances précises dans lesquelles celui-ci a fait l'objet d'une certaine évaluation par le locuteur, le démonstratif traduit (indirectement) la nature de cette évaluation et, par cela même, fait surgir l'effet pragmatique en question. Il s'agit donc d'une indexation du référent sur le contexte, de la même manière que cela se produit lorsque le référent n'est pas de nature « sémantiquement » définie³. Cette évaluation peut être rendue explicite par un modifieur du nom tel qu'un adjectif épithète ou la combinaison « nom de qualité » + *de* :

kataforické) funkce odkazovací; 3. expresivní funkce „přídavné“. První dva okruhy funkcí jsou systémově centrální, třetí je systémově periferní.“

³ Cf. Charolles (2002), Corblin (1987) et Danon-Boileau (1990).

[...] l'énonciation de *ce cher Pellegrin* ou de *cette pauvre Josette* s'appuie sur une situation antérieure qui justifie la qualification opérée par *cher* et par *pauvre*. Il reste que l'interlocuteur doit aussi avoir accès à une telle situation et, puisqu'il s'agit d'une identification démonstrative, qu'il y ait dans la situation d'énonciation des éléments qui y conduisent. (Kleiber, 1991a : 94)

Nous sommes d'avis que souvent, le/l'un des prédicat(s) de l'énoncé accueillant le complexe démonstratif, ainsi qu'une subordonnée relative descriptive apposée à ce dernier reflètent eux aussi cette évaluation. Quant au premier des phénomènes, nous l'avons déjà relevé dans la Section 4.5.1. Nous y avons cité Gary-Prieur et sa remarque au sujet de la nature de ce prédicat lorsqu'il véhicule une qualification subjective d'un référent générique discursif :

Il en découle [...] que la généralité énoncée l'est sous la seule responsabilité du locuteur : c'est lui qui assume l'application du prédicat [...] On a donc affaire à une généralité subjective. C'est pourquoi n'importe quel prédicat ne convient pas à ce type de généralisation. (Gary-Prieur, 2001 : 224)

Nous avons écrit au même endroit que parmi les 11 occurrences « émotionnelles » présentes dans notre échantillon de conversations françaises, 3 s'accompagnent d'un prédicat évaluatif du type *être fous*, comme nous le voyons dans la fameuse réplique d'*Astérix* également citée dans la Section 4.5.4 :

- 1) Ils sont fous, *ces Romains* ! (emprunté à Kleiber, 1991a : 95)

Regardons à présent les exemples 2-5 :

- 2) Regarde *cette pauvre Louise*, elle devient de plus en plus dure d'oreille.

(C. Michelet, emprunté à Kleiber, 1991a : 93)

- 3) L1 : ouais mais bon pour le lire c'est vrai que c'est dur avec euh maîtriser mille cinq cents
L2 : plus même
L1 : plus de m- plus de mille cinq cents ouais c'est mille cinq cents kanjis pour lire un journal
L2 : minimum dans les mille euh plus je crois plus
L1 : hm comment ils font *ces japonais* ils sont fous CEFC

- 4) *Ce salaud de Victor* ! Il m'a encore bu tout le whisky ! (Kleiber, 1991a : 93)

- 5) Et *ce Bush* qui a encore bombardé Bagdad ! (Jonasson, 1994 : 63)

En 2, l'évaluation du référent comme pauvre, opérée au moyen d'un emploi affectif du démonstratif et de la présence de l'adjectif *pauvre* en tant que tel, s'appuie directement sur le fait de devenir de plus en plus dur d'oreille, exprimé, lui, par le prédicat souligné : Louise est qualifié de pauvre, précisément, en tant qu'elle devient de plus en plus dure d'oreille. En 4, Victor est qualifié

par le locuteur de salaud parce qu'il vient de lui boire tout son whisky⁴. Relevons également la présence du datif d'intérêt (*me*) en 4 et la modalité phrastique exclamative dans l'ensemble des énoncés⁵. L'exemple 3 est très proche de 1 : dans les deux cas, nous sommes face à un référent générique (un peuple), et, dans les deux cas, le locuteur qualifie le référent de fou au moyen de la même structure prédicative. En 5, enfin, c'est la subordonnée relative qui reflète de manière explicite l'évaluation dont fait l'objet le référent Bush, celui-ci subissant l'indignation du locuteur parce qu'il a encore bombardé Bagdad. Gary-Prieur conclut au sujet de ce cas de figure que « [le] démonstratif a pour fonction d'attirer l'attention sur le référent du nom propre, et la relative de donner la raison pour laquelle on attire ainsi l'attention sur lui [...] » (1994 : 208). Dans tous les exemples cités *supra*, le démonstratif présente le référent sur fond du contexte étant à l'origine de l'acte d'évaluation.

Rappelons ici encore le fait suivant, également relevé dans la Section 4.5.4, à savoir que ce contexte peut, de même que dans les emplois « désignatifs » du démonstratif, être de plusieurs natures. Ainsi, si le démonstratif dans les énoncés 1 et 3 est responsable du même effet « émotionnel », en 1, l'évaluation se base sur des faits présents au sein de la situation (des représentants du peuple romain viennent d'accomplir, devant le locuteur, une action que ce dernier juge folle), tandis qu'en 3, elle se base sur l'information communiquée par le discours (soit l'obligation, pour les Japonais, de maîtriser un grand nombre de kanjis pour pouvoir lire un journal). En 5, le contexte est le plus probablement de nature mémorielle : il pourrait s'accompagner d'un commentaire du type *tu as entendu ?*, *tu es au courant ?*, etc.

Essayons maintenant de transposer les exemples 1-5 en tchèque :

- 6) *Ti* *Říman-i* *jsou* *blázn-i!*
 ten.M.NOM.PL Romain-NOM.PL être.PRS.3PL fou-NOM.PL
Ils sont fous, ces Romains ! (traduction tchèque par E. Němcová)
- 7) *Podív-ej* *na tu chudink-u Louis-e,*
 regarder-IMP.2SG sur ten.F.ACC.SG (petite) pauvre-ACC.SG Louise-ACC.SG
je *čím dál tím víc nahluchl-á!*
 être.PRS.3SG de plus en plus dur d'oreille-F.NOM.SG
Regarde cette pauvre Louise, elle devient de plus en plus dure d'oreille.

⁴ Remarquons cependant que pour Kleiber, le démonstratif ne s'explique pas ici uniquement par ce comportement ponctuel, car il renverrait aussi à une « qualification générique » : « Nous y voyons également une sorte de généralisation. D'un comportement épisodique du référent, on passe à une qualification stable, habituelle ou générique [...] Victor a-t-il bu tout le whisky du locuteur ? Il sera qualifié de salaud non seulement en tant qu'il a bu le whisky, c'est-à-dire épisodiquement, mais génériquement : c'est un salaud. Le démonstratif marque ainsi le passage d'une manifestation de Victor à une caractérisation de Victor en général. » (Kleiber, 1991a : 95).

⁵ Celle-ci est très souvent présente dans les énoncés accueillant un emploi « émotionnel » du démonstratif, comme nous avons également pu le constater dans la Section 4.5.4 et comme nous le verrons encore dans la Section 7.2.

- 8) Jak to **ty** **Japonc-i** děl-aj?
comment cela.ACC ten.M.NOM.PL Japonais-NOM.PL faire-PRS.3PL

Jsou to blázn-i!
être.PRS.3PL ce fou-NOM.PL

Comment ils font, ces Japonais ? Ils sont fous !

- 9) **Ten** **ničem-a** **Viktor** ! Zas mi vypi-l
ten.M.NOM.SG salaud-NOM.SG Victor encore moi.DAT boire-PST.3SG.M

všechn-u whisky-y!
tout-F.ACC.SG whisky-ACC.SG

Ce salaud de Viktor ! Il m'a encore bu tout le whisky !

- 10) A **ten** **Bush**, zas bombardova-l **Bagdád!**
et ten.M.NOM.SG Bush.NOM.SG encore bombarder-PST.3SG.M Bagdad.ACC.SG

Et ce Bush qui a encore bombardé Bagdad !

L'exemple 6 est directement emprunté aux traductions tchèques existantes d'*Astérix*. En 7 et en 8, c'est un nom qui apparaît : *chudinka* (*petite pauvre*) et *ničema* (*salaud*). L'on voit ainsi que le tchèque connaît une formulation semblable à CE + nom + *de*, à la seule différence près que le nom « de qualité » n'y est pas suivi d'un mot grammatical.

Toutefois, si tous les énoncés français contenant un démonstratif adnominal à valeur « émotionnelle » semblent connaître un équivalent en tchèque, l'inverse n'est pas vrai. Nous avons déjà relevé cette dissymétrie dans la Section 3.5.3, notamment à propos d'une série d'occurrences tchèques où TEN possède une valeur de rappel (Adamec, 1983). Nous en reproduisons une ici :

- 11) L1 : **když** **by** **byl** **[Jindr-a]_i** **voln-ej** . **tak**
si AUX.COND.3SG être.PTCP.SG.M NP-M.NOM.SG libre-M.NOM.SG alors

by **řek-l** **[Jindr-ovi]_i**
AUX.COND.3SG dire-PTCP.SG.M NP-M.DAT.SG

L2 : **no** **protože** **von** **říka-l** **jakoby že že**
MOD parce que lui.NOM dire-PST.3SG.M MOD que

[ten **Jindr-a]_i** .. **von** **vo [něm]_i** **mluvi-l**
ten.M.NOM.SG NP-M.NOM.SG lui.NOM de lui.LOC parler-PST.3SG.M

L1 : si Jindra était libre il demanderait à Jindra

L2 : ouais parce que lui, il disait, genre, que Jindra... il a parlé de lui ORTOFON v1

Rappelons également la lecture qu'à l'instar d'Adamec nous faisons ici de TEN : le locuteur rapporte les paroles d'une personne qui a fait savoir que si un homme prénommé Jindra était libre, elle le contacterait pour lui demander un service. L'emploi du nom propre s'accompagne donc de TEN

pour signaler 1) qu'il a déjà été question de Jindra auparavant, et 2) que c'était précisément en rapport avec le sujet de la conversation en cours, à savoir un service demandé par les interlocuteurs. Comme nous l'avons déjà écrit dans la Section 3.5.3 du Chapitre 3, nous avons conscience que le terme d'« émotionnel » peut ici être contesté, car il n'y va pas de la même émotivité patente que dans les exemples 1-10. Mais nous avons justifié le maintien de ce terme à travers la présence d'une connivence que le recours à TEN permet ici d'installer en ce qu'il active une connaissance partagée. Si cette valeur de TEN est fort courante dans le tchèque parlé informel (voir la Section 7.2.3 *infra*), en français, c'est la question même de son existence qui nous semble se poser. Il nous est pourtant arrivé, ici et là, de rencontrer des occurrences de CE où une telle interprétation nous paraît plausible, voire même la seule possible. Dans la Section 4.5.4, nous avons ainsi identifié, parmi les 11 occurrences « émotionnelles » de notre échantillon de conversations françaises, l'occurrence suivante :

12) c'était quoi *ce nom de cette prof* là

CEFC

Partant du fait qu'aucune pause ne sépare *ce nom* et *de cette prof*, nous sommes arrivé à la conclusion que le démonstratif apparaît sans doute en emploi « émotionnel de rappel ». Récemment, nous avons rencontré un autre usage de CE où une explication en ces termes semble aussi envisageable :

13) Et quand on est en charge d'une globalité, alors ce sont les intérêts de cette nation qui doivent primer. Ce ne sont plus les intérêts de l'humain, ce sont [les intérêts du groupe]. Et [*ces intérêts du groupe*], il faut [les]; prendre en compte. (Wisselmann, 2020)

En 13, le SN démonstratif *ces intérêts du groupe* est coréférent avec le SN défini *les intérêts du groupe*. Or *ces intérêts du groupe* ne saurait, d'après nous, être passible d'une lecture véritablement anaphorique, car son interprétation ne dépend pas du SN défini qui le précède : dans les deux cas, nous sommes face à un emploi générique et indépendant du contexte ; au sein de la seconde occurrence, CE pourrait tout aussi bien être introduit par LE. Par conséquent, si le locuteur a préféré le premier au second, c'est pour insister sur le fait qu'il a déjà été question des intérêts du groupe en même temps que pour thématiser ce référent (à cet égard, il n'est guère étonnant que le SN démonstratif apparaisse en dislocation à gauche)⁶. Malheureusement, hormis ces deux occurrences, nous n'avons pas réussi, jusqu'à présent, à en identifier d'autres.

7.1.2 Les emplois « émotionnels » et la théorie CTD

Comme déjà indiqué à plusieurs reprises dans le présent travail (cf. les Sections 3.5.3 et 4.5.4), les noms introduits par un démonstratif à valeur « émotionnelle » sont soit des noms en emploi

⁶ En 12, le SN démonstratif est placé en dislocation à droite.

générique soit des noms relevant de la définitude « sémantique ». Dans le second cas de figure, il s'agit, par conséquent, de concepts individuels – comme en 2, 4, 5, 7, 9, 10 et 11⁷ – ou fonctionnels – comme en 12, 13 et comme en 14 et 15 *infra* (voir la Section 6.1.3.4 pour la description de ces deux types conceptuels) :

14) L1: a na každ-ý stránc-e je napsán-o dole
 et sur chaque-F.LOC.SG page-LOC.SG être.PRS.3SG écrit-N.NOM.SG en bas

 [moj-e jmén-o a příjmen-í]_i
 mien-N.NOM.SG prénom-NOM.SG et nom-NOM.SG

...
 L2: ale jak tam dosta-l [...] [to **tvoj-e** **jmén-o**
 mais comment là-bas rentrer-PST.3SG.M ten.N.ACC.SG tien-N.ACC.SG prénom-ACC.SG

a příjmen-í_i
 et nom-ACC.SG

L1 : et sur chaque page en bas, il y a [mon prénom et mon nom]_i

...

L2 : mais alors comment il a fait pour y rentrer [ton prénom et ton nom]_i ? ORTOFON v1

15) L1: kdo v-í jestli ty babk-y to to jestli to
 qui.NOM savoir-PRS.3SG si ten.F.NOM.PL petite vieille-NOM.PL cela.NOM si cela.NOM

 tam ještě žij-e
 là-bas encore vivre-PRS.3SG

L2: kter-ý babk-y?
 quel-F.NOM.PL petite vieille-NOM.PL

L1: no **ta** **mám-a** **vod vod ..**
 MOD ten.F.NOM.SG maman-NOM.SG de de

L2: **od Lidk-y** ? . ta ne ta je mrtv-á
 de NP-F.GEN.SG celui-ci.F.NOM.SG non celui-ci.F.NOM.SG être.PRS.3SG mort-F.NOM.SG

L1 : qui sait si toutes ces petites vieilles... si tout ça est encore en vie là-bas

L2 : quelles petites vieilles ?

L1 : ben, la mère de de...

L2 : de Lidka ? pas elle, elle est morte, elle ORTOFON v1

En 14, l'argument possessif des noms fonctionnels *jméno* (*prénom*) et *příjmení* (*nom*) est saturé par le possessif *tvoje* (*ton*). Ce fait nous rappelle que les emplois « émotionnels » autorisent également, en tchèque, la présence de cumuls de TEN avec un autre élément grammatical (« fonctionnel » dans la terminologie d'Abney, 1987). En 15, la saturation est opérée par le complexe prépositionnel au génitif *od Lidky* (*de Lidka*).

⁷ Les noms propres ressortissent à la catégorie des concepts individuels : « Proper names, in particular given names, have a small range of uniform reference. This does not contradict the assumption of semantic uniqueness of proper names since uniqueness is relative to a given context of utterance. » (Löbner, 2011 : 284)

Rappelons que pour ce qui concerne la combinaison du démonstratif avec les noms « sémantiquement » définis, l'effet « émotionnel » a été relevé par Czardybon pour le polonais :

The first reason for the occurrence of *ten* in cases of semantic uniqueness is when the speaker is emotionally affected. This can be a positive or negative emotion such as happiness or anger. In such an emotional context, *ten* can occur regardless of the concept type since *ten* functions to indicate emotional involvement and not unique reference. (2017 : 75)

Voici les deux exemples que l'auteur cite :

16) [...] jak **ta** **Polsk-a** jest pięk-n-a i bogat-a!
comme ten.F.NOM.SG Pologne-NOM.SG être.PRS.3SG beau-F.NOM.SG et riche-F.NOM.SG
Comme la Pologne est belle et riche ! (emprunté à Czardybon, 2017 : 75 ; notre traduction)

17) Że jednak **ta** **Polsk-a** tak-a pojeban-a.
que toujours ten.F.NOM.SG Pologne-NOM.SG tel-F.NOM.SG merdique-F.NOM.SG
Que la Pologne est pourtant si merdique. (emprunté à Czardybon, 2017 : 76 ; notre traduction)

À la fois en 16 et en 17, l'effet « émotionnel » s'apparente à de l'affectivité, tantôt positive (en 16) et tantôt négative (en 17).

Il existe une seule exception, dans les langues slaves occidentales⁸, où un nom à nature conceptuelle « fonctionnelle » peut être introduit par le démonstratif sans que ce dernier soit à l'origine d'un effet « émotionnel ». Cette exception est celle de l'anaphore définie associative, où l'acceptabilité du démonstratif dans son rôle d'article est due à la non-contiguïté syntaxique de l'argument possessif. En effet, comme nous l'avons notamment montré dans la Section 6.2.2.2, cet argument doit être restitué à partir du contexte, ce dernier facteur rapprochant d'ailleurs l'anaphore associative du pôle de la définitude « pragmatique ». En tchèque parlé informel, le cas de l'apparition de TEN avec un nom modifié par un argument possessif (quelle que soit la nature exacte de ce dernier) pose néanmoins des défis considérables, dans la mesure où il n'est pas toujours aisé de trancher entre une interprétation « fonctionnelle » ou non du concept véhiculé par le nom et, par conséquent, entre une lecture « émotionnelle » ou non de TEN. Autrement dit : peut-on dire que TEN contribue à l'identification du référent ? Nous verrons mieux cela dans la Section 7.2 *infra*.

À ces deux types de noms reconnus par la théorie CTD, il faut ajouter les noms génériques (cf. les exemples 1, 3, 6, 8, 13), qui, rappelons-le, ne sont pas traités par cette théorie (voir la Section 4.3). Qu'il s'agisse – comme en français – d'une genericité réduite et souvent mise en place par le discours (voir à cet égard notamment les remarques de Gary-Prieur reproduites dans la Section 4.5.4) ou d'une

⁸ Bien évidemment, à des degrés différents pour chacune de ces langues.

généricité absolue – comme c’est souvent le cas en tchèque –, l’interprétation sémantico-référentielle du complexe nominal contenant ces noms est indépendante du contexte.

7.1.3 Vers un modèle d’explication sémantico-pragmatique des emplois « émotionnels » du démonstratif

Comme annoncé à la fin du Chapitre 3 et au début du chapitre présent, nous tenons à proposer ici une analyse qui s’appuie sur les découvertes de la pragmatique, car cette analyse nous semble opérationnelle et pertinente. Pour ce faire, nous allons nous référer au célèbre article de Grice intitulé « Logic and Conversation », publié en 1975⁹, et aux travaux de Ducrot *Présumés et sous-entendus* (1969), *Dire et ne pas dire, principe de sémantique linguistique* (1972), *Le dire et le dit* (1987 [1984]) et *Les mots du discours* (1980b).

7.1.3.1 La pragmatique gricéenne et ducrosienne : introduction

Avant même d’en venir à constituer une branche à part entière de la linguistique, la pragmatique apparaît au sein de la philosophie analytique en réaction à une crise épistémologique, crise qui est avant tout imputable au clivage qui sépare la logique et les mathématiques du fonctionnement du langage ordinaire, naturel :

L’édifice conceptuel en son entier est mis à mal, contraignant philosophes et théoriciens à se pencher radicalement sur la question du langage et de ses fonctions, condition *sine qua non* de toute activité rationnelle. (Paveau et Sarfati, 2003 : 206)

Dans un premier temps, le principal objectif de la pragmatique est ainsi de réconcilier le langage naturel avec la logique, avec cependant comme axiome de départ que c’est le premier qui doit être adapté à la seconde :

La découverte de Cantor (l’ensemble infini, et l’ensemble des ensembles infinis) a fait la preuve du caractère inévitable de la signification et rendu nécessaire et urgente la construction d’une langue parfaite qui rendrait, à l’avenir, impossible la formulation de pareilles incohérences. Une telle langue assurerait au calcul un fonctionnement stable. (Paveau et Sarfati, 2003 : 206-207)

Ce premier moment de la pragmatique – représenté, entre autres, par Wittgenstein (1986), Frege (1892) et Russell (1905, 1989) – correspond à ce que l’on peut désigner par le terme de « moment logique », voire (non sans une once de dédain) de « réductionnisme logiciste » (Paveau et Sarfati, 2003 : 207). En réaction à ce premier moment, un second moment apparaît qui place au centre de ses investigations le langage ordinaire avec ses formes usuelles. C’est en son sein que s’inscrit Grice¹⁰. Comme le résumait de nouveau Paveau et Sarfati :

⁹ Dans le présent travail, nous citerons l’article dans sa traduction française, parue en 1979 dans la revue *Communications*.

¹⁰ De même qu’Austin.

Entre 1957 et 1969, le philosophe Grice a posé les jalons d'une théorie sémantique et d'une théorie pragmatique complémentaires, l'une et l'autre fondées sur l'hypothèse du caractère intentionnel de la communication. Ces deux contributions se font échos dans la mesure où, une fois définie la notion de signification non naturelle, la théorie de la conversation en spécifie la portée en identifiant des règles fondamentales. (2003 : 216)

Bon nombre des jalons de cette théorie de la conversation sont exposés par Grice dans son étude de 1975. C'est dans ce texte que sont présentés les termes de « principe de coopération », « sens conventionnel » et « implicature », tout comme les quatre « maximes » conversationnelles. Commençons par le premier terme, clé de voûte de l'approche de Grice :

Nos échanges de paroles ne se réduisent pas en temps normal à une suite de remarques décousues, et ne seraient pas rationnels si tel était le cas. Ils sont le résultat, jusqu'à un certain point au moins, d'efforts de coopération ; et chaque participant reconnaît dans ces échanges (toujours jusqu'à un certain point) un but commun ou un ensemble de buts, ou au moins une direction acceptée par tous. Ce but ou cette direction peuvent être fixés dès le départ (par exemple par la proposition initiale de soumettre une question à la discussion), ou bien peuvent apparaître au cours de l'échange ; ils peuvent être relativement bien définis, ou assez vagues pour laisser une latitude considérable aux participants (comme c'est le cas dans les conversations ordinaires et fortuites). Mais à chaque stade certaines manœuvres conversationnelles possibles seraient en fait rejetées comme inappropriées du point de vue conversationnel. Nous pourrions ainsi formuler en première approximation un principe général qu'on s'attendra à voir respecté par tous les participants : que votre contribution conversationnelle corresponde à ce qui est exigé de vous, au stade atteint par celle-ci, par le but ou la direction acceptés de l'échange parlé dans lequel vous êtes engagé. Ce qu'on pourrait appeler principe de coopération (coopérative principle), abrégé en CP. (Grice, 1979 [1975] : 60-61)

En ce qui concerne l'implicature, il convient d'abord de rappeler la nature exacte des distinctions opérées par l'auteur entre 1) la sémantique et la pragmatique et 2) l'acte de *dire* et celui d'*impliquer* (*impliciter*¹¹). Les unités d'un système linguistique sont investies de sens. Du fait d'être codé par le système indépendamment de l'énonciation (du discours), ce sens est appelé « conventionnel » :

Supposons que quelqu'un ait prononcé la phrase suivante : "*He is in the grip of a vice*" : si l'on a connaissance de la langue anglaise, mais pas des circonstances dans lesquelles l'énoncé a été prononcé, on sait bien à peu près ce que le locuteur a dit, en supposant qu'il parle l'anglais standard et n'en fait pas un emploi inhabituel. On saurait qu'il a dit, à propos d'un individu mâle déterminé, humain ou animal, qu'à ce moment précis ce dernier a) soit n'arrivait pas à se débarrasser d'une sorte de mauvais penchant, b) soit qu'une partie de sa personne était coincée dans une sorte d'outil ou de machine (tout cela bien sûr approximativement). Mais pour identifier précisément ce que le locuteur a dit, il faudrait connaître a) l'identité de x, b) le moment d'énonciation

¹¹ Dans le texte original, Grice, afin d'éviter toute ambiguïté propre aux verbes couramment utilisés dans le langage quotidien, et donc en proie à la polysémie, a forgé le néologisme d'*implicate*, rendu, en français, par *impliciter*, le substantif dérivé de ce verbe étant *implicature* (rendu par le traducteur français par *implication*). Il est vrai que ce choix se justifie notamment à travers le fait qu'*impliquer* peut s'entendre comme plus ou moins proche de *dire* (explicitement, c'est-à-dire *affirmer/asserter*) ou de *présupposer*. À cet égard, voir par exemple Strawson : "To say, 'The king of France is wise' is, in some sense of 'imply', to imply that there is a king of France. But this is a very special and odd sense of 'imply'. 'Implies' in this sense is certainly not equivalent to 'entails' (or 'logically implies'). And this comes out from the fact that when, in response to his statement, we say (as we should) 'There is no king of France', we should certainly not say we were contradicting the statement that the king of France is wise." (1950 : 330) De nos jours, c'est le terme d'*implicature* qui est couramment usité en français, d'où notre choix de ce terme plutôt que d'*implication*.

et c) le sens, dans cette situation d'énonciation particulière, de l'expression "*in the grip of a vice*" (une décision entre a et b). (Grice, 1979 [1975] : 59-60)

Or la communication ne se réduit pas à la mobilisation de ce seul sens, conventionnel¹², et c'est là que la notion d'implicature s'introduit. Quant aux implicatures, Grice en distingue deux grands types : l'implicature dite « conventionnelle » et celle dite « conversationnelle ». Voyons d'abord la première :

Dans certains cas, le sens conventionnel des mots utilisés servira non seulement à déterminer ce qui est dit, mais également ce qui est implicite. Si je dis (avec suffisance) : « Il est anglais, il est donc courageux », il ne fait pas de doute que j'affirme, par la vertu du sens des mots que je prononce, que son courage est une conséquence du fait qu'il est anglais. Mais tout en ayant déclaré qu'il est anglais, et qu'il est courageux, je ne veux pas dire par là que j'ai vraiment DIT (au sens fort) que de son anglitude découle son courage, bien que sans aucun doute je l'ai implicite. [...] Donc *certaines* implicatures sont *conventionnelles* [...]

(Grice, 1979 [1975] : 60)

Par conséquent, l'implicature conventionnelle représente une information implicite, bien que véhiculée par le sens linguistique des éléments de la proposition. Contrairement à elle, l'implicature conversationnelle n'est nullement codée par ce dernier, mais se greffe sur lui (l'enrichit) au sein d'une situation d'énonciation concrète. Soit l'exemple suivant, qui sert à Grice d'introduction à la notion :

Supposons que A et B parlent d'un ami commun, C, qui travaille pour le moment dans une banque. A demande à B comment marche le travail de C, et B répond : « Oh, pas mal je crois, il s'entend bien avec ses collègues et on ne l'a pas encore mis en prison. » Là-dessus, A pourrait demander ce que B insinuait, laissait entendre ou même voulait dire quand il déclarait qu'on n'avait pas encore mis C en prison. Beaucoup de réponses seraient possibles, par exemple que C est justement le genre de personne à céder à la tentation que lui offre ce métier, ou que les collègues de C sont vraiment des gens très désagréables et perfides, et ainsi de suite. Il se peut fort bien aussi que A n'ait pas du tout besoin de le demander à B, la réponse étant dans le contexte évidente par avance. Mais il est clair, à mon avis, que ce que B insinuait, laissait entendre, voulait dire, etc., dans cet exemple, est de toute façon distinct de ce que B a dit, à savoir simplement qu'on n'a pas encore mis C en prison. (Grice, 1979 [1975] : 59)

Comme nous pouvons le constater, l'implicature conversationnelle ressortit au discours là où l'implicature conventionnelle ressortit à la langue¹³. De ce fait, elle se distingue de cette dernière par une certaine instabilité, un caractère éphémère, et peut toujours être annulée :

Puisque, pour affirmer qu'il y a une implicature conversationnelle, il nous faut supposer que le principe de coopération est au moins respecté – et puisque, par ailleurs, on peut choisir de ne pas le respecter, il s'ensuit qu'une implicature conversationnelle généralisée peut être annulée éventuellement : elle peut l'être explicitement,

¹² Complété, comme le suggère l'auteur, par la composante contextuelle, indispensable à l'interprétation de la plupart des propositions. C'est là qu'apparaît la singularité des déictiques, qui, comme l'écrivent Ducrot et Todorov en résumant la position de Benveniste, « [...] constituent une irruption du discours à l'intérieur de la langue, puisque leur sens même (la méthode à employer pour trouver leur référent), bien qu'il relève de la langue, ne peut se définir que par allusion à leur emploi [...] ». (Ducrot et Todorov, 1972 : 323)

¹³ Sur le plan de la différence entre la sémantique et la pragmatique, le sens conventionnel et l'implicature conventionnelle ressortissent à la première, alors que l'implicature conversationnelle relève de la seconde. S'agissant de la distinction entre l'acte de *dire* et celui d'*impliciter*, seul le sens conventionnel est *dit*, les deux implicatures étant *implicitées*.

par l'addition d'une proposition affirmant ou sous-entendant que le locuteur s'est mis hors-jeu (du CP) – ou bien elle peut l'être contextuellement, si la forme de l'énoncé qui la produit habituellement est utilisée dans un contexte qui ne laisse pas de doute sur l'intention du locuteur de se mettre hors-jeu. (Grice, 1979 [1975] : 71)

Cette possibilité d'annuler facilement une implicature conversationnelle va directement de pair avec l'option pour le locuteur de la rejeter, et, ce faisant, d'esquiver toute responsabilité de son inférence. Nous reviendrons à cela un peu plus loin, lorsque nous traiterons de l'approche de Ducrot.

Grice explique ensuite que la notion d'implicature conversationnelle ne peut être comprise correctement que si elle est appréhendée en rapport avec l'existence des quatre catégories conversationnelles¹⁴, dont il entreprend la définition dans le même article. Ces catégories – la quantité¹⁵, la qualité¹⁶, la relation¹⁷ et la modalité¹⁸ – constituent autant de fondamentaux du principe de coopération (voir *supra*). L'implicature conversationnelle naît, quant à elle, de la violation ouverte et délibérée de l'une (éventuellement de plusieurs) de ces catégories :

On peut *bafouer* une règle, c'est-à-dire la transgresser *ouvertement* ; si l'on suppose que le locuteur pourrait se conformer à cette règle, et le faire sans pour autant en violer une autre (sans qu'il y ait contradiction), qu'il joue le jeu et qu'il n'est pas en train d'essayer (vu sa façon ouverte d'agir) d'induire les autres en erreur, l'auditeur se voit confronté à un problème de deuxième ordre : comment le fait de dire ce qui a été (effectivement) dit peut-il être concilié avec la présupposition selon quoi le CP a été respecté ? C'est ce type de situation qui donne naissance à l'implication conversationnelle ; et quand une implication est ainsi produite, je dirais qu'il y a, véritablement, *exploitation* de la règle. (Grice, 1979 [1975] : 64)

En résumé : le locuteur procède à la violation d'une catégorie (d'une règle) dans le but de *communiquer* une information tout en ne la *disant* pas. Ainsi, il « joue de la règle du jeu », la soumet à son intention communicationnelle. Au passage, rappelons ici le constat emprunté à Marchello-Nizia et reproduit dans la Section 6.1.2, à savoir qu'une implicature conversationnelle peut, à force d'être réexploitée de façon récurrente, se figer et devenir ainsi de nature conventionnelle :

[...] une implicature dans le discours peut se figer, se conventionnaliser, se 'grammaticaliser' ; ex. : *dès lors que* : valeur temporelle ? valeur logique causale (selon le principe '*post quem, ergo propter quem*', c'est-à-dire 'après X = à cause de X'). (Marchello-Nizia, 2006c : 37)

¹⁴ Ces catégories sont souvent appelées « maximes », mais ce terme est inexact : « Il convient de préciser la tendance d'une certaine vulgate à caractériser ces règles comme "maximes de conversation". Lesdites maximes sont avant tout des catégories de communication. L'emprunt de ces concepts à Kant suggère qu'elles constituent les formes a priori de la logique communicationnelle [...] » (Paveau et Sarfati, 2003 : 217)

¹⁵ « Que votre contribution contienne autant d'information qu'il est requis. Que votre contribution ne contienne pas plus d'information qu'il n'est requis. » (Grice, 1979 [1975] : 61)

¹⁶ « N'affirmez pas ce que vous croyez être faux. N'affirmez pas ce pour quoi vous manquez de preuves. » (Grice, 1979 [1975] : 61)

¹⁷ « Parlez à propos (*be relevant*). » (Grice, 1979 [1975] : 61)

¹⁸ « Évitez de vous exprimer avec obscurité. Évitez d'être ambigu. Soyez bref (ne soyez pas plus prolix qu'il n'est nécessaire). Soyez méthodique. » (Grice, 1979 [1975] : 61-62)

Par conséquent – et comme nous l’avons également souligné dans la Section 6.1.2 –, la conventionnalisation d’une implicature conversationnelle s’avère être un principe particulièrement puissant – sinon *le* principe – de la grammaticalisation. En effet, c’est une fois analysé à l’aune de ce principe que le mécanisme de l’émergence de l’article défini semble gagner en clarté : au départ, le futur défini n’est qu’un démonstratif employé par le locuteur de manière pour ainsi dire redondante. En l’occurrence, cet emploi entre en contradiction avec la « maxime » de quantité, mais l’interlocuteur n’est pas en difficulté pour saisir la pertinence de cette infraction, soit l’insistance sur le statut du référent :

[...] in a language stage where zero determination is the rule for a certain type of nominal reference, in this case definite reference, the use of a demonstrative or another linguistic expression in an article-like role is a conscious strategy by which the speaker intends to externalize for the hearer the discourse status of a referent. (Carlier et De Mulder, 2010 : 253)

Dans la Section 7.1.3.2, nous verrons la façon concrète dont ce procédé opère dans le cas des emplois « émotionnels » du démonstratif adnominal.

Concomitamment au développement de l’approche de H. P. Grice émerge, en France, la « pragmatique intégrée » d’O. Ducrot. Il est d’ailleurs frappant de constater combien les travaux des deux auteurs comportent des ressemblances, mettant souvent au jour les mêmes réalités énoncées dans une terminologie différente.

De même que Grice, Ducrot distingue ce qui relève de la langue de ce qui relève du discours, appelant le premier le « composant linguistique » et le second, le « composant rhétorique » (Ducrot, 1972). De manière générale, ce dernier terme renvoie à l’ancrage d’une proposition dans une situation d’énonciation. L’interprétation des déictiques en fait partie, dans la mesure où elle nécessite la prise en compte du contexte (voir la Section 1.1). Le passage entre le niveau de la langue et celui du discours s’accompagne du passage entre la phrase (proposition)¹⁹ et l’énoncé. Paveau et Sarfati (2003 : 220-221) résumant :

Le composant rhétorique (CR) traite globalement de l’interprétation contextuelle des phrases. L’activité du CR se fonde donc sur une « hypothèse interne » puisqu’à ce stade, la machinerie du sens tient compte de la situation d’énonciation (SE) qui seule permet d’assigner à une phrase (P) le statut d’énoncé à part entière (E) en lui attribuant un sens.

Ducrot procède ensuite à la distinction de la « signification » – propre à la langue – et du « sens » – propre au discours. La « signification » recoupe globalement le « sens conventionnel » de Grice : il s’agit de la composante codée par le système. Le « sens », lui, est obtenu via

¹⁹ Ces deux termes sont ici utilisés comme synonymes, bien qu’une différence les oppose sur le plan de la grammaire.

l'enrichissement de la « signification » par le contexte²⁰. L'analyse du sens met en cause la notion d'« implicite », à savoir, comme chez Grice, de ce qui est communiqué sans pour autant être dit (affirmé). Ici, deux termes sont opposés par l'auteur : le « présupposé » et le « sous-entendu ». La distinction entre ces termes recoupe en bonne partie celle, opérée par Grice, entre l'« implicature conventionnelle » et l'« implicature conversationnelle »²¹ : alors que le présupposé est véhiculé par l'énoncé lui-même, le sous-entendu n'émerge qu'à partir de l'énonciation : « La présupposition est partie intégrante du sens des énoncés. Le sous-entendu, lui, concerne la façon dont ce sens doit être déchiffré par le destinataire. » (Ducrot, 1987 [1984] : 46) Voici l'exemple très bien connu :

18) Pierre a cessé de fumer. (Ducrot, 1972)

Résumons ici l'analyse de cet énoncé, qui, selon Ducrot, véhicule deux informations : 1) Pierre ne fume plus (information communiquée ouvertement, soit le contenu posé) et 2) Pierre fumait auparavant (information communiquée de façon implicite, soit le contenu présupposé)²².

Cependant, d'autres informations peuvent, comme l'explique Ducrot, être inférées par l'interlocuteur à partir de cet énoncé, telles que « C'est bien mieux pour sa santé » ou « Quel dommage, c'était son seul plaisir », inférences dont l'interlocuteur sera le seul à porter la responsabilité directe. En effet, le sous-entendu, pour la raison qu'il ne fait pas partie de la « signification », est annulable par le locuteur (de même que l'implicature conversationnelle gricéenne ; voir *supra*) :

Un locuteur qui aurait prononcé cette phrase [...], peut toujours se retrancher derrière le sens littéral de ses paroles (« je n'ai pas dit ça ») [...]. Le sous-entendu a ainsi la particularité – et l'incalculable avantage – de pouvoir toujours être renié [...] (Ducrot, 1972 : 132)

Voyons à présent comment ces approches peuvent enrichir l'analyse de nos emplois « émotionnels » du démonstratif.

7.1.3.2 La pragmatique gricéenne et ducrosienne : application

Dans la Section 1.2, en traitant des différences les plus flagrantes entre le défini et le démonstratif, nous avons abordé ce qui est communément appelé, en linguistique aussi bien qu'en logique et en philosophie du langage, le « présupposé d'unicité », le « présupposé d'existence » et le « présupposé d'identification » (cf. Donnellan, 1966 ; Hawkins, 1978 ; Kleiber, 1981a ; C. Lyons,

²⁰ Les déictiques appartiennent aux deux à la fois, ce qui permet de voir, une fois de plus, qu'ils articulent langue et discours : bien que possédant une place au sein de la première, leur interprétation ne peut aboutir qu'au sein du second.

²¹ Nous n'allons pas ici entrer dans les différences entre ces deux approches, car cela s'avère sans pertinence pour notre propos.

²² Rappelons que le présupposé échappe à la négation : *Pierre n'a pas cessé de fumer* n'annule pas le présupposé, à savoir que Pierre fumait.

1999 ; Prince, 1981 ; Strawson, 1950). Sans vouloir entrer dans la polémique concernant le statut exact de ces différentes informations véhiculées par le défini et/ou le démonstratif, nous insisterons précisément sur le fait que c'est l'analyse en présupposé qui semble la plus pertinente et, de nos jours, également la plus facilement admise²³. Dans un énoncé tel que 19, l'article défini véhicule le présupposé qu'il existe un et un seul roi de France, éventuellement (mais pas nécessairement) identifiable par l'interlocuteur, au sujet duquel l'on affirme qu'il est chauve :

19) *Le roi de France* est chauve. (Russel, 1905 : 483, traduction)

Ce sens échappe à la négation (voir note 22), car *Le roi de France n'est pas chauve* ne remet pas ouvertement en cause l'existence d'un et d'un seul roi de France, exprimée par LE. De même, en 20, le démonstratif véhicule le présupposé qu'il existe un gâteau (différent à chaque fois) et que ce gâteau peut être identifié par l'interlocuteur²⁴ :

20) Je voudrais *ce gâteau*, et aussi *ce gâteau*. (Gary-Prieur, 2011 : 66)

Et de même qu'en 19, cette information ne tombe pas sous la portée de la négation : *Je ne voudrais pas ce gâteau, ni ce gâteau* ne change rien au fait que le locuteur reconnaît bel et bien l'existence et la possibilité, pour l'interlocuteur, d'identifier un *x* gâteau et un *y* gâteau.

Si ces informations s'appuient, en premier lieu, sur une « signification » des unités CE et LE inscrite dans la langue, ces informations n'émergent qu'en tant que résultat d'une interprétation contextuelle au sein du discours et, de plus, elles ne seraient pas énoncées explicitement mais relèveraient de l'implicite et s'apparenteraient davantage à une implicature conventionnelle dans l'approche de Grice²⁵. Néanmoins, à la fois le défini et le démonstratif sont également susceptibles, dans certains contextes, de porter des implicatures conversationnelles (des sous-entendus). Considérons les deux énoncés suivants, apparaissant à l'intérieur d'un même article :

21) [Matteo Renzi]_i dans le rôle d'un nouveau Machiavel : avec les 2 % d'intentions de vote dont son petit parti, Italia Viva, est crédité, [*le Florentin*]_i a soutenu de l'extérieur le gouvernement de gauche comme la corde soutient le pendu, ne ménageant ni ses recommandations ni ses critiques à l'égard de Giuseppe Conte. (D'Armagnac, 2021 : 36)

22) Il est à présent question de Giuseppe Conte :

²³ Cette analyse n'est pas partagée par les partisans de Russell, ces derniers souscrivant à la thèse exposée par l'auteur dans son étude de 1905 et en vertu de laquelle la communication de ces informations s'apparente davantage à une assertion – un *dit* explicite –, et ferait donc partie, en toute logique, du sens conventionnel vériconditionnel (cf. “Thus when we say ‘*x* was the father of Charles II.’ we not only assert that *x* had a certain relation to Charles II., but also that nothing else had this relation.” (Russell, 1905 : 481-482) Par conséquent, l'énoncé sera considéré comme faux s'il n'y a aucun *x*.

²⁴ La notion d'unicité est, quant à elle, problématique avec le démonstratif, comme nous l'avons vu dans la Section 1.2.1.

²⁵ Ce dernier point n'engage que notre interprétation.

Après avoir dirigé le gouvernement Ligue-M5S, au sein duquel il a promu la loi Salvini sur la sécurité et l'immigration, *ce Rastignac venu des Pouilles* a pris la tête en septembre 2019 d'une coalition diamétralement opposée [...] (D'Armagnac, 2021 : 37)

Dans les deux cas, nous sommes face à un emploi anaphorique infidèle, où l'interprétation des expressions *le Florentin* et *ce Rastignac venu des Pouilles* demande à ce que soit établie par le lecteur la coréférence entre elles et leurs antécédents, *Matteo Renzi* et *Giuseppe Conte*. Or cette coréférence ne va pas de soi, car s'il y a identité de dénotation, il n'y a pas identité de sens (Frege, 1892). Pour cette raison, nous croyons légitime de postuler qu'outre les présupposés mentionnés *supra* (existence, unicité et identification pour le premier SN et existence et identification pour le second), une implicature conversationnelle serait également présente qui porterait sur ce lien de coréférence en tant que tel.

Une analyse possible serait celle selon laquelle cette implicature pourrait découler de l'infraction de la « maxime » de quantité : « Le locuteur me donne l'instruction d'identifier, sur la base des informations contextuelles, un référent correspondant au sens du SN *le Florentin / ce Rastignac venu des Pouilles*. Or il ne me donne pas davantage d'informations sur l'identité de ce référent, et aucun référent parmi ceux qui apparaissent dans le contexte ne correspond *stricto sensu* au sens de ce SN. Tout en agissant de la sorte, le locuteur viole le principe qui lui enjoint de communiquer "autant d'information qu'il est requis" (Grice, 1979 [1975] : 61). Mais je sais aussi que le locuteur est tenu de respecter le "principe de coopération". Par conséquent, il doit tenir pour acquis que je dispose, malgré tout, de tous les outils nécessaires à l'établissement du lien de coréférence. Il se trouve, en même temps, que le discours antérieur a fait mention de la personne de Matteo Renzi / Giuseppe Conte. Après avoir analysé tous les SN éligibles au rôle d'antécédent du SN *le Florentin / ce Rastignac venu des Pouilles* (il se peut qu'il n'y en ait qu'un seul), je constate que c'est celui de *Matteo Renzi / Giuseppe Conte* qui remplit le mieux cette fonction. Le SN *le Florentin / ce Rastignac venu des Pouilles* est donc très probablement, voire certainement, coréférent avec *Matteo Renzi / Giuseppe Conte*²⁶. »

Nous voyons ainsi que le lien de coréférence en question serait, du moins en partie, inféré contextuellement plutôt que communiqué directement, bien que cette inférence se fonde sur le sens implicite présupposé du défini ou du démonstratif. Nous voulons pour preuve de cette inférence la possibilité de son annulation. En effet, il nous semble concevable (bien que difficile), dans une situation comme celle en 21, pour le locuteur de se désolidariser de ladite lecture en coréférence (*Moi,*

²⁶ Cette lecture s'avère la plus opportune (la plus « felicitous » ; cf. Austin, 1962). Par ailleurs, l'implicature conversationnelle qui est ici en jeu requiert davantage d'efforts d'inférence de l'interlocuteur lorsque c'est le défini qui est employé. Cela s'explique à travers le fait que l'ancrage contextuel est plus fort avec le démonstratif (contrairement au défini, celui-ci porte l'instruction « cherche dans le contexte immédiat » ; voir la Section 1.2.1).

je n'ai jamais dit que le Florentin dont je parle était Matteo Renzi !). Le même désengagement de la part du locuteur ne serait probablement pas concevable dans le cas d'une anaphore fidèle.

Un autre contexte où une implicature gricéenne serait en jeu a été mentionné dans le chapitre précédent (voir notamment les remarques de conclusion présentées dans la Section 6.3). Ici, il s'agit de l'effet pragmatique d'insistance sur la nature identifiable d'un référent, au sujet duquel la recherche sur la grammaticalisation semble avoir démontré qu'il caractérise les premières phases du processus (cf. Epstein, 1995 ; Hopper et Traugott, 2003 [1993]). Reproduisons-ici les remarques d'Epstein au sujet de l'évolution du défini en ancien français. Pour ce dernier, l'effet d'insistance que nous décrivons ici est avant tout appréhendé en termes de proéminence du référent²⁷ :

Nous avons montré que l'article défini en AF possède une fonction expressive, à savoir qu'il peut indiquer la proéminence d'un référent. Il peut mettre en valeur un référent dont la proéminence est contextuelle (importance syntagmatique) ou un référent dont la proéminence découle de son statut à l'intérieur d'une catégorie (importance paradigmatique). (1995 : 66)

Là aussi, nous sommes d'avis que l'effet est le fruit d'une inférence (implicature), elle-même le résultat de la violation de la « maxime » de quantité : pourquoi le locuteur utiliserait-il plus d'informations qu'il n'est nécessaire (ici, en l'occurrence, sur la possibilité d'identifier le référent) si ce n'était pour faire passer un autre message, un message du type « veille tout particulièrement à identifier ce référent, qui est d'une importance particulière pour le discours » ?

Passons maintenant aux emplois « émotionnels ». Soit les trois énoncés suivants :

- 23) L1 : ouais mais bon pour le lire c'est vrai que c'est dur avec euh maîtriser mille cinq cents
 L2 : plus même
 L1 : plus de m- plus de mille cinq cents ouais c'est mille cinq cents kanjis pour lire un journal
 L2 : minimum dans les mille euh plus je crois plus
 L1 : hm comment ils font *ces japonais* ils sont fous CEFC

- 24) *Ta* *Ivet-a,* ta mi už jd-e
 ten.F.NOM.SG NP-F.NOM.SG celui-ci.F.NOM.SG moi.DAT déjà aller-PRS.3SG
na nerv-y.
 sur nerf-ACC.PL

(Franchement), Iveta, elle commencer à me casser les pieds.

- 25) to ne v Německ-u jed-u určité cel-ou dob-u
 MOD en Allemagne-LOC.SG roule-PRS.1SG certainement tout-F.ACC.SG temps-ACC.SG
 po dálnic-i tak to není že jo problém
 sur autoroute-LOC.SG MOD ce être-PRS.3SG.NEG MOD problème.NOM.SG

²⁷ Rappelons que dans l'approche d'Epstein vis-à-vis de l'émergence des articles définis, la question de la référence est reléguée au second plan.

do *toho* *Hannover-u*
à ten.M.GEN.SG Hannover-GEN.SG

en Allemagne, c'est sûr que je vais tout le temps rouler sur l'autoroute ; du coup, ça ne sera pas un problème d'aller à Hannover ORTOFON v1

- 26) VĚRA Nechápu, jak ses mohl zaplést s těmi komunisty – BEDŘICH S jakými komunisty? VĚRA No *ten Kohout* a takoví – co ty máš, prosím tě, s nimi společného!

VERA. Qu'est-ce qui a pu te pousser à t'associer avec ces communistes ?! FERDINAND. Ces communistes ? VERA. Eh bien oui... ce Kohout... et bien d'autres...

V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

En 23 (qui est une reprise de l'exemple 3 *supra*), le locuteur recourt au démonstratif, alors même que le référent (les Japonais au sens générique) est identifiable sans ce dernier. En conséquence, il viole la « maxime » de quantité : il fournit à l'interlocuteur plus d'informations que la situation ne l'exige. Néanmoins, le « principe de coopération » veut que cette information ne soit pas redondante (car, si elle l'était, à quoi bon la donner ?). L'inférence qu'en fera l'interlocuteur peut donc être paraphrasée comme suit : « Si le locuteur m'instruit, au moyen du démonstratif, d'identifier le référent en l'indexant sur un contexte alors même que ce dernier est identifiable indépendamment de tout contexte, il doit avoir une raison pertinente de le faire. En effet, il me demande de saisir le référent sur fond d'un contexte qui fonde l'évaluation affective de celui-ci comme étant fou. Tout en m' enjoignant d'appréhender le référent au sein de ce contexte et de nul autre, il me fait part de son jugement affectif concernant les Japonais, jugement qu'il m'invite à partager avec lui²⁸. »

Les exemples 24, 25 et 26 sont analogues, à la seule exception près que les noms y figurant sont des concepts individuels (*unica*). En 24, le contexte sur fond duquel le locuteur invite l'interlocuteur à identifier le référent ne nous est pas accessible ; en 25, l'effet produit par le démonstratif est un effet de rappel, car le démonstratif indexe le référent sur une connaissance contextuelle partagée, à savoir que le locuteur s'apprête à réaliser un voyage à Hanovre ; enfin, en 26, l'effet produit est un effet de distanciation qui surgit en tant que la locutrice donne à son interlocuteur l'instruction de récupérer le référent sur fond du discours, où il est question d'un homme de lettres appelé Kohout²⁹, ami de l'interlocuteur et qu'elle ne connaît pas.

Ces énoncés nous permettent de voir comment le présupposé d'identification, que le démonstratif porte « en temps normal », mue, dans une situation où il s'avère redondant du point de vue de l'identification du référent, en une implicature conversationnelle. Cette mutation s'apparente à un passage du domaine de la sémantique (sens implicite) à celui de la pragmatique. Dans le volet

²⁸ Cette analyse est fortement redevable aux observations de Bonnard (1972 : 1207) et à leur approfondissement de la part de Kleiber (1991a : 94). Ce dernier emploie également le terme d'« implicature gricéenne » (1991a : 85).

²⁹ Auteur communiste tchèque. Opposé à l'intervention des armées du Pacte de Varsovie en 1968, il devient *persona non grata* du régime et finit par s'exiler en Autriche.

suivant, nous présenterons les résultats de l'analyse des emplois « émotionnels » identifiés parmi nos données.

7.2 Les emplois « émotionnels » : analyse des données

Pour rappel, nous avons identifié la présence de 331 occurrences « émotionnelles » dans notre échantillon de conversations tchèques et 50 occurrences « émotionnelles » dans notre sous-corpus des pièces de V. Havel. En termes proportionnels, ces occurrences représentent, respectivement, 33,1 % et 20,75 % de l'ensemble. Hormis 2 occurrences de l'échantillon de conversations, accueillant la forme TENHLETEN, le démonstratif en question est toujours TEN (voir les Sections 5.2.3.1 et 5.2.4.1). Nous avons tâché d'attribuer à chaque occurrence l'une des trois valeurs exposées *supra*. Comme le lecteur s'en doute certainement, cette attribution n'a pas toujours été chose facile, car certaines occurrences ou bien semblent cumuler plus d'une seule valeur, ou bien s'avèrent difficilement analysables faute d'un contexte suffisant (nous avons déjà vu cela dans la Section 6.2.2.3 au sujet des effets « émotionnels » de TEN au sein des « concepts individuels complexes »). Du fait de la haute présence des emplois « émotionnels » dans notre échantillon de conversations, nous nous sommes limité à l'analyse des seules 150 premières occurrences. Le bilan quantitatif eu égard aux différentes valeurs pragmatiques portées par le démonstratif est présenté dans le Tableau 7.1 :

Source de données	Affectivité	Prise de distance	Rappel	Total
ORTOFON v1	16	32	102	150
Sous-corpus de V. Havel	17	14	19	50
Total	33	46	121	200

Tableau 7.1 : Nombre d'occurrences appartenant aux trois types de valeurs pragmatiques exprimées par le démonstratif au sein des emplois « émotionnels » dans l'échantillon de 1 000 occurrences d'ORTOFON v1 (CNT ; 150 premières occurrences) et dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 50 occurrences concernées au total)

Comme déjà écrit *supra*, certaines de ces occurrences posent des défis d'analyse. Outre une hésitation possible entre un emploi « émotionnel » et un autre type d'emploi (voir la Section 5.1.4), la situation peut se présenter où le classement parmi les emplois « émotionnels » ne fait pas de doute, mais où il est difficile d'établir la nature exacte de celui-ci. Si tel était le cas, nous avons essayé d'établir laquelle des trois analyses était la plus plausible. Là où plusieurs valeurs semblaient apparaître à la fois, le classement de l'occurrence a été déterminé par celle qui nous paraissait la plus saillante :

27) a *ten* .. *Cyril* a *Metoděj* přiš-l-i Slovan-ům zvěstovat [...]
et ten.M.NOM.SG Cyril.NOM.SG et Méthode.NOM.SG venir-PST-3PL.M Slave-DAT.PL annoncer
et alors Cyrille et Méthode sont venus pour annoncer aux Slaves [...] ORTOFON v1

- 28) MICHAL A proč vlastně ve skutečnosti tak málo píšeš? Nejde ti to? Nebo jsi snad v nějaké krizi?
BEDŘICH Těžko říct – *ta doba*, to všechno – člověk má takový pocit marnosti.

MICHAEL. *Alors pourquoi est-ce que tu n'écris rien ? Tu as une crise de création ? Tu n'as plus d'idée ?* FERDINAND. *C'est difficile à dire, c'est l'époque, et tout, on a le sentiment que rien n'a un sens, que c'est inutile, que...* V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

En 27, le locuteur, résumant le contenu d'une série sur l'histoire tchèque et faisant part des réflexions que cette série lui inspire, mentionne Cyrille et Méthode, les évangélistes des Slaves occidentaux. Les noms des personnages apparaissent une seule fois dans le dialogue : nous ne pouvons donc pas conclure avec certitude à la présence d'un emploi « de rappel ». Il semble toutefois que cet emploi de TEN en appelle à une information supposée connue par l'interlocuteur (*Cyrille et Méthode, comme tu le sais, ont été missionnés en Moravie pour évangéliser les peuples slaves*) ; en même temps, l'emploi peut également s'expliquer par le profil pas suffisamment accessible du référent pour le locuteur, lequel tente une synthèse d'informations qu'il tient d'une autre source et dont il n'est pas assez familier – nous serions ainsi en présence d'un effet de prise de distance, souvent relatif, comme nous l'avons déjà vu, à la connaissance jugée insuffisante qu'a le locuteur du référent. En 28, l'interprétation se joue entre l'effet de rappel et l'effet d'affectivité. En effet, l'association de TEN avec les noms à sémantique temporelle est susceptible, dans certaines circonstances, de produire un effet « émotionnel », contrairement à l'association de ces mêmes noms avec TENTO/TENHLE, qui, elle, porte souvent la valeur déictique temporelle (cf. la Section 3.3.1)³⁰. En 28, le nom *doba* (*époque*) déboucherait sur cette dernière lecture s'il était introduit par TENTO/TENHLE. La présence de TEN, en revanche, est ici porteuse d'un effet « émotionnel » affectif – le locuteur fait comprendre son désenchantement avec l'époque – et/ou de rappel – *l'époque à laquelle nous vivons, comme tu le sais*³¹.

Si la majorité des occurrences affichent clairement (au moins) une des trois valeurs ici distinguées, pour une minorité d'entre elles (10 occurrences), nous trouvons que si l'une de ces valeurs pourrait être présente, elle serait assez faible, le recours au démonstratif semblant d'abord et avant tout aboutir ici à une focalisation sur le statut thématique-rhématique du référent. L'exemple 29,

³⁰ Lorsqu'ils se prêtent à un emploi déictique temporel, une partie de ces noms peuvent soit être nus soit introduits par TENTO/TENHLE. Ainsi, *tahle doba* ou *tenhle vikend* = *cette époque, ce week-end*, alors que *doba, ponděli* = *l'époque et lundi*. Si les deux peuvent parfois être utilisées de manière interchangeable (cf. également des cas comme *Je le vois lundi* / *ce lundi* en français), la variante avec le démonstratif ressortit à la définitude « pragmatique », tandis que la variante nue appartient à la définitude « sémantique » (le nom est un concept individuel). Cela explique la valeur « émotionnelle » de TEN lorsqu'il apparaît avec ces noms.

³¹ Pour plus d'exemples posant des défis d'analyse, voir la Section 3.5.3 (cf. l'emploi *ta Kamča* dans l'exemple 63).

qui est une reprise de l'exemple 33 du Chapitre 6 et commenté par Kresin, ainsi que l'exemple 30, tiré de nos données, illustrent, d'après nous, ce cas de figure :

29) to je problém tý nevyrovnan-ý společnost-i,
ce être.PRS.3SG problème.NOM.SG ten.F.GEN.SG inégal-F.GEN.SG société-GEN.SG

ta **Prah-a** proud-í rychlejc,
ten.F.NOM.SG Prague-NOM.SG se dynamiser-PRS.3SG plus vite

než se rozvíj-í **ten** **venkov...**
que REFL développer-PRS.3SG ten.M.NOM.SG campagne.NOM.SG

c'est le problème de la société inégale ; Prague se dynamise beaucoup plus vite que se développe la campagne...
(emprunté à Kresin, 2008 : 246 ; modifié)

30) ale vona se mysl-í-m že **ta** **výchov-a** odráž-í
mais elle.NOM.SF REFL penser-PRS-1SG que ten.F.NOM.SG éducation-NOM.SG refléter-PRS-3SG

jakoby ta celkov-á v těch děck-ách jo?
MOD ten.F.NOM.SG global-F.NOM.SG dans ten.N.LOC.PL gamin-LOC.PL MOD

Mais l'éducation, je crois qu'elle laisse des effets, genre, l'ensemble de l'éducation, ça laisse des effets sur les gamins, tu vois ?
ORTOFON v1

Pour 29, reproduisons ici l'analyse de Kresin, à laquelle nous souscrivons :

In this example, *ten* highlights the focal points of the contrast: Prague and the countryside. Word order also reinforces the focus on *Praha*: as in this example, it is usually mentioned either clause-initially, as the point of departure, or clause-finally, as the most communicatively dynamic part of the clause. (Kresin, 2008 : 246)

En 30, c'est l'éducation des enfants qui est présentée par la locutrice comme l'information thématisée, mise au premier plan. Regardons à présent l'une après l'autre les trois valeurs « émotionnelles » que le démonstratif tchèque est susceptible de produire.

7.2.1 Les emplois « émotionnels » présentant un effet d'affectivité

Ce type d'emplois « émotionnels », dont nous estimons la part à 16 occurrences au sein des 150 occurrences « émotionnelles » de l'échantillon de conversations (soit 10,67 % du tout) et à 17 dans le sous-corpus des pièces de Havel (soit 34 % du tout ; voir le Tableau 7.1 *supra*), est le moins représenté des trois dans nos données. Il tend à s'accompagner de certains indices lexicaux et grammaticaux. Ces indices sont notamment : 1) la présence d'adjectifs qualificatifs et d'adverbes exprimant une évaluation positive ou négative – *dobrý* (bon), *strašně nádherný* (terriblement beau/splendide), *pěkně* (joliment), *posranej* (merdeux) ; 2) la présence d'adverbes et de particules modalisateurs – *prostě* (tout simplement), *jakoby* (genre), *dovopravdy* (vraiment), *jako* (genre), *holt* (eh bien), *ty jo* (ça alors), *no jo* (ben ouais) ; 3) la présence de commentaires de modalisation – *jako*

říkám (je dis, genre), jako dobrý jo (très bien, genre) et 4) la présence de la modalité phrastique exclamative (marquée par la prosodie à l'oral)³² :

- 31) Jen kdyby mne ta hlav-a
seulement si.CONJ+AUX.COND.3SG moi.DAT ten.F.NOM.SG tête-NOM.SG

přesta-l-a bolet!
cesser-PTCP-SG.F faire mal.INF

Si seulement je n'avais plus mal à la tête ! (Mathesius, 1926 : 40)

- 32) L1: nó . a to už je vlastně sedm let . přes sedm a půl rok-ů
MOD et ce déjà être.PRS.3SG MOD sept an.GEN.PL plus de sept et demi.NOM.SG an-GEN.PL

L2: no jo no .. holt **ten** čas běž-í no
MOD ten.M.NOM.SG temps.NOM.SG passer-PRS-3SG MOD

L1: ben ouais, et ça, ça fait déjà sept ans, en fait ; plus que sept ans et demi

L2: ben ouais... le temps passe, quoi ORTOFON v1

- 33) protože tam je to strašně nádhern-ý v
parce que là-bas être.PRS.3SG ce terriblement beau-N.NOM.SG dans

těch **Krkonoš-ích**
ten.F.LOC.PL Monts des Géants-LOC

parce que là-bas, c'est terriblement beau, dans les Monts des Géants ORTOFON v1

L'exemple 31 reprend l'exemple 60 du Chapitre 3. La modalité exclamative est présente dans les trois énoncés. Outre cette dernière, les énoncés contiennent les modalisateurs *jen* (*seulement* ; l'exemple 31), *no/nó*³³, *no jo no*, *holt* (l'exemple 32) et l'adjectif évaluatif *nádherný* (*beau, splendide*), précédé de l'intensif *strašně* (*terriblement* ; l'exemple 33). Comme nous avons pu le constater dans la Section 6.2.2.3 lorsque nous avons traité des effets « émotionnels » portés par TEN au sein des « concepts individuels complexes », d'autres indices peuvent également être présents, tels que 5) un lexique à dénotation et/ou connotations positive(s) et 6) une prosodie particulière – une insistance sur le complexe nominal ou sur les adjectifs et adverbes évaluatifs –, voire des éléments non-verbaux. Ainsi, l'énoncé en 31 invite assez intuitivement à insister sur le complexe *ta hlava* (*Jen kdyby mě TA HLAVA přestala bolet!*), alors qu'en 33, la locutrice insiste tout particulièrement sur l'adjectif *nádherný* (allongement de la quantité de la première voyelle /a:/, déjà longue de base, et accent emphatique sur la première syllabe *ná-*).

³² Pour la présence de ces traits, voir également la Section 6.2.2.3.

³³ L'allongement de la voyelle est perçu comme encore plus expressif.

Regardons à présent **la position vis-à-vis du verbe de la proposition** des 16 occurrences de l'échantillon de conversations. Nous recensons : 7 occurrences en position préverbale, 4 en position postverbale, 2 en position de complément de phrase et une seule en position disloquée à droite, à gauche et dans la position inconnue. La position préverbale, celles des deux dislocations et celle du complément de phrase sont clairement les positions qui permettent la thématization du référent (cf. les exemples 31, 32 et 33 *supra*), contrairement à la position postverbale, qui, elle, opère sa rhématisation – nous sommes face à un référent ancré contextuellement (Štícha, 1999 ; cf. l'exemple 34) :

- 34) ne-mě-l-a nic připraven-ýho .. ne-mě-l-a kontakt
NEG-avoir-PST-3SG.F rien.ACC préparé-N.GEN.SG NEG-avoir-PST-3SG.F contact.ACC.SG
- na režisér-y
sur réalisateur-ACC.PL
...
- a na ty co mě-l-a jo .. tak to byl vždycky
et sur celui-ci.M.ACC.PL REL avoir-PST-3SG.F MOD alors ce être.PST.3SG.M toujours
- průser kter-ý to prasi-l .. vyhna-l
fiasco.NOM.SG qui-M.NOM.SG cela.ACC foutre en l'air-PST.3SG.M chasser-PST.3SG.M
- toho** **divák-a** . z toho divadl-a jo
ten.M.ACC.SG spectateur-ACC.SG de ten.N.GEN.SG théâtre-GEN.SG MOD
- elle avait rien de préparé... elle était pas en contact avec les réalisateurs ; et pour ceux avec qui elle était en contact... [avec ceux-là,] ça se terminait toujours par un fiasco qui foutait tout en l'air, qui finissait pas chasser le spectateur du théâtre, tu vois* ORTOFON v1

L'exemple 34 met en scène un acteur et réalisateur critiquant l'incompétence d'une gestionnaire du théâtre où il travaille. Le nom *divák* (*spectateur*) apparaît en emploi générique. La présence de TEN instaure ainsi une focalisation sur ce référent³⁴.

Qu'en est-il de la référentialité des complexes nominaux ? Parmi les occurrences de conversations, toutes sont référentielles sauf une³⁵ ; en revanche, dans le sous-corpus des pièces, nous répertorions 9 occurrences (sur les 17) au sujet desquelles il nous semble éminemment difficile de défendre une quelconque valeur référentielle. De plus, il est à noter que dans 7 cas sur 9, TEN se cumule ici avec un autre élément fonctionnel, le plus souvent NĚJAKÝ (QUELQUE/UN CERTAIN ; 6 occurrences, cf. les exemples 37 et 38 *infra*) :

- 35) takže s tím .. tepl-ým voblečen-ím se to ne-mus-í

³⁴ Idem pour le nom *divadlo* (*théâtre*), employé comme un *unicum* (complexe individuel) : le théâtre en question.

³⁵ Afin de ne pas davantage compliquer les choses, nous considérerons ici les emplois génériques comme référentiels, bien que cela n'aille pas de soi (voir la Section 4.3).

du coup avec ten.F.INS.SG chaud-INS.SG vêtements-INS REFL cela.NOM NEG-devoir-PRS-3SG

přehánět no ale mysl-í-m si že .. že **tu** **bund-u** prostě
exagérer.INF MOD mais penser-PRS-1SG REFL que ten.F.ACC.SG veste-ACC.SG MOD

mít mus-í-te
avoir.INF devoir-PRS-2PL

du coup, pas besoin d'exagérer pour ce qui est des vêtements chauds ; mais je crois qu'une veste, il en faut bien une, quand même ORTOFON v1

- 36) No bóže, **trochu toho deště** na zpáteční cestě!

Oh, mon Dieu, un peu de pluie au retour !

V. Havel, *La fête en plein air*, 1963, trad. par F. Kérel, 1964, InterCorp v10

- 37) A von i ten likvidační úředník má právo na **ten kus toho vopravdu nějak jako plnýho – rozumíte, plnýho – života !**

Et un fonctionnaire de la liquidation a droit lui aussi à une vie, voyons, vous comprenez, à une vie pleine, vraiment pleine ! V. Havel, *La fête en plein air*, 1963, trad. par F. Kérel, 1964, InterCorp v10

- 38) Tuhle Michal je například přímo ideální otec: v podniku se dře, až mi ho je někdy vyloženě líto, jen aby domů přinesl **nějaké ty peníze** – a přitom skoro všechnen volný čas věnuje rodině a domovu!

Michael, par exemple, est tout bonnement le père idéal. Pour ramener un peu d'argent à la maison, il se donne et se dépense dans son entreprise, c'est à faire pitié. Est-ce que ça l'empêche de réserver pratiquement la totalité de ses loisirs à sa famille et son chez soi ? Pas du tout.

V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

Tous ces cas sont fort intéressants. En 35, les locutrices s'entretiennent sur un voyage organisé auquel l'une d'entre elles s'apprête à participer. L'extrait en question s'inscrit dans un passage concernant la météo et les vêtements à emmener. Même si le nom *bunda* (veste) est introduit par TEN, il serait aberrant, selon nous, de plaider ici pour un emploi référentiel défini : en effet, il n'est question d'aucune veste en particulier (fût-ce un référent purement discursif et n'ayant donc aucune existence en dehors du discours³⁶). De plus, le discours ne contient nulle mention du fait qu'il est conseillé d'emmener une veste, ce qui empêche a priori d'analyser l'emploi comme un emploi « de rappel » (voir la Section 7.2.3 *infra*). Par conséquent, l'interprétation la plus plausible serait celle en vertu de laquelle TEN aurait ici pour principale fonction d'insister en la thématisant sur l'information qu'il convient d'emmener une veste.

Une lecture non-référentielle s'impose encore davantage en 36, 37 et 38. Remarquons que dans les trois cas, les traducteurs français ont rendu le complexe nominal par un SN indéfini. En 36, TEN est accompagné du quantificateur TROCHU (*un peu de*). Or, en l'occurrence, nous ne sommes pas ici

³⁶ Voir l'exemple de Prince (1981 : 236-237), déjà cité dans la Section 1.2.1 : *If Mary finds [a pink truck]_i, she'll buy [it]_i, but I don't think she'll find one.*

en présence d'une construction partitive (*un peu de la/cette pluie*) : il s'agit bel et bien d'un complexe nominal indéfini, comme l'atteste la traduction française (*un peu de pluie*), TEN ne servant qu'à marquer l'attitude affective du locuteur. La même chose se produit en 38, où TEN se combine avec NĚJAKÝ : *nějaké ty peníze* correspond à *un peu d'argent*, soit à un emploi indéfini non-spécifique (non-référentiel), enrichi de la marque d'un engagement émotionnel du locuteur. Quant à 37, si TEN y apparaît seul, il participe d'une construction s'inscrivant dans le verbiage creux d'un apparatchik communiste³⁷, rempli de répétitions, de mots « étoffés » et de tours parasites. Le suremploi de TEN adnominal (souvent doté d'une valeur « émotionnelle ») représente un trait proéminent de ce type de discours : remarquons que TEN apparaît également avec *toho vopravdu nějak jako plnýho – rozumíte, plnýho – života*³⁸, figurant en position de complément du nom *kus*, et avec *on ten likvidační úředník* (*fonctionnaire de la liquidation*), où il fait partie de la construction ON TEN. En 37, le complexe nominal marqué en gras figure lui aussi dans un emploi indéfini non-spécifique. Contrairement à 35, la principale fonction de TEN en 36, 37 et 38 n'est pas la focalisation de l'information, bien que cette dernière soit également présente (thématisation en 36 et rhématisation en 37 et 38), mais la production d'un effet « émotionnel ». En 37 et 38, cet effet est, comme nous venons de le voir, encore renforcé par le cumul de TEN avec un autre élément fonctionnel.

S'agissant de la séquence (V)ON TEN³⁹, rappelons qu'elle a déjà été relevée dans le Chapitre 3 (exemple 9, Sections 3.3.1 et 3.5.3), où nous avons cité à son sujet la grammaire de Cvrček *et al.* (2015 [2010] : 175), qui y voit une forme composée du démonstratif apparaissant dans la langue parlée. La construction apparaît aussi dans l'exemple 30 *supra* (*ale vona se myslím že ta výchova odráží [...]*). Il s'agit, ici aussi, d'un élément du « ptydepe » bureaucratique, rendant le discours à la fois verbeux, mal construit et vide. Via son usage excessif dans les répliques de l'apparatchik de *La fête en plein air*, l'auteur parvient à la mise en place d'un effet de l'absurde propre au théâtre ioneskien. L'exemple 39 est une continuation de la réplique de 37 (il est toujours question du profil d'un fonctionnaire de la liquidation) :

39) A von i von musí mít nějaký ty lidský vady – vony i vony k němu nějak jako patřej! Doufám, že i vy máte každej *nějakou tu vadu* – jinak bychom si fakt nějak jako nerozuměli, fakt!

³⁷ Il s'agit du fameux « ptydepe » havlien, dont la pièce est truffée : « [une] langue artificielle que seuls quelques initiés sont censés comprendre mais que tous doivent apprendre, évoqu[ant] alors pour le public une satire de la langue de bois des années 1950. » (RUPNIK J., *Encyclopaedia Universalis*, <https://www.universalis.fr/encyclopedie/havel-v/>).

³⁸ Littéralement : *ten kus toho vopravdu nějak jako*
 ten.M.ACC.SG bout.ACC.SG ten.M.GEN.SG vraiment en quelque sorte MOD
pln-ýho – rozum-í-te, pln-ýho – život-a.
 plein-M.GEN.SG comprendre-PRS-2PL plein-M.GEN.SG vie-GEN.SG

³⁹ La variante VON, typique du « tchèque commun », contient la consonne prothétique /v/.

Et il faut qu'il ait aussi ses faiblesses humaines... des vices, voyons, qui font partie de la vie ! J'espère bien que vous avez vos vices vous aussi, parce qu'autrement, c'est clair, on pourrait pas s'entendre, c'est clair ! V. Havel, *La fête en plein air*, 1963, trad. par F. Kérel, 1964, InterCorp v10

Mais quelle est la nature exacte de ce tour ? Comme nous le voyons dans les exemples 30, 37 et 39, les deux éléments peuvent être séparés par la syntaxe : **vona se myslím že ta výchova odráží [...]**, **A von i ten likvidační úředník má právo na ten kus toho vopravdu nějak jako plnýho – rozumíte, plnýho – života !**. Ensuite, ce n'est pas toujours TEN qui en constitue le second élément : il peut également s'agir d'un complexe nominal sans TEN⁴⁰, voire – même si cela constitue clairement un cas extrême et une violation flagrante du bon usage – d'une reduplication de ON (l'exemple 39). Le tour occupe toujours la fonction sujet et agit sur le plan de la présentation de l'information (la structure thème-rhème), les locuteurs y recourant pour opérer un type spécifique de thématization du référent. L'information ainsi présentée accède au statut de thème, bien qu'à l'échelle de la thématization, elle occupe une position en-dessous de la variante sans ON. En effet, les complexes nominaux *ona ta výchova* et *on ten likvidační úředník* en gras en 30 et 37 sont moins thématiques que ceux en 40 et 41 *infra*, dans la mesure où l'introduction du thème par TEN y est retardée, celui-ci étant d'abord anticipé par ON. Ce procédé aboutit à un effet de suspense :

40) ale	mysl-í-m	že ta	výchov-a	se	odráž-í
mais	penser-PRS-1SG	que ten.F.NOM.SG	éducation-NOM.SG	REFL	réfléter-PRS-3SG
jakoby	ta	celkov-á	v těch	děck-ách	jo?
MOD	ten.F.NOM.SG	global-F.NOM.SG	dans ten.N.LOC.PL	gamin-LOC.PL	MOD

Mais je crois que l'éducation laisse des effets, genre, l'ensemble de l'éducation, ça laisse des effets sur les gamins, tu vois ?

- 41) **i ten likvidační úředník** má právo na ten kus toho vopravdu nějak jako plnýho – rozumíte, plnýho – života !

Nous nous contenterons ici de ces observations, sans entrer dans les détails d'une analyse syntactico-fonctionnelle plus approfondie.

Les exemples 36, 38 et 39 témoignent également de la possibilité, pour TEN, de se combiner avec un autre élément fonctionnel au sein des emplois « émotionnels ». Au total, nous recensons une seule occurrence de ce type parmi les emplois « émotionnels » affectifs de l'échantillon de

⁴⁰ **Von i likvidační úředník má právo na [...]**. Précisons que dans aucun de ces cas, nous ne sommes face à une dislocation, car les deux éléments font partie de la même structure matrice, n'étant séparés ni par un contraste intonatif ni par une pause. La discontinuité syntaxique éventuelle de la séquence tout comme la possibilité de l'absence de TEN avec le second élément nous semblent remettre en question l'analyse d'ON TEN comme forme composée du démonstratif (Cvrček *et al.*, 2015 [2010] : 175).

conversations⁴¹ et 7 occurrences parmi les emplois « émotionnels » affectifs du sous-corpus de Havel. Dans la première source, il s'agit de la combinaison TEN + possessif ; dans la seconde, les occurrences se répartissent entre NĚJAKÝ TEN (5 occurrences, dont les exemples 38 et 39 *supra*), *trochu toho* (une occurrence, soit l'exemple 36 *supra*) et TEN + possessif (une occurrence). Nous avons déjà commenté les cas de NĚJAKÝ TEN et de *trochu toho* ; voyons à présent celui de TEN + possessif :

42) L1: pak mus-í-m zbalit ty hudebn-í nástroj-e
après devoir-PRS-1SG mettre.INF ten.M.ACC.PL de musique-M.ACC.PL instrument-ACC.PL

L2: to já ti mus-í-m sundat tu tašk-u
MOD moi.NOM toi.DAT devoir-PRS-1SG descendre.INF ten.F.ACC.SG sac-ACC.SG

L1: to mus-í-š no a já se k nim pak mus-í-m
cela.ACC devoir-PRS-2SG MOD et moi.NOM REFL jusqu'à eux.DAT après devoir-PRS-1SG

přeskákat v **tom** **naš-em** **obejvák-u**
arriver en sautant.INF dans ten.M.LOC.SG nôtre-M.LOC.SG salon-LOC.SG

L1 : après, je dois mettre les instruments de musique

L2 : moi, je dois te descendre le sac

L1 : c'est ça ; et moi, après, je dois me déplacer jusqu'à eux [= les instruments de musique] en sautant dans **notre salon** ORTOFON v1

43) Uděláš to pro mě? Vid', že to pro mě uděláš? Na jedinej večer – pak mi už bude dobře – pak už bude všechno jiný – pak už budu vědět, že jsem nežil nadarmo – že **ten můj posranej život** nebyl tak úplně posranej [...]

*Tu le feras pour moi, dis ? N'est-ce pas que tu le feras pour moi ? Rien qu'une petite soirée, et j'irai beaucoup mieux, tu verras, tout sera changé, je saurai que je n'avais pas vécu pour rien et **mon existence merdeuse** ne sera pas si pleine de merde après tout [...]*

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

Contrairement aux exemples 35-39, les complexes nominaux en 42 et 43 réfèrent. Pourquoi sommes-nous en présence d'emplois « émotionnels » ? Une fois de plus, c'est la théorie CTD qui permet le mieux, selon nous, de répondre à cette question. Comme les noms *obejvák* (*salon*) et *život* (*vie*) sont des concepts fonctionnels, leur saturation par un argument possessif suffit à elle seule à enclencher l'unicité. Par conséquent, TEN est ici « de trop » à la fois sur le plan de l'instauration de l'unicité dénotative et sur celui de l'identification du référent. En revanche, il est porteur d'une teinte d'affectivité du locuteur, teinte plus ou moins discrète (en 42) ou prononcée (en 43⁴²). C'est dans la Section 6.1.4 que nous avons insisté sur le fait qu'en tchèque, les « concepts individuels et

⁴¹ Rappelons que nous ne considérons ici que les 150 premières occurrences « émotionnelles » analysées en détail. Au total, l'échantillon contient 20 occurrences « émotionnelles » (tous types confondus) où TEN se cumule avec un autre élément fonctionnel (voir la Section 5.2.3.3 pour la proportion des différentes combinaisons).

⁴² Cela dépend de l'ensemble des facteurs contextuels et cotextuels : en 43, l'affectivité est très renforcée par la présence de l'adjectif *posranej* (*merdeuse*).

fonctionnels lexicaux » n'apparaissent pas avec TEN, sauf si ce dernier est doté d'une valeur « émotionnelle ». Nos observations divergent sur ce point de l'échelle de l'unicité proposée par Löbner (2011 : 320 ; voir la Section 6.1.3.4), où les « concepts fonctionnels lexicaux » (“functional CNPs with explicit definite possessor”) sont classés parmi les emplois situés entre les deux types de définitude et où ils sont même représentés comme étant plus proches du pôle de la définitude « pragmatique » que ne le sont les emplois anaphoriques associatifs.

En 43, le complexe nominal est rendu en français par un SN possessif. Nous avons vu, dans la Section 6.2.1.2, que dans les situations où le nom n'est pas un concept fonctionnel, le rôle du possessif français est triple : a) contribuer à l'établissement de l'unicité dénotative par la saturation de l'argument possessif, b) contribuer à l'établissement de l'unicité dénotative et à l'identification du référent par la saturation d'un argument contextuel⁴³ et c) marquer l'unicité dénotative de l'ensemble⁴⁴. Lorsque le nom est de nature fonctionnelle, en revanche, le possessif ne remplit que les rôles a) et c).

En tchèque, dans les situations où le nom n'est pas de nature fonctionnelle, les rôles b) et c) peuvent être assurés par TEN – reproduisons ici l'exemple 47 de la Section 6.2.1.2 :

44) SLÁDEK Co vy jste to vlastně psal, jestli se můžu zeptat? VANĚK [Divadelní hry]_i – SLÁDEK [Divadelní hry]_i? A [to]_i se hrálo někde na divadle? VANĚK Ano - SLÁDEK Hm. No jo, no jo. Tak [divadelní hry]_i, říkáte.

...
SLÁDEK O čem byly [*ty vaše hry*]_i?

SLÁDEK. Sans indiscretion – c'était quoi que vous écriviez ? VANĚK. [Des pièces de théâtre]_i. SLÁDEK. [Des pièces de théâtre]_i... et alors quelque part, dans un théâtre, on [les]_i jouait ? VANĚK. Oui. SLÁDEK. Hum... [des pièces de théâtre]_i... Ouais, ouais, ouais, ouais, ouais, ouais, dites voir...

...
SLÁDEK. Et vous alors, [vos pièces]_i parlaient de quoi ?

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

Ici, TEN contribue à l'unicité dénotative du nom en restreignant son extension via un ancrage contextuel. Il instruit également l'interlocuteur de la possibilité d'identifier le référent au moyen de cet ancrage (il est porteur du présupposé d'identification). Enfin, il serait porteur du présupposé de maximalité : dans le contexte en question, il y a un et un seul (ensemble) *x* tel que *x* est vos pièces. Là où le nom est de nature fonctionnelle, le possessif tchèque seul remplit les rôles a) et c) (cf. les exemples 42 et 43 *supra*). Par conséquent, lorsque TEN apparaît dans ces contextes, il est superflu eu égard à ces fonctions, d'où sa valeur « émotionnelle », dérivée par inférence pragmatique gricéenne.

⁴³ En a) et en b), le possessif restreint ainsi l'extension du nom.

⁴⁴ Le possessif marque l'extensivité extensive du nom (Wilmet, 1986 : 194 ; cf. la Section 1.2.2)

Quid de la traduction française des emplois « émotionnels » à valeur d'affectivité du sous-corpus de Havel ? Les 17 occurrences sont assez également réparties entre une autre transposition (4), le déterminant possessif (3), le déterminant démonstratif (3), l'article indéfini (3), une absence de transposition (3) et l'article défini (1). Les exemples 36 et 38 *supra* relèvent du premier cas de figure : *trochu toho deště* → *un peu de pluie* ; *nějaké ty peníze* → *un peu d'argent* ; les exemples 39 et 43 *supra* appartiennent au second cas de figure. Les exemples 45, 46, 47 et 48 *infra* relèvent, respectivement, de la transposition par le démonstratif, par l'indéfini, de l'absence de transposition et de la transposition par le défini :

- 45) **Ten Tonda Mašků** je fakt slušnej kluk a fakt to potřebuje – tak ho nemůžem přece nechal ve štychu!
Ce Anton Mashek, c'est un type très bien, il en a besoin, je te jure – on ne peut pas le laisser dans la merde. V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10
- 46) Když vidí prostý muž z ulice, že **nějaký ten vévoda** je vlastně ve skutečnosti člověk jako on – z masa a krve, se svými vášněmi, neřestmi a starostmi – zmocní se ho nutně pocit jakési spřízněnosti a hned si ho víc oblíbí!
Quand un pauvre bougre s'aperçoit qu'un gros bonnet est en réalité un homme comme lui de chair et de sang, avec ses passions, ses vices et ses soucis, il se reconnaît en lui et le bonhomme lui devient tout de suite plus sympathique [...] Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10
- 47) Víte, já nesnáším suchaři, jenž strkají hlavu do písek před takovými problémami, jako je například **ten kus citovýho života!** / littéralement : *ce morceau de vie sentimentale*
Vous savez, je supporte pas les pisse-froid qui se fourrent la tête dans le sable devant des problèmes aussi brûlants, par exemple, que la vie sentimentale !
V. Havel, *La fête en plein air*, 1963, trad. par F. Kérel, 1964, InterCorp v10
- 48) I na **ty ženské** si naleznete čas!
Même pour les nanas vous vous trouvez du temps !
Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

Au total, une autre transposition a été utilisée 3 fois pour rendre un cumul de TEN précédé d'un élément fonctionnel indéfini. Comme nous l'avons vu *supra*, l'ensemble du complexe nominal est, dans ce cas, indéfini. Le français n'autorisant pas la reproduction du cumul en question, les traducteurs ont été obligés de sacrifier la valeur émotionnelle apportée par TEN (cf. les exemples 36 et 38). Le possessif a été utilisé une fois pour rendre le cumul TEN + possessif (l'exemple 43) et une fois pour rendre le cumul NĚJAKÝ TEN (l'exemple 39). Dans les deux cas, la valeur affective est également perdue en français. Le démonstratif apparaît une fois avec un nom propre de personne (l'exemple 45). De par son emploi, le locuteur marque son lien affectif avec le personnage Anton Mashek, son ami. Ici, l'effet affectif est maintenu en français. Les deux autres fois, CE transpose un TEN introduisant un nom commun :

- 49) Nikdy jsem tomu nevěřil, ale teď vidím, jak **to dítě** dává člověku úplně novou perspektivu, úplně novou stupnici hodnot – a člověk náhle začíná chápat, že ze všeho nejdůležitější je, co udělá proa člověk náhle začíná chápat, že ze všeho nejdůležitější je, co udělá pro **to dítě** [...]

Je l'ignorais autrefois, mais maintenant j'en suis sûr. [Un enfant]_i, nous donne une perspective plus juste, une échelle de valeur toute neuve. On a la révélation que rien n'est plus important que [ce petit]_i [...]
V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

- 50) Nač myslíš, že jsme **ten byt** vlastně zařizovali ?

Enfin est-ce que tu crois vraiment que c'est rien que pour nous-mêmes que nous avons arrangé cet appartement ?
V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

En 49, il s'agit, en tchèque, de la répétition d'un référent générique défini (*to dítě*). En français, le traducteur a opté pour un référent générique indéfini (*un enfant*), repris par le discours via une expression anaphorique infidèle (*ce petit*). CE ne porte pas ici de valeur « émotionnelle ». En 50, le démonstratif introduirait un concept individuel : *notre appartement*, soit un « permanently established individual concept » de Gerland et Horn (2010). L'emploi de CE dans la traduction aboutit, selon nous, à la situation où le référent est identifié via le contexte, dans lequel il figure néanmoins comme l'unique représentant de sa classe. Deux manières semblent possibles pour rendre compte de cette stratégie : le locuteur use du démonstratif pour communiquer un effet de subjectivité accrue et/ou pour contraster implicitement l'appartement en question à d'autres appartements. Nous avons déjà rencontré ce cas de figure dans la Section 6.2.1.1 (l'exemple 24 de la madone)⁴⁵.

En ce qui concerne l'indéfini, les trois emplois où celui-ci apparaît dans la transposition sont l'exemple 46, l'occurrence de *ten likvidační úředník* en 37 et la première occurrence de *to dítě* en 49. L'exemple en 48 est le seul où TEN soit rendu par l'article défini. Dans ces deux derniers cas de figure, TEN n'est donc pas transposé en français. En 46, la perte de l'effet « émotionnel », véhiculé par la combinaison NEJAKÝ + TEN, est en partie compensée par l'usage du terme *gros bonnet*, fort connoté, contrairement à *vévoda* (*duc*), qui est un terme parfaitement neutre. Enfin, un mot sur l'absence de transposition. Les 3 occurrences concernées – dont celles des exemples 37 et 47 – apparaissent toutes dans les répliques de l'apparatchik communiste. Dans les trois cas, les traducteurs ont opté pour une formulation un peu moins verbeuse, sacrifiant le nom *kus* (*bout* ; cf. les exemples 37 et 47).

7.2.2 Les emplois « émotionnels » présentant un effet de prise de distance

Le second type d'emplois « émotionnels » est représenté à raison de 32 occurrences parmi les 150 occurrences « émotionnelles » de l'échantillon de conversations (soit 21,33 % du tout) et à raison de 14 occurrences dans le sous-corpus des pièces de Havel (soit 28 % du tout ; voir le Tableau 7.1

⁴⁵ MICHAEL : *Tu peux dire ce que tu veux, mais c'est un chef d'œuvre !* FERDINAND : *Quoi ?* MICHAEL : *Cette vierge gothique.* FERDINAND : *Mmh.* MICHAEL : *As-tu remarqué le champ de tension dramatique qui se forme entre la statue et le sabre turc ?*
V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

supra). Comme les emplois porteurs d'un effet d'affectivité, ceux-ci sont également susceptibles de s'entourer d'indices lexicaux et grammaticaux. Nous notons avant tout : 1) des questionnements sur l'identité du référent, à la fois chez le locuteur et l'interlocuteur – cf. *který babky?* (*quelles petites mamies ?*) dans l'exemple 15 *supra*, où la locutrice interroge l'interlocutrice sur l'identité des petites mamies dont parle cette dernière, ainsi que *s jakými komunisty?* (littéralement : *avec quels communistes ?*) dans l'exemple 26, repris *infra*, où l'interlocuteur interroge la locutrice sur l'identité des communistes avec qui il se serait, selon elle, commis. Ces questionnements s'expliquent aisément par la présence d'aspects inconnus, non-familiers, voire troublants du référent :

- 51) VĚRA Nechápu, jak ses mohl zaplést s těmi komunisty – BEDŘICH S jakými komunisty? VĚRA No **ten Kohout** a takový – co ty máš, prosím tě, s nimi společného!

VERA. Qu'est-ce qui a pu te pousser à t'associer avec ces communistes ?! FERDINAND. Ces communistes ? VERA. Eh bien oui... ce Kohout... et bien d'autres...

V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

Ce sont ces mêmes aspects qui encouragent le recours à d'autres procédés, à savoir : 2) l'ajout de précisions et informations supplémentaires au sujet du référent, ayant pour but une identification plus aisée de celui-ci :

- 52) a ještě k tomu mě-l takov-ý řeč-i jako že .. no že
et encore à propos de cela.DAT avoir-PST-3SG.M tel-F.ACC.PL propos-ACC.PL MOD

ten **Alex** jako. šéf **ten** majitel ..
ten.M.NOM.SG NP.M.NOM.SG MOD patron.NOM.SG ten.M.NOM.SG propriétaire.NOM.SG

ten Španěložid .. že ne-m-á rád
ten.M.NOM.SG Espagnol-juif.NOM.SG que NEG-avoir-PRS.3SG content.M.NOM.SG

ty kancelářsk-ý prác-e
ten.F.ACC.PL de bureau-F.ACC.PL travail-ACC.PL

*et en plus de ça, il a tenu des propos du genre que... genre que **cet Alex**, c'est-à-dire le patron, le propriétaire, le Juif espagnol... qu'il apprécie pas les tâches de bureau* ORTOFON v1

Pour référer à la personne prénommée Alex, dont une certaine distance la sépare sur le plan pragmatique, la locutrice recourt ici à TEN, mais, à part cela, elle fournit aussi trois informations additionnelles sur l'identité et la qualité de ce dernier : *c'est-à-dire le patron, le propriétaire, le Juif espagnol*. Il en va de manière similaire en 53, où la locutrice précise l'identité du référent par le biais d'une relative appositive⁴⁶ :

⁴⁶ Notons au passage que la relative est souvent, comme ici, introduite par la « particule relative » *jak* (cf. Haspelmath, 2001).

53) mně se líbí-l-a *ta* *D*Dan-a* jak tancuj-e
 moi.DAT REFL plaire-PST-3SG.F ten.F.NOM.SG NP-F.NOM.SG REL danser-PRS.3SG
 s Jár-ou .. to to vopravdu .. to je teda –
 avec NP-M.INS.SG ce vraiment ce être.PRS.3SG MOD

moi, j'ai trop aimé cette Dana qui danse avec Jára... ça, c'est vraiment, c'est... ORTOFON v1

En 54, où il est question de la technique de plongée appelée « plongée en apnée », l'interlocutrice cherche ouvertement à identifier le référent via la question *c'est bien la plongée en apnée ?* une fois que ce dernier a été introduit par son nom anglais :

54) L1: já jsem právě včera .. večer vidě-l .. ten
 moi.NOM AUX.PST.1SG justement hier soir voir-PTCP.SG.M ten.M.ACC.SG
 dokument vo tom free diving-u
 documentaire-ACC.SG sur ten.M.LOC.SG free diving-LOC.SG
 L2: to je to voln-ý potápě-n-í ?
 ce être.PRS ten.N.NOM.SG libre-N.NOM.SG plongée-NOM.SG
 L1: jo to je to *voln-ý* *potápě-n-í* ..
 ouais ce être.PRS ten.N.NOM.SG libre-N.NOM.SG plongée-NOM.SG

L1 : moi, justement, hier soir, j'ai vu le documentaire sur le free diving

L2 : c'est bien la plongée en apnée ?

L1 : ouais, c'est la plongée en apnée ORTOFON v1

Grâce aux exemples 53 et 54, nous relevons également 3) la présence d'amorces et d'hésitations (*ta D*Dan-a jak tancuje s Jár-ou...*)⁴⁷ et 4) une présence systématique et répétitive de TEN (en partie due, sans doute, à un phénomène d'attraction ; voir la Section 3.5.3) : aussi en 54 TEN apparaît-il chaque fois où il est question de la technique de plongée, mais aussi avec le nom *dokument* (*documentaire*), dont il marque l'unicité dénotative. Nous enregistrons un autre phénomène, à savoir que 5) TEN apparaît deux fois à la forme du neutre singulier, nonobstant le genre et le nombre du nom qu'il introduit. Le même phénomène se reproduit en 55 :

55) L1: m-á-š ty vyživovac-í povinnost vůči mně?
 avoir-PRS-2SG toi.NOM alimentaire-F.ACC.SG obligation.ACC.SG vis-à-vis moi.DAT
 L2: ne .. to není vůči tobě povinnost
 non ce être.PRS.3SG.NEG vis-à-vis toi.DAT obligation.ACC.SG
 L1: to je vůči tomu dítě-ti
 ce être.PRS.3SG vis-à-vis ten.N.DAT.SG enfant-DAT.SG
 L2: no jasně jenom vůči dítě-ti
 MOD bien sûr seulement vis-à-vis enfant-DAT.SG

⁴⁷ Voir aussi *vlastně ta z* Zdena NP.. ona s nama dělá . / en fait, cette Z* cette Zdena X... elle bosse avec nous...*

- L1: a kdybysme ne-mě-l-i dět-i?
et si.CONJ+AUX.COND.1PL NEG-avoir-PTCP-PL.M enfant-ACC.PL
- L2: tak nic
alors rien.NOM
...
- L1: tak co to co to je **to** **vyživovac-í** **povinnost**
alors que.NOM ce être.PRS.3SG ten.N.NOM.SG alimentaire-F.ACC.SG obligation.ACC.SG

L1 : est-ce que toi, tu as une obligation alimentaire vis-à-vis de moi ?

L2 : non... c'est pas une obligation vis-à-vis de toi

L1 : c'est vis-à-vis de l'enfant

L2 : bien sûr, uniquement vis-à-vis de l'enfant

L1 : et si on avait pas d'enfants ?

L2 : alors rien

...

L1 : du coup, c'est quoi, l'obligation alimentaire ORTOFON v1

En 55, le terme de *vyživovací povinnost* (*obligation alimentaire*) fait l'objet d'une interrogation persistante chez la locutrice, qui cherche à comprendre comment cette obligation s'applique concrètement dans les faits. La distance pragmatique qui la sépare du référent fait en sorte que TEN est au neutre, alors que *povinnost* est un nom féminin. Or le genre neutre va jusqu'à traduire ici les interrogations de la locutrice sur le contenu descriptif du terme : dans la mesure où cet emploi du démonstratif tchèque place au premier plan la composante intensionnelle du nom, il introduit dans le discours une certaine dimension métalinguistique.

Ce sont les emplois à valeur de prise de distance qui contiennent les deux occurrences où la forme du démonstratif n'est pas TEN mais TENHLETEN (voir *supra*) :

- 56) dneska prej mě-l-i hudebk-u a .. [...] hrá-l-i ..
aujourd'hui MOD avoir-PST-3PL éducation musicale-ACC.SG et jouer-PST-3PL.M
- písničk-u z Večerníčk-u a oni si uděla-l-i
chanson-ACC.SG de Večerníček-GEN.SG et eux.NOM REFL fabriquer-PST-3PL.M
- papírov-ý čepic-e .. a hrá-l-i si na Večerníčk-y ..
en papier-F.ACC.PL bonnet-ACC.PL et jouer-PST-3PL.M REFL à Večerníček-ACC.PL
- a prostě .. pouště-l-i si tu písničk-u večerníčkov-ou
et MOD passer-PST-3PL.M REFL ten.F.ACC.SG chanson-ACC.SG de Večerníček-F.ACC.SG
- co zpív-á .. ten .. Karel Černoch
REL chanter-PRS-3SG ten.M.NOM.SG NP.M.NOM.SG
- ...
- no a **tenhleten** **mil-ej** **NP** .. sedě-l v
MOD et tenthleten.M.NOM.SG cher-M.NOM.SG NP.M.NOM.SG être assis-PST-3SG.M à
- lavíc-i .. kouka-l na ně a zacpáva-l si uš-i

il paraît qu'aujourd'hui, ils ont eu leur cours d'éducation musicale et... ils ont joué la chanson de Večerníček, et ils se sont fabriqué des bonnets en papier... et ils ont joué au personnage de Večerníček... et, si tu veux, ils passaient la chanson de Večerníček chantée par (ce) Karel Černoch, et ce cher NP restait assis au pupitre en les fixant et en se bouchant les oreilles ORTOFON v1

En 56, la locutrice raconte la réaction d'un garçon ayant du mal à se socialiser avec ses camarades. Plus particulièrement, il est question du contexte d'un cours d'éducation musicale, où l'ensemble de la classe jouait à Večerníček, soit le personnage d'une fameuse émission tchèque pour enfants à l'âge préscolaire. La locutrice s'étonne du comportement du garçon, refusant de se joindre à ses camarades. Le complexe nominal est prononcé sur un ton ironique et contient l'adjectif qualificatif *milej* (*cher*), lui aussi employé de façon ironique⁴⁸. Combiné à ses deux traits, l'emploi du démonstratif TENHLETEN débouche sur un effet de prise de distance plus prononcé que celui de TEN. Cependant, contrairement à l'exemple 55 *supra*, la distance est ici davantage d'ordre du rejet du référent, en partie à cause de son comportement dans la situation décrite.

S'agissant de **la position par rapport au verbe des occurrences** de l'échantillon de conversations, 17 apparaissent en position postverbale, 6 en position préverbale, 4 en complément de phrase, 2 en dislocation à gauche, 2 en position inconnue et une seule en dislocation à droite. Comme déjà indiqué dans la Section 7.2.1, la première de ces positions permet la rhématisation du référent, à la différence des autres (la position inconnue mise à part), qui permettent sa thématization. La séquence ON + TEN est présente 3 fois : 2 fois dans les conversations et une fois dans les pièces. Si les occurrences où le démonstratif est à l'origine d'un effet de distanciation apparaissent plus fréquemment en position postverbale que celles où le démonstratif produit un effet d'affectivité, nulle conclusion quantitative ne peut être tirée de ce bilan, les nombres étant trop faibles⁴⁹.

À l'opposé de ce que nous avons relevé au sein des emplois « affectifs », nous constatons que parmi les emplois « de distanciation », toutes les occurrences sont référentielles (si tant est, une fois de plus, que l'usage générique puisse être considéré comme tel). Le cumul entre le démonstratif et un autre élément fonctionnel est également plus rare ici : seules 2 occurrences de NĚJAKÝ TEN et une occurrence de TEN + possessif parmi les données de conversations et 3 occurrences de TEN + possessif dans le sous-corpus des pièces⁵⁰ :

57) on je nějak-ej [...] nějak-ej . předsed-a ale já ted'ka

⁴⁸ Ainsi, ce type d'emploi « émotionnel » peut également s'associer à des adjectifs qualificatifs sources d'une évaluation globalement négative, que leur sémantisme soit négatif de base (cf. l'exemple 59 *infra*) ou par antiphrase (l'exemple 56).

⁴⁹ L'application de la fonction de vraisemblance ne révèle pas de différence statistiquement significative.

⁵⁰ Proportionnellement à l'ensemble de la population des deux sources, les cumuls représentent 24,24 % des occurrences à valeur d'affectivité et 13,04 % des occurrences à valeur de prise de distance. La différence n'est pas statistiquement significative, les nombres étant, ici aussi, trop faibles.

lui.NOM être.PRS.3SG quelque-M.GEN.PL président-NOM.SG mais moi.NOM maintenant

ne-v-í-m jestli ..
NEG-savoir-PRS-1SG si

...

nějak-ějch **těch** **palírník-ů** **lihovarník-ů**
quelque-M.GEN.PL ten.M.GEN.PL bouilleur-GEN.PL distilleur-GEN.PL

nebo čehosi takov-ýho
ou quelque chose-GEN tel-N.GEN.SG

lui, c'est un espèce de... président... mais moi, je sais pas maintenant si... [le président des] bouilleurs, distillateurs ou de quelque chose de ce genre ORTOFON v1

- 58) Vy jste všichni moc chytrý, moc dobře to máte spočítaný, moc dobře se o sebe umíte postarat! Principy! Principy! Bodejť byste si je nechránili, **ty svý principy** – vám se totiž výborně zhodnotěj, vy si je totiž výborně prodáte, vy si na nich totiž pěkně vyděláte, vás ty principy totiž živěj – ale co já?

Vos principes et vos convictions. Bien sûr, vous les gardez comme la prune de vos yeux. Vous les placez, vous les vendez, vous les faites fructifier, et qu'est -ce qu'ils vous rapportent bien ! Vous en vivez. Et tu sais ce que je suis là-dedans ?

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

- 59) Kdybych ti nevěřil, tak bych se přece s tebou takhle vůbec nebavil – zvlášť po tom, co jsi dnes dopoledne řekl **tomu svému zatracenému Filchovi!**

Si je ne te faisais pas confiance, je ne discuterais même pas avec toi – surtout après ce que tu as dit à ce pauvre Fauche... Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

Contrairement à ce que nous avons pu observer au sujet de la combinaison NĚJAKÝ TEN parmi les emplois à effet d'affectivité, cette même combinaison introduit en 57 un nom en emploi générique : les distillateurs/bouilleurs. Aussi relève-t-on également que la fonction de l'indéfini NĚJAKÝ se limite ici à l'expression de l'effet de prise de distance avec le référent, fonction déjà assurée par TEN seul, mais renforcée par la présence de NĚJAKÝ. De fait, ce dernier apporte une nuance susceptible d'être résumée comme suit : « je ne sais pas exactement lequel, en fait, je ne sais pas lequel, je cherche le terme approprié, je ne trouve pas l'expression exacte, etc. »⁵¹ Uhlířová (1992 : 251). Par conséquent, à l'opposé de ce qui se passe dans les exemples comme 39 ou 46, NĚJAKÝ n'est pas ici un marqueur d'extensité du nom (cf. Wilmet, 1986)⁵², et l'ensemble du complexe nominal reçoit une interprétation définie :

- 60) **nějak-ý** ten vévod-a
quelque-M.NOM.SG ten.M.NOM.SG duc-NOM.SG
un (certain) duc (l'exemple 46)

⁵¹ Nous traduisons : „nevím přesně jaký, vlastně nevím jaký, hledám vhodný atribut, nenapadá mě přesný výraz“ atd.“

⁵² Quant à TEN, il n'est jamais marqueur d'extensité en emploi « émotionnel ».

vs	nějak-ejch	těch	palírník-ů
	quelque-M.GEN.PL	ten.M.GEN.PL	distillateur-GEN.PL
	<i>des (= de + les) distillateurs</i>	(l'exemple 57)	

Les occurrences où TEN se combine avec un possessif ressortissent à deux cas de figure. En 58, la même analyse que celle proposée dans la Section 7.2.1 s'applique : *principy (principes)* est un concept fonctionnel saturé par l'argument possessif *vaše (votre)* – le démonstratif est donc « de trop » sur le plan de l'extension et de l'extensité. Par contre, en 59, le nom utilisé est un nom propre de personne, et donc un concept individuel ; par conséquent, ce n'est pas seulement TEN qui est « de trop » sur le plan de l'extension et de l'extensité, mais aussi le possessif et l'adjectif qualificatif *zatracený (maudit)*, ces trois éléments participant ensemble de la mise en place d'un effet « émotionnel ».

Pour ce qui concerne **les traductions françaises** des occurrences « émotionnelles » de ce type apparaissant dans les pièces de Havel, c'est le déterminant démonstratif qui est le plus utilisé par les traducteurs, soit dans 8 cas, les autres types de transposition utilisés étant : le déterminant possessif (3 fois), l'article défini (2 fois) et une autre transposition (une fois). Nous observons que dans les 8 cas où il est utilisé, le démonstratif traduit un TEN introduisant un nom propre de personne. Quatre fois sur huit, il est question du personnage Kohout (voir aussi l'exemple 26, repris en 51 *supra*) :

61) Kdo je to vlastně **ten Kohout**?

Pendant que j'y pense, qui est ce Kohout ?

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

Dans la pièce *Audience*, le recours à TEN est quasi-systématique chez le personnage du brasseur (Sládek) lorsque celui-ci renvoie à Kohout, ami de Vaněk (un alter-ego de Havel lui-même). En effet, parmi les 4 occurrences où le brasseur parle de Kohout, il introduit son nom via le démonstratif ; une fois, c'est NĚJAKÝ (UN CERTAIN) qu'il emploie, le nom propre apparaissant ainsi en emploi dénominatif (*un X s'appelant Kohout* ; cf. Kleiber, 2004) :

62) Prej vás navštívil **nějakej Kohout** –

On dit qu'un certain Kohout vous a rendu visite, chez nous.

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

L'usage de TEN en 61 va au-delà de l'expression d'un simple manque de familiarité avec le référent : il témoigne d'une méfiance, voire d'une aversion du locuteur à son égard. Le démonstratif est également utilisé, en français, dans la transposition de la combinaison TEN + possessif en 59 (*tomu svému zatracenému Filchovi* → *ce pauvre Fauche*). Une fois, lorsque TEN introduit un nom propre, les traducteurs ont évité l'association entre le démonstratif et ce dernier en recourant à une

« explication » (cf. Popovič, 1983). Nous avons considéré cette solution comme un autre type de transposition :

- 63) Když je někdo křivák, poznám to už na dálku. *Toho Mlynárika ze spilky* – víte, kterej to je? VANĚK
Ano – SLÁDEK Toho jsem odhad, hned jak se tady objevil! / littéralement : *ce Mlynarik de la mise en bouteille*

Un faux jeton, je le vois venir. Comme ce mec de la mise en bouteilles, Mlynarik, vous voyez qui je veux dire ? VANEK. Oui. SLADEK. Le jour qu'il est arrivé, je savais à quoi m'en tenir.

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

Le déterminant possessif a été utilisé dans la traduction de l'exemple 58 et dans un autre exemple analogue (TEN + possessif), mais aussi dans l'exemple suivant, où il est également question d'un nom propre de personne :

- 64) PEACHUM Jenomže v noci zase uprchl – POLLY (vzrušeně) Vážně? Jak se mu to podařilo?
PEACHUM Zorganizovala to prý Lockitova dcera – POLLY Lucy? PEACHUM Takže teď – pokud s **tou Lucy** už neuprchl – určitě bude hledat nějaký úkryt.

*VOLLARD Oui, mais dans la nuit il s'est évadé... ROSELINE, agitée. Vraiment ? Comment a-t-il pu ? VOLLARD C'est la fille de Tourmakoff, paraît-il... ROSELINE Lucie ? VOLLARD Alors, s'il n'est pas déjà bien loin avec sa **Lucie**, il cherche certainement un abri.*

Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

La non-familiarité du locuteur (Vollard) avec le référent (Lucie) transparaît ici à travers le fait que celui-là n'utilise pas le nom propre pour renvoyer à celle-ci, mais la description définie *la fille de Tourmakoff*. Dès lors que la jeune femme est nommée par l'interlocutrice au moyen de son nom propre, Vollard accepte ce nouveau « pacte », mais n'en recourt pas moins au démonstratif, indiquant ainsi « [...] que la connaissance qu'il a du référent n'est pas parfaite ou totale, que celui-ci présente des aspects inconnus, troubles, mystérieux, etc. » (Kleiber, 1991a : 92).

Les transpositions de la combinaison TEN + nom propre au moyen de CE + nom propre (cf. l'exemple 61) tout comme l'association d'un nom propre avec le possessif (l'exemple 64) permettent de maintenir l'effet « émotionnel » de prise de distance propre à l'original, à la différence des autres traductions, comme celles des exemples 58 (transposition par le possessif) et 63 (autre type de transposition) ou celles des exemples 28 *supra* et 65 *infra* (transpositions par le défini). En 65, TEN introduit le défini sémantique *classe ouvrière* :

- 65) Domovník je přeci jen **ta pracující třída**

*Parce que le concierge fait partie de la **classe ouvrière**...*

Havel, *La Grande roue*, 1975, trad. par I. Palec, 1987, InterCorp v10

Rappelons encore que parfois la valeur de prise de distance semble en partie se combiner avec la valeur de rappel, comme nous l'avons vu grâce à l'exemple 27 *supra*.

7.2.3 Les emplois « émotionnels » présentant un effet de rappel

Dans cette dernière section, nous aborderons les emplois « émotionnels » du démonstratif adnominal tchèque au sujet desquels nous estimons qu'ils sont porteurs d'un effet de rappel. Nous constatons que ce type d'emploi est le plus représenté au sein de nos occurrences « émotionnelles » : 102 occurrences parmi les données des conversations (soit 68 % du tout) et 19 occurrences parmi les données des pièces (soit 38 % du tout ; voir le Tableau 7.1 *supra*). Les indices dont nous enregistrons une présence récurrente dans le co(n)texte sont notamment ceux-ci : tout d'abord et avant tout, 1) la présence d'une mention antérieure de l'événement :

66) SLÁDEK Proč myslíte, že trčím v tyhle díře? Ale to vás jistě nezajímá – VANĚK Zajímá, pane sládku – SLÁDEK Víte, co jsem měl dělat? VANĚK Co? SLÁDEK Sládka v Pardubicích! VANĚK Vážně? SLÁDEK No vidíte, a jsem tady! To jsou paradoxy, co? VANĚK A proč jste tam nešel? SLÁDEK Nechme toho –

...

VANĚK Proč jste nešel do **těch Pardubic**? SLÁDEK Nechme toho –

SLÁDEK. Vous vous êtes pas demandé pourquoi je suis là à moisir dans ce trou ? Je suppose que ça ne vous intéresse pas ? VANĚK. Mais si, monsieur, ça m'intéresse. SLÁDEK. Vous savez la place qu'on m'avait promise ? VANĚK. Laquelle ? SLÁDEK. Brasseur en chef à Pardoubiceh. VANĚK. Vraiment ? SLÁDEK. Ah vous voyez, et en attendant me voilà. Ça c'est un paradoxe aussi, hein ? VANĚK. Et pour quelle raison n'y êtes-vous pas allé ? SLÁDEK. Ah, m'en parlez pas.

...

*VANĚK. Pourquoi n'êtes-vous pas allé à **Pardoubiceh** ? SLÁDEK. M'en parlez pas.*

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

En 66, le brasseur fait vaguement savoir à Vanek qu'il avait jadis reçu la proposition d'accepter un poste de brasseur en chef dans la ville de Pardubice. Plus loin dans le dialogue, lorsque Vanek revient sur cette épisode pour en apprendre plus, il introduit le nom de la ville au moyen de TEN (*Proč jste nešel do **těch Pardubic**?*). Malheureusement, comme cela a déjà été évoqué *supra*, il est des cas où le contexte n'est pas suffisamment large pour donner accès à cette mention, en conséquence de quoi l'existence d'un emploi « de rappel » ne peut être établie avec certitude. Malgré cela, il y a parfois des raisons de croire que même si une mention antérieure explicite de l'événement fait défaut dans le contexte, le souvenir en est stocké dans la mémoire commune des interlocuteurs. C'est cela même que nous avons observé dans l'exemple 113 cité dans la Section 6.2.2.3, où TEN introduit un nom modifié par l'adjectif au superlatif⁵³ :

67) to	jsem	byl-a	<u>ted'ka</u>	v Drážďan-ech	když
MOD	AUX.PST.1SG	être.PTCP-SG.F	maintenant	à Dresde-LOC.PL	quand

⁵³ Et où, rappelons-le, TEN marque l'unicité dénotative du nom.

byl-y	ty	<u>nej-větš-í</u>	slejšák-y
être.PST-3PL.M	ten.M.NOM.PL	SUP-plus grand-M.NOM.PL	saucée-NOM.PL

récemment, j'étais à Dresde, au moment des saucées les plus violentes ORTOFON v1

Bien que nulle mention des saucées ne figure dans le discours, l'adverbe temporel *ted'ka* (*maintenant*) – effectuant un renvoi explicite au cadre de l'énonciation – en appelle à la mémoire récente de l'interlocuteur : celui-ci est supposé savoir que des pluies violentes ont récemment eu lieu. Ainsi, c'est également 2) la présence de déictiques temporels qui peut accompagner les emplois « émotionnels » à valeur de rappel.

Le lecteur aura par ailleurs remarqué que nous ne parlons pas ici de référent, mais d'événement. Ce changement terminologique n'est pas anodin, car nous avons pu constater à plusieurs reprises (cf. les Sections 3.5.3 et 5.1.4) que dans le cas des emplois à valeur de rappel, il n'est pas question du référent seul⁵⁴, mais de l'ensemble d'un événement le concernant, fût-ce uniquement le simple fait qu'il a déjà fait l'objet d'une mention (cf. l'exemple 11 *supra*) ou, plus généralement, que l'interlocuteur est supposé le garder dans sa mémoire active (cf. l'exemple 67 *supra* et l'exemple 68 *infra*) :

68) dneska	právě mi	říka-l-a	Páj-a	že že jako	broker	<u>víš?</u>
aujourd'hui	justement moi.DAT	dire-PST-3SG.F	NP-F.NOM.SG	que MOD	courtier-NOM.SG	MOD
ten	Adam	jak	děl-á	ten	NP	<u>víš</u>
ten.M.NOM.SG	NP.M.NOM.SG	REL faire-PRS.3SG	ten.M.ACC.SG	NP.M.ACC.SG	MOD	
jak	byl	s	náma	na tom	na .. na jejich	svatb-ě
REL être.PST.3SG.M	avec	nous.INS	à ten.N.LOC.SG	à	leur.F.LOC.SG	mariage-LOC.SG

aujourd'hui, justement, il y a Pája qui m'a dit que que... le courtier, tu vois ? cet Adam, qui bosse sur le NP, tu vois ? qui était avec nous à ce... à leur mariage ORTOFON v1

C'est cet événement qui trouve sa traduction, au niveau verbal, dans un commentaire mémoriel condensé ou, pour employer la terminologie d'Adamec (1983 : 157 et 165), dans un « connecteur ad-propositionnel » („adpropoziční konektor“) / une « adjonction parenthétique » („parentetický výraz“) de la sorte de « comme on en a parlé », « comme tu le sais ». Grâce à l'exemple 68, nous voyons que ce commentaire mémoriel implicite, condensé dans l'usage de TEN, peut, le cas échéant, être accompagnés de 3) commentaires à caractère mémoriel explicites, dont le modalisateur *víš* (*tu vois ?* / littéralement : *tu sais ?*), de relatives appositives – le plus souvent introduites par la particule relative *jak* (voir *infra* pour une analyse plus détaillée de cette dernière) –, ou d'autres éléments encore, portant sur l'identité et la qualité du référent. De même qu'avec les occurrences à valeur de prise de distance,

⁵⁴ Si référent il y a (voir *infra*).

ici aussi, ces éléments ont pour objectif de faciliter l'accès au référent. Nous enregistrons également 4) la présence de modalisateurs d'insistance comme *teda* (*alors, eh bien*) :

- 69) a kdy po-jede-te teda . na to dřev-o . s tím děd-ú?
 et quand FUT-aller-2PL alors à ten.N.ACC.SG bois-ACC.SG avec ten.M.INS.SG grand-papa-INS.SG
et alors, c'est quand que vous irez chercher le bois avec grand-papa ? ORTOFON v1

Le lexème *teda* possède, comme nous pouvons nous en apercevoir en 69, un fonctionnement anaphorique en ce qu'il renoue avec un contexte précédent ; l'effet d'insistance émerge en tant que ce contexte est (implicitement) mobilisé. Enfin, comme cela apparaît également en 68 et 69, 5) les emplois « émotionnels » à valeur de rappel entraînent souvent une présence systématique et répétitive de TEN (**ten** *Adam jak dělá ten NP* ; *a kdy pojedete teda na to dřevo s tím dědú*), sachant que toutes les occurrences de celui-ci ne sont pas nécessairement dotées d'une valeur « émotionnelle », mais qu'il peut aussi s'agir d'emplois mémoriels (ici aussi, le phénomène d'attraction joue un rôle important ; voir la Section 3.5.3)⁵⁵. D'autres éléments peuvent être présents qui, d'une manière ou d'une autre, interpellent l'interlocuteur et l'incitent à mobiliser ses connaissances mémorielles. En 70, c'est l'interjection *hele* (*tiens*) :

- 70) hele **ten** **Chávez** zase vyhrá-l .. v **tej**
 tiens ten.M.NOM.SG Chávez.NOM.SG de nouveau gagner-PST-3SG.M à ten.F.LOC.SG

Venezuel-e .. po čtvrtý už
 Venezuela-LOC.SG quatrième fois déjà

Tiens, Chávez, il a encore gagné [les élections] au Venezuela... pour la quatrième fois, déjà
 ORTOFON v1

Arrêtons-nous ici un instant sur la présence de *jak* introducteur d'une subordonnée relative, qui est un phénomène apparaissant dans les contextes accueillant TEN à valeur de rappel (cf. l'exemple 68 *supra*). Comme nous l'avons déjà écrit, *jak* fonctionne souvent, dans le registre informel, comme un relatif. Du fait de ne pas s'accorder en genre, en nombre et en cas avec l'antécédent, il valide la définition d'une particule relative (cf. Haspelmath, 2001). Ayant ces origines dans l'adverbe interrogatif (et exclamatif) de la même forme, *jak* est investi d'une multitude de fonctions, notamment dans la langue informelle : en plus de pouvoir être interrogatif, exclamatif ou relatif (adverbe ou particule), il peut également se substituer aux conjonctions de subordination *když* (*quand/lorsque*) et *že* (*que* incolore, introducteur d'une conjonctive pure), etc. D'autres structures encore peuvent être

⁵⁵ Rappelons ici la distinction entre les emplois mémoriels et « émotionnels » : alors que dans les premiers, le démonstratif contribue directement à l'identification du référent, ce n'est pas le cas dans les seconds.

identifiées dans la langue de tous les jours, telles que celle en 71 (qui est une reprise de l'exemple 62 de la Section 6.2.3.1) :

71) jak	<i>tu</i>	<i>jedn-u</i>	<i>noc</i>	<i>jsme</i>	<i>tam [...]</i>	<i>nějak</i>
	comment ten.F.ACC.SG	un-F.ACC.SG	nuit.ACC.SG	AUX.PST.2PL	là-bas d'une façon	
<i>povídá-l-i</i>	<i>já</i>	<i>ne-v-í-m</i>		<i>dokdy</i>		
bavarder-PTCP-PL.M	moi.NOM	NEG-savoir-PRS-1SG		jusque quand		
<i>tu vois, la nuit, là, où on a bavardé, je sais plus jusqu'à quelle heure...</i>						ORTOFON v1

Ici, *jak* semble fonctionner comme une conjonction de subordination porteuse de la valeur que Šimík et Sláma (2020) appellent « mémorielle » (“recognitional”), car elle introduit une perspective partagée via la réactivation d'un souvenir commun. En effet, il s'agirait d'une structure tronquée où seule la relative a été maintenue et pas la principale : *vzpomeň si, jak jsme tam tu noc nějak povídali* (**rappelle-toi la nuit où on a bavardé**). Comme le constatent Šimík et Sláma (2020) et Karlík et Šimík (2017), même lorsque *jak* se comporte comme une particule relative, la subordonnée qu'il introduit affiche un profil singulier. Ainsi, contrairement aux autres relatifs tchèques, la relative introduite par *jak* est toujours censée permettre à l'interlocuteur d'identifier un référent défini et spécifique, récupéré dans la mémoire commune entre lui et l'interlocuteur. En cela, elle serait porteuse d'un sens conatif (Jakobson, 1963b), servant à interpeller l'interlocuteur pour l'inciter à identifier ce référent sur fond de connaissances/souvenirs partagés :

La fonction de base du relatif *jak* n'est pas de délimiter un ensemble de référents possibles, mais d'identifier un référent dont l'existence est présupposée⁵⁶. (Šimík et Sláma, 2020)

La contribution de *jak* à la signification de la proposition est celle que *jak* active une connaissance commune au locuteur et au destinataire au sujet du référent exprimé par la tête [nominale][...] ⁵⁷
(Karlík et Šimík, 2017)

Aussi n'est-il pas étonnant que les relatives introduites par *jak* tendent à apparaître en contextes mémoriels contenant, par ailleurs, une forte présence de TEN à valeur mémorielle ou de rappel (cf. les exemples 68 et 71 *supra*). Šimík et Sláma (2020) montrent sur des données du CNT que les relatives introduites par *jak* contiennent davantage de démonstratifs mémoriels que celles introduites par la particule relative *co*, elle aussi ressortissant au registre (parlé) informel mais non-porteuse de la même sémantique de perspective partagée (contrairement à *jak*, *co* peut modifier un antécédent indéfini, etc.). Dans nos propres données d'ORTOFON v1, nous avons identifié 20 subordonnées introduites

⁵⁶ Nous traduisons : „Základní funkce *jak*-relativa není vymezit množinu možných referentů, nýbrž identifikovat referenta, jehož existence je předpokládána.“

⁵⁷ Nous traduisons : „Příspěvek *jak* k významu věty je takový, že aktivuje společnou znalost referenta vyjádřeného hlavou ze strany mluvčího a adresáta [...]“

par *jak* et véhiculant une perspective partagée avec nos 331 occurrences « émotionnelles » de TEN. Douze parmi ces 20 subordonnées introduites par *jak* contiennent un TEN adnominal « émotionnel » à valeur de rappel. Ce bilan montre que les deux autres valeurs « émotionnelles » peuvent également s’associer avec cette valeur de perspective mémorielle partagée de *jak* (cf. l’exemple 53 *supra*). Parmi les occurrences mémorielles de notre échantillon d’ORTOFON v1, le nombre de subordonnées qui sont introduites par un *jak* porteur de cette sémantique s’élève même à 31. Précision importante : nous avons recensé non seulement les cas où la subordonnée en question contient une occurrence du démonstratif mémoriel, mais aussi ceux où c’est le complexe nominal démonstratif qui est modifié par la subordonnée, voire même quelques cas où la subordonnée se trouve seulement à proximité de l’occurrence du démonstratif. Ensuite, si ces subordonnées sont souvent des relatives, ce n’est pas toujours ainsi : nous recensons également des cas comme dans l’exemple 71 *supra* (où *jak* introduit une conjonctive à valeur mémorielle) ou la construction suivante, appelée par Šimík et Sláma (2020) l’emploi « mémoriel corrélatif » (“correlative recognitional *jak*”) :

72) jak	m-á-m	<i>takov-ej</i>	<i>ten</i>	<i>nov-ej</i>	<i>šátek</i>
	comment avoir-PRS-1 SG	tel-M.ACC.SG	ten.M.ACC.SG	nouveau-M.ACC.SG	foulard.ACC.SG
<i>ten</i>	<i>bílo stříbrno zlat-ej</i>		tak	tam	šl-a
ten.M.ACC.SG	blanc argenté doré-M.ACC.SG			alors là-bas	aller.PST-3SG.F
				nějak-á	quelque-F.NOM.SG
holk-a	a	mě-l-a	ho	taky	
filles-NOM.SG	et	avoir-PST-3SG.F	lui.ACC	aussi	
<i>tu vois mon nouveau foulard ? le foulard blanc, argenté et doré ? alors là-bas, il a une fille qui est passée et qui avait le même</i>					
ORTOFON v1					

Dans cette structure, qui n’a pas d’équivalent direct dans la syntaxe du français, le terme de « corrélation » convient dans la mesure où la proposition introduite par *jak* n’est pas vraiment une subordonnée, se trouvant davantage dans un rapport de dépendance réciproque vis-à-vis de la seconde proposition (la symétrie de la corrélation est marquée par les lexèmes *jak* et *tak*). Le rôle de la première proposition est d’introduire dans le discours un référent – identifié via une connaissance partagée et dont il a déjà été question avant – pour communiquer à son sujet une information nouvelle moyennant la seconde des propositions.

Comme l’expliquent toujours Šimík et Sláma, lorsque *jak* introduit une relative à antécédent nominal, ce n’est pas uniquement le référent qui constitue une information identifiable via la mémoire commune, mais l’ensemble du contenu de la relative⁵⁸. C’est la raison pour laquelle l’ensemble du contenu véhiculé par les relatives *jak dělá ten NP* et *jak byl s náma na tom na .. na jejich svatbě* (qui

⁵⁸ « La sémantique de *jak* influe non seulement sur la tête de la proposition relative, mais sur l’ensemble de la proposition relative. » Nous traduisons : „Sémantika *jak* ovlivňuje nikoliv jen hlavu relativu, nýbrž celou relativní větu.“ (Šimík et Sláma, 2020)

bosse sur le NP et qui était avec nous à... à leur mariage) en 68 doit être connue de l'interlocuteur. Enfin, exactement de la même manière que l'emploi de TEN à valeur de rappel, les relatives introduites par *jak* peuvent, dans certains contextes, ne présenter aucun apport (direct ou indirect) à l'identification du référent, ayant pour unique fonction de rappeler qu'il a déjà été question de ce dernier dans un contexte qui fait partie de l'univers du discours partagé par le locuteur et l'interlocuteur :

73) Jo a *ta* *tvoj-e* *matk-a*, *jak* jsme se o ní
 MOD et ten.F.NOM.SG tien-F.NOM.SG mère-NOM.SG REL AUX.PST.1PL REFL de elle.LOC.SG
včera bavi-l-i, tak ona do toho sbor-u nakonec
 hier parler-PTCP-1PL.M alors elle.NOM.SG à ten.M.GEN.SG chorale-GEN.SG finalement
 ne-přiš-l-a.
 NEG-venir-PST-3SG.F

Et ouais, alors ta mère (tu te rappelles, on a parlé d'elle hier), finalement, elle est pas venue à la chorale.

En 73, le locuteur recourt à la fois à la relative introduite par *jak* et à TEN introduisant le concept fonctionnel *matka* (mère) pour suggérer à l'interlocuteur a) qu'ils ont parlé de la mère de celui-ci la veille et b) que c'était justement en rapport avec sa venue à la séance de la chorale.

Le bilan de **la place du complexe nominal par rapport au verbe de la proposition** dans l'échantillon d'ORTOFON v1 est le suivant : 46 occurrences en position postverbale, 19 en position préverbale, 13 en position de complément de phrase, 11 en position inconnue, 9 en dislocation à droite et 4 en dislocation à gauche. La position inconnue mise à part, les nombres d'occurrences rhématisantes et thématiques sont donc équilibrés (46 et 45). La construction ON + TEN n'apparaît qu'une seule fois dans les données analysées.

Outre sa présence avec des noms sémantiquement uniques (des concepts individuels, avant tout ; cf les exemples 66, 68 et 69 *supra*) et des noms génériques (cf. l'exemple 74 *infra*), nous recensons 9 cas où TEN introduit un nom ne pouvant être analysé comme référentiel (tous ces cas apparaissent dans l'échantillon de conversations). En 74, où les locutrices s'entretiennent du rapport de leurs compagnons avec l'argent, le nom *peníze* (argent) est introduit par TEN, car il a déjà fait l'objet d'une mention deux fois. Au sein de la mention où il est introduit par TEN, nous penchons vers une interprétation générique du nom : *l'argent en général* (voir aussi notre traduction française) :

74) L1: a prostě *prach-y* mu ale ne-chyb-í .. mně
 et MOD pognon-NOM.PL tantum lui.DAT mais NEG-manquer-PRS-3PL moi.DAT
 furt tvrd-í že si mus-í-m jako .. [...] namlouvat že prostě

toujours dire-PRS.3SG que REFL devoir-PRS-1SG MOD convaincre.INF que MOD

m-á-m peněz dost a že to prostě funguj-e
avoir-PRS-1SG argent.GEN.PL tantum assez et que cela.NOM MOD marcher-PRS.3SG

L2: jo tak to m-á-me stejn-ý . styl . chlap-a akorát
ouais alors MOD avoir-PRS-1PL pareil-M.ACC.SG style.ACC.SG mec-GEN.SG seulement

můj [...] on to m-á prostě vsugerovan-é od
mien.M.NOM.SG lui.NOM cela.ACC avoir-PRS-3SG MOD insinué-N.ACC.SG de

babičk-y .. on ty peníz-e m-á furt ..
grand-mère-GEN.SG lui.NOM ten.M.ACC.PL argent-ACC.PL tantum avoir-PRS-3SG toujours

L1: et, genre, le pognon, c'est pas ce qui lui manque... moi, il arrête pas de me dire que je dois genre [...] me convaincre que j'ai assez d'argent, et que ça marche, tout simplement

L2: ouais, je vois qu'on a le même style de mec, seulement, le mien [...], il tient ça de sa grand-mère... lui, l'argent, il en a tout le temps... ORTOFON v1

En revanche, ce que nous observons en 75 – extrait qui ressortit au récit de l'émission sur la plongée en apnée (voir l'exemple 54 *supra*) –, c'est une apparition systématique de TEN à valeur « émotionnelle », quelle que soit la nature du nom. Ainsi, le fait qu'il a déjà été question de ces réalités dans le discours antérieur, combiné au phénomène d'attraction, fait en sorte que TEN apparaît non seulement avec le nom défini *organisme*⁵⁹, mais aussi avec deux autres noms en emploi indéfini non-spécifique : *hloubka* (profondeur) et *kyslík* (oxygène) :

75) Tam byl i nějak-ej profesor z nějak-é ..
là-bas être.PST.3SG.M même quelque-M.NOM.SG professeur.NOM.SG de quelque-F.GEN.SG

fakult-y [...] a oni říka-l-i že .. že vlastně jakmile se potop-í . a
faculté-GEN.SG et eux.NOM dire-PST-3SG.M que MOD dès que REFL plonger-PRS.3SG et

dosáhn-e určít-é té hloubk-y .. [...] tak ten
atteindre-PRS.3SG certain-F.ACC.SG ten.F.ACC.SG profondeur-ACC.SG alors ten.M.NOM.SG

organism-us prostě začn-e šetřit .. ten kyslík .
organisme-NOM.SG MOD commencer-PRS.3SG économiser.INF ten.M.ACC.SG oxygène.ACC.SG

il y avait aussi, là-bas, un prof d'une fac... et ils disaient que... que, genre, dès qu'il plonge et descend dans une certaine profondeur, c'est l'organisme qui, genre, commence à faire des économies d'oxygène... ORTOFON v1

En effet, ni *hloubka* ni *kyslík* ne réfèrent ici, la présence de TEN se limitant à signaler, au travers d'une implicature conversationnelle, que ce sont autant de réalités dont le locuteur a déjà traité dans son récit, et ce à plusieurs reprises : 18 occurrences du nom *hloubka* précèdent son apparition au sein

⁵⁹ Nous pouvons ici hésiter entre une lecture en emploi anaphorique associatif (du type partie-tout : le plongeur et son organisme) et un emploi « émotionnel ». Vu la réticence du type partie-tout à accueillir TEN (voir la Section 6.2.2.2), ainsi que le ton général de l'énoncé, c'est la seconde lecture qui paraît plus plausible.

de l'expression démonstrative marquée en gras ; pour le nom *kyslík*, il s'agit de 6 occurrences⁶⁰. L'on pourrait nous objecter que le nom *kyslík* en 75 pourrait lui aussi s'interpréter comme figurant en emploi générique, car affichant des proximités avec le nom *peníze* en 74. Il n'a pas toujours été évident, il est vrai, de trancher entre une lecture non-spécifique générique et une lecture non-spécifique non-générique. D'autres questions se posent qui portent, plus généralement, sur la notion de référentialité et qui ont déjà été abordées, dans une certaine mesure, dans le présent travail (cf. la Section 5.1.3). À cet égard, nous nous contenterons de renvoyer à la proposition de Himmelmann, fondée sur une approche discursive de la référence, et qui permet en partie d'apaiser le débat sur le sujet tout en prônant une gradualité en la matière, gradualité qui a à voir avec la possibilité de recourir, pour un référent de discours, à une reprise par le démonstratif :

Der diskursbasierte Referenzbegriff erlaubt graduelle Abstufungen in der Referentialität eines Ausdrucks, die damit zusammenhängen, wie klar ein gegebener Ausdruck der Diskursfunktion *tracking* dient. Als operationalisierbares Kriterium für Diskursreferenz schlage ich vor, die Möglichkeit demonstrativer Bezugnahme zu verwenden. (1997 : 118)

Bien que cette approche connaisse aussi des limites – 1) il n'est pas toujours très clair dans quelle mesure une telle reprise se présente comme acceptable et 2) l'auteur parle de la seule possibilité d'une reprise et non pas d'une reprise *de facto*, qui sont des phénomènes différents⁶¹ –, elle s'avère en partie opérationnelle. Ainsi, nous estimons que *peníze* en 74 peut plus naturellement faire l'objet d'une reprise anaphorique par le démonstratif que *hloubka* et *kyslík* en 75.

En 76, l'absence de référence est encore davantage saillante, car le nom est introduit par la préposition *bez* (*sans*) :

76) ale	ne-vyhrá-l	jenom	v	té	jedn-é	disciplín-ě
mais NEG-gagner-PST.3SG.M seulement dans ten.F.LOC.SG un-F.LOC.SG discipline-LOC.SG						
v	té ..	bez těch		ploutv-í ..		
dans ten.F.LOC.SG		sans ten.F.GEN.PL		palme-GEN.PL		

*mais la seule discipline où il a pas gagné, c'était celle... celle **sans palmes**...* ORTOFON v1

⁶⁰ Quantité des occurrences précédentes ne sauraient, elles non plus, être considérées comme référentielles. Ainsi trouvons-nous, par exemple : *nedostatek kyslíku* (un manque d'oxygène), *daleko víc kyslíku* (beaucoup plus d'oxygène), *dvacet metrů hloubky* (vingt mètres de profondeur), etc., tous apparaissant dans des contextes non-référentiels. La présence de TEN peut également être encouragée ici par deux autres facteurs : 1) il s'agit de propos rapportés – le locuteur relaie les explications du professeur, explications dont il est bien moins familier que ce dernier ; ainsi, la présence de TEN peut ici marquer une certaine hésitation et recherche de termes convenables (d'où également la présence des modalisateurs *vlastně* (*en fait*) et *prostě* (*genre*)) ; et 2) sa présence permet au locuteur d'inclure davantage l'interlocuteur dans sa propre perspective en en appelant à une réalité (prétendument) connue de lui. Nous remercions Mme Nádvorníková pour ces deux remarques.

⁶¹ L'approche risquerait, à notre avis, une certaine circularité : dès lors qu'il y a reprise, elle considère, à juste titre, que l'élément qui en fait l'objet réfère (du moins discursivement). En revanche, là où il n'y en a point, il semble quelque peu hasardeux de dire qu'il aurait pu y en avoir une, car les locuteurs disent ce qu'ils disent (cf. Laury, 1997 : 253).

Il ne s'agit pas ici d'une paire particulière de palmes, mais du simple fait qu'on a déjà parlé d'une discipline où les participants ne sont pas autorisés à porter des palmes. En français, la préposition *sans* introduit dans le discours, comme il est de coutume dans ces situations, un nom à extensité nulle.

Il convient à présent de relever ici la présence de plusieurs autres phénomènes, apparaissant, au sein de nos données de référence, exclusivement avec les emplois « émotionnels » à valeur de rappel. Nous commencerons par un type d'emploi que la littérature nomme, le plus souvent, les « emplois définies faibles » (“weak definites” ; cf. Carlson et Sussman, 2005 ; Corblin, 2013, et Schwarz, 2012). Loin de nous l'idée de nous attarder ici sur ce sujet fort épineux de la sémantique référentielle, l'enjeu sera uniquement d'en pointer les singularités par comparaison avec les emplois génériques, dont il se rapproche à première vue. Il s'agit de cas comme en 77 :

77) Mon fils est à *l'école* et le vôtre *y* est aussi. (Corblin, 2013 : 93)

Plusieurs traits caractérisent ces emplois. Comme l'écrit Corblin, résumant les thèses de Carlson et Sussman (2005) :

a) The identity of the individual school is not under discussion. Weak definites are fine when used “out of the blue”, with no previous information regarding the individuals and the situation referred to in the sentence, and it is very likely that no more information will be asked afterwards as to the particular school my son is in. b) In the continuation [...], it is not implied that my son and yours are in the same school [...] c) Weak readings emerge only in some clausal contexts [...] Only some embedding contexts trigger weak readings [...] d) It is implied in the weak reading of *l'école* [...] that the agent of the sentence is not only located to a particular place (a school), but is there in order to receive the kind of service a school is designed for, i.e. it is implied that my son is a school boy. (2013 : 93-94)

Pour ce qui est du premier point (a), une certaine proximité existe avec les emplois génériques, bien que ces derniers « réfèrent » à l'ensemble d'une classe. Par conséquent, la question de l'identité singulière du référent désigné par l'expression en question au sein d'un énoncé *x* (telle l'identité du référent de l'expression *l'école* au sein de l'énoncé en 77) non seulement ne se pose pas, pour les emplois génériques, mais elle n'a pas, tout simplement, lieu de se poser. Pour cette même raison, b) n'est pas non plus pertinent en rapport avec les génériques. Quant aux points c) et d), ils ne sont pas validés par les génériques⁶². En revanche, les noms *nemocnice (hôpital)* et *práce (travail)* en 78 et 79 valident les quatre points :

⁶² Cf. l'exemple 74 *supra*, où l'interprétation générique du nom *peníze* est loin d'être sujette à des contraintes contextuelles aussi strictes que l'interprétation en défini faible du nom *école* en 77. À cet égard, l'un des exemples cités par Corblin (2013 : 93) est la paire d'énoncés *Pierre va à la banque*. vs *Pierre va à l'immeuble*. Seul le SN *la banque* du premier énoncé peut recevoir une lecture en défini faible. Quant à d), cette contrainte n'est pas non plus présente pour les génériques.

78) ani ne-chtě-l před Vánoc-ama nastúpit do **té**
 même NEG-vouloir-PST.3SG.M avant Noël-INS.PLtantum commencer.INF à ten.F.GEN.SG

nemocnic-e že si chc-e ještě posledn-í Vánoc-e
 hôpital-GEN.SG que REFL vouloir-PRS.3SG encore dernier-F.ACC.PL Noël-ACC.PLtantum

užít doma
 profiter.INF à la maison

il voulait même pas commencer son séjour à l'hôpital, en disant qu'il voulait encore profiter de son dernier Noël à la maison ORTOFON v1

79) Když chodím do **té práce**, nemám na to čas – ani soustředění – / littéralement : *quand je vais au travail (comme vous le savez)*

Je fais mes quarante heures, vous savez, et, de toute façon, j'ai du mal à me concentrer.

V. Havel, *Vernissage*, 1975, trad. par S. Meldegg, 1979, InterCorp v10

Ce type d'emploi n'apparaît, parmi les données que nous avons analysées, qu'avec les emplois « émotionnels » à valeur de rappel (9 occurrences au total).

Le second phénomène que nous enregistrons parmi nos « emplois de rappel » est celui des noms à sémantique temporelle. Ces noms sont le plus souvent employés comme des concepts individuels, renvoyant, dans la situation où ils apparaissent, à un seul et unique référent dont l'identification ne dépend pas d'un contexte immédiat. C'est le cas de *ta doba* (l'époque) dans l'exemple 28 *supra*, tout comme de *od toho února* (depuis février) en 80 et *toho dvacátýho prvního [prosince]* (le vingt-et-un [décembre]) en 81 :

80) já zas su rád že vypad-l-a tady prostě z
 moi.NOM MOD être.PRS.1SG content.M.NOM.SG que se sauver-PST-3SG.F ici MOD de
 toho .. kolotoč-e co tady fakt mě-l-a rok .. jo
 ten.M.GEN.SG tumulte-GEN.SG REL ici MOD avoir-PST-3SG.F an.ACC.SG MOD
 vlastně od někdy od **toho** **únor-a** ..
 MOD depuis MOD depuis ten.GEN.SG février-GEN.SG

pour ma part, je suis content, en même temps, qu'elle ait pu se sauver de ce... tumulte qu'elle traversait ici depuis un an, tu vois... en fait, depuis quelque chose comme février

ORTOFON v1

81) já jsem zase slyše-l-a . že **toho** **dvacát-yho**
 moi.NOM AUX.PST.1SG MOD entendre-PTCP-SG.F que ten.M.ACC.SG vingtième-M.ACC.SG

prvn-ího bud-e na nábrež-í obrovsk-á kalb-a ..
 premier- M.ACC.SG AUX.FUT-3SG sur quai-LOC.SG monstrueux-F.NOM.SG buverie-NOM.SG

moi, j'ai entendu que le 21 [décembre], il va y avoir une buverie monstrueuse sur les quais

ORTOFON v1

En 81, *únor* ne saurait référer qu’au mois de février de l’année où a lieu l’énonciation, de même qu’en 82, *dvacátýho prvního [prosince]* ne saurait référer qu’au 21 décembre le plus proche du moment de l’énonciation (pour n’importe quelle année, il n’existe qu’un seul mois de février et qu’un seul 21 décembre)⁶³. Ces emplois aussi n’apparaissent que parmi les emplois « de rappel » (10 occurrences au total)⁶⁴.

Enfin, le dernier phénomène qui mérite, à notre avis, d’être mentionné est celui des « quantificateurs flottants » (“floating quantifiers”). Comme le résume Šimík (2017), il s’agit d’un

[é]lément quantificateur (un déterminant) possédant la distribution syntaxique d’un adverbe ou d’une prédication seconde, mais se rapportant sémantiquement à un syntagme nominal qui le précède dans l’ordre linéaire et se trouve au sein de la même proposition. Le terme métaphorique de quantificateur « flottant » (*floating quantifier* en anglais) est censé évoquer la position non fixe du q. f. dans la syntaxe du syntagme nominal⁶⁵.

Le phénomène des « quantificateurs flottants » existe également en français, où les lexèmes les plus souvent relevés sont les quantificateurs universels *tous/toutes* et *chacun/chacune* :

82) *Tous* les chapitres sont terminés. → Les chapitres sont *tous* terminés.

Chaque étudiant a lu trois livres. → Les étudiants ont *chacun* lu trois livres. /

Les étudiants ont lu trois livres *chacun*.

(Tellier et Valois, 2006 : 99 et 102)

Le tchèque affiche une liberté considérable quant à la position syntaxique d’une série de quantificateurs, fussent-ils universels ou non (la liberté est toutefois plus grande pour les quantificateurs universels ; voir les exemples 83 et 84) :

83) *Všichn-i* ti *uchazeč-i* si chtě-l-i stěžovat.
tout-M.NOM.PL ten.M.NOM.PL participant-NOM.PL REFL vouloir-PST-3PL.M plaindre.INF
Tous les participants voulaient se plaindre.

→ *Ti uchazeči* si *všichni* chtěli stěžovat.

→ *Ti uchazeči* si chtěli *všichni* stěžovat.

→ *Ti uchazeči* si chtěli stěžovat *všichni*. (Šimík, 2017, modifié ; notre traduction)

84) *Zauja-l-o* nás *pět /hodně* *film-ů*.
impressionner-PST-3SG.N nous.ACC cinq beaucoup film-GEN.PL
Cinq / Beaucoup de films nous ont impressionnés.

⁶³ Contrairement aux expressions telles que *aujourd’hui* et *hier*, déictiques temporelles, les expressions qui nous intéressent ici ne relèvent pas de la deixis mais de la définitude « sémantique », leur référent étant exclusivement tributaire d’un index temporel au sens large (Löbner, 2011 : 284).

⁶⁴ L’occurrence de l’exemple 28 *supra* a finalement été classée parmi les emplois « de rappel », bien qu’elle semble également véhiculer une certaine affectivité du locuteur vis-à-vis de l’époque où il vit.

⁶⁵ Nous traduisons : „Kvantifikační element (determinátor), který má adverbialní či sekundárně predikativní syntaktickou distribuci, který se však sémanticky vztahuje k lineárně předcházející nominální frázi vyskytující se v téže větě. Metaforický termín ‚plovoucí‘ kvantifikátor (z angl. *floating quantifier*) má evokovat neukotvenost p.k. v syntaxi nominální fráze.“

→ Filmů nás zaujalo **pět** / **hodně**.

→ ?? Filmů nás **pět** / **hodně** zaujalo. (Šimík, 2017, modifié ; notre traduction)

Dans une approche syntaxique comme celle de Bošković (2004), le nom (accompagné de ses compléments) associé au quantificateur flottant est généré en tant que complément de celui-ci. Par la suite, le nom remonte plus haut dans la structure, laissant derrière lui le quantificateur (ce phénomène est parfois désigné par le terme anglais de « stranding »)⁶⁶. Il existe des points communs entre les « quantificateurs flottants » et le phénomène appelé « left-branch extraction », très répandu dans les langues slaves en général et reflétant bien leurs propriétés syntaxiques (cf. Bošković, 2005). Ainsi, en 85, c'est le nom *filmy* (*films*) qui remonte seul à gauche pour être thématisé, laissant derrière lui l'adjectif *dánské* (*dannois*) ; en 86, qui est un cas de modalité phrastique interrogative, le complexe nominal *jakou bundu* (*quelle veste*) peut soit remonter tout entier (seule option grammaticale en français), soit c'est l'interrogatif *jakou* qui remonte seul, laissant derrière lui le nom :

85) Sledova-l-i jsme dánsk-é **film-y**.
regarder-PTCP-PL.M AUX.PST.1PL dannois-M.ACC.PL film-ACC.PL
*Nous avons regardé des **films** dannois.*

→ **Filmy** jsme sledovali dánské. (Šimík, 2017, modifié ; notre traduction)

86) **Jak-ou** **bund-u** si vybra-l?
quel-F.ACC.SG veste-ACC.SG REFL choisir-PST.3SG.M
***Quelle veste** a-t-il choisie ?*

→ **Jakou** si vybral bundu?

Parmi les occurrences « émotionnelles » que nous avons examinées en détail dans nos sources, nous constatons la présence de « quantificateurs flottants » exclusivement lorsque le démonstratif est porteur d'une valeur de rappel (4 fois au total). Dans nos annotations, ce cas de figure n'a pas été considéré comme un cumul à proprement parler, car alors que TEN précède directement le nom, le quantificateur en est séparé par la syntaxe :

87) no a tak on tam ne-chod-í sám .
MOD lui.NOM là-bas NEG-aller-PRS.3SG seul.M.NOM.SG
ono [jich]_i tam chod-í **víc** [**téch** **děcek**]_i
cela.NOM eux.GEN là-bas aller- PRS.3SG plus ten.N.GEN.PL gamin.GEN.PL
après, il y va pas tout seul... ils sont plus nombreux à y aller, les gamins ORTOFON v1

88) mně už to připad-á že už
moi.DAT déjà cela.NOM sembler-PRS.3SG que déjà

⁶⁶ Pour plus d'informations sur ce sujet, voir par exemple Dowty et Brodie (1984), Krifka (1990), Bošković (2004) ou, plus particulièrement pour le français, Tellier et Valois (2006).

těch *promoc-i* je **moc**
 ten.F.GEN.PL cérémonie de remise des diplômes-GEN.PL être.PRS.3SG trop

moi, il me semble que les cérémonies de remise des diplômes, il y en a déjà trop ORTOFON v1

À la fois en 87 et en 88, le nom est précédé de TEN, procédé qui signale, comme dans tous les autres emplois du même type, qu'il a déjà été question de la réalité décrite par le nom ou que celle-ci est, plus généralement, censée occuper la mémoire active de l'interlocuteur. Mais en plus de cela, les deux énoncés contiennent un quantificateur (non-universel) flottant : *víc jich* (= *těch děcek*) → *jich [...]* *víc* ; **moc** *těch promoci* → *těch promoci* **moc**. Contrairement à TEN, qui n'a aucun impact sur l'extensité du nom, le quantificateur, laissé en fin d'énoncé après la montée du nom, actualise se dernier dans son extensité : *en parlant de cérémonies de remise des diplômes, je trouve il y en a déjà trop en ce moment*. Ce qui est particulier en 87, c'est la présence d'une dislocation à droite : le complexe nominal *těch děcek* est extrait de la proposition matrice, où il est représenté par le pronom cataphorique *jich* (eux.GEN). Nous remarquons également la présence de *ono* (pronom personnel neutre au nominatif singulier)⁶⁷ en début d'énoncé, qui anticipe/retarde le thème *jich* au sein de la structure matrice (voir la Section 7.2.1).

Après avoir relevé ces trois phénomènes touchant uniquement les emplois « de rappel », nous finirons avec **un bilan relatif à la traduction française** des 19 occurrences du corpus de Havel. Avec ses 13 occurrences, c'est de loin la traduction par l'article défini qui prévaut. Parmi le reste, nous trouvons un autre type de transposition (5 occurrences) et une absence de transposition (une occurrence). Par ailleurs, le corpus des pièces ne contient aucune occurrence « émotionnelle de rappel » où TEN se cumulerait avec un autre élément fonctionnel⁶⁸, contrairement aux 150 occurrences analysées de l'échantillon de conversations, qui en contiennent 4, soit 2 occurrences de TEN + possessif, dont l'exemple 14 *supra*, et 2 occurrences de NĚJAKÝ TEN, dont celui cité en 89⁶⁹ :

89) je-l-a	bych	pozděj do prác-e	k
aller-PTCP-SG.F	AUX.COND.1SG	plus tard à travail-GEN.SG	chez

⁶⁷ Faute de mieux, nous l'avons glosé par *cela* en français. Le fait que le pronom apparaisse au neutre singulier est dû à des règles complexes régissant l'accord morphologique en tchèque. En effet, *víc* (*plus*) fait partie des quantificateurs qui imposent sa morphologie à la fois au nom qu'ils introduisent (de même que *plus* en français), mais aussi au verbe et à l'attribut du sujet de la proposition dont ce nom est le sujet. Pour le nom et pour l'attribut du sujet, la morphologie est celle du génitif ; pour le verbe, c'est celle de la 3^{ème} personne du singulier neutre. Le pronom ON est soumis au même accord que le verbe.

⁶⁸ Ce bilan doit, évidemment, être regardé d'un œil réservé vu la taille limitée de l'échantillon.

⁶⁹ Au total, les cumuls représentent seulement 4,96 % de l'ensemble des 121 occurrences où TEN est porteur d'une valeur de rappel. L'application de la fonction de vraisemblance révèle des résultats statistiquement significatifs : $L = 12,5919$ pour la différence entre les emplois « de rappel » et les emplois à valeur d'affectivité ; $L = 4,9313$ pour celle entre les emplois « de rappel » et les emplois à valeur de prise de distance ; la valeur critique = 3,841 (degré de liberté = 1 ; $\alpha = 0,05$).

temu NP na .. *nějak-é* ty *lék-y* ..
 ten.M.DAT.SG NP.M.DAT.SG pour quelque-M.ACC.PL ten.M.ACC.PL médicament-ACC.PL

je partirais plus tard au travail, chez NP, pour chercher quelques médicaments (comme je disais)
 ORTOFON v1

Comme en 38, 39 et 46 *supra*, seul NĚJAKÝ exprime l'extensité du nom, alors que l'usage de TEN suggère que la locutrice a déjà dit vouloir se rendre à l'endroit en question afin de chercher des médicaments.

Parmi celles des occurrences des pièces où TEN a été rendu par le défini, l'effet « émotionnel » n'est maintenu que dans deux cas, et ce au travers de l'introduction par le défini du nom propre d'une actrice tchèque, J. Bohdalová, que le brasseur incite Vanek à amener à la brasserie pour organiser une soirée avec elle :

90) Ale *tu Bohdalku* přitáhneš, vid' že jo?

Tu vas trouver la Bodalova et tu l'amèneras ici, dis, tu le feras, n'est-ce pas ?

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

Certes, l'effet n'est pas tout à fait le même en français, où ce type d'association du défini avec le nom propre est porteuse de connotations populaires, mais où le défini, contrairement à TEN, ne possède aucun fonctionnement déictique.

S'agissant des autres façons de transposer TEN, nous notons la présence d'un phénomène d'« explicitation » (cf. Popovič, 1983) qui se produit à plusieurs reprises dans les traductions :

91) Proč jste nešel do *těch Pardubic*? /

/ littéralement : *Pourquoi est-ce que vous n'êtes pas allé à (ce) Pardoubiceh ?*

Pourquoi est-ce que vous n'avez pas obtenu ce poste à Pardoubiceh ?

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

92) Kdyby se to podařilo – myslím s *tím skladem* – co by pak bylo se starým Šustrem?

/ littéralement : *je pense à ce magasin*

Si jamais ça marchait – je pense à la place de magasinier – que ferait le vieux Sustre ?

V. Havel, *Audience*, 1975, trad. par M. Aymonin et S. Meldegg, InterCorp v10

De même qu'en 63, cette formulation plus étoffée permet de contourner l'association – beaucoup moins acceptable en français – entre le démonstratif et un nom « sémantiquement » défini.

7.3 Résumé et conclusion

Dans le cadre de ce dernier chapitre de notre travail, nous nous sommes fixé pour objectif d'examiner plus avant les emplois « émotionnels » du démonstratif adnominal tchèque se trouvant dans nos données. L'organisation du chapitre a répondu à notre démarche consistant à séparer les aspects davantage théoriques des aspects davantage pratiques de la problématique. Ainsi, suite à une mise au point terminologique et un rappel de la manière dont les emplois « émotionnels » sont appréhendés par la théorie CTD, nous avons tâché de proposer un traitement théorique unifié de la totalité des emplois « émotionnels » en veillant à articuler les niveaux sémantique et pragmatique. En ce qui concerne le premier, nous avons mis l'accent sur le fait que le démonstratif n'a pas d'incidence sur les propriétés extensionnelles du nom et n'intervient pas de façon directe dans la démarche d'identification du référent là où le nom se prête à un emploi référentiel. Pour ce qui concerne le second, nous avons essayé d'exploiter les pistes identifiées dans les études de Bonnard (1972) et de Kleiber (1991a) tout en empruntant une série d'outils liés aux approches de Grice et de Ducrot. Nous croyons avoir montré comment les notions telles que « principe de coopération », « implicature », « présupposé » et « sous-entendu » contribuent à l'élucidation du processus d'inférence à l'œuvre dans l'interprétation des trois grands types de valeurs pragmatiques que les emplois « émotionnels » sont susceptibles de mettre en place. Ces valeurs – dont nous avons avant tout emprunté la typologie aux travaux d'Adamec (1983) et de Kleiber (1991a) – s'avèrent être le résultat d'un présupposé d'identification déictique dévié via une infraction de la « maxime » de quantité gricéenne.

Comme nous l'avons déjà indiqué dans la Section 6.3 au sujet des effets pragmatiques « émotionnels » portés par le TEN introduisant le nom dans les « concepts individuels complexes », l'évaluation des différents types d'effets est parfois un processus délicat. Dans ce chapitre, nous avons aussi identifié les principaux facteurs phonétiques, lexicaux et syntaxiques qui ont tendance à faire pencher la balance du côté de l'une des trois valeurs. Ainsi, la valeur d'affectivité s'accompagne souvent de la présence, au sein de l'énoncé contenant l'occurrence, d'adjectifs qualificatifs de type affectif et évaluatif, d'adverbes de la même sorte, de différents types de modalisateurs, d'une mise en relief prosodique de l'occurrence et de la modalité phrastique exclamative. Si l'affectivité exprimée peut être clairement positive (cf. l'exemple 45) ou clairement négative (cf. l'exemple 43) – surtout si le nom introduit par le démonstratif a pour modifieur un terme à subjectivité forte –, il arrive plus souvent que sa nature soit plus subtile et plus difficile à cerner : de quelle émotion s'agit-il, par exemple, en 34 reproduit *infra* (*toho diváka*) ? C'est notamment dans les cas où la nature de la valeur « émotionnelle » se fait plus discrète qu'il semble souhaitable de tenir compte de l'ensemble du ton général de l'énoncé : en 34 (repris comme 93 *infra*), le locuteur se plaint de l'incompétence d'une de ses collaboratrices ; le ton général est ainsi celui de la désapprobation et de l'indignation, et c'est, à

notre avis, à l'aune de ce ton général qu'il faut évaluer la présence de TEN avec le nom générique *divák* :

93)	ne-mě-l-a	nic	přípraven-ýho ..	ne-mě-l-a	kontakt
	NEG-avoir-PST-3SG.F	rien.ACC	préparé-N.GEN.SG	NEG-avoir-PST-3SG.F	contact.ACC.SG
	na	režisér-y			
	sur	réalisateur-ACC.PL			
	...				
	a na ty		co mě-l-a	jo .. tak to byl	vždycky
	et sur celui-ci.M.ACC.PL		REL avoir-PST-3SG.F	MOD alors ce être.PST.3SG.M	toujours
	průser	kter-ý	to	prasi-l ..	vyhna-l
	fiasco.NOM.SG	qui-M.NOM.SG	cela.ACC	foutre en l'air-PST.3SG.M	chaser-PST.3SG.M
	toho	divák-a .	z toho	divadl-a	jo
	ten.M.ACC.SG	spectateur-ACC.SG	de ten.N.GEN.SG	théâtre-GEN.SG	MOD

elle avait rien de préparé... elle était pas en contact avec les réalisateurs ; et pour ceux avec qui elle était en contact... [avec ceux-là,] ça se terminait toujours par un fiasco qui foutait tout en l'air, qui finissait pas chasser le spectateur du théâtre, tu vois ORTOFON v1

Pour un certain nombre d'occurrences encore (guère élevé), on peut aller jusqu'à se demander si un effet « émotionnel » se produit ou si TEN n'est qu'un outil de marquage du statut informationnel du complexe nominal à l'intérieur de l'énoncé (voir les exemples 29, 30 et 35)⁷⁰. Enfin, rappelons le cas tout à fait singulier des occurrences où TEN, se combinant avec un terme indéfini le précédant, participe d'un complexe nominal indéfini et globalement non-référentiel (cf. les exemples 37, 38 et 39).

La valeur de prise de distance, de son côté, se manifeste fréquemment par la présence de questions et d'ajouts d'informations sur l'identité d'un référent par définition difficile à saisir pour le locuteur. Nous trouvons aussi d'autres marques témoignant de cette difficulté, telles des hésitations ou la forme neutre de TEN adnominal. Quant à la valeur de rappel, nous avons relevé la présence d'une ou de plusieurs mentions antérieures du référent ou de l'information concernée, même si nous ne sommes pas ici face à une règle mais uniquement face à une tendance. Parmi les autres phénomènes identifiés figurent les déictiques temporels ancrant l'information au sein de l'énonciation, des clusters de TEN adnominal à valeur de rappel ou mémorielle, des commentaires incitant le destinataire du message à mobiliser sa mémoire active, éventuellement aussi des moyens d'insistance, etc. Nous avons prêté une attention toute particulière à la présence du relatif *jak*, responsable de la mise en place d'une perspective partagée permettant à l'interlocuteur d'identifier le référent.

⁷⁰ Nous avons par ailleurs constaté que pour ce qui est de la position du complexe nominal eu égard au verbe de la proposition, le partage est assez équilibré entre les positions responsables de la thématisation et la position postverbale, qui, elle, assure la rhématisation du complexe nominal.

Dans la Section 3.5.3 du Chapitre 3, nous avons introduit cette idée qui nous avait été inspirée par Mme Guillot et selon laquelle les deux variantes de la fonction dite « emphatique » par Adamec (1983) – à savoir sa « variante d’émotionnalisation » et sa « variante d’intensification » – seraient davantage orientées sur le locuteur, pour la première, et sur l’interlocuteur (le destinataire du propos), pour la seconde. Répétons ici la définition que l’auteur donne lui-même de ces deux variantes :

- la « variante d’émotionnalisation » sert à « exprimer des émotions diverses du locuteur, telles que l’admiration, l’étonnement, la satisfaction, l’indignation, le désir, etc. »⁷¹
- la « variante d’intensification » serait motivée par « l’effort d’attirer davantage l’attention, d’amplifier l’effet de l’information communiquée, de l’urgence de l’appel, de l’avertissement, etc., [le tout étant] souvent également sous-tendu par des émotions, telles que l’appréhension, la peur, l’impatience, etc. » (1983 : 156)⁷²

Même si la distinction entre ces deux variantes nous paraît assez ténue dans la mesure où l’expression d’émotions est commune aux deux cas de figure, nous estimons que c’est la « variante d’émotionnalisation » qui est largement surreprésentée parmi nos données, au point où nous pouvons nous demander si la « variante d’intensification » est tout simplement présente. En effet, les occurrences porteuses des valeurs d’affectivité ou de prise de distance que nous avons relevées et commentées nous semblent presque toutes traduire l’état émotionnel du locuteur – et donc être centrées sur sa personne – sans lancer un appel urgent au destinataire. Seules les occurrences où la question se pose de la présence d’une valeur « émotionnelle » tout court pourraient éventuellement être taxées de cette fonction. De fait, dans les exemples 29, 35 et 40, la focalisation du nom dans sa position de thème ou de rhème de l’énoncé pourrait s’apparenter à cette volonté « d’amplifier l’effet de l’information communiquée ». Mais cet effet reste modeste sur le plan émotionnel. Il convient ensuite de souligner que les occurrences appartenant à la valeur de rappel représentent, à cet égard, un cas à part, car se tenant en opposition vis-à-vis de l’ensemble de la « fonction emphatique », avec laquelle la « fonction de rappel » constitue les emplois dits « secondaires » (Adamec, 1983). Nous pourrions cependant défier en partie cette analyse, dans la mesure où 1) la valeur « émotionnelle » de rappel est en réalité fortement orientée sur la personne de l’interlocuteur, avec qui le locuteur cherche à instaurer un type de perspective partagée, voire de complicité, car le message implicite qu’il lui destine s’apparente à un genre d’interpellation du type « rappelle-toi, nous en avons déjà parlé / tu es censé être au courant au sujet de ce que je te raconte », etc. Ensuite, cette interpellation (cette dimension appellative au sens de Jakobson, 1963b), qui a sans doute pour conséquence « d’amplifier

⁷¹ Nous traduisons : „[...] vyjadřování různých emocí mluvčího, jako obdiv, překvapení, spokojenost, rozhořčení, touha apod.“

⁷² Nous traduisons : „[...] snaha o upoutání větší pozornosti, zvyšování působivosti sdělení, naléhavosti výzvy, varování apod., spojené ovšem často rovněž s určitými emocemi, jako obava, strach, netrpělivost apod.“

l'effet de l'information communiquée », peut aller plus loin, jusqu'à se transformer en une « urgence », teintée, typiquement, d'une expression d'insistance et d'impatience. C'est ce que nous observons, entre autres, dans l'exemple 69 (repris ici comme l'exemple 94) :

94) a kdy po-jede-te teda . na to dřev-o . s tím děd-ú?
 et quand FUT-aller-2PL alors à ten.N.ACC.SG bois-ACC.SG avec ten.M.INS.SG grand-papa-INS.SG
et alors, c'est quand que vous irez chercher le bois avec grand-papa ? ORTOFON v1

Au total, 200 occurrences ont été analysées dans le second volet du présent chapitre, soit la totalité des 50 occurrences du sous-corpus des pièces de Havel et 150 occurrences du début de l'échantillon de conversations d'ORTOFON v1. Ce sont les valeurs de rappel qui sont les plus nombreuses – 121 occurrences au total, soit 60,5 % de l'ensemble analysé –, les valeurs d'affectivité et de prise de distance représentant, respectivement, 16,5 % et 23 % de l'ensemble analysé (les deux sources de données comprises). Le recours au module « many features – 1 sample » du site du CNT nous permet d'avoir une idée de la représentation des trois types à l'intérieur de l'ensemble des 331 emplois « émotionnels » recensés dans notre échantillon de conversations (plus exactement : à l'intérieur d'un autre échantillon de 150 occurrences) :

Type d'emplois « émotionnels »	Nombre d'occurrences	Intervalle de confiance	
		Borne inférieure	Borne supérieure
Affectivité	16	9	24
Prise de distance	32	22	42
Rappel	102	91	113
Nombre total	150		

Tableau 7.2 : Nombre d'occurrences appartenant aux trois types de valeurs pragmatiques exprimées par le démonstratif au sein des 150 premiers emplois « émotionnels » de l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences) avec les intervalles de confiance anticipant leur fréquence dans un autre échantillon de la même taille au sein de la totalité des 331 emplois « émotionnels » recensés

Les différences entre les taux de représentation des trois valeurs au sein de l'échantillon de 150 occurrences s'avèrent statistiquement significatives après l'utilisation du module « 2 words in 1 corpus » : $L = 112,222$ pour la différence entre les emplois à valeur de rappel et ceux à valeur d'affectivité et $L = 68,9057$ pour la différence entre les emplois à valeur de rappel et ceux à valeur de prise de distance⁷³. Avec la prudence qui est ici de mise à la fois du fait de la taille très limitée de l'échantillon analysé et du profil équivoque d'une partie des occurrences, il nous semble possible de conclure à une surreprésentation de la valeur de rappel relativement aux deux autres valeurs dans la

⁷³ Le taille du « corpus » est ici égale aux 150 occurrences analysées ; la valeur critique = 3,841 (degré de liberté = 1 ; $\alpha = 0,05$).

langue orale informelle contemporaine⁷⁴. Ce constat n'est pas sans importance, compte-tenu, notamment, du profil singulier de ce type d'emploi de TEN. En effet, c'est à l'intérieur de ce type-ci que nous identifions le plus de cas où TEN introduit un nom non-référentiel, l'adjonction parenthétique à valeur mémorielle ayant, dans ce cas, pour rôle de rappeler qu'il a déjà été question d'une information/d'un thème particuliers (cf. les exemples 73 et 74 ; 8,82 % des 102 occurrences)⁷⁵. Enfin, rappelons la force du phénomène de l'attraction, contribuant au fait que dans les énoncés contenant des occurrences de TEN à valeur de rappel, la forme tende à précéder les noms de manière systématique, quelle que soit leur nature sémantico-référentielle.

Il convient également d'insister sur le fait que les emplois « de rappel » se présentent comme une particularité – sinon une anomalie – au sein même de la famille des langues slaves (cf. le constat de leur absence en russe fait par Adamec dans son étude de 1983)⁷⁶. Vu leur profil atypique et leur présence prononcée dans l'oral informel, nous sommes en droit de nous attendre à ce que ces emplois soient investis d'un rôle non négligeable dans l'avancée de la grammaticalisation de TEN adnominal si d'autres conditions et facteurs favorisant cette avancée sont réunies dans l'avenir. Nous écrivions déjà dans la conclusion du chapitre précédent que l'effet pragmatique « émotionnel », découlant de la présence d'un démonstratif adnominal n'ayant aucun impact sur les propriétés extensionnelles du nom qu'il introduit, pourrait avoir une incidence positive sur la grammaticalisation de TEN. À ce propos, nous nous permettons de reproduire à l'identique le passage suivant, emprunté à la conclusion de notre étude de 2020 :

Concerning the presence of *ten* with pure semantic definites, this case is regarded, in this paper, as one situated outside of the grammaticalization process and pertaining to “emotional” uses. The demonstrative still functions as a pure deictic, anchoring the referent explicitly in the context, although it does not play any role in its identification. However, the claim has been made that these two phenomena might get more interwoven in the future. At that stage, the “emotional” uses might accelerate the grammaticalization process. Yet this claim remains speculative, and, therefore, more research needs to be done here. (Dvořák, 2020 : 317)

⁷⁴ Les occurrences des pièces n'ont pas été incluses dans ce calcul afin que seules soient considérées les conversations et que soit visée l'homogénéité des données.

⁷⁵ C'est aussi ce type d'emploi « émotionnel » qui, parmi nos données, est le seul à afficher la présence des trois phénomènes que nous avons essayé de décrire dans la Section 7.2.3, à savoir 1) un nom à valeur temporelle (9,8 % des 102 occurrences), 2) un nom appartenant à la catégorie sémantico-référentielle des « définis faibles » et 3) un « quantificateur flottant ».

⁷⁶ À en croire les jugements de nos deux informateurs, ces emplois seraient inacceptables en polonais. Si la valeur « émotionnelle de rappel » de TEN semble présente en slovaque, les locuteurs seraient beaucoup plus réticents à y recourir, même dans la langue parlée informelle. Ainsi, dans les énoncés des exemples 70 et 75 *supra* (rappelons qu'aucun des deux noms marqués en gras en 75 ne réfère), la présence de TEN est jugée à la limite possible mais fortement connotée sur le plan dialectal. En effet, la plupart des locuteurs slovaques préféreraient l'éviter, à l'exception des locuteurs dialectaux de l'ouest du pays, limitrophe de la République tchèque. Toujours en slovaque, la plus grande réticence à employer TEN semble également vraie pour les deux autres valeurs « émotionnelles ». Autre remarque intéressante de nos trois informateurs slovaques : à l'opposé du tchèque parlé informel, très éloigné de la norme soutenue, le slovaque parlé informel serait davantage respectueux de cette norme.

Il est un dernier point important qu'il convient ici d'aborder. En parlant de l'ensemble de la classe de nos emplois « émotionnels », nous avons insisté sur le fait qu'au sein de ces derniers, le démonstratif ne contribue pas directement à l'identification du référent du nom référentiel qu'il introduit dans le discours (cf. les Sections 3.5.3 et 4.5.4). Si nous avons systématiquement eu soin d'insister sur le terme « directement », c'est qu'il existe probablement des emplois de ce type où le démonstratif facilite l'identification du référent *lato sensu*, même s'il n'agit pas sur les propriétés extensionnelles du nom. Regardons de nouveau les deux exemples suivants qui sont des reprises, respectivement, de l'exemple 68 *supra* et de l'exemple 63 de la Section 3.5.3 :

95) dneska právě mi říka-l-a Páj-a že že jako broker víš?
aujourd'hui justement moi.DAT dire-PST-3SG.F NP-F.NOM.SG que MOD courtier-NOM.SG MOD

ten *Adam* jak děl-á ten NP víš
ten.M.NOM.SG NP.M.NOM.SG REL faire-PRS.3SG ten.M.ACC.SG NP.M.ACC.SG MOD

jak byl s náma na tom na .. na jejich svatb-ě
REL être.PST.3SG.M avec nous.INS à ten.N.LOC.SG à leur.F.LOC.SG mariage-LOC.SG

aujourd'hui, justement, il y a Pája qui m'a dit que que... le courtier, tu vois ? cet Adam, qui bosse sur le NP, tu vois ? qui était avec nous à... à leur mariage ORTOFON v1

96) L1 : pak jsou holč-y tady z toho pokoj-e
après être.PRS.3PL fille-NOM.PL ici de ten.M.GEN.SG chambre-GEN.SG

L2 : tady z tohodle prvn-ího ? .. to
ici de tenhle.M.GEN.SG premier-M.GEN.SG ce.NOM

jsou všechno čtvrtáčk-y ?
être.PRS.3PL tout.N.NOM.SG élève de terminale-NOM.PL

L1 : no Elišk-a .. Míš-a a *ta* *Kamč-a*
MOD NP-F.NOM.SG NP-F.NOM.SG et ten.F.NOM.SG NP-F.NOM.SG

L1 : après, y a les filles de cette chambre-là

L2 : de cette première chambre ? ce sont toutes des filles de terminale ?

L1 : bah, Eliška... Míša et (cette) Kamča ORTOFON v1

Dans ces deux exemples, il est plus ou moins évident que, d'une certaine manière, TEN facilite ou pourrait faciliter l'accès de l'interlocuteur au référent. En 95, la situation ne laisse aucun doute que l'interlocuteur peine à identifier la personne nommée Adam, qu'elle n'aurait rencontrée qu'une seule fois, le contexte laissant entendre que cette rencontre fut de nature épisodique. L'interlocutrice, consciente de cette difficulté, emploie à la fois le démonstratif et les relatives introduites par *jak*. En 96, il est possible que la locutrice utilise TEN pour introduire le référent Kamča non pas (ou pas uniquement) pour marquer son propre manque de familiarité avec le référent (Kleiber, 1991a : 92), mais (aussi) parce qu'elle anticiperait la difficulté que l'identification de cette jeune fille représente

pour son interlocutrice. Dans ce cas, le recours à TEN aurait pour objectif de pallier cette insuffisance en matière d'accès au référent supposée chez l'interlocutrice.

Une fois de plus, les remarques de Šimík et Sláma (2020) au sujet des relatives introduites par *jak* sont ici révélatrices. Les auteurs écrivent que bien que ces dernières ne soient pas toujours de nature restrictive, elles peuvent, malgré cela, faciliter l'accès au référent à travers la mobilisation d'un cadre mémoriel partagé. C'est précisément ce qui se produit en 94, où les relatives sont de nature appositive, mais où elles rendent clairement possible l'identification de la personne prénommée Adam. En vérité, ce cas de figure pourrait en partie être rapproché de l'exemple suivant, cité dans la Section 1.2.3, où la relative, bien qu'appositive, semble nécessaire pour la construction de la référence (rappelons qu'il est ici question de tous les programmes de télé et non seulement d'un certain type de programmes de télé) :

97) *Ces programmes de télé qui exaspèrent les français.* (emprunté à Kleiber et Sock, 2006 : 259)

La ressemblance n'est toutefois que partielle, car contrairement à 95, le référent de 97 est identifié sans la relative. Pour les énoncés comme 95, en revanche, le rôle de la/des relative(s) s'avère incontournable dans l'identification du référent. Nous empruntons l'explication sémantique du fonctionnement de ce type de relatives introduites par *jak* à Šimík et Sláma (2020) et à Šimík (communication personnelle). Cette explication est la suivante : en 95, la locutrice supposerait a) que si l'interlocuteur est (plus ou moins) familier avec le référent, il a du mal à l'identifier via son nom propre. En plus de cela, b) elle saurait aussi que l'interlocuteur est au courant du fait que la personne appelée Adam travaille sur le projet mentionné dans la première relative et qu'il a participé au mariage mentionné dans la seconde des relatives⁷⁷. Ces deux relatives auraient ainsi pour but d'aider l'interlocuteur à éliminer tous les candidats possibles au référent désigné par la locutrice. En ce qui concerne ces candidats, ils ressortiraient, selon Šimík (communication personnelle), à un ensemble de mondes possibles :

Techniquement parlant, il est question d'individus ayant la même extension dans le monde réel, mais qui diffèrent de par leur identité dans une variété de mondes possibles admis par le destinataire. La fonction de la relative introduite par *jak* est d'éliminer tous les mondes possibles dans lesquels le référent en question ne correspond pas au référent réel, et ce via un appel à l'expérience du destinataire [...]⁷⁸

Dans les contextes comme celui de 95, TEN adnominal à valeur « émotionnelle » semble partager avec ce type de relatives introduites par *jak* une part de leur fonctionnement (d'où également

⁷⁷ Rappelons que le contenu des relatives introduites par *jak* doit être connu de l'interlocuteur.

⁷⁸ Nous traduisons : „Technicky vzato se jedná o individua, která mají identickou extenzi ve skutečném světě, liší se však ve své identitě v různých možných světech, které připouští posluchač. *Jak*-relativní věta má pak za úkol vyřadit všechny možné světy, v nichž daný referent neodpovídá tomu skutečnému a dělá to právě tím, že apeluje na posluchačovy zkušenosti [...]“

la tendance de ces deux phénomènes à apparaître ensemble dans le même contexte). En effet, il nous semble que la présence de TEN produit un effet semblable à celui de la relative introduite par *jak*. En 96, si effectivement la locutrice agit dans le but de faciliter à l'interlocutrice l'accès au référent, elle pourrait également ajouter une relative introduite par *jak*, moyennant laquelle elle rappellerait à l'interlocutrice une réalité au sujet de la jeune fille dont celle-ci est au courant : *ta Kamča, jak jsem ti o ní říkala* (*cette Kamča, que je t'ai parlé d'elle*). Si, contrairement à la relative, TEN ne peut décrire cette réalité faute de contenu descriptif aussi fourni, il peut toutefois, en effectuant un renvoi déictique à un univers du discours partagé, lui aussi contribuer à la traque du référent au sein des mondes possibles de l'interlocuteur. Nous tenons néanmoins à séparer ce cas de figure de celui où TEN intervient dans l'identification du référent de façon directe, c'est-à-dire en agissant sur les propriétés extensionnelles du nom.

Conclusion générale

Conclusion générale

Dans cette thèse présentée pour l'obtention du titre de docteur en sciences du langage, nous avons proposé une analyse des emplois adnominaux du démonstratif défini dans l'oral informel des langues tchèque et française. À présent, le moment est venu de procéder à la synthèse des constats faits tout au long de ce travail tout comme à l'évaluation de notre question de recherche principale. La présente conclusion vise à s'inscrire dans un dialogue et une continuité avec l'Introduction générale, où ont été présentées l'orientation du projet, ses principales hypothèses et sa structure. C'est là que nous avons annoncé que nous souhaitions, avant tout, mettre en lumière les tendances proéminentes et les spécificités de l'oral informel des deux systèmes eu égard au démonstratif défini adnominal. Conformément à l'orientation contrastive de notre projet, nous avons prévu de nous focaliser en particulier sur les phénomènes par lesquels ces deux systèmes divergent. Malgré cela, nous avons également réussi à identifier un certain nombre de points communs, et ce notamment dans la perspective diachronique.

Récapitulation

La Partie I s'est ouverte sur le Chapitre 1, qui se voulait une entrée en matière théorique. Nous y avons rappelé les différents angles d'approche de la question de la deixis et des démonstratifs. Plusieurs modèles explicatifs de la deixis ont ainsi été brièvement présentés : le modèle fondé sur la notion d'énonciation, celui fondé sur le site de présence du référent, l'ensemble des modèles articulés autour de la notion d'accessibilité référentielle et le modèle de la « *token-réflexivité* », développé à partir de la notion de « token » de Reichenbach (1966 [1947]) et représenté, avant tout, par Vuillaume (1980) et par Kleiber (1984a et 1986d). Nous avons notamment argumenté en faveur de ce dernier modèle, partant d'une approche sémasiologique des déictiques. En même temps, le modèle onomasiologique basé sur le site de présence du référent a fourni à nos analyses une ossature terminologique, dont notamment la distinction entre emplois exophoriques – a-mémoriel et mémoriel – et endophoriques – anaphorique et cataphorique.

Dans le second et le troisième volet de ce chapitre, c'est le démonstratif en tant qu'unité linguistique qui a été abordé. D'abord, sur le plan de la sémantique référentielle et, ensuite, sur les plans de la morphosyntaxe, de la typologie linguistique et de la diachronie. Il a été rappelé que le démonstratif est un représentant saillant de la classe des expressions indexicales en tant que son interprétation nécessite, dans une situation discursive, le recours à son occurrence apparue à l'intérieur de cette situation. Le démonstratif (défini) fait également partie de la catégorie

hétérogène des expressions définies, bien que les linguistes et les logiciens semblent globalement s'accorder sur le fait que sa sémantique définie est d'une autre nature que celle du reste des expressions de la classe avec, pour représentant emblématique, les descriptions définies. En effet, la définitude associée au démonstratif est d'abord et avant tout tributaire du mécanisme par lequel ce dernier rend possible une identification univoque du référent par son indexation explicite sur un contexte immédiat (d'où le terme de « désignateur contextuel ») : nous sommes ici face à une définitude par définition « pragmatique », car fruit des circonstances et non pas du sens descriptif du nom et de ses compléments (définitude « sémantique » ; Löbner, 1985 et 2011). En revanche, et contrairement aux autres expressions définies, le démonstratif ne semble pas être porteur du présupposé d'unicité, lequel serait en contradiction directe (à la fois sur le plan logico-sémantique et pragmatique) avec son mode référentiel par « contraste interne » (cf. Corblin, 1987). C'est également la raison pour laquelle le démonstratif ne réagit pas avec la même univocité au test de la négation (Löbner, 2011 : 293). La définitude du démonstratif semble donc découler directement de l'acte de référence (l'identification d'une entité extralinguistique), fait qui rend difficilement imaginable tout emploi non référentiel du démonstratif¹, à la différence du défini, dont les emplois attributifs ne devraient plus faire aucun doute après l'« assassin de Smith » de Donnellan (1966). Nous avons ensuite vu en quoi le démonstratif est un véritable « primitif » du langage humain. A été posée, à ce moment-là, la question des critères définitoires de la catégorie. À ce titre, nous nous sommes rallié à la position de Himmelmann (1997) et de Diessel (1999), à savoir que l'expression de la distance n'est pas une condition nécessaire à l'appartenance d'un élément à la classe des démonstratifs². Nous avons rappelé que les systèmes du démonstratif dans les langues du monde sont d'habitude classés entre ceux qui sont articulés autour de l'opposition entre les différentes sphères de distance du référent par rapport au locuteur et ceux qui sont articulés autour de l'opposition entre les acteurs de l'énonciation. Aussi problématique et approximative que s'avère cette binarité (cf. Meira, 2003), elle permet de voir que les deux systèmes qui nous intéressent affichent une structuration en trois sphères de distance. Bien entendu, cette conclusion « naïve » pêche, comme nous l'avons vu plus loin (dans les Chapitres 3 et 4), par un excès de simplification. En effet, alors qu'en tchèque, la quasi-éviction de TEN de l'emploi situationnel aboutit de fait à une opposition à deux termes, en français, l'opposition est en grande partie neutralisée autant par l'absence fréquente des particules localisatrices qui en sont responsables que par la raréfaction de la particule *-ci*. S'agissant de la syntaxe, nous avons d'emblée insisté sur l'unité ou, au contraire, la différence paradigmatique des formes en emplois adnominal et pronominal.

¹ À l'exception des formes déterminatives du démonstratif (cf. la proforme CELUI suivie d'un complément en français ; Diessel, 1999).

² Et pas non plus suffisante, d'ailleurs.

Le Chapitre 2 a été consacré à la présentation des sources de données et de la méthodologie. Nous y avons justifié notre décision de travailler avec ORTOFON v1 (CNT), le CEFC (Orféo) et, d'une manière davantage complémentaire, avec InterCorp (CNT). Il s'est vite avéré à travers l'examen des propriétés des deux principales sources que c'est plus par défaut que notre choix s'était porté sur elles, la situation relative à l'état d'avancement de la linguistique de corpus dans les deux milieux linguistiques restant, de nos jours encore, difficilement comparable. À ce titre, nous constatons l'urgence pour le milieu francophone de développer des corpus oraux d'envergure, aussi bien au niveau de la taille qu'au niveau de la lemmatisation, de l'annotation en catégories grammaticales et des outils d'exploitation des données. Quelque disparates que soient les natures des projets ORTOFON v1 et CEFC, malgré une certaine hétérogénéité des sources et un manque de représentativité du CEFC, nous avons conclu à la comparabilité de ces deux ressources sur le plan des usages du démonstratif adnominal. Les échanges constituant les deux ensembles de données ont pour dénominateur commun leur caractère spontané et informel : les cadres dans lesquels le démonstratif adnominal est employé sont donc comparables. En outre, ce démonstratif atteint dans les deux ensembles une fréquence d'emploi suffisante pour autoriser des généralisations. L'intérêt du choix du sous-corpus des quatre pièces de V. Havel était double : 1) renforcer l'aspect contrastif de notre étude via l'analyse des stratégies de traduction des occurrences tchèques en français et 2) comparer le genre des conversations authentiques avec une représentation littéraire de la langue parlée informelle. Répétons ici que notre projet ne visait ni à étudier la situation relativement à différents paramètres concernant les locuteurs (âge, sexe, niveau de formation, région d'origine, etc.) ni à proposer une analyse traductologique du phénomène. Toutefois, ces deux pistes sont loin d'être dépourvues d'intérêt de recherche et elles pourraient certainement déboucher sur une série de constats précieux. La première repose sur la nécessité de disposer d'une source française équivalente à ORTOFON v1, c'est-à-dire équilibrée eu égard aux paramètres susmentionnés. Quant à la seconde, elle est conditionnée par une double injonction : celle de travailler avec une masse de données plus conséquente et celle d'examiner la situation dans les deux sens (originaux tchèques et traductions françaises et vice versa).

L'importance de disposer d'informateurs natifs s'est révélée vitale, malgré les inconvénients que cette pratique représente. Le principal inconvénient est le suivant : les questionnaires destinés aux informateurs ne peuvent jamais, à notre avis, mettre en place la situation d'une pratique discursive parfaitement spontanée et naturelle, ne serait-ce que par leur cadre artificiel et leur nature écrite. Ainsi, il n'est pas possible d'exclure chez les informateurs une tendance subconsciente à ne jamais s'éloigner complètement de la norme soutenue, aussi explicite que soit l'injonction au « relâchement » et à la spontanéité. De même, l'utilité de l'application à nos données de tests

statistiques ainsi que le recours à des sources de données complémentaires, deux pratiques qui ont été présentées dans le Chapitre 2, ne fait aucun doute. C’est grâce à leur apport que notre travail peut – du moins l’espérons-nous – aspirer à une certaine légitimité quantitative.

Dans le reste du Chapitre 2, nous avons présenté nos requêtes, nos catégories d’annotation ainsi que nos démarches d’annotation et d’échantillonnage. Ces dernières étaient indispensables dans le cas des deux corpus oraux et elles ont opéré en plusieurs temps. En ce qui concerne ORTOFON v1, l’échantillonnage a d’abord opéré au niveau du premier résultat obtenu à travers la requête, trop long pour pouvoir être traité dans son intégralité : une occurrence sur dix a été retenue. Par la suite, les mille premières occurrences ont été annotées. Pour le CEFC, l’échantillonnage est d’abord intervenu avant même l’application de la requête et il a reposé sur le choix de textes ressortissant à l’oral informel (« conversations du secteur privé »). Il a été poursuivi via la maximisation, dans l’échantillon finalement retenu pour l’annotation, des données autres que celles empruntées au sous-corpus TUFS, soit la source de loin la plus importante de données, et ce pour augmenter la diversité du matériel analysé.

Dans le Chapitre 3, dédié au démonstratif tchèque, il a d’abord importé de fournir une description des grands traits de l’ensemble du système. Cela nous a obligé à aborder les oppositions entre les différentes formes, oppositions qui se situent sur plusieurs plans : plan de l’expression de la distance, plans diastatique et diatopique et plan du régime d’indexicalité. Cet exposé s’est en partie appuyé sur une série de comparaisons quantitatives, effectuées par nous-même sur des corpus du CNT ou empruntées à trois dictionnaires de fréquences. Il résulte d’abord de ces comparaisons que la fréquence de TEN va croissant avec le caractère informel des textes. Ainsi, TEN est beaucoup plus représenté dans les conversations et même dans les entretiens dirigés que dans les textes écrits : sa fréquence relative dans l’ensemble d’ORTOFON v1 est cinq fois supérieure à sa fréquence relative dans l’ensemble de SYN2015. Selon le *Frekvenční slovník mluvené češtiny* (Dictionnaire des fréquences du tchèque parlé ; 2007), basé sur le PMK (Pražský mluvený korpus / Corpus parlé de Prague), nous avons aussi pu comparer les conversations pures à des échanges contenant à la fois conversations et entretiens dirigés (les seconds étant plus formels que les premières) : la fréquence relative de TEN dans *Frekvenční slovník mluvené češtiny* est 3,5 fois supérieure à celle de SYN2015. Au sein des textes écrits, TEN est clairement plus représenté dans les textes de fiction que dans les textes de presse et les textes de non-fiction, comme le montrent également les données de SYN2015 (voir *infra* pour l’explication de ce phénomène). Ensuite, il s’avère que la fréquence des formes TENHLE et TENHLETEN (avec leurs variantes *ten(d)le* et *tendle(nc)ten*, etc. ; emplois adnominal et pronominal confondus) constitue un indicateur assez fiable à la fois du registre parlé informel du tchèque en général et, plus particulièrement, de la variété inter-dialectale appelée « tchèque

commun »³. En effet, cette forme est particulièrement répandue sur le territoire de la Bohême, d'où elle se propage dans les autres régions. Vu le statut socio-culturel privilégié de la Bohême – et notamment celui de Prague –, c'est le « tchèque commun » qui est très souvent utilisé dans la littérature pour évoquer le registre informel⁴. Cela apparaît assez nettement à travers l'étude des textes de fiction du corpus SYN2015. Des trois parties qui constituent ce corpus, c'est justement dans la partie fiction – type de textes de nature par définition hybride – que la fréquence relative de TENHLE est la plus élevée. La fréquence relative atteinte par TENHLE dans la partie fiction de SYN2015 dépasse même celle de l'ensemble d'ORTOFON v1. S'agissant de ce dernier corpus, nous avons enregistré une nette différence entre les productions de locuteurs de Bohême et celles de locuteurs de Moravie-Silésie. C'est uniquement dans les premières que la fréquence relative de TENHLE est très proche de celle qu'il a dans l'ensemble de la partie fiction de SYN2015. Mais le constat de loin le plus inattendu pour nous était la possibilité, en Silésie et en Moravie de l'Est, d'un usage accentué de TEN en deixis situationnelle pour identifier un référent plus proche du locuteur, par opposition à la forme d'éloignement TAMTEN. Sur ces territoires, la grammaticalisation de TEN – forme neutre eu égard à la distance – est ainsi susceptible de reproduire (en partie) la situation du polonais et du slovaque⁵, avec une réalisation accentuée (démonstrative) et une réalisation inaccentuée (articulaire) de TEN en emploi situationnel.

La question du modèle syntaxique à appliquer au démonstratif adnominal tchèque est une question majeure qui continue, de nos jours, à faire polémique au sein de la communauté linguistique. Si la tentation est grande d'appliquer au démonstratif tchèque (et aux autres mots grammaticaux fonctionnels) la même analyse que celle basée sur les langues à articles définis (au-dedans ou en dehors du paradigme de la grammaire générative), nous avons pris le parti de la prudence et du scepticisme. À l'heure actuelle, le comportement syntaxique et distributionnel du démonstratif tchèque continue à trop défier cette analyse. Cependant, il serait tout à fait dans l'ordre des choses de s'attendre, dans l'avenir, à un changement de la configuration actuelle à mesure que la grammaticalisation de TEN avancera (si tant est qu'elle avance). Ainsi, un resserrement des compléments et des actualisateurs autour du nom pourrait se produire, débouchant sur une

³ Il ne faut toutefois pas oublier que ces formes appartiennent également à la variété parlée du tchèque soutenu („hovorová čeština“). Quant au « tchèque commun », rappelons ici la conclusion de Berger au sujet d'une possible réfection de l'opposition proche vs lointain : „In der gesprochenen Sprache könnte dieser Entwicklung in einem System resultieren, in dem neben dem universalen und daher zur Verwendung als Deiktikum nicht geeigneten Pronomen TEN ein Nahdeiktikum TENHLETEN und ein Ferndeiktikum TAMHLETEN stehen.“ (1993 : 510) Malheureusement, nos propres données se sont avérées insuffisantes pour étayer ou infirmer cette hypothèse, les formes autres que TEN y étant trop peu représentées.

⁴ Si les autres variétés inter-dialectales et dialectales sont elles aussi utilisées, c'est dans un but bien plus précis : pour connoter explicitement l'origine et/ou l'identité régionale des personnages.

⁵ La situation semble plus proche du slovaque, où TEN peut être accentué, mais où d'autres formes existent qui s'opposent à TAMTEN (TENTO, TOŤTEN, etc.), tandis qu'en polonais, la forme TEN semble quasi-exclusive pour désigner un référent plus proche.

organisation plus stricte et plus hiérarchisée de l'ensemble et voyant l'émergence, en tchèque, d'un syntagme nominal, exactement comme cela s'est produit dans les langues configurationnelles et dotées d'articles (cf. Himmelmann, 1997).

Nous avons tâché de retracer la mise en place du système du tchèque actuel en diachronie, en insistant notamment sur un basculement plausible entre un système basé sur la personne et un système basé sur la distance, basculement qui semble rapprocher l'évolution du tchèque de celle du français. Ensuite, nous sommes passé à une première ébauche du fonctionnement du démonstratif adnominal dans les différents contextes. Outre les régimes d'indexicalité, nous avons relevé la nominalisation d'un adjectif introduit par TEN en l'absence de nom et le cas des emplois « émotionnels », soit un contre-exemple au fonctionnement désignatif habituel du démonstratif. D'autres singularités liées au démonstratif adnominal tchèque ont été abordées et commentées brièvement, dont les emplois baptisés par nous « métalinguistiques » et le cas d'un TEN déterminatif anticipant une relative « établissante » de nature « classificatoire ». Mais, surtout, nous avons d'emblée focalisé notre attention sur les contextes où TEN semble se prêter à un emploi proche des définis davantage que des démonstratifs.

La structure du Chapitre 4 reflète celle du Chapitre 3. Il y a été question du système du démonstratif français. Dans un premier temps, ce sont les distinctions entre la série des déterminants et celle des pronoms qui ont été rappelées. Contrairement à ce qui se passe en tchèque, l'analyse en déterminants de la première des séries est à l'abri de toute contestation, car elle est appuyée par un nombre suffisant de preuves, à commencer par le fait morphologique. A ensuite été traitée la question de l'opposition de distance entre les formes. À cet égard, les retours de nos informateurs francophones étaient riches en constats. Non seulement nous avons eu une confirmation des deux thèses désormais bien connues – celle selon laquelle les formes sans particule sont globalement préférées aux formes pourvues d'une particule, et celle selon laquelle les formes en *-ci* sont sous-employées –, mais il est également apparu que la présence de *-là* peut, dans certains contextes, quitter le domaine de l'expression de la distance pour le marquage d'une sémantique indexicale plus prononcée. Nous avons ensuite procédé à une comparaison succincte de la fréquence du déterminant démonstratif dans différents types de textes. Le constat majeur sur lequel il convient ici de revenir est que ce sont nos conversations qui se sont avérées les plus pauvres en déterminants démonstratifs. Ce constat pointe vers une tendance exactement inverse à celle révélée par les données tchèques. Il semble donc que contrairement au tchèque, où le démonstratif est, de manière générale, un marqueur (parmi d'autres, certes) de l'oralité informelle, en français, il s'agirait d'un marqueur de la complexité structurelle propre à certains types d'écrit, notamment les textes scientifiques et de fiction. Inutile de

dire que cette conclusion très rudimentaire demanderait une étude quantitative et qualitative beaucoup plus poussée d'une masse de données appartenant à plusieurs types de textes oraux et écrits.

Une synthèse de la constitution du système en diachronie a été présentée, avec une insistance sur les parallèles entre nos deux langues. À cet égard, nous avons relevé la ressemblance entre le latin et le tchèque en tant que langues largement non-configurationnelles. Il est intéressant que la syntaxe du complexe nominal du latin classique soit plus souple que celle du complexe nominal du tchèque contemporain, comme cela ressort de nos données et des travaux de Goyens et de Veselovská. Pourtant, il n'en a pas toujours été ainsi : les données du corpus Diakorp v6 montrent que la postposition du démonstratif au nom était jadis tout à fait acceptable⁶. Ainsi, comparé à ses états antérieurs, le complexe nominal tchèque semble déjà avoir évolué vers davantage de resserrement et de rigidité, ce qui conforte la thèse d'un article défini naissant (voir *supra*).

Un autre point abordé a été la sémantique indexicale du déterminant démonstratif français en synchronie. Autrement dit, on s'est demandé si cette forme se serait affaiblie sur le plan déictique, d'une part du fait de sa tendance à ne plus marquer la distance en français actuel et, d'autre part, du fait de changements apparus dans le fonctionnement de l'article défini. Si cette thèse semble séduisante à première vue, nous nous sommes employé à montrer ses incohérences pour conclure qu'a priori, rien ne permettait, de nos jours, de traiter CE comme un article défini (fort). C'est exactement ce qu'atteste son fonctionnement dans les différents types d'emploi : contrairement à ce qui a été observé pour TEN en tchèque, une différence de nature sépare CE du défini français. Le seul régime d'indexicalité où cette différence semble davantage tenue est celui de l'emploi mémoriel, quoiqu'ici aussi, il apparaisse assez clairement que le recours au démonstratif s'accompagne du marquage d'un « coût d'activation » et d'une accessibilité référentielle plus élevés. Nous avons aussi vu que le phénomène des emplois « émotionnels » du démonstratif adnominal est sans commune mesure dans les conversations françaises et dans les conversations tchèques. Non seulement sa fréquence y est beaucoup plus basse, mais son apparition y obéit à des contraintes d'ordre sémantique et pragmatique plus fortes.

Le Chapitre 5 de la Partie II a apporté une série de constats quantitatifs. Nous avons d'abord tenu à porter à la connaissance du lecteur les problèmes et dilemmes majeurs rencontrés lors du travail d'annotation, autant de traductions concrètes de l'ambiguïté inhérente aux productions langagières authentiques et, a fortiori, aux productions orales informelles. Le principal obstacle auquel nous nous sommes heurté était l'absence d'un contexte suffisamment large, malgré notre accès aux

⁶ « Pour l'ensemble de la période historique du tchèque, la position de l'épithète est caractérisée par une oscillation forte entre l'antéposition [...] et la postposition [...], souvent dépourvue de distinction fonctionnelle apparente. » (Navrátilová et Zikánová, 2017) Nous traduisons : „Pro slovosled shodného atributu je po celé historické období č. typická silná oscilace mezi antepozicí [...] a postpozicí [...], a to často bez zřejmějšího funkčního rozlišení.“

transcriptions de l'ensemble des conversations. Ensuite, un certain nombre d'occurrences – notamment tchèques – se sont révélées ambiguës indépendamment de la question du contexte (il y avait souvent une hésitation entre un emploi anaphorique associatif, mémoriel ou « émotionnel »). En conséquence de ce caractère problématique, une partie des occurrences ont été annotées comme ambiguës et elles ont été écartées de la plupart des analyses. S'il est une leçon majeure qu'il convient de tirer de cet état de fait, c'est que dans une langue affichant une haute fréquence du démonstratif adnominal, il est souvent difficile d'identifier les différents cas de figure dont relève chaque emploi. Cela rend l'étude de la grammaticalisation particulièrement délicate, car les différents types d'emploi peuvent chacun relever d'une phase différente (au sens de plus ou moins avancée) du processus, mais ils peuvent aussi se situer complètement en dehors de ce processus (le cas des emplois « émotionnels »).

Le bilan quantitatif en tant que tel a commencé par la présentation des fréquences du démonstratif adnominal dans les corpus de conversations des deux langues et dans le sous-corpus des pièces de V. Havel. Les fréquences relatives des deux ensembles de conversations affichent un clivage sans commune mesure : il y a à peu près huit fois de plus de démonstratifs adnominaux en tchèque qu'en français. Le lemme TEN représente à peu près 96,47 % des occurrences tchèques, ce qui en fait de loin le lemme le plus fréquemment utilisé dans l'oral informel. Le sous-corpus des pièces de V. Havel affiche lui aussi une fréquence relative plus élevée du démonstratif adnominal que le corpus de conversations françaises (à peu près dans le rapport de 2,5 à 1). Cependant, cette fréquence est largement inférieure à celle des conversations tchèques (à peu près dans le rapport de 1 à 3,3). Il s'est malheureusement avéré tout au long du travail que la langue du sous-corpus des pièces est moins informelle que nous ne le croyions, notamment à cause de la présence de la pièce *La grande roue*⁷. Il sied de rappeler ici les maillons faibles de ces comparaisons quantitatives initiales, à savoir : l'impossibilité d'identifier toutes les occurrences du démonstratif adnominal, le caractère approximatif des chiffres obtenus pour le corpus de conversations tchèques⁸ et la taille très modeste du sous-corpus des pièces de V. Havel⁹. Nous réitérons ici notre grief lié à l'impossibilité de trier, dans les corpus du CNT, les emplois adnominaux et pronominaux¹⁰. En effet, ce tri aurait considérablement facilité notre travail.

⁷ Cela s'est vérifié via la comparaison avec notre corpus de pièces de théâtre de référence, défini au sein de la partie fiction de SYN2015.

⁸ Du fait de la grande taille du résultat initial et de la présence de bruit, nous avons été obligé de recourir à l'interface « 1 feature – many samples » pour arriver à la fréquence relative du démonstratif adnominal, comprise dans un intervalle de confiance.

⁹ Rappelons aussi que tous les textes des pièces ne sont pas non plus tout à fait conformes à la définition de la synchronie (voir la Section 2.1.3).

¹⁰ Répétons que si cette confusion se justifie sur le plan syntaxique, elle rend le travail avec les données d'autant plus délicat que le lemme inclut aussi des emplois où la forme est plus proche d'une particule (*to*), comme nous l'avons vu dans le Chapitre 3.

Nous avons également comparé la fréquence relative de TEN dans ORTOFON v1 avec la fréquence relative de l'article défini français dans plusieurs types de textes, soit les textes scientifiques, de presse, de fiction et les conversations. Le bilan de cette comparaison incite au scepticisme et à la prudence : toute élevée qu'elle est, la fréquence relative de TEN dans les conversations tchèques se situe assez loin en-deçà de la fréquence relative de LE (à peu près dans le rapport de 1 à 3,2 par comparaison avec la fréquence relative de LE dans les conversations du CEFC et à peu près dans le rapport de 1 à 4,5 par comparaison avec la fréquence relative de LE dans notre corpus de fiction défini au sein de Frantext). À cela s'ajoute le fait que dans nombre de ces occurrences, TEN fonctionne comme un démonstratif « émotionnel ». Ainsi, il est possible de tirer de ces bilans une première conclusion, purement quantitative : de par sa fréquence, TEN adnominal se situe, dans le tchèque parlé informel de nos jours, entre un vrai démonstratif et un défini¹¹.

Dans le second volet du même chapitre, nous avons présenté les résultats de l'annotation des trois ensembles de données. S'agissant d'abord des types et sous-types d'emploi du démonstratif, l'échantillon de conversations tchèques se caractérise, avant tout, par la prédominance des emplois « émotionnels » et mémoriels. L'échantillon de conversations françaises, lui, est le plus riche en emplois endophoriques et déictiques temporels. À l'intérieur des emplois endophoriques, le français affiche une forte présence des emplois anaphoriques résomptif et temporel, contrairement au tchèque, où ce sont la cataphore (intra-syntagmatique) et l'anaphore associative qui priment. S'agissant ensuite de la position par rapport au verbe de la proposition, les données françaises révèlent une prédominance de la position postverbale et de la position de complément de phrase. En tchèque, ce sont les positions postverbale et préverbale qui prévalent. À l'aide du calcul de spécificités, des affinités ont été constatées, en français, entre le complément de phrase et les emplois déictique temporel et anaphorique temporel et entre la dislocation à droite et l'endophore en général. Alors que la première affinité s'explique très clairement à travers la fonction de circonstant extra-prédicatif des déictiques et anaphoriques temporels, la seconde est moins intuitive et s'expliquerait notamment par la nature discursivement établie et saillante des référents anaphoriques figurant en dislocation à droite. Dans les deux ensembles de conversations, une affinité a été constatée entre la position postverbale et l'exophore mémorielle. Cette affinité semble s'expliquer à travers le caractère globalement rhématique des référents mémoriels, ces référents ne se recrutant ni à partir de la situation d'énonciation ni à partir du discours, mais à partir de connaissances de plus longue date. Nous savons

¹¹ Force est aussi de souligner que l'emploi du défini est particulièrement élevé en français (et dans les langues romanes en général). En effet, comme nous l'avons vu dans le Chapitre 4, ces langues représentent un état très avancé de la grammaticalisation du démonstratif en défini. Par conséquent, il serait intéressant de voir la fréquence relative du défini dans une langue plus « conservatrice » sur ce plan, telle que l'anglais. Pourtant, même en anglais, *the* serait le lexème le plus fréquent, y compris dans les textes écrits (Leech, Rayson et Wilson, 2001). Rappelons que le lemme TEN est, de nos jours, le plus fréquent, mais uniquement dans la langue (parlée) informelle, et lorsque nous tenons compte des deux positions syntaxiques (adnominale et pronominale).

grâce au calcul de spécificités que la grammaticalisation de TEN en tchèque est plus étroitement liée aux emplois mémoriel, cataphorique, anaphorique associatif et, dans une moindre mesure, situationnel. D'autres phénomènes ont été étudiés, dont les cumuls, en tchèque, du démonstratif avec un autre actualisateur du nom. Nous avons insisté sur l'agrammaticalité de ce type de cumuls en français, où les deux éléments pouvant se combiner en tchèque occupent tous les deux la position du déterminant central, définie par des critères syntactico-distributionnels. À nos yeux, cette possibilité du tchèque constitue une preuve forte de l'absence, dans cette langue, de l'unité syntaxique qu'est le syntagme nominal. Les emplois « métalinguistiques » – une autre spécificité du tchèque – ont eux aussi été étudiés, sur fond de l'ensemble des occurrences contenant des disfluences.

Enfin, c'est le démonstratif introduisant dans le discours les noms à sémantique temporelle qui a fait l'objet d'une investigation dans les données de conversations. L'idée nous en est venue lorsque nous nous sommes aperçu que ce phénomène se trouvait davantage dans les conversations françaises que dans les conversations tchèques. Outre une différence structurelle (en tchèque, la référence au temps se fait davantage au moyen d'expressions adverbiales), nous avons relevé la présence en français de connecteurs argumentatifs employant un nom à sémantique temporelle, le plus souvent le nom *moment*. Cependant, ce ne sont pas uniquement les noms temporels introduits par un démonstratif qui apparaissent plus souvent dans nos conversations françaises, mais aussi les noms temporels tout court, comme l'a montré une brève recherche dans nos deux corpus. Par conséquent, d'autres facteurs contribuent sans doute à cet état de fait et dépassent le cadre du présent travail. Un ensemble d'observations intéressantes ont aussi été faites sur les conversations françaises, à commencer par la très faible présence de *-ci*, y compris en emploi déictique temporel. Ce constat conforte la thèse sur la faible utilisation de *-ci* dans la langue parlée informelle (voir *supra*). Le seul type de contextes où la place de *-ci* serait encore relativement solide est justement celui où *-ci* accompagne un nom à sémantique temporelle, notamment le nom *fois*. Dans le reste des cas, l'emploi de *-ci* se présente comme hypercorrect (ce sentiment est confirmé par nos informateurs). Toutefois, sa faible présence dans l'ensemble de nos données ne nous permet guère d'avancer des conclusions fortes (si ce n'est cette très faible présence), une étude plus poussée serait nécessaire. Il y a ensuite la forte représentation de la particule *-là* en emploi endophorique, surtout avec les noms employés dans des anaphores résomptive et temporelle. Pour ces deux derniers types d'anaphore, le calcul de spécificités indique une affinité. Mais de manière plus générale, nos données pour le français vont dans le sens de la thèse principale¹², à savoir celle de la capacité de la langue à se passer des formes à particule localisatrice, car dans notre échantillon de 1 000 occurrences de CE, ces formes ne constituent que 20,2 % de l'ensemble (soit 202 occurrences).

¹² Cf. Anderson et Keenan, 1985.

La Partie III a été toute entière dédiée aux deux spécificités de l'oral informel tchèque au sujet desquelles nous avons émis l'hypothèse qu'elles seraient responsables de la haute fréquence de TEN adnominal dans ce registre du tchèque. Le Chapitre 6 a apporté une analyse des différents contextes où TEN est plus ou moins en passe de devenir un article défini. Ainsi, alors que le Chapitre 5 s'est contenté de donner, d'une manière quelque peu dogmatique et arbitraire à première vue, les chiffres relatifs à la grammaticalisation présumée de TEN pour chacun des (sous-)types d'emploi, l'objectif du Chapitre 6 était précisément de justifier nos choix en matière de grammaticalisation. Comme déjà annoncé dans l'Introduction générale, la structure de l'ensemble des analyses présentées dans ce chapitre a reposé d'abord sur la distinction löbnérienne entre les deux types de définitude. Après une introduction très générale à la notion de grammaticalisation, nous nous sommes penché plus particulièrement sur celle du démonstratif en défini, tout en essayant de nous situer au sein du débat linguistique relatif à ce sujet éminemment complexe. Trois approches en partie concurrentes ont été rappelées à cette occasion : celle qui accentue le rôle, dans ce processus, de l'emploi anaphorique, celle qui insiste sur la présence obligatoire de l'ancien démonstratif en contextes génériques et celle qui souligne l'importance de l'emploi mémoriel. Nous avons pris parti pour la troisième de ces approches, la plus à même, à notre avis, de rendre compte de la transition entre la définitude « pragmatique » et la définitude « sémantique ». En ce qui concerne la théorie CTD, outre les notions fondatrices de Löbner (1985 et 2011), nous avons brièvement présenté les travaux d'Ortmann (2014) et de Czardybon (2017). Ceux-ci ont pour nous l'incontestable vertu de porter sur les langues slaves occidentales, notamment le polonais et le haut sorabe. Avant d'entamer les analyses à proprement parler, nous avons passé en revue les positions d'une série d'auteurs concernant le statut de TEN adnominal en tchèque (parlé) informel, parmi lesquelles il est possible d'identifier deux partis-pris majeurs : 1) si TEN est en voie de grammaticalisation en défini, le processus n'est pas abouti et 2) le statut de TEN serait, dans le tchèque parlé informel de nos jours, équivalent à celui du défini des langues qui en disposent (le processus serait donc largement abouti). Les analyses présentées dans le Chapitre 6 donnent largement raison à la première des positions. Nous avons ensuite présenté les hypothèses relatives à la grammaticalisation et les critères permettant d'évaluer la réalité de cette dernière.

Dans le domaine de la définitude « pragmatique », la grammaticalisation de TEN est le plus évidente en emploi situationnel. Ici, toute occurrence semble pouvoir être évaluée comme grammaticalisée, à l'exception de celles où la présence du lexème permet au locuteur de témoigner de sa subjectivité accrue (les situations où le référent est l'unique représentant de sa classe dénotative) et à l'exception des cas où il est accentué et fonctionnerait comme un déictique (cas qui n'apparaît qu'une seule fois dans notre échantillon de conversations). Dans ces deux derniers cas, le français

optera pour CE¹³. En régime anaphorique, il est plus difficile de trancher, de nombreux facteurs étant en jeu, dont certains dépassent clairement le cadre de ce travail. Ces facteurs sont liés à la structure fonctionnelle de l'énoncé, aux contraintes pragmatiques et au degré d'ancrage et d'activation du référent dans le discours. Nos critères d'évaluation de la grammaticalisation en régime d'anaphore se fondent en premier lieu sur la notion d'activation du référent et sur son statut (non) établi dans le discours. À ce titre, nous avons distingué deux fonctions de TEN, soit la fonction « établissante » et la fonction « réactivante », toutes deux mettant à profit l'indexicalité du démonstratif. Ainsi, toute occurrence de TEN anaphorique ne relevant ni de l'une ni de l'autre de ces fonctions était susceptible d'être analysée comme étant en cours de grammaticalisation¹⁴, excepté les cas où TEN était nécessaire pour le maintien d'une coréférence en situation d'anaphore infidèle. Dans ce dernier cas, c'est CE qui constitue la première, voire l'unique option en français, excepté les cas où la recatégorisation du référent via le nom de la reprise est si discrète que le défini ne met pas en péril la coréférence¹⁵. Le jugement de nos informateurs natifs nous a également été d'une grande aide. En emploi mémoriel, toutes les occurrences sauf celles où la présence de TEN manifestait une subjectivité accrue du locuteur¹⁶ et celles où TEN se combinait avec les actualisateurs indéfinis NĚJAKÝ et TAKOVÝ ont automatiquement été considérées comme en voie de grammaticalisation. Cette décision est étayée par les faits suivants : 1) dans tous ces cas, TEN peut être traduit en français (ou une autre langue à articles définis) par le défini¹⁷ ; 2) dans les langues slaves où le démonstratif n'a pas entamé sa grammaticalisation, tel le russe, c'est le nom nu qui constitue le premier choix des locuteurs dans ce contexte¹⁸ ; enfin, 3) du moment où l'emploi du nom était interprété comme mémoriel, l'écrasante majorité de nos informateurs a préféré la variante avec TEN, ce dernier étant ainsi ressenti comme quasi-obligatoire. Rappelons aussi qu'un lien fort a été identifié et décrit dans la littérature entre la grammaticalisation d'un démonstratif et une hausse marquée de sa fréquence en emploi mémoriel (cf. Laury, 1997).

¹³ Cela se confirme sur le matériel de notre sous-corpus des pièces, bien que le nombre d'occurrences concernées y soit très bas. Inversement, nous avons aussi vu que les occurrences où TEN fonctionne comme un défini sont parfois rendues en français par le démonstratif. L'explication que nous avons suggérée est que le traducteur peut viser la mise en place d'un contraste implicite entre le référent en question et d'autres référents de la même catégorie.

¹⁴ Les données de notre sous-corpus des pièces appuient cette conclusion : les occurrences considérées par nous comme en voie de grammaticalisation sont le plus souvent rendues par LE et celles considérées comme non touchées par la grammaticalisation sont souvent rendues par CE.

¹⁵ Globalement, les données du sous-corpus des pièces vont également dans ce sens.

¹⁶ Dans ces situations, c'est de nouveau CE qui sera spontanément utilisé en français.

¹⁷ Rappelons cependant que parmi les données du sous-corpus des pièces, les occurrences évaluées comme en voie de grammaticalisation sont presque autant transposées par CE que par LE, même si, comme nous l'avons montré, LE aurait également pu être employé. Nous attribuons ce fait à la ligne de démarcation plutôt floue entre le défini et le démonstratif en emploi mémoriel, phénomène maintes fois mentionné dans nos analyses.

¹⁸ Cette différence notable dans le comportement du tchèque et du russe a été relevée par Adamec (1983) et confirmée par nos deux informateurs pour le russe.

Quant aux emplois situés entre les définitudes « pragmatique » et « sémantique », le travail avec nos données a démontré que la présence de TEN adnominal marquant l'unicité dénotative du nom est ici aussi loin d'être une rareté. En nous référant aux travaux d'Ortmann (2014) et de Czardybon (2017), nous avons distingué trois cas au sein de cette classe : 1) ceux où TEN marque l'unicité dénotative d'un nom rendu unique par une subordonnée relative « établissante », une subordonnée conjonctive ou un complément prépositionnel/adverbial, 2) ceux où TEN marque l'unicité dénotative d'un nom rendu unique grâce à la présence d'une anaphore associative et 3) ceux où l'unicité marquée par TEN est atteinte par la présence d'un superlatif, d'un ordinal ou d'un complément analogue (les « concepts individuels complexes »). La préférence pour l'option avec TEN variant d'un cas à l'autre, elle est particulièrement forte lorsque le nom est modifié par une relative « établissante » et plutôt forte dans certains cas d'anaphore associative de type situationnel et dans certains contextes accueillant des « concepts individuels complexes ». Pour ce qui est de ces derniers, TEN est très favorisée a) par l'expressivité du locuteur et b) par l'absence de nom. Alors que l'importance du premier facteur a été confirmée par nos informateurs, celle du second facteur a été confirmée par une recherche sur l'ensemble d'ORTOFON v1. Nous nous sommes également penché sur la totalité des facteurs contribuant à la mise en place d'un effet pragmatique « émotionnel » (d'ordre affectif, de prise de distance ou de rappel) et nous avons relevé la possibilité d'une influence positive de cet effet sur le processus de grammaticalisation en tant que tel, hypothèse qui tire sa légitimité des travaux soulignant l'importance de la pragmatique et de la subjectivité dans les premiers moments de ce processus (cf. Epstein, 1995 ; Traugott et Hopper, 2003 [1993]). Nous avons fini par conclure que la situation du tchèque parlé informel de nos jours se rapproche d'une « scission du type I » (Ortmann, 2014), à savoir d'une configuration où la définitude « pragmatique » est marquée par un article défini en devenir et la définitude « sémantique » n'est pas marquée.

Dans le septième et dernier chapitre, nous avons pleinement abordé le sujet des emplois « émotionnels » du démonstratif adnominal tchèque. Dans le volet théorique du chapitre, nous avons proposé un traitement unifié de l'ensemble de ces emplois en empruntant des concepts à la sémantique et à la pragmatique. Nous avons montré que si le démonstratif n'influe pas sur les propriétés extensionnelles du nom, sa composante sémantique indexicale¹⁹ continue d'agir et, du fait de sa caducité dans la démarche active d'identification du référent par les participants à l'énonciation, elle se transforme en une implicature conversationnelle porteuse d'un effet « émotionnel ». Plus précisément, cet effet naît en tant que le démonstratif oblige l'interlocuteur à appréhender le référent sur fond d'un contexte fondant son évaluation par le locuteur. Nous avons également vu comment l'émergence de l'effet en question est conforme au « principe de coopération » (Grice, 1979 [1975]).

¹⁹ Cette composante s'apparente, comme nous l'avons également montré, à un présupposé.

Dans le second volet du chapitre, nous avons présenté les résultats des analyses d'un échantillon de 200 occurrences du démonstratif adnominal tchèque porteur d'une valeur « émotionnelle ». Cet échantillon était composé des 50 occurrences du sous-corpus des pièces de V. Havel et des 150 premières occurrences de l'échantillon de conversations d'ORTOFON v1. Parmi les trois valeurs fondamentales – affectivité, prise de distance et rappel –, c'est la troisième qui est la plus représentée dans cet échantillon (60,5 % des 200 occurrences examinées). Le recours au module « many features – 1 sample » a appuyé la prévalence de la valeur de rappel à l'intérieur de la totalité des 331 occurrences « émotionnelles » de notre échantillon de conversations tchèques. Nous avons procédé à l'étude des différents traits phonétiques, lexicaux et syntaxiques qui tendent à accompagner chacune des trois valeurs reconnues par nous. Par ailleurs, nous avons attiré l'attention sur les deux faits suivants : 1) certaines des occurrences combinent en elles plusieurs valeurs « émotionnelles », avec, cependant, presque toujours une valeur se présentant comme dominante et 2) pour une poignée d'occurrences, il semble difficile d'affirmer la présence d'une quelconque valeur « émotionnelle », si ce n'est une simple thématisation ou focalisation de l'information. L'écrasante majorité des occurrences concernées sont des occurrences de TEN, sauf deux cas où c'est TENHLETEN qui est utilisé (les deux occurrences mettent en place un effet de prise de distance avec le référent). Quant à TEN adnominal porteur de l'effet de rappel, non seulement ce cas de figure est le plus représenté dans nos données, mais c'est aussi, comme nous l'avons montré, celui qui affiche le plus de singularités autant au niveau sémantico-référentiel que syntaxique. Au niveau sémantico-référentiel, d'abord, le nom introduit de la sorte est loin d'être systématiquement référentiel, ni défini, d'ailleurs²⁰. Ensuite, contrairement à ce qui se produit lorsque TEN véhicule les deux autres valeurs « émotionnelles », le nom peut être un défini « faible » ou un concept individuel à sémantique temporelle. Enfin, sur le plan de la syntaxe, nous avons observé que TEN apparaît quatre fois dans une structure à « quantificateur flottant ». Nous avons relevé la présence, notamment parmi les occurrences à valeur de rappel, des subordonnées (notamment relatives) introduites par *jak*, lui aussi porteur d'une valeur mémorielle et d'une perspective partagée entre le locuteur et l'interlocuteur. Ce dernier phénomène est également répandu parmi les occurrences mémorielles du démonstratif tchèque. Nous avons relevé le fait que bien que le démonstratif « émotionnel » ne contribue jamais à l'identification du référent directement, c'est-à-dire en agissant sur l'extension du nom, il peut tout de même le faire indirectement, en facilitant l'accès au référent dans un ensemble de mondes possibles imaginés par l'interlocuteur. C'est une autre propriété que le démonstratif « émotionnel » partage avec les subordonnées introduites par *jak*.

²⁰ Toutefois, la référentialité ne saurait être considérée comme une propriété systématique du complexe nominal dans aucune des trois valeurs « émotionnelles ».

Nous avons constaté que l'écrasante majorité de nos occurrences « émotionnelles » porteuses des valeurs d'affectivité ou de prise de distance relèvent de la « variante d'émotionnalisation » de la fonction dite « emphatique » par Adamec (1983), dans la mesure où elles sont centrées sur la personne du locuteur. En revanche, les occurrences affichant la valeur « de rappel » nous semblent avant tout orientées vers le destinataire du message, et ce via la mise en place d'une perspective partagée. De même qu'au sujet des « concepts individuels complexes » étudiés dans le Chapitre 6, l'hypothèse qui accentue un éventuel impact positif de l'ensemble des emplois « émotionnels » sur la grammaticalisation de TEN a été avancée. En effet, même si c'est pour des raisons très différentes que TEN y est aujourd'hui présent, la haute fréquence de ces emplois pourrait aider à faire pencher la balance vers un usage systématique et obligatoire de la forme, du moins dans ceux des contextes de définitude « sémantique » où la majorité des langues à articles définis recourt à l'article (pas nécessairement donc devant les noms propres et génériques, mais, éventuellement, devant les noms communs fonctionnant comme des concepts individuels et fonctionnels).

Enfin, sur le plan de la traduction française des occurrences du sous-corpus des pièces de V. Havel, nous avons remarqué que le démonstratif avait le plus souvent été transposé par un démonstratif lorsqu'il s'agissait de TEN introduisant un nom propre de personne et marquant une prise de distance avec son référent. Par conséquent, c'est là que les traducteurs ont le plus aisément conservé le même effet. En revanche, le démonstratif tchèque ainsi que l'effet « émotionnel » qu'il portait ont été beaucoup moins maintenus dans les autres occurrences, notamment dans celles où il véhiculait un effet de rappel (ici, c'est le défini qui a été le plus fréquemment employé).

D'une manière globale, l'étude des emplois « émotionnels » dans les deux langues a révélé que si un modèle d'analyse universel comme celui que nous avons proposé dans le premier volet du Chapitre 7 peut être appliqué aux deux systèmes, ce modèle doit ensuite être à la fois élargi et raffiné afin de rendre compte des emplois « émotionnels » tchèques, plus diversifiés. Dans tous les cas, nous pouvons conclure que les approches des deux milieux linguistiques – dont celle pour le français s'appuie notamment sur les travaux de Bonnard (1972) et de Kleiber (1991a) et celle pour le tchèque s'appuie notamment sur l'étude d'Adamec (1983) – sont en bonne partie complémentaires et qu'elles permettent un enrichissement mutuel. À nos yeux, ce constat montre clairement l'intérêt d'une approche contrastive comme la nôtre.

Évaluation de la question de recherche

Rappelons d'abord **notre question de recherche**, telle que nous l'avons formulée dans l'Introduction générale :

Y a-t-il des différences substantielles – aussi bien sur le plan quantitatif que qualitatif – en matière d'emploi des démonstratifs adnominaux au sein du registre parlé informel du français et du tchèque ? Si oui, alors comment expliquer ces différences ?

L'hypothèse principale directement liée à cette question et dont nous sommes parti est la suivante :

L'oral informel tchèque emploierait le démonstratif adnominal beaucoup plus fréquemment et dans davantage de contextes que l'oral informel français, et ce notamment pour deux raisons. La première serait la grammaticalisation en cours de TEN en article défini. À cet égard, il convient de mentionner qu'un lien de causalité a été mis au jour par des chercheurs étudiant la grammaticalisation entre, d'une part, la grammatisation d'un démonstratif et, d'autre part, une forte hausse de la fréquence des emplois de ce dernier (cf. Kroch, 1989 ; Laury, 1997 ; Zíková, 2017). La seconde raison, ce seraient les emplois dits « émotionnels » du démonstratif adnominal tchèque. Ces derniers connaîtraient à la fois une fréquence et un nombre de valeurs bien plus élevés que les emplois « émotionnels » du démonstratif adnominal français.

En nous basant sur l'ensemble des analyses présentées jusqu'ici, nous pouvons déclarer valide cette hypothèse. En effet, nos données pour le tchèque parlé informel contiennent beaucoup plus d'occurrences du démonstratif adnominal que nos données pour le français parlé informel : entre 16 107 et 16 745 occurrences pour un million de positions contre seulement 2 003 occurrences pour un million de positions. Il est également vrai que les deux facteurs susmentionnés – à savoir les occurrences de TEN où son fonctionnement est bien plus proche de celui d'un article défini que d'un démonstratif ainsi que les occurrences du démonstratif (dans la quasi-totalité des cas, c'est de nouveau TEN qui est concerné) où il introduit dans le discours un nom sans avoir un impact sur ses propriétés extensionnelles et sans ainsi contribuer directement à l'acte de référence dont le nom forme le noyau – sont massivement présents dans les données tchèques. Au total, nous avons identifié, à l'intérieur de notre échantillon de 1 000 occurrences, 451 occurrences où TEN subit, à notre avis, un processus de grammaticalisation en article défini et 331 occurrences qui relèvent, selon nous, du second phénomène, c'est-à-dire des emplois « émotionnels »²¹. De toute évidence, la prudence est toujours

²¹ L'usage de l'interface « many features – 1 sample » a révélé l'intervalle de confiance 421-481 pour les emplois grammaticalisés et 302-361 pour les emplois « émotionnels ». Ces intervalles anticipent la fréquence des deux phénomènes dans un autre échantillon de 1 000 occurrences prélevées sur la liste des concordances originelle (en écartant les occurrences bruitées ; $\alpha = 0,05$).

de mise, car comme nous l'avons souligné à plusieurs reprises, l'analyse et l'évaluation des occurrences étaient un processus difficile, marqué de doutes et d'incertitudes. Notre propre subjectivité a ainsi pesé d'une manière non négligeable dans les jugements sur les occurrences (faillie à laquelle nous espérons avoir en partie remédié via le recours à nos informateurs). Ceci dit, la totalité des occurrences qui ont été analysées comme appartenant à l'un ou à l'autre de ces deux cas de figure s'élève au nombre de 782, soit 78,2 % de l'échantillon. Ce chiffre est plus que significatif. Si nous soustrayons cette proportion à la moyenne de l'intervalle de confiance 16 107-16 745 (le nombre total d'occurrences du démonstratif adnominal), nous obtenons le chiffre de 3 581, soit un chiffre bien plus proche de la fréquence du démonstratif adnominal français (2003 occurrences). Il s'agit donc d'une preuve que les deux phénomènes participent largement à la haute fréquence du démonstratif dans l'oral informel tchèque²².

D'une manière quelque peu inattendue, c'est sa contribution à l'étude de la grammaticalisation de TEN adnominal qui constitue sans conteste l'un des apports les plus significatifs (sinon le plus significatif) de ce travail. Inattendu, car si nous étions bien au courant au sujet de ce phénomène, nous ne soupçonnerions pas son ampleur.

Outre l'emploi mémoriel, beaucoup plus représenté en tchèque, les phénomènes qui ont directement parti lié avec la grammaticalisation de TEN sont, dans l'oral informel tchèque, tous les emplois où TEN marque une unicité dénotative du nom ayant ses origines à l'intérieur du complexe nominal lui-même. Parmi nos catégories d'annotation, ces phénomènes correspondent à la cataphore (intra-syntagmatique) et à l'anaphore (éventuellement la cataphore) associative. Au total, le nombre d'occurrences de notre échantillon tchèque relevant de ces phénomènes s'élève à 117.

Parmi les autres singularités propres au seul oral informel tchèque, nous relevons notamment : 1) la présence d'emplois dits « métalinguistiques », représentés, dans notre échantillon tchèque, à raison de 14 occurrences, ainsi que 2) la présence de cumuls entre le démonstratif et un autre mot fonctionnel (70 occurrences dans l'échantillon au total, soit 7 %). À l'inverse, les phénomènes davantage associés à l'oral informel français sont : 1) l'anaphore résomptive, 2) l'anaphore temporelle, 3) la deixis temporelle et 4) la deixis textuelle. Sur le plan de la position de l'occurrence par rapport au verbe de la proposition, les deux différences majeures sont : 1) une sous-représentation de la position préverbale en français par comparaison avec le tchèque (19 vs 187 occurrences, soit 1,9 % vs 18,7 %) et 2) une sous-représentation de la position de complément de phrase en tchèque par comparaison avec le français (100 vs 356 occurrences, soit 10 % vs 35,6 %). La seconde différence est essentiellement due à la sous-représentation, en tchèque, des emplois temporels

²² Rappelons que la part des occurrences « émotionnelles » dans l'ensemble de l'échantillon français est extrêmement basse (11 occurrences). Outre la question quantitative, c'est la batterie d'effets produits par ce type de démonstratif qui est, comme nous l'avons vu, plus riche en tchèque qu'en français.

(déictique et anaphorique), car une affinité existe entre ceux-ci et la position de complément de phrase. Enfin, mentionnons encore 3) une plus forte représentation, en tchèque, de la position postverbale, ce qui est en partie dû à une plus forte représentation de l'emploi mémoriel (une affinité a été révélée entre cet emploi et la position postverbale). Les phénomènes que nous venons de rappeler sont autant de différences quantitatives et/ou qualitatives entre les oraux informels des deux langues. Si l'ensemble de ce qui vient d'être dit dans ce paragraphe ne relève pas de l'hypothèse principale, il n'en demeure pas moins qu'il s'agit, ici aussi, de constats s'apparentant à des différences qualitatives substantielles en matière d'emploi des démonstratifs adnominaux (voir *supra*).

S'agissant de la configuration des deux systèmes, nous sommes, dans les deux cas, face à un système reposant sur des oppositions de distance du référent par rapport au locuteur. Cependant, les deux systèmes diffèrent en ce sens que la forme CE française peut, le cas échéant, s'opposer aux deux formes comprenant la particule localisatrice CE *N-ci* et CE *N-là*, alors que le tchèque recourt à d'autres formes pour rendre les mêmes oppositions de distance. Pour la proximité, ce sont les formes TENTO/TENHLE, TENHLETEN, TADYTEN, TADYHLEN, etc. La capacité de la forme TEN à marquer le non-éloignement (plutôt que la proximité) en situation de deixis situationnelle est très limitée sur le plan diatopique. Pour l'éloignement, ce sont les formes TAMTEN, TAMHLETEN, etc. Ensuite, les formes françaises modifiées étant relativement rares (notamment avec la particule *-ci*), il semble possible de parler d'un système minimal, ne comprenant que CE tout seul²³. S'ajoute à cela le fait que *-là* ne marque pas toujours l'éloignement, sa fonction pouvant être de renforcer la sémantique indexicale de CE. En deixis situationnelle, l'opposition CE vs CE *N-là* semble d'ailleurs préférée à CE *N-ci* vs CE *N-là*. En revanche, TEN ne peut pas, la plupart du temps, fonctionner comme un démonstratif en deixis situationnelle, d'où la simplification, dans ce régime d'indexicalité, du système à deux termes au lieu de trois (voir *supra*). Or cette réduction ne concerne pas exclusivement la deixis situationnelle, car, en vérité, aucun régime d'indexicalité ne semble (du moins dans son fonctionnement courant) user d'une opposition à trois termes : en emploi mémoriel, les formes autres que TEN apparaissent très rarement (un seul cas dans nos données) ; en anaphore, les formes d'éloignement apparaissent très rarement, voire quasiment jamais (nos données n'en contiennent aucune), alors que les formes de proximité possèdent, avant tout, un rôle particulier qui consiste à maintenir le lien de coréférence en anaphores infidèle et indirecte²⁴ ou à résumer le contenu d'un segment plus large en anaphore résomptive ; enfin, en cataphore propre (renvoyant au-delà de la frontière d'un complexe nominal) et

²³ "Another possible example of a 'one-term' deictic system is French *ce* (/cette/cet), which does not encode any sort of distance distinction." (Anderson et Keenan, 1985 : 280)

²⁴ Plus généralement : « [Moyennant l'usage de *tento/tenhle*], la réalité nommée par l'antécédent est, au sein du nouvel énoncé, mise en relief en tant que son thème explicite et accentué, destiné à un nouveau jugement. » (Štícha, 2001 : 92) Nous traduisons : „[Identifikátorem *tento/tenhle* se] skutečnost pojmenovaná antedementem v nové výpovědi aktuálně vytýká jako její zdůrazněné a vytčené téma pro nový soud.“

en deixis temporelle, seules les formes de proximité semblent admises. En français, la situation n'est guère plus symétrique, car seule la deixis situationnelle semble pouvoir recourir à une opposition à trois termes, mais pas les autres régimes : l'emploi mémoriel n'utilise jamais, parmi nos données, de la particule *-ci* et assez rarement de la particule *-là* (12,90 % des occurrences mémorielles de notre échantillon) ; l'anaphore semble recourir très peu à *-ci* (seulement 4 occurrences anaphoriques temporelles dans notre échantillon) ; la deixis temporelle s'avère très pauvre aussi bien en *-ci* qu'en *-là*. Au sein de la catégorie des emplois émotionnels, les formes tchèques autres que TEN sont rares (seulement 2 occurrences dans nos données) ; en français, les formes modifiées par les particules semblent elles aussi rares (bien qu'il nous soit difficile d'appréhender cet aspect du fait de la représentation extrêmement faible de ce type d'emploi dans nos données françaises). Dans les contextes où le démonstratif intervient directement dans l'identification du référent (où il garde sa fonction « désignative ») tout en étant à l'origine d'un effet de subjectivité accrue, le tchèque emploie TEN (quel que soit le régime d'indexicalité), et le français emploie CE sans particule.

Rappelons encore les correspondances sur le plan de la valeur linguistique, dans les registres parlés informels de nos deux langues, entre les formes tchèques TEN et TENHLE (TENTO) et la forme française CE telles que le travail avec nos données les a mises au jour. Pour des raisons évidentes, il faut aussi inclure, dans le schéma, le défini français LE (le schéma est, bien entendu, une simplification, les autres possibilités des deux systèmes n'étant pas prises en compte) :

TEN → LE/CE

TENHLE (TENTO) → CE

CE → TEN/TEHLE (TENTO)

LE → Ø/TEN

Nos données du sous-corpus des pièces de V. Havel appuient en partie ce schéma : 54,55 % des occurrences de TEN considérées comme grammaticalisées sont rendues en français par LE, alors que 30,36 % de celles considérées comme non-grammaticalisées sont rendues par CE ; 47,62 % des occurrences des formes fortes du démonstratif tchèque sont rendues en français par CE. Les principales raisons du caractère partiel de ces correspondances ont aussi été identifiées, à savoir : une marge de liberté certaine des traducteurs, l'ambiguïté de la valeur de TEN en emploi mémoriel et l'absence d'un équivalent, en français, pour certains types d'emploi « émotionnel » de TEN.

Apports, limites et ouvertures du projet

Nous utiliserons cette dernière section aussi bien pour résumer les apports du présent travail que pour rappeler ses incontestables limites et failles. Nous finirons en suggérant des pistes pour une recherche future. D’abord, nous pensons que l’orientation contrastive de notre projet a permis d’approfondir notre compréhension du fonctionnement sémantico-référentiel de la classe des démonstratifs dans une perspective universelle. Cette perspective est, nous l’espérons bien, autant synchronique que diachronique. Pour le niveau synchronique, que soit uniquement mentionnée ici la nature de *désignateur* contextuel du démonstratif, dont découle toute une série de caractéristiques logico-sémantiques et de fonctions au sein du discours. Pour le niveau diachronique, c’est notamment le démonstratif en tant que source de l’article défini qu’il convient de relever, cette réalité universelle se manifestant aujourd’hui en tchèque sous notre regard de chercheur. Notre travail avec des productions langagières authentiques a ainsi permis d’étayer à sa manière l’existence de traits universels dans le comportement du démonstratif et, ce faisant, il contribue à une meilleure connaissance de cette unité linguistique forte intrigante. Mais outre ces traits universels, le travail a aussi dévoilé des différences marquées entre l’oral informel des deux langues, que ce soit dans la structuration du système ou dans le champ des emplois du démonstratif (son extension).

Ensuite, notre projet présente un intérêt sur le plan de la linguistique typologique. En effet, le tchèque et le français appartiennent chacun à un type de langue différent, synthétique pour le tchèque et analytique pour le français. Nous avons ainsi pu voir comment cette différence typologique se manifeste dans le domaine des démonstratifs, à commencer par l’existence d’une scission paradigmaticque, à la fois syntaxique et morphologique, dans le cas du français, qui s’oppose à l’absence d’une telle scission en tchèque. Cette différence se manifeste aussi par l’appartenance du démonstratif adnominal français à la classe des déterminants, avec tout ce que cela signifie en termes de contraintes syntactico-distributionnelles. À notre avis, le contraste entre les deux langues en matière de comportement syntaxique du démonstratif adnominal, mis au jour par ce travail, nous en dit long sur les divergences syntaxiques entre langues à articles et langues sans articles, nos recherches apportant une voix de plus à l’approche typologique représentée par des auteurs tels que Himmelmann et Diessel, à l’opposé de ceux qui défendent la possibilité d’appliquer la même analyse syntaxique à toutes ou à la plupart des langues du monde (cf. Hlavsa , Uhlířová, Veselovská). Cette comparaison syntaxique des deux systèmes, nous osons également y voir une dimension novatrice de notre travail. Enfin, nous espérons sincèrement que le présent projet a apporté une preuve supplémentaire de l’utilité d’une démarche linguistique basée sur des données de corpus de langues (et, plus spécifiquement encore, sur des données de corpus de langue parlée informelle). Nous avons nous-

même dû reconnaître après une phase initiale de scepticisme qu'aucune autre source ne vaut l'authenticité de ces dernières, a fortiori lorsqu'il est question d'étudier un phénomène étroitement lié à une langue informelle et en plein mouvement.

Les points faibles du projet sont malheureusement nombreux. Tout au long du travail, ils ont été relevés un par un et rappelés. La première série de points faibles a trait aux différences de nature qui opposent nos deux principales sources de données et, avant tout, aux failles du corpus créé à partir du CEFC. Ces failles sont notamment le manque d'équilibre, avec la surreprésentation d'une source particulière, et l'absence de nombreuses métadonnées relatives aux conversations. Mais il faut également rappeler un certain écart dans les tailles des deux corpus de conversations ainsi que les différences dans les choix des annotateurs. Il y a ensuite la nature plus ou moins heuristique de nos requêtes, notamment celle appliquée à ORTOFON v1. Celle-ci a inévitablement entraîné la non inclusion de certaines occurrences recherchées. Il y a aussi le caractère approximatif des chiffres portant sur la fréquence absolue et relative du démonstratif adnominal dans le corpus ORTOFON v1, du fait que ces chiffres sont compris dans un intervalle de confiance. Le sous-corpus des pièces de V. Havel comporte au moins deux faiblesses à la fois : sa très petite taille et ses écarts considérables par rapport à l'oral informel authentique. Nous avons également remarqué que l'auteur recourait parfois à des usages spécifiques de TEN adnominal (notamment à ses emplois « émotionnels ») comme à un moyen de caractérisation des personnages, d'où la nature en partie déséquilibrée de la répartition de ses occurrences dans les textes. La démarche ayant consisté à comparer les originaux tchèques avec leurs traductions françaises a notamment péché par son unidirectionnalité et par les écarts entre les deux versions, dus à une traduction libre.

D'une manière plus générale et indépendamment des sources utilisées, travailler avec des quantités de données conséquentes – ce qui était notre cas – comporte en soi-même des inconvénients. Il s'agit surtout du fait que le volume important de données rend rarement possible une bonne familiarisation avec celles-ci, ce qui exclut toute approche détaillée et oblige à un ciblage précis. Concrètement, dans notre cas, il a été nécessaire de cibler les seules occurrences du démonstratif adnominal, sans tenir suffisamment compte des autres expressions référentielles présentes dans les textes. Par conséquent, s'il était tout à fait possible d'appréhender les occurrences du démonstratif dans leur contexte, il aurait été irréaliste, en revanche, de vouloir examiner l'ensemble des chaînes de référence dont elles faisaient partie. Or l'impossibilité de cette approche soucieuse du détail empêche d'aborder certains facteurs importants, à commencer par la proportion des occurrences introduites par le démonstratif au sein de l'ensemble d'un régime d'indexicalité. Nous avons avant tout fait les frais de ces limites en étudiant la grammaticalisation de TEN en emploi anaphorique, où il aurait fallu, dans l'idéal, examiner la place des expressions démonstratives au sein de toutes les expressions se

rapportant au référent en question, tout en tenant compte d'une série de facteurs qui relèvent de l'analyse textuelle : topicalité du référent, fonction syntaxique, nature ontologique du référent, nombre de candidats potentiels à la reprise anaphorique, « cadres » ("frames") d'identification du référent (Ariel, 1988), distance exacte entre les maillons de la chaîne, etc. De même, ces limites nous ont empêché de dresser un bilan cohérent en ce qui concerne un éventuel lien, parmi les occurrences anaphoriques de notre échantillon de conversations tchèques, entre la présence de TEN adnominal et la fonction syntaxique de l'occurrence. C'est là qu'apparaissent les avantages de la démarche adoptée par exemple par Laury (1997) et par Zíková (2017), visant un seul type d'emploi du démonstratif. Une telle circonscription de la thématique de recherche permet de travailler avec un corpus de taille beaucoup plus limitée, dont le chercheur peut prendre connaissance jusque dans ses moindres détails et où il peut naviguer facilement. Mais nous persistons à croire que la démarche utilisée dans ce travail présente elle aussi des atouts, qui font défaut à la démarche mentionnée précédemment. Ces avantages sont avant tout une plus grande représentativité des phénomènes étudiés et, par conséquent, des conclusions. Ensuite, cette démarche permet d'analyser le phénomène dans une multitude de ses aspects et expressions plutôt que de devoir se cantonner à un seul. Dans le cas de la grammaticalisation de TEN, cela s'apparente à la possibilité d'acquérir une certaine idée holistique de la façon dont les choses se passent dans l'ensemble des emplois susceptibles d'être touchés par le processus. Dans l'idéal, il serait donc désirable de combiner les deux approches ici mentionnées, ce que nous tenterons de faire dans nos futurs projets de recherche. Dans le cadre de ce travail à venir, nous ne manquerons pas non plus de nous focaliser davantage sur une série de pistes identifiées ici, tels les principes – plus complexes qu'il n'y paraît – qui semblent régir l'emploi de TEN dans les anaphores associatives (la solidité du lien associatif, les différentes manières dont opère la restitution de l'argument établissant la définitude ainsi que les différents profils de ces arguments).

Les inconvénients cités se trouvent évidemment amplifiés dans notre travail par son orientation contrastive : outre les écueils déjà évoqués et qui ont trait à la comparabilité des sources, le fait de se pencher sur deux systèmes à la fois oblige à encore plus de ciblage et de sélectivité. Enfin, il est aussi des limites propres à la linguistique de corpus en général. Ces limites ont pour origine la nature même des productions langagières que les corpus contiennent : des productions achevées et figées dans leur forme, dont les auteurs restent pour nous anonymes et non joignables. Par conséquent, il est autant impossible d'interroger ces auteurs sur leurs choix et d'avoir un retour sur ces choix qu'il est impossible de conclure à l'absence d'un emploi particulier dans l'ensemble du système à partir de sa seule absence dans le corpus étudié, ce dernier permettant uniquement l'inverse, à savoir d'attester, à l'appui d'une quantité suffisante d'occurrences, qu'un emploi est bel et bien existant dans le système. Pour toutes ces raisons, une étude sur corpus comme la nôtre doit, à notre avis, être complétée par

des enquêtes menées auprès de locuteurs natifs, qu'il est possible d'interroger sur leurs choix et dont on peut en partie recueillir et analyser les réflexions. Dans nos futures activités de recherche, il nous tarde donc aussi de développer les pistes identifiées grâce aux retours de nos informateurs tirés des trois questionnaires ici utilisés²⁵.

La grammaticalisation de TEN adnominal : quelles perspectives ?

Revenons ici brièvement à la question déjà abordée dans la conclusion du Chapitre 6, à savoir celle des perspectives concernant la suite du processus de grammaticalisation de TEN adnominal en tchèque parlé informel de nos jours. Nos données nous ont permis de confirmer que TEN adnominal apparaît, dans ce registre, dans les contextes relevant de la définitude « pragmatique » et dans les contextes de transition entre les définitudes « pragmatique » et « sémantique », le tableau général se rapprochant vaguement d'une « scission du type I » (Ortmann, 2014). En même temps, TEN n'a pas encore vraiment atteint le statut de lexème obligatoire, si ce n'est en contexte situationnel (hormis les cas exceptionnels susmentionnés), voire aussi, en grande partie, en contexte mémoriel. Ce fait constitue évidemment un bémol non négligeable, dans la mesure où le caractère obligatoire de l'emploi de la forme est un critère important de son analyse en défini. Cette règle s'appliquerait à la totalité des unités issues d'un processus de grammaticalisation, ce dernier pouvant être considéré comme achevé lorsque

[...] pour exprimer une notion grammaticale, il y a obligation à employer le nouveau morphème, et seulement celui-ci [...].
(Marchello-Nizia, 2006c : 53)

[...] un morphème cesse d'exprimer la notion grammaticale qu'il marquait à l'origine [...]
(Marchello-Nizia, 2006c : 53)

S'agissant du second critère, il supposerait que TEN perde complètement son fonctionnement de démonstratif, son extension se limitant désormais aux seuls contextes qui accueillent l'article défini²⁶. Or ce n'est pas, tant s'en faut, ce que nous pouvons observer aujourd'hui, car TEN garde encore très souvent son sémantisme déictique, et ce non uniquement dans les emplois « émotionnels »²⁷. Inévitablement, se pose la question de la restructuration du système actuel dans

²⁵ Un exemple parmi d'autres : dans la Section 6.2.2.1, nous avons vu grâce au retour d'une informatrice que la préférence pour la présence ou l'absence de TEN devant un nom modifié par une relative « établissante » semble en partie dépendre du statut du référent pour l'interlocuteur : ce dernier est-il au courant de son existence ?

²⁶ Précisons cependant que ce critère cité par Marchello-Nizia est discutable, car il a été démontré par Hagège (1993) que le sens d'arrivée et le sens de départ peuvent continuer à exister dans le système côte à côte, phénomène que l'auteur nomme la « preuve par anachronie ».

²⁷ Même si force est de constater que si nous enlevons de l'ensemble des 967 occurrences de TEN de notre échantillon les occurrences en voie de grammaticalisation et les occurrences « émotionnelles », il ne reste que 185 occurrences de TEN non concernées par la grammaticalisation et n'affichant pas une valeur « émotionnelle », soit une quantité qui est loin

pareille conjoncture. Nous avons vu qu'à deux cas périphériques près – lorsque TEN anticipe une relative « établissante » du type « classificatoire » en contexte de cataphore intra-syntagmatique (emploi déterminatif) et lorsqu'il permet, dans certains dialectes, de désigner un référent présent dans le site d'énonciation par contraste avec un autre référent de la même classe (emploi déictique situationnel) – TEN adnominal n'est pas accentué. Dans l'hypothèse où il deviendrait un défini à part entière, comment serait-il, dès lors, distingué du démonstratif au sein des contextes où il peut également fonctionner comme tel²⁸ ? Nous croyons que deux scénarii peuvent ici être envisagés avec le plus de plausibilité. Ou bien la forme perdrait entièrement la possibilité d'être employée comme un démonstratif, ce qui entraînerait la généralisation de l'opposition entre deux sphères de distance à travers les formes renforcées, exactement comme c'est déjà le cas, la plupart du temps, en contexte situationnel ; ou bien la forme TEN se maintiendrait, avec cependant deux réalisations sonores différentes dans tous les contextes, à savoir une réalisation accentuée – démonstrative – et une réalisation non-accentuée – articulaire. Dans le premier scénario, il se produirait ce qui s'est produit en bulgare, où le démonstratif TEH (TEN), hérité du vieux slave, a cessé de l'être une fois grammaticalisé en défini, et où le système du démonstratif s'est structuré autour de l'opposition proche vs lointain (ТОЗИ vs ОНЗИ)²⁹. Dans le second scénario, la situation de TEN serait analogue à celle de la forme DER en allemand. Rappelons qu'une telle situation serait conforme aux observations faites à propos des articles définis dans les langues d'Europe centrale :

Schroeder (2006) shows that there is geographical continuum: in western European languages, the definite article is formally distinct from the demonstrative (e.g. English, Spanish, French); in central Europe the definite article has the same form as the demonstrative pronoun (e.g. German) [...] (Carlier et De Mulder, 2011 : 523)

Quoi qu'il en soit, il nous semble difficilement envisageable qu'une fois le processus arrivé au point où il serait effectivement possible de traiter le tchèque comme un système à articles définis, TEN puisse à la fois continuer de fonctionner comme un démonstratif et rester non-accentué, car un tel brouillage des valeurs défierait les lois structuralistes du langage³⁰.

d'être considérable ! Il semble ainsi possible d'en conclure que ce sont les emplois « émotionnels » qui constituent une part très élevée, dans le tchèque parlé informel d'aujourd'hui, de l'ensemble des emplois démonstratifs de TEN.

²⁸ Soit le contexte anaphorique (hormis l'anaphore associative), le contexte mémoriel et les contextes où il introduit un concept individuel ou fonctionnel lexical.

²⁹ *Idem* pour le passage entre le latin et l'ancien français : une fois que ILLE a donné le défini LE, il a perdu sa valeur originelle de démonstratif, étant évincé, dans ce rôle, par *ecce* + ILLE → CIL. Il semble cependant logique de supposer que cette réfection du système ne s'est pas faite du jour au lendemain, et qu'il fut ainsi un temps où ILLE fut ambigu sur le plan de la valeur sémantique, tantôt (déjà) plus proche du défini et tantôt (encore) plus proche du démonstratif. C'est cela même qui se produit dans le tchèque parlé informel de nos jours, où TEN est, selon le contexte, plus proche de l'un ou de l'autre. Toutefois, cette micro-confusion est tolérée par le système tant que celui-ci est en mouvement : une fois qu'il se sera de nouveau figé, le rétablissement de l'homéostasie exigera la levée de cette ambiguïté.

³⁰ Dans l'hypothèse où la forme serait tantôt accentuée et tantôt non, la situation serait suffisamment transparente, ce qui rendrait la polysémie de la forme acceptable. Une autre hypothèse, quelque peu extrême, certes, pourrait prévoir, pour les usages articulaires de TEN, un affaiblissement phonétique à ce point radical qu'il aboutirait à un affaiblissement morphologique de certaines de ses formes via la neutralisation ou la perte de la voyelle (/tən/, /tn/, etc.). À l'heure actuelle,

Rappelons nos conclusions 1) au sujet du schéma de grammaticalisation en « courbe en S » proposé par Kroch (1989) et 2) au sujet du schéma de grammaticalisation proposé par Heine (2002). Pour ce qui est du premier, la grammaticalisation de TEN semble, à l’heure actuelle, ne pas avoir avancé au-delà du début de la courbe (la fréquence des emplois est en hausse lente, la forme se répandant progressivement de contexte en contexte, avant de connaître une montée en flèche dans tous les contextes à la fois, incluant aussi ceux de définitude « sémantique »). En ce qui concerne le second schéma (A, Ab, Ba, B), TEN a entamé son parcours vers le stade Ba (le « contexte de passage »), où

[l]e sens initial est relégué à l’arrière-plan (mais toujours accessible), et le nouveau sens permet au mot d’apparaître dans des contextes tout à fait nouveaux, et pas seulement dans les contextes ambigus qui avaient permis son apparition. (Marchello-Nizia, 2006c : 23)

Voici à présent le Tableau 8.1, où nous proposons une synthèse des différentes étapes que TEN a atteintes et qu’il pourrait encore atteindre si sa grammaticalisation se poursuit dans l’avenir. Les stades atteints à l’heure actuelle sont marqués en rouge. Le séparateur indique le seuil que TEN doit franchir pour être considéré comme un article défini pleinement constitué :

STADE ATTEINT	SITUATION
DEMONSTRATIF PUR	En contextes de définitude « pragmatique », il permet d’identifier un référent via son indexation sur le contexte immédiat. La définitude qu’il marque procède directement de cette identification (donc d’un acte de référence). En contextes de définitude « sémantique », il présente un référent déjà défini (identifiable) sur fond du contexte immédiat, opération qui est à l’origine d’un effet « émotionnel ». ³¹
ARTICLE DEFINI « FORT » D’APPARITION NON-SYSTEMATIQUE	Son apparition est limitée aux contextes de définitude « pragmatique » et aux contextes transitoires entre les définitudes « pragmatique » et « sémantique ». Il ne contribue pas à l’identification du référent, mais marque l’unicité dénotative du nom et

une tendance de TEN en voie de grammaticalisation à un certain degré de réduction phonétique a déjà été relevée par Machač et Zíková (2014 : 99) : “Given that apparent phonetic reduction only appears in the last stage of grammaticalization, we cannot expect to observe any massive changes in the phonetic substance in this case. The demonstrative *ten* is quite obviously still used in its phonologically full form in formal speech. However, we may notice that it is frequently reduced in casual speech.”

³¹ En absence de référence (le plus souvent en emploi « émotionnel » de rappel), c’est la seule information qui est présentée sur fond du contexte.

	ses compléments. La définitude qu'il marque procède de l'unicité. Son usage n'est pas obligatoire au sein de (tous) ces contextes.
ARTICLE DEFINI « FORT » D'APPARITION SYSTEMATIQUE	Son apparition est limitée aux contextes de définitude « pragmatique » et aux contextes transitoires entre les définitudes « pragmatique » et « sémantique ». Il ne contribue pas à l'identification du référent, mais marque l'unicité dénotative du nom et ses compléments. La définitude qu'il marque procède de l'unicité. Son usage est obligatoire au sein de tous ces contextes.
ARTICLE DEFINI UNIVERSEL D'APPARITION NON-SYSTEMATIQUE	Son apparition a atteint les contextes de définitude « sémantique » pure (concepts individuels et fonctionnels lexicaux). Il ne contribue pas à l'identification du référent, mais marque l'unicité dénotative du nom et ses compléments. La définitude qu'il marque procède de l'unicité. Son usage n'est pas obligatoire au sein de (tous) ces contextes.
ARTICLE DEFINI UNIVERSEL D'APPARITION SYSTEMATIQUE	Son apparition a atteint les contextes de définitude « sémantique » pure (concepts individuels et fonctionnels lexicaux). Il ne contribue pas à l'identification du référent, mais marque l'unicité dénotative du nom et ses compléments. La définitude qu'il marque procède de l'unicité. Son usage est obligatoire au sein de tous ces contextes.
DES STADES PLUS AVANCES (?)	Son apparition pourrait s'étendre aux cas de la définitude « sémantique » rarement marqués par l'article dans les langues du monde (noms propres de personnes) et/ou à l'emploi générique pluriel. Une « scission du type II » (Ortmann, 2014) pourrait apparaître...

Tableau 8.1 : Étapes du processus de grammaticalisation de TEN adnominal avec les étapes actuelles (rouge) et le seuil de l'article défini constitué (séparateur gras)

Ainsi, à l'heure actuelle, dans les contextes où il est grammaticalisé, TEN fonctionne comme un article défini « fort » (cf. Studler, 2014). Toutefois, le terme d'article n'est pas tout à fait approprié, car la forme n'est pas toujours obligatoire³².

Cependant, la poursuite de la grammaticalisation risque-t-elle d'avoir lieu ? Voilà qui est une question fondamentale. Autrement dit, quelle est la probabilité pour TEN de devenir, dans l'avenir, un marqueur systématique de l'unicité dénotative du nom dans l'ensemble des contextes de définitude « pragmatique » ainsi que dans l'essentiel des contextes de définitude « sémantique » pure, et ce dans tous les registres de la langue ? Comme nous l'avons écrit dans la conclusion du Chapitre 6, la conjoncture du tchèque actuel ne semble guère favorable à ce scénario. Dans ce qui suit, nous allons reprendre l'essentiel des faits présentés à ce sujet dans la Section 6.3. D'abord, le tchèque est une langue dotée d'une flexion très riche et qui ne manifeste pas de signes de déclin : il n'y a pas d'effritement de la flexion, contrairement par exemple au latin tardif (cf. Herman, 1985). Au vu du lien constaté entre l'émergence des articles et le passage d'un système synthétique et non configurationnel à un système analytique et configurationnel (cf. Marchello-Nizia, 1995 ; Carlier et Combettes, 2015³³), ce fait n'est point encourageant pour le plein aboutissement du processus de grammaticalisation. Certes la position du démonstratif adnominal à l'intérieur du complexe nominal a connu un figement considérable par comparaison aux états plus anciens de la langue (voir *supra*). Pourtant, le tchèque reste un système largement non-configurationnel, où l'ordre des mots reste en grande partie « [...] régi par des facteurs pragmatiques plutôt que par la syntaxe [...] » (Carlier et Combettes, 2015 : 17). De même, les rapports entre les constituants sont plus souples que dans les langues analytiques³⁴.

Ensuite, il y a l'autorité déjà mentionnée de la norme soutenue, qui n'est pas sans avoir un impact retardataire sur l'évolution de l'article défini à partir de TEN. Rappelons que :

Dans le tchèque parlé, le déterminant *ten* possède une fréquence et des fonctions similaires, en position adnominale, à l'article défini *the* en anglais, même si dans le tchèque écrit, il n'est que rarement utilisé dans la détermination nominale. Les deux formes du tchèque s'influencent mutuellement, d'où probablement le fait

³² Contrairement à l'équivalent de TEN en haut-sorabe (TÓN), obligatoire aussi dans tous les emplois anaphoriques, devant un nom modifié par une relative « établissante » ainsi que (chez la jeune génération) dans tous les types d'anaphore associative et avec les « concepts individuels complexes » en position rhématique (Trovesi, 2002 ; Nědolužko, 2004 ; Czardybon, 2017).

³³ « [...] dans la perspective d'une cohérence quant au mode d'expression de la grammaire, l'affaiblissement progressif de la morphologie suffixale conduit à grammaticaliser davantage les marqueurs grammaticaux antéposés à la tête, à leur faire perdre le poids pragmatique. » (Carlier et Combettes, 2015 : 20) Au vu de ce constat, il n'est pas surprenant que les deux langues slaves disposant d'articles définis – le bulgare et le macédonien – soient aussi les seules avec un système casuel très rudimentaire.

³⁴ Pour ne citer qu'un exemple parmi d'autres, la coordination entre un possessif (élément fonctionnel) et un syntagme prépositionnel (complément du nom) reste possible :

<i>lékař</i>	můj	<i>a můch</i>	dět-í.
médecin.NOM.SG	mien.M.NOM.SG	et mien.N.GEN.PL	enfant-GEN.PL
<i>mon médecin et celui de mes enfants</i>			

que la détermination au moyen du pronom *ten* n'est pas figée, dans le tchèque parlé, de la même façon qu'elle l'est dans les langues à articles définis [...] ³⁵ (Cvrček *et al.*, 2015 [2010] : 176)

Par conséquent, la poursuite du processus de grammaticalisation semble très clairement conditionnée par une remise en question de cette autorité, au moins dans le domaine de l'usage informel de la langue. Dans la conjoncture où la norme soutenue « lâcherait complètement prise » sur le « tchèque commun », ce dernier pourrait davantage s'émanciper et poursuivre pleinement les tendances évolutives qui lui sont propres, entérinant ainsi la tendance à la diglossie qui caractérise le clivage entre cette variété et la langue soutenue (Čermák, 2014). Le rôle du « tchèque commun » est ici crucial pour la simple raison qu'il s'agit de la variété du registre informel utilisée par le plus grand nombre de locuteurs tchèques (voir la Section 3.1.1 du Chapitre 3).

Il n'est un secret pour personne que par comparaison au « tchèque commun », la langue soutenue – y compris dans sa forme parlée („hovorová čeština“) – constitue une variété en partie archaïque et globalement assez conservatrice, car procédant d'une tentative savante de figer la langue à un moment de son évolution, voire de revenir à des états passés :

Standard Czech (*spisovná čeština*) is then a semi-artificial creation, archaic in many respects, while the vernacular has continued to evolve since the norms (whether of the sixteenth or the nineteenth century) were set. There is a consequent tension between the modern literary language and the spoken Czech, usually known as Common Czech (*obecná čeština*), in which natural development has culminated. Between these two poles there are transitional strata, notably Colloquial Czech (*hovorová čeština*) [...] (Short, 1993 : 455-456)

De fait, la norme soutenue du tchèque moderne fut largement définie au moment de la Renaissance nationale tchèque, soit au tournant du 18^{ème} et du 19^{ème} siècle, moment où la langue tchèque fut en grande partie ressuscitée du coma où elle se trouvait suite à plus de cent cinquante ans de germanisation du pays. Ce mouvement fut essentiellement porté, du moins dans ses premières phases, par des intellectuels et des savants, avec pour figures emblématiques J. Dobrovský³⁶ pour la première phase, défensive, et J. Jungmann³⁷ pour la deuxième phase, offensive. Afin de pallier aux insuffisances des moyens lexicaux dont la langue souffrait, les auteurs (à commencer par Jungmann) recouraient le plus souvent à des emprunts à d'autres langues slaves qu'ils soumettaient ensuite aux règles morpho-phonétiques du tchèque, s'appliquaient à ranimer des termes ressortissant à des états

³⁵ Nous traduisons : „V muvené češtině má determinátor *ten* v pozici před substantivem podobnou frekvenci i podobné funkce jako určitý člen *the* v angličtině, v psané češtině se ale k určení substantiv užívá jen zřídka. Obě formy češtiny se ovlivňují, a patrně proto není determinace zájmenem *ten* v mluvené češtině ustálena jako v jazycích s určitým členem [...]“

³⁶ Avec notamment ses deux ouvrages pionniers consacrés à la grammaire tchèque : *Deutsch-böhmisches Wörterbuch* (1802 et 1821) et *Ausführliches Lehrgebäude der Böhmischen Sprache zur gründlichen Erlernung derselben für Deutsche, zur vollkommenen Kenntniss für Böhmen* (1809). Dobrovský ne croyait pas que le tchèque puisse jamais recouvrer l'ensemble de son domaine d'usage d'avant la germanisation.

³⁷ Avec notamment son immense œuvre lexicographique, soit un dictionnaire tchéco-allemand en cinq tomes (1835-1839), et ses traductions des œuvres phares de la littérature européenne (cf. *Atala* de F.-R. de Chateaubriand).

plus anciens de la langue et tombés dans l'oubli, ou bien ils forgeaient des néologismes (par dérivation, par composition ou à l'aide de calques). Dans le domaine de la morphologie et de la syntaxe, c'est le tchèque de la fin du 16^{ème} siècle, dit « humaniste », soit la variété soutenue et littéraire employée dans les œuvres des Frères tchèques, qui fut en grande partie pris pour modèle³⁸. Cette variété fut choisie pour son niveau de développement élevé et son très grand prestige, le tchèque « humaniste » représentant sans conteste, au moment de la Renaissance nationale, la dernière variété de référence en date :

Le tchèque moderne se constitua au début du 19^{ème} siècle. Dans les premières phases, ce fut notamment en tant qu'objet des tentatives de culture de la part des savants, qui prétendirent à la codification de son nouveau système orthographique, à la stabilisation de sa forme conformément à la norme de la langue évoluée de la fin du 16^{ème} siècle ainsi qu'à son introduction subséquente dans les domaines d'utilisation d'où il était sorti dans la période suivant la bataille de la Montagne-Blanche³⁹. (Kosek, 2017)

Les auteurs de la norme moderne prirent ainsi pour principale référence une norme certes hautement évoluée, mais assez conservatrice, car ne tenant pas compte d'un grand nombre d'évolutions plus récentes. Par conséquent, dès le moment où la nouvelle norme fut entérinée, une tension s'installa entre cette norme et la langue parlée, tension qui fut le plus marquée dans le domaine de la morphologie⁴⁰.

Dans la Section 6.3, nous avons écrit que l'autorité que la norme exerce sur le bon usage est particulièrement saillante dans l'éducation nationale, les médias de service public, les journaux et le monde de l'édition. Dans le Chapitre 3, nous avons vu qu'au sein des productions écrites, c'est la fiction contemporaine qui est, de loin, la plus ouverte à l'innovation et à l'accueil de la langue parlée informelle. L'influence de la norme soutenue est clairement moins prononcée dans des types de textes situés entre la communication officielle et officieuse, tels que les « textes écrits de la communication quotidienne » („schriftliche Texte der Alltagskommunikation“ ; Berger, 1993), les entretiens dirigés, les réunions de travail, etc. C'est dans tous ces types de textes, où le brassage entre la norme soutenue et la norme parlée informelle est le plus intense, que les emplois articulaires de TEN adnominal

³⁸ Il s'agit de la variété historique dans laquelle furent rédigés les travaux tchèques de Comenius ainsi que la traduction humaniste de la Bible (*La bible de Kralice*), réalisée par les Frères tchèques. Pour plus d'informations, voir Šlosar et Kosek (2017).

³⁹ Nous traduisons : „Nová č. se konstituovala na poč. 19. stol., v prvních fázích především jako objekt kultivační péče vzdělavců (obrozenců), kteří se snažili o kodifikaci jejího nového pravopisného systému, o stabilizaci její podoby podle normy vyspělého jaz. konce 16. stol. a o její následné uvedení do funkčních sfér, které v pobělohorském období opustila.“ Au sujet de la constitution de la norme linguistique à l'époque de la Renaissance nationale tchèque, voir aussi Šlosar (2017) ou Garvin (1983).

⁴⁰ À titre d'exemple, nous pouvons mentionner la diphtongue *-ej*, présente dans une série de morphèmes, dont notamment les désinences des nominatif et vocatif masculins singuliers des adjectifs durs (*mlad-ej*, contrairement à *mlad-ý* de la langue soutenue ; *jeune*). Cette diphtongue évolua, sur le territoire de la Bohême, à partir de la voyelle *y* dès le début du 15^{ème} siècle. Ainsi, le changement fut largement achevé au moment de la Renaissance nationale tchèque, dont les auteurs préférèrent toutefois revenir à la norme « humaniste », plus conservatrice sur ce point, à savoir le maintien de *y*. C'est comme cela que *-ej* devint un trait du « tchèque commun » plutôt que d'être élevé au statut de nouvelle norme.

pourraient s'imposer davantage avec le temps. Ensuite, dans une conjoncture qui reste, pour le moment, d'ordre purement hypothétique, ces types de textes pourraient un jour servir de pont entre la langue parlée informelle de tous les jours et la langue soutenue parlée et écrite. Mais il est aussi tout à fait envisageable que ce cap ne soit jamais franchi, les emplois articulaires continuant d'être majoritairement confinés à la langue parlée informelle⁴¹. Si tel est le cas, deux scénarii peuvent se produire. Soit, comme nous l'avons écrit *supra*, l'oral informel parviendra à s'affranchir pleinement de l'autorité de la norme soutenue, ce qui permettrait une poursuite « décomplexée » de la grammaticalisation de TEN, mais uniquement au sein du registre informel du tchèque⁴². Ce premier scénario aurait pour conséquence logique un renforcement du clivage déjà existant entre la langue soutenue et la langue parlée informelle (dont le « tchèque commun » représente la variété la plus influente⁴³). Soit l'oral informel ne parvient pas à s'émanciper du poids de la norme soutenue, ce qui débouchera sans doute sur le gel du processus dans sa phase actuelle, si bien que la grammaticalisation de TEN en restera, au sein même de l'oral informel, à la phase 2 du Tableau 8.1 (celle d'un article défini « fort » d'apparition non-systématique).

Dans l'éventualité où le processus de grammaticalisation se poursuivrait – que ce soit uniquement au sein de la langue (parlée) informelle ou dans l'ensemble des registres du tchèque –, les emplois « émotionnels » seraient susceptibles, répétons-le, de jouer un rôle important. En effet, nous avons déjà pu observer au sujet des « concepts individuels complexes » (notamment les noms rendus uniques par la présence d'un adjectif au superlatif) que les valeurs articulaire et « émotionnelles » de TEN peuvent en partie se combiner et se renforcer. Ainsi, il n'est pas exclu que plus tard, les valeurs « émotionnelles » de TEN puissent également, du fait de leur présence abondante dans la langue parlée informelle, encourager l'extension de sa valeur articulaire aux contextes de définitude « sémantique » pure. D'une part, ce scénario serait conforme à l'approche mettant l'accent sur le rôle joué par la subjectivité et par l'expressivité dans le processus de grammaticalisation (cf. Epstein, 1995 ; Traugott et Hopper, 2003 [1993]). D'autre part, il obéirait au principe d'attraction

⁴¹ C'est, évidemment, la politique linguistique du pays qui jouera un rôle décisif à cet égard.

⁴² La langue informelle écrite, présente notamment sur les réseaux sociaux, déjà bien plus « décomplexée » par rapport à la norme soutenue que ne l'est la langue des productions écrites ou orales de la communication semi-officielle, pourrait également contribuer à cet affranchissement.

⁴³ En tant que premier inter-dialecte du pays, le « tchèque commun » est, contrairement aux autres inter-dialectes et dialectes, de plus en plus perçu comme découplé d'une quelconque appartenance territoriale et/ou sociologique. Ainsi, la sous-variété du « tchèque commun » épurée d'influences régionales est de plus en plus reconnue par les linguistes (notamment praguais) comme le « second standard linguistique national » après le tchèque soutenu. Conformément à ce statut, il existe un certain continuum entre cette sous-variété du « tchèque commun » et la variété parlée du tchèque soutenu („hovorová čeština“), sachant que les traits de la première ont une influence considérable sur la seconde (cf. Krěmová et Chloupek, 2017). Précisons enfin que le phénomène de la grammaticalisation de TEN adnominal ne semble nullement se trouver aujourd'hui plus avancé ou moins avancé dans le « tchèque commun » que dans les autres inter-dialectes et dialectes, même si seule une étude bien centrée sur ces aspects diatopiques permettrait de l'affirmer avec certitude.

(« les emplois fréquents d'une forme attirent d'autres emplois de cette même forme »), mentionné dans les Sections 3.5.3 et 7.2. Cependant, tout cela reste, pour l'instant, de l'ordre de la spéculation.

Bibliographie, index et annexes

Bibliographie

LE DÉMONSTRATIF EN FRANÇAIS

EN SYNCHRONIE

BARBÉRIS J.-M., 1988, « Identité, ipséité dans la deixis spatiale : *ici* et *là*, deux appréhensions concurrentes de l'espace ? », *L'Information grammaticale*, [En ligne], 77/1, 28-32, consulté le 16 novembre 2017, URL : <http://www.persee.fr/doc/igram=0222-9838=1998=num=77=1=2872>.

BÉNARD J., 1998, « Démonstratifs insolents. De quelques emplois du démonstratif dans le texte célinien », *Langue française*, [En ligne], 120/1, 110-124, consulté le 25 juin 2016, URL : https://www.persee.fr/doc/lfr_0023-8368_1998_num_120_1_6273.

BONNARD H., 1972, « Les mots démonstratifs », in *Grand Larousse de la langue française*, 2, Paris, Librairie Larousse, 1206-1213.

CORBLIN F., 1983, « Défini et démonstratif dans la reprise immédiate », *Le français moderne*, 51/2, 118-133.

CORBLIN F., 1987, *Indéfini, défini et démonstratif : constructions linguistiques de la référence*, Genève/Paris, Droz.

DE MULDER W., 1998, « Du sens des démonstratifs à la construction d'univers », *Langue française*, [En ligne], 120/1, 21-32, consulté le 15 août 2017, URL : http://www.persee.fr/doc/lfr_0023-8368_1998_num_120_1_6266.

GARY-PRIEUR M.-N., 2001, « GN démonstratifs à référence générique : une généralité discursive », *Journal of French Language Studies*, 11/2, 221-239.

GARY-PRIEUR M.-N., 1998, « La dimension cataphorique du démonstratif. Étude de constructions à relative », *Langue française*, [En ligne], 120/1, 44-50, consulté le 13 septembre 2017, URL : http://www.persee.fr/doc/lfr_0023-8368_1998_num_120_1_6268.

GARY-PRIEUR M.-N. et NOAILLY M., 1996, « Démonstratifs insolites », *Poétique*, 105, 111-121.

GUÉNETTE L., 1995, *Le démonstratif en français : essai d'interprétation psychomécanique*, Paris, H. Champion.

GUILLOT C., 2004, « Entre anaphore et deixis : l'anaphore démonstrative à fonction résomptive », in TROTTER D. (dir.), *Actes du XXIV^e Congrès international de linguistique et de philologie romanes*, Aberystwyth, M. Niemeyer, 307-315.

GUILLOT C. (dir.), 2006b, « Le démonstratif en français », *Langue française*, 152/4.

GUILLOT C., 2006c, « Le démonstratif en français. Présentation », *Langue française*, 152/4, 3-8.

JONASSON K., 1998a, « Ce Marc nous fait bien bosser ! Sur le rôle du démonstratif devant le nom propre », in ENGLEBERT A. (dir.), PIERRARD M. (dir.), ROSIER L. (dir.) et VAN RAEMDONCK D. (dir.), *La ligne claire. De la linguistique à la grammaire. Mélanges offerts à Marc Wilmet à l'occasion de son 60^e anniversaire*, Louvain-la-Neuve, De Boeck Supérieur, 75-85.

JONASSON K., 2006, « Démonstratifs indéfinis en suédois et leurs correspondants en français », *Langue française*, [En ligne], 152/4, 24-38, consulté le 27 août 2017, URL : https://www.persee.fr/doc/lfr_0023-8368_2006_num_152_4_6633.

JONASSON K., 1998b, « Le déterminant démonstratif en français : un marqueur de quoi ? », in WILLEMS D. (dir.) et WILMET M. (dir.), *Travaux de linguistique*, 36, *Les marqueurs de hiérarchie et la grammaticalisation*, 59-70.

- KLEIBER G., 1986a, « Adjectif démonstratif et article défini en anaphore fidèle », in DAVID J. (dir.) et KLEIBER G. (dir.), *Déterminants : syntaxe et sémantique*, Paris, Klincksieck, 169-185.
- KLEIBER G., 2003, « Adjectifs démonstratifs et point de vue », *Cahiers de praxématique*, [En ligne], 41, 33-54, consulté le 26 août 2017, URL : <http://praxématique.revues.org/2638>.
- KLEIBER G., 1986b, « À propos de l'analyse Adjectif démonstratif = Article défini + Élément déictique ou Sur l'irréductibilité des symboles indexicaux », in *Actes du XVIIe Congrès International de linguistique et philologie romanes, Aix-en-Provence, 29 août-3 septembre 1983, Marseille, Université de Provence, Morphosyntaxe des langues romanes*, 4, 193-210.
- KLEIBER G., 1990, « Article défini et démonstratif : approche sémantique versus approche cognitive », in KLEIBER G. (dir.) et TYVAERT J.-E. (dir.), *L'anaphore et ses domaines*, Paris/Metz, Klincksieck, 199-227.
- KLEIBER G., 2006, « Démonstratifs : emplois à la mode et mode(s) d'emploi », *Langue française*, [En ligne], 152/4, 9-23, consulté le 15 septembre 2017, URL : <https://www.cairn.info/revue-langue-francaise-2006-4-page-9.htm>.
- KLEIBER G., 2007, « Des démonstratifs mémoriels aux démonstratifs de point de vue », in BEGIONIL. (dir.) et MULIER C. (dir.), *Problèmes de sémantique et de syntaxe. Hommage à André Rousseau*, Lille, Éditions du Conseil Scientifique de l'Université Charles de Gaulle-Lille 3, 373-392.
- KLEIBER G., 1995a, « D'ici à là et vice versa : pour les aborder autrement », *Le Gré des Langues*, 8, 8-27.
- KLEIBER G., 1991a, « Du nom propre non modifié au nom propre modifié : le cas de la détermination des noms propres par l'adjectif démonstratif », *Langue française*, [En ligne], 92/1, 82-103, consulté le 12 septembre 2017, URL : https://www.persee.fr/doc/lfr_0023-8368_1991_num_92_1_6213.
- KLEIBER G., 1995b, « Ici on ne peut pas utiliser là », in FIGUEIRA A. (dir.) et LAGO J. (dir.), *Estudios en homenaje a las profesoras Françoise Jourdan Pons e Isolina Sánchez Regueira*, Université de Saint-Jacques de Compostelle, Département de Philologie Française et Italienne, 133-146.
- KLEIBER G., 1983b, « Les démonstratifs (dé)montrent-ils ? Sur le sens référentiel des adjectifs et pronoms démonstratifs », *Le français moderne*, 51/2, 99-117.
- KLEIBER G., 1986c, « Pour une explication du paradoxe de la reprise immédiate », *Langue française*, [En ligne], 72/1, 54-79, consulté le 15 août 2017, URL : http://www.persee.fr/doc/lfr_0023-8368_1986_num_72_1_4658.
- KLEIBER G., 2004, « Sémantique, référence et discours : le cas des démonstratifs cataphoriques spécifiques », in AUCHLIN A. (dir.), *Structures et discours. Mélanges offerts à Eddy Roulet*, Québec, Éditions Nota bene, 231-245.
- KLEIBER G., 1984a, « Sur la sémantique des descriptions démonstratives », *Linguisticae Investigationes*, 8/1, 63-85.
- KLEIBER G. et SOCK R., 2006, « Ces + N + relative : Sémantique et prosodie », *Linguisticae Investigationes*, 29/2, 251-273.
- MAILLARD M., 1974, « Essai de typologie des substituts diaphoriques. Supports d'une anaphore et/ou d'une cataphore », *Langue française*, [En ligne], 21/1, 55-71, consulté le 17 juin 2018, URL : https://www.persee.fr/doc/lfr_0023-8368_1974_num_21_1_5665.
- PIERRARD M., 1991, « Observations sur le statut syntaxique et sémantique de CELUI + relative », *Travaux de linguistique*, 22, 69-91.
- PIERRARD M., 1990, « “Prenez votre bonheur en patience, de peur que celui qui est à venir ne vous consume tout à fait” : à propos de l'emploi phorique de celui », *Revue de linguistique romane*, 54, 189-210.
- PIERRARD M., 1989, « “Que celui qui n'a jamais péché ainsi me jette la première pierre” : emploi générique ou spécifique des relatives nominalisées introduites par celui », *Travaux de linguistique*, 19, 187-211.

REICHLER-BÉGUELIN M.-J., 1998, « L'usage des SN démonstratifs dans *les Fables* de La Fontaine », *Langue française*, [En ligne], 120/1, 95-109, consulté le 13 avril 2018, URL : https://www.persee.fr/doc/lfr_0023-8368_1998_num_120_1_6272.

SCHNEDECKER C., 2006, « SN démonstratifs “prédicatifs” : qu'est-ce qui limite leur apport informatif ? », *Langue française*, 152/4, 39-55.

THEISSEN A., 2008, « Le SN démonstratif *cette fois* (-ci/-là) », *Journal of French Language Studies*, [En ligne], 18/2, 209-226, téléchargé le 20 mai 2019, URL : <https://www.cambridge-org.acces.bibliotheque-diderot.fr/core/journals/journal-of-french-language-studies/article/le-sn-demonstratif-cette-fois-cila/5C3409B666AC60C5863D3500E6B90797#>.

VANDERBAUWHEDE G., 2012, « Les emplois référentiels du SN démonstratif en français et en néerlandais : pas du pareil au même », *French Language Studies*, [En ligne], 22/2, 273-294, téléchargé le 24 août 2017, URL : <https://www.cambridge.org/core/journals/journal-of-french-language-studies/article/les-emplois-referentiels-du-sn-demonstratif-en-francais-et-en-neerlandais-pas-du-pareil-au-meme/1D18DE82114D35AEF1AF7DFAA1205445>.

EN DIACHRONIE

AEBISCHER P., 1948, « Contribution à la protohistoire des articles ILLE et IPSE dans les langues romanes », *Cultura Neolatina*, 8, 181-203.

BANNIARD M., 1992, « *Ille* et son système : chronologie du développement (IIIe-VIIIe s.) », in CALLABAT L. (dir.), *Latin vulgaire, latin tardif*, IV, Hildesheim, Ols-Weidmann, 313-321.

BOLLÉE A., « Le développement du démonstratif dans les créoles de l'Océan Indien », *Creolica*, [En ligne], consulté le 3 mars 2020, URL : <http://www.creolica.net/article-34.html>.

BUZZONI M., 2012, *The Development of the Definite Article in Old High German and Old French: A Comparative Analysis*, thèse de doctorat, Université de Florence, [En ligne], consulté le 6 août 2018, URL : <http://dspace.unive.it/bitstream/handle/10579/2584/815421-1165049.pdf?sequence2>.

CABOLI G., 1990, « Les pronoms démonstratifs latins et la formation de l'article roman », *Revue de philologie, de littérature et d'histoire anciennes*, 64, 71-88.

CARLIER A. et COMBETTES B., 2015, « Typologie et catégorisation morphosyntaxique : du latin au français moderne », *Langue française*, 157/3, 15-58.

CARLIER A. et DE MULDER W., 2006, « Du démonstratif à l'article défini : le cas de *ce* en français moderne », *Langue française*, 152/4, 96-113.

CARLIER A. et DE MULDER W., 2007, « Les premiers stades de développement de l'article défini », in COMBETTES B. (dir.) et MARCHELLO-NIZIA C. (dir.), *Études sur le changement linguistique en français. Communications du Colloque « Diachro 2 »*, Paris, 15, 16, 17 janvier 2004, Nancy, Presses universitaires de Nancy, 85-113.

CARLIER A. et DE MULDER W., 2010, « The Emergence of the Definite Article in Late Latin : “Ille” in Competition with “Ipse” », in CUYKENS H. (dir.), DAVIDSE K. (dir.) et VAN DE LANOTTE L. (dir.), *Subjectification, Intersubjectification and Grammaticalization*, The Hague, De Gruyter Mouton, 241-276, [En ligne], téléchargé le 10 novembre 2017, URL : https://www.researchgate.net/publication/274256241_Carlier_A_W_De_Mulder_2010_The_emergence_of_the_definite_article_in_Late_Latin_ille_in_competition_with_ipse_In_H_Cuykens_K_Davidse_L_Van_de_Lanotte_Eds_Subjectification_intersubjectification_and_gram.

CARLIER A. et GUILLOT C., 2015, « Evolution des démonstratifs du latin : le passage d'un système ternaire à un système binaire », in CARLIER A. (dir.), GOYENS M. (dir.) et LAMIROY B. (dir.), *Le français en diachronie. Nouveaux objets et méthodes*, Bern, Peter Lang, 337-371.

- CARLIER A. et SIMONENKO A., 2020, « Between Demonstrative and Definite: A Grammar Competition Model of the Evolution of French *L*-Determiners », *Canadian Journal of Linguistics*, [En ligne], 65/3, 393-437, téléchargé le 17 mai 2021, URL : <https://muse-jhu-edu.acces.bibliotheque-diderot.fr/article/764992/pdf>.
- CHRISTOL A., 1994, « Ipse : articuloïde ou article dans la Peregrinatio ? » *Lalies*, 13, 143-153.
- DEES A., 1971, *Étude sur l'évolution des démonstratifs en ancien et moyen français*, Groningen, Wolters-Noordhoff Publishing.
- EPSTEIN R., 1995, « L'article défini en ancien français : l'expression de la subjectivité », *Langue française*, [En ligne], 107/1, 58-71, consulté le 121 novembre 2017, URL : http://www.persee.fr/doc/lfr_0023-8368_1995_num_107_1_5305.
- EPSTEIN R., 1994, « The Development of the Definite Article in French », in PAGLIUCA W. (dir.), *Perspective on Grammaticalization*, Amsterdam/Philadelphia, J. Benjamins Publishing Company.
- FOULET L., 1954, « L'effacement des adverbes de lieu : *Ici, là* et leur groupe », *Romania*, 75/300, 433-456.
- FRUYT M., 2003, « Anaphore, cataphore et deixis dans l'Itinerarium d'Egérie », in SOLIN H. (dir.), LEIWO M. (dir.) et HALLA-AHO H. (dir.), *Latin vulgaire, latin tardif VI: Actes du VI^e colloque international sur le latin vulgaire et tardif (Helsinki, 29 août – 2 septembre 2000)*, Hildesheim, Olms-Weidmann, 99-119.
- FRUYT M., 2010a, « Deictics and Endophors in the Diachrony of Latin », *Revue de Linguistique Latine du Centre Alfred Ernout De lingua Latina*, [En ligne], 5, consulté le 11 mars 2020, URL : <http://www.paris-sorbonne.fr/revue-de-linguistique-latine-du>.
- FRUYT M., 2010b, « L'emploi de *is, hic, iste, ille, ipse* en latin archaïque et classique », *Revue des Etudes latines*, 87, 44-75.
- GOUGENHEIM G (dir.), 1964, *L'élaboration du français fondamental (1^{er} degré)*, Philadelphia-New York.
- GUILLOT C., 2005a, « “Ceste parole” et “ceste aventure” dans *La Queste del Saint Graal* : Marques de structuration discursive et transitions narratives », *L'Information grammaticale*, [En ligne], 103/1, 23-27, consulté le 26 février 2017, URL : http://www.persee.fr/doc/igram_0222-9838_2004_num_103_1_2549.
- GUILLOT C., 2006a, « Démonstratif et deixis discursive : analyse comparée d'un corpus écrit en français médiéval et d'un corpus oral de français contemporain », *Langue française*, 152/4, 56-69.
- GUILLOT C., 2010a, « Le démonstratif de notoriété de l'ancien français : approche textuelle », in COMBETTES B. (dir.), *Le changement en français. Etudes de linguistique diachronique*, Bern, Peter Lang, 217-233.
- GUILLOT C., 2005b, « Le rôle des démonstratifs dans la cohésion textuelle au XVe siècle. Éléments de grammaire textuelle », *L'Information grammaticale*, [En ligne], 106/1, 45, consulté le 26 février 2017, URL : http://www.persee.fr/doc/igram_0222-9838_2005_num_106_1_3774.
- GUILLOT C., 2010b, « Les démonstratifs de l'ancien français : un système encore personnel ? », in NEVEU F. (dir.), MUNI TOKE V. (dir.), DURAND J. (dir.), KLINGLER T. (dir.), MONDANA L. (dir.) et PREVOST S. (dir.), *2^{ème} Congrès Mondial de Linguistique Française, 12-15 juillet 2010, La Nouvelle-Orléans*, 237-248.
- GUILLOT C., 2013, « Système des démonstratifs médiévaux et exemples de stratégies communicatives », *Journal of French Language Studies*, 23/2, 221-242.
- GUILLOT-BARBANCE C., 2017, *Le démonstratif en français : étude de sémantique grammaticale diachronique (9^{ème} – 15^{ème} siècles)*, Paris/Leuven, Peters.
- GUILLOT-BARBANCE C. et MARCHELLO-NIZIA C., 2015, « Spécialisation morpho-syntaxique et changement sémantique : le cas du démonstratif français », *Langue française*, 157/3, 79-109.

HARRIS M., 1977, « “Demonstratives”, “Articles” and “Third Person Pronouns” in French: Changes in Progress », *Zeitschrift für Romanische Philologie*, 93/3-4, 249-261.

KLEIBER G., 1984b, « L'opposition *cist / cil* en ancien français ou comment analyser les démonstratifs ? », *Revue de linguistique romane*, 51, 271-295.

KURZOVÁ-JEDLIČKOVÁ H., 1963, « Die Demonstrativa im Vulgärlatein, (4.-6. Jahrhundert) », *Acta Antiqua Academiae Scientiarum Hungaricae*, 11/1-2, 121-143.

MAILLARD M., 1990, « Du “ça” français au “sa” seychellois », in *Prépublication du Colloque de linguistique La deixis (Paris III, 8-9 Juin 1990)*, Paris, Paris-Sorbonne III, 483-494.

MARCHELLO-NIZIA C., 2006a, « Du subjectif au spatial : l'évolution des formes et du sens des démonstratifs en français », *Langue française*, [En ligne], 152/4, 114-126, consulté le 15 septembre 2017, URL : <https://www.cairn.info/revue-langue-francaise-2006-4-page-114.htm>.

MARCHELLO-NIZIA C., 2006b, « From Personal Deixis to Spatial Deixis: The Semantic Evolution of Demonstratives from Latin to French », in HICKMAN N. (dir.) et ROBERT S. (dir.), *Space in Languages, Linguistic Systems And Cognitive Categories*, Amsterdam, J. Benjamins Publishing Company, 103-120.

MARCHELLO-NIZIA C., 2004, « La sémantique des démonstratifs en ancien français : une neutralisation en progrès », *Langue française*, [En ligne], 141/1, 69-84, consulté le 16 septembre 2017, URL : <https://www.cairn.info/revue-langue-francaise-2004-1-page-69.htm>.

MARCHELLO-NIZIA C., 1995, *L'évolution du français : ordre des mots, démonstratifs, accent tonique*, Paris, A. Colin.

MARCHELLO-NIZIA C., 2003, « “Se voz de ceste ne voz poez oster, Je voz ferai celle teste coper.” (*Ami et Amile*) : la sphere du locuteur et la deixis en ancien français », in VANNESTE A. (dir.), DE WILDE P. (dir.), KINDT S. (dir.) et VLEMINGS J. (dir.), *Memoire en temps advenir. Hommage à Theo Vennikeer*, 22, Louvain, Paris, Peeters, 413-427.

ORLANDINI A., 1981, « Wesen und Entwicklung des Artikels vom Lateinischen zu den Romanischen Sprachen », *Indogermanische Forschungen*, [En ligne], 86, 223-247, consulté le 16 juin 2018, URL : <https://www.degruyter.com/view/journals/indo/86/1981/indo.86.issue-1981.xml>.

PRICE G., 1969, « La transformation du système français des démonstratifs », *Zeitschrift für Romanische Philologie*, 85/5-6, 459-505.

PRICE G., 1968, « Quel est le rôle de l'opposition CIST/CIL en ancien français ? », *Romania*, 89/354, 240-253.

SELIG M., 1992, *Die Entwicklung der Nominaldeterminanten im Spätlatein: romanischer Sprachwandel und lateinische Schriftlichkeit*, Tübingen, Gunter Narr.

VINCENT N., 2014, « The Emergence of the D System in Romance », in VAN KEMENADE A., *Parameters of Morphosyntactic Change*, Cambridge, Cambridge University Press, 149-169, [En ligne], téléchargé le 10 novembre 2017, URL : <https://www.researchgate.net/publication/240462109=Théemergencéof=théD=system=inRomance>.

LE DÉMONSTRATIF EN TCHÈQUE ET DANS D'AUTRES LANGUES

ADAMEC P., 1983, « České zájmeno *ten* a jeho ruské ekvivalenty », in HRABĚ V. (dir.) et ŠIROKOVÁ A. G. (dir.), *Konfrontační studium ruské a české gramatiky a slovní zásoby*, II, Prague, Université Charles, 153-170.

BĚLIČ J., 1970, « Hele, tenhle tudle hnedle takhle... », *Naše řeč*, [En ligne], 53/4-5, 228-235, consulté le 28 mai 2018, URL : <http://nase-rec.ujc.cas.cz/archiv.php?art=5512>.

BERGER T., 1993, *Das System der tschechischen Demonstrativpronomina. Textgrammatische und stilspezifische Gebrauchsbedingungen*, thèse d'habilitation non publiée, München, Ludwigs-Maximilians-Universität, [En ligne], téléchargé le 14 février 2018 URL :

https://www.academia.edu/3398045/Das_System_der_tschechischen_Demonstrativpronomina_-_Textgrammatische_und_stilspezifische_Gebrauchsbedingungen_unpublizierte_Habilitationsschrift.

BRUGMANN K., 1904, *Die Demonstrativpronomina der indogermanischen Sprachen*, Leipzig, Teubner, [En ligne], consulté le 14 juin 2018, URL : <http://digital.slub-dresden.de/werkansicht/dlf/4268/16/>.

CORBLIN F. et ASIC T., 2006, « Une nouvelle approche de l'opposition *ici/là* et *ovde/tu* », *Travaux de linguistique*, 72/1, 29-48.

DIESEL H., 1999, *Demonstratives. Form, Function, and Grammaticalization*, Amsterdam, J. Benjamins Publishing Company.

DIESEL H., 2006, « Demonstratives, Joint Attention, and the Emergence of Grammar », *Cognitive linguistics*, 17/4, 463-489.

DIXON R. M., 2003, « Demonstratives. A cross-linguistic typology », *Studies in Language*, 27/1, 61-112.

DVOŘÁK J., à paraître, « L'emploi du démonstratif TEN avec l'adjectif superlatif en tchèque : entre simple marquage de la définitude et effets pragmatiques », *Scolia*, 35, 13-31.

DVOŘÁK J., 2020, « The Emerging Definite Article *Ten* in (Informal Spoken) Czech: A Further Analysis in Terms of Semantic and Pragmatic Definiteness », *Naše řeč*, 103/4, 297-319.

ENFIELD N. J., 2003, « Demonstratives in Space and Interaction: Data from Lao Speakers and Implications for Semantic Analysis », *Language*, 79/1, 82-117.

FOSSARD M., 2006, « Aspects psycholinguistiques du traitement des démonstratifs : résultats "croisés" en français et en anglais », *Langue française*, 152/4, 82-95.

HAUENSCHILD C., 1982, « Demonstrative Pronoun in Russian and Czech – Deixis and Anaphora », in WEISENBORN J. (dir.) et KLEIN W. (dir.), *Here and There. Crosslinguistic Studies on Deixis and Demonstration*, Amsterdam, J. Benjamins Publishing Company, 167-186.

HIMMELMANN N., 1997, *Deiktikon, Artikel, Nominalphrase. Zur Emergenz syntaktischer Struktur*, thèse de doctorat, Tübingen, Niemeyer.

HIMMELMANN N. P., 1996, « Demonstratives in Narrative Discourse: A Taxonomy of Universal Use », in FOX B. (dir.), *Studies in Anaphora*, Amsterdam/Philadelphia, J. Benjamins Publishing Company, 205-254.

HIRSCHOVÁ M., 1988, « Netypické případy užití ukazovacích výrazů *takový, ta* », *Naše řeč*, [En ligne], 71/2, 57-61, consulté le 12 mai 2018, URL : <http://nase-rec.ujc.cas.cz/archiv.php?art=6750>.

HRDLÍČKA M., « On a ten », 1997, *Naše řeč*, [En ligne], 80/3, 60-61, consulté le 26 février 2017, URL : <http://nase-rec.ujc.cas.cz/archiv.php?art=7389>.

JODAS J., 2010, « Regionální difference v užívání ukazovacích zájmen », in KOMÁREK M. (dir.), *Bohemica Olomucensia*, 3, Olomouc, Universitě Palacký, 88-93.

KOMÁREK M., 1978, « Sémantická struktura deiktických slov v češtině », *Slovo a slovesnost*, [En ligne], 39/1, 5-14, consulté le 12 mai 2018, URL : <http://sas.ujc.cas.cz/archiv.php?art=2497>.

KORTLANDT F., 1983, « Demonstrative Pronouns in Balto-Slavic, Armenian, and Tocharian », *Studies in Slavic and General Linguistics*, [En ligne], 3, 311-322, téléchargé le 29 mai 2018, URL : <https://www.jstor.org/stable/40986726?seq=1>.

KRESIN S., 2008, « Demonstrative Modification of Proper Nouns: A Corpus-based Study », in ŠTÍCHA F. (dir.) et FRIED M. (dir.), *Gramatika a korpus 2007 (Sborník příspěvků ze stejnojmenné konference 25.-27. 9. 2007, Liblice)*, Prague, Academia, 243-249, [En ligne], téléchargé le 8 janvier 2020, URL : <https://escholarship.org/content/qt5zw5v3d4/qt5zw5v3d4.pdf>.

- KŘÍŽKOVÁ H., 1971, « Zájmena typu *ten* a *takový* v současných slovanských jazycích », *Slavica Slovaca*, 6/1, 15-30.
- LAKOFF R., 1974, « Remarks on This and That », *Proceedings of the Chicago Linguistics Society*, Chicago, the University of Chicago, 345-356.
- LAURY R., 1997, *Demonstratives in Interaction. The Emergence of a Definite Article in Finnish*, Amsterdam/Philadelphia, J. Benjamins Publishing Company.
- MATHESIUŠ V., 1926, « Přívlastkové *ten*, *ta*, *to* v hovorové češtině », *Naše řeč*, [En ligne], 10/2, 39-41, consulté le 26 février 2017, URL : <http://nase-rec.ujc.cas.cz/archiv.php?art=1316>.
- MEYERSTEIN Z. P., 1972, « Czech Deictics: Pronouns and Articles? », *Linguistics*, [En ligne], 91/10, 17-30, téléchargé le 21 mars 2018, URL : <https://www.degruyter.com/downloadpdf/j/ling.1972.10.issue-91/ling.1972.10.91.17/ling.1972.10.91.17.pdf>.
- NĚDOLUŽKO A., 2004, « Několik poznámek k používání ukazovacího zájmena *tón* ve funkci určitého členu v běžně mluvené horní lužičtině v konfrontaci s češtinou », in POŘÍZKA P. (dir.) et POLÁCH V. P. (dir.), *Vztah langue a parole v perspektivě „interaktivního obratu“ v lingvistickém zkoumání*, Olomouc, Universitě Palacký, 155-158.
- PORTELE Y. et BADER M., 2016, « Accessibility and Referential Choice: Personal Pronouns and D-pronouns in Written German », *Discours*, [En ligne], 18, consulté le 5 janvier 2019, URL : <https://journals.openedition.org/discours/9188>.
- POTTS C. et SCHWARZ F., 2009, « Affective “this” », *Linguistic Issues in Language Technology*, [En ligne], 3/5, 1-30, téléchargé le 10 novembre 2017, URL : <https://pdfs.semanticscholar.org/9e55/74b126b632be16720f9f09e114ad00daa6df.pdf>.
- PRINCE E. F., 1981, « On the Inferencing of Indefinite-this NPs », in JOSHI A. (dir.), WEBER B. H. (dir.) et SAG I. A. (dir.), *Elements of Discourse Understanding*, Cambridge, Cambridge University Press, 231-250.
- SCHNEIDEROVÁ E., 1993, « K užívání zájmena *ten* (v přívlastkové pozici) v mluvených projevech », *Naše řeč*, [En ligne], 76/1, 31-37, consulté le 24 février 2017, URL : <http://nase-rec.ujc.cas.cz/archiv.php?art=7107>.
- ŠIMÍK R., 2015, « On Pragmatic Demonstratives: The Case of Pragmatic Discourse Anaphora in Czech », *Proceedings of Sinn und Bedeutung*, [En ligne], 20, téléchargé le 26 février 2019, URL : <https://semanticsarchive.net/sub2015/SeparateArticles/Simik-SuB20.pdf>.
- SOUČEK M., 2011, « Má čeština určitý člen? Komparativní analýza funkcí určitého členu *the* v angličtině a zájmen *ten*, *ta*, *to* v češtině », in ČERMÁK F. (dir.), *Korpusová lingvistika Praha 2011. 1: InterCorp*, Prague, Nakladatelství Lidové noviny, 64-72.
- ŠTÍCHA F., 2001, « Anaforické koreferenční substantivum », in HLADKÁ Z. (dir.) et KARLÍN P. (dir.), *Čeština. Univerzália a specifiká*, 3, Brno, MU, 87-97.
- ŠTÍCHA F., 1999, « K deikticko-anaforickým funkcím lexému *ten* », *Slovo a slovesnost*, [En ligne], 60/2, 123-135, consulté le 23 février 2017, URL : <http://sas.ujc.cas.cz/archiv.php?art=3836>.
- TROVESI A., 2002, « Určenost nominální skupiny v češtině, horní lužické srbštině a slovinštině », in KRAUSOVÁ A. (dir.), SLEZÁKOVÁ M. (dir.) et SVOBODOVÁ Z. (dir.), *Setkání s češtinou (Sborník z konference Setkání s češtinou konané v Praze 6.-7. září 2001)*, Prague, ÚJČ AV ČR, 129-136.
- UHLÍŘOVÁ L., 1992, « *Ten* nějaký // nějaký *ten* a případy podobné », *Naše řeč*, [En ligne], 75/5, 247-254, consulté le 8 décembre 2017, URL : <http://nase-rec.ujc.cas.cz/archiv.php?art=7098>.
- UHLÍŘOVÁ L., 2006, « Tento vs. *ten*: co vypovídá korpus? », in ŠTÍCHA F. (dir.), *Možnosti a meze české gramatiky*, Prague, Academia, 268-281.
- VORÁČ J., 1949, « *Ten*, *tento* », in HAVRÁNEK B. (dir.), JEDLIČKA A. (dir.) et VÁHALA F. (dir.), *Jazykový koutek československého rozhlasu, první výběr*, Prague, Orbis, 229-231.

ZÍKOVÁ M., 2018, « Anaforická funkce výrazu ten v gramatikalizační perspektivě », *Studie z aplikované lingvistiky/Studies in Applied Linguistics*, 9, 106–123.

ZÍKOVÁ M., 2017, *Gramatikalizační potenciál anaforické funkce lexému ten v mluvených narativech*, thèse de doctorat non publiée, Prague, Université Charles de Prague, [En ligne], téléchargé le 15 mars 2019, URL : <https://dspace.cuni.cz/handle/20.500.11956/91732>.

ZÍKOVÁ M. et MACHAČ P., 2014, « Variability in Phonetic Realization of the Demonstrative Ten in Terms of Its Informational Relevance in the Sentence », in VESELOVSKÁ L. (dir.) et JANEBOVÁ M. (dir.), *Nominal Structures: All in Complex DPs*, Olomouc, Palacký University, 98-109.

ZÍKOVÁ M. et VOLÍN J., 2012, « Přízvučné a nepřízvučné *ten* při odkazování k vedlejším přívlastkovým větám », *Naše řeč*, 95/1, 33-44.

ZUBATÝ J., 1916, « Ten », *Naše řeč*, 1/10, 289-294.

ZUBATÝ J., 1920, « Ten nejlepší člověk », *Naše řeč*, [En ligne], 4/3, 74-76, consulté le 8 décembre 2017, URL : <http://nase-rec.ujc.cas.cz/archiv.php?art=1639>.

INDEXICALITÉ

ANDERSON S. R. et KEENAN E. L., 1985, « Deixis », in SHOPEN T. (dir.), *Language Typology and Syntactic Description. Volume 3. Grammatical Categories and the Lexicon*, Cambridge, Cambridge University Press.

APOTHÉLOZ D., 1995, *Rôle et fonctionnement de l'anaphore dans la dynamique textuelle*, Genève, Droz.

BARBÉRIS J.-M., 1992, « Sous la grammaire des expressions spatiales : la ville vécue », *Cahiers de praxématique*, 18, Montpellier, Pulm, 105-124.

BAR-HILLEL Y., 1954, « Indexical Expressions », *Mind*, 63, 359-379.

BEAUZÉE N. et BARTLETT B. (dir.), 1974 [1767], *Grammaire générale : ou exposition raisonnée des éléments nécessaires du langage, pour servir de fondement à l'étude de toutes les langues*, Stuttgart-Bad Cannstatt, F. Frommann.

BENVENISTE E., 1974, « Appareil formel de l'énonciation », in *Problèmes de linguistique générale*, 2, Paris, Gallimard.

BÜHLER K., 2009 [1934], *Théorie du langage : la fonction représentationnelle*, Marseille, Agone.

BURKS A. W., 1948, « Icon, Index, Symbol », *Philosophy and Phenomenological Research*, 9/4, 673-689.

CLARK E. V., 1978, « From Gesture to Word: On the Natural History of Deixis in Language Acquisition » in BRUNER J.S. (dir.) et GARTON A. (dir.), *Human Growth and Development*, Oxford, Oxford University Press, 85-120.

CONEIN B., 1984, « L'enquête sociologique et l'analyse du langage : les formes linguistiques de la connaissance sociale », *Arguments ethnométhodologiques, Problèmes d'épistémologie en sciences sociales*, 3, Paris, Écoles des Hautes Études en Sciences Sociales, 5-30.

CORBLIN F., 1995, *Les formes de reprise dans le discours*, Presses Universitaires de Rennes, Rennes.

CORBLIN F., 1985, « Remarques sur la notion d'anaphore », *Revue québécoise de linguistique*, 15/1, 173-195.

CORNISH F., 2010, « Indexicaux, discours et mémoire discursive : ce que les premiers révèlent du second et de la troisième », *Lynx*, [En ligne], 62-63, 111-133, consulté le 26 février 2017, URL : <https://linx.revues.org/1377>.

CORNISH F., 2001, « L'anaphore pronominale indirecte : une question de focus », in DE MULDER W. (dir.), VET C. (dir.) et VETTERS C. (dir.), *Anaphores pronominales et nominales : études pragma-sémantiques*, Amsterdam/New York, Rodopi, 1-25.

- DAMOURETTE J. et PICHON É., 1936, *Des mots à la pensée : Essai de grammaire de la langue française. Tome 5*, Paris, J. L. L. d'Artrey.
- DANON-BOILEAU L., 1990, « Il y a deixis et deixis : considérations cursives sur les limites du fonctionnement déictique de *le* et du fonctionnement anaphorique de *ce* », in KLEIBER G. (dir.) et TYVAERT J.-E. (dir.), *L'anaphore et ses domaines*, Paris, Klincksieck, 97-109.
- DECOOL-MERCIER N. et AKINCI M.-A., 2010, *Le fonctionnement des anaphores dans les textes oraux et écrits en français d'enfants bilingues et monolingues, Psycholinguistique et acquisition*, in NEVEU F. (dir.), MUNI TOKE V. (dir.), DURAND J. (dir.), KLINGLER T. (dir.), MONDANA L. (dir.), PREVOST S. (dir.), *2^{ème} Congrès Mondial de Linguistique Française, 12-15 juillet 2010, La Nouvelle-Orléans*, [En ligne], téléchargé le 26 mai 2019, URL : https://www.linguistiquefrancaise.org/articles/cmlf/abs/2010/01/cmlf2010_000090/cmlf2010_000090.html.
- DE MULDER W. (dir.), VET C. (dir.) et VETTERS C. (dir.), 2001, *Anaphores pronominales et nominales : études pragma-sémantiques*, Amsterdam/New York, Rodopi.
- ERBENOVÁ A., 2016, « Volba determinantů v nominální anafóře – srovnávací situace ve španělštině a češtině », *Časopis pro moderní filologii*, [En ligne], 98/2, 181-198, téléchargé le 30 mars 2018, URL : https://dspace.cuni.cz/bitstream/handle/20.500.11956/97202/1482617_adela_erbenova_181-198.pdf?sequence=1&isAllowed=y.
- ERKÜ F. et GUNDEL J.K., 1987, « The Pragmatics of Indirect Anaphors », in VERSCHUEREN J. (dir.) et BERTUCCELLI-PACI (dir.), *The Pragmatic Perspective: Selected Papers from the 1985 International Pragmatics Conference*, Amsterdam, J. Benjamins Publishing Company, 533-545.
- FILLMORE C., 1997, *Lectures on Deixis*, Stanford, CSLI Publications.
- FILLMORE C. J., 1982, « Towards a Descriptive Framework for Spatial Deixis », in JARVELLA R. (dir.) et KLEIN W. (dir.), *Speech, Place and Action. Studies in Deixis and Related Topics*, Chichester, John Wiley, 31-59.
- FRASER T. et JOLY A., 1980, « Le système de la deixis. Endophore et cohésion discursive en anglais », *Modèles linguistiques*, 2/2, 22-51.
- FRASER T. et JOLY A., 1979, « Le système de la deixis. Esquisse d'une théorie d'expression en anglais », *Modèles linguistiques*, 1/2, 97-157.
- FREI H., 1944, « Systèmes de déictiques », *Acta Linguistica*, 4/1, 111-129.
- FRICKE E., 2003, « Origo, Pointing, and Conceptualization – What Gestures Reveal about the Nature of the Origo in Face-to-face Interaction », in LENZ F. (dir.), *Deictic conceptualisation of Space, Time and Person*, Amsterdam/Philadelphia, J. Benjamins Publishing Company, 69-93.
- HALLIDAY M. et HASAN R., 1976, *Cohesion in English*, London, Longman.
- HANKS W., 1992, « The Indexical Ground of Deictic Reference », in DURANTI A. (dir.) et GOODWIN C. (dir.), *Rethinking Context as an Interactive Phenomenon*, Cambridge, Cambridge UN, 43-76.
- HLAVSA Z., 1975, *Denotace objektu a její prostředky v současné češtině*, Prague, Academia.
- JAKOBSON R., 1963a, « Les embrayeurs, les catégories verbales et le verbe russe », in *Essais de linguistique générale*, Paris, Éditions de Minuit, 176-196.
- JESPERSEN O., 1922, *Language: its Nature, Development and Origin*, London, G. Allan & Unwin.
- JESPERSEN O., 1992 [1924], *The Philosophy of Grammar*, Chicago/London, University of Chicago Press.

- JUNGBLUTH K., 2003, « Deictics in the Conversational Dyad: Findings in Spanish and Some Cross-linguistic Outlines », in LENZ F. (dir.), *Deictic Conceptualisation of Space, Time and Person*, Amsterdam/Philadelphia, J. Benjamins Publishing Company, 13-40.
- JUNGBLUTH K. (dir.) et DA MILANO F. (dir.), 2015, *Manual of Deixis in Romance Languages*, Berlin/Boston, De Gruyter.
- KARA M. et WIEDERSPIEL B., 2011, « Anaphore résomptive conceptuelle et mémoire discursive : entre identité et altérité », *Itinéraires*, [En ligne], 2, 79-93, consulté le 13 novembre 2017, URL : <https://itinéraires.revues.org/134>.
- KERBRAT-ORECCHIONI C., 2014 [1999], *L'énonciation. De la subjectivité dans le langage*, Paris, A. Colin.
- KESIK M., 1989, *La Cataphore*, Paris, Presse universitaire de France.
- KLEIBER, G., 1991b, « Anaphore-deixis : où en sommes-nous ? », *L'Information grammaticale*, [En ligne], 51/1, 3-18, consulté le 18 septembre 2017, URL : http://www.persee.fr/doc/igram_0222-9838_1991_num_51_1_3231.
- KLEIBER G., 1986d, « Déictiques, embrayeurs, “token-réflexives”, symboles indexicaux, etc. : comment les définir ? », *L'Information grammaticale*, [En ligne], 30/1, 3-22, consulté le 18 septembre 2017, URL : http://www.persee.fr/doc/igram_0222-9838_1986_num_30_1_2122.
- KLEIBER G., 1988, « Peut-on définir une catégorie générale de l'anaphore ? », *Vox Romanica*, [En ligne], 47, 1-14, téléchargé le 13 novembre 2017, URL : <https://www.e-periodica.ch/cntmng?pid=vxr-001:1988:47::349>.
- KLEIBER G., 1989, « Références indirecte ou de la divergence sur les anaphores divergentes », *Cahiers de praxématique*, [En ligne], 12, téléchargé le 26 mai 2019, URL : <https://journals.openedition.org/praxématique/3450>.
- LENZ F. (dir.), 2003, *Deictic Conceptualisation of Space, Time and Person*, Amsterdam/Philadelphia, J. Benjamins Publishing Company.
- LYONS J., 1977, « Deixis, Space and Time », in LYONS J., *Semantics*, 2, Cambridge, Cambridge University Press, 636-724.
- LYONS J., 1990 [1978], *Sémantique linguistique*, Paris, Larousse.
- MEIRA S., 2003, « “Addressee Effects” in Demonstrative Systems. The Case of Tiriyo and Brazilian Portuguese », in LENZ F. (dir.), *Deictic Conceptualisation of Space, Time and Person*, Amsterdam/Philadelphia, J. Benjamins Publishing Company, 3-11.
- MEOLA C. di, 2003, « Non-deictic Uses of the Deictic Motion Verbs *Kommen* and *Gehen* in German », in LENZ F. (dir.), *Deictic Conceptualisation of Space, Time and Person*, Amsterdam/Philadelphia, J. Benjamins Publishing Company, 41-67.
- PERDICOYANNI-PALÉOLOGOU H., 2001, « Le concept d'anaphore, de cataphore et de deixis en linguistique française », *Revue québécoise de linguistique*, [En ligne], 29/2, 55-77, téléchargé le 15 novembre 2018, URL : <https://www.erudit.org/fr/revues/rql/2001-v29-n2-rql3575/039441ar.pdf>.
- PEŠEK O., 2014, « Nominální anafora a determinace – kontrastivní analýza francouzských a českých systémových možností », *Časopis pro moderní filologii*, 96/2, 147-164.
- PEŠEK O., 2012, « Progression thématique et anaphorisation. L'apport informationnel des procédés de reprise », *Écho des études romanes*, [En ligne], 8/1, České Budějovice, l'Université de Bohême du sud, 217-228, téléchargé le 20 juillet 2018, URL : <https://www.eer.cz/files/2012-1/2012-1-17-Pesek.pdf>.
- REICHENBACH H., 1966 [1947], *Elements of Symbolic Logic*, London, the Macmillan Company.
- REICHLER-BÉGUELIN M.-J., 1988, « Anaphore, cataphore et mémoire discursive », *Pratiques*, [En ligne], 57/1, 15-43, consulté le 16 février 2020, URL : https://www.persee.fr/doc/prati_03382389_1988_num_57_1_1470.

RICŒUR P., 1975, *La métaphore vive*, Paris, Éditions du Seuil.

VUILLAUME M., 1980, *La déixis en allemand*, thèse de doctorat d'État, Paris, Université Paris IV.

WIEDERSPIEL B., 1989, « Sur l'anaphore : du modèle "standard" au modèle "mémoriel" », *Travaux de linguistique et de philologie*, 27, 95-113.

ZIMOVA L., 1994, *Způsoby vyjadřování větných členů v textu: konkurence pojmenování, pronominalizace a elize*, Ústí nad Labem, Université Jan Evangelista Purkyně.

GRAMMATICALISATION

ANDERSEN H., 2001, « Actualization and the (Uni)directionality of Change », in ANDERSEN (dir.), *Actualization: Linguistic Change in Progress. Papers from a Workshop Held at the 14th International Conference on Historical Linguistics, Vancouver, B.C., 14 August 1999*, 225-248.

CARLIER A. et DE MULDER W., 2011, « The Grammaticalization of Definite Articles », in HEINE B. (dir.) et NARROG H. (dir.), *The Oxford Handbook of Grammaticalization*, Oxford, Oxford University Press, 522-535.

EPSTEIN R., 1993, « The Definite Article: Early Stages of Development », in VAN MARLE J. (dir.), *Historical Linguistics 1991. Papers from the 10th International Conference on Historical Linguistics, Amsterdam, 12-16 August 1991*, Amsterdam, J. Benjamins Publishing Company, 111-134.

GIVÓN T., 1971, « Historical Syntax and Syntactic Morphology: an Archaeologist's Field Trip », *Chicago Linguistic Society*, 7, 394-415.

GIVÓN T., 2015, *The Diachrony of Grammar. Volumes 1 et 2*, Amsterdam/Philadelphia, J. Benjamins Publishing Company.

GREENBERG J., 1978a, « How Does a Language Acquire Gender Markers? » in GREENBERG J. (dir.), *Universals of Human Language. Word Structure*, 3, Stanford, Stanford University Press, 47-82.

HAGEGE, C., 1993, *The Language Builder*, Amsterdam/Philadelphia, J. Benjamins Publishing Company.

HEINE B., 2002, « On the Role of Context in Grammaticalization », in WISCHER I. (dir.) et DIEWALD G. (dir.), *New Reflections on Grammaticalisation*, Amsterdam/Philadelphia, J. Benjamins Publishing Company, 83-101.

HEINE B. et KUTEVA T., 2002, *World Lexicon of Grammaticalization*, New York, Cambridge University Press.

HEINE B. et REH M., 1984, *Grammaticalization and Reanalysis in African Languages*, Hamburg, Helmut Buske Verlag.

KROCH A., 1989, « Reflexes of Grammar in Patterns of Language Change », *Language Variation and Change*, 1/3, 199-244.

KURYLOWICZ J., 1965, « The Evolution of Grammatical Categories », *Diogenes*, [En ligne], 13/51, 55-77, téléchargé le 3 septembre 2020, URL : <https://journals-sagepub-com.acces.bibliotheque-diderot.fr/doi/pdf/10.1177/039219216501305105>.

LANGACKER R. W., 1985, « Observations and Speculations on Subjectivity », in HAIMAN J. (dir.), *Iconicity in Syntax*, Amsterdam, J. Benjamins Publishing Company, 109-150.

LEHMANN C., 2015[1982], *Thoughts on Grammaticalization, 3rd Edition*, Berlin, Language Science Press.

MARCHELLO-NIZIA C., 2009, *Grammaticalisation et Changement Linguistique*, Louvain-La-Neuve, De Boeck Supérieur.

MARCHELLO-NIZIA C., 2000, « Les grammaticalisations ont-elles une cause ? Le cas de *beaucoup*, *moult* et *très* en moyen français », *L'Information grammaticale*, [En ligne], 87/1, 3-9, consulté le 12 février 2021, URL : https://www.persee.fr/doc/igram_0222-9838_2000_num_87_1_2738.

MEILLET A., 1982 [1912], « L'évolution des mots grammaticales », in MEILLET A., *Linguistique historique et linguistique générale*, Paris/Genève, H. Champion-Slatkine, 131-148.

PEYRAUBE A., 2002, « L'évolution des structures grammaticales », *Langages*, 36/146, 46-58.

TRAUGOTT E. C., 2003, « From Subjectification to Intersubjectification », in HICKEY R. (dir.), *Motives for Language Change*, Cambridge, Cambridge University Press, 124-139.

TRAUGOTT E. C., 2010, « (Inter)subjectivity and (Inter)subjectification: A reassessment », in CUYKENS H. (dir.), DAVIDSE K. (dir.) et VAN DE LANOTTE L. (dir.), *Subjectification, Intersubjectification and Grammaticalization*, The Hague, De Gruyter Mouton, 29-74.

TRAUGOTT E. C., 1980, « Meaning-change in the Development of Grammatical Markers », *Language Sciences*, 2/1, 44-61.

TRAUGOTT E. C., 2005 [1995], « Subjectification and Grammaticalisation », in STEIN D. (dir.) et WRIGHT, S. (dir.), *Subjectivity and Subjectivisation*, Cambridge, Cambridge University Press, 31-54.

TRAUGOTT E. C., 1992, « Syntax », in HOGG R. M. (dir.), *The Cambridge History of the English Language. Vol. 1. The Beginnings to 1066, 168-289*, Cambridge, Cambridge University Press.

TRAUGOTT E. C. et DASHER R. B., 2002, *Regularity in Semantic Change*, Cambridge, Cambridge University Press

TRAUGOTT E. C. (dir.) et HEINE B. (dir.), 1991, *Approaches to Grammaticalization. Volume 1. Theoretical and Methodological Issues*, Amsterdam/Philadelphia, J. Benjamins Publishing Company.

TRAUGOTT E. C. et HOPPER P. J., 2003 [1993], *Grammaticalization*, Cambridge, Cambridge University Press.

TRAUGOTT E. C. et KOENIG E., 1991, « The Semantics-pragmatics of Grammaticalization Revisited », in TRAUGOTT E. C. (dir.) et HEINE B. (dir.), *Approaches to Grammaticalization. Volume 1. Theoretical and Methodological Issues*, Amsterdam/Philadelphia, J. Benjamins Publishing Company.

TYPLOGIE LINGUISTIQUE

CROFT W., 2002, *Typology and Universals*, Cambridge, Cambridge University Press.

DANYLENKO A. (dir.) et NOMACHI (dir.), 2019, *Slavic on the Language Map of Europe: Historical and Areal-Typological Dimensions*, Berlin/Boston, Mouton de Gruyter.

DRYER M. S. (dir.) et HASPELMATH M. (dir.), 2013, « Czech », in *The World Atlas of Language Structures Online*, Leipzig, Max Planck Institute for Evolutionary Anthropology, [En ligne], consulté le 7 octobre 2020, URL : https://wals.info/languoid/lect/wals_code_cze.

GREENBERG J. H. (dir.), 1978b, *Universals of Human Language. Volume 4. Syntax*, Stanford, Stanford University Press.

HALE K., 1983, « Warlpiri and the Grammar of Non-configurational languages », *Natural Language and Linguistic Theory*, 1, 5-47

HASPELMATH M., 2001, *The European Linguistic Area: Standard Average European*, in HASPELMATH M. (dir.), KOENIG E. (dir.), OESTERREICHER W. (dir.) et RAIBLE W. (dir.), *Language Typology and Language Universals. Volume 2*, Berlin/New York, Mouton de Gruyter, 1492-1510.

HAWKINS J. A., 1983, *Word Order Universals*, New York, Academic Press.

DÉFINITUDE, SÉMANTIQUE RÉFÉRENTIELLE, PRAGMATIQUE, SYNTAXE (APPROCHES THÉORIQUES)

ABNEY, S. P., 1987, *The English Noun Phrase in its Sentential Aspect*, thèse de doctorat non publiée, Massachusetts Institute of Technology, Cambridge MA, [En ligne], téléchargé le 7 janvier 2020, URL : <http://www.vinartus.net/spa/87a.pdf>.

ARIEL M., 1988, « Referring and Accessibility », *Journal of Linguistics*, 24/1, 65-87.

AUSTIN J. L., 1962, *How to Do Things With Words*, Oxford, Oxford University Press.

BARTON D., KOLB N. et KUPISCH T., 2015, « Definite Article Use with Generic Reference in German: An Empirical Study », *Zeitschrift für Sprachwissenschaft*, [En ligne], 34/2, 147-173, consulté le 18 mars, URL : <https://www.degruyter.com/view/j/zfs.2015.34.issue-2/zfs-2015-0009/zfs-2015-0009.xml>.

BAUER B., 2007, « The Definite Article in Indo-European: Emergence of a New Grammatical Category », in STARK E. (dir.), LEISS E. (dir.) et ABRAHAM W. (dir.), *Nominal Determination: Typology, Context Constraints, and Historical Emergence*, Amsterdam, J. Benjamins Publishing Company, 103-139.

BLANCHE-BENVENISTE C. et CHERVEL A., 1966, « Recherches sur le syntagme substantif », *Cahiers de lexicologie*, 9/2, 3-37.

BOŠKOVIĆ Ž., 2004, « Be Careful Where You Float Your Quantifiers », *Natural Language & Linguistic Theory*, 22/4, 681-742.

BOŠKOVIĆ Ž., 2005, « Left Branch Extraction, Structure of NP, and Scrambling », *Studia Linguistica*, 59, 1-45.

BROWN C., 1983, « Topic Continuity in Written English Narrative », in GIVÓN T. (dir.), *Topic Continuity in Discourse: A quantitative Cross-language Study*, Amsterdam, J. Benjamins Publishing Company, 313-341.

CARLSON G., SUSSMAN R., 2005, « Seemingly Indefinite Definites », in KEPSAR S. (dir.) et REIS M. (dir.), *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*, 71-86, Berlin, Mouton de Gruyter.

CHAROLLES M., 2002, *La référence et les expressions référentielles en français*, Paris, Orphys.

CHAROLLES M. (dir.) et PERY-WOODLEY M.-P. (dir.), 2005, *Langue française*, [En ligne], 148/4, *Les adverbiaux cadratifs*, consulté le 12 février 2020, URL : https://www.persee.fr/issue/lfr_0023-8368_2005_num_148_4.

CHAROLLES M. et VIGIER D., 2005, « Les adverbiaux en position préverbale : portée cadrative et organisation des discours », *Langue française*, [En ligne], 148/4, 9-30, consulté le 12 février 2020, URL : https://www.persee.fr/doc/lfr_0023-8368_2005_num_148_4_6604.

CHOMSKY N., 1970, *Le langage et la pensée*, Paris, Payot.

CHRISTOPHERSEN P., 1939, *The Articles: A Study of Their Theory and Use in English*, Copenhagen, Munksgaard.

CLARK H. H., 1975, « Bridging », in SCHANK R. C. (dir.) et NASH-WEBBER (dir.), *Theoretical Issues in Natural Language Processing*, Association for Computing Machinery, New York.

CORBLIN F., 2013, « Weak Definites as Bound Relational Definites », *Recherches linguistes de Vincennes*, [En ligne], 42, 91-122, consulté le 24 février, URL : <https://journals.openedition.org/rlv/2184>.

CORVER N., 1990, *The Syntax of Left Branch Extractions*, thèse de doctorat non publiée, Tilburg, Tilburg University, [En ligne], téléchargé le 27 février 2018, URL : <http://norbert.abelcorver.com/wp-content/uploads/2010/10/CorverPhDdissertation19901.pdf>.

- CZARDYBON A., 2017, *Definiteness in a Language without Articles – A Study on Polish*, thèse de doctorat, Düsseldorf, Düsseldorf University Press, [En ligne], téléchargé le 26 février 2017, URL : http://dup.oa.hhu.de/579/1/DiLaC_Bd.%203_Czardybon_Inhalt_9783957580474_final.pdf.
- CZARDYBON A., 2010, *Die Verwendung des definiten Artikels im Oberschlesischen im Sprachvergleich*, unpublished MA thesis, University Düsseldorf.
- DONNELLAN K., 1966, « Reference and Definite Descriptions », *Philosophical Review*, [En ligne], 75/3, 281-304, téléchargé le 10 novembre 2018, URL : [https://www.uvm.edu/~lderosse/courses/lang/Donnellan\(1966\).pdf](https://www.uvm.edu/~lderosse/courses/lang/Donnellan(1966).pdf).
- DOWTY D. et BRODIE B., 1984, *The Semantics of Floated Quantifiers in a Transformationless Grammar*, in COBLER M. (dir.) et MACKAYE S. (dir.), *Proceedings of WCCFL*, 3, 75-90.
- DRYER M. S., 1989, « Article-noun Order », *Chicago Linguistic Society*, 25, 83-97.
- DUCROT O., 1980a, « Descriptions définies et présupposés existentiels », in *Dire et ne pas dire. Principes de sémantique linguistique*, Paris, Hermann, 221-240.
- DUCROT O., 1972, *Dire et ne pas dire, principe de sémantique linguistique*, Paris, Hermann.
- DUCROT O., 1987 [1984], *Le dire et le dit*, Paris, Éditions de Minuit.
- DUCROT O., 1980b, *Les mots du discours*, Paris, Éditions de Minuit.
- DUCROT O., 1969, « Présupposés et sous-entendus », *Langue Française*, [En ligne], 4/1, 30-43, consulté le 19 février 2021, URL : https://www.persee.fr/doc/lfr_0023-8368_1969_num_4_1_5456.
- DUCROT O. et TODOROV T., 1972, *Dictionnaire encyclopédique des sciences du langage*, Paris, Éditions du Seuil.
- EBERT K., 1971, *Referenz, Sprachsituation und die bestimmten Artikel in einem nordfriesischen Dialekt (Fehring)*, Bredstedt (Bräist), Nordfriisk Institut.
- EPSTEIN R., 2002, « The Definite Article, Accessibility, and the Construction of Discourse Referents », *Cognitive Linguistics*, 12, 333-378.
- FIRBAS J., 1999, *Functional Sentence Perspective in Written and Spoken Communication*, Cambridge, Cambridge University Press.
- FLAUX N. (dir.), VAN DE VELDE D. (dir.) et DE MULDER W. (dir.), 1997, *Entre général et particulier : les déterminants*, Arras, Artois Presses Université.
- FODOR, J. D. et SAG I., 1982, « Referential and Quantificational Indefinites », *Linguistics and Philosophy*, [En ligne], 5, 355-398, téléchargé le 30 novembre 2018, URL : <https://link.springer.com/content/pdf/10.1007%2FBF00351459.pdf>.
- FREGE F. L. G., 1892, « Über Sinn und Bedeutung », *Zeitschrift für Philosophie und philosophische Kritik*, [En ligne], 100, 22-50, consulté le 20 février 2021, URL : <http://www.gavagai.de/HHP31.htm>.
- FRETHEIM T. (dir.) et GUNDEL J. K. (dir.), 1996, *Reference and Referent Accessibility*, Amsterdam/Philadelphia, J. Benjamins Publishing Company.
- FUCHS C. 1987, « Les relatives et la construction de l'interprétation », *Langages*, [En ligne], 22/88, 95-127, consulté le 28 octobre 2017, URL : http://www.persee.fr/doc/lgge_0458-726x_1987_num_22_88_2069.
- FUGIER H., 1983, « Le syntagme nominal en latin classique », in HAASE W. (dir.), *Aufstieg und Niedergang der römischen Welt*, 2, 29/1, Berlin/New York, Mouton de Gruyter, 212-269.
- GALMICHE M., 1989, « À propos de la définitude », *Langages*, [En ligne], 24/94, 7-37, consulté le 5 août 2018, URL : https://www.persee.fr/doc/lgge_0458-726x_1989_num_24_94_1544.

- GALMICHE M., 1983, « Les ambiguïtés référentielles ou les pièges de la référence », *Langue française*, [En ligne], 57, 60-86, consulté le 9 août 2018, URL: <http://www.jstor.org/stable/41558134>.
- GALMICHE M., 1979, « Quelques remarques sur l'exploitation linguistique de la notion de description définie », *Linx*, [En ligne], 1/1, 1-78, consulté le 11 octobre 2017, URL : http://www.persee.fr/doc/linx_0246-8743_1979_num_1_1_925.
- GALTON H., 1973, « The Function of the Definite Article in some Indo-European Languages: the Grammatical Category of Determinacy », *Linguistics*, [En ligne], 11/107, 5-13, téléchargé le 23 novembre 2017, URL : <https://www.degruyter.com/downloadpdf/j/ling.1973.11.issue-107/ling.1973.11.107.5/ling.1973.11.107.5.pdf>.
- CARLIER A. et GOYENS M., 1998, « De l'ancien français au français moderne : régression du degré zéro de la détermination et restructuration du système des articles », *Cahiers de l'Institut de Linguistique de Louvain*, 24/3-4, 77-122.
- GARY-PRIEUR M.-N., 1994, *Grammaire du nom propre*, Paris, P.U.F.
- GARY-PRIEUR M.-N., 2011, *Les déterminants du français*, Paris, Éditions Ophrys.
- GARY-PRIEUR M.-N., 2005 « Où il est montré que le nom propre n'est (presque) jamais "modifié" », *Langue française*, [En ligne], 146/2, 53-66, consulté le 11 octobre 2017, URL : <https://www.cairn.info/revue-langue-francaise-2005-2-page-53.htm>.
- GEACH P. T., 1962, *Reference and Generality: An Examination of some Medieval and Modern Theories*, London, Cornell University Press.
- GERLAND D. et HORN C., 2010, « Referential Properties of Nouns Across Languages », KANG Y.-S. (dir.), YOON J.-Y. (dir.), HONG J. (dir.), WU J.-S. (dir.), RHEE S. (dir.), KIM K.-A. (dir.), CHOI D.-H. (dir.), KIM K.-H. (dir.) et KANG H.-K. (dir.), *Universal Grammar and Individual Languages. Proceedings of SICOL*, Seoul, University of Korea, 133-150, [En ligne], téléchargé le 13 octobre 2018, URL : <http://dup.oa.hhu.de/503/2/BeitragGerland.pdf>.
- GIVÓN T., 1978, « Definiteness and Referentiality », in GREENBER J. H. (dir.), *Universals of Human Language. Volume 4 : Syntax*, Stanford, Stanford University Press.
- GIVÓN T., 2001, *Syntax: An Introduction. Volume 1*, Amsterdam, J. Benjamins Publishing Company.
- GIVÓN T. (dir.), 1983, *Topic Continuity in Discourse: A Quantitative Cross-language Study. Typological Studies in Language*, 3, Amsterdam/Philadelphia, J. Benjamins Publishing Company.
- GOYENS M., 1994, *Émergence et évolution du syntagme nominal en français*, Bern, Peter Lang.
- GRICE H.P., 1975, « Logic and Conversation », COLE P. (dir.) et MORGAN J. (dir.), *Syntax and Semantics*, 3, Academic Press, 22-40.
- GRICE H. P., 1979 [1975], « Logique et conversation », *Communications*, [En ligne], 30/1, 57-72, consulté le 18 février 2021, URL : https://www.persee.fr/doc/comm_0588-8018_1979_num_30_1_1446.
- GUILLAUME G., 1919, *Le problème de l'article et sa solution dans la langue française*, Paris, Hachette.
- GUIRAUD P., 1967, « L'assiette du nom dans *La chanson de Roland* », *Romania*, 88/349, 59-83.
- HARRIS M., 1978, *The Evolution of French Syntax: A Comparative Approach*, London, Longman.
- HARRIS M., 1980, « The Marking of Definiteness in Romance », in FISIÁK J. (dir.), *Historical Morphology*, 141-156, The Hague, Mouton.
- HARTSHORN C. (dir.) et WEISS P. (dir.), 1960-1966, *Collected Papers of Charles Sanders Peirce*, Cambridge, Harvard University press.

- HAWKINS J. A., 1978, *Definiteness and Indefiniteness*, London, Croom Helm.
- HAWKINS J. A., 1991, « On (In)definite Articles: Implicatures and (Un)grammaticality Prediction », *Journal of Linguistics*, [En ligne], 27/2, 405-422, téléchargé le 26 octobre 2018, URL : <https://www.jstor.org/stable/pdf/4176125.pdf>.
- HEIM I. R., 2011 [1982], *The Semantics of Definite and Indefinite Noun Phrases*, thèse de doctorat, Abingdon-on-Thames, Taylor & Francis Group. [En ligne], téléchargé le 13 novembre 2018, URL : <https://semanticsarchive.net/Archive/jA2YTJmN/Heim%20Dissertation%20with%20Hyperlinks.pdf>.
- HERMAN J., 1985, « La disparition de la déclinaison latine et l'évolution du syntagme nominal », in C. TOURATIER (dir.), *Latin et syntaxe*, Aix-en-Provence, Presses de l'Université de Provence, 345-360.
- HIMMELMANN N., 2001, « Articles », in HASPELMATH M. (dir.), KOENIG E. (dir.), OESTERREICHER W. (dir.) et RAIBLE W. (dir.), *Language Typology and Language Universals. Volume 2*, Berlin/New York, Mouton de Gruyter, 831-841.
- HOFHERR P. C., 2014, « Reduced Definite Articles with Restrictive Relative Clauses », in HOFHERR P. C. (dir.) et ZRIBI-HERTZ A. (dir.), 2014, *Crosslinguistic Studies on Noun Phrase Structure and Reference*, Leiden, Brill.
- HOFHERR P. C. (dir.) et ZRIBI-HERTZ A. (dir.), 2014, *Crosslinguistic Studies on Noun Phrase Structure and Reference*, Leiden, Brill.
- JACKENDOFF R., 1977, *X-Bar Syntax: A Study of Phrase Structure*, Cambridge, MIT Press.
- JOHNSEN L.-A., 2010, « Les pronoms “neuters” et leur référence à des procès en français parlé », *Lynx*, [En ligne], 62-63, 153-178, consulté le 4 mars 2017, URL : <https://linx.revues.org/1382>.
- JONASSON K., 1994, *Le Nom propre : Constructions et interprétations*, Louvain-la-Neuve, Duculot.
- KARLÍK P. et ŠIMÍK R., 2017, « Vztažná věta s jmennou hlavou », in NEKULA M. (dir.) et PLESKALOVÁ (dir.), *Nový encyklopedický slovník češtiny*, [en ligne], consulté le 29 mars 2021, URL : <https://www.czechency.org/slovník/VZTA%C5%BDN%C3%81%20V%C4%9ATA%20S%20JMENNOU%20HLAVOU> U.
- KALBERTODT J., PRIMUS B. et SCHUMACHER P. B., 2015, « Punctuation, Prosody, and Discourse: Afterthought vs. Right Dislocation », *Frontiers in Psychology*, 1803/6, 1-12.
- KAMP H. et REYLE U., *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*, Dordrecht, London, Boston, Kluwer Academic Publishers.
- KARTTUNEN L., 1976, « Discourse Referents », in McCAWLEY J. D. (dir.), *Syntax and Semantics, 7: Notes from the Linguistic Underground*, New York/London/San Francisco, Academic Press, 363-385, [En ligne], téléchargé le 15 novembre 2018, URL : https://www.researchgate.net/publication/2482414_Discourse_Referents.
- KARTTUNEN L., 1968a, « What Do Referential Indices Refer to? », publié sur RAND Corporation, Report No. P3854, [En ligne], téléchargé le 14 novembre 2018, URL : <https://www.rand.org/pubs/papers/P3854.html>.
- KARTTUNEN L., 1968b, « What Makes Definite Noun Phrases Definite? », publié sur RAND Corporation, Report No. P3871, [En ligne], téléchargé le 15 novembre 2018, URL : <https://www.rand.org/content/dam/rand/pubs/papers/2008/P3871.pdf>.
- KLEIBER G., 2004, « Peut-on sauver un sens de dénomination pour les noms propres ? », *Functions of Language*, 11/1, 115-145.
- KLEIBER G., 1981a, *Problèmes de référence : descriptions définies et noms propres*, thèse de doctorat, Paris, Klincksieck.

- KLEIBER G., 1994, « Qu'est-ce qui est (in)défini ? », *Faits de langues*, [En ligne], 2/4, 81-87, consulté le 10 septembre 2018, URL : https://www.persee.fr/doc/flang_1244-5460_1994_num_2_4_942.
- KLEIBER G., 1981c, « Relatives spécifiantes et relatives non spécifiantes », *Le Français moderne*, 49/3, 216-233.
- KLEIBER G. (dir.), 1983a, *Rencontre(s) avec la généricité*, Metz, Université de Metz.
- KLEIBER G., 1997, « Sens, référence et existence : que faire de l'extra-linguistique ? », *Langage*, [En ligne], 31/127, consulté le 14 avril 2018, 9-37, URL : https://www.persee.fr/doc/lgge_0458-726x_1997_num_31_127_2123.
- KRÁMSKÝ J., 1972, *The Article and the Concept of Definiteness in Language*, The Hague, Mouton.
- KRIFKA M., 1990, « Four Thousand Ships Passed Through the Lock », *Linguistics and Philosophy*, 13, 487-520.
- KRIPKE S., 1980, *Naming and Necessity*, Oxford. B. Blackwell.
- LAMBRECHT K., 2001, « Dislocation », in HASPELMATH M. (dir.), KOENIG E. (dir.), OESTERREICHER W. (dir.) et BAIBLE W. (dir.), *Language Typology and Language Universals. Volume 2*, Berlin/New York, Mouton de Gruyter, 1050-1078.
- LE PESANT D., 2002, « La détermination dans les anaphores fidèles et infidèles », *Langages*, [En ligne], 36/145, 33-59, consulté le 18 août 2018, URL : https://www.persee.fr/doc/lgge_0458-726x_2002_num_36_145_906.
- LOEBNER S., 2011, « Concept Types and Determination », *Journal of Semantics*, [En ligne], 28/3, 279-333, téléchargé le 18 septembre 2018, URL : <https://academic-oup-com.acces.bibliotheque-diderot.fr/jos/article/28/3/279/1660018>.
- LOEBNER S., 1998, « Definite Associative Anaphora », in BOTLEY S. P. (dir.), *Approaches to Discourse Anaphora: Proceedings of DAARC96 – Discourse Anaphora and Resolution Colloquium, Lancaster University, July 17th–18th 1996*, Lancaster, University Centre for Computer Corpus Research on Language.
- LOEBNER S., 1985, « Definites », *Journal of Semantics*, [En ligne], 4/4, 279-326, téléchargé le 28 novembre 2018, URL : <https://academic-oup-com.acces.bibliotheque-diderot.fr/jos/article/4/4/279/1612653>.
- LOUCKÁ H., 1990, *Francoúzké vytýkací konstrukce a aktuální členění*, Prague, Université Charles.
- LYONS C., 1999, *Definiteness*, Cambridge, Cambridge University Press.
- MARTIN R., 1983, *Pour une logique du sens*, Paris, Presse universitaire de France.
- MARUŠIČ F. et ŽAUCER R., « On the Adjectival Definite Article in Slovenian », *Pismo – Časopis za jezik i književnost*, [En ligne], 15, téléchargé le 3 juin 2018, URL : <http://www.ung.si/~rzaucer/papers/MarusicZaucer-2007-Adjectival-definite-article-in-Slo.pdf>.
- MC GINN C., 1981, « The Mechanisms of Reference », *Synthese*, 49/2, 157-186.
- MENDOZA I., 2004, *Nominaldetermination im Polnischen. Die primären Ausdrucksmittel*, thèse de doctorat non publiée, München, Ludwig-Maximilians-Universität.
- MILNER J.-C., 1978, *De la syntaxe à l'interprétation*, Paris, Éditions du Seuil.
- MILNER J.-C., 1976, « Réflexions sur la référence », *Langue française*, [En ligne], 30/1, consulté le 25 juin 2018, 63-73, URL : https://www.persee.fr/doc/lfr_0023-8368_1976_num_30_1_6112.
- NAVRÁTILIVÁ O. et ZIKÁNOVÁ Š., 2017, « Vývoj slovosledu nominální skupiny », in NEKULA M. (dir.) et PLESKALOVÁ (dir.), *Nový encyklopedický slovník češtiny*, [en ligne], consulté le 11 avril 2021, URL : <https://www.czechency.org/slovník/V%C3%9DVOJ%20SLOVOSLEDU%20NOMIN%C3%81LN%C3%8D%20SKUPINY>.

- NOALLY M., 1995, « Un nom propre, deux morphologies : pour quoi dire ? », in NOALLY M. (dir.), *Nom propre et nomination*, Paris, Klincksieck, 77-83.
- ORTMANN A., 2014, « Definite Article Asymmetries and Concept Types: Semantic and Pragmatic Uniqueness », in GAMERSCHLAG T. (dir.), GERLAND D. (dir.), RAINER O. (dir.) et PETERSEN W. (dir.), *Concept Types and Frames – Applications in Language, Cognition and Philosophy*, Dordrecht, Springer, 293-322.
- PREVOST S., 2003, « Détachement et topicalisation : des niveaux d'analyse différents », *Les cahiers de praxématique*, [En ligne], 40, 97-126, consulté le 1 juin 2020, URL : <https://journals.openedition.org/praxématique/2707>.
- PRICE G., 1971, *The French Language: Present and Past*, London, E. Arnold
- ROBERTS C., 2003, « Uniqueness in Definite Noun Phrases », *Linguistics and Philosophy*, [En ligne], 26/3, 287-350, téléchargé le 25 février 2020, URL : <https://link-springer-com.acces.bibliotheque-diderot.fr/article/10.1023/A:1024157132393>.
- RUNIĆ M., 2015, « The Definite Article in an Articleless Language », in MUELLER-REICHAU O. (dir.), YASTREBOVA M. (dir.), ZYBATOW G. (dir.), BISKUP P. (dir.), GUHL M. (dir.), et HURTIG C. (dir.), *Slavic Grammar from a Formal Perspective, the 10th Anniversary FDSL Conference, Leipzig 2013*, Frankfurt am Main, Peter Lang, 410-424.
- RUSSELL B., 1989, *Écrits de logique philosophique*, Paris, Presse universitaire de France.
- RUSSELL B., 1905, « On Denoting », *Mind*, [En ligne], 114/456, 479-493, téléchargé le 28 septembre 2018, URL : <https://academic-oup-com.acces.bibliotheque-diderot.fr/mind/article/114/456/873/1071184>.
- SCHLOSLE L., 2018, « How Useful is Case Morphology?: from Latin to French », in CARLIER A. (dir.) et GUILLOT-BARBANCE C. (dir.), *Latin tardif, français ancien: Continuités et ruptures*, Berlin, Walter de Gruyter, 127-170.
- SCHLOSLE L., 1984, *La déclinaison bicasuelle de l'ancien français, son rôle dans la syntaxe de la phrase, les causes de sa disparition*, Odense, Odense university press.
- SCHNEDECKER C., 2002, « Présentation : les adjectifs “inclassables”, des adjectifs du troisième type ? », *Langue française*, [En ligne], 136/1, 3-19, consulté le 16 mars 2020, URL : https://www.persee.fr/doc/lfr_0023-8368_2002_num_136_1_6468.
- SCHNEDECKER C., 1997, *Nom propre et chaînes de référence*, Metz, Université de Metz.
- SCHROEDER C., 2006, « Articles and Article Systems in Some Languages in Europe », in BERNINI G. (dir.) et SCHWARTZ M. L. (dir.), *Pragmatic Organization of Discourse in the Languages of Europe*, Berlin, Mouton de Gruyter, 545-601.
- SCHWARZ F., 2012, « How Weak and How Definite are Weak Definites? », *Revised Draft, December 21, 2012*, University of Pennsylvania, [En ligne], téléchargé le 4 mars 2021, URL : https://www.researchgate.net/publication/266281597_How_weak_and_how_definite_are_Weak_Definites.
- SCHWARZ F., 2009, *Two Types of Definites in Natural Language*, thèse de doctorat non publiée, Amherst, University of Massachusetts, [En ligne], téléchargé le 28 janvier 2020, http://florianschwarz.net/FSDiss/FS-Diss_singlespace.pdf.
- ŠIMÍK R., 2017, « Plovoucí kvantifikátor », in KARLÍK P. (dir.), NEKULA M. (dir.) et PLESKALOVÁ J. (dir.), *Nový encyklopedický slovník češtiny*, [en ligne], consulté le 5 mars 2021, URL : <https://www.czechency.org/slovník/PLOVOUC%C3%8D%20KVANTIFIK%C3%81TOR>.
- ŠIMÍK R. et DEMIAN C., 2020, « Definiteness, Uniqueness, and Maximality in Languages with and without Articles », *Journal of Semantics*, [En ligne], 37/2, téléchargé le 24 février 2020, URL : <https://ling.auf.net/lingbuzz/004730>.
- ŠIMÍK R. et SLÁMA J., 2020, « Relativní věty uvozené výrazem *jak* », exposé non publié, présenté à l'Association de linguistique (Jazykovědné sdružení), Prague, 8 octobre 2020, [En ligne], consulté le 29 mars 2021, URL : <https://www.radeksimik.eu/manuscripts.html>.

SPEVAK O., 2010, *Constituent Order in Classical Latin prose*, Amsterdam, J. Benjamins Publishing Company.

STANLEY J. et SZABÓ Z. G., 2000, « On Quantifier Domain Restriction », *Mind & Language*, [En ligne], 15/2-3, 219-261, consulté le 24 février 2020, URL : <https://onlinelibrary-wiley-com.acces.bibliotheque-diderot.fr/doi/epdf/10.1111/1468-0017.00130>.

STRAWSON P. F., 1950, « On referring », *Mind*, [En ligne], 10/235, 320-344, téléchargé le 30 septembre 2018, URL : <https://www.jstor.org/stable/pdf/2251176.pdf>.

STUDLER R., 2014, « The Morphology, Syntax and Semantics of Definite Determiners in Swiss German », in HOFHERR P. C. (dir.) et ZRIBI-HERTZ A. (dir.), 2014, *Crosslinguistic Studies on Noun Phrase Structure and Reference*, Leiden, Brill.

TELLIER C. et VALOIS D., 2006, « Quantification à distance et quantificateurs flottants », in TELLIER C. et VALOIS D., *Constructions méconnues du français*, Montréal, Les Presses de l'université de Montréal, 93-105.

VAN DEYCK R., 1998, « La détermination nominale en ancien français », in ENGLEBERT A. (dir.), PIERRARD M. (dir.), ROSIER L. (dir.) et VAN RAEMDONCK D. (dir.), *La ligne claire. De la linguistique à la grammaire. Mélanges offerts à Marc Wilmet à l'occasion de son 60^e anniversaire*, Louvain-la-Neuve, De Boeck Supérieur, 125-136.

VESELOVSKÁ L., 2014, « Universal DP-Analysis in Articleless Language: A Case Study in Czech », in VESELOVSKÁ L. et JANEBOVÁ M., *Nominal Structures: All in Complex DPs*, Olomouc, Olomouc Editors, 12-28, [En ligne], téléchargé le 27 février 2019, URL : https://www.researchgate.net/publication/264397487_Universal_DP-Analysis_in_Articleless_Language_A_Case_Study_in_Czech.

VON HEUSINGER K., 2002, « Specificity and Definiteness in Sentence and Discourse Structure », *Journal of Semantics*, [En ligne], 19/3, 245-274, téléchargé le 15 septembre 2018, URL : <https://academic-oup-com.acces.bibliotheque-diderot.fr/jos/article/19/3/245/1629749>.

WILMET M., 1986, *La détermination nominale : quantification et caractérisation*, Paris, Presses universitaires de France.

WITSCHKO M., 2012, « What Does it Take to Host a (Restrictive) Relative Clause? », *Working Papers of the Linguistics Circle of the University of Victoria*, [En ligne], 21/2, 100-145, téléchargé le 14 juin 2020, URL : <file:///C:/Users/Lenovo/Downloads/7815-Article%20Text-6418-1-10-20120305.pdf>.

YULE, G., 1981, « New, Current and Displaced Entity Reference », *Lingua*, [En ligne], 55/1, 41-52, téléchargé le 23 février, URL : <https://www.sciencedirect-com.acces.bibliotheque-diderot.fr/science/article/pii/0024384181900541>.

ZRIBI-HERTZ A., 2008, « La syntaxe de l'article défini : de la structure à l'interprétation », in VAN RAEMDONCK D. et PLOOG K., *Acte du colloque « Modèles syntaxiques. La syntaxe à l'aube du XXI^{ème} siècle »*, 63-84. Bruxelles/Berlin/Oxford/Vienne, Peter Lang, 1-15.

AUTRES TRAVAUX (GRAMMAIRES, DICTIONNAIRES ET AUTRES)

ADAM J.-M., 1997, « Genres, textes, discours : pour une reconception linguistique du concept de genre », *Revue belge de philologie et d'histoire*, [en ligne], 75/3, 1997, 665-681, téléchargé le 3 novembre 2020, URL : https://www.persee.fr/doc/AsPDF/rbph_0035-0818_1997_num_75_3_4188.pdf.

ADAM J.-M., 2005 [1992], *Les textes, types et prototypes : récit, description, argumentation, explication et dialogue*, Paris, Nathan.

ADAM J.-M., 1999, *Linguistique textuelle : des genres de discours aux textes*, Paris, Nathan.

BALLY C., 1926, *Le Langage et la Vie*, Paris, Payot.

BALLY C., 1932, *Linguistique générale et linguistique française*, Berne, Francke.

BALLY C., 1951 [1909], *Traité de stylistique française*, Heidelberg/Paris, C. Winter/Klincksieck.

- BARBÉRIS J.-M. et GARDES-MADRAY F., 1988, « Praxématique et analyse textuelle : les concepts de discours et d'énonciation historique », *Cahiers de praxématique*, 10, Montpellier, Pulm, 121-146.
- BAUER W., 1993, *Maori*, London, Routledge.
- BĚLIČ J., 1972, *Nástin české dialektologie*, Prague, Státní pedagogické nakladatelství.
- BĚLIČ J., HAVRÁNEK B., JEDLIČKA A., 1962, « Problematika obecné češtiny a jejího poměru k jazyku spisovnému », *Slovo a slovesnost*, [En ligne], 23/2, 108-126, consulté le 7 décembre 2017, URL : <http://sas.ujc.cas.cz/archiv.php?art=1162>.
- BERGER T., 2008, « Deutsche Einflüsse auf das grammatische System des Tschechischen », in *Studien der Historischen Grammatik des Tschechischen. Bohemistische Beiträge zur Kontaktlinguistik*, Brno, Travaux linguistiques de Brno.
- BIBER D., 2014, « Using Multi-Dimensional Analysis to Explore Cross-Linguistic Universals of Register Variation », *Languages in Contrast*, 14/1, 7-34.
- BIBER D. et CONRAD S., 2009, *Register, Genre, and Style*, Cambridge, Cambridge University Press.
- BLANCHE-BENVENISTE C., 1990, *Le français parlé. Études grammaticales*, Paris, CNRS Éditions.
- BOPP F., 1857-1861, *Vergleichendes Grammatik des Sanskrits, Send, Armenischen, Griechischen, Lateinischen, Litauischen, Altslavischen, Gothischen und Deutschen*, Berlin, F. Dümmler's Verlagsbuchhandlung.
- BRUNET E. (dir.), 1981, *Le vocabulaire français de 1789 à nos jours, d'après les données du Trésor de la langue française*, Paris/Genève, Champion/Slatkine.
- ČERMÁK F. (dir.), 2007, *Frekvenční slovník mluvené češtiny*, Prague, Karolinum.
- ČERMÁK F., 2014, « Obecná čeština: je součástí české diglosie? », in ČERMÁK F., *Jazyk a slovník. Výbrané lingvistické studie*, Prague, Karolinum.
- ČERMÁK F. et KŘEN M., 2011, *A Frequency Dictionary of Czech. Core Vocabulary for Learners, 1st Edition*, Abingdon-on-Thames, Routledge.
- ČERMÁK F. (dir.) et KŘEN M. (dir.), 2004, *Frekvenční slovník češtiny*, Prague, Nakladatelství Lidové noviny.
- CHAFE W., 1994, *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*, Chicago, University of Chicago Press.
- CHLOUPEK J., 1986, *Dichotomie spisovnosti a nespisovnosti*, Brno, Universitě J.E. Purkyně.
- CORBETT G. (dir.) et COMRIE B. (dir.), 1993, *The Slavonic Languages (Routledge Language Family Series), 1st Edition*, London/New York, Routledge.
- CVRČEK V. (dir.), 2015 [2010], *Mluvnice současné češtiny*, 1, *Jak se píše a jak se mluví*, Prague, Karolinum.
- CVRČEK V. (dir.), 2020, *Registry v češtině*, Prague, Academia.
- DANEŠ F., 1974, *Papers on Functional Sentence Perspective*, Hague/Paris, Mouton.
- DANEŠ F. (dir.), FILIPEC J. (dir.), MACHAŘ J. (dir.) et MEJSTRÍK V. (dir.), 2009 [1994], *Slovník spisovné češtiny pro školu a veřejnost*, Prague, Academia.
- DANEŠ F. (dir.), HLAVSA Z. (dir.) et GREPL M. (dir.), 1987, *Mluvnice češtiny*, 3, *Skladba*, Prague, Academia.

- DESCLÉS J.-P. et GUENTCHÉVA Z., 2000, « Énonciateur, locuteur, médiateur », in MONOD-BECQUELIN A. (dir.) et ERIKSON P. (dir.), *Les rituels du dialogue*, Nanterre, Société d'ethnologie, 79-112.
- DESCLÉS J.-P. et GUENTCHÉVA Z., 1997, « Énonciateur, locuteur, médiateur dans l'activité dialogique », in *Colloque international des Américanistes, Quito, Equateur, juillet 1997*, [En ligne], téléchargé le 18 avril 2017, URL : <http://lalic.paris-sorbonne.fr/PUBLICATIONS/1997-1998/Descles/Quito.pdf>.
- DOBROVSKÝ J., 1809, *Ausführliches Lehrgebäude der Böhmisches Sprache, zur gründlichen Erlernung derselben für Deutsche, zur vollkommenern Kenntniss für Böhmen*, Prague, Herrl, [en ligne], consulté le 19 avril 2021, URL : <https://play.google.com/store/books/details?id=V5YBAAAAMAAJ&rdid=book-V5YBAAAAMAAJ&rdot=1>.
- DOBROVSKÝ J., 1802 et 1821, *Deutsch-böhmisches Wörterbuch*, Prague, Herrl, [en ligne], consulté le 19 avril 2021, URL : <https://bara.ujc.cas.cz/slovniky/dobrovsky/dbsearch.php?databaségbsgb653sd>.
- DUBOIS J. (dir.), 1972, *Dictionnaire de linguistique*, Paris, Larousse.
- FRÉDÉGAIRE et WALLACE-HADRILL (dir.), 1960, *Fredegarii Chronicurum liber quartus cum continuationibus / The Fourth Book of the Chronicle of Fredegar, translated from the Latin with Introduction and Notes by J.M. Wallace-Hadrill*, London, Thomas Nelson & Sons.
- GARVIN P. L., 1983, « Le rôle des linguistes de l'École de Prague dans le développement de la norme linguistique tchèque », in BÉDARD E. (dir.) et MAURAI J. (dir.), *La norme linguistique, textes colligés et présentés par Édith Bédard et Jacques Maurais*, Québec/Paris, Gouvernement du Québec, Conseil de la langue française/Le Robert, 141-152.
- GEBAUER J., 1960 [1896], *Historická mluvnice jazyka českého*, III, *Tvarosloví*, I. *Skloňování*, Prague, l'Académie tchéco-slovaque des sciences, [en ligne], consulté le 5 février 2020, URL : <https://vokabular.ujc.cas.cz/moduly/literatura/edice/00492e83-c6eb-4dc1-9995-e0c130eb8100/plny-text/s-apatem/folio/451>.
- GEBAUER J. et TRÁVNÍČEK F. (dir.), 1929, *Historická mluvnice jazyka českého*, IV, *Skladba*, Prague, l'Académie tchèque des sciences et des arts, [en ligne], consulté le 4 janvier 2020, URL : <https://vokabular.ujc.cas.cz/moduly/literatura/edice/f66effc8-4a20-4157-b6f2-4fec090c2e7c/plny-text/s-apatem/folio/167>.
- HAUSENBLAS K., 1955, « K základním pojmům jazykové stylistiky », *Slovo a slovesnost*, [en ligne], 16/1, 1-15, consulté le 24 décembre 2019, URL : <http://sas.ujc.cas.cz/archiv.php?art=776>.
- HAUSENBLAS K., 1991, *Výstavba jazykových projevů a styl*, Prague, Acta Universitatis Carolinae.
- HAVRÁNEK B. (dir.), 1960-1971, *Slovník spisovného jazyka českého*, Prague, Academia, [en ligne], consulté le 15 janvier 2020, URL : <https://ssjc.ujc.cas.cz/search.php?db=ssjc>.
- HAVRÁNEK B. et JEDLIČKA A., 1960, *Česká mluvnice*, Prague, Státní pedagogické nakladatelství.
- HEYSE K. W. L., 1956, *System der Sprachwissenschaft*, Berlin, F. Dümmler, [en ligne], consulté le 14 juin 2018, URL : <https://archive.org/details/systemdersprachw00heys>.
- HORÁLEK K., 1962, *Úvod do studia slovanských jazyků*, Prague, Maison d'édition de l'Académie tchéco-slovaque des sciences.
- HUJER O. (dir.), SMETÁNKA E. (dir.) WEINGART M. (dir.) et HAVRÁNEK B. (dir.), 1935-1957, *Příruční slovník jazyka českého*, [en ligne], consulté 21 janvier 2020, URL : <https://psjc.ujc.cas.cz/>.
- HOFFMANNOVÁ J. (dir.) et HOMOLÁČ J. (dir.), 2016, *Stylistika mluvené a psané češtiny*, Prague, Academia.
- HOFFMANNOVÁ J. (dir.), HOMOLÁČ J. (dir.) et MRÁZKOVÁ K. (dir.), 2019, *Syntax mluvené češtiny*, Prague, Academia.

- IMBS P. (dir.) et MARTIN R. (dir.), 1971, *Dictionnaire des fréquences : vocabulaire littéraire des XIXe et XXe siècles*, I, II, III et IV, Paris, Klincksieck.
- IMBS P. (dir.) et QUEMADA B. (dir.), 1971-1994, *Trésor de la langue française : dictionnaire de la langue du XIXe et du XXe siècle (1789-1960)*, Paris, Éditions du Centre national de la recherche scientifique/Gallimard.
- JAKOBSON R., 1963b, « Linguistique et poétique », in *Essais de linguistique générale*, Paris, Éditions de Minuit, 209-248.
- JELÍNEK J., BEČKA J. V., TĚŠITELOVÁ M., 1961, *Frekvence slov, slovních druhů a tvarů v českém jazyce*, Prague, Státní pedagogické nakladatelství.
- JUNGSMANN J., 1989-1990 [1835-1839], *Slovník česko-německý, díly 1.-5*, Prague, Academia.
- KARLÍK P. (dir.), NEKULA M. (dir.) et PLESKALOVÁ J. (dir.), 2016, *Nový encyklopedický slovník češtiny*, I et II., Brno, Universitě Masaryk.
- KERBRAT-ORECCHIONI C., 1978, *La connotation*, Lyon, Presses universitaires de Lyon.
- KOŘENSKÝ J. (dir.) et KOMÁREK M. (dir.), 1986, *Mluvnice češtiny, 2, Tvarosloví*, Prague, Academia.
- KOSEK P., 2017, « Periodizace vývoje češtiny », in KARLÍK P. (dir.), NEKULA M. (dir.) et PLESKALOVÁ J. (dir.), *Nový encyklopedický slovník češtiny*, [en ligne], consulté le 19 avril 2021, URL : <https://www.czechency.org/slovník/PERIODIZACE%20V%C3%9DVOJE%20%C4%8CE%C5%A0TINY>.
- KRČMOVÁ M., 2017, « Obecná čeština », in KARLÍK P. (dir.), NEKULA M. (dir.) et PLESKALOVÁ J. (dir.), *Nový encyklopedický slovník češtiny*, [en ligne], consulté le 18 mai 2017, URL : <https://www.czechency.org/slovník/OBECN%C3%81%20%C4%8CE%C5%A0TINA>.
- KRČMOVÁ M. et CHLOUPEK J., 2017, « Národní jazyk », in KARLÍK P. (dir.), NEKULA M. (dir.) et PLESKALOVÁ J. (dir.), *Nový encyklopedický slovník češtiny*, [en ligne], consulté le 20 avril 2021, URL : <https://www.czechency.org/slovník/N%C3%81RODN%C3%8D%20JAZYK#%E2%80%9Eobecn%C3%A1%20morav%C5%A1tina%E2%80%9C>.
- KRZESZOWSKI T. P., 1990, *Contrasting Languages: the Scope of Contrastive Linguistics*, Berlin/New York, Mouton de Gruyter.
- LAMPRECHT A., BAUER J. et ŠLOSAR D., 1986, *Historická mluvnice češtiny*, Prague, Státní pedagogické nakladatelství.
- LE GOFFIC P., 1993, *Grammaire de la phrase française*, Paris, Hachette.
- LEECH G., RAYSON P. et WILSON A., 2001, *Word Frequencies in Written and Spoken English: based on the British National Corpus*, London, Longman.
- MAGLIONE C., 2001, « Nové uplatnění obecné češtiny v literatuře », *Naše řeč*, [En ligne], 84/2, 74-80, consulté le 17 mai 2017, URL : <http://nase-rec.ujc.cas.cz/archiv.php?art=7612>.
- MATHESIUŠ V., 1947, *Čestina a obecný jazykozpyt*, Prague, Melantrich.
- MEILLET A., 1982, *Linguistique historique et linguistique générale*, Paris/Genève, H. Champion-Slatkine.
- MOIGNET G., 1973, *Grammaire de l'ancien français*, Paris, Klincksieck.
- NEBESKÁ I., 2017, « Spisovná čeština, hovorová čeština », in KARLÍK P. (dir.), NEKULA M. (dir.) et PLESKALOVÁ J. (dir.), *Nový encyklopedický slovník češtiny*, [en ligne], consulté le 18 mai 2017, URL : <https://www.czechency.org/slovník/SPISOVN%C3%81%20%C4%8CE%C5%A0TINA>.

- PAVEAU M.-A. et SARFATI G.-E., 2003, *Les grandes théories de la linguistique. De la grammaire comparée à la pragmatique*, Paris, Armand Colin.
- PERRIN L., 2004, « La notion de polyphonie en linguistique et dans le champ des sciences du langage », *Questions de communication* [En ligne], 6, 265-282, consulté le 31 mars 2017, URL : <https://questionsdecommunication.revues.org/4445>.
- PETIT G., 2001, « Dénomination et lexique », *Journal of French Language Studies*, [En ligne], 11/1, 89-121, téléchargé le 20 avril 2018, URL : https://halshs.archives-ouvertes.fr/file/index/docid/464109/filename/Denomination_et_lexique.pdf.
- PETR J., 1984, *Základy slavistiky*, Brno, Státní pedagogické nakladatelství.
- QUIRK R. (dir.), 1972, *A Grammar of Contemporary English*, London, Longman.
- RIEGEL M. (dir.), PELLAT J.-C. (dir.) et RIOUL R. (dir.), 2016 [1994], *Grammaire méthodique du français*, 6^{ème} édition, Paris, Presses universitaires de France.
- ROSCH E., 1977, « Human categorization », in WARREN N. (dir.), *Advances in Cross-cultural Psychology*, I, London, Academic Press.
- ŠABRŠULA J., 1964, « Ke srovnávacím studiím slovosledným (Některé otázky větné perspektivy ve francouzštině ve srovnání s češtinou) », *Slovo a slovesnost*, [En ligne], 25/3, 161-166, consulté le 12 novembre 2017, URL : <http://sas.ujc.cas.cz/archiv.php?art=1298>.
- ŠABRŠULA J., 1990, *Problèmes de la stylistique comparée du français et du tchèque*, Prague, Université Charles.
- SAUSSURE F. de, BALLY C. (dir.) et SECHEHAYE A. (dir.), 1960 [1916], *Cours de linguistique générale*, Paris, Payot.
- SGALL P., 1973, « Kontextové zapojení a otázková metoda », *Slovo a slovesnost*, [En ligne], 34/3, 202-211, consulté le 8 décembre 2017, URL : <http://sas.ujc.cas.cz/archiv.php?art=3965>.
- SGALL P. (dir.), 1992, *Variation in Language. Code Switching in Czech as a Challenge for Sociolinguistics*, Amsterdam/Philadelphia, J. B. Publishing Company.
- SGALL P. et TRNKOVÁ A., 1963, « K metodám zkoumání běžně mluvené češtiny », *Naše řeč*, [En ligne], 46/1, 28-35, consulté le 6 décembre 2017, URL : <http://nase-rec.ujc.cas.cz/archiv.php?art=4927>.
- SHORT D., 1993, « Czech », in CORBETT G. (dir.) et COMRIE B. (dir.), *The Slavonic Languages (Routledge Language Family Series), 1st Edition*, London/New York, Routledge, 455-532.
- ŠLOSAR D., 2017, « Obrozenecká čeština », in KARLÍK P. (dir.), NEKULA M. (dir.) et PLESKALOVÁ J. (dir.), *Nový encyklopedický slovník češtiny*, [en ligne], consulté le 19 avril 2021, URL : <https://www.czechency.org/slovník/OBROZENSK%C3%81%20%C4%8CE%C5%A0TINA>.
- ŠLOSAR D. et KOSEK P., 2017, « Humanistická čeština », in KARLÍK P. (dir.), NEKULA M. (dir.) et PLESKALOVÁ J. (dir.), *Nový encyklopedický slovník češtiny*, [en ligne], consulté le 19 avril 2021, URL : <https://www.czechency.org/slovník/HUMANISTICK%C3%81%20%C4%8CE%C5%A0TINA>.
- ŠONKOVÁ J., 2008, *Morfologie mluvené češtiny: frekvenční analýza*, Prague, Nakladatelství Lidové noviny.
- ŠTÍCHA F. (dir.), 2013, *Akademická gramatika spisovné češtiny*, Prague, Academia.
- ŠTÍCHA F., 1987, « Komunikativní funkce umístění východiska (tématu) na konci věty », *Naše řeč*, [En ligne], 70/2, 69-75, consulté le 8 décembre 2017, URL : <http://nase-rec.ujc.cas.cz/archiv.php?art=6683>.
- TĚŠITELOVÁ M. (dir.), 1985, *Kvantitativní charakteristiky současné češtiny*, Prague, Academia.

TOWNSEND C. E., 1990, *A Description of Spoken Prague Czech*, Ohio, Columbus.

TRÁVNÍČEK F., 1951, *Mluvnice spisovné češtiny*, II, Skladba, Prague, Slovanské nakladatelství.

UHLÍŘOVÁ L., 1980, « K emocionalitě v rovině aktuálního členění », in *Otázky slovanské syntaxe*, 4/2, Brno, Opera Universitatis Purkynianae Brunensis, 167-171.

VAILLANT A., 1958, *Grammaire comparée des langues slaves. Tome II. Deuxième partie. Morphologie/Flexion pronominale*, Lyon, Paris, IAC.

VEY M., 1946, *Morphologie du tchèque parlé*, Paris, C. Klincksieck.

VINAY J.-P. et DARBELNET J., 1960, *Stylistique comparée du français et de l'anglais : méthode de traduction*, Paris, Didier.

VOLKOVÁ B., 1992, « K emotivní sémantice a sémiotice », *Slovo a slovesnost*, [En ligne], 53/1, 11-21, consulté le 8 décembre 2017, URL : <http://sas.ujc.cas.cz/archiv.php?art=3443>.

WILMET M., 2010 [1997], *Grammaire critique du français*, 5^{ème} réédition, Bruxelles, De Boeck-Duculot.

WITTGENSTEIN L., 1986, *Tractatus logico-philosophicus*, suivi de *Investigations philosophiques*, Paris, Gallimard.

ZIMA J., 1961, *Expresivita slova v současné češtině*, Prague, Maison d'édition de l'Académie tchéco-slovaque des sciences.

OUTILS DE LINGUISTIQUE DE CORPUS, MÉTHODOLOGIE, SOURCES D'EXEMPLES, CARTES

ALBERT S., 2009, « L'informatisation du *Trésor de la langue française* : perspectives offertes pour l'étude des emprunts », *Études de linguistique appliquée*, [En ligne], 156/4, 491-498, consulté le 31 octobre 2020, URL : <https://www.cairn.info/revue-ela-2009-4-page-491.htm#>.

AVANZI M., 2005, « Autour du Corpus de référence du français parlé », *Lidil*, [En ligne], 32, consulté le 31 mars 2021, URL : <https://journals.openedition.org/lidil/120>.

AVANZI M., BÉGUELIN M.-J. et DIÉMOZ F., 2016, « De l'archive de parole au corpus de référence. Le corpus oral de français de Suisse romande (OFRON) », *Corpus*, [En ligne], 15, 309-342, consulté le 30 octobre 2020, URL : <https://journals.openedition.org/corpus/3060>.

BENZITOUN C., DEBAISIEUX J.-M., DEULOFEU J., 2016, « Le projet ORFÉO : un corpus d'étude pour le français contemporain », *Corpus*, [En ligne], 15, 91-114, consulté le 30 octobre 2020, URL : <https://journals.openedition.org/corpus/2936>.

BLANCHE-BENVENISTE C., 2007, « Corpus de langue parlée et description grammaticale de la langue », *Langage et société*, [En ligne], 121-122/3-4, 129-141, consulté le 15 mai 2018, URL : <https://www.cairn.info/revue-langage-et-societe-2007-3-page-129.htm>.

CARRUTHERS J., 2008, « Annotating an Oral Corpus Using the Text Encoding Initiative. Methodology, Problems, Solutions », *Journal of French Language Studies*, 18/1, 103-119.

ČERMÁK P. (dir.) et NÁDVORNÍKOVÁ O. (dir.), 2015, *Románské jazyky a čeština ve světle paralelních korpusů*, Prague, Karolinum.

ČERMÁK F. et ROSEN A., 2012, « The Case of InterCorp, a Multilingual Parallel Corpus », *International Journal of Corpus Linguistics*, [En ligne], 13/3, 411-427, téléchargé le 14 octobre 2018, URL : https://www.researchgate.net/publication/234118417_The_Case_of_InterCorp_a_multilingual_parallel_corpus.

CHLUMSKÁ L., 2017, *Překládová čeština a její charakteristiky*, Prague, Nakladatelství lidové noviny.

- CRESTI E., BACELAR DO NASCIMENTO F., MORENA S. A., VERONIS J., MARTIN P., KALID C., 2004, « The C-ORAL-ROM CORPUS: A Multilingual Resource of Spontaneous Speech for Romance Languages », in *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*, Lisbon, European Language Resources Association, 575-578, [En ligne], téléchargé le 20 avril 2018, URL : <http://www.lrec-conf.org/proceedings/lrec2004/pdf/357.pdf>.
- D'ARMAGNAC M., 2021, « Le retour par effraction de Matteo Salvini », *Valeurs actuelles*, 4395, 36-39.
- DEBAISIEUX J.-M., 2003, « Les corpus oraux : situation, exploitation linguistique, bilan et perspectives », *Scolia, Université des sciences humaines Strasbourg*, [en ligne], 9-40, téléchargé le 21 avril 2018, URL : <https://halshs.archives-ouvertes.fr/halshs-00149141/document>.
- DOSTIE G., 2016, « Le Corpus de français parlé au Québec (CFPQ) et la langue des conversations familiales : Exemple de mise à profit des données à partir d'un examen lexico-sémantique de la séquence *je sais pas* », *Corpus*, [en ligne], 15, consulté le 15 février 2020, URL : <https://journals.openedition.org/corpus/2945>.
- EISELE A. et CHEN Y., 2010, *MultiUN: A Multilingual Corpus from United Nation Documents*, [En ligne], consulté le 6 mars 2020, URL : <http://opus.nlpl.eu/MultiUN.php>.
- FALAISE A., TUTIN A. et KRAIF O., 2011, « Une interface pour l'exploitation de corpus arborés par des non informaticiens : la plate-forme ScienQuest du projet Scientext », *Revue TAL*, 52/3, 103-128.
- GARRIGUES M., 1992, « Dictionnaires hiérarchiques du français. Principes et méthode d'extraction », *Langue française*, 96/1, 88-100.
- Groupe ICOR, (BALDAUF-QUILLIATRE H., COLON DE CARVAJAL I., ETIENNE C., JOUIN-CHARDON E., TESTON-BONNARD S. et TRAVERSO V.), 2016, « CLAPI, une base de données multimodale pour la parole en interaction : apports et dilemmes », in *Corpus*, [En ligne], 15, consulté le 30 octobre 2020, URL : <https://journals.openedition.org/corpus/2991#bibliography>.
- HEIDEN S., MAGUE J.-P. et PINCEMIN B., 2010, « TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement », in BOLASCO I. C. S. (dir.), *Proc. of 10th International Conference on the Statistical Analysis of Textual Data, JADT, 2010, 2, 1021-1032*, Rome, Edizioni Universitarie di Lettere Economia Diritto, [En ligne], téléchargé le 24 février 2019, URL : <https://halshs.archives-ouvertes.fr/halshs-00549779>.
- KOPŘIVOVÁ M., KOMRSKOVÁ Z., LUKEŠ D., POUKAROVÁ P. et ŠKARPOVÁ M., *ORTOFON : Corpus de tchèque parlé informel avec un système de transcription à plusieurs niveaux*, Institut du Corpus national tchèque, Faculté des Lettres de l'Université Charles, Prague, [En ligne], consulté le 7 avril 2021, URL : <https://wiki.korpus.cz/doku.php/en:cnk:ortofon>.
- KOPŘIVOVÁ M., KOMRSKOVÁ Z., LUKEŠ D., POUKAROVÁ P., WACLAWIČKOVÁ M., BENEŠOVÁ L. et KŘEN M., *ORAL: corpus de tchèque parlé informel, version 1 du 2. juin 2017*, Institut du Corpus national tchèque, Faculté des Lettres de l'Université Charles, Prague, [En ligne], consulté le 7 novembre 2020, URL : <https://wiki.korpus.cz/doku.php/en:cnk:oral>.
- KŘEN M., CVRČEK V., ČAPKA T., ČERMÁKOVÁ A., HNÁTKOVÁ M., CHLUMSKÁ L., JELÍNEK T., KOVÁŘÍKOVÁ D., PETKEVIČ V., PROCHÁZKA P., SKOUMALOVÁ H., ŠKRABAL M., TRUNEČEK P., VONDŘIŠKA P. et ZASTINA A., 2015, *SYN2015: corpus représentatif du tchèque écrit en synchronie*, Institut du Corpus national tchèque de la Faculté des Lettres de l'Université Charles, Prague, [En ligne], consulté le 7 novembre 2020, URL : <https://wiki.korpus.cz/doku.php/en:cnk:syn2015>.
- KUČERA K., ŘEHOŘKOVÁ A. et STLUKA M., *DIAKORP: Diachronic corpus of Czech, version 6 from 18 Feb 2015*, Institut du Corpus national tchèque de la Faculté des Lettres de l'Université Charles, Prague, [En ligne], consulté le 7 novembre 2020, URL : <https://wiki.korpus.cz/doku.php/en:cnk:diakorp>.
- LAGARDE A. et MICHARD L., 1964, *XVII^e siècle : les grands auteurs du programme anthologie et histoire littéraire. Collection littéraire Lagarde et Michard*, Paris, Bordas.
- McENERY T. et HARDIE A., 2012, *Corpus Linguistics*, Cambridge, Cambridge University Press.

- NÁDVORNÍKOVÁ O., 2017, « Pièges méthodologiques des corpus parallèles et comment les éviter », *Corela*, [En ligne], 15/1, consulté le 16 mars 2018, URL : <http://journals.openedition.org/corela/4810>.
- NÁDVORNÍKOVÁ O., 2020, « The Use of English, Czech and French Punctuation Marks in Reference, Parallel and Comparable Web Corpora: A Question of Methodology », *Linguistica Pragensia*, 30/2, 30-50.
- NÁDVORNÍKOVÁ O. et VAVŘÍN M., *Corpus InterCorp – langue française, version 10 du 1^{er} décembre 2017*, Institut du Corpus national tchèque de la Faculté des Lettres de l'Université Charles, Prague, [En ligne], consulté le 7 novembre 2020, URL : <https://wiki.korpus.cz/doku.php/en:cnk:intercorp:verze10>.
- NASR A., BÉCHET F. et REY J.-F., 2009, « MACAON. Une chaîne linguistique pour le traitement de graphes de mots », [En ligne], téléchargé le 7 novembre 2020, URL : <https://www.lattice.cnrs.fr/sites/itellier/articles/TALN2011.pdf>.
- ROSEN A., VAVŘÍN M. et ZASINA A. J., 2017, *Corpus InterCorp – langue tchèque, version 10 du 1^{er} décembre 2017*, Institut du Corpus national tchèque de la Faculté des Lettres de l'Université Charles, Prague, [En ligne], consulté le 7 novembre 2020, URL : <https://wiki.korpus.cz/doku.php/en:cnk:intercorp:verze10>.
- RUPNIK J., « HAVEL VÁCLAV (1936-2011) », *Encyclopaedia Universalis*, [En ligne], consulté le 24 février 2021, URL : <https://www.universalis.fr/encyclopedie/havel-v/>.
- STRAKOVÁ (dir.), 2014, « Open-Source Tools for Morphology, Lemmatization, POS Tagging and Named Entity Recognition », in *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, Baltimore, Maryland, June 2014*, Association for Computational Linguistics, 13-18, [En ligne], téléchargé le 15 février 2020, URL : <https://www.aclweb.org/anthology/P14-5003/>.
- TUTIN A. (dir.) et GROSSMANN F. (dir.), 2014, *L'écrit scientifique : du lexique au discours. Autour de Scientext*, Rennes, Presses de l'Université de Rennes.
- WISSELMANN B., 2020, « Je ne veux pas qu'on vienne défigurer ou pire, nier la culture française », *entretien avec F. Ferrand, Valeurs actuelles*, [En ligne], consulté le 13 février 2021, URL : <https://www.valeursactuelles.com/societe/franck-ferrand-je-ne-veux-pas-quon-vienne-defigurer-ou-pire-nier-la-culture-francaise-125447>.
- Analyse et traitement informatique de la langue française - UMR 7118 (ATILF), Cognition, Langue, Langages, Ergonomie - UMR 5263 (CLLE) (2020). Corpus journalistique issu de l'*Est Républicain* [Corpus]. ORTOLANG (Open Resources and TOols for LANGuage), [en ligne], consulté le 18 novembre 2020, URL : https://www.ortolang.fr/market/corpora/est_republicain/v4.
- ATILF. Base textuelle Frantext, 1998-2020, ATILF-CNRS & Université de Lorraine, [En ligne], consulté le 7 novembre 2020, URL : <https://www.frantext.fr>.
- Corpus Calculator, interface du Corpus national tchèque, [En ligne], consulté le 12 novembre 2020, URL : <https://www.korpus.cz/calc/>.
- Corpus d'Étude pour le Français Contemporain (CEFC), Orfeo (Outils et recherches pour le français écrit et oral), [En ligne], consulté le 7 novembre 2020, URL : <https://www.projet-orfeo.fr>.
- Corpus national tchèque (CNT), [En ligne], consulté le 17/03/2021, URL : <https://www.korpus.cz/>.
- Manuel de TXM 0.7, [En ligne], consulté le 13 octobre 2020, URL : <http://txm.sourceforge.net/doc/manual/0.7.7/fr/manual1.xhtml>.
- « Nářečí češtiny », *Wikipédia*, [En ligne], consulté le 13 octobre 2020, URL : https://cs.wikipedia.org/wiki/N%C3%A1%C5%99e%C4%8D%C3%AD_%C4%8De%C5%A1tiny#/media/Soubor:Czech_dialects.PNG.
- Page personnelle de Denis Apothéloz, ATILF, « Corpus divers, Corpus de français parlé et écrit », [En ligne], consulté le 12 février 2020, URL : <https://perso.atilf.fr/apotheloz/corpus/>.

« République tchèque », [En ligne], consulté le 13 octobre 2020, URL : <https://www.axl.cefan.ulaval.ca/europe/reptcheque-1General-demo.htm>.

ScienQuest, [En ligne], consulté le 28 novembre 2020, URL : <https://corpora.aiakide.net/scientext20/?do=SQ.setView&view=corpora>.

« Silésie tchèque », *Wikipédia*, [En ligne], consulté le 13 octobre 2020, URL : https://fr.wikipedia.org/wiki/Sil%C3%A9sie_tch%C3%A8que#/media/Fichier:CZ_Cechy_Morava_kraje.gif.

Site du logiciel d'annotation Morče, [En ligne], consulté le 10 février 2020, URL : <http://ufal.mff.cuni.cz/morce/index.php>.

Site du logiciel d'annotation MorphoDiTa, [En ligne], consulté le 10 février 2020, URL : <http://ufal.mff.cuni.cz/morphodita>.

Site du logiciel d'annotation TreeTagger, [En ligne], consulté le 10 février 2020, URL : <https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>.

Site Wikipédia du Corpus national tchèque, [En ligne], consulté le 13 mars 2018, URL : <http://wiki.korpus.cz/doku.php>.

The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses, 2015, [En ligne], téléchargé le 12 septembre 2019, URL : <https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf>.

TLFi : Trésor de la langue Française informatisé, ATILF – CNRS & Université de Lorraine, [En ligne], consulté le 7 novembre 2020, URL : <http://www.atilf.fr/tlfi>.

Index des auteurs

- Abney, 7, 8, 20, 114, 163, 165, 166, 168, 222, 515
Adam, 84, 231, 547, 548, 566, 567
Adamec, 7, 8, 16, 21, 22, 23, 162, 180, 198, 205, 207, 211, 212, 213, 214, 262, 264, 268, 283, 285, 349, 390, 420, 422, 433, 470, 497, 499, 508, 509, 512, 513, 547, 561, 563, 565, 582, 584, 585, 651
Andersen, 394
Anderson, 18, 40, 68, 69, 70, 73, 148, 225, 580, 587
Ariel, 50, 51, 52, 53, 252, 261, 310, 313, 442, 591
Austin, 517, 524
Bar-Hilley, 53
Bauer, 174, 177, 399, 420
Beauzée, 42
Bělič, 145, 174, 176, 178
Benveniste, 15, 42, 61, 124, 399, 518
Berger, 5, 7, 8, 17, 21, 45, 65, 85, 141, 142, 145, 150, 153, 162, 170, 172, 173, 174, 178, 180, 184, 194, 195, 198, 206, 214, 390, 400, 404, 420, 421, 422, 423, 426, 433, 469, 505, 574, 599
Biber, 84, 85
Bonnard, 17, 78, 204, 209, 210, 220, 221, 222, 224, 225, 263, 509, 526, 561, 585, 636
Bošković, 557
Brodie, 557
Brown, 443
Brugmann, 42
Bühler, 18, 41, 43, 44, 46, 203
Burks, 53
C. Lyons, 18, 65, 390
Caboli, 18, 79
Carlier, 1, 5, 18, 25, 47, 72, 73, 77, 79, 167, 169, 175, 208, 209, 211, 222, 233, 234, 235, 239, 240, 241, 243, 244, 245, 247, 248, 261, 270, 390, 395, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 471, 503, 520, 594, 597, 604, 637
Čermák, 144, 155, 156, 504, 597
Chafe, 261, 468
Charolles, 17, 40, 62, 66, 67, 129, 204, 205, 206, 262, 263, 510
Chervel, 61
Chloupek, 600
Christophersen, 18, 52, 64, 77, 244, 404, 442
Clark, 68, 480
Combettes, 167, 169, 222, 248, 270, 409, 410, 503, 597
Corbett et Comrie, 174
Corblin, 17, 40, 56, 61, 62, 67, 78, 117, 119, 120, 184, 204, 205, 244, 251, 252, 254, 255, 263, 286, 405, 442, 443, 448, 510, 554, 555, 571
Cornish, 48, 49, 50, 51
Cvrček, 15, 16, 83, 84, 85, 212, 289, 426, 504, 533, 597
Czardybon, 5, 17, 18, 27, 73, 79, 120, 176, 177, 184, 188, 189, 193, 283, 390, 400, 404, 406, 408, 415, 416, 417, 422, 423, 425, 426, 471, 478, 479, 481, 482, 484, 488, 490, 495, 497, 499, 502, 504, 515, 580, 582, 596, 615, 652
Damourette et Pichon, 43
Danon-Boileau, 205, 263, 510
De Mulder, 17, 18, 25, 47, 73, 77, 79, 175, 209, 211, 233, 234, 239, 240, 241, 243, 244, 245, 247, 261, 390, 395, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 471, 520, 594
Dees, 237
Diessel, 18, 36, 40, 51, 63, 68, 69, 73, 74, 75, 76, 77, 79, 80, 115, 166, 168, 169, 172, 173, 177, 191, 192, 215, 217, 218, 219, 220, 221, 222, 259, 260, 390, 393, 400, 401, 404, 408, 442, 475, 571, 590
Dixon, 18, 40, 56, 68, 69, 76, 77, 140, 236
Dobrovský, 178, 598
Donnellan, 57, 60, 61, 242, 411, 472, 475, 522, 571, 615
Dowty, 557
Dryer, 399, 420
Ducrot, 46, 59, 66, 80, 216, 507, 516, 518, 519, 520, 521, 522, 561
Dvořák, 6, 21, 416, 471, 490, 566
Ebert, 244, 410, 419
Epstein, 18, 64, 79, 403, 432, 440, 441, 502, 524, 582, 600
Erbenová, 17
Fillmore, 18, 46, 205, 207, 433, 509
Fodor, 60, 242, 262, 485
Foulet, 226
Fraser, 21, 44, 45, 46, 47, 50, 55, 117, 118, 245, 249
Frege, 517, 523
Fruyt, 79, 233

Fugier, 248
 Galmiche, 60
 Gary-Prieur, 17, 23, 40, 57, 63, 64, 65, 66, 67, 203, 205, 209, 263, 264, 266, 405, 510, 511, 516, 522
 Gebauer, 152, 162, 174, 175, 177, 178, 234
 Gerland, 79, 211, 461, 538
 Givón, 50, 52, 126, 390, 395, 404, 442
 Gougenheim, 226
 Goyens, 167, 217, 222, 233, 245, 246, 247, 248, 333, 576
 grammaticalisation, 292
 Greenberg, 18, 73, 77, 240, 241, 390, 401, 402, 404, 408
 Grice, 28, 216, 397, 507, 516, 517, 518, 519, 520, 521, 523, 524, 561, 583
 Guénette, 17, 72, 233, 237
 Guillot, 1, 4, 15, 17, 18, 49, 72, 117, 213, 233, 234, 235, 239, 563, 637
 Hagège, 390, 393, 394, 593
 Hakulinen, 16
 Halliday, 18, 44, 45, 47, 55
 Hanks, 18, 72
 Hardie, 31
 Harris, 73, 218, 219, 220, 234, 239, 240, 241, 243, 245, 407, 409
 Hasan, 18, 44, 45, 47, 55
 Haspelmath, 35, 420, 539, 548
 Hausenblas, 83
 Hawkins, 18, 21, 57, 58, 189, 211, 260, 410, 412, 415, 455, 472, 522
 Heim, 64, 617
 Heine, 390, 393, 394, 399, 401, 505, 594
 Herman, 503, 597
 Heusinger, von, 60
 Himmelmann, 18, 24, 28, 40, 47, 49, 50, 52, 58, 60, 68, 73, 74, 77, 79, 117, 118, 165, 166, 167, 168, 169, 191, 192, 209, 214, 215, 217, 220, 222, 224, 244, 245, 246, 259, 260, 279, 390, 399, 400, 401, 402, 405, 406, 407, 408, 409, 410, 428, 442, 471, 553, 571, 575, 590
 Hirschová, 163
 Hlavsa, 18, 21, 163, 180, 425, 590
 Hoffmannová, 83, 84
 Hofherr, 79, 211, 243, 418, 471
 Homoláč, 83, 84
 Hopper, 390, 395, 397, 402, 502, 524, 582, 600
 Horálek, 174
 Horn, 79, 211, 461, 538
 Imbs, 101, 229
 J. Lyons, 18, 44, 45, 46, 59, 63, 117, 118, 238
 Jakobson, 549, 564
 Jodas, 5, 142, 145, 147, 149, 150, 178, 181, 421, 434
 Joly, 21, 44, 45, 46, 47, 50, 55, 117, 118, 245, 249
 Jonasson, 7, 8, 17, 23, 209, 210, 263, 510
 Jungbluth, 17, 71
 Jungmann, 598
 Kalbertodt, Primus et Schumacher, 126
 Kara, 49, 117
 Karlík, 549
 Karttunen, 60
 Keenan, 18, 40, 68, 69, 70, 73, 148, 225, 580, 587
 Kerbrat-Orecchioni, 42, 203, 264, 464
 Kleiber, 5, 7, 8, 17, 22, 23, 29, 40, 41, 43, 44, 46, 47, 48, 49, 53, 54, 55, 56, 57, 59, 62, 63, 65, 66, 67, 68, 78, 79, 80, 117, 119, 120, 180, 185, 202, 207, 209, 210, 218, 224, 233, 235, 238, 244, 251, 252, 255, 256, 259, 263, 264, 268, 286, 442, 474, 496, 506, 509, 510, 511, 522, 526, 544, 545, 561, 567, 570, 585
 Komárek, 150, 163
 König, 397
 Kořenský, 150
 Kosek, 598, 599
 Krámský, 57, 77, 405, 420
 Krčmová, 143, 600
 Křen, 4, 112, 155, 156
 Kresin, 16, 23, 441, 528, 529
 Krifka, 557
 Kripke, 56, 209
 Křížková, 16
 Kroch, 25, 396, 505, 585, 594
 Kuteva, 390, 399, 401
 Lambrecht, 126
 Lamprecht, 174
 Langacker, 390, 395
 Laury, 25, 53, 194, 260, 261, 279, 310, 390, 391, 399, 404, 408, 423, 443, 444, 457, 459, 461, 462, 468, 553, 582, 585, 591
 Le Pesant, 62
 Lehmann, 390, 395
 Löbner, 5, 7, 8, 18, 21, 22, 25, 47, 58, 59, 79, 120, 188, 209, 242, 243, 244, 267, 273, 283, 390, 405, 410, 411, 412, 413, 414, 415, 417, 418, 455, 471, 472, 473, 474, 479, 480, 481, 500, 514, 535, 556, 571, 580, 639
 Loucká, 171
 Machač, 16, 176, 420, 594
 Maillard, 44

Marchello-Nizia, 18, 68, 71, 79, 80, 128, 203,
 222, 233, 235, 236, 237, 238, 239, 240, 269,
 390, 392, 393, 394, 395, 396, 397, 398, 503,
 505, 520, 593, 594, 597
 Martin, 61, 101, 229
 Mathesius, 7, 8, 16, 21, 22, 23, 125, 162, 179,
 180, 181, 189, 197, 206, 207, 208, 213, 390,
 420, 421, 424, 425, 428, 491, 495, 502, 508,
 509, 529
 McEnery, 31
 Meillet, 390, 392, 393, 395, 396
 Meira, 71, 234, 571
 Meyerstein, 7, 8, 16, 21, 420
 Milano, da, 17
 Milner, 61, 119, 184, 264, 437
 Nádvorníková, 1, 4, 15, 139, 204, 212, 294,
 489, 553
 Nebeská, 144
 Nédoluzko, 16, 17, 177, 184, 422, 596
 Noailly, 17
 Orlandini, 7, 8, 18, 21, 79, 175, 420
 Ortmann, 47, 73, 79, 120, 191, 209, 244, 390,
 415, 416, 417, 418, 420, 429, 471, 476, 502,
 580, 582, 592, 596
 Paveau et Sarfati, 516
 Pešek, 1, 16, 17, 62, 119, 183, 184, 185, 186,
 255, 648, 649
 Petr, 174
 Peyraube, 390, 395, 398
 Prévost, 127
 Prince, 57, 60, 242, 279, 522
 Quémada, 101
 Quirk, 191, 220
 Reichenbach, 53, 570
 Reichler-Béguelin, 61
 Rioul, Pellat et Riegel, 124, 128, 129
 Roberts, 64
 Russell, 57, 262, 411, 517, 522
 Sag, 60, 242, 262, 485
 Schnedecker, 49, 225
 Schneiderová, 17, 122, 163
 Schøsler, 409
 Schwarz, 18, 120, 244, 419, 480, 485, 554
 Selig, 79, 460
 Short, 15, 144, 598
 Šimík, 5, 17, 57, 211, 248, 442, 549, 550, 551,
 556, 557, 567
 Simonenko, 409
 Sláma, 549, 550, 551, 567
 Šlosar, 174, 177, 598, 599
 Šonková, 16, 158
 Souček, 16, 17, 21, 184, 420, 422, 426
 Spevak, 248
 Štícha, 17, 21, 22, 144, 150, 152, 162, 173,
 179, 184, 185, 186, 187, 207, 422, 437, 438,
 439, 448, 508, 530, 588
 Strawson, 518
 Studler, 244, 419, 420, 596, 639
 Tellier, 557
 Theissen, 258
 Todorov, 46, 518
 Traugott, 167, 390, 392, 394, 395, 396, 397,
 398, 402, 404, 502, 506, 524, 582, 600
 Trávníček, 22, 151, 162, 171, 173, 182, 184,
 188, 207, 469, 470, 509
 Trovesi, 15, 17, 21, 122, 195, 199, 209, 260,
 400, 420, 421, 422, 426, 441, 460, 491, 504,
 596
 Uhlířová, 17, 163, 165, 184, 367, 428, 452,
 453, 465, 543, 590
 Vaillant, 145, 174, 176, 178
 Valois, 557
 Veselovská, 114, 163, 164, 248, 428, 576, 590
 Vey, 7, 8, 21, 144, 180, 197, 390, 420, 425
 Vigier, 129
 Vinay et Darbelnet, 369
 Volín, 190, 191, 192, 475
 Voráč, 17, 22, 184, 207, 509
 Vuillaume, 40, 41, 42, 53, 570
 Wiederspiel, 49
 Wilmet, 17, 63, 67, 78, 222, 241, 242, 246,
 333, 499, 536, 543, 602, 620
 Wittgenstein, 517
 Yule, 50, 52, 53, 252
 Zíková, 7, 8, 15, 16, 21, 25, 125, 126, 128, 130,
 162, 176, 180, 184, 190, 191, 192, 194, 279,
 309, 390, 391, 408, 420, 421, 422, 426, 432,
 441, 442, 443, 444, 457, 459, 475, 585, 591,
 594
 Zimová, 18, 184, 185
 Zribi-Hertz, 79, 243
 Zubatý, 190, 423, 496, 655

Annexes

Cartes



Figure 3.1 : Trois régions historiques de l'actuelle République tchèque - la Bohême (vert), la Moravie (bleu) et la Silésie (orange). Source : https://fr.wikipedia.org/wiki/Sil%C3%A9sie_tch%C3%A8que#/media/Fichier:CZ_Cechy_Morava_kraje.gif.



Figure 3.2 : Découpage régional de l'actuelle République tchèque. Source : <https://www.axl.cefan.ulaval.ca/europe/reptcheque-1General-demo.htm>.

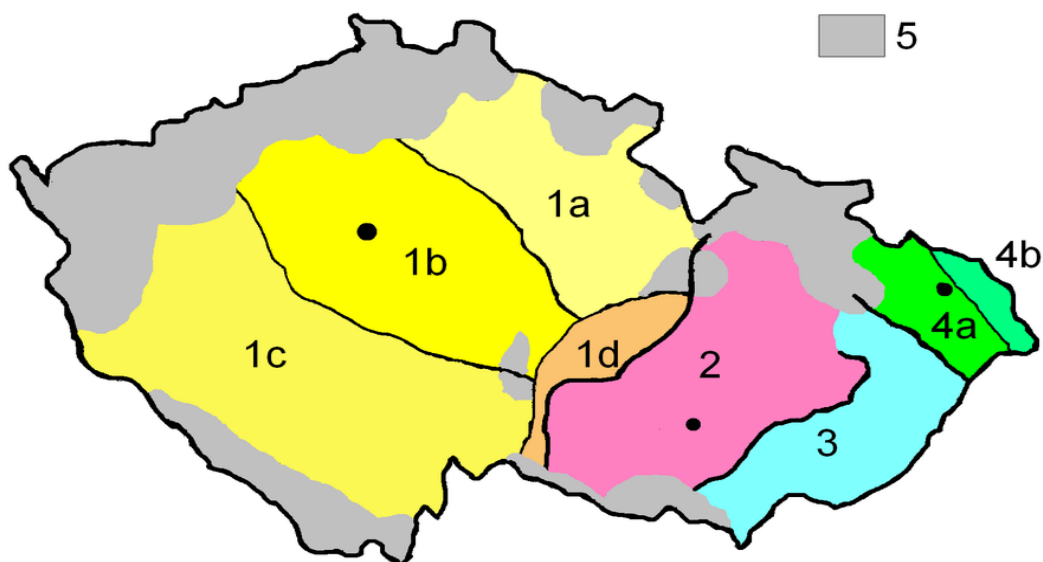


Figure 3.3 : Carte des dialectes de l'actuelle République tchèque.

1. groupe de Bohême : 1a – sous-groupe de Bohême du Nord et de l'Est, 1b – sous-groupe de Bohême centrale, 1c – sous-groupe de Bohême du Sud et de l'Ouest, 1d – sous-groupe de Bohême-Moravie
2. groupe de Moravie centrale
3. groupe de Moravie de l'Est
4. groupe silésien : 4a – sous-groupe moravo-silésien, 4b – sous-groupe siléso-polonais
5. zones de cohabitation de dialectes divers (anciens territoires germanophones)

Source :

https://cs.wikipedia.org/wiki/N%C3%A1%C5%99e%C4%8D%C3%AD_%C4%8De%C5%A1tiny#/media/Soubor:Czech_dialects.PNG ; notre traduction

Liste des abréviations

ACC = accusatif

AUX = verbe auxiliaire

CNT = Corpus national tchèque

CONJ = conjonction

CZ = tchèque

DAT = datif

DE = allemand

ENG = anglais

ES = espagnol

F = féminin

FR = français

FUT = futur

GEN = génitif

IMP = impératif

INF = infinitif

INS = instrumental

IT = italien

LOC = locatif

M = masculin

MOD = terme modal (adverbe, particule)

N = neutre

NEG = négation

NOM = nominatif

NP = nom propre de personne

PL = pluriel

PLtantum = noms étant employés exclusivement au pluriel

PRS = présent

PST = passé

PTCP = participe passé

SG = singulier

REFL = pronom réfléchi

REL = particule relative

SUFFIXE DEICTIQUE = *-hle/-dle* dans certains pronoms et adverbes déictiques tchèques

SUP = le préfixe *nej-* du superlatif tchèque

TEN = le lexème tchèque TEN en emploi adnominal

VOC = vocatif

1-3 = personnes (1^{ère}, 2^{ème} et 3^{ème})

Liste des tableaux

Tableau 1.1 : Système du démonstratif nominal en français (une simplification)

Tableau 1.2 : Système du démonstratif nominal en tchèque (une simplification)

Tableau 1.3 : Système du démonstratif en japonais

Tableau 1.4 : Système de l'article défini en macédonien

Tableau 2.1 : Propriétés du corpus ORTOFON v1 (CNT)

Tableau 2.2 : Sous-corpus du CEFC (Orféo) retenus pour notre corpus de français oral informel avec, pour chacun, le nombre de conversations et le nombre de positions de l'ensemble

Tableau 2.3 : Propriétés de notre corpus de français oral informel constitué à partir des conversations privées du CEFC (Orféo)

Tableau 2.4 : Propriétés du sous-corpus des quatre pièces de V. Havel constitué à partir d'InterCorp v10 (CNT)

Tableau 2.5 : Informateurs tchécophones natifs avec leurs métadonnées

Tableau 2.6 : Informateurs francophones natifs avec leurs métadonnées

Tableau 2.7 : Catégories et valeurs d'annotation appliquées à l'échantillon de 1 000 occurrences du CEFC (Orféo)

Tableau 2.8 : Catégories et valeurs d'annotation appliquées à l'échantillon de 1 000 occurrences d'ORTOFON v1 (CNT)

Tableau 2.9 : Catégories et valeurs d'annotation appliquées à l'échantillon de 241 occurrences des pièces de V. Havel (InterCorp v10 ; CNT)

Tableau 3.1 : Paradigme morphologique de TEN (tchèque soutenu)

Tableau 3.2 : Paradigme morphologique de TENTO (tchèque soutenu)

Tableau 3.3 : Paradigme morphologique de TENHLE (tchèque soutenu)

Tableau 3.4 : Paradigme morphologique de TEN (« tchèque commun »)

Tableau 3.5 : Démonstratif en tchèque : synthèse des paradigmes

Tableau 3.6 : Position des 8 lemmes du démonstratif tchèque les plus représentés dans *A Frequency Dictionary of Czech* (2011) avec leur rang et leur fréquence réduite moyenne normalisée indiquée en positions

Tableau 3.7 : Position des 7 lemmes du démonstratif tchèque les plus représentés dans *Frekvenční slovník mluvené češtiny* (2007) avec leur rang et leur fréquence relative indiquée en positions

Tableau 3.8 : Fréquences absolues et relatives des lemmes TEN, TENTO, TENHLE et TENHLETEN dans les corpus ORTOFON v1 (CNT ; 1 236 508 de positions) et SYN2015 (CNT ; 120 748 715 de positions)

Tableau 3.9 : Scores de la fonction de vraisemblance appliquée aux différences entre les fréquences absolues des lemmes TEN, TENTO, TENHLE et TENHLETEN dans les corpus ORTOFON v1 (CNT) et SYN2015 (CNT) ; valeur critique = 3,841

Tableau 3.10 : Fréquences absolues et relatives des lemmes TEN, TENTO, TENHLE et TENHLETEN dans les trois parties du corpus SYN2015 (CNT) : fiction (41 619 587 de positions), non-fiction (39 384 709 de positions) et presse (39 744 419 de positions)

Tableau 3.11 : Scores de la fonction de vraisemblance appliquée aux différences entre les fréquences absolues des lemmes TEN, TENTO, TENHLE et TENHLETEN dans a) la partie non-fiction du corpus SYN2015 et la partie fiction du corpus SYN2015 et b) la partie presse du corpus SYN2015 et la partie fiction du corpus SYN2015 ; valeur critique = 3,841

Tableau 3.12 : Fréquences absolues et relatives des lemmes TEN, TENTO, TENHLE et TENHLETEN dans l'ensemble du corpus ORTOFON v1 (CNT ; 1 236 508 de positions), la partie Bohême du corpus d'ORTOFON v1 (CNT ; 739 963 positions) et le corpus du *Frekvenční slovník mluvené češtiny* (2007 ; 780 322 positions)

Tableau 3.13 : Scores de la fonction de vraisemblance appliquée aux différences entre les fréquences absolues des lemmes TEN, TENTO, TENHLE et TENHLETEN dans a) la partie Bohême du corpus d'ORTOFON v1 et l'ensemble du corpus ORTOFON v1 et b) le corpus du *Frekvenční slovník mluvené češtiny* (2007) et l'ensemble du corpus ORTOFON v1 ; valeur critique = 3,841

Tableau 3.14 : Formes nominatives des pronoms personnels tchèques de la 3^{ème} personne

Tableau 4.1 : Système du démonstratif en français (Bonnard, 1972 : 1206)

Tableau 4.2 : Rang et fréquence relative des formes du déterminant démonstratif CE pour la période 1946-1964 dans le *Dictionnaire des fréquences : vocabulaire littéraire des XIX^e et XX^e siècles* (1971)

Tableau 4.3 : Fréquences absolue et relative du déterminant démonstratif CE dans la tranche 1946-1964 du corpus du *Dictionnaire des fréquences* (1971), le corpus Écrits scientifiques en français (ScienQuest), le corpus de l'Est républicain (ScienQuest), l'échantillon de textes de fiction publiés entre 1961 et 2014 (Frantext) et le corpus de conversations du CEFC (Orféo)

Tableau 4.4 : Scores de la fonction de vraisemblance appliquée aux différences entre les fréquences absolues du lemme CE dans a) le corpus Écrits scientifiques en français (ScienQuest) et le corpus de l'Est républicain (ScienQuest), b) le corpus Écrits scientifiques en français (ScienQuest) et l'échantillon de textes de fiction publiés entre 1961 et 2014 (Frantext), c) le corpus de l'Est républicain (ScienQuest) et l'échantillon de textes de fiction publiés entre 1961 et 2014 (Frantext) et d) le corpus de conversations du CEFC (Orféo) et l'échantillon de textes de fiction publiés entre 1961 et 2014 (Frantext) ; valeur critique = 3,841

Tableau 4.5 : Passage du système du démonstratif latin au système du démonstratif en ancien français (Carlier et Guillot, 2015 : 358)

Tableau 4.6 : Système des démonstratifs au début de l'ancien français

Tableau 5.1 : Fréquence absolue du démonstratif en emploi adnominal avec et sans bruit et fréquence relative dans les conversations du CEFC (Orféo), ORTOFON v1 (CNT) et les pièces de V. Havel (InterCorp v10, CNT)

Tableau 5.2 : Scores de la fonction de vraisemblance appliquée aux différences entre les fréquences absolues du démonstratif en emploi adnominal dans a) ORTOFON v1 (CNT) et le corpus de conversations du CEFC (Orféo) et b) ORTOFON v1 (CNT) et le corpus des pièces de V. Havel (InterCorp v10, CNT) ; valeur critique = 3,841

Tableau 5.3 : Rang et fréquence relative des formes du lemme LE pour la période 1946-1964 dans le *Dictionnaire des fréquences : vocabulaire littéraire des XIXe et XXe siècles* (1971 ; entre 4 et 6 millions de mots)

Tableau 5.4 : Fréquence absolue et relative de l'article défini LE dans la tranche 1946-1964 du corpus du *Dictionnaire des fréquences* (1971), le corpus Écrits scientifiques en français (ScienQuest), le corpus de l'Est républicain (ScienQuest), l'échantillon de textes de fiction publiés entre 1961 et 2014 (Frantext) et le corpus de conversations du CEFC (Orféo)

Tableau 5.5 : Fréquence relative de l'article défini LE dans l'Est républicain (ScienQuest), l'échantillon de textes de fiction publiés entre 1961 et 2014 (Frantext), le corpus Écrits scientifiques en français (ScienQuest), la tranche 1946-1964 du corpus du *Dictionnaire des fréquences* (1971) et le corpus de conversations du CEFC (Orféo) ; fréquence relative de TEN adnominal dans le corpus de conversations d'ORTOFON v1 (CNT)

Tableau 5.6 : Types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences) avec les intervalles de confiance anticipant leur fréquence dans un autre échantillon de la même taille

Tableau 5.7 : Présence de *-ci* et *-là* avec les différents types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)

Tableau 5.8 : Position de l'occurrence du déterminant démonstratif CE par rapport au verbe de la proposition dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)

Tableau 5.9 : Rapport entre la position postverbale et les différents types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 359 occurrences)

Tableau 5.10 : Rapport entre la position de complément de phrase et les différents types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 356 occurrences)

Tableau 5.11 : Rapport entre la position de la dislocation à droite et les différents types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 144 occurrences)

Tableau 5.12 : Rapport entre la position inconnue et les différents types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 70 occurrences)

Tableau 5.13 : Rapport entre la position de la dislocation à gauche et les différents types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 48 occurrences)

Tableau 5.14 : Rapport entre la position préverbale et les différents types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 19 occurrences)

Tableau 5.15 : Proportion des occurrences du déterminant démonstratif CE affichant des disfluences pour chaque type et sous-type d'emploi dans l'échantillon de conversations du CEFC (Orféo ; 69 occurrences)

Tableau 5.16 : Proportion des occurrences du déterminant démonstratif CE affichant des disfluences pour chaque position de l'occurrence par rapport au verbe dans l'échantillon de conversations du CEFC (Orféo ; 69 occurrences)

Tableau 5.17 : Types et sous-types d'emploi du démonstratif adnominal tchèque (tous lemmes confondus) dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences) avec les intervalles de confiance anticipant leur fréquence dans un autre échantillon de la même taille

Tableau 5.18 : Présence des lemmes du démonstratif tchèque autres que TEN en emploi adnominal dans l'échantillon de 1 000 occurrences d'ORTOFON v1 (CNT) avec leurs sous-types d'emploi

Tableau 5.19 : Proportion des occurrences de TEN adnominal évaluées comme engagées dans le processus de grammaticalisation pour chaque type et sous-type d'emploi dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Tableau 5.20 : Proportion des occurrences du démonstratif adnominal se cumulant avec un autre mot fonctionnel pour chaque type et sous-type d'emploi dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 70 occurrences)

Tableau 5.21 : Différents types de cumuls du démonstratif adnominal avec un autre mot fonctionnel et les différents types et sous-types d'emploi au sein desquels ils apparaissent dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Tableau 5.22 : Position de l'occurrence du démonstratif adnominal par rapport au verbe de la proposition dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Tableau 5.23 : Rapport entre la position postverbale et les différents types et sous-types d'emploi du démonstratif adnominal dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Tableau 5.24 : Rapport entre la position préverbale et les différents types et sous-types d'emploi du démonstratif adnominal dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Tableau 5.25 : Rapport entre la position inconnue et les différents types et sous-types d'emploi du démonstratif adnominal dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Tableau 5.26 : Rapport entre la position de complément de phrase et les différents types et sous-types d'emploi du démonstratif adnominal dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Tableau 5.27 : Rapport entre la position de la dislocation à droite et les différents types et sous-types d'emploi du démonstratif adnominal dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Tableau 5.28 : Rapport entre la position de la dislocation à gauche et les différents types et sous-types d'emploi du démonstratif adnominal dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Tableau 5.29 : Proportion des occurrences du démonstratif adnominal affichant des disfluences pour chaque type et sous-type d'emploi dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Tableau 5.30 : Proportion des occurrences du démonstratif adnominal affichant des disfluences pour chaque position de l'occurrence par rapport au verbe dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Tableau 5.31 : Types et sous-types d'emploi du démonstratif adnominal tchèque (tous lemmes confondus) dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences) avec les intervalles de confiance anticipant leur fréquence dans un autre échantillon de la même taille

Tableau 5.32 : Présence des lemmes du démonstratif tchèque en emploi adnominal autres que TEN dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Tableau 5.33 : Fréquence absolue du lemme TEN et celle des autres lemmes du démonstratif en emploi adnominal dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Tableau 5.34 : Proportion des occurrences de TEN adnominal évaluées comme engagées dans le processus de grammaticalisation pour chaque type et sous-type d'emploi dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Tableau 5.35 : Proportion des occurrences du démonstratif adnominal se cumulant avec un autre mot fonctionnel pour chaque type et sous-type d'emploi dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Tableau 5.36 : Différents types de cumuls du démonstratif adnominal avec un autre mot fonctionnel et les différents types et sous-types d'emploi avec lesquels ils apparaissent dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Tableau 5.37 : Fréquence relative (nombre d'occurrences pour un million de positions) des cumuls TAKOVÝ TEN, « démonstratif (sauf TAKOVÝ) + possessif » et NEJAKÝ TEN dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 48 568 positions), le corpus de théâtre de référence (CNT ; 1 301 447 de positions), ORTOFON v1 (CNT ; 1 236 508 de positions) et la partie « fiction » de SYN2015 (CNT ; 41 619 587 de positions)

- Tableau 5.38** : Différents types de transposition du démonstratif adnominal tchèque dans les traductions françaises du sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT)
- Tableau 5.39** : Différents (sous-)types d'emploi des occurrences du démonstratif adnominal tchèque dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT) rendues, en français, par le déterminant démonstratif (InterCorp v1, CNT ; 241 occurrences)
- Tableau 5.40** : Différents (sous-)types d'emploi des occurrences du démonstratif adnominal tchèque dans les traductions françaises du sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT) rendues, en français, par l'article défini (InterCorp v1, CNT ; 241 occurrences)
- Tableau 5.41** : Différents (sous-)types d'emploi des occurrences du démonstratif adnominal tchèque dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT) rendues, en français, par l'article indéfini (InterCorp v1, CNT ; 241 occurrences)
- Tableau 5.42** : Différents (sous-)types d'emploi des occurrences du démonstratif adnominal tchèque dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT) rendues, en français, par le déterminant possessif (InterCorp v1, CNT ; 241 occurrences)
- Tableau 5.43** : Différents (sous-)types d'emploi des occurrences du démonstratif adnominal tchèque dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT) rendues, en français, par l'article défini (InterCorp v1, CNT ; 241 occurrences)
- Tableau 5.44** : Différents (sous-)types d'emploi des occurrences du démonstratif adnominal tchèque dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT) rendues, en français, par l'article indéfini (InterCorp v1, CNT ; 241 occurrences)
- Tableau 5.45** : Noms à sémantique temporelle en emploi endophorique (fréquence, sous-type(s) d'emploi et position(s) par rapport au verbe) dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)
- Tableau 5.46** : Les 7 noms à sémantique temporelle en emploi endophorique les plus fréquents (fréquence, sous-type(s) d'emploi et position(s) par rapport au verbe) dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)
- Tableau 5.47** : Présence de particules localisatrices avec les 7 noms à sémantique temporelle les plus fréquents en emploi endophorique (fréquence, sous-type(s) d'emploi) dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)
- Tableau 5.48** : Les 7 noms à sémantique temporelle les plus fréquents en emploi déictique temporel (fréquence, position par rapport au verbe, présence de particules localisatrices) dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)
- Tableau 5.49** : Données utilisées pour le calcul de spécificités appliqué au lien entre *CE + fois + -ci* et *CE + moment + -là* dans l'ensemble des 1 871 occurrences générées par la requête au sein des conversations du CEFC (Orféo)
- Tableau 6.1** : Les quatre types de noms dans la théorie CTD (Löbner, 2011)
- Tableau 6.2** : Le démonstratif et le double paradigme de l'article défini dans les dialectes allemands de Suisse (Studler, 2014 : 144)
- Tableau 6.3** : Nombre d'occurrences de TEN adnominal apparaissant en emplois anaphoriques fidèle, infidèle et indirect en fonctions de sujet, COD, COI, complément du nom et complément circonstanciel dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)
- Tableau 6.4** : Nombre d'occurrences de TEN adnominal apparaissant en emplois anaphoriques fidèle, infidèle et indirect en fonctions de sujet, COD, COI, complément du nom et complément circonstanciel dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)
- Tableau 6.5** : Nombre d'occurrences où l'unicité dénotative du nom est instaurée par a) une subordonnée relative « établissante », b) une subordonnée conjonctive pure et c) un complément prépositionnel/adverbial dans l'échantillon de 1 000 occurrences d'ORTOFON v1 (CNT) et dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)
- Tableau 6.6** : Nombre de personnes ayant choisi d'employer TEN avec les exemples 80, 81, 82 et 85 (sur 10 informateurs natifs)
- Tableau 6.7** : Nombre de personnes ayant choisi d'employer TEN avec les exemples 89, 90, 93, 100 et 103 (sur 10 informateurs natifs)
- Tableau 6.8** : Nombre d'occurrences où l'unicité dénotative du nom est instaurée par a) un adjectif au superlatif, b) un numéral ordinal et c) un autre complément analogue dans l'échantillon de 1 000 occurrences d'ORTOFON v1 (CNT) et dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)
- Tableau 6.9** : Nombre de personnes ayant choisi d'employer TEN avec les exemples 105, 112 et 113 (sur 10 informateurs natifs)
- Tableau 7.1** : Nombre d'occurrences appartenant aux trois types de valeurs pragmatiques exprimées par le démonstratif au sein des emplois « émotionnels » dans l'échantillon de 1 000 occurrences d'ORTOFON v1 (CNT ; 150 premières occurrences) et dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 50 occurrences concernées au total)
- Tableau 7.2** : Nombre d'occurrences appartenant aux trois types de valeurs pragmatiques exprimées par le démonstratif au sein des 150 premiers emplois « émotionnels » de l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences) avec les intervalles

de confiance anticipant leur fréquence dans un autre échantillon de la même taille au sein de la totalité des 331 emplois « émotionnels » recensés

Tableau 8.1 : Les étapes du processus de grammaticalisation de TEN adnominal avec les étapes actuelles (rouge) et le seuil de l'article défini constitué (séparateur gras)

Liste des figures

Figure 2.1 : Interface d'ORTOFON v1 (CNT)

Figure 2.2 : Métadonnées du corpus ORTOFON v1 (CNT)

Figure 2.3 : Interface du CEFC (Orféo) : choix des textes

Figure 2.4 : Métadonnées des conversations du CEFC (Orféo)

Figure 2.5 : Interface du CEFC (Orféo)

Figure 2.6 : Interface du sous-corpus des quatre pièces de V. Havel (InterCorp v10, CNT) et de leur traduction française

Figure 2.7 : Interface « 2 words in 2 corpora » (CNT) ; application de la fonction de vraisemblance aux fréquences absolues du lemme TEN dans les corpus ORTOFON v1 et SYN2015

Figure 2.8 : Interface « 2 words in 1 corpus » (CNT) ; application de la fonction de vraisemblance aux fréquences absolues des lemmes TENHLE et TENTO dans le corpus ORTOFON v1

Figure 2.9 : Interface « 1 feature – many samples » (CNT) ; évaluation de la représentativité de 24 échantillons du démonstratif adnominal (choisis au hasard) eu égard à la proportion de bruit dans l'ensemble des occurrences (24 096) générées par la requête

Figure 3.1 : Trois régions historiques de l'actuelle République tchèque – la Bohême (vert), la Moravie (bleu) et la Silésie (orange).

Figure 3.2 : Découpage régional de l'actuelle République tchèque.

Figure 3.3 : Carte des dialectes de l'actuelle République tchèque.

Figure 4.1 : « Quelle forme emploieriez-vous si vous étiez la personne à gauche et que vous référiez a) à la boîte rouge et b) à la boîte verte ? »

Figure 4.2 : « Quelles formes emploieriez-vous si vous étiez la personne à gauche et que vous référiez, dans un même énoncé, au livre vert et au livre bleu ? »

Figure 5.1 : Types et sous-types d'emploi du déterminant démonstratif CE dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)

Figure 5.2 : Présence des différents (sous-)types d'emploi du déterminant démonstratif CE au sein des occurrences contenant *-là* dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)

Figure 5.3 : Calcul de spécificités appliqué au lien entre les différents (sous-)types d'emploi du déterminant démonstratif CE et la présence de particules localisatrices dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences) ; rouge = sans particule, vert = *-là*, bleu = *-ci*

Figure 5.4 : Calcul de spécificités appliqué au lien entre les types d'emploi endophorique, exophorique a-mémoriel et exophorique mémoriel du déterminant démonstratif CE et la présence de particules localisatrices dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences) ; rouge = sans particule, vert = *-là*, bleu = *-ci*

Figure 5.5 : Représentation en % des différentes positions de l'occurrence du déterminant démonstratif CE par rapport au verbe de la proposition dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)

Figure 5.6 : Représentation en % des différents (sous-)types d'emploi du déterminant démonstratif CE au sein de la position postverbale dans l'échantillon de conversations du CEFC (Orféo ; 359 occurrences)

Figure 5.7 : Représentation en % des différents (sous-)types d'emploi du déterminant démonstratif CE au sein de la position de complément de phrase dans l'échantillon de conversations du CEFC (Orféo ; 356 occurrences)

Figure 5.8 : Représentation en % des différents (sous-)types d'emploi du déterminant démonstratif CE au sein de la position de la dislocation à droite dans l'échantillon de conversations du CEFC (Orféo ; 144 occurrences)

Figure 5.9 : Représentation en % des différents types d'emploi du déterminant démonstratif CE au sein de la position inconnue dans l'échantillon de conversations du CEFC (Orféo ; 70 occurrences)

Figure 5.10 : Représentation en % des différents (sous-)types d'emploi du déterminant démonstratif CE au sein de la position de la dislocation à gauche dans l'échantillon de conversations du CEFC (Orféo ; 48 occurrences)

Figure 5.11 : Représentation en % des différents (sous-)types d'emploi du déterminant démonstratif CE au sein de la position préverbale dans l'échantillon de conversations du CEFC (Orféo ; 19 occurrences)

Figure 5.12 : Calcul de spécificités appliqué au lien entre les différentes positions de l'occurrence du déterminant démonstratif CE par rapport au verbe de la proposition et les différents (sous-)types d'emploi (hormis les emplois ambigus, « émotionnels », cataphoriques et indéfinis) dans l'échantillon de conversations du CEFC (Orféo ; 938 occurrences)

Figure 5.13 : Représentation en % des différents (sous-)types d'emploi du déterminant démonstratif CE parmi les occurrences affichant des disfluences dans l'échantillon de conversations du CEFC (Orféo ; 69 occurrences)

Figure 5.14 : Représentation en % des différentes positions du déterminant démonstratif CE par rapport au verbe parmi les occurrences affichant des disfluences dans l'échantillon de conversations du CEFC (Orféo ; 69 occurrences)

Figure 5.15 : Calcul de spécificités appliqué au lien entre les types d'emploi du déterminant démonstratif CE et la présence de disfluences dans l'échantillon de conversations du CEFC (Orféo ; 1 000 occurrences)

Figure 5.16 : Types et sous-types d'emploi du démonstratif adnominal (tous lemmes confondus) dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Figure 5.17 : Représentation en % des différents (sous-)types d'emploi parmi les occurrences de TEN adnominal encourageant le processus de grammaticalisation dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Figure 5.18 : Calcul de spécificités appliqué au lien entre les différents (sous-)types d'emploi et les occurrences de TEN adnominal encourageant le processus de grammaticalisation dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Figure 5.19 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal parmi les occurrences où le démonstratif se cumule avec un autre mot fonctionnel dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 70 occurrences)

Figure 5.20 : Représentation en % des différentes positions de l'occurrence du démonstratif adnominal par rapport au verbe de la proposition dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Figure 5.21 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal au sein de la position postverbale dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Figure 5.22 : Calcul de spécificités appliqué au lien entre les différentes positions de l'occurrence du démonstratif adnominal par rapport au verbe de la proposition et les différents types d'emploi dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Figure 5.23 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal au sein de la position préverbale dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Figure 5.24 : Représentation en % des différents types d'emploi du démonstratif adnominal au sein de la position inconnue dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Figure 5.25 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal au sein de la position de complément de phrase dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Figure 5.26 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal au sein de la position de la dislocation à droite dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Figure 5.27 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal au sein de la position de la dislocation à gauche dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Figure 5.28 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal parmi les occurrences affichant des disfluences dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Figure 5.29 : Représentation en % des différentes positions du démonstratif adnominal par rapport au verbe parmi les occurrences affichant des disfluences dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Figure 5.30 : Calcul de spécificités appliqué au lien entre les types d'emploi du déterminant démonstratif CE et la présence de disfluences dans l'échantillon de conversations d'ORTOFON v1 (CNT ; 1 000 occurrences)

Figure 5.31 : Types et sous-types d'emploi du démonstratif adnominal (tous lemmes confondus) dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Figure 5.32 : Représentation en % des différents (sous-)types d'emploi parmi les occurrences de TEN adnominal encourageant le processus de grammaticalisation dans les pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Figure 5.33 : Calcul de spécificités appliqué au lien entre les différents (sous-)types d'emploi et les occurrences de TEN adnominal encourageant le processus de grammaticalisation dans les pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Figure 5.34 : Représentation en % des différents (sous-)types d'emploi du démonstratif adnominal parmi les occurrences où le démonstratif se cumule avec un autre mot fonctionnel dans le sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT ; 241 occurrences)

Figure 5.35 : Différents types de transposition du démonstratif adnominal tchèque dans les traductions françaises du sous-corpus des quatre pièces de V. Havel (InterCorp v1, CNT)

Figure 5.36 : Calcul de spécificités appliqué à CE + *fois* + *-ci* et CE + *moment* + *-là* dans l'ensemble des 1 871 occurrences générées par la requête au sein des conversations du CEFC (Orféo)

Figure 6.1 : Comment l'érosion de la flexion nominale a pu encourager le développement de l'article défini

Questionnaires

Questionnaire n°1

Questionnaire servant à tester certaines thèses relatives au fonctionnement des pronoms démonstratifs en tchèque dans le cadre d'un projet de thèse (1^{ère} partie ; traduction¹)

Tout d'abord, je tiens à vous remercier d'avoir pris le temps de remplir ce bref questionnaire qui m'aidera à tester la validité de certains constats généraux portant sur les pronoms démonstratifs en tchèque. Je voudrais vous demander de choisir, chaque fois, la forme qui vous paraît la plus naturelle et que vous emploieriez **au quotidien, dans une conversation informelle avec vos amis**. Avant de commencer à remplir le questionnaire, veuillez renseigner votre nom, votre année de naissance, ainsi que le/les endroit(s) où vous avez vécu jusqu'à l'âge de vos 18 ans et l'endroit où vous vivez actuellement², ensemble avec le nombre d'années (ou de mois) passés à cet endroit.

Nom et prénom(s) :

Date de naissance :

Lieu de résidence jusqu'à l'âge de vos 18 ans (ville(s) et région(s)) :

Lieu de résidence actuel (ville et région) :

Nombre d'années (de mois) passés à cet endroit :

A Choisissez la forme de pronom démonstratif que vous trouvez la plus appropriée dans le contexte donné

La personne à gauche s'adresse à la personne à droite.



a) Elle lui demande de lui passer la boîte **rouge**, qu'elle pointe simultanément du doigt :

„ <i>Pod-ej</i>	<i>mi</i>	<i>prosim tě</i>	TU	/ TUTO	/ TUHLE	/ ONU
passer-IMP.2SG	moi.DAT	s'il te plaît	ten.F.ACC.SG	tento.F.ACC.SG	tenhle.F.ACC.SG	onen.F.ACC.SG
/ TAMTU		<i>krabic-i.</i> “				
tamten.F.ACC.SG		boîte-ACC.SG				

¹ Les trois questionnaires ont été envoyés aux informateurs dans un document Word. Après les avoir remplis sur ordinateur, ces derniers nous les ont retournés en version électronique.

² L'endroit où vous résidez la plupart du temps.

- b) Elle lui demande de lui passer la boîte **verte**, qu'elle pointe simultanément du doigt :

„*Pod-ej* *mi* *prosim tě* **TU** / **TUTO** / **TUHLE** / **ONU**
 passer-IMP.2SG moi.DAT s'il te plaît ten.F.ACC.SG tento.F.ACC.SG tenhle.F.ACC.SG onen.F.ACC.SG

/ **TAMTU** *krabic-i*. “
 tamten.F.ACC.SG boîte-ACC.SG

B **Laquelle des formes suivantes considérez-vous comme une éventuelle bonne alternative, dans une conversation informelle avec vos amis, pour la forme**

- a. Employée dans la phrase **Aa**: ONEN / TENHLETEN / TAMTEN / TADYHLETEN / TUHLETEN / TEN / TADYHLECTEN / TÁMHLETEN / TÁMLENCTEN / TUTEN / TENTO / TENHLE / TADYTEN
- b. Employée dans la phrase **Ab**: ONEN / TENHLETEN / TAMTEN / TADYHLETEN / TUHLETEN / TEN / TADYHLECTEN / TÁMHLETEN / TÁMLENCTEN / TUTEN / TENTO / TENHLE

Si vous considérez plusieurs formes comme appropriées, marquez-les toutes. Dans le cas où vous marquez plusieurs formes dont certaines vous paraissent plus naturelles que d'autres, essayez de hiérarchiser vos préférences à l'aide d'une échelle numérique (par ex. TAMTEN = 1, TUHLETEN = 2, etc.).

Questionnaire n°2

Questionnaire servant à appuyer certaines thèses relatives au fonctionnement des déterminants démonstratifs en français (dans le cadre d'un projet de thèse)

Tout d'abord, je tiens à vous remercier d'avoir pris le temps de remplir ce bref questionnaire qui, je l'espère, m'aidera à tester la validité de certains constats portant sur le système des déterminants démonstratifs en français. Je voudrais vous demander de choisir, chaque fois, la forme qui vous paraît la plus naturelle et que vous emploieriez **au quotidien, dans une conversation informelle avec vos amis**. Avant de commencer à remplir le questionnaire, veuillez renseigner votre nom, votre année de naissance, ainsi que l'endroit où vous avez vécu jusqu'à l'âge de vos 18 ans et l'endroit où vous vivez actuellement¹, ensemble avec le nombre d'années (ou de mois) passés à cet endroit.

Nom et prénom(s) :

Date de naissance :

Lieu de résidence jusqu'à l'âge de vos 18 ans (ville(s) et région(s)) :

Lieu de résidence actuel (ville et région) :

Nombre d'années (de mois) passés à cet endroit :

A **Choisissez la forme de déterminant démonstratif que vous trouvez la plus appropriée dans le contexte donné**

La personne à gauche s'adresse à la personne à droite.



- a) Elle lui demande de lui passer la boîte **rouge**, qu'elle pointe simultanément du doigt :

*« Passe-moi, s'il te plaît, **cette** boîte / **cette** boîte-ci / **cette** boîte-là. »*

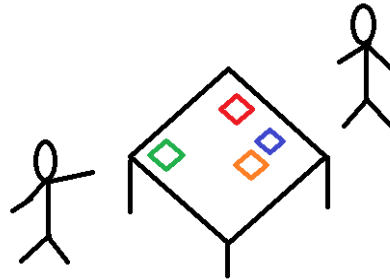
- b) Elle lui demande de lui passer la boîte **verte**, qu'elle pointe simultanément du doigt :

*Passe-moi, s'il te plaît, **cette** boîte/ **cette** boîte-ci / **cette** boîte-là.*

¹ L'endroit où vous résidez la plupart du temps.

B Choisissez la forme de déterminant démonstratif que vous trouvez la plus appropriée dans le contexte donné

Le contexte est celui d'une librairie. Le client, à gauche, s'adresse au libraire, à droite, pour lui demander de lui passer le livre **vert** et le livre **bleu**, qu'il a l'intention d'acheter.



« Je vais vous prendre **ce** livre / **ce** livre-ci / **ce** livre-là et **ce** livre / **ce** livre-ci / **ce** livre-là ».

C Avez-vous des commentaires / remarques à faire au sujet de vos choix ? (facultatif)

Questionnaire n°3

Questionnaire servant à appuyer certaines thèses relatives au fonctionnement des pronoms démonstratifs en tchèque dans le cadre d'un projet de thèse (2nde partie ; traduction¹)

Tout d'abord, je tiens à vous remercier d'avoir encore pris le temps de remplir ce bref questionnaire qui m'aidera à tester la validité de certains constats généraux portant sur les pronoms démonstratifs en tchèque. Je voudrais vous demander de choisir, chaque fois, la forme qui vous paraît la plus naturelle et que vous emploieriez **au quotidien, dans une conversation informelle avec vos amis, à l'exception des énoncés n°4 et 21, où il est question du registre formel. Marquez toujours la forme que vous trouvez la plus appropriée dans le contexte donné (c'est-à-dire toujours une seule).** Si vous avez des remarques à me présenter (relativement, par ex., aux circonstances dans lesquelles vous choisiriez plutôt telle forme ou telle autre), je vous saurai gré de me les mettre dans les commentaires.

- 1) La personne à gauche s'adresse à la personne à droite pour lui demander de lui passer le ballon :



- a) *D-ej* *mi* *míč!*
donner-IMP.2SG moi.DAT ballon.ACC.SG
- b) *D-ej* *mi* ***tenhle*** *míč!*
donner-IMP.2SG moi.DAT *tenhle*.M.ACC.SG ballon.ACC.SG
- c) *D-ej* *mi* ***ten*** *míč!*
donner-IMP.2SG moi.DAT *ten*.M.ACC.SG ballon.ACC.SG
Passe-moi le ballon !

- 2) La personne parle du temps qu'il fait du 11 au 17 janvier 2021² :

- a) ***Ten*** *tejden* *je* *hrozn-ý* *počas-í.*
ten.M.ACC.SG semaine.ACC.SG être.PRS.3SG horrible-N.NOM.SG temps-NOM.SG
- b) ***Tenhle*** *tejden* *je* *hrozný* *počasí.*
tento.M.ACC.SG semaine.ACC.SG être.PRS.3SG horrible-N.NOM.SG temps-NOM.SG

¹ Dans les traductions des exemples, nous avons marqué en gras les démonstratifs dans les options qui en contiennent. Ce choix a évidemment pour seul but de faciliter l'orientation du lecteur francophone et n'a pas été adopté dans la version originale, destinée à nos informateurs tchèques.

² Soit la semaine de l'emploi de l'énoncé ; cette précision est présente pour donner aux informateurs l'instruction d'interpréter l'énoncé comme relevant de la deixis temporelle.

Cette semaine, il fait un temps horrible.

3)

- a) *Koupi-l* *jsem* *si* *aut-o* *a* *motork-u*
 acheter-PTCP.SG.M AUX.PST.1SG REFL voiture-ACC.SG and moto-ACC.SG
- To* *aut-o* *jezd-í* *výborně, ale s tou*
 ten.N.NOM.SG voiture-NOM.SG rouler-PRS.3SG parfaitement mais avec ten.F.INS.SG
- motork-ou* *jsou* *pořád problém-y.*
 moto-INS.SG être.PRS.3PL toujours problème-NOM.PL

- b) *Koupi-l* *jsem* *si* *aut-o* *a* *motork-u*
 acheter-PTCP.SG.M AUX.PST.1SG REFL voiture-ACC.SG and moto-ACC.SG
- Aut-o* *jezd-í* *výborně, ale s*
 voiture-NOM.SG rouler-PRS.3SG parfaitement mais avec
- motork-ou* *jsou* *pořád problém-y.*
 moto-INS.SG être.PRS.3PL toujours problème-NOM.PL

J'ai acheté une voiture et une moto. La voiture roule parfaitement, mais il y a tout le temps des soucis avec la moto. (exemple inspiré par Pešek, 2014 : 156)

4) Les phrases a), b) et c) enchaînent directement sur la phrase citée ci-dessous :

Z cest *po jižn-í* *Americ-e* *si* *přivez-l*
 de voyage.GEN.PL après du sud-F.LOC.SG Amérique-LOC.SG REFL ramener-PST.3SG.M

dv-ě *žáb-y* *rod-u* *dendrobates*
 deux-F.ACC.PL grenouille-ACC.PL genre-GEN.SG dendrobates.NOM.SG

- a) *Tito* *obojživelníc-i* *m-a-jí* *ve sv-ých* *žláz-ách*
 tento.M.NOM.PL batracien.NOM.PL avoir-PRS-3PL dans leur.F.LOC.PL gland-LOC.PL
- smrteln-ý* *jed* *kurare.*
 mortel-M.ACC.SG poison.ACC.SG curare.ACC.SG
- b) *Ti* *obojživelníc-i* *m-a-jí* *ve sv-ých* *žláz-ách*
 ten.M.NOM.PL batracien.NOM.PL avoir-PRS-3PL dans leur.F.LOC.PL gland-LOC.PL
- smrteln-ý* *jed* *kurare.*
 mortel-M.ACC.SG poison.ACC.SG curare.ACC.SG
- c) *Obojživelníc-i* *m-a-jí* *ve* *sv-ých* *žláz-ách*
 batracien.NOM.PL avoir-PRS-3PL dans leur.F.LOC.PL gland-LOC.PL
- smrteln-ý* *jed* *kurare.*
 mortel-M.ACC.SG poison.ACC.SG curare.ACC.SG

De ses voyages en Amérique du Sud, il a ramené deux grenouilles du genre dendrobates. Les glands de ces batraciens contiennent le poison mortel curare.
(Pešek, 2014 : 159)

5) Les phrases a) et b) enchaînent directement sur l'extrait cité ci-dessous :

L1 : ti soused-i jak jsem ti vo nich vyprávě-l-a
ten.M.NOM.PL voisin-NOM.PL REL AUX.PST.1SG toi.DAT de eux.LOC raconter-PTCP-SG.F

jak tam bydle-l-i u mě . ty NP jak . voni
REL là-bas habiter-PST-3PL.M chez moi.ACC ten.M.NOM.PL NP REL eux.NOM

mě-l-i tapety tam padesát let
avoir-PST-3PL.M papier peint-NOM.PL là-bas cinquante an.GEN.PL

*L2 : no tak to já je ne**
MOD moi.NOM eux.ACC NEG

L1 : a ona ještě byl-a kuřák . a vona rád-a
et elle.NOM.SG en plus être.PST-3SG.F fumeur.NOM.SG et elle.NOM.SG content-F.SG

děla-l-a tydlet-y panenky víš
fabriquer-PST-3SG.F tenhleten-F.ACC.PL poupée-ACC.PL MOD

L2 : já v-í-m
moi.NOM savoir-PRS-1SG

L1 : ona skoro byl-a už na smrt nemocn-á .
elle.NOM.SG presque être.PST-3SG.F déjà à mort.ACC.SG malade-F.NOM.SG

L1 : já jsem tam příš-l-a ona mi
moi.NOM AUX.PST.1SG là-bas arriver-PTCP-SG.F elle.NOM.SG moi.DAT

ukazova-l-a panenky .
montrer-PST-3SG.F poupée-ACC.PL

já řík-á-m no Dáš-o ale je to na úkor .
moi.NOM dire-PRS-1SG MOD NP-F.SG.VOC mais être.PRS.3SG cela.NOM à détriment.ACC.SG

se urazi-l-a . no poněvač . a ted' tam příše-l .
REFL vexer-PST-3SG.F MOD parce que et maintenant là-bas arriver-PST.3SG.M

nov-ej majitel .
nouveau-M.NOM.SG propriétaire.NOM.SG

a ty představ si ty tapet-y muse-l
et ten.F.ACC.PL imaginer.IMP.2SG REFL ten.F.ACC.PL papier peint-ACC.PL devoir-PST.3SG.M

strhat . to byl-y ještě takov-ý ze západn-ího
arracher.INF ce être.PST-3PL.F encore tel-F.NOM.PL de occidental-N.GEN.SG

Německ-a rozumíš .
Allemagne-GEN.SG MOD

proto mi tam děla-l-y sbíječč-y
 c'est pour ça que moi.DAT là-bas faire-PST-3PL.F marteau-piqueur-NOM.PL

to vše-cko šl-o dolů to byl-o příšern-ý
 cela.NOM tout-N.NOM.SG aller.PST-3SG.N en bas ce être.PST-3SG.N horrible-N.NOM.SG

a) *tak dlouho ty tapet-y tam byl-y že jo*
 si longtemps ten.F.NOM.PL papier peint-NOM.PL là-bas être.PST-3PL.F MOD

b) *tak dlouho tapet-y tam byl-y že jo*
 si longtemps papier peint-NOM.PL là-bas être.PST-3PL.F MOD

L1 : ces voisins que je t'ai parlé d'eux, ceux qui vivaient chez moi, ces NP, ceux qui avaient chez eux des papiers peints depuis cinquante ans

L2 : bon, moi, je les connais pas

L1 : et en plus de ça, elle fumait et elle aimait faire ces espèces de poupées, tu vois

L2 : je vois

L1 : elle était presque mortellement malade, j'y suis arrivée, elle m'a montré des poupées, je lui dis mais Dáša, ça, c'est au détriment de... elle s'est vexée parce que bon... et maintenant, il y a un nouveau propriétaire qui est arrivé, et imagine qu'il a été obligé d'arracher ces papiers peints, en plus c'était des papiers peints d'Allemagne de l'Ouest, tu vois, c'est pour ça que j'avais des marteaux-piqueurs à côté de chez moi, il fallait bien enlever tout ça, c'était horrible, ça faisait si longtemps que les papiers peints étaient là, tu vois

ORTOFON v1

6) Les phrases a), b) et c) enchaînent directement sur la phrase citée ci-dessous :

ta tet-a ona tam m-á [...] právě nějak-ou
 ten.F.NOM.SG tante.NOM.SG elle.NOM.SG là-bas avoir-PRS.3SG MOD quelque-F.ACC.SG

kamarádk-u a vona byd-l-í ve Švýcarsk-u a m-á
 copine-ACC.SG et elle.NOM.SG habiter-PRS-3SG en Suisse-LOC.SG et avoir-PRS.3SG

jakoby ten .. její přítel je Švýcar
 MOD ten.M.NOM.SG sien.M.NOM.SG copain.NOM.SG être.PRS.3SG Suisse.NOM.SG

a) *a ta kamarádk-a normálně mluv-í i česky jo*
 et ten.F.NOM.SG copine-NOM.SG MOD parler-PRS.3SG même tchèque MOD

což je jako fajn
 ce qui.NOM être.PRS.3SG MOD cool.N.NOM.SG

b) *a tahle kamarádk-a normálně mluv-í i česky jo*
 et tenhle.F.NOM.SG copine-NOM.SG MOD parler-PRS.3SG même tchèque MOD

což je jako fajn
 ce qui.NOM être.PRS.3SG MOD cool.N.NOM.SG

c) *a kamarádk-a normálně mluv-í i česky jo*
 et copine-NOM.SG MOD parler-PRS.3SG même tchèque MOD

což je jako fajn
 ce qui.NOM être.PRS.3SG MOD cool.N.NOM.SG

cette tante, elle a [...] une copine là-bas et elle habite en Suisse et elle a [...] genre son copain, il est Suisse, et cette copine-là elle parle tchèque et tout, ce qui est cool
 ORTOFON v1

7) Dans une situation avec deux morceaux de gâteau :

L1 : Kter-ý kousek chc-e-š?
 quel-M.ACC.SG morceau.ACC.SG vouloir-PRS-2SG

a) *L2 : Ten menš-í.* b) *L2: Menš-í.*
 ten.M.ACC.SG plus petit-M.ACC.SG plus petit-M.ACC.SG

L1 : Quel morceau veux-tu ?

L2 : Le plus petit (des deux). (Adamec, 1983 : 155)

8) Deux locuteurs regardent des photos prises lors de leurs dernières vacances passées ensemble :

a) *L1 : tady jsou promenád-y ne ? kolem chodník-y*
 ici être.PRS.3PL promenade-NOM.PL MOD autour trottoir-NOM.PL

L2 : jo tady-hlenc je promenád-a
 MOD ici-SUFFIXE DEICTIQUE être.PRS.3SG promenade-NOM.SG

b) *L1 : tady jsou ty promenád-y ne ? kolem*
 ici être.PRS.3PL ten.F.NOM.PL promenade-NOM.PL MOD autour

ty chodník-y
 ten.M.NOM.PL trottoir-NOM.PL

L2 : jo tady-hlenc je ta promenád-a
 MOD ici-SUFFIXE DEICTIQUE être.PRS.3SG ten.F.NOM.SG promenade-NOM.SG

L1 : ici, il y a les/ces promenades et les/ces trottoirs tout autour, non ?

L2 : ouais, là, il y a la/cette promenade ORTOFON v1

9) Deux locuteurs se souviennent d'une connaissance commune :

a) *Pamatuj-eš si na divn-ýho chlápka,*
 rappeler-PST-2SG REFL à étrange-M.ACC.SG type-ACC.SG

co jsme ho potka-l-i v Řím-ě?
 REL AUX.PST.1PL lui.ACC rencontrer-PTCP-PL.M à Rome-LOC.SG

b) *Pamatuj-eš si na toho divn-ýho chlápka,*
 rappeler-PST-2SG REFL à ten.M.ACC.SG étrange-M.ACC.SG type-ACC.SG

co jsme ho potka-l-i v Řím-ě?
REL AUX.PST.1PL lui.ACC rencontrer-PTCP-PL.M à Rome-LOC.SG

Tu te rappelles le/ce type étrange qu'on a croisé à Rome ?

10)

a) *Po kin-ě se ještě bavi-l-i o tom film-u.*
après cinéma-LOC.SG REFL encore discuter-PST-3PL.M de ten.M.LOC.SG film-LOC.SG

b) *Po kin-ě se ještě bavi-l-i o film-u.*
après cinéma-LOC.SG REFL encore discuter-PST-3PL.M de film-LOC.SG
(Czardybon, 2017 : 90)

11)

a) *Když jsem veše-l do toho krám-u, tak se mě*
quand AUX.PST.1SG entrer-PTCP-SG.M dans ten.M.GEN.SG magasin-GEN.SG alors REFL moi.ACC
prodavačk-a hned pta-l-a, co chc-i.
vendeuse-NOM.SG tout de suite demander-PST-3SG.F que.ACC vouloir-PRS-1SG

b) *Když jsem veše-l do toho krám-u, tak se mě*
quand AUX.PST.1SG entrer-PTCP-SG.M dans ten.M.GEN.SG magasin-GEN.SG alors REFL moi.ACC
ta prodavačk-a hned pta-l-a, co chc-i.
ten.F.NOM.SG vendeuse-NOM.SG tout de suite demander-PST-3SG.F que.ACC vouloir-PRS-1SG

Quand je suis entré dans le magasin, la vendeuse m'a tout de suite demandé ce que je cherchais.

12)

a) *Když jsem posledně nastoupi-l v Praz-e do autobus-u,*
quand AUX.PST.1SG la dernière fois monter-PTCP.SG.M à Prague-LOC.SG dans bus-GEN.SG
tak mě ten řidič ani ne-pozdravi-l.
alors moi.ACC ten.M.NOM.SG chauffeur même NEG-saluer-PST.3SG

b) *Když jsem posledně nastoupi-l v Praz-e do autobus-u,*
quand AUX.PST.1SG la dernière fois monter-PTCP.SG.M à Prague-LOC.SG dans bus-GEN.SG
tak mě řidič ani ne-pozdravi-l.
alors moi.ACC chauffeur même NEG-saluer-PST.3SG

La dernière fois que je suis monté dans un bus à Prague, le chauffeur m'a même pas dit bonjour.

13)

a) *Petr si koupi-l nov-ý aut-o.*
NP.M.NOM.SG REFL acheter-PST.3SG.M nouveau-N.ACC.SG voiture-ACC.SG
Hned druh-ej den se mu rozbi-l
tout de suite deuxième-M.ACC.SG jour.ACC.SG REFL lui.DAT tomber en panne-PST.3SG.M

ten *motor.*
ten.M.NOM.SG moteur.NOM.SG

- b) *Petr* *si* *koupi-l* *nov-ý* *aut-o.*
NP.M.NOM.SG REFL acheter-PST.3SG.M nouveau-N.ACC.SG voiture-ACC.SG

Hned *druh-ej* *den* *se* *mu* *rozbi-l*
tout de suite deuxième-M.ACC.SG jour.ACC.SG REFL lui.DAT tomber en panne-PST.3SG.M

motor.
moteur.NOM.SG

Petr s'est acheté une nouvelle voiture. Dès le lendemain, le moteur est tombé en panne.

14) Les phrases a) et b) enchaînent directement sur l'extrait citée ci-dessous :

L1: nahoře nad tím byl-y tak obrovitánsk-ý pavučin-y
en haut au-dessus cela.INS être.PST-3PL.F si géant-F.NOM.PL toile d'araignée-NOM.PL

L2: jako že se tam rok-y ne-vysáva-l-o jo ?
MOD REFL là-bas an-ACC.PL NEG-passer l'aspirateur-PST-3SG.N MOD

- a) *ne on ten týpek říká-l že ti*
non lui.NOM ten.M.NOM.SG gars.NOM.SG dire-PST.3SG.M que ten.M.NOM.PL

pavou-ci to uděla-l-i za hodin-u takov-ou
araignée-NOM.PL cela.ACC faire-PST-3PL.M en heure-ACC.SG tel-F.ACC.SG

obrovsk-ou pavučin-u
géant-F.ACC.SG toile d'araignée-F.ACC.SG

- b) *ne on ten týpek říká-l že*
non lui.NOM ten.M.NOM.SG gars.NOM.SG dire-PST.3SG.M que

pavou-ci to uděla-l-i za hodin-u takov-ou
araignée-NOM.PL cela.ACC faire-PST-3PL.M en heure-ACC.SG tel-F.ACC.SG

obrovsk-ou pavučin-u
géant-F.ACC.SG toile d'araignée-F.ACC.SG

L1 : là-haut, juste au-dessus, il y avait des toiles tellement géantes...

L2 : genre on y avait pas passé l'aspirateur pendant des années, c'est ça ?

L1 : non, le gars, lui, il a dit que les araignées avaient fait ça en l'espace d'une heure, une toile géante comme ça

ORTOFON v1

15)

- a) *Byl tu ten kluk, co s nim*
être.PST.3SG.M ici ten.M.NOM.SG garçon.NOM.SG REL avec lui.INS

chod-í naše Věr-a.
sortir-PRS.3SG nôtre.F.NOM.SG NP-F.NOM.SG

- b) *Byl tu kluk, co s nim*
 être.PST.3SG.M ici garçon.NOM.SG REL avec lui.INS
- chod-í naše Věr-a.*
 sortir-PRS.3SG nôtre.F.NOM.SG NP-F.NOM.SG

Il y a le gars que notre Věra sort avec lui qui est passé.

16)

- a) *ale on samozřejmě m-á potom psychick-ej*
 mais lui.NOM évidemment avoir-PRS.3SG après mental-M.ACC.SG
- problém že jako se znova přípravu-je kdyby*
 problème.ACC.SG that MOD REFL de nouveau préparer-PRS.3SG si.CONJ+AUX.COND.3SG
- byl náhod-ou dalš-í hlad*
 être. PTCP.SG.M par hasard autre-M.NOM.SG famine.NOM.SG
- b) *ale on samozřejmě m-á potom ten psychick-ej*
 mais lui.NOM évidemment avoir-PRS.3SG après ten.M.ACC.SG mental-M.ACC.SG
- problém že jako se znova přípravu-je kdyby*
 problème.ACC.SG that MOD REFL de nouveau préparer-PRS.3SG si.CONJ+AUX.COND.3SG
- byl náhod-ou dalš-í hlad*
 être. PTCP.SG.M par hasard autre-M.NOM.SG famine.NOM.SG

mais après, lui, le problème mental [qu'il a], c'est qu'il continuer à se préparer au cas où il y aurait une autre famine
 ORTOFON v1

17)

- a) *jo a todleto jsi vidě-l-a ? .. jak na*
 MOD cela.ACC AUX.PST.2SG voir-PTCP-SG.F comment sur
- tý nej-rušnějš-í ulic-i v Los Angeles*
 ten.F.LOC.SG SUP-plus animé-F.LOC.SG avenue-LOC.SG à Los Angeles.LOC.SG
- jsou takov-ý-dle upoutávk-y*
 être.PRS.3PL tel-F.NOM.PL-SUFFIXE DEICTIQUE panneau-réclame-NOM.PL
- b) *jo a todleto jsi vidě-l-a ? .. jak na*
 MOD cela.ACC AUX.PST.2SG voir-PTCP-SG.F comment sur
- nej-rušnějš-í ulic-i v Los Angeles*
 ten.F.LOC.SG SUP-plus animé-F.LOC.SG avenue-LOC.SG à Los Angeles.LOC.SG
- jsou takov-ý-dle upoutávk-y*
 être.PRS.3PL tel-F.NOM.PL-SUFFIXE DEICTIQUE panneau-réclame-NOM.PL

et ça, tu l'as vu ? ces espèces de panneaux-réclames qu'il y a dans l'avenue la plus animée de Los Angeles ?
 ORTOFON v1

18)

- a) *U nás na krám-ě m-á-me jenom*
 chez nous.GEN sur magasin-LOC.SG avoir-PRS-1PL uniquement
to nej-kvalitnějš-í zbož-í.
 ten.N.ACC.SG SUP-de meilleure qualité-N.ACC.SG marchandise-ACC.SG

- b) *U nás na krám-ě m-á-me jenom*
 chez nous.GEN sur magasin-LOC.SG avoir-PRS-1PL uniquement
nej-kvalitnějš-í zbož-í.
 SUP-de meilleure qualité-N.ACC.SG marchandise-ACC.SG

Chez nous, dans notre magasin, nous n'avons que la meilleure marchandise.

(exemple inspiré par Zubatý, 1920)

19)

- a) *Wilshire Boulevard je*
 Wilshire Boulevard.NOM.SG être.PRS.3SG
nej-rušnějš-í ulic-e v Los Angeles.
 SUP-plus animé-F.NOM.SG avenue-NOM.SG à Los Angeles-LOC.SG

- b) *Wilshire Boulevard je*
 Wilshire Boulevard.NOM.SG être.PRS.3SG
ta nej-rušnějš-í ulic-e v Los Angeles.
 ten.F.NOM.SG SUP-plus animé-F.NOM.SG avenue-NOM.SG à Los Angeles-LOC.SG

Wilshire Boulevard est l'avenue la plus animée de Los Angeles.

20)

- a) *A nejvýznamnějšími epochami v dějinách jsou v důsledku toho ty epochy, které se dovedou opřít o střední vrstvy a položit do jejich rukou své ideály.*
 b) *A nejvýznamnějšími epochami v dějinách jsou v důsledku toho epochy, které se dovedou opřít o střední vrstvy a položit do jejich rukou své ideály.*

En conséquence de quoi les plus grandes époques de l'histoire sont les époques qui savent s'appuyer sur les classes moyennes et leur remettre leurs idéaux en main propre [...]

V. Havel, *La fête en plein air*, 1963, trad. par F. Kérel et modifiée par nous, 1964, InterCorp v10

21)

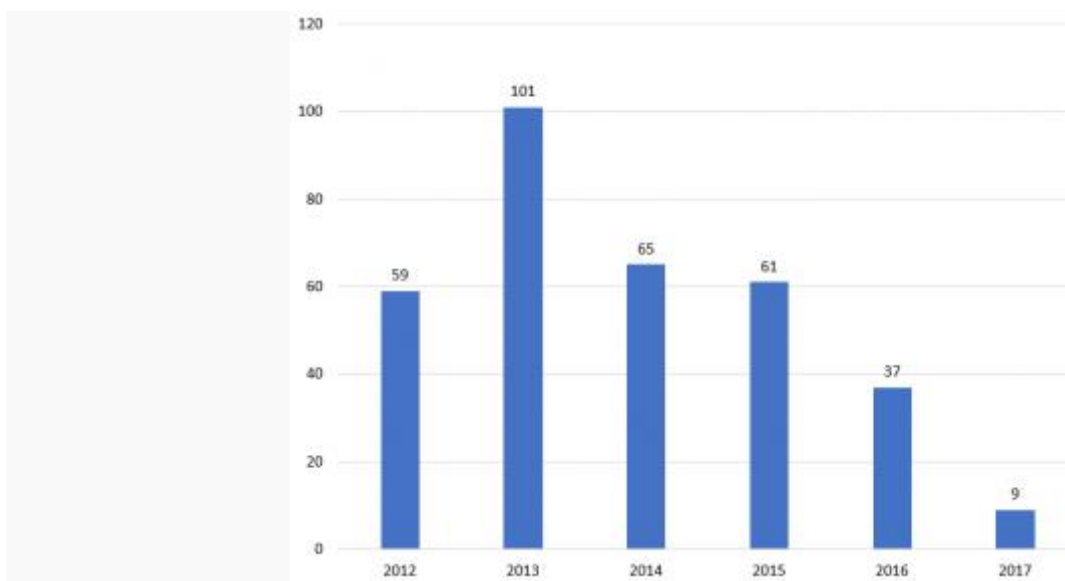
- tam v těch zem-ích kde už je eur-o*
 là-bas dans ten.F.LOC.PL pays-LOC.PL où déjà être.PRS.3SG euro-NOM.SG
tam v zem-ích kde už je eur-o
 là-bas dans pays-LOC.PL où déjà être.PRS.3SG euro-NOM.SG

là-bas, dans les pays où ils ont déjà l'euro...

ORTOFON v1

Structure du corpus ORTOFON v1 (traduction¹)

Tab. 1 : Nombre d'enregistrements par année d'enregistrement



Tab. 2 : Nombre d'enregistrements par nombre de locuteurs

Nombre de locuteurs	Nombre d'enregistrements
2	270
3	43
4	11
5	7

Tab. 3 : Nombre d'enregistrements par nombre de générations

Nombre de générations	Nombre d'enregistrements	Explications
1	285	personne de même âge
2	46	par ex. mère et fils
3	1	par ex. grand-mère, père, fille

Tab. 4 : Nombre de locuteurs et de mots par niveau de formation atteint

Formation		Nombre de locuteurs	Nombre de mots
B (élémentaire)	total	340	598 243
	école primaire	31	54 596
	formation d'apprentissage	73	123 334
	études secondaires	236	420 313

¹ https://wiki.korpus.cz/doku.php/cnk:struktura_ortofon.

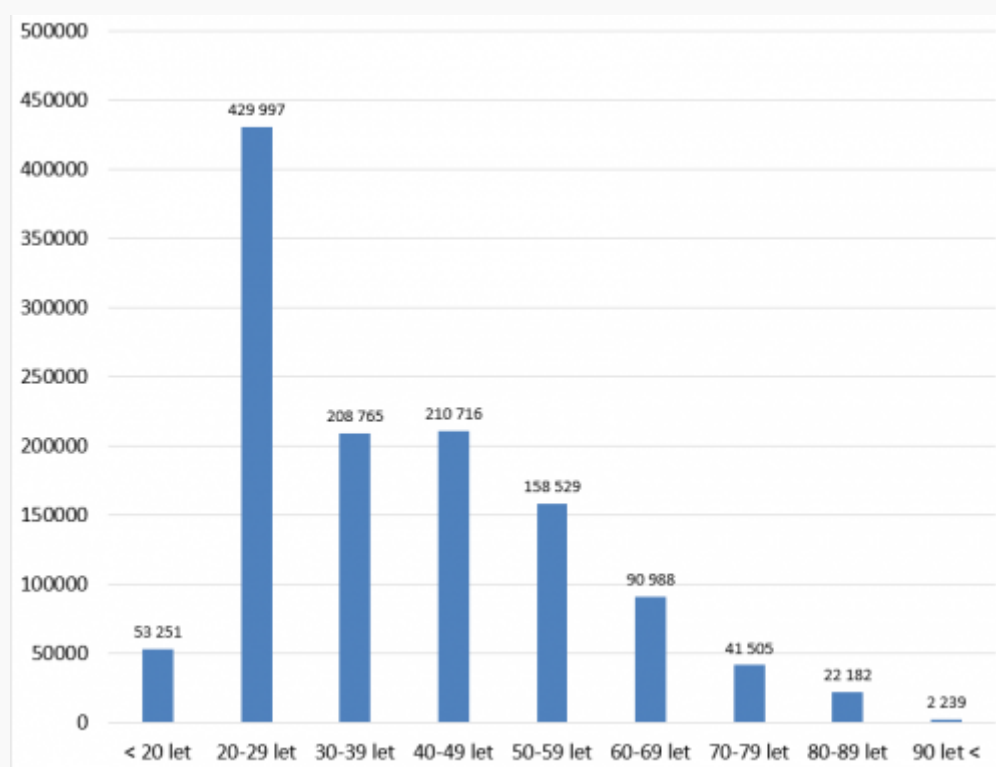
Formation		Nombre de locuteurs	Nombre de mots
A (supérieure)	études supérieures	284	619 929

Tab. 5 : Nombre de locuteurs et de mots par sexe

Sexe	Nombre de locuteurs	Nombre de mots
F (femmes)	322	607 464
H (hommes)	302	610 708

Tab. 6 : Nombre de locuteurs et de mots par âge

Age	Nombre de locuteurs	Nombre de mots
I (18-34 ans)	314	605 200
V (35 ans et au-delà)	310	612 972



Nombre de mots par tranche d'âge

Tab. 7 : Nombre de locuteurs et de mots par région de résidence principale avant l'âge de 18 ans

Région de résidence principale	Nombre de locuteurs	Nombre de mots
Bohême-Moravie	53	122 081
Bohême du Sud	68	128 187
Région frontalière de Bohême	59	118 101
Région frontalière de Moravie	54	118 562
Bohême du Nord et de l'Est	62	128 475

Région de résidence principale	Nombre de locuteurs	Nombre de mots
Silésie	55	120 828
Bohême centrale	79	123 121
Moravie centrale	70	119 142
Moravie de l'Est	62	117 659
Bohême de l'Ouest	60	119 998
Etranger	2	2 018

Tab. 8 : Nombre de locuteurs et de mots par taille de la commune de résidence avant l'âge de 18 ans

Taille de la commune	Nombre de locuteurs	Nombre de mots
village	124	245 984
ville de moins de 10 000 habitants	66	162 806
ville de plus de 10 000 habitants	105	211 637
ville de plus de 50 000 habitants	49	106 891
ville de plus de 100 000 habitants	280	490 854

Table des matières

Introduction générale	13
Présentation du projet.....	15
Problématique, hypothèses et principaux axes du projet	15
Structure du travail.....	19
Partie I	19
Partie II.....	20
Partie III.....	20
Conclusion	24
Question de recherche	24
Méthodologie : une approche sémantico-référentielle contrastive basée sur un corpus exploité à l'aide du logiciel textométrique TXM.....	28
Modes de citation et de présentation des exemples	31
1 Le démonstratif : présentation générale.....	41
1.1 La deixis : un grand débat linguistique	41
1.1.1 La double « découverte » de la deixis et de l'énonciation	42
1.1.2 Les approches « localisantes » de la deixis	44
1.1.3 Les approches fondées sur l'accessibilité référentielle.....	48
1.1.4 Les approches fondées sur la « token-réflexivité ».....	54
1.2 Le démonstratif : singularités sémantico-référentielles	56
1.2.1 Le démonstratif : un <i>désignateur</i> contextuel.....	56
1.2.2 Le démonstratif : référence par contraste interne et prédication (re)classifiant le référent	61
1.2.3 Le démonstratif : un indexical opaque.....	63
1.3 Le démonstratif : un « primitif » du langage humain	68
1.3.1 Le démonstratif dans la variabilité typologique	69
1.3.2 La sémantique du démonstratif	73
1.3.3 La morphosyntaxe du démonstratif	75

1.3.4	Le démonstratif dans une perspective diachronique.....	77
1.4	Résumé.....	79
2	Présentation des corpus et de la méthodologie du travail avec les données.....	81
2.1	Présentation des sources de données utilisées et des informateurs sollicités	81
2.1.1	Types de textes, registres, genres... quels choix terminologiques ?.....	82
2.1.2	Les corpus de langue orale informelle : le CEFT et ORTOFON v1.....	85
2.1.3	Le corpus multilingue InterCorp	95
2.1.4	Les autres corpus et bases de données	97
2.1.5	Présentation des informateurs	101
2.1.6	Présentation des outils statistiques employés.....	105
2.2	Travail avec les corpus et annotation des données	108
2.2.1	Interroger les corpus : le CEFC et ORTOFON v1.....	109
2.2.2	Interroger les corpus : InterCorp v10.....	112
2.2.3	Annoter les données : catégories et valeurs d'annotation.....	113
2.2.4	L'annotation des données : échantillonnage et taille des échantillons	132
2.3	Résumé.....	134
3	Le démonstratif en tchèque.....	137
3.1	Le démonstratif tchèque en synchronie.....	137
3.1.1	Le système du démonstratif en tchèque : une diversité (à première vue) vertigineuse	138
3.1.2	Le démonstratif tchèque : comparaison des fréquences des lemmes dans différents types de textes.....	151
3.1.3	Quel modèle d'analyse syntaxique pour les démonstratifs adnominaux tchèques ?	158
3.1.4	Les emplois adnominaux et pronominaux : peut-on parler de polysémie/homonymie ?.....	166
3.2	La constitution du système en diachronie	170
3.3	La distribution des démonstratifs adnominaux tchèques dans les différents régimes d'indexicalité	175

3.3.1	Les emplois exophoriques situationnels, temporels et textuels.....	175
3.3.2	Les emplois anaphoriques	179
3.3.3	Les emplois cataphoriques	184
3.3.4	Les emplois exophoriques mémoriels.....	188
3.4	La fonction de substantivation de TEN adnominal.....	192
3.5	Les emplois « émotionnels »	197
3.5.1	Introduction : le démonstratif et la subjectivité par défaut	198
3.5.2	Les emplois « désignatifs » porteurs d'une subjectivité accrue	199
3.5.3	Les emplois « émotionnels » propres	201
3.6	Résumé.....	209
4	Le démonstratif en français.....	211
4.1	Le démonstratif français en synchronie	211
4.1.1	Remarques morphosyntaxiques et sémantiques nécessaires	212
4.1.2	Les démonstratifs adnominaux français sont des déterminants	216
4.1.3	Le système du démonstratif français et l'expression de la distance	219
4.1.4	Le déterminant démonstratif français et sa fréquence dans différents types de textes	222
4.2	Le démonstratif français en diachronie	227
4.3	Le déterminant démonstratif français : un article défini fort ?.....	233
4.4	L'émergence du SN en français.....	238
4.5	La distribution des déterminants démonstratifs français dans les différents régimes d'indexicalité	242
4.5.1	Les emplois exophoriques situationnels, temporels et textuels.....	242
4.5.2	Les emplois endophoriques.....	245
4.5.3	Les emplois exophoriques mémoriels.....	252
4.5.4	Les emplois « émotionnels ».....	255
4.6	Résumé.....	261
5	Le démonstratif adnominal : bilan quantitatif.....	267

5.1	Les difficultés de l'annotation	267
5.1.1	Endophore, exophore situationnelle ou exophore mémorielle ?	268
5.1.2	Deixis textuelle, deixis situationnelle ou emploi indéfini.....	271
5.1.3	Les différents types d'anaphore.....	273
5.1.4	Emploi anaphorique/mémoriel ou emploi « émotionnel » ?	276
5.1.5	Autres difficultés d'annotation	280
5.2	Bilan quantitatif	281
5.2.1	Constats généraux.....	282
5.2.2	Le Corpus d'Étude pour le français contemporain	289
5.2.2.1	Premiers constats	289
5.2.2.2	Position de l'occurrence de CE par rapport au verbe de la proposition	295
5.2.2.3	Les disfluences dans les occurrences de CE.....	307
5.2.3	ORTOFON v1	313
5.2.3.1	Premiers constats	313
5.2.3.2	La grammaticalisation de TEN adnominal	318
5.2.3.3	Les propriétés combinatoires du démonstratif adnominal avec d'autres mots fonctionnels	320
5.2.3.4	Position de l'occurrence du démonstratif adnominal par rapport au verbe de la proposition	327
5.2.3.5	Les disfluences dans les occurrences du démonstratif adnominal	341
5.2.4	Le sous-corpus des pièces de V. Havel (InterCorp v10).....	348
5.2.4.1	Premiers constats	348
5.2.4.2	La grammaticalisation de TEN adnominal	352
5.2.4.3	Les propriétés combinatoires du démonstratif adnominal avec d'autres mots fonctionnels	355
5.2.4.4	La transposition du démonstratif adnominal dans les traductions françaises	358
5.3	Les noms à sémantique temporelle : comparaison des deux langues	367
5.3.1	Les noms à sémantique temporelle en emploi endophorique	367
5.3.2	Les noms à sémantique temporelle en emploi déictique temporel.....	371
5.4	Résumé et conclusion.....	374
6	La grammaticalisation du démonstratif TEN en article défini	381
6.1	Grammaticalisation et émergence des articles définis : présentation générale	382

6.1.1	La grammaticalisation.....	382
6.1.2	Grammaticalisation, expressivité et subjectivité	386
6.1.3	Le processus de grammaticalisation des démonstratifs en articles définis	389
6.1.3.1	Constats généraux	390
6.1.3.2	Le rôle de la subjectivité et de l'intersubjectivité	392
6.1.3.3	La grammaticalisation du démonstratif en article défini : trois approches concurrentes ...	394
6.1.3.4	Définitude « pragmatique » et définitude « sémantique »	400
6.1.4	Le statut du lexème TEN en tchèque : démonstratif ou article ?	410
6.2	L'étude des données.....	419
6.2.1	Les contextes de définitude « pragmatique » pure	419
6.2.1.1	L'emploi exophorique situationnel	419
6.2.1.2	L'emploi anaphorique	426
6.2.1.3	L'emploi exophorique mémoriel	448
6.2.2	Les contextes à la charnière entre les définitudes « pragmatique » et « sémantique »	458
6.2.2.1	Les noms sortaux ou relationnels avec des compléments établissant l'unicité dénotative.	459
6.2.2.2	L'anaphore associative.....	467
6.2.2.3	Les « concepts individuels complexes »	478
6.3	Résumé et conclusion.....	486
7	Les emplois « émotionnels » du démonstratif TEN	495
7.1	Les emplois « émotionnels » : pour un modèle unifié de traitement	496
7.1.1	Les emplois « émotionnels » : introduction	496
7.1.2	Les emplois « émotionnels » et la théorie CTD	501
7.1.3	Vers un modèle d'explication sémantico-pragmatique des emplois « émotionnels » du démonstratif	504
7.1.3.1	La pragmatique gricéenne et ducrosienne : introduction.....	504
7.1.3.2	La pragmatique gricéenne et ducrosienne : application	509
7.2	Les emplois « émotionnels » : analyse des données	514
7.2.1	Les emplois « émotionnels » présentant un effet d'affectivité.....	516
7.2.2	Les emplois « émotionnels » présentant un effet de prise de distance	525

7.2.3 Les emplois « émotionnels » présentant un effet de rappel	533
7.3 Résumé et conclusion.....	547
Conclusion générale.....	555
Récapitulation	557
Évaluation de la question de recherche	572
Apports, limites et ouvertures du projet	576
La grammaticalisation de TEN adnominal : quelles perspectives ?.....	579
Bibliographie	590
Index des auteurs	618
Annexes.....	621
Cartes	621
Liste des abréviations	623
Liste des tableaux.....	624
Liste des figures	628
Questionnaires	631
Questionnaire n°1	631
Questionnaire n°2.....	633
Questionnaire n°3.....	635
Structure du corpus ORTOFON v1 (traduction).....	644