

Table des matière

RÉSUMÉ	III
ABSTRACT	IV
TABLE DES MATIÈRE	V
LISTE DES TABLEAUX	VIII
CONVENTIONS DE TRANSCRIPTION	IX
CONVENTIONS DE NOTATION.....	X
LISTE DES ABRÉVIATIONS	XI
DÉDICACE.....	XII
REMERCIEMENTS	XIII
1 INTRODUCTION	1
2 ÉTAT DE LA QUESTION ET APPROCHE.....	4
2.1 ALTERNANCE CODIQUE : ÉLÉMENTS DE DÉFINITION	4
2.1.1 <i>Code</i>	4
2.1.2 <i>Alternance</i>	6
2.2 DÉFINITIONS RETENUES ET IMPLICATIONS POUR L'ANALYSE.....	9
2.2.1 <i>Code, langue et altérité codique</i>	9
2.2.2 <i>Répertoire du locuteur et catégorisation de l'altérité codique</i>	12
2.2.3 <i>Récapitulatif des définitions retenues</i>	13
2.3 ALTERNANCE CODIQUE : ÉLÉMENTS D'EXPLICATION.....	13
2.4 AUTRES CONSIDÉRATIONS THÉORIQUES PERTINENTES	17
2.4.1 <i>Non-équivalence des synonymes et alternance codique</i>	17
2.4.2 <i>Paradigme lexical multicode</i>	18
2.4.3 <i>Ressources linguistiques</i>	20
2.4.4 <i>Sélection lexicale</i>	20
2.4.5 <i>Communauté de pratique</i>	21
2.4.6 <i>Sémantisation, habitudes d'usage et recontextualisation</i>	22
2.4.7 <i>Sens canonique, sens en usage et visée référentielle</i>	23
2.5 RÉCAPITULATIF DU CHAPITRE 2.....	25
3 LE CORPUS ANALYSÉ	27
3.1 CRITÈRES DE SÉLECTION DES CORPUS À ANALYSER	27
3.1.1 <i>Critères favorisant une quantité suffisante d'occurrences d'alternance codique</i>	27
3.1.1.1 Corpus existants et accessibles	27
3.1.1.2 Corpus de conversation libre ou semi-dirigée.....	27
3.1.1.3 Locuteurs bilingues.....	28
3.1.1.4 Enquêtes menées auprès des membres des communautés de langue minoritaire sur le plan régional	31

3.1.2	<i>Critères concernant l'analysabilité des occurrences</i>	32
3.1.2.1	Corpus sur support audio ou audiovisuel.....	32
3.1.2.2	Locuteurs bilingues anglais-français.....	32
3.1.2.3	Communautés en Amérique du Nord	33
3.1.3	<i>Exclusion des corpus d'apprenants</i>	33
3.1.4	<i>Vue d'ensemble : caractéristiques des corpus retenus</i>	35
3.2	CORPUS RETENUS	36
3.2.1	<i>Description générale</i>	36
3.2.2	<i>Contexte sociohistorique pertinent des communautés en question</i>	39
3.2.2.1	L'Alberta francophone.....	39
3.2.2.2	Le Maine francophone	43
3.3	CONSTITUTION DU CORPUS D'ANALYSE	48
3.4	RÉCAPITULATIF DU CHAPITRE 3.....	49
4	PRÉ-ANALYSES : REPÉRAGE D'OCCURRENCES, CATÉGORISATION ET TRANSCRIPTION	50
4.1	DÉFINITION PRATIQUE DE LA NOTION D'OCCURRENCE.....	50
4.1.1	<i>Identification des codes représentés</i>	50
4.1.2	<i>Juxtaposition des codes</i>	51
4.1.3	<i>La notion d'unité lexicale</i>	53
4.1.4	<i>Identification et exclusion des unités ambiguës quant à leur association à un code donné : la question des emprunts</i>	53
4.1.5	<i>Identification et exclusion des noms propres</i>	64
4.1.6	<i>Exclusion sur la base de la réalisation phonétique</i>	64
4.2	OCCURRENCES CIBLÉES POUR L'ANALYSE	65
4.2.1	<i>Les occurrences isolables</i>	65
4.2.2	<i>Les occurrences en anglais</i>	68
4.3	REPÉRAGE DES OCCURRENCES D'ALTERNANCE CODIQUE DANS LE CORPUS	68
4.4	CATÉGORISATION DES OCCURRENCES ISOLABLES REPÉRÉES	71
4.5	PROTOCOLE DE TRANSCRIPTION.....	72
4.6	RÉCAPITULATIF DU CHAPITRE 4.....	73
5	ANALYSES ET RÉSULTATS	75
5.1	COMPARAISON GLOBALE : LES CATÉGORIES FRÉQUENTES EN FRANÇAIS	75
5.2	CATÉGORIES FAISANT L'OBJET D'OBSERVATIONS GÉNÉRALES SEULEMENT	79
5.2.1	<i>Les chiffres</i>	79
5.2.2	<i>Les expressions exclamatives</i>	79
5.2.3	<i>Les formules</i>	80
5.2.4	<i>Les formes d'adresse</i>	80
5.2.5	<i>Les termes de relation familiale</i>	80
5.2.6	<i>Les marqueurs de position épistémique</i>	81
5.3	YEAH / YES / YEP.....	81
5.3.1	<i>Aperçu global et statut</i>	82
5.3.2	<i>Fréquence d'usage par locuteur</i>	83
5.3.3	<i>Tendances thématiques et cooccurrences</i>	87

5.3.4	<i>Sens en usage et fonctions discursives</i>	89
5.3.4.1	Établissement des catégories d'analyse.....	89
5.3.4.2	Résultats et discussion	95
5.4	<i>RIGHT / THAT'S RIGHT / IS THAT RIGHT?</i>	101
5.4.1	<i>Aperçu</i>	101
5.4.2	<i>Tendances thématiques et cooccurrences</i>	102
5.4.3	<i>Sens en usage et fonctions discursives</i>	103
5.5	LES MARQUEURS PRAGMATIQUES	109
5.5.1	<i>Aperçu</i>	110
5.5.2	<i>So</i>	112
5.5.3	<i>Anyway</i>	116
5.5.4	<i>Now</i>	118
5.6	DISCUSSION GÉNÉRALE	123
6	CONCLUSION	127
	RÉFÉRENCES	133

Liste des tableaux

TABLEAU 3.1 : CARACTÉRISTIQUES DES SOUS-CORPUS	38
TABLEAU 4.1 : OCCURRENCES D'ALTERNANCE CODIQUE DANS LE CORPUS D'ANALYSE PAR TYPE D'OCCURRENCE ..	69
TABLEAU 4.2 : LES CATÉGORIES FRÉQUENTES D'OCCURRENCES ISOLABLES DANS LE CORPUS D'ANALYSE	72
TABLEAU 5.1 : LES CATÉGORIES FRÉQUENTES DES OCCURRENCES ISOLABLES EN ANGLAIS ET EN FRANÇAIS,	76
TABLEAU 5.2 : LES MARQUEURS PRAGMATIQUES ISOLABLES EN FRANÇAIS ET EN ANGLAIS,	78
TABLEAU 5.3 : FRÉQUENCE D'OCCURRENCE DE <i>YEAH</i> , <i>YES</i> ET <i>YEP</i> PAR CORPUS,	82
TABLEAU 5.4 : FRÉQUENCE D'OCCURRENCE DE <i>YEAH</i> , <i>YES</i> ET <i>YEP</i> PAR LOCUTEUR,	84
TABLEAU 5.5 : FRÉQUENCE D'OCCURRENCE DE <i>YEAH</i> / <i>YES</i> / <i>YEP</i> ET <i>OUAIS</i> / <i>OUI</i> PAR LOCUTEUR,	87
TABLEAU 5.6 : FRÉQUENCES ABSOLUE ET RELATIVE DES OCCURRENCES DE <i>YEAH</i> ET DE <i>YES</i> PAR CATÉGORIE D'ANALYSE	96
TABLEAU 5.7 : FRÉQUENCES ABSOLUE ET RELATIVE DES OCCURRENCES DE <i>YEAH</i> ET DE <i>YES</i> PAR CATÉGORIE D'ANALYSE, CORPUS FAC SEULEMENT	96
TABLEAU 5.8 : FRÉQUENCES ABSOLUE ET RELATIVE DES OCCURRENCES DE <i>OUAIS</i> ET DE <i>OUI</i> PAR CATÉGORIE D'ANALYSE, CORPUS FAC SEULEMENT	98
TABLEAU 5.9 : FRÉQUENCES ABSOLUE ET RELATIVE DES OCCURRENCES DE LA CATÉGORIE <i>YEAH</i> / <i>YES</i> / <i>YEP</i> ET <i>OUAIS</i> / <i>OUI</i> , CORPUS FAC SEULEMENT	98
TABLEAU 5.10 : FRÉQUENCE DES OCCURRENCES ISOLABLES DE <i>RIGHT</i> , <i>THAT'S RIGHT</i> ET <i>IS THAT RIGHT?</i>	102
TABLEAU 5.11 : LES COOCCURRENCES FRÉQUENTES DE <i>RIGHT</i> / <i>THAT'S RIGHT</i> / <i>IS THAT RIGHT?</i>	102
TABLEAU 5.12 : FRÉQUENCE OBSERVÉE DES CATÉGORIES D'ANALYSE DE <i>RIGHT</i> , <i>THAT'S RIGHT</i> ET <i>IS THAT RIGHT?</i>	106
TABLEAU 5.13 : FRÉQUENCE DES MARQUEURS PRAGMATIQUES DE L'ANGLAIS ISOLABLES, ANALYSE INDIVIDUELLE,	111
TABLEAU 5.14 : FRÉQUENCE DE PRODUCTION DE <i>SO</i> , <i>ANYWAY</i> ET <i>NOW</i> PAR LOCUTEUR, TOUS LES CORPUS	111
TABLEAU 5.15 : LES OCCURRENCES ISOLABLES DE <i>SO</i> EN CONTEXTE	113
TABLEAU 5.16 : LES OCCURRENCES DE <i>ÇA FAIT (QUE)</i> EN CONTEXTE	114
TABLEAU 5.17 : CATÉGORIES D'ANALYSE OBSERVÉES POUR <i>SO</i> ET <i>ÇA FAIT (QUE)</i> , CORPUS FAC SEULEMENT	115
TABLEAU 5.18 : LES OCCURRENCES ISOLABLES DE <i>ANYWAY</i> EN CONTEXTE	117
TABLEAU 5.19 : CATÉGORIES D'ANALYSE OBSERVÉES POUR <i>ANYWAY</i>	118
TABLEAU 5.20 : LES OCCURRENCES ISOLABLES DE <i>NOW</i> EN CONTEXTE	119
TABLEAU 5.21 : LES OCCURRENCES DE <i>ASTHEURE (LÀ)</i> EN CONTEXTE	119

Conventions de transcription

<abc>	intervention d'un autre locuteur sans changement de tour de parole (ex. signal d'écoute; interruption ou chevauchement qui ne résulte pas dans un changement de tour de parole)
{abc}	chevauchement
[abc]	événement paralinguistique ou non linguistique qui influe sur l'interprétation de l'énoncé ou sur le déroulement de l'interaction (rire, toux, certains bruits de l'environnement, certains gestes, etc.)
(.)	pause (perceptuelle) ¹
ab--c	hésitation ou auto-interruption
:	prolongement d'un segment phonologique (voyelle ou consonne) (perceptuel) ^{1,2}
.	intonation déclarative (perceptuelle) ^{1,2}
?	intonation interrogative (perceptuelle) ^{1,2}
!	intonation emphatique (perceptuelle) ^{1,2}
,	séparation prosodique non finale (perceptuelle) ¹
...	le locuteur ne produit pas d'intonation terminale mais ne continue pas son énoncé (généralement avec <i>ou...</i> ou <i>so...</i>)
XXX	incompréhensible, un « X » par syllabe entendue
***	information supprimée afin de protéger l'anonymat des locuteurs (ex. nom propre). Au besoin, la première lettre d'une unité ainsi anonymisée a été conservée afin de mettre en évidence des récurrences de cette même unité.
[abc]	information contextuelle supplémentaire importante pour l'interprétation de l'exemple

Tout extrait des transcriptions est suivi d'une référence selon le modèle suivant : (FAC 01, 22:50). Cette référence mentionne, dans l'ordre, le corpus dont l'extrait provient, le numéro de fichier de l'enregistrement en question, et l'indice temporel du début de l'extrait (en format [hh:]mm:ss).

Notes :

(1) En lien avec l'approche et les objectifs de l'analyse, tous les phénomènes prosodiques marqués dans la transcription (pauses, prolongements vocaliques, intonations) sont relevés en fonction de la perception lors de l'écoute de l'enregistrement. Aucune prise de mesure de longueur de pause ou de prolongement ni de la pente mélodique des énoncés n'a été effectuée.

(2) Le marqueur d'intonation attribué à un segment de discours est déterminé selon l'intonation en conjonction avec la fonction du segment en question, sans considération de la forme syntaxique. Ainsi, un segment qui se termine par une intonation montante mais qui n'a pas la fonction d'interrogation ne sera pas indiqué par un point d'interrogation. En revanche, un segment qui se termine par une intonation montante et qui a la fonction d'interrogation mais qui n'est pas une question syntaxique sera quand même indiqué par un point d'interrogation.

Conventions de notation

Les italiques : Lorsqu'il est question de désigner une unité lexicale (par exemple, un terme ou une unité lexicale repérée dans le corpus) de façon métalinguistique, cette unité est représentée en caractères italiques.

N.B. Les italiques conservent également leurs fonctions d'accentuation et de mise en relief des expressions savantes extracodiques au français (par exemple, les locutions latines).

Les guillemets : Outre les citations directes, les guillemets sont employés lorsqu'il est question de citer une formulation en particulier, par exemple : un emploi effectif d'une unité lexicale dans le corpus, un titre de tableau ou une formulation terminologique qui ne relève pas d'une citation particulière mais est généralisée dans les travaux d'un auteur donné.

Liste des abréviations

AC : alternance codique

AN : animateur/animateur (locuteur/locutrice dans le corpus)

C : chercheur (locutrice dans le corpus)

CB : corpus Beaulieu

CB-IND : corpus Beaulieu, entrevues individuelles (résidents au foyer pour personnes âgées)

CB-GRP : corpus Beaulieu, entrevues en groupe (résidents au foyer pour personnes âgées)

CB-PRS : corpus Beaulieu, entrevues avec du personnel soignant, individuelles ou en groupe

EN : anglais

FAC : corpus du Franco-American Center

FR : français

INT : interviewé(e) (locuteur/locutrice dans le corpus)

PFC : corpus du projet de Phonologie du français contemporain

PLM : paradigme lexical multicodique

*For Patrick —
we are an effective team*

Remerciements

*To those gentle ones my memory runs, to the laughter we shared at the meals –
I filled their kitchens and living rooms with my schemes and my broken wheels.
It was never clear how far or near the gate to my citadel lay.
They were cutting from stone some dreams of their own, but they listened to mine anyway.*

(Jackson Browne, « Farther On »)

Ce mémoire est le fruit d'un projet de recherche en préparation – dans une forme ou une autre – depuis l'automne 2013. C'est à ce moment-là, alors que je commençais la dernière année de mon baccalauréat en sciences du langage et que je visais encore une maîtrise en orthophonie, que M. Courbon m'a encouragée de préparer une proposition de projet de recherche en linguistique afin de postuler pour des bourses, « juste pour voir ». And the rest, as they say, is history. Ces quatre ans et demi ont été pour moi une expérience extraordinaire. Je n'arriverai jamais à décrire en ces quelques lignes tout ce que cette expérience m'a apporté : tout ce que j'ai appris, sur mon sujet et sur moi-même; tout ce que j'ai acquis; tout ce que j'ai vécu; et surtout, tous les liens de mentorat et d'amitié qui ont pris racine et qui se sont approfondis en cours de route. Je reste toujours émerveillée par le fait que cette vie ainsi remplie soit réellement la mienne.

Arrivée à ce point dans mon parcours singulier, au terme de ce projet à la fois scientifique et personnel, je pense surtout aux personnes qui m'ont accompagnée et qui ont fait en sorte que cette recherche – et cette chercheuse – devienne ce qu'elle est aujourd'hui. Tout d'abord, à Bruno Courbon, qui a su dès le début m'encadrer tout en me permettant d'y arriver par moi-même. Il a été non seulement mon directeur de recherche mais bien mon mentor, veillant à ce que j'obtiens des opportunités d'expérience professionnelle et intellectuelle, tout en créant un environnement de travail chaleureux et humain, et en m'offrant des réflexions et des conseils qui me sont toujours précieux. Aussi à Suzie Beaulieu, dont la curiosité intellectuelle et l'enthousiasme sont contagieux et qui m'a elle aussi prise sous son aile et encouragée à avoir la même croyance qu'elle en mes capacités. Un énorme merci à vous deux pour toute votre générosité d'esprit, pour toute votre expertise, pour toutes les échanges passionnantes – et pour toute votre patience avec mes maintes erreurs de français! Merci également à M. Patrick Duffley, qui a gracieusement accepté d'évaluer mon projet de recherche ainsi que le présent mémoire, pour ses commentaires judicieux.

Thanks to my parents, who always knew something like this would happen. You made sure I knew that nothing I *really* wanted was out of my reach, and taught me that, sometimes, seemingly impossible problems can be solved, with determination and creativity (just how many grass seeds are in that bag, anyway?). To Grandmom, who paved the way. And to Grandma and Grandpa (miss you, Grandpa) who were always so proud of whatever I was up to now. A very special thank you to Miss Midna, who kept me company during all those late nights at my desk, and to Miss Pepper, who kept me entertained!

À mes camarades de bureau au 2252, Myriam Paquet-Gauthier, Rémi Belleau, Anthony Lemire, et Anne-Sophie Bergeron : vous avez rendu le bureau plein de vie et la vie pleine d'amitié. Merci de votre compagnie, de vos sourires, de votre écoute et de votre aide. Merci à Rémi pour la musique ambiante! Et tout particulièrement à Myriam, qui m'a accompagnée du début à la fin – merci d'avoir pris en charge une petite nouvelle encore au premier cycle et de lui aider à trouver sa place et sa confiance dans le grand monde des études supérieures. Et d'avoir dégagé du temps pour relire et réviser soigneusement tous les 140 pages de mon texte – c'est grâce à toi qu'il est aussi beau à lire! Et pour le ver d'oreille comme cerise à la fin.

À Marie Steffens, merci d'avoir été une collègue et une amie hors pair pendant les longs hivers québécois et les longues journées des JDL, et à Julie Rouaud, pour ton énergie et ta convivialité effervescentes. À vous deux, pour tous les moments chaleureux et bien amusants passés ensemble. To Liz Lebovic, the best virtual thesis partner a girl could ask for! I'm so glad I met you and so glad to have gone through this experience with you. Thank you. Merci à Anne-Marie Beaudoin-Bégin d'avoir été non seulement mon enseignante mais aussi mon cheerleader et mon amie. Merci également à Pierre-Luc Langevin et à François Tremblay pour le soutien moral pendant mes derniers mois à Québec.

To Andrea Spencer, thank you for keeping me from being lost at sea during my first teaching gig. And for introducing me to all your friends, who are now my friends, and for all those fantastic evenings at your and Max's place that are some of my favourite memories out of this whole thing. Thanks too to Stephanie Fraser and Aileen Ruane for hanging out with me and swapping teaching stories with me. I had a blast thanks to you!

À Elisabeth Barbeau – sans toi, je n'aurais jamais survécu mon baccalauréat. Thank you for your steady confidence in me, for making sure I went outside sometimes and did other things than study, for all the late night talks and all the delicious food! Merci tout spécialement d'avoir été l'inspiration pour ma question de recherche : c'était nos conversations, où on passait librement du français à l'anglais et back again, qui m'ont amenée à demander « mais pourquoi donc on fait ça? »

To Gabriel McKee : « as iron sharpens iron, so a friend sharpens a friend ». Thank you, my friend. You pushed me to be the best that I could be, kept me honest – and kept me caffeinated. I always looked forward to our chats over some beverage or other and I look forward to there being more in the future. I'm glad we were on this road together; I am the richer for it. Got time for coffee?

Et puis à Laurence Poulin. What can I say? You. Are. The. Best. Point. Thank you for always being there (même aux petites heures du matin), pour ta générosité inégalée (de ton temps, de ton écoute, de ton chez-vous, de ton café), for always knowing how to make me laugh and get me back to feeling like myself. I would never have made it to the finish line without you. Team Awesome Sauce Cyborg Mutants forever!

And finally, to Patrick. Who saw what I could not, believed when I did not. Who loved me when. Who made everything happen so that I could find my path, and who walked it alongside me. « Like the moonlight without the moon / I can't dance without you » (Alan Doyle, « I Can't Dance Without You »). Thank you.

* * * * *

Ce mémoire a bénéficié du soutien financier du Conseil de recherche en sciences humaines du Canada ainsi que le Fonds de recherche du Québec – Société et culture.

Je remercie Mme Suzie Beaulieu de m'avoir permis l'utilisation des enregistrements faits dans le cadre de sa thèse doctorale ainsi que Mme Susan Pinette et le Franco-American Center de l'Université du Maine, Orono de m'avoir permis l'utilisation de leurs archives vidéo. Merci aussi à tous les locuteurs qui ont participé aux enregistrements qui sont devenus mon corpus d'analyse. Sans eux, ce mémoire n'aurait pas pu se réaliser.

1 Introduction

Lors des premières études sur le bilinguisme, l'alternance codique était considérée comme un comportement marginal, soit une erreur d'acquisition ou de production, soit une phase de la transition vers une nouvelle langue dominante (Haugen, 1956 ; Weinreich, 1970 [1953]). Ce point de vue va de pair avec l'attitude prédominante au début du 20^e siècle envers le bilinguisme en général : un fléau social associé aux immigrants et autres populations « indésirables », pouvant mener à un retard dans le développement langagier ou psychosocial, à des maladies mentales et à des difficultés affectives (Hakuta, 1986 ; Haugen, 1972 [1970] ; Pavlenko, 2011a). Depuis ce temps, le bilinguisme est passé d'une « maladie » à un avantage, d'un problème social à un objet intrinsèquement digne d'étude. Le « redoragage » de l'image du bilinguisme a légitimé également l'intérêt pour ses « symptômes » ; ainsi, l'alternance codique est devenue un objet d'étude en soi, et l'on sait maintenant qu'elle n'est pas du tout marginale. Étant donné que le multilinguisme¹ est sans doute une situation plus commune que le monolinguisme à l'échelle mondiale², et étant donné qu'il y a de l'alternance codique, sous une forme ou une autre, partout où il y a du multilinguisme (voir par exemple Gardner-Chloros, 2009 pour une idée de l'étendue des communautés dans laquelle l'alternance codique a déjà été attestée), l'alternance codique devrait être considérée comme un comportement langagier incontournable.

De la prise de conscience de l'importance du phénomène découle la question : qu'est-ce qui fait que le phénomène d'alternance codique est si répandu ? Autrement dit, quels sont les processus et les motivations, apparemment inhérents au bilinguisme, qui rendent autant de locuteurs susceptibles de le produire ? Bref, pourquoi produit-on de l'alternance codique ? C'est à travers les réponses à cette question fournies dans la littérature que la question de recherche de la présente étude a pris forme.

L'étude de l'alternance codique par différentes disciplines de la linguistique a permis d'apporter différents éléments de réponse. Les études menées en syntaxe ont permis d'éliminer la possibilité d'un manque de compétence comme explication (par exemple, MacSwan, 2013 ; Poplack, 1980). L'approche de type psycholinguistique se concentre sur les processus cognitifs qui rendent possible, voire provoquent, l'alternance

¹ Le terme *multilinguisme* est employé avec plusieurs significations dans la littérature : pour désigner une communauté où plusieurs langues sont en contact, par opposition à *plurilinguisme* qui réfère au fait chez l'individu de parler plusieurs langues ; pour désigner le fait de parler plus de deux langues, par opposition à *bilinguisme* et à *monolinguisme* ; ou pour désigner toute situation, individuelle ou communautaire, où plus d'une langue est présente (Kemp, 2009). C'est cette dernière acception que j'entends dans la présente recherche. Également, quoique je privilégie le terme *bilinguisme* lorsque la discussion porte spécifiquement sur la présence de deux langues, à l'instar de Grosjean (1997 ; 2013), je n'entends pas exclure par ce terme la présence de plus que deux langues.

² Il est difficile d'avoir des données chiffrées sur le multilinguisme à l'échelle mondiale. Le recenser nécessiterait, entre autres, un consensus sur la définition de « langue » et de « parler une langue », et une connaissance avancée de toutes les langues parlées dans le monde. Cependant, on n'a qu'à penser aux régions fortement multilingues qui sont aussi densément peuplées, comme l'Afrique et l'Asie du Sud, ou à la prévalence des conditions donnant lieu au multilinguisme, comme l'immigration et la globalisation, pour se rendre compte de son importance (Grosjean, 2010 ; Lüdi et Py, 2009).

codique (par exemple, Abutalebi et Green, 2008 ; Heredia et Altarriba, 2001). L'approche de type sociolinguistique tient compte de la valeur identitaire et sociale des différents codes en tant que motivation pour alterner entre eux (par exemple, Blom et Gumperz, 2000 [1972] ; Heller, 1995). Dans la perspective de l'analyse conversationnelle, on étudie les effets que le locuteur peut souhaiter produire, au moyen de l'alternance codique, sur le déroulement de l'interaction (par exemple, Auer, 1995 ; Li, 1998).

La diversité et la complémentarité de ces éléments de réponse démontrent que l'alternance codique est un phénomène complexe dont une description juste ne sera possible que par la mise en commun d'une diversité de traditions empiriques et épistémologiques. Cependant, elle n'est qu'exceptionnellement abordée dans une perspective où l'on tient compte de ce que le locuteur exprime sur le plan sémantique au moyen de ses alternances : à ma connaissance, seuls Backus (2001) et King et Nadasdi (1999) le font de façon explicite. Or, exprimer des sens est une partie essentielle de l'activité langagière. En même temps, différentes langues fournissent différents moyens d'exprimer des sens, c'est-à-dire différentes lexicalisations, dont certains peuvent convenir plus ou moins à l'idée que le locuteur a à l'esprit, à sa visée d'expression. Même dans une seule langue, on reconnaît que la visée d'expression, c'est-à-dire la *visée référentielle* du locuteur – la réalité conceptuelle, telle qu'il se l'approprie et s'y oriente, qu'il souhaite communiquer à son interlocuteur au moyen des ressources linguistiques – puisse avoir une influence sur le choix de mots (par exemple, dans le cas de synonymes, cf. Courbon, 2012). Lorsqu'un locuteur compose avec plus qu'une langue, il est tout à fait concevable que sa visée référentielle puisse influencer aussi le choix de langue; Heredia et Altarriba (2001) l'admettent, par exemple, sans toutefois s'y attarder. La prise en compte de la dimension de l'expression sémantique s'avère alors nécessaire pour la description d'un tel phénomène.

La présente étude vise à contribuer à ce que la dimension de l'expression sémantique soit prise en compte dans l'étude de l'alternance codique. Ainsi, elle prend comme point de départ la question suivante : des facteurs relatifs à la visée référentielle du locuteur figurent-ils, de manière observable et analysable, parmi les raisons qui pourraient motiver la production de l'alternance codique? Cependant, comme il y a relativement peu de littérature existante sur ce sujet, il n'y a pas de démarche ou de méthode établie qui permettrait de répondre à une telle question. L'objectif premier de la présente étude est donc mettre sur pied une méthode d'analyse des occurrences d'alternance codique qui permettra de prendre en compte les possibles intentions expressives du locuteur. Cela implique d'établir un cadre d'analyse dans lequel le locuteur bilingue est considéré comme un usager pleinement impliqué dans l'emploi de ses ressources linguistiques particulières et d'analyser dans ce cadre des occurrences d'alternance codique observées dans un corpus convenable à cette tâche – en l'occurrence, des corpus de conversations semi-dirigées enregistrées auprès des locuteurs bilingues anglais-français de l'Alberta et du Maine. Ce faisant, la présente étude contribue non seulement à l'approfondissement

des connaissances théoriques, mais aussi à une meilleure prise de conscience de ce qu'est produire de l'alternance codique et, donc, de ce qu'est être bilingue.

2 État de la question et approche

2.1 Alternance codique : éléments de définition

La question au centre de la présente étude est essentiellement « pourquoi » : pourquoi — c'est-à-dire, pour quelles raisons — produit-on l'alternance codique? Plus particulièrement, la visée référentielle que permettent de traduire des unités lexicales de « l'autre langue » pourrait-elle figurer parmi ces raisons? Toutefois, l'analyse de ce « pourquoi » se fonde nécessairement sur le « quoi ». Alors qu'est-ce que l'alternance codique? Qu'est-ce que le code qui est alterné, et en quoi consiste l'alternance? En fait, il n'est pas aisé de répondre à ces questions. Comme Alvarez-Cáccamo (1998) le remarque, le terme *alternance codique* (« code-switching ») est conceptuellement riche et se prête facilement à différentes interprétations selon l'orientation avec laquelle on l'aborde. Or, ce phénomène est étudié par une diversité de disciplines de la linguistique et selon une diversité d'approches. Il y a donc effectivement une diversité d'interprétations, ou en d'autres mots, une absence de consensus quant à sa définition. Cependant, comme le soulignent Gafaranga et Torras (2002), ce manque de consensus n'est pas « simplement » un débat terminologique; la manière dont on conçoit *code*, *alternance* et donc *alternance codique* relève de questions fondamentales portant sur la langue, sur le bilinguisme et même sur le phénomène psycho-cognitif qu'est la perception.

2.1.1 Code

On peut classer les conceptualisations de la notion de *code* dans la littérature en trois « familles » (cf. Alvarez-Cáccamo, 1998), compte tenu du fait que ce ne sont pas des catégories formelles ni, en fin de compte, des définitions mutuellement exclusives. Une première conception de *code* est celle qu'on retrouve en cryptographie : un mécanisme de transduction qui permet d'établir des correspondances entre les éléments de deux ensembles (Alvarez-Cáccamo, 1998). Cette conception se retrouve aussi en linguistique, en particulier en phonologie structurale. On établit des correspondances entre des réalisations phonétiques et un phonème; c'est ce qui permet, par exemple, la compréhension des accents étrangers (Hockett, 1987 ; Jakobson, Fant, et Halle, 1963). Changer de code, dans cette conception, veut dire changer de « clé » d'interprétation, et n'implique pas nécessairement la production d'énoncés.

Une deuxième conception, la plus courante, est celle du code en tant que système de conventions partagées. C'est le plus souvent à partir de cette conception qu'on fait l'équivalence entre *code* et *langue* dans le contexte de l'alternance codique, où la langue est conçue comme un système objectif, distinct d'autres langues, plutôt homogène et plutôt stable (voir à ce sujet Milroy, 2001). *Alternance codique* réfère donc au fait de changer de système, c'est-à-dire de changer de langue. On retrouve cette conception surtout dans des approches structuralistes de la langue telles qu'on trouve en grammaire générative et en psycholinguistique (par exemple, Abutalebi et Green, 2008 ; Grosjean, 1997 ; MacSwan, 2013 ; Poplack, 1980).

Cette conception se heurte toutefois à certaines difficultés. Elle occulte le fait qu'une langue n'est pas une entité objectivement déterminée ayant une existence propre, mais avant tout un ensemble flexible de pratiques sociales et créatives qui varient avec le temps, selon les communautés et même selon les individus (Gardner-Chloros et Edwards, 2004 ; Heller, 2007 ; Kemp, 2009 ; Le Page et Tabouret-Keller, 1985 ; Lüdi et Py, 2009). De plus, par cette objectivation des langues, elle implique que l'individu bilingue dispose de deux systèmes, distincts et autonomes. Autrement dit, cette conception de *code* en tant que langue-système immanente implique que le bilinguisme est une duplication du monolinguisme, et que l'individu bilingue est, ou devrait être, l'équivalent de deux individus monolingues en un³ (Lüdi et Py, 2009). Cependant, il est très rare, voire même impossible, qu'un locuteur bilingue atteigne dans sa deuxième langue une compétence identique à celle d'un locuteur monolingue⁴ de cette langue (Abrahamsson et Hyltenstam, 2009). En même temps, le fait de connaître une deuxième langue a des effets observables sur l'ensemble des habitudes langagières de l'individu bilingue, y compris ses habitudes dans sa langue maternelle (Cook, 1992 ; Kroll *et al.*, 2015 ; Pavlenko, 2011b). Considérés dans leur ensemble, ces constats semblent indiquer que le bilinguisme n'est pas la simple addition de monolinguismes — ce n'est pas un deuxième monolinguisme que le bilingue acquiert, et ce qu'il acquiert ne fonctionne pas comme un module autonome, mais s'intègre à l'ensemble de ses connaissances linguistiques. L'idée de langue-système reflétée dans cette conception de *code* prédit donc un fonctionnement du langage chez le bilingue qui ne se voit toutefois pas confirmé dans les faits.

Selon une troisième conception, vue en sociolinguistique ainsi qu'en analyse conversationnelle, *code* s'applique à tout ensemble de comportements auquel est attribuée une signification par les membres d'une communauté donnée (voir par exemple Alvarez-Cáccamo, 1998 ; Auer, 1995 ; Blom et Gumperz, 2000 [1972] ; Heller, 1995 ; Li, 1998). Étant socialement significatif, un code est un mode de présentation de soi, permettant de signaler son appartenance à un groupe donné, d'indiquer comment on voit ou construit la situation de communication (ex. formelle, décontractée, etc.), de signaler le rôle qu'on adopte dans cette situation, de faire allusion à d'autres rôles ou d'autres situations, et ainsi de suite. *Code*, en ce sens, peut s'appliquer tant à une langue qu'à un dialecte, à une variété de langue, à un registre ou à un style, si l'on se restreint au plan linguistique (Alvarez-Cáccamo, 1990). *Alternance codique* réfère donc à l'emploi d'un autre ensemble significatif, ce qui représente un changement dans la façon dont on se présente aux autres. La perception et

³ Force est d'admettre que plusieurs auteurs travaillant avec cette deuxième conception de *code* se distancient quand même d'un modèle de double monolinguisme. On le voit dans le principe de complémentarité de Grosjean (1997), par exemple, et dans les travaux sur l'influence translinguistique (« cross-linguistic influence », parfois appelé *transfert*) (Cook, 1992 ; Jarvis et Pavlenko, 2008 ; Treffers-Daller, 2009). Cependant, il s'agit toujours dans ces cas d'une conception des langues en tant qu'entités objectives, immanentes et relativement homogènes.

⁴ Dans la littérature, on fait plutôt référence aux locuteurs natifs ou ayant une compétence (quasi-)native, mais en réalité, le fait d'être un locuteur natif est souvent confondu avec le fait d'être monolingue, comme démontre par exemple la sélection des groupes témoins qui exclut les locuteurs natifs non-monolingues (Abrahamsson et Hyltenstam, 2009 en est un exemple).

l'interprétation d'une occurrence d'alternance codique dépendent donc d'une connaissance des normes de la communauté quant à l'interaction, aux codes présents et aux valeurs qui y sont associées.

En mettant l'accent sur la nature sociale de la construction et de l'emploi des codes, cette conception rend mieux compte de la créativité et de la variété inhérentes à l'activité langagière, qui est menée par des individus singuliers⁵, membres de communautés particulières, avec des répertoires et des normes qui leur sont propres. De cette perspective, le bilinguisme n'est pas le simple ajout d'un système modulaire, mais un changement dans la constitution même du répertoire du locuteur, soit l'addition de ressources linguistiques⁶ (Lüdi et Py, 2009 ; voir aussi dans ce sens Heller, 2007 ; Lamarre, 2002).

Cela n'implique toutefois pas l'abandon du terme *langue*, ni l'abandon du terme *code* pour référer à une langue, mais plutôt la reconnaissance qu'une langue n'est une langue (et un code n'est un code) que parce que les membres d'une communauté donnée la désignent comme telle⁷. Ce qu'on appelle une langue est en fait un ensemble socialement construit de comportements linguistiques; elle n'a pas d'existence propre en dehors de ses locuteurs. Ce sont les locuteurs d'une langue qui établissent ses conventions et ses normes et qui engendrent l'évolution de celles-ci, et ce sont les locuteurs qui forment la représentation collective de leurs pratiques langagières comme étant une langue. Cela est clairement illustré par le fait qu'il n'y a pas de critère objectif permettant de trancher, dans l'ensemble des contextes qu'on pourrait avoir à considérer, entre une langue et une variété de langue (diatopique ou autre) ou un sociolecte (Kemp, 2009 ; Lewis, Simons, et Fennig, 2015). Comme l'explique Kemp (2009) : « People, including researchers, abstract this social construct, reify it, and understand a language as existing in fact, not just as utterances. The 'fact' is much easier to understand and refer to than the complexity of the reality » (p. 16). Autrement dit, une langue, ainsi qu'un code, est plutôt une catégorie en soi, une simplification (nécessaire) qu'on impose sur la réalité.

2.1.2 Alternance

Tout comme la notion de *code*, représentant la matière qui est manipulée, la notion d'*alternance*, représentant la manipulation qui en est faite, est conceptualisée différemment dans la littérature selon l'approche. La conceptualisation d'*alternance* en tant qu'un basculement tranché (« switching ») découle de la conceptualisation de *code* en tant que langue-système. Comme ces systèmes sont distincts, un locuteur devrait se servir soit de l'un, soit de l'autre, et ce, à tout moment. Même si l'énoncé produit est mixte, les systèmes qui

⁵ La notion de singularité et sa relation avec la notion de communauté et d'intersubjectivité seront discutées plus en détail à la section 2.4.5.

⁶ La notion de ressources linguistiques sera discutée plus en détail à la section 2.4.3

⁷ Il est à noter que cela implique également (1) qu'une langue est un type spécifique de code, avec un statut social et politique particulier, et (2) la possibilité de dissociation entre *langue* et *code* : un code peut être constitué de plus d'une langue (cf. le « bilingual mode » de Grosjean, 1997 ; ou le « bilingual medium » de Gafaranga et Torras, 2002), et en ce cas, une alternance de langues ne constituerait pas en fait une alternance de codes.

l'ont produit ne le sont pas. Plutôt, le locuteur basculerait effectivement entre systèmes, accédant d'abord au lexique et à la grammaire d'un système en délaissant l'autre pour formuler une partie de l'énoncé, et vice versa pour formuler une autre partie.

Dans cette conception, il devient important de poser les frontières entre l'alternance codique et d'autres phénomènes de contact linguistique qui lui ressemblent, en particulier l'emprunt et le transfert⁸ (voir par exemple Clyne, 2000 [1987] ; Haugen, 1956 ; Sankoff et Poplack, 1981 ; Weinreich, 1970 [1953]). Dans cette perspective, l'emprunt, c'est-à-dire l'intégration d'une unité du code B au code A, ne constituerait pas une alternance codique, parce que le locuteur ne change pas de système en synchronie. Plutôt, on utilise toujours le code A qui, au cours de son histoire⁹, a subi une influence du code B. Ainsi, si on peut démontrer qu'une unité est « intégrée » au code A au moment de son emploi, cet emploi ne constitue pas une alternance codique, peu importe l'association de l'unité au code B. Turpin (1998) donne l'exemple du syntagme *trois policewoman*, où *policewoman* est considéré un emprunt, car la morphologie du pluriel suit, à l'oral, le système français plutôt que le système anglais. Cependant, ce qui constitue une intégration effective demeure pour l'instant une question sans réponse claire (voir la section 4.1.4 pour une discussion de quelques critères possibles).

Tout comme pour l'alternance codique, le sujet du transfert connaît une prolifération de terminologie et de conceptualisations (Jarvis et Pavlenko, 2008 ; Treffers-Daller, 2009); pour en discuter, il s'avère donc nécessaire de faire une parenthèse terminologique. Le transfert a suscité des réactions négatives depuis l'Antiquité (Jarvis et Pavlenko, 2008), et même lorsque Weinreich (1970 [1953]) a présenté ce phénomène dans son travail pionnier (sous le terme *interférence*) comme étant une conséquence tout à fait compréhensible du bilinguisme, le jugement négatif a persisté. Pour Weinreich, l'interférence réfère à toute différence des normes monolingues et est causée par la réorganisation des éléments d'une langue-système due à la connaissance d'une autre. Toutefois, et tout comme son contemporain Haugen, il considère que le bilingue devait reproduire les normes monolingues. Donc, malgré la ressemblance de sa définition avec celle du concept moderne de *cross-linguistic influence* (voir ci-dessous), son application a été restreinte, de façon générale, à des influences inconscientes et négatives de la langue maternelle sur la langue seconde – autrement dit, un échec de la part d'un apprenant qui n'a pas veillé à garder ses deux langues bien séparées (« keeping the two languages apart », Haugen, 1956, p. 11). Cette conception demeure dans certains travaux sur l'alternance codique (voir par

⁸ Quoiqu'il ne soit pas possible d'étudier les définitions de l'alternance codique sans discuter de ces phénomènes « voisins », il importe de noter qu'autant l'emprunt que le transfert sont des domaines d'étude vastes et très actifs, dont une discussion complète dépasse le cadre du présent travail. Le transfert est discuté brièvement dans les prochaines pages, et une discussion plus détaillée relative à l'emprunt est présentée à la section 4.1.4.

⁹ Il est à noter que l'histoire évoquée ici peut référer à un très court laps de temps. Selon certains chercheurs, notamment Poplack et ses collègues, les emprunts peuvent même être ponctuels, en autant qu'ils fassent preuve d'intégration au code A (mais voir la section 4.1.4).

exemple Poplack, 1987, p. 72 : « speech errors which involve both languages, and which may properly be considered “interference” »). De nos jours, le terme *interférence* est généralement rejeté dans le domaine de l'acquisition des langues secondes, vu sa connotation négative, mais il est toujours utilisé en psycholinguistique, où il réfère aux effets de l'activation des connaissances dans plusieurs langues sur les processus cognitifs de production de la parole (Treffers-Daller, 2009). En acquisition des langues secondes, *transfert* est interchangeable avec *influence translinguistique* (« cross-linguistic influence »), qui réfère à l'influence qu'a la connaissance d'une langue donnée sur la connaissance et l'emploi de toute autre langue (Jarvis et Pavlenko, 2008 ; Treffers-Daller, 2009). Ce terme représente un élargissement conceptuel par rapport à l'interférence, car il englobe des influences négatives et positives, ponctuelles ou permanentes, dont le locuteur est inconscient ou dont il fait l'emploi stratégique, ouvertement détectables ou subtiles, opérant dans toutes les directions entre toutes les langues connues par le locuteur (Jarvis et Pavlenko, 2008 ; Lüdi et Py, 2009 ; Treffers-Daller, 2009). Avec cette conceptualisation, le transfert pourrait englober tout phénomène de contact des langues, y compris l'alternance codique. Dans le cadre de la présente recherche, cependant, j'utilise *transfert* dans le sens typiquement employé dans les études qui le distinguent de l'alternance codique : une influence inconsciente et « erronée » d'une langue sur l'autre, qui éloigne cette dernière de la norme monolingue à viser (versus la juxtaposition en discours de deux codes monolingues bien maîtrisés). Dans cette perspective, et dans une approche qui considère l'alternance codique comme un basculement entre deux codes distincts, le transfert ne constituerait pas un changement de système; on utilise toujours le code A, mais ce code a subi l'influence du code B chez un locuteur donné.

L'attention ainsi portée à la distinction entre les concepts d'*emprunt*, d'*alternance codique* et de *transfert* n'est pas sans raison. Effectivement, lorsqu'il est question de l'alternance codique, il est nécessaire de déterminer ce qui constitue et ne constitue pas un changement de code. Toutefois, les distinctions habituellement posées semblent relever plutôt des a priori de l'analyse (en particulier quant à l'existence de plusieurs systèmes de production objectivement distincts) que des phénomènes réellement distincts dans la réalité des locuteurs (Gardner-Chloros, 1995).

Cette conception de l'alternance codique comme un basculement nécessite aussi la modélisation d'un mécanisme d'interaction entre deux systèmes autrement autonomes et modulaires. Les travaux de Myers-Scotton (par exemple, 1995) tentent d'établir les critères d'une langue matrice (« matrix language ») et les règles d'insertion des éléments provenant d'une langue enchâssée (« embedded language »). Les travaux de Poplack (Poplack, 1980 ; Sankoff et Poplack, 1981) proposent un modèle où l'interaction des deux systèmes constitue un troisième système, à partir duquel on produit l'alternance codique. Elle propose également des contraintes universelles sur le lieu d'occurrence d'une alternance particulière dans la phrase, soit la contrainte d'équivalence en ordre de mots entre les deux langues, et la contrainte contre l'alternance des morphèmes liés (Poplack,

1980). En effet, beaucoup d'études ont été réalisées dans le but d'établir de telles règles universelles (voir Clyne, 2000 [1987] ; MacSwan, 2013 pour un résumé), mais les résultats ne semblent pas aboutir à un consensus permettant d'affirmer leur universalité (Alvarez-Cáccamo, 1998 ; Gardner-Chloros et Edwards, 2004).

Cependant, il existe une autre façon de concevoir le phénomène perçu comme de l'alternance codique : le locuteur bilingue ne disposerait pas de deux systèmes, mais bien d'un seul : son répertoire de ressources linguistiques (Cook, 1992 ; Franceschini, 1998 ; Lüdi et Py, 2009). Sachant que des ressources données sont plus appropriées dans certains contextes que dans d'autres, le locuteur peut choisir de puiser ou non dans ces ressources selon le contexte (Franceschini, 1998 ; Gardner-Chloros, 1995). On retrouve ici le lien entre code et normes évoqué dans la troisième conception de *code* décrite à la section 2.1.2 : ce qui est considéré comme un code serait une « zone de focus », c'est-à-dire un ensemble de ressources dans le répertoire, plutôt qu'un module autonome; la perception de cet ensemble comme étant un code distinct dépend du statut que lui attribue la communauté et l'emploi qu'on peut ou ne peut pas en faire est déterminé par la situation de communication (Franceschini, 1998). Ainsi, l'alternance codique n'est pas un comportement particulier aux bilingues¹⁰, mais relève des forces fondamentales à l'activité langagière : « variability in [language] use » et « flexibility in behavior » (Franceschini, 1998 p. 52). Comme l'alternance codique bilingue (c'est-à-dire, entre deux codes reconnus comme des langues différentes) est un phénomène très répandu parmi les locuteurs du monde, un modèle qui la relie à un aspect fondamental du langage semble être le plus logique.

2.2 Définitions retenues et implications pour l'analyse

2.2.1 *Code, langue et altérité codique*

Il est clair que la perspective dans laquelle on aborde les concepts de langue et de fonctionnement du langage, entre autres, a un impact sur la manière dont on identifie et analyse l'alternance codique. En effet, le simple fait de dire qu'il y a ou n'y a pas d'alternance codique dans un énoncé donné représente déjà une sorte d'analyse, une prise de position sur ces concepts. Cependant, il y a évidemment un noyau commun autour duquel toutes ces différentes perspectives se rassemblent, c'est-à-dire qu'il y a effectivement un phénomène qu'on remarque et qu'on tente d'analyser. Tout en reconnaissant que la perception n'est pas neutre mais reflète déjà, au minimum, un filtre attentionnel sinon une certaine conception préexistante de ce qu'on observe, il devient alors pertinent de prendre du recul par rapport à toutes ces théories et ces modèles et de revenir vers

¹⁰ Effectivement, dans la langue courant, on trouve plus souvent le terme *code-switch* employé pour faire référence à un changement de style ou de registre, notamment en lien avec l'expression ou la suppression de son identité ethnique (voir par exemple <https://www.npr.org/sections/codeswitch>). Toutefois, dans la présente étude, ce terme sera réservé pour des changements de langue.

le phénomène lui-même. Qu'est-ce qu'on perçoit dans les énoncés des individus bilingues qui suscite l'attention et invite à l'analyse?

Fondamentalement, il s'agit du même phénomène qui est à l'origine de toute perception : le contraste. On perçoit visuellement un objet au moyen du contraste entre figure et fond, c'est-à-dire par le fait qu'il y a une différence entre les deux, par exemple de couleur ou de texture, qui nous permet de tracer les contours de l'objet. L'identification autant de la figure que du fond se fait en raison de l'altérité de chacun par rapport à l'autre, lorsque juxtaposés. Dans le cas de l'alternance codique, on perçoit la juxtaposition de différents codes linguistiques, ou plutôt d'éléments linguistiques (unités lexicales, phonèmes, etc.) qui fonctionnent comme des représentants de ces différents codes. Le phénomène d'alternance codique repose donc sur la perception de l'altérité d'un code par rapport à l'autre, lorsque juxtaposés. Cela renvoie à la conception de *code* en tant qu'ensemble socialement significatif : on perçoit cette altérité parce qu'on reconnaît les ensembles représentés comme des codes, auxquels on attribue des valeurs différentes, en particulier la valeur de « langue ».

Puisque la notion de *langue*, qui est d'une importance capitale pour l'alternance codique, est une catégorie imposée sur la réalité de l'activité langagière humaine, les critères définitoires varient selon les objectifs de la personne qui les impose. De ce fait, comme le souligne Auer (1995), ainsi que Gafaranga et Torras (2002), il importe de distinguer entre des phénomènes de contact des langues tels que repérés et interprétés par l'analyste et un phénomène tel que produit et perçu par les locuteurs (voir aussi à ce sujet Kerbrat-Orecchioni, 2009). Ainsi, j'utilise le terme *alternance codique* pour référer à une juxtaposition des codes, c'est-à-dire un « changement de langue », dans laquelle les membres de la communauté en question perçoivent effectivement une juxtaposition de codes et reconnaissent ces codes comme étant des langues. On peut également décrire ce phénomène comme la présence d'altérité codique¹¹. Cette conceptualisation de l'alternance codique diffère de la conceptualisation la plus courante, au point où l'emploi même du terme pourrait être remis en question. En effet, l'approche que j'adopte dans la présente étude cadre plutôt avec le concept de *translanguaging* (Otheguy, García, et Reid, 2015). Cependant, comme le terme *translanguaging* n'est pas encore bien connu en français, j'ai trouvé préférable de continuer à utiliser le terme *alternance codique*, déjà établi dans la littérature pour parler de ce genre de phénomène et plus facilement reconnu par association avec le terme de l'anglais, *code-switching*, en précisant bien la conceptualisation qu'il représente.

De cette conceptualisation découle le fait que, pour la présente étude, la distinction entre « l'anglais » et « le français » devra être définie autant que possible selon les perceptions des locuteurs. Cependant, la

¹¹ Comme *langue* et *code* sont dissociables (cf. la section 2.1.1), il peut y avoir altérité codique, au sens large, à l'intérieur d'une même langue (par exemple, lors d'un changement de registre ou de variété régionale). Cependant, comme la visée de la présente recherche est l'alternance entre des codes socialement construits comme étant des langues différentes, l'altérité codique dont il est question ici est également celle qui existe entre des langues différentes.

perception est difficile à saisir en pratique. Elle n'est qu'en partie partagée, c'est-à-dire que quoique la perception qu'ont les locuteurs de l'appartenance d'une unité lexicale donnée à l'une ou l'autre langue peut être relativement stable à travers une communauté, elle peut aussi varier pour différents membres de cette communauté et selon différents contextes de communication, par exemple dans le cas des emprunts (cf. la section 4.1.4). De plus, la perception est forcément un phénomène interne à celui qui perçoit; en l'absence de méta-commentaires émis par les participants à l'interaction – qui sont peu fréquents en discours relativement au nombre d'unités que l'analyste peut avoir à classer – elle n'est pas directement observable pour un analyste externe, mais doit être inférée à partir des comportements linguistiques et paralinguistiques.

En même temps, il ne faut pas nier le rôle de l'intuition de l'analyste dans son analyse; il faudrait plutôt admettre qu'on ne peut jamais en faire totalement abstraction. Ainsi, contrairement à Gafaranga et Torras (2002), je ne considère pas qu'il faille exclure de la définition d'*alternance codique* tout cas remarqué par l'analyste, mais auquel les locuteurs participants ne réagissent pas. D'une part, quoiqu'il soit vrai que « analysts' own knowledge of the world, in this case, their ability to assign some stretches of talk to a particular language and certain others to a particular other language [...] need not correspond to the reality participants themselves orient to while talking » (Gafaranga et Torras, 2002, p. 11), il ne s'ensuit pas que l'analyste n'a aucun accès à la perception des locuteurs, surtout s'il est aussi membre de la communauté étudiée. D'autre part, et plus significativement, cela revient à nier la possibilité que les locuteurs perçoivent effectivement une alternance de codes, sans toutefois y réagir. L'alternance codique peut être tout à fait perceptible aux locuteurs sans violer les normes communautaires ou les normes de l'interaction en cours, et une alternance codique qui ne suscite pas de réaction n'indique pas nécessairement l'existence d'un code mixte. Dans le contexte nord-américain, notamment au Canada, l'anglais et le français sont des codes fortement investis de signification sociale. Donc, même si l'alternance codique est plus ou moins acceptée — c'est-à-dire, suscite plus ou moins de réaction — d'une communauté à l'autre (selon les normes, le taux de bilinguisme dans la communauté, le degré de prescriptivisme linguistique, etc.), les membres des communautés où ces codes sont présents sont portés à percevoir une distinction marquée entre les deux, et donc, à percevoir de l'altérité codique lorsque ces codes sont juxtaposés.

Il est vrai que si les locuteurs ne réagissent pas ouvertement, leur perception d'une alternance codique n'est pas démontrable, ce qui laisse l'identification du phénomène à l'intuition de l'analyste. Ce dernier se doit dans ce cas de s'interroger attentivement sur la justesse de son intuition. Cependant, si, comme dans le cas de la présente recherche, l'objectif de l'analyse est d'étudier l'emploi que font les locuteurs des ressources linguistiques à leur disposition, il convient mal d'exclure « whatever else participants may know as community members » (Gafaranga et Torras, 2002, p. 11) afin de se concentrer uniquement sur ce qui provoque une tentative de réparation quelconque (voir aussi à cet égard Kerbrat-Orecchioni, 2009).

À la lumière de l'ensemble de ces constats, je considère qu'il est justifiable de se fier, lorsque nécessaire, à son intuition dans l'identification de l'alternance codique, en autant qu'on tente d'abord de rapprocher cette intuition de celle des locuteurs à l'étude — par exemple, en basant l'analyse le plus souvent possible sur l'usage observable — tout en reconnaissant qu'on ne peut jamais récupérer complètement l'intuition des locuteurs spécifiques dans un contexte spécifique.

2.2.2 Répertoire du locuteur et catégorisation de l'altérité codique

Suivant Franceschini (1998) et Gardner-Chloros (1995), je considère que l'alternance codique ne relève pas du basculement entre systèmes distincts, mais plutôt du recours à un ensemble multicodeur de ressources dans le répertoire du locuteur. Si le critère de « changement de système » au moment de la production d'une alternance codique n'est plus applicable, cela implique que la distinction traditionnelle entre l'alternance codique et l'emprunt, ainsi que le transfert, ne relève pas d'une différence essentielle phénoménologique, mais bien de catégories d'analyse. Autrement dit, ces phénomènes de contact des langues souvent considérés distincts par les linguistes ne le sont pas nécessairement pour les locuteurs. Dans une analyse qui vise à tenir compte de l'alternance codique telle que perçue par les locuteurs, il ne convient évidemment pas d'imposer des distinctions que les locuteurs eux-mêmes n'imposent pas. Il semble cependant y avoir, au moins dans certains cas, une véritable différence perçue entre ces phénomènes qui pourrait justifier des catégories d'analyse différentes. Par exemple, quoiqu'il soit très probable qu'un locuteur perçoive le *trois policewoman* de Turpin (1998, cf. la section 2.1.2) comme manifestant de l'altérité codique, presque aucun locuteur anglophone ne percevrait d'altérité dans *attorney general*¹², même si cette unité ne se conforme pas aux normes du placement de l'adjectif en anglais (Poplack, 1987).

S'il n'y a pas de frontière nette entre l'alternance codique et l'emprunt, on pourrait plutôt y voir une sorte de continuum, selon le degré d'altérité codique perçue par les locuteurs. Une unité qui ne suscite aucune perception d'altérité représente le degré minimal d'altérité perçue; on pourrait inclure dans cette catégorie les emprunts historiques bien établis comme *attorney general*. L'alternance codique « classique », clairement identifiable et délimitée (« unambiguous code-switches », Poplack, 1987 ; Turpin, 1998), représente le degré maximal d'altérité perçue. Entre ces deux extrêmes, l'altérité perçue est plutôt gradable que catégorielle, c'est-à-dire qu'une occurrence comportant de l'altérité linguistique peut être plus ou moins « autre » en fonction de différents facteurs, comme l'intégration en usage, les connaissances métalinguistiques des locuteurs en présence et leurs attitudes envers l'emprunt. De là découle la difficulté qu'éprouve l'analyste de catégoriser l'altérité, car il faut imposer des catégories d'analyse sur une réalité qui ne présente pas de divisions nettes et

¹² Cette unité lexicale, de l'ancien français *atourné (général)* (Oxford English Dictionary online, Anon, 2016), est souvent citée comme un exemple de l'influence du français sur le langage juridique anglais, justement en raison de sa préservation de l'adjectif postposé (Tiersma, 2008).

évidentes. Compte tenu de ce continuum d'altérité, il ne convient pas d'exclure a priori une unité de l'analyse parce qu'elle n'est pas maximale distincte du code auquel elle est juxtaposée (c'est-à-dire, qui pourrait représenter, selon l'approche, un emprunt plutôt qu'une alternance codique). Dans les deux cas, il s'agit d'une altérité codique, plus ou moins forte, perçue et exploitée par les locuteurs. Cependant, toute recherche requiert la délimitation de ses objets d'étude et donc l'imposition de catégories d'analyse. Pour la présente étude, il faudrait idéalement imposer une distinction entre les emplois qui constituent un changement de code et ceux qui n'en constituent pas dans l'esprit des locuteurs. Comme les perceptions des locuteurs ne sont pas directement observables, par contre, il faut se fier à des indices indirects du statut codique d'une unité donnée. Cette question est discutée plus en détail à la section 4.1.4.

2.2.3 Récapitulatif des définitions retenues

En résumé, le principe directeur de la présente étude est de se rapprocher le plus possible des perceptions des locuteurs, autant de ce qu'expriment les ressources qu'ils emploient que du statut codique qu'ils leur associent. Pour ce faire, cette recherche se base sur une conceptualisation de la langue non pas comme un système immanent, mais comme un ensemble de comportements linguistiques socialement construit comme une entité. Cela implique d'identifier les occurrences d'alternance codique par la juxtaposition des langues *telles qu'identifiées par les locuteurs eux-mêmes*, sans nier complètement l'intuition de l'analyste dans l'identification, mais au moins en la mettant au deuxième plan. La présente étude se base également sur une conceptualisation de l'alternance non pas comme un basculement entre des modules distincts, conditionnés par l'existence de systèmes distincts, mais comme un recours à un ensemble multicode de ressources linguistiques, dont l'existence est conditionnée par l'exposition, dans l'usage, à plusieurs codes. Cela implique que la distinction des différents phénomènes de contact des langues ne se fonde pas sur l'intégration ou non des unités dans la langue-système (individuel ou collectif), mais selon la perception qu'ont les locuteurs du degré de contraste que ces unités évoquent par rapport à un code communicatif donné.

2.3 Alternance codique : éléments d'explication

Sans oublier les conséquences importantes de la définition de l'objet sur l'analyse, la question au cœur de la recherche proposée n'en est toutefois pas une de définition, mais de motivation : pourquoi produit-on de l'alternance codique? Encore là, les réponses dépendent, dans une large mesure, de la perspective à partir de laquelle on aborde ce phénomène. De façon très générale, on peut distinguer dans la littérature cinq grandes familles d'approches à l'étude de l'alternance codique : les approches syntaxiques, les approches psycholinguistiques, les approches sociolinguistiques, les approches conversationnelles, et les approches en linguistique appliquée.

Une partie importante des études ayant pour objet l'alternance codique sont réalisées dans une perspective syntaxique, tout particulièrement dans le cadre de la grammaire générative et universelle (voir par exemple Clyne, 2000 [1987] ; MacSwan, 2013 ; Muyksen, 1995 ; Myers-Scotton, 1995 ; Poplack, 1980 ; Sankoff et Poplack, 1981). Tel que mentionné à la section 2.1.2, ce domaine vise en général à déterminer quelles sont les contraintes universelles régissant l'alternance codique (MacSwan, 2013). Même si l'universalité de ces règles s'avère difficile à démontrer, ces études ont permis de faire état de l'alternance codique en tant que phénomène répondant à certaines règles et nécessitant un degré élevé de compétence. Cela permet d'éliminer « la paresse » et « la confusion » comme étant des explications suffisantes. Toutefois, puisque ces études syntaxiques cherchent surtout à confronter des théories grammaticales aux données contenant de l'alternance codique, elles ne tentent pas d'en expliquer les causes.

Dans une perspective psycholinguistique, plusieurs raisons sont suggérées pour expliquer le recours à l'alternance codique : combler un manque de connaissances dans une des deux langues, répondre à un problème de récupération en mémoire dans la « bonne » langue ou encore le degré de dominance (c'est-à-dire, de maîtrise et de fréquence d'usage) d'une langue relativement à l'autre (Heredia et Altarriba, 2001). Le concept de *déclenchement* (« triggering ») de Clyne (2000 [1987]) relève également d'une perspective psycholinguistique. Cet auteur propose que des éléments lexicaux ambigus, présents à la fois dans chacune des deux langues (tels que des homophones ou des noms propres) peuvent déclencher un changement de langue en activant un peu les deux langues en même temps. Dans cette perspective, il est effectivement souvent question du niveau d'activation ou d'inhibition d'un système linguistique ou de l'autre, et des processus de production de parole en temps réel. Ainsi, les raisons pour lesquelles on a recours à l'alternance codique se résument à des questions de capacité, de facilité et de rapidité d'accès. Cependant, le regard sur les processus met de côté le locuteur en tant qu'individu qui s'exprime dans une situation de communication donnée. L'alternance codique dans cette approche n'est pas considérée sous l'angle de ses fonctions communicatives, mais sous l'angle — incontournable, mais insuffisant en lui-même — du fonctionnement du cerveau d'une personne qui connaît deux langues.

Les approches sociolinguistiques, que je distingue ici des approches conversationnelles, attribuent des motivations pour l'alternance codique qui sont d'ordre « macro-situationnelle », relatives aux groupes et aux rôles sociaux. L'étude de Blom et Gumperz (2000 [1972]), souvent citée comme étant fondatrice en matière d'alternance codique, adopte une approche macro-situationnelle. Les auteurs associent les deux codes entre lesquels leurs participants alternent à différents rôles sociaux : le dialecte local représente la famille et la communauté locale, la langue standard représente tout ce qui possède un caractère officiel ou national. Les locuteurs alternent entre ces deux codes selon la situation de communication, afin de signaler laquelle de ces identités ils s'attribuent. Cette correspondance entre le code d'un groupe et ce que ce groupe symbolise est

également présente dans Heller (1995). Cette perspective a le mérite de tenir compte du fait que l'alternance codique sert à accomplir un but communicatif. Toutefois, en mettant l'accent sur le code utilisé pour véhiculer le message et sur les enjeux sociopolitiques qui régissent l'emploi de l'un ou l'autre code, les approches sociolinguistiques mettent au second plan la signification de l'alternance codique en soi ainsi que le contenu verbal du message. De plus, malgré l'objectif de ces approches de prendre en compte les locuteurs lors de l'analyse linguistique, elles peuvent faire abstraction de l'agentivité des locuteurs individuels au profit des circonstances qui conditionnent leurs comportements langagiers.¹³

À la différence de la perspective sociolinguistique macro-situationnelle, une approche conversationnelle attribue des motivations « micro-situationnelles », relatives à la situation de communication, aux buts communicatifs, aux tours de parole, etc. On s'intéresse donc plus à la valeur que prend chaque acte d'alternance dans l'interaction en cours qu'à la valeur identitaire associée à chaque code dans l'absolu. L'étude de l'alternance codique au moyen de l'analyse conversationnelle (ou analyse séquentielle) exemplifie cette approche. Selon Li (1998), « [b]ilingual speakers change from one language to another in conversation not because of some external value attached to those particular languages, but because the alternation itself signals to their co-participants how they wish their utterances to be interpreted on that particular occasion » (p. 161). Dans cette perspective, les locuteurs produisent de l'alternance codique pour accomplir certains buts de communication, par exemple attirer l'attention sur un énoncé ou une partie d'énoncé, annoncer un changement de sujet, inclure ou exclure certains participants, créer de la cohésion, etc. (Angermeyer, 2003 ; Li, 1998). Cependant, dans cette perspective, ce n'est en général ni le code utilisé ni le contenu verbal qui est significatif, mais le contraste entre les différents codes, c'est-à-dire le fait d'avoir produit une alternance, abstraction faite de sa direction. De plus, semblables aux approches sociolinguistiques, les approches conversationnelles mettent généralement l'accent sur les circonstances (internes à l'interaction, dans ce cas) qui conditionnent le comportement des locuteurs, ce qui peut laisser en arrière-plan l'agentivité des locuteurs dans l'emploi de leurs ressources linguistiques dans ces circonstances.

Finalement, les études portant sur l'alternance codique en linguistique appliquée ne sont pas axées sur une approche théorique en particulier, mais sont centrées sur des thèmes comme l'enseignement et l'apprentissage des langues, le développement de différentes compétences (sociale, pragmatique, etc.) en langue seconde ou des facteurs sociaux et psychologiques qui influent sur quand et comment on emploie ses

¹³ Bien sûr, on peut avancer – avec raison – qu'un choix de langue qui semble être libre à l'échelle immédiate ou locale demeure en fait contraint par plusieurs facteurs, si on le situe sur une échelle plus grande (enjeux de pouvoir et de statut, accès aux ressources [linguistiques et autres], idéologies, normes culturelles, etc.). Les travaux de Blommaert (ex. 2007 ; 2010) et de Heller (ex. 1992 ; 1995 ; 2011) présentent cet argument de façon très convaincante. Cependant, on ne peut pas de là affirmer que toute alternance codique qu'un locuteur produit est nécessairement *motivée* par des considérations sociopolitiques ou sociohistoriques, c'est-à-dire, que le sentiment de liberté à l'échelle immédiate n'est pas valide.

différentes langues. Elles relèvent d'une variété d'approches, notamment la psycholinguistique et l'analyse conversationnelle. Ces études considèrent l'emploi de l'alternance codique comme étant une stratégie communicative pour maintenir le déroulement de l'interaction et assurer la communication lorsque la compétence dans une des langues fait défaut, pour exprimer une identité culturelle ou son identité en tant que locuteur multicompétent (voir par exemple Canagarajah, 2011 ; Cook, 1992 ; Jarvis et Pavlenko, 2008 ; Liebscher et Dailey-O'Cain, 2005), ainsi que pour exprimer ses émotions ou contourner des tabous dans l'une ou l'autre des cultures associée à une langue donnée (Dewaele, 2010). Cependant, la possibilité que l'alternance codique soit une stratégie d'expression sur le plan *lexical*, non pas pour combler une lacune dans les connaissances, mais bien pour traduire une visée référentielle spécifique, n'est généralement pas prise en compte.

Ainsi, dans l'ensemble de ces approches, on ne retrouve que quelques mentions de l'emploi de différentes options de lexicalisation, c'est-à-dire de ressources lexicales différentes, que l'alternance codique rend possible, et très peu d'études qui traitent expressément du sujet. En fait, dans toute la littérature examinée dans la préparation de la présente étude, seules deux études abordant la signification en alternance codique de façon méthodique ont été trouvées. L'étude de King et Nadasdi (1999) porte sur l'alternance codique en tant que fonction discursive pour organiser la conversation, en particulier au moyen de verbes de croyance, d'opinion et de pensée (collectivement, des « expressions d'évidentialité »). Les auteurs observent que des distinctions sémantiques entre le français et l'anglais pour ces verbes ont une influence sur les alternances codiques produites par les participants. Spécifiquement, ils concluent que des expressions telles que *I think* ou *I guess* font l'objet d'alternance codique parce qu'elles expriment une idée d'incertitude absente des expressions comparables en français comme *je pense* ou *je crois*. Ainsi, ce serait la visée du locuteur qui aurait provoqué l'alternance codique : celui-ci peut exprimer de manière plus juste l'incertitude qu'il ressent envers des éléments de son propos grâce aux ressources linguistiques que lui fournit le code anglais. Cette étude démontre effectivement l'importance de la dimension sémantique de l'alternance codique; toutefois, les types d'alternances examinées sont très restreints et l'objectif ultime de cette étude n'est pas en fait une analyse sémantique, mais l'analyse de discours. L'étude de Backus (2001) porte sur le concept de la spécificité sémantique et son rôle dans la sélection des unités alternées. La notion de spécificité sémantique renvoie à la difficulté relative de remplacer un mot sans recourir à la paraphrase, et comprend autant le ou les référents ontologiques que les connotations et associations culturelles. En effet, les résultats de son étude permettent à l'auteur de conclure d'une part que plus un mot est sémantiquement spécifique, plus il est susceptible de faire l'objet d'une alternance codique, et d'autre part que les connotations, en particulier, permettent d'expliquer une grande partie des alternances codiques produites par ses participants. Il fait remarquer également l'influence sur l'alternance codique de la mise en relief (« focusing »), qui attire l'attention sur un item lexical particulier. La mise en relief s'accomplit soit explicitement (hésitations, prosodie, mention de recherche lexicale, etc.), soit

implicitement (structures de phrase emphatiques), et indique que le locuteur « is maximally aware of lexical selection, which means that the language of provenance of the word he/she wants to use is irrelevant, as long as the interlocutors know it too » (Backus, 2001, p. 148). Quoique cette étude présente malheureusement des faiblesses méthodologiques et conceptuelles significatives,¹⁴ sa contribution à une réflexion d'ordre sémantique sur l'alternance codique est appréciable.

Les études de King et Nadasdi (1999) et Backus (2001) sont toutefois des exceptions notables. Plus souvent, la possibilité que la visée référentielle (cf. la section 2.4.7) du locuteur joue un rôle dans l'alternance codique est généralement reléguée à des mentions anecdotiques sans approfondissement théorique.¹⁵ On sait cependant que des différences de sens entre des traductions dites équivalentes dans différentes langues peuvent en fait exprimer des sens légèrement différents, c'est-à-dire que la conceptualisation que représente un élément linguistique donné n'est pas identique d'une langue à l'autre (Backus, 2001). Pour un individu bilingue ayant connaissance de ces conceptualisations non identiques, produire de l'alternance codique est, entre autres, une manière de tirer profit de l'ensemble de ses ressources linguistiques et de s'exprimer pleinement et avec justesse.

2.4 Autres considérations théoriques pertinentes

2.4.1 Non-équivalence des synonymes et alternance codique

Un des présupposés sur lequel se base la présente étude est la non-équivalence des synonymes. En effet, « les occurrences de mots dits 'synonymes' ne co-'signifient' jamais complètement 'une même chose' » (Courbon, 2012, p. 330, reprenant à une citation de Vaugelas). La synonymie est plutôt une mise en relation de substitution possible des unités sur la base d'une convergence — et non pas une identité — conceptuelle. Les unités dites synonymes ne sont donc pas des équivalents, mais des « concurrents potentiels d'expression » (Courbon, 2012, p. 331). De plus, utiliser une forme différente implique de prendre une perspective différente sur ce dont on parle, c'est-à-dire, d'avoir une visée référentielle différente : « This perspective, marked by word-choice, allows [speakers] to present to their addressees a specific conceptualisation of an object, property, relation or event » (Clark, 1997, p. 1).

Appliqué au répertoire de ressources linguistiques bilingue, ce principe implique qu'il n'y a pas de synonymes parfaits entre langues pour l'individu bilingue, tout comme il n'y a pas de synonymes parfaits entre

¹⁴ Par exemple : son corpus consiste en trois enregistrements d'un total de neuf locuteurs; les connaissances encyclopédiques sont incluses dans la catégorie de « connotations »; il y a confusion entre l'acte de mettre un élément en relief et les indices qu'un tel acte s'est produit.

¹⁵ Par exemple, Heredia et Altarriba (2001) citent le cas de *cariño* en espagnol, qui comporte une idée d'affection qui n'est pas nécessairement présente dans *like* en anglais, en disant que cela pourrait engendrer de l'alternance codique afin de mieux se faire comprendre, sans plus.

les différents codes d'un individu monolingue (cf. la section 2.2.1). Autrement dit, ce qu'on appelle des « équivalents de traduction » ne sont pas des équivalents parfaits, mais des concurrents potentiels d'expression extracodiques, les uns par rapport aux autres, que l'on juge avoir suffisamment de convergence conceptuelle pour être une traduction acceptable (Backus, 2001)¹⁶. Cela est confirmé par les défis posés par la tâche de traduction et par l'existence même de la traductologie. Si les équivalents de traduction étaient effectivement équivalents, un simple dictionnaire bilingue suffirait pour traduire; il ne serait pas question d'une « bonne » ou « mauvaise » traduction, ni du jugement expert du traducteur, ni d'aucune théorie de traduction.

Sans vouloir impliquer que le simple fait d'être bilingue suffit pour faire d'un individu un traducteur compétent (voir à ce sujet Grosjean, 2013), l'intuition d'un locuteur bilingue comprend quand même la reconnaissance de la non-identité — sous un certain seuil de convergence — des conceptualisations des unités avec lesquelles il est familier à travers les langues qu'il connaît. L'importance de cette non-équivalence pour l'analyse sémantique de l'alternance codique vient du fait que, lorsque ses interlocuteurs comprennent et acceptent une alternance de codes, la priorité du locuteur bilingue n'est plus de maintenir la « monocodicité » de son énoncé, mais de bien représenter le sens particulier qu'il tient à exprimer (Backus, 2001)¹⁷. Si ce sens particulier est lexicalisé dans un autre code qu'il connaît et représente, pour lui, une divergence conceptuelle suffisamment grande par rapport aux concurrents d'expression disponibles monocodiquement, le locuteur peut être disposé à produire une alternance codique. Ainsi, la production de l'alternance codique peut signaler non pas un lapsus ou un manque de connaissances, mais une prise de perspective, une conceptualisation spécifique de ce à quoi l'unité ainsi produite fait référence.

2.4.2 Paradigme lexical multicodique

En conjonction avec la notion de concurrents potentiels d'expression dans différents codes, plusieurs cas ont été dégagés parmi les occurrences d'alternance codique relevées dans le corpus d'analyse où une unité

¹⁶ Comme Backus (2001) mentionne, cette convergence peut être suffisamment grande pour que les locuteurs ne remettent généralement pas en question l'équivalence des deux unités, comme dans le cas de *arbre / tree*, *sel / salt* ou *bol / bowl*. Cependant, même dans ces cas, l'équivalence n'est toujours pas parfaite. Notamment, ces mots appariés peuvent toutefois différer quant à leur combinatoire et quant aux locutions ou expressions auxquels ils contribuent et dont l'usage contribue en retour à la sémantisation des mots que font les locuteurs (voir la section 2.4.6). Par exemple, *arbre* est employé dans *arbre de transmission*, mais *tree* n'entre pas dans la désignation du même objet en anglais, à savoir *drive shaft*; il en est de même pour *worth one's salt* par rapport à *digne de ce nom*. Également, ces mots peuvent différer quant aux frontières qu'ils entretiennent avec des catégories conceptuelles voisines : la frontière entre *bol* et *tasse* pourrait être différente de celle entre *bowl* et *cup* de façon qu'un même objet qu'on désignerait par *bol* en français pourrait être appelé *cup* en anglais (voir à ce sujet Malt et Ameel, 2011). À tout le moins, ces mots vont différer pour l'individu bilingue quant à leurs contextes d'usage (quand, où, pourquoi et avec qui on utilise une langue plutôt que l'autre), et donc quant aux associations formées à travers son expérience.

¹⁷ Cela est comparable à la situation qui peut mener, dans un contexte monolingue, à l'insertion d'éléments d'un registre donné (qui est lui aussi un type de code) dans un énoncé d'un registre différent; par exemple, des mots formels dans un discours informel ou vice versa.

d'une langue ainsi que son « équivalent » dans l'autre langue font l'objet de l'alternance codique. Par exemple, on trouve dans le corpus autant *yeah* et *ouais*, ou *so* et *ça fait que*, parfois dans les productions d'un même locuteur. Ce phénomène est particulièrement pertinent pour la présente étude, d'une part parce qu'il élimine la possibilité que l'alternance soit produite par un manque de connaissances dans l'un ou l'autre code, et d'autre part parce qu'une comparaison directe des contextes d'emploi de ces « équivalents » pourrait apporter des éclaircissements sur des différences dans les visées référentielles que chaque unité est employée pour traduire.

La présence de ces unités appariées (« équivalents ») conduit par contre à un problème terminologique. Le terme *équivalent de traduction* est inadéquat, puisque ce ne sont pas, en fait, des équivalents (cf. la section 2.4.1). L'appellation *paire contrastante* a été considérée, afin d'évoquer le contraste entre les codes en présence comme la propriété saillante de ces unités en concurrence. Toutefois, ce terme aussi est inadéquat. D'une part, *paires* restreint indûment la catégorie, car rien n'oblige qu'une telle concurrence s'établisse entre deux unités seulement. D'autre part, comme le contraste est un phénomène à la base de toute perception, « paire contrastante » pourrait aussi bien référer à toute sorte de contraste – celui entre des phonèmes d'une même langue, pour ne mentionner qu'un exemple. Les appellations *paire transcodique* (à l'instar des *marques transcodiques* de Lüdi, 1987) et *paire intercodique*, en plus de ne pas résoudre le problème posé par *paire*, présentent d'autres difficultés. *Transcodique* implique un mouvement de traversée, alors que ce ne sont pas les paires repérées elles-mêmes qui « traversent » la frontière des codes; elles sont les traces d'un acte langagier déjà effectué et sont donc statiques. *Intercodique*, pour sa part, implique que les unités existeraient dans un espace intermédiaire entre les deux codes qui n'est ni l'un, ni l'autre. Quoiqu'il y ait peut-être lieu de parler d'un tel espace dans la construction d'une identité multilingue (voir par exemple Li, 2011), l'idée de « ni l'un, ni l'autre » dans la représentation des codes eux-mêmes neutralise la notion de contraste entre codes, c'est-à-dire de l'altérité codique.

Pour résoudre ce problème terminologique, il importe de se poser la question suivante : que perçoit-on dans ces unités qui fait qu'on les remarque? D'abord, il s'agit d'un ensemble (et pas nécessairement d'une paire) d'unités qui entrent en concurrence potentielle d'expression — autrement dit, il s'agit d'un paradigme, et en particulier d'un paradigme sur le plan lexical. En même temps, ce qui attire l'attention sur ces paradigmes est effectivement un contraste, soit le contraste entre différents codes à l'intérieur du paradigme — ou plutôt entre des unités associées à ces différents codes. Ce contraste est d'autant plus marqué que les codes en question sont reconnus comme des langues, une langue étant un code auquel on accorde un statut social et politique particulier. À la lumière de cette réflexion, c'est le terme *paradigme lexical multicode* (PLM) qui sera employé pour désigner cette relation entre unités dans la présente étude.

2.4.3 Ressources linguistiques

Le concept de *ressources*, évoqué à quelques reprises déjà dans ce chapitre, est un concept central à la présente étude. Il provient d'une approche interactionniste : « linguistic elements, signs and actions are used to accomplish or structure activities [...]: they therefore constitute resources for the organisation of interaction » (Doehler, dans Lüdi et Py, 2009, p. 164). L'emploi de ce terme dans la présente étude rejoint, dans l'ensemble, la conceptualisation de *ressources* de Lüdi et Py (2009) : un ensemble d'éléments linguistiques — constitué autant d'unités pré-formées que de processus de formation — partiellement stabilisé et organisé, mais ouvert; partagé, mais laissant place à l'individualité et à l'innovation. La conceptualisation de « ce que le locuteur sait sur sa langue » en termes de ressources plutôt que de compétence a l'avantage de mettre de l'avant tant le locuteur, qui emploie activement ses ressources pour accomplir ses buts, que la nature située, dynamique et co-construite des connaissances et compétences langagières (Lüdi et Py, 2009). Cependant, une approche purement interactionniste risque de délaissier les signes linguistiques en tant qu'outils pour traduire une visée référentielle; on remarque, par ailleurs, que Lüdi et Py (2009) ne parlent pas directement du lexique. Ainsi, j'utilise le concept de *ressources* de Lüdi et Py (2009) dans une acception plus large : l'ensemble des unités ou structures pré-formées (morphologiques, lexicales, syntaxiques) et des processus de formation, ainsi que des autres comportements et connaissances langagiers, notamment des connaissances pragmatiques, dont le locuteur dispose pour accomplir ses buts communicatifs (référer, structurer l'interaction, accomplir des fonctions perlocutoires, etc.).

2.4.4 Sélection lexicale

Lorsqu'on parle d'alternance codique en tant que stratégie d'expression sur le plan lexical, on parle nécessairement de la sélection d'une unité lexicale plutôt que d'une autre pour l'emploi en discours, soit une unité qui traduit le mieux sa visée référentielle. Pour cette raison, il s'avère nécessaire de clarifier ce qui est entendu par *sélectionner une unité lexicale*.

La présente étude ne vise pas à théoriser de manière approfondie sur les processus de production de parole chez les locuteurs bilingues, comme le moment où le choix de langue a lieu ou encore la nature du lexique chez les bilingues (pour des perspectives sur le sujet, voir Abutalebi et Green, 2008 ; de Bot, 1992 ; Kroll, Bobb, et Wodniecka, 2006 ; Lowie et Verspoor, 2011). Cependant, il importe de préciser que l'emploi d'une unité plutôt qu'une autre peut relever d'un choix conscient ou non, en fonction de plusieurs facteurs. Il n'est pas possible de distinguer, par des observations faites sur un corpus, entre des emplois qui sont conscients et réfléchis, et des emplois qui sont plutôt automatiques. Ainsi, le terme *choix*, qui implique un acte délibéré de choisir, sera dorénavant évité dans la présente étude. À la place, le terme *sélection* sera préféré, car il n'implique en soi que ce qui est observable dans les faits : on peut présumer que, si une unité lexicale a été produite, c'est qu'elle a été sélectionnée pour la production (au minimum, quant à la programmation moteur nécessaire pour

la prononcer) au cours du processus d'énonciation. À quel moment la sélection s'est faite et si elle a été plutôt consciente ou plutôt inconsciente sont des questions qui dépassent le cadre de la présente étude.

2.4.5 Communauté de pratique

La notion de *communauté de pratique* est fondamentale à la présente étude, notamment parce qu'elle est impliquée dans l'identification des codes et dans la perception de l'altérité codique (cf. la section 2.2.1), ainsi que dans les pratiques de sémantisation des locuteurs (cf. la section 2.4.6). Cette notion vient de Lave et Wenger (1991) et a été développée ensuite par Wenger (1998). À son origine, le terme *communauté de pratique* (« community of practice ») a été proposé dans le cadre de la théorie de l'apprentissage situé, qui met en valeur la communauté de personnes pratiquant une même activité ainsi que les connaissances et compétences de ses membres. Dans cette perspective, la communauté de pratique est envisagée à la fois comme la source de motivation pour l'apprentissage (afin de s'y intégrer davantage), comme une partie constitutive de l'identité des individus pratiquant cette activité, comme la source des connaissances et compétences à apprendre et comme le contexte d'apprentissage ainsi que de mise en œuvre de celles-ci (Wenger, 1998).

Communauté de pratique a ensuite été adopté dans le domaine de la linguistique, notamment en sociolinguistique et autres domaines connexes (Holmes et Meyerhoff, 1999). Comparé au concept mieux connu de *communauté linguistique*, le concept de *communauté de pratique* a le mérite de rendre compte des groupes identitaires comme étant socialement construits et constitués autour des pratiques partagées plutôt que de les délimiter « in virtue of shared abstract characteristics (e.g. class, gender) or simple copresence (e.g. neighbourhood, workplace) » (Eckert, 2006, p. 683). Cette conceptualisation de la notion de communauté tient compte aussi de la relation complexe entre l'identité personnelle du locuteur et la constellation de groupes identitaires auxquels il participe, comme l'illustre Eckert (2006) :

A white working-class Italian-American woman does not develop her ways of speaking directly from the larger categories 'working class,' 'Italian-American' and 'female,' but from her day-to-day experience as a person who combines those three (and other) memberships. Her experience will be articulated by her participation in activities and communities of practice that are particular to her place in the social order. It is in these communities of practice that she will develop an identity and the linguistic practices to articulate this identity (p. 684).

Communauté de pratique permet ainsi de lier les plans micro- et macro-situationnels, les pratiques individuelles et celles de différents groupes, de même que l'expérience subjective singulière et l'intersubjectivité (Holmes et Meyerhoff, 1999). Pour les raisons énumérées ci-dessus, le terme *communauté de pratique* sera préféré au terme *communauté linguistique* dans la présente étude.

La notion de *communauté de pratique* telle qu'employée dans la présente étude peut être comprise comme suit : chaque regroupement de personnes ayant des réalités saillantes qui lui sont particulières, et donc

des façons de parler de ces réalités (c'est-à-dire, des pratiques langagières partagées), constitue une communauté de pratique (voir à ce sujet Courbon, 2015, sur les concepts de *référentiel* et d'*univers de référence*). Ainsi, une communauté de pratique peut aller de relativement restreinte ou spécifique (par exemple, les pratiques langagières des chercheurs en biochimie dans une université donnée; les pratiques langagières des amateurs de la cuisine mexicaine) à relativement large (par exemple, les pratiques langagières de l'ensemble des francophones).

2.4.6 Sémantisation, habitudes d'usage et recontextualisation

Une autre notion importante pour la présente recherche est celle de la sémantisation. Suivant Courbon (2015), je considère que, comme tous les aspects de l'activité langagière, le sens linguistique n'est pas complètement objectif ni homogène. En effet, le sens des unités n'est que rarement enseigné de manière explicite et encore moins de manière formelle, c'est-à-dire que le sens n'est pas appris de façon uniformisée et pré-structurée, mais bien (co-)construit à travers les interactions qu'on entretient avec autrui (Brown, 2005). Ainsi, chaque locuteur sémantise les unités qu'il emploie, c'est-à-dire les investit de sens, selon son expérience du monde et des pratiques de sémantisation des autres, en particulier des différentes communautés de pratique auxquelles il appartient ou avec lesquelles il entre en contact.

Différentes communautés de pratique ont différentes habitudes quant à la sémantisation des unités, auxquelles le locuteur doit s'adapter et qu'il doit intégrer à ses propres habitudes afin de pouvoir communiquer. Ainsi, la sémantisation – et le sens en général – n'est pas seulement une relation à deux termes, entre mot et concept, mais à trois termes, entre mot, concept, et interlocuteur : « Words are symbols, cultural conventions, learned through a social process of interaction. The child¹⁸ has to learn what the community accepts as appropriate usage; what is conventional, what is creative, what are the boundaries to the category to which a word applies » (Brown, 2005, p. 295) Les pratiques de sémantisation d'un locuteur donné ont donc une part d'intersubjectivité, en ce que les concepts et les façons d'en parler (c'est-à-dire les pratiques référentielles) sont acquis par la socialisation et sont normés dans les communautés qui les utilisent, tout en comprenant une part de singularité, en ce que chaque locuteur a un parcours d'expériences et d'appartenances – qu'on peut appeler son parcours sémantique – qui lui est propre.

Effectivement, chaque locuteur entre en contact de façon plus ou moins intensive avec plusieurs communautés de pratique au cours de sa vie, en fonction de la région où il habite et des régions où il a déjà habité, de ses emplois, de son éducation, de ses passe-temps, de ses amis, et ainsi de suite. Le parcours de

¹⁸ Quoique l'ouvrage de Brown (2005) traite de l'apprentissage du sens chez les enfants pendant le développement initial du langage, la plupart de ses constats s'appliquent également aux locuteurs de tous âges, compte tenu du fait que l'apprentissage de nouvelles habitudes langagières, y compris des pratiques de sémantisation, ne cesse en fait jamais.

vie singulier d'un locuteur à travers différentes communautés de pratique fonde son univers de référence, c'est-à-dire l'ensemble des concepts dont il est susceptible de parler et les façons dont il est susceptible d'en parler (Courbon, 2015). Selon ce qui fait ou non partie de son parcours, et donc de son univers de référence, un locuteur aura des habitudes d'usage – y compris des habitudes de conceptualisation – différentes.

De ce fait, tout comme pour la perception de l'altérité linguistique, la perception qui importe quant aux différences dans la visée référentielle qu'une unité permet de traduire n'est pas celle du linguiste, mais bien celle du locuteur. Il est cependant impossible d'accéder directement à la sémantisation d'un locuteur donné à moins de lui demander directement, et ce, à condition que le locuteur lui-même soit en mesure de l'explicitier, ce qui n'est pas souvent le cas. Il faut plutôt se fier aux informations contextuelles : le cotexte immédiat, les habitudes d'usage de ce locuteur observées ailleurs dans le corpus, la situation de communication, le niveau de scolarisation des participants, les attitudes des participants, leurs parcours sémantiques, le sujet de discussion, les réactions des participants (cf. Gafaranga et Torras, 2002), les mises en relation des unités effectuées par les participants (cf. Courbon, 2012), et ainsi de suite.

Évidemment, il est encore impossible d'accéder directement à la contextualisation d'un énoncé lors de l'analyse, cette dernière s'effectuant nécessairement sur des traces décontextualisées. On peut, par contre, tenter de minimiser l'écart en recontextualisant l'énoncé à l'étude, c'est-à-dire en tenant maximale compte des facteurs contextuels dont on dispose (voir à ce sujet Kerbrat-Orecchioni, 2007). Une telle recontextualisation ne peut être que partielle. Dès le moment où on n'est plus dans l'interaction, une partie du contexte est perdue à jamais – même en travaillant à partir d'un enregistrement vidéo de haute qualité, on ne recaptura jamais toutes les informations contextuelles qui étaient disponibles aux locuteurs (Kerbrat-Orecchioni, 2007). Même si l'analyste lui-même a participé à l'interaction, il n'aurait plus accès à l'entièreté du contexte lorsqu'il commence son analyse. Enfin, même avec des informations contextuelles idéales, on n'aura jamais accès à ce qui se passait dans l'esprit des locuteurs au moment de produire leurs énoncés. Cependant, le fait d'être toujours partielle ne rend pas la recontextualisation moins essentielle pour l'étude des pratiques de sémantisation.

2.4.7 *Sens canonique, sens en usage et visée référentielle*

Avec la notion de sémantisation vient aussi le concept de *sens en usage*. Comme chaque locuteur sémantise les unités qu'il emploie selon son parcours sémantique singulier, il s'ensuit qu'on ne peut pas postuler une signification inhérente ou uniforme pour une unité donnée. Plutôt, dans une telle approche, on doit observer le ou les *sens en usage* d'une unité donnée, c'est-à-dire le sens avec lequel elle a été employée. En observant les contextes dans lesquels un locuteur emploie une unité donnée – y compris le cotexte, la thématique de discussion, et toute autre information pertinente – on peut accumuler des indices indirects du ou des sens dont le locuteur a investi l'unité en question dans un emploi donné. Par exemple, Vogh et Courbon (à paraître) ont

pu relever 4 catégories principales de sens en usage de l'adjectif *genetic* en anglais dans le discours en ligne en lien avec le diabète : relatif à la matière génétique physique, relatif à une prédisposition à la maladie, relatif à l'hérédité et aux antécédents familiaux, et relatif à l'inévitabilité.

Dans la présente recherche, j'étends l'objet d'étude du *sens en usage* à la *visée référentielle*, comme il a déjà été mentionné dans ce chapitre. Dans l'introduction, le terme *visée référentielle* a été défini comme la réalité conceptuelle que le locuteur souhaite communiquer à son interlocuteur au moyen de ressources linguistiques telle qu'il se l'approprie et s'y oriente. Pour compléter cette définition, il faut rappeler que la référence, comme tout aspect du sens (au sens large), est une relation à trois termes : le locuteur, le signe linguistique et l'interlocuteur. Autrement dit, la visée référentielle opère en fonction non seulement de ce que le locuteur entend, mais ce qu'il pense que son interlocuteur va comprendre. Ainsi, ce terme sera employé pour comprendre tout ce que l'emploi de l'unité en question semble communiquer, notamment :

- le ou les sens en usage, tel que défini ci-dessus. Le sens en usage peut ou non comprendre le sens « canonique » : le sens communément partagé, qu'on pourrait s'attendre à trouver dans l'usage général. Les guillemets sont utilisés pour souligner que ce sens « canonique » n'est qu'une abstraction d'un certain ensemble des pratiques sémantiques et n'est peut-être pas le sens réellement attribué en contexte par les locuteurs;
- les fonctions discursives accomplies par l'unité elle-même; cela exclut, par exemple, des effets perlocutoires attribuables à l'énoncé au complet.

L'inclusion des fonctions discursives n'implique pas que celles-ci soient confondues avec le sens des unités étudiées - c'est en partie pour cela que *sens en usage* et *visée référentielle* ne sont pas utilisés de façon interchangeable dans la présente étude, même si on peut y voir un degré de synonymie. Cependant, ce qu'une unité permet régulièrement d'accomplir en discours fait partie des connaissances intersubjectives du locuteur qui emploie cette unité, autrement dit de ce qu'il peut s'attendre que cette unité communique ou puisse communiquer à son interlocuteur, et donc de la visée référentielle qu'il associe à l'unité en question.

Il importe de souligner qu'il n'est pas dans les objectifs de la présente étude de prendre position dans une discussion portant sur la sémantique lexicale ou sur la nature de la signification. Je n'entends pas par là que ces questions sont sans intérêt ou sans importance, ni que la conceptualisation qu'on adopte de la signification ou du lexique n'a pas d'influence profonde sur les données sélectionnées pour l'analyse, sur la méthodologie utilisée pour les analyser et ultimement, donc, sur les résultats et les conclusions qu'il est possible d'en tirer. Toutefois, la présente étude vise la prise en compte de ce que les locuteurs font réellement avec les unités lexicales qu'ils emploient, au moyen d'une analyse en contexte des emplois effectifs repérés dans un

corpus. Comme la nature de la signification ne fait pas partie de ce qui est observable dans le corpus d'analyse, elle ne constitue pas l'objet de la présente étude.

Il s'ensuit que la méthode d'analyse utilisée dans la présente étude ne me permet pas de tirer des conclusions de mon corpus d'analyse quant à la division entre la sémantique et la pragmatique, par exemple où se trace la frontière entre le sens d'une unité et les connaissances du monde en lien avec cette unité, ou entre le sens d'une unité et sa fonction ou son interprétation dans le discours. La méthode d'analyse ne permet pas non plus de trancher sur la nature même de la signification, par exemple si la signification d'une unité est plutôt vague ou schématique et est précisée en contexte par des mécanismes quelconques, ou si la signification est plutôt une liste plus ou moins détaillée des sens possibles qu'une unité peut prendre en contexte (pour un aperçu de la complexité de ces questions, fondamentales pour le domaine de la sémantique, et des réponses qui sont proposées par différents auteurs, voir par exemple Evans, 2009 ; Langacker, 2008 ; Ravin et Leacock, 2000 ; Récanati, 1997 ; Ruhl, 1989). Compte tenu du fait que la distinction entre sens et fonction importe pour certaines théories de la signification et pas pour d'autres, il ne convient donc pas d'imposer une distinction catégorielle entre les deux là où cela n'est pas pertinent à l'analyse des stratégies d'expression employées par les locuteurs dont on étudie les productions¹⁹. L'emploi du terme *visée référentielle* pour englober tant les sens en usage que les fonctions discursives d'une unité se veut une manière de parler des deux sans devoir imposer une frontière tranchée entre eux sans toutefois les confondre, c'est-à-dire en laissant ouverte la possibilité d'éventuellement établir et exploiter une distinction entre les deux, si besoin en est.

Il importe également de souligner ici que, comme l'inclusion des fonctions discursives et le rôle important du contexte le laissent croire, la présente étude s'appuie fortement sur des travaux et des notions issues de l'analyse conversationnelle. Toutefois, il s'agit bien d'une étude de l'alternance codique comme stratégie d'expression sur le plan lexical, non pas d'une étude d'analyse conversationnelle en tant que telle. Son objet principal demeure la visée référentielle traduite par des unités lexicales; des notions comme les tours de parole, la rupture conversationnelle et les fonctions discursives sont plutôt des outils utiles pour l'analyse de cet objet principal dans un corpus de l'oral.

2.5 Récapitulatif du chapitre 2

Ce chapitre a présenté des concepts, des termes et des définitions nécessaires pour situer la présente étude parmi les différentes traditions épistémologiques et empiriques qui traitent de l'alternance codique et de la signification. On peut résumer l'approche de la présente étude comme suit : un *code* est un ensemble de comportements (langagiers, en l'occurrence) socialement significatif. Une *langue* est un code; elle se constitue

¹⁹ Dorénavant « les locuteurs étudiés ».

d'un ensemble de ressources linguistiques considéré socialement significatif par ses locuteurs. Une langue est donc une réalité mentale et sociale plutôt qu'un système immanent. L'*alternance codique* renvoie à la juxtaposition des éléments de différentes langues de façon à ce que les locuteurs en présence la perçoivent comme une juxtaposition de codes-langues. Afin d'analyser l'alternance codique en tant qu'une possible stratégie d'expression avec une telle approche, il est nécessaire de prendre en compte les perceptions des locuteurs et leurs pratiques de sémantisation des unités lexicales qu'ils emploient. Cela implique d'étudier des emplois effectifs repérés dans un corpus, et de tenir compte des informations disponibles dans et autour du corpus qui permettent de recontextualiser ces emplois.

Le prochain chapitre décrit les caractéristiques recherchées dans la sélection d'un corpus approprié pour ce genre d'analyse. Il présente ensuite le corpus qui a ultimement été constitué pour la présente étude ainsi que des informations autour de ce corpus, soit la nature des enquêtes qui ont donné lieu aux productions analysées et le contexte socio-historique des régions où ces enquêtes ont eu lieu.

Rapport-Gratuit.com

3 Le corpus analysé

Pour pouvoir réaliser une étude portant sur ce que des unités lexicales communiquent en usage, il faut nécessairement des usages à analyser, et plus précisément, un corpus. La présente recherche nécessite en particulier un corpus où figurent un bon nombre de productions d'alternance codique. Ces productions doivent être identifiables comme telles par l'analyste, et leurs sens en usage doivent lui être accessibles. De ces faits découlent un ensemble de critères qui doivent caractériser le ou les corpus à analyser; ces critères sont détaillés dans la section 3.1. La section 3.2 décrit les corpus sélectionnés selon ces critères (les « sous-corpus » du corpus d'analyse), y compris les contextes sociohistoriques des communautés dans lesquelles ils ont été collectés. La section 3.3 décrit le processus d'échantillonnage opéré sur ces corpus et son résultat : le corpus réellement exploité dans la présente recherche (« corpus d'analyse » ou « corpus étudié »).

3.1 Critères de sélection des corpus à analyser

3.1.1 Critères favorisant une quantité suffisante d'occurrences d'alternance codique

3.1.1.1 *Corpus existants et accessibles*

Pour pouvoir analyser le plus grand nombre possible d'occurrences en un temps raisonnable, j'ai choisi d'exploiter des corpus existants — donc, des corpus publiquement accessibles ou auxquels on accepte de m'accorder l'accès aux fins de la présente recherche — plutôt que de faire une collecte de données originales. Il fallait également que ces corpus soient dans un format facilement exploitable, idéalement numérique. Ce choix implique toutefois certaines limites sur les données disponibles pour l'analyse; ces limites seront discutées lorsqu'elles se présentent.

3.1.1.2 *Corpus de conversation libre ou semi-dirigée*

Parce qu'elle représente une déviation des normes prescriptives monolingues, et parce qu'elle implique souvent l'emploi d'une langue ou d'une variété de langue qui a peu de prestige, l'alternance codique est souvent un phénomène stigmatisé (Boztepe, 2003; Gardner-Chloros, 2009; Grosjean, 1997; Pfaff, 1979; Toribio, 2001). Même lorsqu'elle est socialement acceptable, voire valorisée, au sein d'une communauté donnée, elle l'est généralement moins dans des situations de distance communicative²⁰, situations plutôt formelles où les normes tendent à s'aligner vers le standard monolingue écrit (Boztepe, 2003; Gardner-Chloros, 2009; Kootstra, van Hell, et Dijkstra, 2009; Paquet-Gauthier et Beaulieu, 2015). Dans de tels contextes, « multilinguals must rely on what is supposed to be shared by all participants, that is, a single language used appropriately for situations of communicative distance » (Paquet-Gauthier et Beaulieu, 2015, p. 174). Inversement, les situations les plus

²⁰ Le terme vient de Koch et Oesterreicher (2001).

propices à la production de l'alternance codique sont celles avec des caractéristiques typiques de l'immédiat communicatif²¹ : l'interaction est spontanée et informelle, le choix de thématique est libre et les transitions fluides, les interlocuteurs sont co-présents et ils sont relativement à l'aise les uns avec les autres. Dans de tels contextes, les normes sont assouplies, d'une part, et d'autre part les locuteurs bilingues sont plus en mesure de déterminer les ressources et les attitudes qu'ils partagent avec leurs interlocuteurs, et dont ils peuvent se servir pour produire un énoncé qui sera compris et accepté. Le fait de partager des ressources dans plus d'une langue « allows them to make use more freely of the creative combinatory possibilities of these codes » (Paquet-Gauthier et Beaulieu, 2015, p. 175).

En ce qu'il ressemble le plus à la situation prototypique de proximité communicative, le protocole d'enquête idéal serait une conversation libre entre deux ou plusieurs membres d'une même communauté de pratique, qui participent tous à l'enquête au même titre. Dans les faits, ce genre de protocole est très rare. On trouve plus souvent des enquêtes employant un protocole de conversation semi-dirigée, où un enquêteur anime la conversation en posant des questions ou en amenant des sujets de discussion pré-établis. Ce sont donc des corpus de conversation libre ou semi-dirigée qui ont été recherchés. Il est à noter qu'aucun des deux protocoles ne reproduit parfaitement les conditions de proximité : la présence d'un enregistreur et, le cas échéant, d'un enquêteur ou d'un animateur diminue plus ou moins l'intimité et la spontanéité.

3.1.1.3 *Locuteurs bilingues*

Il est tout à fait possible que des locuteurs considérés monolingues produisent de l'alternance codique. Par exemple, l'expression courante « go, t'es capable! » est souvent produite par des francophones considérés monolingues au Québec²². Toutefois, la présente recherche portant sur la visée référentielle des individus bilingues, il est donc nécessaire que le corpus étudié soit constitué d'énoncés produits par des locuteurs bilingues.

Cependant, le bilinguisme — et, au même titre, le monolinguisme — est un statut difficile à délimiter. Si on s'accorde généralement pour dire que le bilinguisme est le fait, pour un locuteur, de connaître deux (ou plusieurs) langues, la question de ce qu'est « connaître une langue » reste ouverte. Quel niveau précis de compétence doit-on avoir atteint dans une L2, et selon quel(s) observateur(s), pour être considéré bilingue? Par quelles mesures évalue-t-on cette compétence? Évidemment, répondre à ces questions, qui renvoient à des

²¹ Le terme vient de Koch et Oesterreicher (2001).

²² Cela arrive en particulier si, comme c'est le cas pour l'anglais dans des communautés francophones en Amérique du Nord, la L2 fait partie de la « culture ambiante ». Les locuteurs dits monolingues dans une telle situation ont souvent en effet au moins quelques ressources dans un autre code, même si on ne dirait pas qu'ils sont tout à fait bilingues.

débats encore ouverts²³, dépasserait le cadre de la présente recherche. Il faut néanmoins une définition de travail qui permet de délimiter un ensemble de données pertinentes pour cette recherche.

Il serait peu réaliste de fonder cette définition sur l'image populaire du bilinguisme, soit la maîtrise parfaite de deux langues : la plupart des individus bilingues ou multilingues n'atteignent pas les mêmes compétences dans toutes leurs langues, entre autres parce qu'ils ne les utilisent pas dans tous les mêmes contextes (voir à ce sujet les travaux de Grosjean, par exemple, 1997; 2013). Une personne qui utilise couramment la langue A à la maison et la langue B à l'école pourrait ne pas avoir beaucoup de ressources pour parler de sujets savants en langue A, ou pourrait même ne pas savoir lire ou écrire en langue A; on ne dirait pas pour autant que cette personne n'est pas bilingue. Or, le degré d'habileté dans les deux langues n'est pas sans importance : « while almost everyone knows at least a few words in other languages, we generally require a little more competence than that before we are willing to acknowledge bilingual or multilingual ability. » (Edwards, 2013, p. 5).

Afin de prendre en compte « the vast majority of people who use two or more languages regularly but who do not have native-like fluency in each » (Grosjean, 1997, p. 164), Grosjean propose de définir le bilinguisme comme le fait d'employer plus d'une langue dans sa vie de tous les jours. Cette définition met intentionnellement à l'écart des questions de quantification ou de catégorisation des compétences d'un tel locuteur sans toutefois évacuer complètement la notion de compétence, car ce locuteur doit être en mesure d'employer les langues en question dans son quotidien. On trouve une définition semblable pour la notion de « bilinguality » chez Hamers et Blanc (2000, p. 6) : « the psychological state of an individual who has access to more than one linguistic code as a means of social communication ». Ces auteurs précisent que « the degree of access will vary along a number of dimensions which are psychological, cognitive, psycholinguistic, social psychological, social, sociological, sociolinguistic, sociocultural and linguistic » sans toutefois tenter d'établir un seuil quantifiable du degré d'accès qui ferait d'un locuteur donné un individu bilingue.

Cette mise à l'écart n'implique pas que la quantification et la catégorisation du bilinguisme – par exemple, selon les niveaux de référence du Cadre européen commun de référence pour les langues (ou CECR; Conseil d'Europe, 2005), selon le moment d'acquisition des deux langues (le bilinguisme simultané, séquentiel précoce ou séquentiel tardif), le contexte et la fréquence d'exposition à ou d'emploi de l'une ou l'autre langue, la langue dominante et le degré de dominance, etc. (Grosjean, 1997; Mougeon et Beniak, 1996) – ne sont pas importantes. Par exemple, il est nécessaire pour une école de langues de pouvoir évaluer et classer les étudiants selon leur niveau de compétence afin de les placer dans le cours approprié. La quantification et la catégorisation

²³ Pour une discussion des défis entourant la question de la définition, voir Edwards (2013), Kemp (2009), et Li (2000). Pour une discussion des défis entourant la question de l'évaluation, voir Anderson *et al.* (2018).

sont nécessaires en recherche aussi, selon les objectifs de l'étude : par exemple, si on souhaite étudier, comme l'ont fait Mougeon et Beniak (1996), l'influence de la langue majoritaire sur les pratiques langagières d'une communauté de langue minoritaire, on doit nécessairement contrôler la quantité et les contextes d'exposition aux langues en question.

Cependant, comme le font ressortir Edwards (2013), de Bot (2016) et Anderson *et al.* (2018), le bilinguisme est essentiellement non catégoriel, relevant d'un ensemble complexe d'expériences et d'habiletés dont aucun critère, seuil ou catégorie ne pourra rendre compte de façon adéquate. Le définir n'implique donc pas de trouver un ensemble de critères qui sera décisif dans tous les contextes, mais bien de sélectionner les critères les mieux adaptés aux objectifs de recherche en question. Quant à la présente recherche, la définition fonctionnelle de Grosjean est tout à fait adaptée. Il est vrai que des facteurs tels que le moment d'acquisition ou le degré de dominance de l'une ou l'autre langue ont effectivement une incidence sur les ressources linguistiques des locuteurs et sur la propension à l'alternance codique. Toutefois, ils n'ont pas d'incidence sur le fait qu'un individu bilingue ainsi défini soit en mesure de juxtaposer de manière originale²⁴ des ressources linguistiques de deux codes-langues différents dans ses énoncés.

De plus, comme la présente recherche exploite des corpus préexistants, il n'est pas possible d'évaluer les compétences des locuteurs au moment où ils s'y exprimaient. Même s'il était possible d'entrer en contact avec certains – ceux qui ne sont pas décédés ou devenus introuvables – les évaluer aujourd'hui ne donnerait pas d'informations fiables sur leurs compétences au moment de l'enregistrement du corpus : ils auraient eu entretemps des expériences langagières qui influent sur leurs pratiques, leurs compétences, leurs préférences, etc. Il n'est donc possible d'évaluer leur niveau de bilinguisme que par leurs comportements langagiers mis en évidence dans le corpus en question, c'est-à-dire, que par des critères fonctionnels. Cela va dans le sens d'Auer (1995), pour qui le bilinguisme « has its foremost reality in the interactive exchanges between members of a bilingual speech community (as well as between them and monolingual outsiders), by which they display to each other, and ascribe to each other, their bilingualism » (p. 115).

Appliqué à la sélection des corpus, le critère des locuteurs bilingues – selon la définition fonctionnelle donnée ci-dessus – fait en sorte que le ou les corpus recherchés devaient être composés majoritairement²⁵ des enregistrements dans lesquels on peut observer des locuteurs s'exprimant dans au moins deux codes-langues.

²⁴ Contrairement aux productions d'alternance codique des individus monolingues, qui, comme l'exemple de « go, t'es capable! » ci-dessus, tendent à être des formules plus ou moins conventionnalisées (dont certaines pourraient même être considérées des emprunts, *cf.* la section 4.1.4).

²⁵ Il est possible qu'une enquête menée principalement auprès de locuteurs bilingues inclue toutefois des participants monolingues (*cf.* la section 3.1.1.4).

3.1.1.4 Enquêtes menées auprès des membres des communautés de langue minoritaire sur le plan régional

Pour la présente étude, l'intérêt de viser des membres d'une communauté de langue régionalement minoritaire est que ces locuteurs ont davantage de chances de répondre à la définition fonctionnelle du bilinguisme énoncée à la section 3.1.1.3. Puisque les possibilités sont élevées que certaines activités (école, emploi, loisirs, interactions avec les amis ou la famille étendue, etc.) se déroulent dans la langue majoritaire, et que d'autres se déroulent dans la langue minoritaire, il est permis de supposer que de tels locuteurs ont la capacité d'utiliser plus d'une langue dans la vie de tous les jours. Autrement dit, ces locuteurs sont presque assurément exposés à plus d'un code-langue, et ce, à un degré d'intensité relativement élevé.

C'est cette exposition, et non le statut minoritaire de la langue en soi, qui augmente les chances que ces locuteurs produisent de l'alternance codique dans leurs énoncés (Mougeon et Beniak, 1996). En effet, on peut être membre d'une communauté de langue minoritaire au sein de laquelle les activités de tous les jours se déroulent majoritairement ou entièrement dans la langue minoritaire et par conséquent démontrer relativement peu d'influence de la langue majoritaire dans ses énoncés. C'est le cas, par exemple, de la communauté franco-ontarienne de Hawkesbury étudiée par Mougeon et Beniak (1996, entre autres). Cette communauté constitue une minorité linguistique à l'échelle provinciale, mais à l'intérieur de la communauté elle-même, 85 % des locuteurs sont francophones et la plupart des domaines de la vie courante sont associés au français (Mougeon et Beniak, 1996). Par conséquent, les auteurs ont observé une faible influence de l'anglais chez ces locuteurs; leurs usages linguistiques sont relativement semblables à des locuteurs québécois monolingues.

Également, une enquête qui porterait sur une telle communauté peut très bien inclure des participants qui ont très peu d'exposition à la langue minoritaire. En situation minoritaire, l'identité et l'héritage prennent des dimensions d'autant plus importantes que les conditions sociales, politiques et économiques nuisent à la transmission de la langue d'une génération à l'autre (Schmidt, 2008 ; cf. aussi dans ce sens Mougeon et Beniak, 1996). Ainsi, un individu ayant un fort attachement à son histoire familiale ou à son héritage culturel pourrait s'identifier comme membre d'une communauté de langue minoritaire même s'il ne partage pas les pratiques linguistiques autour desquelles ladite communauté est définie.

Il s'ensuit que le ou les corpus recherchés devaient provenir d'une enquête menée auprès d'une communauté de langue minoritaire, avec la reconnaissance toutefois que ce statut de langue minoritaire sert à favoriser mais non à garantir que les locuteurs qui y figurent soient aptes à produire de l'alternance codique.

3.1.2 Critères concernant l'analysabilité des occurrences

3.1.2.1 *Corpus sur support audio ou audiovisuel*

De préférence, le ou les corpus analysés devaient être sur un support audio ou audiovisuel. Ces formats fournissent des indices utiles pour la recontextualisation et donc pour l'interprétation et l'analyse de la visée référentielle : le ton de voix, les pauses et hésitations, les gestes et mimiques, etc. (Kerbrat-Orecchioni, 2005; 2007). Ces éléments ne sont pas toujours notés dans un corpus d'oral transcrit, selon les buts de la recherche originale pour laquelle la transcription a été faite (cf. Kerbrat-Orecchioni, 2005). Avoir recours directement au support audio ou audiovisuel permet d'assurer que ces indices sont bien intégrés à l'analyse.

3.1.2.2 *Locuteurs bilingues anglais-français*

Étant donné que l'analyse porte en partie sur le sens en usage des occurrences d'alternance codique, l'analyste doit être en mesure d'apercevoir les nuances de sens et de contexte qui influencent le sens. Les langues entre lesquelles les locuteurs alternent dans le corpus doivent donc être des langues dans lesquelles l'analyste a un niveau de compétence suffisamment élevé; en l'occurrence, ces langues sont l'anglais et le français. Ainsi, seront retenus des corpus dont les enquêtes visent des locuteurs bilingues anglais-français seulement.

Ce critère n'empêche pas en soi que certains locuteurs participant à une telle enquête aient été exposés auparavant à d'autres langues en plus de l'anglais et du français — et comme il n'y a pas d'accès direct aux locuteurs dans des corpus préexistants, il n'est pas possible de contrôler d'avance cette exposition. Force est d'admettre qu'avoir dans son répertoire des ressources linguistiques d'une troisième langue (ou plus) pourrait avoir une incidence non seulement sur les ressources des codes anglais et français que le locuteur est disposé à utiliser, mais aussi sur le sens avec lequel il les investit, au moyen de l'influence translinguistique (Jarvis et Pavlenko, 2008; Cook, 1992). Cependant, puisque l'analyse porte justement sur le sens en usage, c'est-à-dire le sens avec lequel un locuteur spécifique investit une unité lexicale donnée selon un contexte personnel et conversationnel donné, elle prend déjà en compte de telles variations. Autrement dit, l'analyse prévoit déjà que chaque locuteur investit différemment les unités lexicales qu'il emploie selon son parcours référentiel singulier, que ce parcours contienne ou non une deuxième ou troisième (etc.) langue (Courbon, 2015 ; Vogh et Courbon, à paraître).

Il a été prévu d'exclure de l'analyse, le cas échéant, une occurrence d'alternance codique dont le contexte conversationnel nécessaire à son interprétation contient des éléments d'une langue autre que l'anglais et le français, puisqu'une telle occurrence serait difficilement analysable. Toutefois, un tel cas ne s'est pas réellement produit au cours de l'analyse.

3.1.2.3 Communautés en Amérique du Nord

Puisque ce qui compte ou non comme un code distinct est socialement défini (cf. la section 2.1.1), il faut avoir accès aux normes codiques communautaires pour bien rapprocher son identification des occurrences d'alternance codique de celle des locuteurs. Au même titre, ce sont les normes et pratiques communautaires qui permettent de distinguer un emprunt et une occurrence d'alternance codique (cf. la section 4.1.4). En l'occurrence, mes connaissances portent surtout sur les normes et usages du français et de l'anglais nord-américains. Il s'ensuit que le ou les corpus retenus devaient être constitués à partir d'enquêtes menées en Amérique du Nord.

Force est de constater que l'Amérique du Nord constitue une « communauté » au sens très large, et que le corpus est au final assez hétérogène. Cependant, étant donné la relative pénurie de corpus correspondant aux critères déjà énumérés et la relative rareté des occurrences d'alternance codique dans un corpus donné (comparé à des phénomènes phonologiques ou morphologiques, par exemple), il ne convient pas de limiter les enquêtes considérées à une ou des communautés plus restreintes.

3.1.3 Exclusion des corpus d'apprenants

On pourrait penser que les corpus dits « d'apprenants » d'une langue seconde ou étrangère (corpus de cours de langue, d'enfants en bas âge, etc.) constituent une source potentielle d'occurrences d'alternance codique. En effet, on peut parfois y trouver des locuteurs qui s'expriment dans plus d'une langue. Ces corpus présentent cependant quelques difficultés par rapport aux critères énumérés ci-dessus.

Tout d'abord, il y a de bonnes chances que les locuteurs qui s'expriment dans un corpus d'apprenants ne répondent pas à la définition de bilinguisme présentée à la section 3.1.1.3, c'est-à-dire qu'ils n'utilisent pas plus d'une langue dans la vie de tous les jours. Bien sûr, il n'est pas garanti que ce soit le cas simplement parce qu'il s'agit d'un corpus d'apprenants, mais il y a plus de chances que les locuteurs dans un tel corpus n'aient pas encore la compétence fonctionnelle évoquée en 3.1.1.3 dans les langues mises en évidence dans le corpus, ou qu'ils ne vivent pas dans un contexte où cette langue est utilisée dans la vie de tous les jours, c'est-à-dire, en dehors de la salle de classe (Granger, 2008).

Deuxièmement, les protocoles d'enquête utilisés pour constituer des corpus d'apprenants ne favorisent généralement pas l'alternance codique. D'une part, puisque les locuteurs apprenants n'ont habituellement pas beaucoup d'exposition à l'interaction spontanée dans la langue qu'ils apprennent (Beaulieu *et al.*, 2016; Beaulieu et Dupont Rochette, 2014), les enquêtes s'appuient souvent sur des tâches de sollicitation comme remplir des textes à trous, résumer un texte, traduire un texte, etc. (Granger, 2008). Ce genre de tâche donne lieu à des situations plutôt de distance communicative qui sont moins propices à la production de l'alternance codique (cf. la section 3.1.1.2). D'autre part, puisque les enquêtes cherchent spécifiquement à évaluer ou à documenter les

productions dans une langue en particulier, c'est-à-dire la langue que les participants apprennent, ces derniers sont dissuadés de produire de l'alternance codique, soit implicitement par la nature des tâches, par le comportement de l'enquêteur ou par les normes établies lors de l'instruction, soit explicitement par les directives de l'enquête.

Une troisième difficulté que les corpus d'apprenants pourraient introduire relève plutôt de l'analysabilité des occurrences qu'on pourrait trouver dans un corpus d'apprenants. Par rapport à d'autres types de corpus d'interaction bilingue, l'alternance codique relevée dans un corpus d'apprenants est plus susceptible d'avoir été produite pour résoudre un problème d'accès lexical ou pour pallier le manque de ressources dans la langue d'apprentissage; l'alternance serait en quelque sorte nécessaire pour que l'interaction se poursuive. Comme l'analyse porte sur les différences d'investissement de sens entre des unités de langues différentes, il est préférable que les occurrences d'alternance codique examinées soient plutôt le résultat de la sélection de l'unité qui correspond le mieux à la visée référentielle du locuteur.

Il est à noter qu'aucune de ces difficultés n'est exclusive aux corpus d'apprenants. Quant au bilinguisme, comme il a été mentionné à la section 3.1.1.4, une enquête ciblant principalement des locuteurs bilingues peut toutefois inclure des locuteurs qui ne sont pas en mesure d'employer plus d'une langue dans leur vie de tous les jours. Pour ce qui est des protocoles d'enquête, tout corpus sera influencé par les objectifs et la nature de l'enquête de laquelle il découle. Par exemple, bon nombre de corpus où s'expriment des locuteurs bilingues découlent d'enquêtes qui ont pour but de documenter les productions dans une langue donnée. C'est un objectif particulièrement fréquent dans les enquêtes sur des langues minoritaires, et il peut influencer les tâches de production qui sont utilisées, ou les questions d'entrevue, et ainsi les données disponibles dans le corpus. Tout corpus sera aussi influencé par le fait même qu'il découle d'une enquête : même avec un protocole de conversation libre ou semi-dirigée entre locuteurs qui se connaissent bien, le simple fait de se savoir enregistré et ultérieurement étudié influe sur la situation de communication et ainsi sur les productions. Par exemple, le fait de vouloir « bien répondre aux questions » ou de « bien paraître sur l'enregistrement » peuvent passer au premier plan et supplanter les objectifs premiers de la communication – communiquer efficacement, bien gérer l'interaction, etc. Autrement dit, aucun corpus n'arrivera à capter véritablement un contexte de production tout à fait naturel ou normal pour les locuteurs en question.

De plus, à l'égard de l'alternance codique en tant que stratégie de contournement, tout locuteur bilingue peut recourir à la langue B pour continuer l'interaction s'il ne connaît pas une unité lexicale de la langue A, l'a oubliée ou ne peut pas y accéder temporairement. En fait, les chances que cette situation se produise sont encore plus élevées en contexte d'enquête, où la présence de l'enquêteur ou de l'enregistreur peut mettre le locuteur plus ou moins mal à l'aise.

Toutefois, même si les difficultés énumérées ci-dessus ne sont pas exclusives aux corpus d'apprenants, elles y sont particulièrement présentes. En choisissant d'exclure ce type de corpus, on peut minimiser de façon significative leur impact potentiel sur l'analyse. La réduction de leur impact est d'autant plus important dans le cas des stratégies de contournement qu'il n'y a pas de critère empiriquement observable qui permette d'identifier les occurrences qui en relèvent. Par exemple, Poplack (1987) suggère que les occurrences dysfluides (c'est-à-dire, avec pauses et hésitations) indiqueraient « switching for lack of lexical or syntactic availability » (p. 54), mais, tandis que la dysfluidité peut effectivement indiquer une recherche lexicale en cours, cette recherche lexicale pourrait indiquer la sélection du « mot juste » pour exprimer sa visée référentielle aussi bien qu'elle pourrait indiquer un problème d'accès lexical. On ne peut pas non plus se fier à la conscience des locuteurs en ce qui a trait à leurs propres processus de production pour trancher cette question, puisque ces processus sont souvent inconscients (Gardner-Chloros, 2009). De plus, en ce qui concerne la présente analyse, il n'est tout simplement pas possible d'avoir accès à la conscience linguistique des participants : même s'il était envisageable de recontacter les locuteurs enregistrés et de les interroger sur leurs choix lexicaux, le temps écoulé depuis la production des énoncés en question affaiblirait considérablement la fiabilité des indices fournis par leurs réponses. On doit plutôt admettre qu'il n'est tout simplement pas possible de déterminer, dans l'état actuel des connaissances, si une occurrence d'alternance codique est effectivement le résultat d'une sélection lexicale en fonction de la visée conceptuelle du locuteur. Ainsi, il importe que les conditions d'enquête favorisent autant que possible la participation de locuteurs ayant les compétences fonctionnelles recherchées et le déroulement de l'interaction dans une situation de proximité communicative, non seulement pour augmenter les chances de repérer une bonne quantité d'occurrences d'alternance codique, mais aussi pour que ces occurrences soient susceptibles de relever d'une stratégie d'expression.

3.1.4 Vue d'ensemble : caractéristiques des corpus retenus

En résumé, les corpus retenus pour l'analyse devaient être des corpus existants et accessibles, de conversation libre ou semi-dirigée (à l'exclusion de corpus d'apprenants), enregistrés en format audio ou audiovisuel auprès de locuteurs nord-américains bilingues anglais-français appartenant à des communautés de langue minoritaire sur le plan régional. Ces critères excluent les enquêtes menées au Québec auprès des communautés identifiées comme francophones, les enquêtes menées ailleurs en Amérique du Nord auprès des communautés identifiées comme anglophones, et les enquêtes menées auprès des francophones apprenants de l'anglais ou des anglophones apprenants du français. La recherche de corpus répondant à l'ensemble de ces critères a donné trois résultats; ces corpus, ainsi que le procédé d'échantillonnage utilisé pour constituer le corpus final sous analyse, sont décrits dans le reste de ce chapitre.

3.2 Corpus retenus

3.2.1 Description générale

Les trois corpus correspondants aux critères de sélection énumérés à la section 3.1 sont les sous-corpus qui, ensemble, constituent le corpus analysé dans le présent mémoire. Ces sous-corpus sont :

1. Les enregistrements faits en 2011 pour la collecte de données qualitatives dans le cadre de la thèse de doctorat de la professeure Suzie Beaulieu, co-directrice de la présente recherche (dorénavant le « corpus Beaulieu »)²⁶;
2. Les enregistrements de la partie « conversation libre » du protocole du projet *Phonologie du français contemporain*, spécifiquement ceux de l'enquête à Peace River, Alberta en 2001 (dorénavant le « corpus PFC »)²⁷;
3. Une sélection de vidéos des archives du Franco-American Center, à l'Université du Maine, Orono (dorénavant le « corpus FAC »)²⁸.

Il est à noter que les critères de facilité d'accès et d'exploitation (cf. la section 3.1.1.1) a mené à l'exclusion de certains corpus bien connus dans la littérature sur le français en situation minoritaire en Amérique du Nord, tels le corpus Ottawa-Hull (Poplack, 1989) et les corpus de Mougeon et ses collègues recueillis auprès des francophones de l'Ontario (voir par exemple Mougeon et Beniak, 1996; Mougeon, Nadasdi, et Rehner, 2009). La question de facilité d'exploitation, en particulier, exclut malheureusement le corpus Maheux-Pelletier, recueilli en Alberta, auquel on m'a toutefois généreusement accordé l'accès²⁹.

Une présélection a dû être opérée sur les archives vidéo du Franco-American Center en raison de la taille des archives : elles contiennent 170 vidéos en tout, enregistrées entre 1989 et 2014, la plupart d'une durée de deux à trois heures. Les critères de cette présélection portent sur :

- la date d'enregistrement : pour rester comparables avec les deux autres sous-corpus, seules les vidéos enregistrées entre 2001 et 2012 ont été retenues³⁰. Toute vidéo pour laquelle une date d'enregistrement n'était pas disponible a aussi été exclue.

²⁶ cf. Beaulieu (2012).

²⁷ cf. Walker (2012); Durand, Laks et Lyche (2002; 2009). Le projet PFC est accessible sur <http://www.projet-pfc.net>. Au moment de rédiger ces lignes, le site web du projet est en refonte majeure depuis la fin de l'année 2016, ce qui pourrait entraver l'accès aux données de l'enquête de Peace River. Je tiens donc à préciser que ces données étaient publiquement disponibles à des fins de recherche et d'enseignement au moment de les consulter en septembre 2015.

²⁸ Accessible sur <http://francoamericanarchives.org>.

²⁹ Geneviève Maheux-Pelletier, communication personnelle.

³⁰ Il n'y a pas de vidéo dans les archives qui a été enregistrée en 2011.

- le type d'enregistrement : seules les vidéos d'un format d'histoire orale (« oral history ») ont été retenues. Les conférences, concerts, performances préparées d'avance, etc. n'ont pas été retenus pour analyse.
- les langues de l'enregistrement : la plupart des vidéos dans les archives indiquent, parmi leurs métadonnées, la ou les langues utilisées dans l'enregistrement. Les vidéos indiquant n'utiliser que l'anglais ou le français ont été exclues parce qu'elles ne contenaient fort probablement pas d'occurrences d'alternance codique. Les vidéos pour lesquelles les langues n'étaient pas indiquées ont également été exclues.

Le tableau 3.1 présente les principales caractéristiques de chaque sous-corpus. Comme ce tableau l'indique, et malgré le descripteur « conversation libre » attribué au corpus PFC, les trois sous-corpus utilisent un protocole de conversation semi-dirigée. Dans le corpus Beaulieu, les conversations sont animées par l'enquêtrice, qui demande aux participants de discuter de leurs réponses à un questionnaire sur leurs attitudes linguistiques. Les éléments du questionnaire structurent et orientent donc les conversations. Quoique l'enquêtrice parle français et anglais, elle n'est pas membre de la communauté de pratique visée par l'enquête, et son statut d'enseignante du français à l'université fait d'elle une représentante, en quelque sorte, du français standard. De plus, son rôle d'enquêtrice fait en sorte qu'elle ne participe pas à l'interaction au même titre que ses interlocuteurs, et que ses interlocuteurs ne voient peut-être pas l'interaction tout à fait comme une conversation ordinaire de tous les jours.

Le volet « conversation libre » du PFC est animé par un ami ou un membre de la famille des participants de l'enquête, dans le but de les mettre à l'aise et de pouvoir ainsi observer leur prononciation lorsqu'ils relâchent leur auto-surveillance. Cependant, quoiqu'elle ne soit pas strictement guidée, la conversation n'est pas tout à fait libre non plus, dans la mesure où l'animateur a une liste de sujets à aborder avec son interlocuteur qui relèvent de sa vie quotidienne et de son vécu³¹. De plus, quoique l'animateur soit un membre de la même communauté de pratique que son interlocuteur, il ne participe pas à l'interaction au même titre que ce dernier, puisqu'il est conscient d'avoir la responsabilité de guider la conversation et d'inciter son interlocuteur à produire de la parole pour l'analyse phonologique. Comme avec le corpus Beaulieu, la situation de communication ne correspond pas à l'immédiat communicatif idéal ou prototypique, mais puisque les locuteurs sont membres de la même communauté de pratique et se connaissent personnellement, elle s'en rapproche généralement plus que dans le corpus Beaulieu. Toutefois, certains animateurs moins expérimentés avec ce genre d'enquête sont mal à l'aise avec leur rôle et créent de ce fait une situation encore moins naturelle.

³¹ Dans un des enregistrements analysés, on peut même entendre l'animatrice consulter cette liste pour trouver de nouveaux sujets de discussion.

TABLEAU 3.1 : CARACTÉRISTIQUES DES SOUS-CORPUS

	Corpus Beaulieu	Corpus PFC	Corpus FAC
Lieu d'enquête	Alberta (Edmonton et Peace River)	Alberta (région de Peace River)	Maine (plusieurs régions, principalement le comté d'Aroostook)
Objectif de l'enquête	recueillir des données qualitatives en lien avec un protocole de locuteur masqué	documenter des traits phonologiques de la variété de français parlée dans cette région	documenter l'histoire et la culture de la communauté francophone à travers le vécu de ses membres
Protocole d'enquête (dans les enregistrements analysés)	conversation semi-dirigée (animée par l'enquêtrice)	conversation semi-dirigée (animée par un proche des participants)	conversation semi-dirigée (animée par des bénévoles du Franco-American Center)
Date d'enregistrement	2011	2001	2001-2012
Population(s) enregistrée(s)	résidents de foyers pour personnes âgées, personnel médical	personnes de plusieurs tranches d'âge (selon le protocole du PFC)	personnes âgées
langue dominante³²	français	français	mixte
type d'alternance dominant³³	brève (insertionnelle)	brève (insertionnelle)	longue
Type d'enregistrements	audio	audio	vidéo
Nombre de fichiers audio ou vidéo	31	8	17
Heures d'enregistrement	19	2	31
Informations disponibles au préalable sur les enregistrements	durée d'enregistrement; année d'enregistrement; sexe, nombre et type d'interviewés (résident ou personnel)	durée d'enregistrement; année d'enregistrement; âge(s), sexe et nombre d'interviewés	durée d'enregistrement; année d'enregistrement; sexe et nombre d'interviewés

Le corpus FAC se rapproche le plus de la situation idéale de l'immédiat communicatif, en ce qu'il n'y a pas de sujets précis préparés à l'avance, les locuteurs de chaque enregistrement se connaissent personnellement et sont, sauf exception³⁴, membres de la même communauté de pratique. De plus, les locuteurs participent à des entretiens dans le cadre d'un projet d'histoire communautaire, plutôt que dans le cadre d'une étude scientifique. La situation est donc encore plus décontractée que dans les autres sous-corpus et il y a moins de pression de « bien répondre » aux questions de l'animateur. Cependant, il y a effectivement un animateur pour chaque conversation, c'est-à-dire que les locuteurs ne participent pas tous à l'interaction au même titre. Bien que le choix de thématique soit plus libre que dans les autres sous-corpus, cet animateur est conscient que son rôle est de guider la conversation vers des sujets d'histoire locale et personnelle, sur lesquels porte le projet. Pour chaque sous-corpus, donc, il y a des facteurs qui imposent plus ou moins de limites à la

³² Globalement, selon ce qui est le plus fréquent dans les interactions observées.

³³ Globalement, selon ce qui est le plus fréquent dans les interactions observées.

³⁴ Dans l'échantillon analysé (cf. la section 3.3) il y a deux exceptions à noter. Il s'agit en premier lieu de l'enregistrement FAC 02, dont l'interviewé n'habite plus à temps plein au Maine, et ne parle pas du tout français pendant son entretien, ce qui représente un écart par rapport aux pratiques langagières des autres locuteurs du même corpus. En second lieu, la locutrice principale de l'enregistrement FAC 09, qui raconte son vécu directement et agit parfois comme intervieweuse envers ses proches, apprend le français en tant que langue d'héritage et cherche activement à rapprocher ses pratiques langagières de celles des membres plus expérimentés de la communauté, ce qui indique qu'elle ne les partage pas entièrement encore.



situation de communication, l'éloignant plus ou moins de l'immédiat communicatif idéal (cf. la section 3.1.1.2) et influant nécessairement sur les usages qu'on y trouve. Toutefois, tous les sous-corpus ont été enregistrés dans des situations relevant suffisamment de l'immédiat communicatif pour être propices à l'alternance codique, comme en témoigne le nombre d'occurrences relevé (cf. le tableau 4.1 à la section 4.3).

3.2.2 Contexte sociohistorique pertinent des communautés en question

Les trois sous-corpus proviennent de deux grandes régions où le français est une langue minoritaire : l'Alberta et le Maine. Chacune de ses régions a un contexte sociohistorique singulier qui influe sur les pratiques et les attitudes linguistiques ainsi que sur les parcours référentiels des locuteurs enregistrés; un survol de ces contextes est donc de mise afin de compléter la description des sous-corpus.

3.2.2.1 *L'Alberta francophone*

Le français est présent en Alberta depuis le 18^e siècle, quoique les auteurs diffèrent sur l'origine de ces premiers francophones et de la proportion de la population qu'ils représentaient (Allaire, 2014; Mougeon, 1998; Walker, 2012). Quoiqu'il y ait certainement eu une période, à l'époque de la traite des fourrures, où les francophones étaient majoritaires sur le territoire, les personnes « d'origine française³⁵ » ne constituaient déjà en 1885 que 10,1 % de la population des Territoires du Nord-Ouest (qui comprenaient le territoire qui, en 1905, allait devenir l'Alberta), selon Statistique Canada (Allaire, 2014).

Aux 18^e et 19^e siècles, la francophonie de l'Ouest « [était] réunie par la religion, la langue et, pour la grande majorité, la ruralité » (Allaire, 2014, p. 37). La vie sociale en particulier était organisée autour des paroisses et de la vie religieuse : l'église, les fêtes, les rites. Les paroisses et les congrégations religieuses s'occupaient également de l'éducation, autant primaire que secondaire et post-secondaire. Les communautés étaient essentiellement autonomes et pouvaient établir leurs propres écoles; la scolarisation dans des communautés francophones se faisait donc en français (Mougeon, 1998). Cependant, après la Confédération du Canada, il y a eu un mouvement visant « à définir le dominion du Canada et l'Ouest en particulier comme britannique » (Allaire, 2014, p. 36). L'infrastructure gouvernementale (anglophone) a pris de l'expansion et les systèmes de santé et d'éducation y ont été incorporés (Mougeon, 1998). Notamment, en 1892, les Territoires du Nord-Ouest faisaient de l'anglais « la seule langue de la législature et des tribunaux », et un peu plus tard, la seule langue de scolarisation, mis à part les toutes premières années de la petite école dans des communautés francophones (Allaire, 2014, p. 36). Cela a eu pour effet non seulement de réduire l'exposition

³⁵ Cette appellation tirée des sondages démographiques comprend tous les individus de souche francophone, qu'ils étaient originaires de la France ou non.

au français et de restreindre ses domaines d'emploi à la faveur de l'anglais, mais aussi de porter atteinte à la performance, au rendement et aux aspirations scolaires des élèves et étudiants francophones (Mougeon, 1998).

Face à cette volonté de dominance anglophone, le Canada français, au début du 20^e siècle, prône *La Survivance*, une stratégie culturelle et politique misant sur la solidarité. Le Québec et les communautés minoritaires renforcent leurs liens sociaux, politiques et financiers. Dans ce contexte, les Franco-Albertains créent en 1926 l'Association canadienne-française de l'Alberta (ACFA) dans le but de préserver et de valoriser leur langue et leurs traditions. Cet organisme met sur pied un concours annuel de français et veille à la publication de l'hebdomadaire de langue française *La Survivance*, entre autres projets. Plus tard, l'ACFA fait la promotion des coopératives financières et commerciales dans les communautés francophones, et lutte pour qu'il y ait une programmation radio en français. Tout cela avait pour but d'illustrer « la volonté des communautés de suivre le courant », tout en montrant leur appartenance au Canada français (Allaire, 2014, p. 44).

Après la Deuxième Guerre mondiale, les communautés francophones de l'Alberta ressentent les impacts de la modernisation. L'agriculture se mécanise, la population s'urbanise, les réseaux de transport s'améliorent. La voiture devient de plus en plus accessible, ainsi que la télévision (et avec elle, la culture anglo-américaine). Tout cela défait l'isolation — et l'identité — rurale. En même temps, la pratique religieuse baisse, ainsi que les effectifs des congrégations religieuses, et le rôle social de l'Église diminue. Notamment, plusieurs écoles des communautés francophones ferment ou passent au système public, réduisant ainsi l'accès des jeunes au peu d'enseignement en français et du français qui leur était permis (Allaire, 2014).

Au milieu du 20^e siècle, la population francophone constitue autour de 6 % de la population de la province, et le taux de transmission du français comme langue d'usage d'une génération à une autre varie entre 50 et 60 % (Allaire, 2014). Un sondage ethnographique mené en 1984 par Beniak, Carey et Mougeon témoigne d'ailleurs de ce faible taux de transmission : les locuteurs francophones nés deux générations auparavant (donc autour des années 1940) rapportent que leurs enfants étaient bilingues et que la plupart de leurs petits-enfants étaient anglo-dominants ou unilingues anglais (Beniak, Carey, et Mougeon, 1984). Selon Mougeon (1998), cette non-transmission serait due au moins en partie à l'exogamie, le mariage entre francophones et anglophones faisant en sorte que la langue parlée à la maison et la langue d'usage des enfants issus de ces unions serait plutôt l'anglais (voir aussi à ce sujet Daveluy, 2008). Cependant, Mougeon (1998) souligne que la non-transmission du français n'est pas uniquement le résultat d'un accommodement en faveur du conjoint anglophone — entre autres, elle a également lieu dans des mariages entre francophones. Il s'agit plutôt d'un choix délibéré dans lequel entrent en jeu des facteurs économiques et sociaux comme la valeur instrumentale ou bien le prestige des langues en question.

Dans la deuxième moitié du 20^e siècle, on voit le Québec se dissocier des communautés francophones de l'Ouest. Dans le but de se montrer un état indépendant, le Québec mise sur son identité distincte — y compris des autres communautés francophones au Canada, qui, pour leur part, n'approuvent pas l'idée de séparation et d'indépendance politique (Allaire, 2014; Mougeon, 1998). Ainsi, la Révolution tranquille voit « la grande rupture, voire la fin du Canada français » (Allaire, 2014, p. 47). En même temps, elle fait en sorte que le gouvernement fédéral, et dans le même courant, les gouvernements provinciaux, commencent à s'occuper des questions des langues officielles et des droits linguistiques des francophones. En 1968, le gouvernement provincial de l'Alberta accorde aux francophones le droit à l'enseignement en français pendant jusqu'à la moitié du temps d'instruction. Les droits et services accordés aux Franco-Albertains continuent à se bonifier jusqu'à la fin du siècle : le droit à un procès judiciaire dans la langue officielle de sa préférence (l'arrêt Beaulac); le droit à l'éducation publique en français pour des élèves de souche francophone; la formation de conseils scolaires francophones (l'arrêt Mahé); l'établissement de réseaux de santé francophones et des programmes de formation en français de professionnels de la santé et de l'enseignement, etc. (Allaire, 2014). Du côté du gouvernement fédéral, on reconnaît et encadre officiellement les francophones canadiens, notamment par la *Loi sur les langues officielles* (dont une première version a été adoptée en 1969), mais aussi par d'autres mesures comme l'exigence que l'emballage des produits commerciaux soit traduit en français et que tout matériel gouvernemental imprimé soit disponible en français et en anglais (Allaire, 2014; Mougeon, 1998). Malgré cela, les perspectives ne sont pas tout à fait réjouissantes : il faut lutter pour la mise en pratique de certains de ces droits, le taux de transmission reste dangereusement faible, et « le spectre de l'assimilation [fait] poser beaucoup de questions et couler beaucoup d'encre » (Allaire, 2014, p. 51).

C'est aussi à cette époque, à partir des années 1970, que des programmes d'immersion française voient le jour à travers le pays, programmes qui « [vont] modifier en profondeur les provinces de l'Ouest et rapprocher l'Autre anglo-canadien du francophone » (Allaire, 2014, p. 53). L'immersion présente l'apprentissage du français comme désirable pour les anglophones, augmentant ainsi le prestige du français. Dans l'Ouest, bon nombre d'anglophones issus de ces programmes « se sont intégrés à la communauté [francophone] et, dans plusieurs cas, en sont devenus des promoteurs et de grands défenseurs » (Allaire, 2014, p. 59)³⁶.

³⁶ Il est à noter que ce « bon nombre » est relatif : dans bien des contextes, on observe que la plupart des diplômés des programmes d'immersion utilisent peu ou pas le français dans leur vie adulte (voir par exemple Husum et Bryce, 1991 pour la Saskatchewan; Vanderveen, 2015 pour la région de Toronto; Wesche, 1993 pour la région d'Ottawa). Cependant, la littérature indique que la minorité qui continue à utiliser le français d'une manière ou d'une autre est non négligeable (par exemple, 33 % des participants dans l'étude de Vanderveen). De plus, on observe que de façon générale les diplômés de ces programmes perçoivent positivement le fait de connaître le français, pour les possibilités d'emploi et de voyage que cela offre et pour les liens que cela permet de tisser avec la communauté francophone au Canada – ce qui va dans le même sens que l'affirmation d'Allaire (2014).

Finalement, pendant cette période, l'Église perd son statut central dans la structuration de la vie sociale des communautés francophones au profit des activités civiles et culturelles. Les écoles bilingues ou francophones, les centres communautaires, les festivals annuels et les compétitions sportives (notamment les Jeux franco-albertains) et musicales sont devenus les principaux lieux de célébration et de renforcement de l'identité francophone à la place des paroisses (Allaire, 2014).

Aujourd'hui, la francophonie de l'Ouest est caractérisée par une diversité grandissante. La population francophone en Alberta ne constitue que 2,1 % de la population, soit environ 67 000 personnes (Allaire, 2014; Walker, 2012). Cependant, cette population comprend non seulement les Franco-Albertains « de souche », mais des anglophones issus des programmes d'immersion, de nouveaux immigrants du Nouveau-Brunswick et du Québec — ce dernier « [s'étant] davantage rapproché de la francophonie canadienne » depuis les années 1960 (Allaire, 2014, p. 58) — et des immigrants de l'extérieur du pays, notamment des pays francophones de l'Afrique. Ces immigrants sont attirés par l'essor de l'industrie pétrolière et des domaines connexes (industrie et services), ainsi que par la demande augmentée pour des professionnels d'enseignement et de santé bilingues. Quoique son héritage agricole soit encore très présent (Walker, 2012), la communauté francophone en Alberta s'est diversifiée aussi dans ses occupations : « commerce et affaires, services hôteliers et hospitaliers, fonction publique, politique, enseignement, divertissements, écriture, exploitation des ressources naturelles » (Allaire, 2014, p. 60).

L'Alberta francophone est aussi et toujours caractérisée par la volonté non seulement de survivre, mais aussi de perdurer (« lasting »; Daveluy, 2008). Cette volonté implique un fort sentiment de responsabilité individuelle de connaître et de valoriser ses droits en tant que francophone au Canada. En particulier, les « ayants-droits » (des jeunes pouvant accéder à l'enseignement public en français en raison de leur héritage) sont socialisés à considérer leur statut comme un privilège et une responsabilité envers la communauté et la langue (Daveluy, 2008). Cette même volonté de survivre et de perdurer implique aussi d'entretenir des relations non conflictuelles avec la majorité anglophone, ce qui touche aux normes de quand et comment employer le français : « speaking French that does not blend in the political context may trigger repercussions and jeopardize the achieved state of equilibrium » (Daveluy, 2008, p. 34). On pourrait être tenté d'y voir des traits du « caractère national » du Canada (des traits relevés notamment par Lipset, par exemple 1990) : la subordination des intérêts de l'individu aux intérêts du groupe, le respect pour l'État de droit, la réserve et l'intolérance envers le conflit et le désordre civil. Cependant, il faut prendre cet apparent lien avec beaucoup de précaution. Non seulement les conclusions de Lipset quant au caractère des Canadiens ont été remises en cause par des études empiriques subséquentes (Baer, Grabb, et Johnston, 1990; Baer *et al.*, 1995; Raney, 2010), mais la validité de la notion même de caractère national est discutable : selon McCrae et Terracciano (2006), quoiqu'on peut souvent trouver un consensus d'opinion sur le caractère des citoyens d'un pays donné, cette opinion ne

correspond pas nécessairement aux véritables traits de personnalité de ces citoyens. Plus important encore, même dans la mesure où l'on peut parler de traits caractéristiques des Canadiens, on doit se demander à quel point la communauté francophone d'Alberta a une identité « typiquement » canadienne. Ses membres font partie d'une minorité linguistique qui a vécu des relations changeantes, et pas toujours positives, avec l'État canadien (le gouvernement provincial et fédéral) et avec le Québec, se retrouvant ainsi à la périphérie non seulement par rapport au centre anglophone, mais aussi par rapport au centre francophone. Cette idéologie de « non-confrontational persistence » (Daveluy, 2008, p. 33) est bien le produit des conditions historiques particulières aux francophones de l'Ouest canadien : la situation linguistique des Franco-Albertains est beaucoup plus précaire que celle des francophones du Québec et même du Nouveau-Brunswick, mais ne comprend pas le bagage d'expériences traumatisantes avec la majorité anglophone de l'ordre du Grand Déplacement en Acadie (Daveluy, 2008) ni de l'ordre de la xénophobie vécue au Maine (cf. la section 3.2.2.2). Il leur est donc à la fois possible et nécessaire de coexister sans trop d'antagonisme.

Sur le plan linguistique, le sentiment de responsabilité envers sa propre communauté et l'approche de non-confrontation avec la communauté majoritaire mènent à une pratique communautaire d'emploi stratégique des deux langues. Il faut effectivement se prévaloir de ses droits linguistiques et des occasions d'afficher et de promouvoir l'identité franco-albertaine, non seulement pour son propre bénéfice, mais aussi pour assurer que les gains actuels soient maintenus pour d'autres qui vont suivre. En même temps, la connaissance de l'anglais (et la volonté de l'utiliser quand la situation l'exige) est considérée nécessaire afin de bien s'intégrer et de s'accommoder au reste de la société, et de ne pas être une source de problèmes ou de conflit (Daveluy, 2008; sentiment également exprimé par des participants dans Beaulieu, 2012). Cependant, on voit également l'héritage de la stratégie de *La Survivance*, en ce que le français a sa place, et l'anglais la sienne. Cet héritage s'observe non seulement dans l'attitude exprimée par certains locuteurs qu'un « bon français » ne comporte pas de mots anglais (Beaulieu, 2012), mais aussi à travers les pratiques langagières relatives à l'alternance codique, qui seront présentées plus en détail à la section 4.3.

3.2.2.2 *Le Maine francophone*

L'histoire du français au Maine commence avec des Acadiens qui s'installent dans la vallée de la Rivière St-Jean au 18^e siècle. Plus tard, lors de l'établissement définitif de la frontière entre les États-Unis et le Canada, ceux qui s'étaient installés sur la rive sud se trouvent dorénavant sur le territoire américain. Cette région reste relativement autonome et isolée par la suite, sa population augmentant en raison d'un taux de naissance élevé (Allen, 1974). Au 19^e siècle, le Maine accueille des francophones d'une deuxième source : l'immigration québécoise vers les villes industrielles du sud de l'État. Cette immigration, due à un manque de terres cultivables et d'emplois industriels pour soutenir la population croissante du Québec, ne ralentira que dans les années 1920

et 1930 (Richard, 2009). « By the onset of the Great Depression, nearly one million French Canadians had migrated to Maine and the northeastern United States » (Richard, 2009, p. 286).

Comme en Alberta, les communautés francophones du Maine se sont maintenues pendant longtemps en restant entre elles. Les francophones s'organisent en paroisses et envoient leurs enfants à des écoles paroissiales. La démarcation géographique entre francophones et anglophones dans le comté d'Aroostook (qui comprend la vallée de la St-Jean) se fait naturellement en raison de l'isolement rural, mais même dans les villes industrielles, les francophones établissent des « Petits Canadas ». Au travail, aussi, ils sont souvent avec d'autres francophones dans les usines de coton, de souliers ou de papier, ou bien dans les camps de bûcherons (Allen, 1974; Richard, 1998). Cette tendance à se tenir dans des environnements avec une haute concentration de francophones est due en partie aux circonstances géographiques et économiques, mais elle fait également partie de la stratégie de *La Survivance*, le Maine francophone se considérant et étant considéré comme faisant partie des communautés canadiennes-françaises. *La Survivance* relie la langue, la religion et la culture francophone comme composantes fondamentales d'une société bien distincte, et prône la solidarité entre et à l'intérieur des communautés francophones comme la clé de sa préservation. Tout cela a contribué au fait que les Franco-Américains ont conservé leur langue, leur religion et leurs traditions culturelles beaucoup plus longtemps que d'autres groupes ethniques immigrants aux États-Unis (Pinette, 2002). Aussi récemment que dans les années 1970, certaines communautés étaient toujours presque à 100 % francophones (Allen, 1974)³⁷.

Pendant une longue période de leur histoire, il n'est toutefois pas possible de connaître avec précision le taux de transmission linguistique dans ces communautés, ni même la proportion de francophones au Maine, en raison d'un manque de données démographiques. Le recensement fédéral américain n'a posé aucune question sur la langue maternelle des répondants (autre que celle des immigrants récents) avant 1940; aucune question sur les langues actuellement parlées par les répondants n'a été posée avant 1990 (Allen, 1974 ; United States Census Bureau, 2016). Toutefois, la proportion de catholiques français identifiés dans un recensement privé de 1908 et la proportion de personnes disant avoir le français comme langue maternelle dans le recensement fédéral de 1970 sont relativement semblables (Allen, 1974). On peut en déduire que le taux de transmission est relativement élevé pendant cette période, quoique Allen remarque en 1974 que les jeunes générations sont de plus en plus bilingues, ce qui pourrait indiquer le début d'une transition vers l'anglo-dominance.

³⁷ Encore de nos jours, ces mêmes communautés — Eagle Lake, Fort Kent, Madawaska et Van Buren — sont de 37 % à 64 % francophones, et la ville de Lewiston est à 15 % francophone. Bien sûr, la majorité de la population francophone dans ces endroits relève de la portion la plus âgée de la population, tandis que relativement peu de jeunes disent parler le français à la maison (United States Census Bureau, 2015).

Comme ce taux de transmission semble l'indiquer, les Franco-Américains ont de façon générale longtemps perçu la langue française (ainsi que la religion catholique) comme étant un élément essentiel à la préservation de leur identité (Allen, 1974; Richard, 1998). En même temps, et particulièrement dans les villes industrielles, l'assimilation culturelle et le maintien d'une identité ethnique distincte ne sont pas considérés comme des objectifs nécessairement incompatibles jusqu'au milieu du 20^e siècle³⁸. En particulier, ne pas avoir de bonnes connaissances en anglais entraîne souvent de l'intimidation et la discrimination à l'école et au travail, limite le rendement et les aspirations scolaires, restreint les opportunités de travail et entrave généralement la mobilité sociale ascendante (Richard, 1998 ; St-Pierre, 2011). Ainsi, dans certaines paroisses, l'apprentissage de l'anglais est accepté et même préconisé dès les premières décennies du 20^e siècle, au point où l'école paroissiale enseigne l'anglais, sans qu'il y ait des préoccupations quant au maintien du français (Richard, 1998).

Par contre, le sentiment ultra-patriotique courant pendant et après les deux guerres mondiales amplifie et popularise les attitudes xénophobes déjà présentes aux États-Unis. Si, avant les guerres, le fait de parler une langue autre que l'anglais n'a pas de place dans la conception dominante du caractère du pays³⁹, dans les années 1920 et au-delà, il est devenu nettement anti-américain de le faire (Wiley, 2014). En 1922, 34 États avaient procédé au « hasty passage of official English policies » (Wiley, 2014, p. 26). Le Maine en fait partie : en 1919, l'anglais est déclaré la langue d'instruction dans l'ensemble des écoles de son territoire (Richard, 1998). Comme en Alberta, ces lois ont eu pour résultat de faire diminuer l'exposition au français en faveur de l'anglais non seulement à l'école, mais aussi à la maison, car la fréquentation de l'école publique favorise l'assimilation et les mariages exogames (Allen, 1974). Ces lois ont aussi eu pour effet de faire des langues autres que l'anglais des symboles encore plus saillants de l'étranger et du supposé anti-américanisme (Wiley 2014). Les réactions anti-immigrantes contre les Franco-Américains s'intensifient : harcèlement, intimidation, ridiculisation, discrimination et exclusion dans toutes les sphères de la vie. Le point focal est souvent le fait de parler français ou de « mal » parler l'anglais (Albert *et al.*, 2013; Richard, 1998). Les Franco-Américains se trouvent même dans la mire du Ku Klux Klan, selon lequel ils sont sans légitimité aux États-Unis. Plus particulièrement, le Ku Klux Klan émet des menaces, fait brûler des croix et exploser des bombes dans le but d'intimider les Franco-Américains, en particulier lors des activités politiques, religieuses ou civiles entreprises par ces derniers (Richard, 2009).

Dans ce climat hostile, l'assimilation culturelle et le maintien d'une identité ethnique distincte sont devenus des objectifs manifestement incompatibles, particulièrement en ce qui a trait à la langue. Le maintien

³⁸ Ce qui n'est pas sans rappeler la stratégie actuelle des Franco-Albertains : être un promoteur et défenseur de son héritage, tout en s'intégrant dans la culture dominante de façon à ne pas attirer de l'attention indésirable.

³⁹ Idée qu'on retrouve tant chez Benjamin Franklin au 18^e siècle (Wiley, 2014) que chez les membres du Ku Klux Klan au 20^e siècle (Richard, 2009).

de la culture était une affaire interne à la communauté franco-américaine, se passant en famille – par exemple, le choix de servir des plats traditionnels ou de respecter des traditions religieuses comme le Réveillon – et au sein de l'église et des organismes communautaires (Richard, 1998). Cependant, le fait d'être un locuteur non natif de l'anglais met en évidence l'identité distincte des Franco-Américains dès lors qu'ils sortent de la communauté francophone. Ainsi, il y avait beaucoup de pression non seulement pour parler l'anglais, mais aussi pour *ne pas parler* le français (Allen, 1974; Albert *et al.*, 2013; Richard, 1998; Richard, 2009). Beaucoup de familles ont fait le choix délibéré de ne pas transmettre le français aux enfants, non seulement pour leur conférer des avantages scolaires, sociaux et économiques, mais aussi pour les protéger de discrimination, d'intimidation et du « handicap » d'être francophone (voir les commentaires à l'annexe D de Albert *et al.*, 2013). Dans la dernière moitié du 20^e siècle, il n'y a plus la masse critique de locuteurs du français nécessaire pour assurer la vitalité de la langue dans beaucoup de communautés, et même les organismes communautaires traditionnellement francophones mènent leurs activités de plus en plus en anglais afin d'attirer et de retenir de nouveaux membres, en particulier des jeunes (Allen, 1974 ; Richard, 1998 ; St-Pierre, 2011).

Ironiquement, *La Survivance*, qui a bien aidé à soutenir le français au Canada, a plutôt contribué à son déclin dans le Maine. Malgré son apport positif au départ, l'insistance sur une culture homogène commune à toute la francophonie de descendance canadienne « hindered the ability of Franco-Americans to address the specific pressures that their community faced in American society » en se fiant trop sur des stratégies « that no longer appealed to a secularized and increasingly Anglophone population » (Pinette, 2002, p. 11).

En 1969, la loi sur la langue d'éducation au Maine a été amendée pour permettre l'enseignement bilingue dans certains cas, notamment l'enseignement transitionnel dans des régions où beaucoup d'élèves ne parlaient pas anglais à la maison. Cela coïncide avec un renouvellement d'intérêt pour le bilinguisme aux États-Unis et le renouvellement parallèle d'intérêt chez beaucoup de Franco-Américains pour leur héritage et leur langue (Allen, 1974). Cependant, on voit que l'objectif de cet amendement n'est pas tout à fait la promotion des langues autres que l'anglais, mais de l'accommodement de celles-ci en vue de faciliter l'assimilation. Encore aujourd'hui, le français est pris en compte au sein du système scolaire en tant que langue étrangère, et non comme une langue d'instruction (The Task Force on Franco-Americans, 2012, annexe D).

Quoiqu'il n'y ait pas eu de gains significatifs dans les droits linguistiques des francophones au Maine, le gouvernement de l'État établit ponctuellement des comités de réflexion sur la communauté franco-américaine et sur des questions sociales et économiques qui touchent les Franco-Américains (The Task Force on Franco-Americans, 2012). Il y a de plus en plus de ressources communautaires « that focus on promoting and preserving Franco-American language and culture » (The Task Force on Franco-Americans, 2012, p. 2) : des archives historiques et généalogiques, des groupes d'arts de la scène, des festivals patrimoniaux, des programmes

universitaires d'études franco-américaines et des programmes de langue d'origine (*heritage language programs*). Malgré cela, on peine toujours à reconnaître le caractère distinct de la culture franco-américaine. Par exemple, dans le recensement fédéral américain, les personnes s'identifiant comme franco-américaines, acadiennes ou québécoises sont toutes regroupées sous la dénomination *French Canadian*, et les personnes qui s'identifient comme *French* (qui peut signifier autant « Français » que « francophone » dans le contexte américain) sont comptées dans une autre catégorie (Albert *et al.*, 2013). On estime que les Franco-Américains constituent environ 24 % de la population du Maine, mais ils demeurent une « minorité silencieuse » dont les contributions historiques et économiques à l'état sont méconnues (Pinette, 2002).

On peine également à caractériser la communauté franco-américaine, quoiqu'on affirme qu'une telle communauté existe (Albert *et al.*, 2013; The Task Force on Franco-Americans, 2012). Même si la langue française est considérée « a determining mark of culture » qui définit « [Franco-Americans'] cultural space in contrast to English-speaking institutions » (Pinette, 2002, p. 10), l'appartenance à cette communauté est, comme en Alberta, plus complexe que le simple fait de parler français. Notamment, il n'est pas nécessaire d'avoir des compétences en français pour s'identifier comme Franco-Américain : dans un sondage de l'American Community Survey, seulement 58,1 % des personnes s'identifiant ainsi ont dit qu'il est « très important » ou « plutôt important » pour un Franco-Américain de parler français, et seulement 28,5 % des mêmes répondants ont indiqué qu'ils parlent français couramment⁴⁰ (Albert *et al.*, 2013). En fait, l'identité franco-américaine se rattache à plusieurs marqueurs⁴¹, mais aucun n'est définitif. Ceux identifiés comme les plus importants sont les réunions familiales, la nourriture, et l'héritage familial (Albert *et al.*, 2013), reflétant la place qu'a prise la sphère de la famille dans la transmission de la culture au début et au milieu du 20^e siècle (Richard, 1998).

Il n'en reste pas moins que la majorité des Franco-Américains considèrent le français une composante importante de leur identité. Malheureusement, il ne semble pas exister de données claires sur les préférences des Franco-Américains du Maine quant à la norme linguistique. Il est donc impossible de dire à quel point ils valorisent un « bon français » et pourquoi, ni ce à quoi ce bon français pourrait ressembler selon eux. Fox (2007) relate que, au moins pour ses répondants francophones du Massachusetts, certains sont d'avis qu'un bon français ne doit pas comporter de mots anglais ou être mélangé avec l'anglais. Cependant, ces mêmes

⁴⁰ Comme la question permettait plus d'une réponse, il est impossible d'additionner les réponses pour connaître le pourcentage total de répondants qui ont déclaré avoir des compétences en français. Toutefois, le comité de réflexion de 2012 rapporte que, dans un autre sondage, 63 % des répondants ont affirmé parler au moins un peu français (The Task Force on Franco-Americans, 2012).

⁴¹ Au moins dans les cas où les répondants étaient capables d'en identifier – 20,2 %, d'entre eux n'ont pas pu penser à un marqueur spécifique de l'identité franco-américaine (Albert *et al.*, 2013). Il semble y avoir « a bifurcation of the Franco-American experience » (The Task Force on Franco-Americans, 2012, p. 6), avec au moins une partie de la communauté se sentant « déracinée » des marqueurs traditionnels comme la langue, la religion, les convictions politiques, etc. (Albert *et al.*, 2013 ; The Task Force on Franco-Americans, 2012).

répondants rapportent des réactions négatives à cette attitude de la part des autres locuteurs francophones de la communauté. Ces réactions comprennent notamment le refus de parler à ces répondants en français, ou une réticence à le faire. Il semble donc que la plupart des membres de la communauté s'attachent à une variété locale du français qui n'est pas trop restrictive quant à l'incorporation des ressources en anglais. On peut présumer que des attitudes semblables existent aussi au Maine, étant donné : l'absence d'exposition à un français dit standard (c'est-à-dire l'absence de scolarisation en français); le potentiel d'association de ce français standard à des francophones non membres de la communauté locale; ainsi que l'échec, ultimement, de *La Survivance*, dont le manque d'adaptation aux circonstances locales et l'insistance sur l'homogénéité (c'est-à-dire la standardisation) linguistique et culturelle ont plutôt nui à la vitalité du français. Les données de mon corpus vont également dans ce sens (cf. les sections 4.3 et 5.3.2).

Les répondants de Fox (2007) rapportent également qu'il y a des espaces appropriés pour le français (en privé et à l'intérieur de la communauté) et des espaces appropriés pour l'anglais (en public et en présence d'au moins un anglophone). Ces réponses correspondent à ce qu'on observe dans les corpus examinés de l'Alberta. Cependant, contrairement à l'Alberta, il semble que pour les Franco-Américains les espaces appropriés pour le français ne sont pas nécessairement des espaces *inappropriés* pour l'anglais; cette ambivalence s'observe non seulement dans les données de Fox (2007), mais aussi dans le corpus analysé pour la présente étude (cf. la section 4.3).

3.3 Constitution du corpus d'analyse

L'ensemble des trois corpus représentant une cinquantaine d'heures d'enregistrement non transcrites, il n'a pas été possible d'en étudier la totalité en profondeur dans le cadre de la présente étude. Le corpus d'analyse est dans les faits constitué d'un échantillon représentant environ 50 % de cet ensemble.⁴² Les sous-corpus étant très hétérogènes, notamment pour ce qui est des contextes d'interaction (nature de l'enquête, contexte sociohistorique de la communauté en question, etc.) et des expériences, habitudes et préférences de chaque locuteur, un échantillonnage complètement aléatoire n'aurait pas convenu. Les enregistrements dans l'échantillon ont donc été sélectionnés en fonction de leur représentativité par rapport aux informations disponibles avant de les écouter⁴³.

Pour le corpus Beaulieu, l'échantillon a été sélectionné d'abord en fonction des trois sous-groupes de locuteurs : les résidents des foyers pour personnes âgées interviewés individuellement (CB-IND), les résidents

⁴² Dans la mesure du possible, autant 50 % du nombre total d'enregistrements que 50 % de la durée totale de chaque sous-corpus a été retenu.

⁴³ Ces informations peuvent être exactes ou non dans la réalité (par exemple, l'un des entretiens du corpus FAC a été mené entièrement en anglais malgré l'indice « French and English » qui lui a été attribué).

interviewés en groupe (CB-GRP), et le personnel médical (seul ou en groupe; CB-PRS). Approximativement 50 % du contenu de chaque sous-groupe a été sélectionné, en tenant compte de la distribution du sexe des interviewés et de la durée de l'enregistrement. Pour les entretiens en groupe, la taille du groupe a aussi été prise en compte. En tout, 16 enregistrements, totalisant 9,5 heures, ont été sélectionnés.

Pour le corpus PFC, la représentativité de l'échantillon a été déterminée en fonction de la distribution du sexe et de l'âge des interviewés. La durée d'enregistrement est égale dans tous les cas, selon le protocole du projet PFC. Quatre enregistrements, soit une heure au total, ont été retenus. Force est d'admettre que ce sous-corpus est de taille substantiellement plus petite que les autres sous-corpus; toutefois, il contient une quantité non négligeable d'occurrences d'alternance codique – en fait, plus d'occurrences par heure en moyenne que le corpus Beaulieu (*cf.* la section 4.3).

Pour le corpus FAC, 50 % du sous-corpus présélectionné (*cf.* la section 3.2.1) a été retenu en fonction de la distribution dans le sous-corpus du sexe des interviewés et de leur nombre (seul ou en couple), des années d'enregistrement entre 2001 et 2011 et de la durée des enregistrements. Ce sous-corpus compte un total de 9 enregistrements, soit 15 heures enregistrées. Le corpus d'analyse comporte donc 29 enregistrements en tout et est d'une durée totale de 25,5 heures.

3.4 Récapitulatif du chapitre 3

Ce chapitre a porté sur la sélection des corpus dont le dépouillement fait l'objet du prochain chapitre. Un ensemble de critères de sélection a été établi dans le but d'assurer que le corpus d'analyse contienne suffisamment d'occurrences d'alternance codique pour que l'analyse puisse être fructueuse. Ces critères sont : un corpus existant, sur support audio ou visuel, contenant des interactions libres ou semi-dirigées entre locuteurs bilingues anglais-français dans des communautés de français minoritaire en Amérique du Nord. De plus, il a fallu notamment établir une définition fonctionnelle du bilinguisme, afin de pouvoir définir ce qui constitue un « corpus d'interaction entre locuteurs bilingues ».

Ensuite, les caractéristiques des trois corpus sélectionnés selon ces critères ont été présentés : le corpus PFC, le corpus Beaulieu et le corpus FAC. Comme la recontextualisation permettant de prendre en compte les perceptions des locuteurs implique non seulement le contexte étroit ou le cotexte, mais aussi le contexte large, un résumé du contexte sociohistorique des régions de français minoritaire d'où provient ces corpus – en l'occurrence l'Alberta et le Maine – a également été présenté. Ce chapitre ayant couvert la sélection et la contextualisation des sous-corpus analysés, le prochain chapitre se concentre sur les choix méthodologiques nécessaires pour procéder à l'identification et la catégorisation des occurrences d'alternance codique dans le corpus d'analyse.

4 Pré-analyses : repérage d'occurrences, catégorisation et transcription

Une fois le corpus d'analyse délimité, la prochaine étape était de faire la préparation nécessaire pour pouvoir mener une analyse en contexte des unités lexicales faisant l'objet d'une alternance codique. En ce qu'elle repose sur des choix méthodologiques qui ont une incidence sur les données prises en considération, et donc ultimement sur les résultats qui en ressortent, cette préparation représente une analyse préalable – en fait un ensemble d'analyses préalables – du corpus. Ces pré-analyses consistent à déterminer ce qui sera considéré une occurrence d'alternance codique (section 4.1) et les types d'occurrences qui seront ciblés (section 4.2), repérer les occurrences dans le corpus d'analyse (section 4.3), identifier des catégories d'unités fréquentes parmi les occurrences ainsi relevées (section 4.4), et transcrire et annoter les occurrences relevant de ces catégories (section 4.5).

4.1 Définition pratique de la notion d'occurrence

L'alternance codique a déjà été définie à la section 2.2.1 comme la juxtaposition des codes, ou plus spécifiquement, la juxtaposition d'éléments fonctionnant comme des représentants de différents codes, où ces codes ont une valeur sociale de langue, de sorte que les locuteurs en présence perçoivent de l'altérité codique – un « changement de langue ». Cependant, afin de repérer dans un corpus des occurrences effectives d'alternance codique ainsi définie, il est nécessaire de rendre la plus concrète possible cette définition plutôt abstraite. Cela implique d'effectuer plusieurs choix méthodologiques, en fonction des données disponibles et des objectifs de l'analyse.

4.1.1 Identification des codes représentés

Il a déjà été établi à la section 2.2.1 que ce qui fait partie ou non d'un code-langue donné – et par opposition, ce qui constitue ou non de l'altérité codique – dépend des normes en vigueur dans la ou les communautés de pratique en question. Autrement dit, l'identification des éléments comme représentant le code anglais ou le code français relève de la perception des locuteurs qui les emploient et les entendent. Pour identifier des occurrences d'alternance codique, la perception de l'analyste quant aux langues en présence doit alors se rapprocher le plus possible de la perception des locuteurs des communautés cibles. C'est justement dans ce but que les critères de sélection de corpus à la section 3.1.2 ont été imposés : afin que l'analyse porte sur des locuteurs avec lesquels je partage, en tant qu'analyste, une partie du bagage culturel et linguistique, pour que leur jugement et le mien quant aux codes en présence se rapprochent. Il est vrai, cependant, que ce rapprochement reste partiel ou approximatif, parce qu'il se base sur une recontextualisation qui est toujours partielle (cf. la section 2.4.5). Elle reste partielle même si l'analyste est membre d'une ou plusieurs des mêmes

communautés de pratique précises que les locuteurs étudiés (cf. la section 2.4.5), car il ne partagera quand même pas la totalité de leurs perceptions (entre autres, en raison de son rôle d'analyste), et même s'il a participé à l'interaction analysée, car l'analyse est nécessairement au moins en partie décontextualisée. En conséquence, même dans des conditions idéales à l'égard de la recontextualisation, l'analyste doit toujours se fier en partie à sa propre perception, en fonction de ses expériences antérieures ainsi que des informations cotextuelles et contextuelles (Kerbrat-Orecchioni, 2005 ; 2007).

Ainsi, l'identification à l'un ou l'autre code des éléments dans le corpus analysé a été faite selon ma perception, c'est-à-dire selon ma connaissance de ce qui fait et ne fait pas partie, pour les locuteurs bilingues anglais-français en Amérique du Nord, des codes en question. Selon les informations disponibles (par exemple, les pratiques, commentaires et réactions observables des locuteurs), les normes locales ont aussi été prises en compte, notamment par rapport à l'identification des unités à exclure en raison de leur statut plutôt ambigu quant au code auquel elles sont associées (voir la section 4.1.4).

4.1.2 Juxtaposition des codes

Lorsqu'on ne considère que des énoncés isolés, la juxtaposition d'un code à un autre s'identifie par rapport aux autres codes employés dans chaque énoncé analysé. Cependant, lorsqu'on considère une interaction au complet, la juxtaposition peut avoir lieu non seulement à l'intérieur d'un même énoncé, mais aussi par rapport à d'autres énoncés dans la même conversation, y compris ceux produits par différents locuteurs, c'est-à-dire entre différents tours de parole. On peut donc envisager la juxtaposition sur différents plans : en fonction des productions d'un même locuteur, que ce soit à l'intérieur d'un même tour de parole ou par rapport au code utilisé dans son tour précédent, ou en fonction du déroulement de la conversation, c'est-à-dire par rapport au choix de langue immédiatement précédant le segment de discours en question, peu importe le locuteur qui l'a produit.

L'échange suivant permet de mieux illustrer l'effet de cette distinction sur le repérage des occurrences :

- (1) AN : C'est: um (.) moi j'ai entendu parler in the fifties and the forties g-- uh {this was going on that the}
- (2) INT: {Oh dans ce temps-là j'étais pas là moi.} Ben pas icitte moi.
- (3) AN : W-- we're talking uh early là! Nineteen thirties huh?
- (4) INT: I'm talking the nineteen thirties.

(FAC 01, 22:50)

Si l'on ne considère que les productions d'un même locuteur, le début de l'énoncé en (3) ne constituerait pas une occurrence d'alternance codique, mais le début de l'énoncé en (4) en serait une. Si l'on ne considère que le déroulement de la conversation, l'énoncé en (3) débiterait par une occurrence d'alternance codique mais pas

l'énoncé en (4). Le changement de langue à l'intérieur de (1) pourrait être pris en compte ou non (voir ci-dessous). L'impact de ce choix méthodologique sur l'analyse est manifestement significatif : chaque façon de procéder fait en sorte que certaines données seront retenues comme des occurrences tandis que d'autres seront exclues. C'est pour cette raison qu'il importe d'explicitier qu'il s'agit bien d'un choix, une catégorisation préalable des données effectuée selon l'approche et les objectifs de l'analyse.

La présente étude s'intéresse à ce qui pourrait motiver la sélection d'une unité produite par un locuteur donné, lorsque cette sélection relève de l'altérité codique. Tandis que les études sur la négociation du choix de langue (par exemple Auer, 1995 ; Heller, 1995) démontrent effectivement que chaque locuteur peut avoir des raisons qui lui appartiennent pour maintenir ou ne pas maintenir le choix de langue qu'il a fait dans son tour de parole précédent, chaque nouveau choix de langue s'effectue toujours dans le contexte de la conversation *in situ*, par rapport aux informations disponibles au moment de faire le choix (informations dont son choix de langue précédent ne fait qu'une partie). Par exemple, on pourrait dire que maintenir ou non son choix de langue précédent revient aussi à rejeter ou à accepter le choix de langue que son interlocuteur a fait entretemps. Comme le souligne Auer, « the preceding verbal activities provide the contextual frame for a current utterance », et ainsi, « the meaning of code-alternation depends in essential ways on its 'sequential environment' »⁴⁴ (1995, p. 116).

Cependant, la notion du déroulement *séquentiel* en analyse conversationnelle (par exemple, Auer, 1995; Li, 1998) met l'accent surtout sur les changements de langue qui ont lieu d'un tour de parole à un autre, en réponse au choix de langue précédent ou à un changement dans la situation de communication. Elle relève de l'intérêt, dans cette approche, pour la signification que l'enchaînement des choix de langue porte pour l'interaction dans son ensemble. En revanche, la présente étude s'intéresse non seulement à la juxtaposition des codes en lien avec la séquence des tours de parole, mais à toute juxtaposition qui a lieu par rapport au discours qui lui précède, y compris à l'intérieur d'un même tour de parole. Je préfère ainsi parler du déroulement *linéaire* de la conversation. Cela va dans le sens de Kerbrat-Orecchioni (2009), qui souligne qu'une approche purement séquentielle de l'interprétation d'une interaction néglige le fait que la perception des locuteurs est fondamentalement ancrée dans la linéarité temporelle du déroulement de la conversation. Ainsi, sera comptée comme une occurrence d'alternance codique toute juxtaposition des codes qui a lieu par rapport au déroulement

⁴⁴ L'approche d'Auer (1995) est basée sur une conceptualisation de l'interaction en termes de séquences de tours de parole, d'où son insistance sur le rôle principal de l'environnement séquentiel dans la signification même de l'alternance codique. Il serait plus juste de dire qu'une partie importante de l'interprétation de cette signification, par les interlocuteurs dans un premier temps et par l'analyste plus tard, dépend de l'environnement séquentiel (ou linéaire, voir plus loin dans cette section) – sans oublier toutefois le rôle important du contexte extralinguistique. L'environnement linéaire peut donner des indices quant aux intentions significatives du locuteur, mais ne donne pas un accès direct à ces intentions. Autrement dit, il ne faut pas perdre de vue qu'il peut y avoir d'autres raisons pour lesquelles le locuteur produit de l'alternance codique que la relation de cette production à l'activité verbale précédente.

linéaire de la conversation, que ce soit à l'intérieur d'un même tour de parole ou entre le tour de parole d'un locuteur et le tour d'un autre locuteur qui le suit immédiatement, à l'exception de la juxtaposition qui a lieu à la suite d'une occurrence portant sur une seule unité. Dans ces derniers cas, on peut concevoir l'occurrence comme l'insertion d'une unité du code B dans un segment de discours qui est, sinon, en code A (cf. la section 4.2). La reprise du code A après l'insertion d'une seule unité n'a pas été considérée comme une nouvelle occurrence d'alternance codique mais bien la fin prévisible de l'insertion.

4.1.3 La notion d'unité lexicale

La délimitation théorique et pratique de ce qui constitue un mot est un sujet de débat depuis l'Antiquité, autant quant aux questions de forme et de sens qu'à celles de traitement cognitif. Ces débats sont loin d'être résolus par la linguistique moderne (voir Branca-Rosoff, 1998 pour une discussion des approches de la notion de mot et des problèmes associés ; voir aussi Parent, 2015). Tout comme la notion de langue, il semble que la notion de mot renvoie à une réalité mentale et sociale dont il peut être effectivement utile de parler, même s'il faut admettre la difficulté d'en établir une définition précise qui permet de la distinguer d'autres réalités d'un genre semblable (notamment les morphèmes et les locutions), surtout d'un point de vue sémantique. L'emploi, dans la présente étude, du terme *unité lexicale* plutôt que *mot* se veut une tentative de référer à cette réalité mentale et sociale sans prendre de position tranchée quant à sa nature ou sa définition. En effet, l'élaboration d'une telle position dépasserait nettement le cadre de la présente étude. Toutefois, en raison du rôle important de la notion d'unité lexicale dans le repérage des occurrences d'alternance codique à analyser (voir la section 4.2), elle doit être rendue suffisamment concrète pour servir d'outil dans une analyse sur corpus.

Ainsi, pour les fins de la présente étude, une unité lexicale a été définie comme suit : un mot⁴⁵, une locution, une expression ou un petit groupe de mots qui est mémorisé et partagé en tant que ressource d'expression sur le plan lexical. Cela donne lieu à l'inclusion de phrases grammaticalement complètes, par exemple *c'est dommage* ou *how are you*, ainsi que des séquences de mots où l'emploi de la séquence au complet est plutôt prévisible à partir de l'emploi de l'un d'entre eux, par exemple *la fille à [untel]* ou *ninety-one years old*.

4.1.4 Identification et exclusion des unités ambiguës quant à leur association à un code donné : la question des emprunts

La présente étude accorde une place importante aux occurrences d'alternance codique d'une unité seule (cf. la section 4.2). Or, Poplack et ses collègues (par exemple Poplack, 1987 ; Poplack et Meechan, 1995 ; Turpin, 1998) ont soulevé à juste titre que ce sont justement ces occurrences qui sont les plus susceptibles

⁴⁵ « Mot » ici englobe toute forme lexicale qui est syntaxiquement simple, y compris, le cas échéant, les troncations, les acronymes, et les sigles (en l'occurrence, ces types de formes sont très peu ou pas présents dans le corpus d'analyse).

d'être ambiguës quant à leur « statut codique », c'est-à-dire le code auquel l'unité s'associe ou qu'elle représente dans l'esprit des locuteurs. Autrement dit, il y a lieu de se demander si l'emploi d'une telle unité relève vraiment de l'alternance codique ou s'il s'agit plutôt d'un emprunt. La présente étude ne se veut pas un traité compréhensif sur l'emprunt, qui est un sujet d'étude à part entière. Cependant, il importe de se pencher sur la question de l'emprunt telle qu'elle se pose par rapport à l'alternance codique, c'est-à-dire de voir comment on peut (si même on doit) distinguer les deux catégories d'analyse.

Toutefois, des critères distinctifs qui seraient applicables en pratique au corpus d'analyse sont difficiles à trouver. La littérature sur l'emprunt en soi s'intéresse généralement aux raisons et aux conditions favorisant l'acte d'emprunter (voir par exemple Bouchard, 1999), ou bien à la catégorisation des emprunts effectifs, par exemple selon le type et le degré d'intégration dans le lexique de la langue « d'accueil » (voir par exemple Humbley, 1974 ; voir aussi pour une discussion Paquet-Gauthier, 2015). Ces études ne se penchent généralement pas sur la question de comment on peut identifier les emprunts dans un corpus, étant donné qu'elles visent plutôt des emprunts identifiés a priori comme tels. Une exception à cela est Pflanz (2014), qui applique une échelle de marqueurs d'altérité (originellement élaborée par Josette Rey-Debove dans son ouvrage *Le métalangage*⁴⁶) à un corpus. Cependant, les marqueurs identifiés relèvent exclusivement des textes écrits préparés et sont peu ou pas utiles pour un corpus d'oral spontané. Par exemple, plusieurs d'entre eux sont des marqueurs visuels, comme les italiques ou les guillemets. Le reste des marqueurs consistent à des séquences spécifiques qui qualifient l'unité en question, comme des gloses, des traductions ou des explications. Or, ces séquences peuvent ne pas figurer dans un corpus d'interactions spontanées, où la mise à distance par rapport à des usages moins formels que représente une telle séquence est moins ou pas nécessaire, et où les locuteurs n'ont pas la distance critique par rapport à leurs propres usages qu'auraient des scripteurs d'un texte préparé. Si la constitution du corpus d'analyse n'a pas visé pas spécifiquement ce genre de séquence comme objectif, il est peu probable qu'il y en aurait suffisamment pour servir d'indice fiable du degré d'altérité perçue d'une unité donnée.

Suivant Rey-Debove, Pflanz (2014) propose que l'absence de tout commentaire ou signe métalinguistique indique l'intégration totale. En effet, l'absence de commentaire métalinguistique sur un usage « étranger » indique qu'on présume que son lecteur ou son interlocuteur comprend cet usage sans balisage et l'accepte⁴⁷ – et c'est exactement ce qui fait que l'alternance codique non plus ne comporte pas de marques métalinguistiques dans bien des cas (voir par exemple Poplack, 1980). Un texte écrit préparé s'inscrit dans une

⁴⁶ Rey-Debove, Josette. 1997 (1978). *Le métalangage*. Paris : Armand-Collin.

⁴⁷ On peut également voir dans un commentaire métalinguistique un positionnement du locuteur par rapport à son propre usage, c'est-à-dire un positionnement énonciatif. Par exemple, le locuteur peut vouloir signaler de manière explicite sa connaissance de ressources d'une autre langue, ou atténuer un usage là où l'emploi d'une unité extracodique serait jugé (par le locuteur ou l'interlocuteur) peu ou pas acceptable dans le contexte de communication en question.

situation de distance communicative où on cherche à rapprocher ses usages d'un standard partagé par le plus grand nombre de lecteurs (cf. la section 3.1.1.2); ainsi un usage « étranger » sans commentaire suggère de fait que cet usage est intégré aux pratiques du lectorat envisagé. Dans la plupart des contextes d'interaction propices à l'alternance codique, les locuteurs en présence savent ou présument qu'ils partagent des ressources similaires dans les deux langues, ainsi que des normes d'usage. Le marquage métalinguistique peut ne pas être nécessaire entre eux, sans que ce soit indicatif de l'intégration de l'unité lexicale en question dans la communauté plus large. Ainsi, quoique l'étude de Pflanz (2014) ne manque pas d'intérêt comme approche de l'altérité, l'échelle des marqueurs qu'elle emploie est peu applicable à la présente analyse en raison de la nature fondamentalement différente du corpus analysé.

D'autres possibilités sont mises de l'avant par des chercheurs considérant l'emprunt spécifiquement dans sa relation avec l'alternance codique dans le discours spontané. Poplack et ses collègues (par exemple Poplack, 1987 ; Poplack, 2012 ; Poplack et Meechan, 1998 ; Poplack, Sankoff, et Miller, 1988 ; Turpin, 1998) mettent particulièrement l'accent sur l'intégration, notamment l'intégration morphosyntaxique⁴⁸, de l'unité en question dans la langue « d'accueil » comme critère distinctif. Pour ces chercheurs, une unité de la langue A dans un énoncé en langue B qui ne présente pas de marqueurs morphologiques ou syntaxiques de la langue B (marque du pluriel, présence d'un déterminant approprié, etc.) est considérée comme une alternance codique puisqu'elle ne s'est pas intégrée à la grammaire de la langue B. Si de tels marqueurs sont présents, alors la grammaire de la langue B a nécessairement participé à la production de l'unité en question, indiquant son intégration à cette même grammaire. Si l'unité ainsi intégrée n'est pas diffusée dans la communauté, il s'agit quand même d'un emprunt, soit un emprunt ponctuel (« nonce borrowing »). Cette approche n'est toutefois pas satisfaisante. En admettant la possibilité d'un emprunt ponctuel, employé par un seul locuteur une seule fois (Poplack, Sankoff, et Miller, 1988), l'importance pour la notion d'emprunt de la dimension sociale est écartée. Or, le statut d'emprunt implique que l'unité en question fait partie du code « emprunteur », et parce qu'un code est un ensemble socialement défini, faire partie d'un code implique de faire l'objet d'un certain consensus social. Même si tout code manifeste de la variation interindividuelle, en ce que les expériences, ressources et habitudes d'usage qui constituent un code donné peuvent varier d'un locuteur à un autre, tout code a aussi nécessairement

⁴⁸ Voir la section 4.1.5 pour une discussion relative à l'intégration phonologique. Turpin (1998) considère aussi l'intégration discursive, soit que la présence de dysfluidités ou de marqueurs comme « tsé » ou « like » pourrait signaler qu'un changement de langue s'effectue. Cependant, de telles dysfluidités peuvent également signaler une recherche lexicale en cours ou représenter une stratégie discursive, telles que la mise en relief d'une unité lexicale ou le maintien du tour de parole, entre autres (Backus, 2001; Brennan et Williams, 1995; Clark et Fox Tree, 2002; Fraundorf et Watson, 2011). De toute façon, comme les autres formes d'intégration (voir ci-dessous), l'intégration discursive relève des informations idiolectales et ponctuelles, qui ne permettent pas de répondre à une question qui porte plutôt sur les pratiques collectives d'une communauté.

une part d'intersubjectivité, sans quoi il ne pourrait pas signifier ou servir à communiquer quoi que ce soit.⁴⁹ Le degré de diffusion et de partage nécessaire pour établir qu'une unité appartient à un code donné est, dans l'état actuel des connaissances, impossible à déterminer avec précision; cependant, on peut être assez certain qu'un seul emploi par un seul individu ne suffit pas (Backus, 2009 ; Humbley, 1974).

De plus, l'intégration en tant que critère distinctif de l'emprunt a été critiquée par plusieurs autres chercheurs. Quoique Poplack et ses collègues indiquent à juste titre que l'intégration phonologique n'est pas un indice fiable (cf. la section 4.1.6), même l'intégration morphosyntaxique pose problème. Pour Myers-Scotton, dans son modèle du cadre de la langue matrice (« Matrix Language Frame [MLF] model »; par exemple Myers-Scotton, 1992a ; 1992b ; 1995), il est tout à fait normal que tant des occurrences d'alternance codique que des emprunts fassent preuve d'une telle intégration. Dans ce modèle, c'est la langue matrice (« matrix language [ML] ») qui fournit le cadre syntaxique et tous les « morphèmes de système » (en gros, les morphèmes dits fonctionnels ou grammaticaux) dans des constituants mixtes; une unité de la langue enchâssée (« embedded language [EL] ») insérée dans un tel constituant se verra nécessairement intégrée à la morphosyntaxe de la ML.⁵⁰

Muysken (2000) soutient aussi que les occurrences d'alternance codique insertionnelle font preuve de plus souvent d'intégration morphosyntaxique, donc que l'intégration ne permet pas de distinguer entre l'alternance codique et l'emprunt. Cependant, il considère que ces deux phénomènes sont fondamentalement différents. Pour Muysken, l'alternance codique a lieu sur le plan supralexical, c'est-à-dire qu'elle consiste en la combinaison, à l'intérieur d'une même proposition, de mots avec des indices de langue différents. L'emprunt, en revanche, aurait lieu au niveau sublexical : un ou des éléments de formation de la langue A sont insérés dans une sorte d'enveloppe lexicale vide de la langue B. Cette enveloppe conférerait à son contenu le statut d'unité lexicale de la langue B, lui permettant ainsi de participer dans des relations morphosyntaxiques propres à la langue B. L'approche de Muysken a le mérite de reconnaître le rôle de la dimension sociale, en ce que Muysken soutient que l'emprunt est généralement partagé et conventionnalisé, tandis que l'alternance codique est plus souvent perçue comme novatrice ou créative. Toutefois, il distingue l'alternance codique insertionnelle de l'emprunt en s'appuyant sur des a priori théoriques (l'indice de langue, l'interaction des grammaires) et

⁴⁹ Un code est un objet *socialement* défini, mais qui se construit chez chaque individu à travers un ensemble *singulier* d'expériences. Cette relation – on pourrait même dire cette tension – entre la variation interindividuelle dans les ressources et les habitudes d'usage, d'une part, et d'autre part la nécessité d'intersubjectivité pour la compréhension et la signification, fait qu'on ne peut pas oublier ou écarter ni l'une, ni l'autre dimension, de la définition ou de la délimitation d'un code.

⁵⁰ Le modèle MLF est l'un des modèles théoriques de l'alternance codique les plus élaborés, et Myers-Scotton et ses collègues y ont apporté plusieurs révisions et ajustements avec le temps. Il n'est toutefois pas exempt de critiques, notamment par rapport aux notions de morphèmes de système et de contenu, ainsi qu'aux méthodes pour identifier la langue matrice dans un segment de discours donné. Comme il ne servira pas pour la présente étude, une discussion plus détaillée ne s'avère pas pertinente, mais on peut consulter par exemple MacSwan (2005) ou Tracy (2000).

notamment sur une construction théorique (l'enveloppe lexicale) dont la validité théorique et empirique restent à démontrer. Cette construction, de plus, n'est pas directement observable, et Muysken (2000) ne propose pas d'indices indirects qui pourraient remplacer l'intégration morphosyntaxique. Autrement dit, même si l'on acceptait son explication du fonctionnement des deux phénomènes, son approche n'aide pas à déterminer dans quelles circonstances on peut attribuer le statut d'emprunt à une unité donnée.

Pour sa part, Myers-Scotton (1992a) trouve que le seul critère exploitable pour distinguer les deux phénomènes est la fréquence d'occurrence. Puisque les emprunts font partie de la ML, il y a moins de restrictions sur leur emploi dans une phrase mixte. Ils devraient donc apparaître plus souvent dans un corpus, autant en fréquence absolue (nombre total de tokens) qu'en fréquence relative par rapport à leurs concurrents d'expression dans la langue « d'accueil ». Myers-Scotton elle-même souligne que le critère de fréquence absolue nécessite que soit fixé un seuil de fréquence arbitraire (par exemple, on pourrait considérer une unité comme étant un emprunt si elle figure plus de trois fois dans le corpus, employée par des locuteurs différents dans des contextes différents). Ce critère est d'autant plus arbitraire en ce qu'il présuppose que la fréquence d'emploi observée dans le corpus en question serait représentative de la fréquence d'emploi des unités en question dans la communauté de façon générale, et que la fréquence en général est représentative de l'appartenance à un code donné. Notamment, on pourrait facilement penser à une unité qui serait très peu fréquente dans un corpus de conversation (par exemple, un mot « savant » comme *focalisation* ou un terme technique ou spécialisé comme *myéline*) dont l'association à un code donné n'est quand même pas en question. Le critère de fréquence relative, quant à lui, est également problématique en ce qu'il repose non pas sur l'acceptation dans la communauté de pratique, mais sur le fait de supplanter des concurrents d'expression existants, phénomène qui n'arrive pas forcément pour tout emprunt.

Toutes ces façons de distinguer l'emprunt de l'alternance codique présupposent l'existence de deux systèmes de production distincts — c'est-à-dire qu'on ne peut pas produire de la morphologie ou de la syntaxe de la langue B si l'on n'est pas en train de « parler la langue B », d'une façon ou d'une autre. Elles relèvent d'un objectif premier qui est d'explicitier les mécanismes et les procédures par lesquels les phénomènes observés pourraient être produits, tout en préservant cette notion de deux systèmes distincts. Par contre, lorsqu'on considère qu'une langue n'est pas un module grammatical distinct, mais un ensemble de ressources socialement défini, une distinction sur la base de l'interaction ou non de deux grammaires ne peut pas être opérée : il est tout à fait possible pour les locuteurs de combiner – ou non – des bases de la langue A avec des affixes ou des flexions de la langue B, ou des noms de la langue A avec des déterminants de la langue B, etc., compte tenu du fait que des processus de formation et de combinaison de mots figurent parmi les ressources linguistiques dans les répertoires des locuteurs au même titre que des morphèmes et des unités lexicales qui font l'objet de ces processus (Lüdi et Py, 2009 ; voir aussi de Bot, 2016).

Ce qui importe alors pour la présente analyse n'est pas les mécanismes de production ou d'incorporation dans la structure d'une phrase, mais plutôt les valeurs attribuées – c'est-à-dire le statut codique attribué – à un élément donné au sein d'une communauté de pratique donnée. Il a déjà été signalé à la section 2.2.2 qu'on peut considérer l'alternance codique et l'emprunt comme se situant sur un continuum d'altérité perçue par rapport à un code donné. Cette altérité est une fonction du statut codique que l'unité en question a au sein de la communauté de pratique en question. Si une unité est identifiée par les locuteurs comme faisant partie d'un autre code, ils percevront un degré élevé d'altérité lors de son emploi, qui constituera pour eux la juxtaposition d'un code à un autre, c'est-à-dire de l'alternance codique. Si l'unité est identifiée comme faisant partie du même code malgré ses origines extracodiques (auxquelles les locuteurs peuvent ou non être sensibles), les locuteurs percevront un degré moindre – allant jusqu'à un degré zéro – d'altérité, et on peut la qualifier plutôt d'emprunt.

Ainsi, à un extrême de ce continuum, il y a ces occurrences d'alternance codique qui sont perçues comme des innovations, ne faisant pas du tout partie des habitudes d'usage de la communauté de pratique du code en question (Backus, 2009). À l'autre extrême, il y a des emprunts qui sont complètement intégrés et ne soulèvent aucun sentiment d'altérité, dont l'association à un autre code n'est pas perceptible pour la majorité des locuteurs (à moins d'avoir des connaissances sur leur étymologie). Près des deux pôles du continuum, donc, le statut codique d'une unité donnée est relativement facile à déterminer, et l'emprunt et l'alternance codique semblent être des phénomènes clairement distincts. Plus une unité s'éloigne de ces cas de figure, toutefois, plus son statut devient flou. On peut à juste titre se demander quel seuil d'altérité ou d'intégration est nécessaire, et selon quelles mesures, pour pouvoir affirmer que l'emploi d'une unité donnée relève de l'alternance codique ou de l'emprunt, respectivement.

Par exemple, il n'est pas nécessaire qu'une unité ne soulève aucune perception d'altérité pour qu'elle soit considérée comme un emprunt : on peut percevoir une certaine altérité dans une unité donnée sans avoir le sentiment que son emploi constitue tout à fait un changement de code (tel est le cas, par exemple, de *parking*, *jeans*, ou *fun* en français canadien). On pourrait toutefois y voir un emploi de ressources relevant d'une autre langue. C'est notamment le cas des emprunts à l'anglais, qui ont une place particulière en français (et surtout en Amérique du Nord), comme l'abondance d'écrits sur « les anglicismes » en témoigne. Ce n'est pas l'objectif de la présente étude de traiter en profondeur de la notion d'anglicisme (on peut voir à ce sujet Courbon et Paquet-Gauthier, 2014 ; Paquet-Gauthier, 2015), mais il s'agit d'un sujet incontournable dans la francophonie nord-américaine qui a une importance particulière pour la présente étude. Brièvement, les emprunts à l'anglais, ou même les unités lexicales qui ont une ressemblance formelle ou sémantique fortuite à des usages de l'anglais, font l'objet de nombreuses chroniques linguistiques, de critiques, de marques lexicographiques spécifiques, d'ouvrages dits correctifs, et ainsi de suite, par souci du respect de la « qualité » du français

(Courbon et Paquet-Gauthier, 2014 ; Paquet-Gauthier, 2015). Cette attitude négative envers l'anglais se trouve également chez des Franco-Albertains (Beaulieu, 2012) et chez des Franco-Américains (Fox, 2007). Ainsi, selon les attitudes et les connaissances linguistiques et métalinguistiques du locuteur quant à l'unité en question, et selon le contexte d'emploi, un emprunt à l'anglais pourrait (ou non) susciter un sentiment d'altérité marquée, quel que soit son degré de diffusion et d'intégration aux habitudes d'usage au sein de la communauté de pratique francophone.

Établir une distinction entre l'alternance codique et l'emprunt est d'autant plus compliqué qu'une unité peut ne pas appartenir exclusivement et définitivement à un code ou à un autre dans l'esprit des locuteurs (de Bot, 2016), surtout dans le cas de l'anglais et du français, qui partagent beaucoup de formes similaires, dont beaucoup sont aussi similaires quant à leur(s) signification(s) (par exemple *traff[ff]ic, information, -ism[e]*, etc.). Également, le statut codique d'une unité peut varier d'un locuteur à l'autre, selon le contexte et selon ses expériences, croyances et habitudes.

Ainsi, plus on s'éloigne des extrêmes du continuum d'altérité perçue, moins la division conceptuelle entre l'emprunt et l'alternance codique paraît correspondre à un phénomène perçu par les locuteurs, et plus elle paraît comme des catégories d'analyse, imposées selon les intérêts et les objectifs de l'analyse. Autrement dit, quoique le statut d'emprunt ait une certaine réalité dans l'esprit des locuteurs, il est aussi en partie un artefact de l'analyse. On peut même, à juste titre, se questionner sur la pertinence de distinguer deux catégories d'analyse différentes lorsque le statut codique n'est pas clairement évident. Cette pertinence est d'autant plus incertaine lorsqu'on considère que l'analyste n'a pas d'accès direct au sentiment d'altérité codique, ou à l'absence de ce sentiment, qu'éprouvent les locuteurs qu'il étudie. Sans des enquêtes visant spécifiquement à cerner le statut de chaque unité en question dans les communautés de pratique étudiées⁵¹ – et idéalement auprès de chaque locuteur spécifique dont les productions dans le corpus sont analysées – toute décision quant au statut d'une unité ambiguë donnée est, d'une façon ou d'une autre, l'imposition du jugement de l'analyste, dans la mesure où il ne repose pas sur la perception réelle des locuteurs.

Au vu des difficultés indiquées ci-dessus, de la petite taille du corpus et du manque d'enquêtes sur le statut codique des unités en question dans les communautés en question, il est hors de la portée de la présente

⁵¹ Encore là, la question de *comment* cerner ce statut au moyen d'une enquête demeurerait. On pourrait penser à la méthode que Sablayrolles (2003) a utilisée pour sonder le sentiment néologique, qui est de demander à plusieurs locuteurs d'identifier dans un texte les unités qui suscitent ce sentiment en eux. Toutefois, même une méthode aussi pertinente pour la question du sentiment d'altérité n'est qu'un indice indirect, en ce qu'elle mesure les réactions des locuteurs envers des éléments au moins en partie décontextualisés. Elle ne mesure pas ce que les locuteurs sont susceptibles de produire, ni le sentiment d'altérité qu'ils pourraient éprouver au moment de la production. De plus, elle mesure des réactions ouvertes et conscientes, donc susceptibles d'être biaisées par les attitudes susmentionnées envers les unités d'origine anglaise en français, ou par les attitudes envers l'alternance codique ou l'emprunt en général.

analyse d'attribuer le statut d'emprunt à une unité donnée. Il ne convient donc pas d'exclure des unités sur la base d'une distinction entre l'alternance codique et l'emprunt. Toutefois, certaines unités dans le corpus d'analyse ont effectivement un statut codique ambigu (ou du moins, plus ambigu que d'autres unités)⁵². Cette ambiguïté n'exclut pas en elle-même que l'emploi de ces unités corresponde à l'objet de la présente étude – un recours à des ressources identifiées à un autre code dans le but d'exprimer une visée référentielle donnée – mais elle complique l'analyse des unités en question. Or, la présente est, à ma connaissance, la première étude de l'alternance codique qui aborde sa dimension d'expression ou de signification dans l'approche spécifique que j'utilise, axée sur la perception des locuteurs et basée sur l'observation des sens en usage et des fonctions discursives. Elle vise à démontrer la pertinence de cette même approche, et ne prétend pas, du même coup, pouvoir résoudre de telles complexités. Dans le but de restreindre l'analyse à des « lieux (plus) sûrs » d'observation d'un changement de code en tant que stratégie d'expression, il convient donc d'exclure de la présente étude des unités manifestant de l'ambiguïté quant au statut codique que leur attribuent les locuteurs étudiés. Il est à souligner que cette exclusion est un choix méthodologique fait en fonction des objectifs et des contraintes de l'analyse, et n'implique pas que les unités en question soient nécessairement ou assurément des emprunts, ni qu'elles soient sans intérêt pour l'étude de l'alternance codique. Elle n'implique pas non plus qu'on puisse ni qu'on doive supposer qu'une frontière nette entre les deux langues existe dans les faits.

On revient donc à la question des critères d'identification, non pas de la catégorie d'analyse « emprunt », mais de l'ambiguïté quant à la catégorisation qui pourrait être faite d'une unité (anglais ou français, juxtaposition de codes ou non). Cette ambiguïté relève dans un premier temps de ma perception en tant qu'analyste : en dépouillant le corpus d'analyse, certaines juxtapositions apparentes de codes ont occasionné plus de doutes quant à leur classification que d'autres. Toutefois, il ne faut pas se fier à cette perception sans recul critique, ni prétendre qu'on peut en faire entièrement abstraction, mais encore faut-il l'ancrer dans les données empiriquement observables qui l'informent. Il s'agit, autrement dit, de formuler concrètement et explicitement ce sur quoi se fonde cette perception : quels indices sont présents dans le corpus d'analyse et dans les informations contextuelles disponibles qui me font remettre en question le statut codique attribué par les locuteurs étudiés à une unité donnée?

L'ambiguïté codique peut être constatée par la présence, par rapport à une unité qui pourrait autrement être associée au code A, des indices d'intégration aux habitudes d'usage au sein de la communauté de pratique du code B⁵³. Or, comme la revue de la littérature sur l'emprunt le laisse croire, il n'y a pas un seul critère ou

⁵² Je réfère ici et dans ce qui suit aux unités dont la nature extracodique est formellement manifeste (cf. Courbon et Paquet-Gauthier, 2014), c'est-à-dire à l'exclusion des emprunts sémantiques et des calques.

⁵³ Les formulations bien connues de « langue emprunteuse » et « langue donneuse » sont toutefois malheureuses en ce qu'elles personnifient la langue comme l'agent de l'emprunt et, par le fait même, mettent en arrière-plan les locuteurs, qui sont les véritables agents de tout changement dans le statut d'une unité donnée par l'usage qu'ils en font. « Langue

indice de cette intégration qui est suffisant à lui seul dans tous les cas. Il semble plutôt y avoir un ensemble de facteurs qui *suggèrent* tous l'intégration au sein de la communauté, et lorsque plusieurs de ces facteurs convergent autour d'une même unité, cela met en doute le statut codique que les locuteurs étudiés lui attribuent. Ainsi, plusieurs facteurs, y compris certains mentionnés dans la littérature, ont été pris en considération dans la décision d'exclure certaines unités de l'analyse en raison de l'ambiguïté de leur statut codique :

- la diffusion observée dans la communauté de pratique : plus une unité potentiellement extracodique est reprise spontanément (c'est-à-dire dans des emplois qui ne sont pas des répétitions ou des reprises), par différents locuteurs, dans différents contextes, plus cette unité est diffusée dans la communauté de pratique, et moins elle devrait susciter un sentiment d'altérité codique (*cf.* Backus, 2009 ; Myers-Scotton, 1992a). Il faut rappeler que, sans une enquête de terrain sur une telle unité précise, la diffusion observée dans un corpus n'indique que rarement à quel point cette unité est vraiment partagée et acceptée dans la communauté (si c'est un emprunt établi ou un emprunt en devenir, si l'unité est employée dans la communauté en général ou seulement par un groupe très restreint, si sa fréquence dans le corpus est représentative ou un hasard de l'échantillonnage, etc.);
- la préférence observée dans la communauté de pratique : plus les locuteurs étudiés semblent reprendre une unité potentiellement extracodique pour exprimer une réalité donnée plutôt que de possibles concurrents dans le code B, plus les chances sont élevées que cette unité corresponde à « la façon dont on parle » de la réalité en question, et donc que l'unité soit au moins en partie intégrée au code B. Il est à noter que la concurrence doit porter sur une *même réalité* sur le plan de la référence (ou sur des réalités qu'on pourrait juger convergentes au point d'être la même), ce qui implique de tenir compte dans la mesure du possible de la valeur d'expression associée au registre de langue ou au style, ainsi que du degré de généralité. On ne pourrait donc pas dire que *pussycat* et *chat domestique*, par exemple, ou *pussycat* et *félin*, sont des concurrents, du moins pas dans tous les contextes. Il est à noter également qu'il n'est pas question ici du remplacement ou de la supplantation totale des concurrents par l'unité d'origine extracodique, comme Myers-Scotton (1992a) l'entend, mais bien d'une préférence pour l'unité en question, au regard de la fréquence d'emploi.

d'accueil » n'a pas le même sens d'agentivité, mais participe tout de même à la métaphore des mots qui « voyagent » lors de l'emprunt, et en plus n'a pas d'expression homologue pour exprimer la source de l'emprunt. Afin d'éviter ces formulations, « code B » sera dorénavant utilisé pour signifier le code par rapport auquel on évalue le degré d'intégration ou d'altérité que présente une unité d'origine extracodique, et « code A » pour signifier le code qui est, relatif au code B, le code d'origine de cette unité.

- l'intégration phonologique : plus une unité est prononcée selon la phonologie du code B (c'est-à-dire sans un « accent étranger »), plus les chances sont élevées qu'elle soit au moins en partie intégrée aux habitudes d'usage des locuteurs du code B (mais voir la section 4.1.6);
- l'intégration morphosyntaxique : plus une unité suit la morphologie du code B et est employée selon la syntaxe de ce même code (lorsque celle-ci est distinctive, comme dans la position des adjectifs relatif au nom, cf. par exemple Pfaff, 1979 ; Poplack, 1980), plus les chances sont élevées qu'elle soit au moins en partie intégrée aux habitudes d'usage des locuteurs du code B. Toutefois, comme il a déjà été mentionné, il est tout à fait possible qu'un locuteur intègre à la morphosyntaxe de son énoncé en cours une unité qui n'est aucunement intégrée au code en question;
- l'intégration discursive : plus une unité est employée sans dysfluidité (pauses, hésitations, bribes et reformulations) et sans réaction observable des participants (cf. Gafaranga et Torras, 2002 ; Turpin, 1998), plus les chances sont élevées que cette unité soit conventionnalisée pour les locuteurs en question (c'est-à-dire que son emploi ne représente pas une innovation ou une violation des normes) et donc soit intégrée à leurs habitudes d'usage dans le code B. Il faut cependant noter que toute unité qui ne suscite pas de réaction de la part des locuteurs en présence n'est pas nécessairement intégrée au code en question, surtout dans des communautés où l'alternance codique est une pratique usuelle. Aussi, il faut rappeler que les dysfluidités sont des phénomènes caractéristiques de l'oral (cf. Blanche-Benveniste, 2010) qui peuvent être produits en lien avec une unité parfaitement bien intégrée;
- la modification sémantique : si la signification de l'unité potentiellement extracodique dans le code en question diffère de la signification associée à cette même forme dans le code A, cela suggère que les locuteurs du code en question se sont en quelque sorte approprié cette unité (cf. Courbon, 2015 ; Vogh et Courbon, à paraître), l'intégrant à leurs habitudes d'usage. C'est le cas, par exemple, de *parking* ou *fucké* en français québécois. Cependant, quoiqu'on puisse observer les sens en usage et les fonctions discursives de l'unité en question dans le code B, il est rarement possible de connaître la signification qu'elle aurait dans le code A tel qu'il se constitue dans la communauté de pratique étudiée. En l'absence de telles connaissances, cet indice a été appliqué sur la base des connaissances des usages plutôt généraux en anglais et en français;
- l'attestation dans des ouvrages lexicographiques : l'attestation dans un ouvrage lexicographique est un bon indice qu'une unité soit diffusée et intégrée auprès d'une large communauté de pratique. Toutefois, cet indice a plusieurs limites. Premièrement, les ouvrages lexicographiques ne représentent pas fidèlement l'ensemble des usages externes à leur principale communauté cible (par exemple, la France pour *Le Petit Robert*, le Québec pour *Usito*). Or, à ma connaissance, il n'existe pas d'ouvrages

lexicographiques ayant pour communauté cible l'Alberta francophone ou le Maine francophone. De plus, malgré l'emploi des unités de lexique caractéristiques du français acadien par plusieurs locuteurs étudiés, le *Dictionnaire du français acadien* (Cormier, 2009) n'a pas pu être utilisé en raison du choix du lexicographe, explicité dans la préface, de ne pas repérer d'anglicismes plus récents que le 19^e siècle. Ainsi, il a fallu recourir aux deux premiers ouvrages susmentionnés, en sachant que seules les unités diffusées de façon à retenir l'attention des lexicographes dans les grandes communautés francophones québécoise (pour *Usito*) et française (pour *Le Petit Robert*) y seraient attestées. Deuxièmement, même au sein de cette communauté cible, toute unité en usage ne figure pas forcément dans le dictionnaire; notamment, les usages des registres familiers et de la langue orale sont généralement sous-représentés. En plus, beaucoup de temps peut s'écouler entre l'attestation d'une unité lexicale (empruntée ou forgée) dans une communauté de pratique et son inclusion dans le dictionnaire.⁵⁴ L'absence d'une unité de ces ouvrages ne dit donc rien de son statut au sein de la communauté cible, et encore moins dans les communautés étudiées. Toutefois, en l'absence d'ouvrages lexicographiques spécifiques aux régions représentées dans le corpus d'analyse, des unités attestées dans le *Petit Robert* (2016, version en ligne) ou dans *Usito* (2014) comme étant des emprunts à l'anglais ou des anglicismes (critiqués ou non) ont été exclues de l'analyse.

Plusieurs points sont à souligner ou à rappeler ici. D'abord, la liste d'indices ci-dessus a été énumérée en fonction de ce qui était évaluable selon les données et les ressources qui m'étaient disponibles, et ne se veut pas un ensemble définitif de critères établissant l'ambiguïté codique. L'exclusion ou non d'une unité en raison de ces indices n'implique pas non plus qu'il y ait une différence sur le plan phénoménologique entre les unités qu'ils permettent d'exclure et des unités qu'ils n'excluent pas. Il s'agit plutôt d'une distinction opérée en fonction des besoins de l'analyse. De plus, comme il l'a déjà été signalé ci-dessus, chaque indice a des limites importantes, et aucun d'entre eux ne suffit en soi pour exclure une unité de l'analyse. Plutôt, lorsqu'une unité fait l'objet d'une convergence de plusieurs indices, il y a lieu de penser que son statut codique est plutôt ambigu, soit dans l'esprit des locuteurs, soit par rapport aux données disponibles pour catégoriser l'unité aux fins de l'analyse.

Finalement, l'exclusion des unités de l'analyse en raison de l'ambiguïté de leur statut codique n'implique pas que toute ambiguïté potentielle a été éliminée de cette analyse (cf. les sections 5.3.1 et 5.5.3). Ce ne sont que les unités présentant un degré élevé d'ambiguïté, selon les indices énumérés ci-dessus, qui ont été écartées

⁵⁴ Par exemple, Boulanger et Malkowska (2008) ont trouvé des écarts de 15 ans et plus, dans les dictionnaires les plus usuels, pour des emprunts « récents » dans le domaine de l'alimentation.

pour l'instant. Concrètement, la liste d'indices ci-dessus a mené à l'exclusion de l'analyse des unités comme *job, party, fun, rider (v.), driver (v.), watcher (v.),* et [*mouvoir, revenir, etc.*] *back*.

4.1.5 Identification et exclusion des noms propres

La signification des noms propres – ce qu'ils expriment dans l'esprit des locuteurs et les mécanismes sous-jacents par lesquels ils l'expriment – est un domaine de recherche à part entière, complexe, avec de multiples sujets à débat qui dépassent le cadre de la présente étude (voir à ce sujet van Langendonck, 2007 ; Vaxelaire, 2009 ; cf. aussi par exemple Kleiber, 1996 ; 2004 ; 2007). Cela n'exclut pas en soi la possibilité d'une analyse de la visée référentielle qui y est associée (cf. la section 2.4.7), mais d'autres facteurs sont aussi à considérer. Premièrement, il serait difficile de rapprocher mes perceptions en tant qu'analyste de celles des locuteurs lorsqu'il s'agit des noms propres : je ne partage pas les connaissances encyclopédiques d'une bonne partie des référents que les locuteurs étudiés sont susceptibles de désigner avec un nom propre (des personnes de leur entourage, des entreprises locales, des villages voisins, etc.). Cela nuirait par conséquent à la recontextualisation nécessaire pour identifier la visée référentielle.

Il faut aussi se demander s'il est même possible d'identifier un nom propre à un code donné. S'il est relativement clair que *London* et *Londres*, par exemple, relèvent de deux codes différents, que faire avec des cas comme *Marc Messier* ou *Edmonton* où « c'est comme ça » que le référent est désigné, peu importe le code utilisé? Comment faire pour savoir s'il s'agit ou non d'un changement de code dans l'esprit des locuteurs? De fait, les noms propres figurent parmi les « exemples notables » des mots ambigus mentionnés par Clyne (2000 [1987]), qui semblent évoquer les deux codes en même temps dans l'esprit des locuteurs.

Ainsi, pour toutes ces raisons, tout nom de personne, entreprise, institution, ville ou autre lieu géographique, marque⁵⁵, etc. n'a pas été repéré comme une occurrence d'alternance codique et ne figure pas dans les analyses.

4.1.6 Exclusion sur la base de la réalisation phonétique

Les travaux de Poplack et ses collègues (Poplack, 1980 ; Poplack, 1987 ; Poplack, Sankoff, et Miller, 1988 ; Turpin, 1998) ont fait ressortir clairement la non-fiabilité de la réalisation phonétique dans l'identification de l'alternance codique. D'une part, ce sont les locuteurs bilingues qui maintiennent le plus souvent la prononciation « étrangère » d'une unité d'origine extracodique (Poplack, Sankoff, et Miller, 1988), et d'autre part, les réalisations phonétiques des bilingues diffèrent de celles des monolingues, et ce, dans les deux langues (Flege, 2007). En plus, tout locuteur peut commettre une « erreur de production » qui influe sur sa prononciation. Il est donc difficile d'établir avec certitude dans quel code le locuteur avait l'intention de s'exprimer, si la seule

⁵⁵ À l'exception des noms de marque qui se sont généralisés dans l'usage (ex. *kleenex, thermos*).

distinction entre une unité d'un code et une unité de l'autre est la réalisation d'un ou deux phonèmes. Pour cette raison, de telles unités n'ont pas été repérées. Cela exclut des unités comme *no* (par similarité avec *non*) et *hey* (par similarité avec *hé* ou *eille*). Des unités comme *agent*, *library* [*librairie*], ou *modern[e]*, qui se distinguent par plusieurs phénomènes phonétiques (phonèmes, accentuation, etc.), ont toutefois été retenues, car leur association à l'un ou l'autre code dans l'esprit des locuteurs est plus sûre.

4.2 Occurrences ciblées pour l'analyse

Plusieurs façons de classifier la juxtaposition des codes sont proposées dans la littérature, dans le but d'orienter les analyses à effectuer ou bien de restreindre l'analyse à une ou des manières spécifiques de juxtaposer. Bien entendu, cette classification se fait en fonction de l'approche du ou des chercheurs en question et des objectifs de leur analyse. Par exemple, toute étude qui analyse l'alternance codique comme une interaction entre deux grammaires distingue entre l'alternance codique interphrastique et intraphrastique : l'alternance codique interphrastique est écartée de ces analyses explicitement (voir par exemple MacSwan, 2013 ; Myers-Scotton, 1992b ; Sankoff et Poplack, 1981) ou implicitement, parce que ce type d'alternance n'est pas un lieu d'observation des interactions syntaxiques. La présente étude n'est pas exemptée du besoin de classifier et de restreindre les occurrences analysées selon les objectifs de l'étude. Ainsi, quoique toute juxtaposition de codes conforme aux critères en 4.1 ait été considérée comme une occurrence d'alternance codique lors du dépouillement du corpus (voir le tableau 4.1), seules certaines de ces occurrences ont été analysées. Notamment, l'analyse se concentre sur les occurrences « isolables » (la section 4.2.1) et qui ont été produites en anglais (la section 4.2.2).

4.2.1 Les occurrences isolables

Afin d'analyser l'alternance codique en tant que stratégie d'expression sur le plan lexical, il faut pouvoir identifier, au sein de toute occurrence étudiée, la ou les unités lexicales qui ont potentiellement fait l'objet d'une sélection à des fins référentielles. Pour s'assurer qu'une unité donnée a bien ce potentiel, il faut pouvoir l'identifier comme l'unité lexicale dont l'emploi aurait « occasionné » le changement de code. Dans ce sens, et contrairement à l'approche de Poplack (par exemple, 1987), ce sont les occurrences de plusieurs unités qui sont les plus ambiguës, car il est plus difficile (voire impossible) d'y identifier avec certitude l'unité « cible », s'il y en a une, qui aurait motivé ou occasionné l'alternance.⁵⁶ On ne peut pas simplement repérer la première unité de l'occurrence comme étant cette unité cible, car le locuteur aurait pu changer de code en anticipation de l'emploi d'une unité donnée plus loin dans son énoncé. Cela peut notamment être une stratégie pour éviter des

⁵⁶ Il est à souligner ici qu'une unité lexicale consiste en une association entre une forme et un ou des sens. Autrement dit, lorsqu'une occurrence d'alternance codique a été motivée par la sélection d'une unité donnée, cette motivation tient non seulement de la forme matérielle qu'on observe mais aussi – voire plutôt – du ou des sens que cette forme permet de représenter en discours.

changements de code à des endroits particulièrement inhabituels dans la phrase, comme entre un nom et un adjectif attribut (Pfaff, 1979 ; Poplack, 1980).

On pourrait donc penser à restreindre l'analyse à des occurrences d'alternance codique de type insertionnel, c'est-à-dire qui relèvent de l'insertion d'une seule unité du code B dans un discours ayant lieu autrement en code A (aussi appelé « single-word switches » ou « lone other-language items [LOLIs] », cf. par exemple Poplack, 1980 ; Poplack, 2012). La notion d'insertion est déjà bien établie dans la littérature, notamment dans les travaux de Myers-Scotton (par exemple, Myers-Scotton, 1995 ; Myers-Scotton et Jake, 1995). Le modèle MLF (déjà discuté à la section 4.1.4) envisage toute alternance codique à l'intérieur d'un même constituant syntaxique comme un processus d'insertion : l'insertion d'un ou plusieurs « morphèmes de contenu » (« content morphemes ») de la langue enchâssée dans un cadre de « morphèmes de système » (« system morphemes ») fourni par la langue matrice. L'insertion d'une unité seule représente le cas non marqué ou prototypique du modèle; un mécanisme d'exception est nécessaire pour expliquer des séquences plus longues (« EL islands »), à l'intérieur desquelles la langue enchâssée peut elle aussi fournir une partie du cadre.

On retrouve également la notion d'insertion dans la typologie de Muysken (2000), basée en partie sur le modèle de Myers-Scotton. Muysken distingue trois processus de juxtaposition des codes : l'insertion d'une unité du code A dans une phrase construite selon la grammaire du code B (l'insertion prototypique dans le modèle MLF), l'alternance entre les grammaires du code A et du code B (semblable à la notion d'îlot dans le modèle MLF), et la lexicalisation congruente (« congruent lexicalisation »), où les deux grammaires sont suffisamment similaires pour que les deux codes puissent contribuer du matériel lexical en même temps pour « vêtir » une même structure grammaticale.

L'insertion figure également dans des approches qui ne sont pas axées sur la syntaxe, comme celle de Auer (1995) en analyse conversationnelle. Auer distingue quatre patrons de changement de langue, selon ce que chacun implique pour l'interaction en cours :

- le nouveau choix de langue, où un locuteur change de langue pendant son tour de parole, et de là, l'interaction continue dans cette nouvelle langue;
- la négociation du choix de langue, où les deux locuteurs⁵⁷ utilisent chacun une langue différente. Ils pourraient ou non finir par converger sur une seule langue à utiliser comme « langue de l'interaction »;

⁵⁷ Auer (1995) envisage spécifiquement un échange dialogal.

- le non-choix de langue, où le premier locuteur change souvent de langue pendant son tour de parole, de sorte qu'il n'y a pas clairement une langue de l'interaction. Le deuxième locuteur peut ensuite choisir la langue qu'il préfère ou bien continuer lui aussi dans le non-choix;
- l'insertion, c'est-à-dire une alternance de codes visant une unité en particulier (« unit-focused switching », p. 126), qui représente une interruption momentanée dans la langue de l'interaction plutôt qu'un changement de celle-ci.

Cependant, aucune de ces approches n'est tout à fait adéquate pour la présente étude. Une définition de l'insertion comme étant la production de matériel extracodique sans changer de grammaire de production n'est pas compatible avec l'approche de la présente étude, selon laquelle il n'y a pas deux grammaires distinctes dont l'interaction doit être expliquée, mais bien un répertoire de ressources lexicales, morphologiques, syntaxiques, discursives, etc., acquises par l'expérience avec l'usage et associées par cette même expérience à un ou des codes communicatifs. La définition de Auer (1995) pourrait convenir, mais il n'en reste pas moins que la notion d'insertion elle-même ne capte pas bien la distinction qui est recherchée ici : il importe peu qu'une unité soit insérée ou non dans un cadre grammatical ou discursif quelconque. Plutôt, il importe qu'il soit possible d'identifier l'unité en question comme celle étant effectivement l'objet – ou plutôt, pour ainsi dire, l'objectif, s'il y en a un – de l'occurrence d'alternance codique en question.

Pour ces raisons, je vais plutôt parler d'unités « isolables ». Par là, j'entends des unités pour lesquelles il y a peu de chances qu'elles aient fait l'objet d'une alternance codique par anticipation ou par enchaînement, autrement dit, des unités qui, on peut en être relativement certain, ont occasionné l'occurrence d'alternance codique dont elles sont l'objet. Cela comprend les insertions selon la définition de Auer (1995), c'est-à-dire les occurrences d'une unité seule, ainsi que les unités qui amorcent une occurrence plus longue, mais qui sont séparées du reste de l'occurrence d'une façon qui réduit les chances que ce soit une alternance par anticipation. Les longues pauses, les fins de phrase syntaxique (indiquées par une prosodie de fin de phrase déclarative, exclamative ou interrogative) et les changements de tour de parole ont tous été considérés comme des indicateurs d'isolabilité. Cependant, les bribes, les répétitions, les reformulations, les interruptions et les chevauchements n'ont pas été considérés comme suffisants pour constituer l'isolabilité, car ils n'indiquent pas avec le même degré de certitude la séparation de l'unité en question de ce qui suit ou qui aurait pu suivre, dans l'esprit du locuteur.

Il importe de souligner ici que la notion d'isolabilité est une catégorie d'analyse imposée sur les données. Tout comme l'exclusion des unités dont l'association par les locuteurs à l'un ou l'autre code est plutôt difficile à établir (voir la section 4.1.4), l'isolabilité est employée dans le but de restreindre la présente étude à l'analyse des occurrences les moins ambiguës possible – tout en admettant qu'une certaine ambiguïté est

incontournable, car il est impossible pour un analyste externe de déterminer avec certitude pourquoi et comment un locuteur donné a procédé à la formulation d'un énoncé donné. Restreindre l'analyse à des unités isolables n'implique en rien que de tels usages ont un statut particulier auprès des locuteurs, ou que des usages non isolables ne donneraient pas lieu à des observations significatives. Force est d'admettre que cette restriction, quoique nécessaire en vue des objectifs de la présente étude, a nécessairement des conséquences sur l'analyse, notamment en ce qu'elle donne une prépondérance à des unités qui sont plus souvent employées comme insertion ou en début de phrase dans le corpus étudié.

4.2.2 Les occurrences en anglais

Comme il a été mentionné dans la section 4.1.2, pour l'insertion d'une unité seule du code B dans un discours du code A, le retour à l'utilisation du code A par la suite a été considéré comme la fin de l'insertion et non comme une nouvelle juxtaposition de codes. Ainsi, comme les occurrences d'alternance codique repérées dans le corpus Beaulieu et le corpus PFC consistent uniquement en des insertions d'une unité de l'anglais dans un discours autrement en français (*cf.* la section 4.3), il n'y a dans les faits pas d'occurrences d'alternance codique vers le français dans ces deux corpus. Pour que l'analyse soit la plus comparable possible pour les trois sous-corpus, il a été décidé de cibler les occurrences d'alternance codique vers l'anglais. Des occurrences en français figurent toutefois dans l'analyse lorsqu'il s'agit de comparer les usages des unités ciblées avec ceux des éléments français du même paradigme lexical multicode (*cf.* la section 2.4.2). Il est à noter, cependant, qu'en raison du manque d'occurrences comparables d'alternance codique en français pour les deux corpus de l'Alberta, ces comparaisons ont pu être faites avec le corpus FAC seulement.

4.3 Repérage des occurrences d'alternance codique dans le corpus

Le repérage des occurrences d'alternance codique a été accompli en écoutant ou en visionnant chaque enregistrement sélectionné dans la constitution du corpus d'analyse (*cf.* la section 3.3). Toute occurrence correspondant à la définition fonctionnelle établie à la section 4.1 a été notée avec ses informations de référence (le sous-corpus et l'enregistrement dans lesquels l'occurrence se trouve, le locuteur qui l'a produite, et le temps écoulé dans l'enregistrement au moment de sa production). Le tableau 4.1 présente un survol quantitatif des occurrences d'alternance codique ainsi repérées dans le corpus d'analyse.

Ainsi, parmi 2 993 occurrences d'alternance codique en tout, ce sont les 1 227 occurrences isolables en anglais (dont 954 proviennent du corpus FAC) qui constituent l'objet principal de la présente étude. Avant d'aller plus en détail sur les occurrences ciblées, cette vue d'ensemble permet de soulever quelques points, déjà évoqués au chapitre 3, sur lesquels il importe d'insister ici.

L'observation la plus frappante qui ressort du tableau 4.1 est la quantité d'occurrences produites par les locuteurs du corpus FAC, soit 2 746, par rapport à un total de 247 dans les deux sous-corpus de l'Alberta. Les types d'occurrences produites sont aussi remarquables : le corpus FAC est le seul sous-corpus où on observe de l'alternance codique vers le français ainsi qu'une quantité significative d'occurrences non isolables (1 376 en tout, dont 590 en anglais et 786 en français). Tandis que les locuteurs du corpus Beaulieu et du corpus PFC parlent principalement en français, avec des insertions en anglais plus ou moins fréquentes, les locuteurs du corpus FAC ont tendance, dans l'ensemble, à parler tantôt en français avec parfois de l'alternance codique vers l'anglais, tantôt en anglais avec parfois de l'alternance codique vers le français.⁵⁸

TABLEAU 4.1 : OCCURRENCES D'ALTERNANCE CODIQUE DANS LE CORPUS D'ANALYSE PAR TYPE D'OCCURRENCE

	# occurrences d'AC	# occurrences d'AC en anglais	# occurrences d'AC en anglais, isolables	# occurrences d'AC en français	# occurrences d'AC en français, isolables
Corpus Beaulieu (16 enregistrements, 9,5 heures)	162 (17,1 occ./heure)	162	150	0	0
Corpus PFC (4 enregistrements, 1 heure)	85 (85 occ./heure)	85	83	0	0
Corpus FAC (9 enregistrements, 15 heures)	2 746 (99,5 occ./heure)	1 591	954	1 157	461
Total (29 enregistrements, 25,5 heures)	2 993	1 838	1 227	1 157	461

Plusieurs facteurs pourraient contribuer à ces différences. Il a déjà été évoqué à la section 3.2 que le bagage sociohistorique de l'Alberta francophone a renforcé une séparation entre les espaces sociaux francophones et anglophones, ce qui a fait en sorte que les locuteurs franco-albertains préfèrent employer le français et l'anglais de façon complémentaire. Autrement dit, ils valorisent l'emploi des connaissances en anglais dans le contexte approprié, et ils valorisent aussi, dans des contextes où le français est la langue appropriée, l'emploi d'un « bon français » avec le moins de recours à l'anglais possible. Le tableau 4.1 semble confirmer quantitativement cette attitude : les locuteurs étudiés de l'Alberta ont préféré mener leurs interactions enregistrées principalement en français, tout en incorporant certaines ressources de l'anglais lorsqu'ils ont jugé pertinent de le faire. Les Franco-Américains, pour leur part, ont historiquement fait face à une pression d'assimilation plus forte et plus violente, contre laquelle la volonté de séparation des sphères sociales a plutôt

⁵⁸ Quoique déjà soulevées dans la littérature, les difficultés possibles dans l'identification d'une langue de base d'un énoncé ou d'une partie d'énoncé, de laquelle l'alternance codique représente un départ, ne sont pas discutées ici (mais on pourra voir Auer, 2000 ; Gafaranga et Torras, 2002), car dans le corpus d'analyse actuel, de telles difficultés ne sont pas survenues. Cela n'implique pas toutefois que ce serait le cas pour un autre corpus semblable ou même un autre échantillon des mêmes corpus utilisés dans la présente étude.

nui qu'aïdé. Ils ne valorisent donc pas particulièrement l'emploi du français sans recours à l'anglais; en fait, les locuteurs étudiés expriment, de même que les participants de Fox (2007), une préférence pour le mélange de codes plutôt libre.⁵⁹ Cette préférence semble effectivement être présente chez les locuteurs étudiés, comme le reflète le tableau 4.1.

Un autre facteur à considérer est la nature des enquêtes qui ont donné lieu au corpus analysé (cf. la section 3.1.1.2 et la section 3.2.1). Le corpus FAC résulte d'une série d'entretiens plutôt que d'une enquête structurée, et ces entretiens ne portent pas directement sur le français mais sur l'histoire de la communauté, sous la forme du vécu de l'interviewé. L'animateur est un membre de la communauté qui semble connaître et être connu de ses interlocuteurs. Les locuteurs parlent ensemble de leurs expériences et des sujets qui leur sont significatifs ou qui suscitent leur intérêt en prenant souvent des détours. La nature de la conversation est donc la plus libre des trois sous-corpus, et la situation de communication la plus proche de l'immédiat communicatif idéal. Cela contribue sans doute au fait que le corpus FAC comporte le plus d'occurrences dans l'absolu et le taux d'occurrences par heure le plus élevé des trois sous-corpus.

Le corpus PFC a également un taux relativement élevé d'occurrences d'alternance codique par heure, malgré le fait que sa fréquence absolue d'occurrences soit la plus basse des trois en raison de la petite taille de ce sous-corpus. Comme dans le corpus FAC, chaque animatrice dans le corpus PFC est membre de la communauté, et elle connaît et est connue de ses interlocuteurs. Quoique l'enquête porte directement sur le français, les sujets de discussion portent plutôt sur la vie quotidienne. Ainsi, la situation de communication s'approche aussi de l'immédiat communicatif, mais dans une moindre mesure que dans le corpus FAC. La conversation est également légèrement moins libre, avec une liste de sujets de discussion préparée à l'avance, mais peut prendre des détours à l'intérieur des sujets prévus.

Finalement, le corpus Beaulieu a le taux d'occurrences par heure le plus bas des trois sous-corpus. Cela peut s'expliquer en bonne partie par le fait que le contexte de l'enquête s'éloigne le plus de l'immédiat communicatif idéal (sans toutefois sortir de l'immédiat communicatif en général). La relation entre l'enquêtrice et ses interlocuteurs, quoique cordiale, est la plus hiérarchique des trois sous-corpus. En tant qu'universitaire et enseignante de français, l'enquêtrice pourrait représenter le « bon français » pour les locuteurs. En voulant faire bonne impression, il se peut que les locuteurs aient surveillé leur langage et évité d'employer l'anglais plus qu'à leur habitude. De plus, les questions posées aux locuteurs portent directement sur le sujet de la langue et ce qui constitue ou non, selon eux, un français approprié dans un contexte de soins médicaux. La conversation est

⁵⁹ « On parle anglais on parle français ça fait pas de différence » (FAC 01). « Moi je ne parle ni français ni anglais. Je parle français » (FAC 01). « Half French half English [...] is typical of the St. John Valley. We speak whichever way we want » (FAC 08).

donc la moins libre des trois, avec une thématique restreinte et des questions précises. Qui plus est, le fait que la discussion relève en particulier de la langue et de la façon de s'exprimer aurait pu mener les locuteurs à être encore plus conscients de leurs usages et à se surveiller encore plus.

4.4 Catégorisation des occurrences isolables repérées

Parmi les 1 227 occurrences isolables en anglais dans le corpus d'analyse, on trouve 597 unités différentes (c'est-à-dire, des occurrences types⁶⁰). Il a donc été nécessaire de restreindre les analyses détaillées à une sélection de ces unités. On pourrait penser à sélectionner les unités les plus fréquentes; cependant, il a déjà été mentionné à la section 3.1.2.3 que la présence de l'alternance codique est en général beaucoup plus rare dans un corpus que la présence, par exemple, d'un phonème particulier ou même d'une unité lexicale particulière. La récurrence d'une même unité spécifique faisant objet de l'alternance codique est encore moins fréquente. Effectivement, parmi les 597 occurrences types, seulement 11 sont produites 10 fois ou plus dans l'entièreté du corpus (et seulement quatre 20 fois et plus), tandis que 422 ne sont produites qu'une seule fois. La fréquence d'occurrence des unités individuelles pourrait donc occulter des tendances significatives dans l'emploi des ressources multicodeuses par les locuteurs étudiés.

Ainsi, les occurrences ont aussi été regroupées en catégories qui font ressortir certaines de ces tendances observées. Les catégories ont été constituées autour de ressemblances lexicales (champs lexicaux) et/ou de ressemblances dans la fonction discursive. Les catégories potentielles ont été identifiées de manière inductive (Thomas, 2006). Seules les catégories avec plus que 10 occurrences correspondantes ont été retenues comme des catégories potentiellement significatives relativement au corpus d'analyse. Il est à noter toutefois que chaque unité type différente à l'intérieur de ces catégories pourrait présenter des différences importantes quant à ses contextes d'usage (y compris des différences quant aux communautés et aux locuteurs qui l'utilisent le plus) et à la stratégie d'expression qu'elle pourrait représenter. Lorsque les données disponibles le permettent, ces différences sont explorées au chapitre 5, où les catégories sont présentées plus en détail. Le tableau 4.2 présente une vue d'ensemble des catégories identifiées dans le corpus d'analyse.

⁶⁰ Afin d'opérer ce calcul, il a d'abord été nécessaire d'uniformiser les formes variantes des unités isolables présentes dans le corpus afin de créer des occurrences types. Ainsi, toutes les variantes d'une même base lexicale ont été réduites à cette base. Les noms ont été mis au singulier, à l'exception de ceux qui ne sont pas employés au singulier dans l'usage ordinaire (ex. *pants*) et ceux pour lesquels le pluriel représente une signification distincte (ex. *greens*). Les verbes ont été mis à leur forme infinitivale (« bare infinitive »). Le cas échéant, un indice a été employé pour désambiguïser des homographes hétérocatégoriels : nom (n), verbe (v) ou adjectif (a). Les unités lexicales constituées de plus d'un mot (cf. la section 4.1.3) ont été uniformisées en fonction (1) de l'expression ou groupe de mots qui constitue l'unité, dans son entièreté, ou (2) le cas échéant, en fonction du mot ou groupe de mots qui rend le reste de la séquence prévisible. Ainsi, l'occurrence type de « la fille à [untel] » serait *fille*, et l'occurrence type de « nineteen thirties » serait *nineteen thirties*, et non *nineteen*, *thirties* ou *thirty*.

TABLEAU 4.2 : LES CATÉGORIES FRÉQUENTES D'OCCURRENCES ISOLABLES DANS LE CORPUS D'ANALYSE

Nom de la catégorie	Exemples des unités types incluses dans la catégorie	# occurrences correspondant à la catégorie (sur l'ensemble du corpus d'analyse)	% des occurrences isolables en anglais représenté par la catégorie
<i>yeah / yes / yep</i>	<i>yeah, yes, yep</i>	225	18,3
marqueurs pragmatiques	<i>so, now, anyway, but, well</i>	68	5,5
chiffres	<i>eighty-one, eight years, nineteen fifty</i>	60	4,9
expressions exclamatives	<i>holy boys!, oh my word!, oh geez!, oh gosh!</i>	53	4,3
formules (« chunks »)	<i>matter of time, boys and girls, room and board, back at it</i>	46	3,7
<i>right / that's right / is that right?</i>	<i>right, that's right, is that right?</i>	46	3,7
formes d'adresse	<i>you, darling, mister, honey</i>	36	2,9
termes de relation familiale	<i>stepfather, first cousins, grandfather, boyfriend, mom</i>	18	1,5
marqueurs de position épistémique	<i>oh I see, I think so, I wouldn't know, I heard</i>	15	1,2

4.5 Protocole de transcription

Les occurrences correspondant à chaque catégorie identifiée à la section 4.4 et sélectionnée pour l'analyse en détail (voir le chapitre 5) ont ensuite été transcrites en préparation pour l'analyse, ainsi que, le cas échéant, les occurrences isolables en français correspondant à des éléments du paradigme lexical multicodeur des unités visées. La transcription comprend le cotexte allant de quelques secondes avant l'occurrence jusqu'à quelques secondes après. Une limite précise n'a pas été fixée quant à la quantité du cotexte à prendre en compte; plutôt, la transcription comprend le cotexte nécessaire pour bien interpréter l'occurrence. Les transcriptions ont aussi été annotées avec des informations sur la thématique de discussion et toute autre information contextuelle (au sens large de *contexte*, voir Kerbrat-Orecchioni, 2007) jugée pertinente pour l'interprétation de l'occurrence et pour l'identification des habitudes d'usage relatives à l'unité en question. Les conventions de transcription sont présentées à la page ii.

Aucune transcription ne peut être une reproduction complètement fidèle de l'interaction dont elle constitue la trace écrite (Blanche-Benveniste, 2010). Notamment, il faut réduire l'immense quantité d'information qui est transmise dans une interaction orale à une quantité qui sera lisible sur la page, ou autrement dit il faut sélectionner d'avance les données auxquelles on veut porter une attention particulière, selon l'approche et les objectifs de l'analyse. En cela, la transcription constitue elle aussi une forme de pré-analyse.

Dans le cas de la présente étude, les transcriptions visent un maximum de simplicité tout en essayant d'éviter d'imposer des règles de l'écrit sur l'oral spontané, et tout en prenant en compte certains phénomènes de l'oral qui peuvent être des indices observables de la perception ou des intentions des locuteurs. Ainsi,

l'orthographe standard a été utilisée, à la fois pour la simplicité de la lecture (Blanche-Benveniste, 2010) et pour éviter de folkloriser les prononciations non standards. Cependant, lorsque l'orthographe standard ne permet pas de rendre compte des traits de l'oral potentiellement significatifs, des formes communément utilisées dans l'écrit spontané ont été adoptées, par exemple *dunno*, *chuis*, *tsé*, etc. Notamment, la forme *y* est utilisée pour transcrire à la fois *il*, *ils* et *il y*, lorsque réalisés en /i/. Les signes de ponctuation n'ont pas été imposés sur les transcriptions selon les règles de l'écrit, qui présupposerait l'existence de certaines structures syntaxiques, mais selon la prosodie en tant qu'indice des intentions de communication du locuteur. Ainsi le point, le point d'interrogation et le point d'exclamation représentent une pente mélodique déclarative, interrogative et emphatique respectivement, sans que ce qui précède soit nécessairement une phrase grammaticalement complète de forme syntaxique déclarative, interrogative ou emphatique selon les normes de l'écrit. La virgule, quant à elle, représente une séparation prosodique faite par le locuteur qui ne correspond pas à l'une des autres marques prosodiques prévues, et est utilisée seulement là où l'absence d'un tel séparateur visuel pourrait gêner l'interprétation de l'énoncé lors de la lecture.

D'autres éléments retenus comme potentiellement significatifs pour l'interprétation en contexte des occurrences sont les tours de parole, les chevauchements et les énoncés de *backchannel* (aussi appelés *rétroactions* ou *signes/signaux d'écoute*) ainsi qu'un ensemble de phénomènes souvent identifiés comme des dysfluidités : les pauses, les hésitations ou auto-interruptions et les prolongements de segments phonologiques. Les dysfluidités, notamment, peuvent accomplir plusieurs fonctions dans le discours, dont certaines sont d'un intérêt particulier pour la présente étude : indiquer une recherche lexicale en cours, différencier un nouvel événement discursif, indiquer qu'un segment de son discours nécessite une attention particulière de son interlocuteur, indiquer une rupture avec les normes ou les attentes sur l'interaction et même faciliter spécifiquement l'intégration dans l'énoncé d'une occurrence d'alternance codique (Brennan et Williams, 1995 ; Clark et Fox Tree, 2002 ; Fraundorf et Watson, 2011 ; Hlavac, 2011). Toutefois, les transcriptions ne se veulent pas des transcriptions fines de la phonétique, de la phonologie ou de la prosodie, ni du visuel ou du mimogestuel; ces éléments sont toutefois pris en note lorsqu'ils sont jugés pertinents pour l'interprétation de l'occurrence visée. Ainsi, tous les éléments prosodiques indiqués dans les transcriptions ont été repérés perceptuellement lors de l'écoute; aucune prise de mesure de la longueur des pauses ou des prolongements ni de la pente mélodique des énoncés n'a été effectuée.

4.6 Récapitulatif du chapitre 4

Ce chapitre a porté sur les choix méthodologiques nécessaires afin de procéder au traitement des données dans le corpus d'analyse, autrement dit sur les pré-analyses. Notamment, des critères de définition de la notion d'occurrence d'alternance codique ont été établis de façon à pouvoir appliquer cette notion à un corpus de productions effectives, en excluant les unités desquelles le statut codique est trop ambigu pour traiter dans

la présente étude. Une fois cette définition en place, il a été possible de repérer les occurrences d'alternance codique dans le corpus. Ensuite, il a été décidé de se concentrer sur les occurrences en anglais, par souci de la comparabilité à travers les trois sous-corpus, et sur les occurrences en position isolable, par souci de lever l'ambiguïté quant à l'unité dont le sens aurait pu motiver l'alternance codique. Cette présélection faite, il a été possible d'identifier des catégories d'unités à analyser et de transcrire les occurrences associées à ces catégories. Le prochain chapitre présente les analyses proprement dites qui ont été effectuées sur les données ainsi sélectionnées et préparées.

5 Analyses et résultats

Ce chapitre présente les analyses qui ont été effectuées sur les données du corpus et les résultats qui en découlent. La nature qualitative d'une bonne partie de ces analyses implique nécessairement l'interprétation des résultats quantifiés; ainsi ce chapitre présente également la discussion des résultats à mesure qu'ils sont présentés. Avant de se tourner vers les catégories d'unités lexicales identifiées à la section 4.4, la section 5.1 présente un aperçu des catégories fréquentes identifiées en français au cours des préparations pour les analyses comparatives; quoique les catégories en français ne sont pas l'objet principal d'analyse du présent chapitre, une comparaison avec les catégories fréquentes en anglais permet de faire ressortir quelques observations qui méritent d'être discutées ici. À la section 5.2 sont présentées les catégories identifiées en anglais qui ne sont pas analysées en profondeur ici pour diverses raisons. Les sections suivantes présentent les analyses en profondeur des autres catégories : *yeah / yes / yep* (section 5.3), *right / that's right / is that right?* (section 5.4) et les marqueurs pragmatiques (section 5.5.). Enfin, la section 5.6 résume les résultats de la présente étude et les points de discussion qu'ils soulèvent.

Il est à noter que chaque catégorie pourrait en soi faire l'objet d'une étude à part entière – autrement dit, ce qui est présenté ici ne se veut pas une étude maximale approfondie de chaque forme ou chaque catégorie, mais plutôt une observation des tendances dans le corpus d'analyse. Il importe également de garder à l'esprit certaines limites globales importantes. Premièrement, les analyses ci-dessous portent uniquement sur des occurrences *isolables* (cf. la section 4.2), non pas sur tout emploi de l'unité en question dans le corpus d'analyse. Cette restriction pourrait occulter certaines tendances et donner prépondérance à d'autres. Deuxièmement, ces analyses, comme le repérage et la transcription, ont été effectuées selon le jugement d'un seul analyste, c'est-à-dire le mien. Comme avec toute analyse qualitative, il se peut qu'un autre analyste établirait des catégories d'analyse différentes, et qu'il observerait des tendances différentes avec les mêmes données. Finalement, il est à rappeler que des occurrences d'alternance codique vers le français, y compris les occurrences isolables en français, ne surviennent que dans le corpus FAC. Toute comparaison anglais-français se limite donc à ce sous-corpus et ne reflète que les habitudes d'usage des locuteurs franco-américains étudiés.

5.1 Comparaison globale : les catégories fréquentes en français

Comme la présente étude vise principalement les occurrences isolables en anglais (cf. la section 4.2), seules les occurrences isolables en français faisant l'objet d'une comparaison à une ou des unités de l'anglais étudiées ont été analysées en détail. Cependant, afin de pouvoir effectuer ces comparaisons, toute occurrence isolable en français a dû être repérée, se voir attribuée une occurrence-type et, le cas échéant, une catégorie. Ce faisant, des tendances sont ressorties quant aux catégories et aux unités fréquentes faisant l'objet

d'alternance codique en position isolable en français, en particulier par rapport aux catégories et unités fréquentes en anglais. Le tableau 5.1 présente les catégories fréquentes identifiées en français et en anglais. Le nombre d'occurrences associées à chaque catégorie est indiqué ainsi que le pourcentage du total des occurrences isolables dans le FAC pour la langue en question (n = 461 en français, n = 954 en anglais) que cette catégorie représente. Pour que la comparaison soit juste, seules les occurrences isolables en anglais repérées dans le corpus FAC ont été prises en considération.

Ainsi, on voit que les locuteurs étudiés du corpus FAC ont recours à l'alternance codique en français et en anglais, quoiqu'à des fréquences différentes, pour des catégories comme *yeah / yes / yep* et *ouais / oui*, des marqueurs pragmatiques, des expressions exclamatives, des formules, des termes de relation familiale et des marqueurs de position épistémique. Le nombre de catégories communes entre l'anglais et le français est probablement au moins en partie un artefact de la restriction à des occurrences en position isolable, qui donne

TABLEAU 5.1 : LES CATÉGORIES FRÉQUENTES DES OCCURRENCES ISOLABLES EN ANGLAIS ET EN FRANÇAIS, CORPUS FAC SEULEMENT

Nom de la catégorie (français)	Exemples des unités types incluses dans la catégorie	occurrences correspondant à la catégorie (FAC seulement)		Nom de la catégorie (anglais)	Exemples des unités types incluses dans la catégorie	# occurrences correspondant à la catégorie (FAC seulement)	
		#	% occ. isol fr.			#	% occ. isol. angl.
marqueurs pragmatiques	<i>là, pis, mais, ben,</i>	215	46,6	<i>yeah / yes / yep</i>	<i>yeah, yes, yep</i>	174	18,2
<i>ouais / oui</i>	<i>ouais, oui</i>	55	11,9	chiffres	<i>eighty-one, eight years, nineteen fifty</i>	60	6,3
expressions présentatives	<i>ça, y avait, on avait</i>	38	8,2	marqueurs pragmatiques	<i>so, now, anyway, but, well</i>	59	6,2
pronoms d'insistance (pronoms détachés)	<i>moi, toi, nous autres</i>	35	7,6	expressions exclamatives	<i>holy boys!, oh my word!, oh geez!</i>	48	5,0
formules (« chunks »)	<i>venu au monde, dans le bois</i>	33	7,2	<i>right / that's right / is that right</i>	<i>right, that's right, is that right?</i>	46	4,8
termes de relation familiale	<i>matante, mononcle, pèpère, mémère</i>	32	6,9	formules (« chunks »)	<i>matter of time, room and board, back at it</i>	27	2,8
expressions exclamatives	<i>voyons, mon dieu</i>	17	3,7	termes de relation familiale	<i>first cousins, grandfather, mom</i>	16	1,7
marqueurs de position épistémique	<i>sais pas, pour moi, pas au courant</i>	12	2,6	marqueurs de position épistémique	<i>oh I see, I think so, I wouldn't know</i>	14	1,5
				formes d'adresse	<i>mister, Father</i>	5	0,5

prépondérance à des unités qui sont plus propices à l'insertion ou à l'isolation syntaxique et discursive. Il est probablement aussi dû en partie à la situation de communication plutôt immédiate et informelle et à la nature de l'enquête, en particulier aux thématiques abordées (la vie quotidienne, les expériences personnelles). Ces facteurs rendent plus prévalentes des unités qui aident à structurer la conversation (*yeah / yes / yep, ouais / oui*, et autres marqueurs discursifs, les marqueurs de position épistémique) ainsi que des unités relevant de la famille (associée à la vie quotidienne), des émotions (associées à la subjectivité et les expériences personnelles) et des formules et expressions courantes (associées à l'oral spontané).

Cependant, il est intéressant de noter que, malgré le nombre total moindre d'occurrences isolables en français dans ce corpus (461 par rapport à 954 en anglais), plusieurs catégories ressortent qui ne sont pas fréquentes en anglais, voire pas présentes du tout : les expressions présentatives et les pronoms d'insistance. Dans le cas des expressions présentatives, Leclercq (2009) suggère qu'elles permettent de soulever une partie de l'ambiguïté inhérente dans les verbes conjugués au présent en français, en ancrant l'énoncé dans une situation spécifique (tandis qu'en anglais, on peut se servir du présent progressif [-*ing*] pour cet ancrage). Elle remarque, par conséquent, que l'emploi des présentatifs est beaucoup plus fréquent en français qu'en anglais, voire obligatoire, lors de la narration. De fait, le corpus FAC contient beaucoup de narration, en ce que les locuteurs racontent des événements vécus. Cependant, l'interprétation de Leclercq (2009) simplifie excessivement la signification et les emplois des présentatifs en français (sans parler de la signification des formes verbales en anglais; voir à ce sujet Hirtle, 2007), et ses données ne lui permettent pas de comparer des présentatifs en français et en anglais. Il serait intéressant de se pencher plus en détail sur la question des visées référentielles traduites par des présentatifs en français à l'oral (comme l'a fait Rabatel, 2001 pour l'écrit) et une éventuelle comparaison avec les présentatifs en anglais présents dans le corpus, ce qui dépasse toutefois les objectifs de la présente étude.

Dans le cas des pronoms d'insistance, il n'y a à ma connaissance pas d'études comparatives entre le français et l'anglais. De façon générale, cependant, on peut constater que dans l'usage général il est plus habituel, c'est-à-dire plus fréquent, de produire des pronoms détachés en français oral qu'en anglais oral et dans un plus grand nombre de positions dans l'énoncé. On peut penser, par exemple, à la fréquence et à la saillance relative des phrases comme *En tout cas, moi, j'aime pas ça* par rapport à *Anyway, me, I don't like that*, ou encore *Je pense pas, moi* par comparaison à *I don't think so, me*. Cela pourrait indiquer que les pronoms détachés du français constituent une ressource permettant d'insister sur la subjectivité et sur la position énonciative que prend le locuteur par rapport à ses propres paroles, ressource qui n'est pas disponible de la même façon dans le code anglais. Encore ici, une analyse fine de ces unités en français serait nécessaire pour déterminer les habitudes d'usage, y compris les visées référentielles exprimées, chez les locuteurs étudiés.

TABLEAU 5.2 : LES MARQUEURS PRAGMATIQUES ISOLABLES EN FRANÇAIS ET EN ANGLAIS,
COMPARAISON INDIVIDUELLE

Marqueur (français)	# occurrences ⁶¹	Marqueur (anglais)	# occurrences
<i>là</i>	116	<i>so</i>	14
<i>pis</i>	39	<i>now</i>	10
<i>ben</i>	17	<i>anyway</i>	10
<i>parce que</i>	12	<i>but / yes but</i>	7
<i>mais</i>	12	<i>you know / y'know</i>	6
<i>bon</i>	7	<i>too</i>	5
<i>ça fait (que)</i>	6	<i>well</i>	5
<i>ça</i>	5	<i>see</i>	4
<i>astheure</i>	4	<i>of course</i>	3
<i>hein</i>	3	<i>if / and if</i>	2
<i>avec</i>	1	<i>like</i>	2
<i>ici</i>	1	<i>because</i>	1
<i>pour moi</i>	1	<i>first of all</i>	1
<i>tout d'un coup</i>	1		

Une autre observation que permettent les résultats du tableau 5.1 est la prévalence des marqueurs pragmatiques en français, qui dépassent en fréquence même la catégorie *yeah / yes* dans le corpus FAC. À des fins de comparaison, le tableau 5.2 présente les marqueurs individuels qui constituent la catégorie en français et celle en anglais. On voit que la prévalence des marqueurs pragmatiques en français faisant l'objet de l'alternance codique est principalement due à la fréquence élevée de *là*, *pis*, et *ben*. Les marqueurs pragmatiques feront l'objet d'une analyse plus détaillée ci-dessous (voir la section 5.5), mais on peut déjà remarquer que, de façon générale, les plus fréquents en français sont ceux qui permettent d'exprimer des sens qui ne sont pas lexicalisés de la même façon en anglais : *là* peut se traduire entre autres par *now*, *then*, et *there*, en plus de sa fonction en tant que « ponctuant de la langue » (Vincent, 1993), dans laquelle il sert à démarquer des groupes prosodiques; *pis* peut se traduire par *and* et *then* ou encore par *so what?*; et *ben* peut se traduire par *well*, mais peut aussi contribuer à des marqueurs composites, par exemple, *ah ben oui!* dont le sens ressemble *of course*.

Globalement, alors, cet aperçu de l'alternance codique vers le français observée dans le corpus permet déjà de confirmer l'hypothèse qui sous-tend la présente étude : lorsqu'ils parlent en anglais, les locuteurs étudiés du corpus FAC se servent de certaines ressources en français qui leur permettent de traduire des visées référentielles qui seraient moins facilement traduisibles en anglais. De plus, le fait que ces unités sont employées dans des contextes variés suggère que l'alternance codique dans ces cas ne relève pas d'un problème de

⁶¹ À la différence des marqueurs pragmatiques observés en anglais, les marqueurs pragmatiques en français sont souvent eux-mêmes composés de plusieurs formes constituant elles aussi des marqueurs pragmatiques (par exemple, *pis là*, *et pis*, *mais bon*, *non mais*, etc.). Une seule occurrence peut alors correspondre à plusieurs formes récurrentes identifiées à l'intérieur de la catégorie. Ainsi, les chiffres rapportés dans cette colonne s'additionnent à un total qui dépasse le nombre total d'occurrences repérées correspondant à la catégorie.

connaissances lexicales ou d'accès lexical. Autrement dit, les locuteurs semblent en effet se servir de l'alternance codique comme stratégie d'expression.

5.2 Catégories faisant l'objet d'observations générales seulement

Certaines des catégories d'unités identifiées comme fréquentes dans le corpus d'analyse (*cf.* le tableau 4.2) se sont toutefois révélées moins propices à une analyse détaillée, et ce, pour diverses raisons. Cette section donne un bref aperçu de ces catégories et des raisons de leur exclusion avant de passer, dans la prochaine section, aux catégories analysées plus en profondeur.

5.2.1 Les chiffres

La catégorie des chiffres comprend à la fois des années (par exemple, « nineteen-seventy-three », FAC 01), des décennies (par exemple, « early forties », FAC 06), des âges (« sixty-four », FAC 08), des durées de temps (« ten years », FAC 01), des mesures de poids (« one-seventy-five », FAC 07) et d'autres expressions diverses comportant des chiffres (« three-ply », FAC 01; « one room », FAC 07). La prévalence de cette catégorie est fort intéressante car elle n'était pas prévue dans la littérature sur l'alternance codique.⁶² Toutefois, il n'y a que 4 occurrences isolables en français dans la catégorie des chiffres, et compte tenu de l'hétérogénéité de la catégorie, il serait impossible de constituer un paradigme lexical multicodique (PLM) analysable. Il serait toutefois intéressant, dans de futures recherches, d'analyser l'ensemble des productions de chiffres dans le corpus, y compris ceux en position non isolable, pour avoir un meilleur aperçu du rôle des chiffres en anglais dans les habitudes d'usage de ces locuteurs et des raisons potentielles pour leur prévalence.

5.2.2 Les expressions exclamatives

La prévalence des expressions exclamatives en anglais dans le corpus est également intéressante. On pourrait y voir le rôle de l'alternance codique dans l'expression des émotions ou dans le contournement des tabous (Dewaele, 2010), ou l'association du code anglais à l'enthousiasme et à la franchise (voir aussi la section 5.3.4). Cependant, une étude productive de ces unités nécessiterait des informations détaillées sur les valeurs qu'attribuent les locuteurs étudiés aux langues en question, aux tabous sociaux impliqués et aux unités particulières analysées. Il faudrait aussi des informations détaillées sur les expériences potentiellement pertinentes de chaque locuteur (l'éducation religieuse, les valeurs des parents, les expériences formatrices quant à l'expression de telles émotions, etc.). Ces informations ne sont évidemment pas disponibles dans le cas de la présente étude.

⁶² Quoique des études sur la cognition numérique, par exemple Spelke et Tsivkin (2001), ont déjà soulevé l'intérêt d'étudier le traitement des chiffres par les bilingues.

5.2.3 Les formules

La catégorie des formules regroupe des ensembles d'unités qui fonctionnent eux-mêmes comme des unités, tels que « back and forth », « massive heart attack », « jack of all trades master of none » (FAC 01), « by the way », « to make a long story short » (FAC 07), « go-to guy » (PFC 03), etc. Les questions de définition et même de désignation de cette catégorie (« chunk », « bundle », « cluster », « expression ») sont complexes et ne seront pas discutées en détail ici (mais voir Wray, 2002 pour une revue de la littérature sur ce sujet). Le fait que cette catégorie soit si prévalente pourrait être vu comme une confirmation de l'hypothèse de départ de la présente étude, à savoir que l'alternance codique pourrait être produite parce que des ressources dans l'autre langue permettent de mieux traduire une visée référentielle donnée. Spontanément, il est difficile de trouver pour beaucoup des formules en anglais repérées dans le corpus des séquences en français dans l'usage général dont la signification est convergente. Même lorsque cela est possible (par exemple, « back and forth » est souvent traduit par « va-et-vient »), on peut postuler qu'il y a un sentiment de non-convergence entre les deux suffisamment fort pour engendrer de l'alternance codique dans le but de traduire la visée référentielle spécifique voulue (Backus, 2009). Toutefois, compte tenu que la plupart des formules de l'anglais repérées en position isolable dans le corpus d'analyse n'apparaissent qu'une fois chacune et ne s'apparient pas à l'une des formules repérées en français, une analyse plus poussée n'est pas envisageable pour l'instant.

5.2.4 Les formes d'adresse

Les formes d'adresse comprennent des pronoms d'adresse (notamment « you », dans plusieurs enregistrements du corpus Beaulieu) et des appellatifs (« mister », « doctor », FAC 06; « honey », « sweetheart », CB-GRP 01; « guys », CB-IND 01; « dude », CB-IND 03). Toutefois, la prévalence de cette catégorie est plutôt un artefact de l'enquête qui a donné lieu au corpus Beaulieu : Beaulieu (2012) étudiait les attitudes des locuteurs franco-albertains envers, entre autres, différentes formes d'adresse. En même temps, certains des appellatifs pourraient s'associer plutôt à des noms propres, notamment *mister*, *doctor* et *Father*, au point où on pourrait considérer que leur statut codique est ambigu. Cette catégorie n'a donc pas fait l'objet d'une analyse plus détaillée.

5.2.5 Les termes de relation familiale

Les termes de relation familiale ou de parenté comprennent des unités comme « stepfather » (FAC 01), « my wife » (FAC 06), « first cousins » (FAC 07), ou « boyfriend » (PFC 04). Cette catégorie a été exclue pour deux raisons : premièrement, semblable aux appellatifs, certains de ces termes pourraient être plutôt des noms propres dans l'esprit des locuteurs (par exemple, « Mom », FAC 09, et « Dad », PFC 02). Deuxièmement, il y a une forte possibilité que certaines de ces occurrences relèvent plutôt de l'alternance codique en lien avec ce que Fishman (2000 [1965]) appelle le domaine (qui trouve des parallèles dans le « situational switching » de

Blom et Gumperz, 2000 [1972] ; voir aussi les « language modes » de Grosjean, par exemple 2013). Autrement dit, certains locuteurs pourraient utiliser l'anglais ou le français pour faire référence à des membres de leur famille, soit parce qu'il s'agit de leur langue dominante, soit parce qu'elle est celle utilisée pour communiquer avec ces membres de la famille ou pour parler d'eux.

5.2.6 Les marqueurs de position épistémique

Une analyse en profondeur de la catégorie de marqueurs de position épistémique comme « I see », « I can't believe that », « I dunno » (FAC 01), « I think so » ou « I forget where » (FAC 08) aurait été très intéressante, car cette catégorie est semblable aux expressions évidentielles de King et Nadasdi (1999). Comme il en a déjà été question à la section 2.3, ces auteurs ont montré que l'alternance codique était motivée par des différences dans les visées référentielles traduites par de telles expressions en anglais par rapport aux ressources semblables disponibles en français. Cependant, le corpus d'analyse ne contient que 15 occurrences, dont seulement une (« I think », PFC 04) qui a aussi été analysée par King et Nadasdi (1999); il est donc impossible de comparer les données ou de répliquer leur analyse. On peut toutefois constater que, en accord avec les observations de King et Nadasdi (1999), ces expressions sont effectivement productives dans le corpus, en ce que la catégorie figure parmi les plus fréquentes. On constate également une plus grande variété de marqueurs de position épistémique dans le corpus à l'étude que dans les résultats présentés par King et Nadasdi (1999); ce fait démontre la pertinence d'étudier cette catégorie plus en profondeur dans des recherches ultérieures, car elle semble refléter un thème commun entre différentes communautés mais qui varie beaucoup dans sa réalisation.

Les catégories exclues d'une analyse en profondeur ayant été énumérées et les raisons pour leur exclusion ayant été résumées brièvement, on peut maintenant se tourner vers l'analyse en détail des catégories *yeah / yes / yep*, *right / that's right / is that right?* ainsi que de la catégorie des marqueurs pragmatiques.

5.3 *Yeah / yes / yep*

En raison de la fréquence d'occurrence des unités qui la constituent, la catégorie *yeah / yes / yep* peut être analysée sous plusieurs angles. À cet effet, elle est abordée de plusieurs façons dans cette section. La section 5.3.1 présente un aperçu global de la fréquence et de la distribution de cette catégorie dans le corpus d'analyse, ainsi qu'une discussion des implications de celles-ci pour son statut codique. La section 5.3.2 approfondit cet aperçu en présentant la fréquence et la distribution de *yeah*, *yes* et *yep* par locuteur, avec une discussion des facteurs pouvant influencer sur les pratiques langagières des locuteurs étudiés. La section 5.3.3 aborde la question de contexte d'usage sous l'angle des cooccurrences et des tendances dans la thématique de discussion en cours lorsqu'un emploi de *yeah*, *yes* ou *yep* est survenu (ou plus précisément, l'absence de

telles tendances). Enfin, la section 5.3.4 discute des sens en usage et les fonctions discursives de *yeah*, *yes* et *yep*, notamment en comparaison aux emplois de *ouais* et *oui*.

5.3.1 Aperçu global et statut

La catégorie *yeah / yes / yep* est constituée, comme on peut s’y attendre, de toutes les occurrences isolables des unités *yeah*, *yes* et *yep*. Au total, 225 occurrences isolables (ou potentiellement isolables, voir ci-dessous) de ces unités ont été repérées, faisant de *yeah / yes / yep* la catégorie la plus fréquente en anglais dans tout le corpus, dépassant de loin celle des marqueurs pragmatiques de l’anglais (cf. le tableau 4.2). Toutefois, 38 de ces 225 occurrences sont précédées par une unité dont le statut codique pourrait potentiellement être remis en question : des segments que la qualité d’enregistrement ne permet pas d’identifier avec certitude (par exemple *see* par rapport à *ici*), des noms propres francophones et des unités phonologiquement semblables entre les deux langues (par exemple *oki* ou *'ké* par rapport à *okay* ou *'kay*). Ces occurrences ont donc fait l’objet d’une évaluation supplémentaire visant à déterminer s’il y avait suffisamment d’indices du statut codique des unités précédentes (par exemple un nom propre francophone clairement prononcé selon la phonologie française) pour affirmer raisonnablement qu’il s’agissait d’un changement de code entre celle-ci et l’occurrence en question. En définitive, 22 occurrences sur 38 ont été considérées comme de véritables changements de code, tandis que 16 autres ont été jugées de statut trop ambigu. Ces 16 occurrences ont donc été exclues du reste des analyses par souci de restreindre l’analyse aux occurrences représentant le plus certainement possible un changement de code. Au total, 209 occurrences de *yeah*, *yes* et *yep* font l’objet des analyses qui suivent dans cette section. Le tableau 5.3 présente un aperçu de ces 209 occurrences, triées par unité et par sous-corpus.

TABLEAU 5.3 : FRÉQUENCE D’OCCURRENCE DE *YEAH*, *YES* ET *YEP* PAR CORPUS, OCCURRENCES ANALYSÉES SEULEMENT

Sous-corpus	# occurrences <i>yeah</i>	# occurrences <i>yes</i>	# occurrences <i>yep</i> ⁶³	# occurrences (toute la catégorie)
FAC	143	15	4	162
CB (tous)	40	0	0	40
PFC	7	0	0	7
Tous	190	15	4	209

Le tableau 5.3 indique que la prévalence de la catégorie *yeah / yes / yep* dans le corpus d’analyse est principalement le résultat de la prévalence de l’unité *yeah*, qui constitue 77,5 % de l’ensemble de la catégorie et dont des occurrences ont été repérées dans les trois sous-corpus. En revanche, il y a peu d’occurrences de

⁶³ Il est à noter que, quoique seulement 4 occurrences définitives de *yep* ont été observées, la qualité des enregistrements fait en sorte qu’il est parfois difficile de distinguer entre *yep* et une réalisation de *yeah* avec coup de glotte ([jæʔ]). Il se peut alors que le nombre d’occurrences de *yep* soit plus élevé en réalité.

yes et de *yep*, et elles proviennent toutes du corpus FAC. Étant donné la fréquence et la diffusion de *yeah* dans le corpus, il y a lieu de se demander s'il peut être considéré plutôt comme un emprunt. Qui plus est, la nature de l'unité *yeah* fait en sorte qu'on ne peut pas se servir de plusieurs autres critères identifiés à la section 4.1.4 qui pourraient donner une idée de son degré d'intégration. *Yeah* ne présente normalement pas d'intégration morphologique ou syntaxique. Quant à l'intégration discursive, on observe effectivement souvent des dysfluidités (pauses, hésitations, reformulations, etc.) précédant *yeah*, mais cela pourrait être en lien avec sa fonction discursive de réparation (décrite plus en détail à la section 5.3.4) plutôt qu'avec son statut codique. Finalement, *yeah* n'est attestée ni dans *Le Petit Robert* (2016, version en ligne) ni dans *Usito* (2014), mais comme il l'a été expliqué à la section 4.1.4, ce fait ne confirme ni n'infirme rien quant à son statut codique. Force est d'admettre qu'il n'est pas possible de garantir le statut codique de *yeah* dans l'esprit des locuteurs, mais plusieurs facteurs ont fait pencher la balance en faveur de son inclusion dans les unités à analyser. Premièrement, la fréquence et la diffusion ne sont pas suffisantes en elles-mêmes pour indiquer qu'une unité aussi polyvalente que *yeah*, et aussi fréquente en conversation spontanée, soit un emprunt. Deuxièmement, il n'y a pas de modification sémantique observable dans le corpus : *yeah* est employée dans des positions et avec des fonctions ressemblant à la façon dont l'unité est employée en anglais en général. Son utilisation peut être contrastée avec celle de *yes* par des locuteurs francophones au Québec pour exprimer spécifiquement l'affirmation enthousiaste, voire exagérée ou sarcastique. Troisièmement, il y a concurrence entre *yeah* et *ouais*. Cette concurrence se remarque notamment dans le corpus FAC, où on relève des occurrences isolables de *ouais* permettant la comparaison (voir le tableau 4.1 à la section 4.3). On observe que malgré des différences parfois marquées dans l'emploi des deux unités, *yeah* et *ouais* sont employées par tous les locuteurs du FAC (à une exception près) et avec tous les sens en usage et toutes les fonctions discursives identifiés à la section 5.3.4. *Ouais* est également souvent employée par les locuteurs dans le corpus Beaulieu et le corpus PFC, quoiqu'en position non isolable. Pris ensemble, ces facteurs indiquent qu'on peut raisonnablement attribuer un statut codique à *yeah*, à savoir qu'il s'associe au code anglais dans l'esprit des locuteurs, malgré sa fréquence dans le corpus d'analyse et son apparente diffusion dans les communautés étudiées.

5.3.2 Fréquence d'usage par locuteur

Le tableau 5.4 présente la fréquence d'usage de *yeah*, *yes* et *yep* par locuteur, en ordre décroissant de fréquence, ainsi que le pourcentage d'occurrences produites par le locuteur sur le total de 209 occurrences retenues pour l'analyse. Chaque locuteur est identifié d'abord par l'enregistrement auquel il participe et ensuite par son rôle : interviewé ou interviewée (INT, suivi d'un chiffre lorsqu'il y a plusieurs personnes interviewées), animateur ou animatrice (AN) ou chercheure (C). Les locuteurs qui n'ont produit aucune occurrence des unités en question ne sont pas représentés dans le tableau. Trois locuteurs ont participé à plusieurs entretiens, soit l'animateur de huit enregistrements sur neuf du corpus FAC (FAC : AN), l'animatrice des trois entretiens du PFC

où des occurrences de *yeah* ont été produites (PFC : AN) et la chercheuse du corpus Beaulieu (CB :C). En plus des données sur leurs productions des unités en question dans chaque entretien individuel, une ligne supplémentaire a été ajoutée pour chacun de ces trois locuteurs totalisant leurs productions sur l'ensemble du corpus d'analyse (indiquée par « tous »).

TABLEAU 5.4 : FRÉQUENCE D'OCCURRENCE DE *YEAH*, *YES* ET *YEP* PAR LOCUTEUR, OCCURRENCES ANALYSÉES SEULEMENT

Locuteur	# <i>yeah</i>	# <i>yes</i>	# <i>yep</i>	Total	% <i>yeah / yes / yep</i> (sur 209)
FAC : AN (tous)	56	2	0	58	27,8
FAC 08 : INT	25	8	1	34	16,3
FAC 01 : AN	31	1	0	32	15,3
FAC 06 : INT1	24	1	0	25	12,0
FAC 01 : INT	16	3	0	19	9,1
FAC 03 : INT1	12	1	0	13	6,2
CB-PRS 02AB : INT	13	0	0	13	6,2
FAC 07 : AN	12	0	0	12	5,7
CB-PRS 03 : INT1	12	0	0	12	5,7
FAC 08 : AN	5	1	0	6	2,9
CB-IND 04 : INT	6	0	0	6	2,9
FAC 06 : AN	5	0	0	5	2,4
FAC 09 : INT2	5	0	0	5	2,4
PFC : AN (tous)	5	0	0	5	2,4
FAC 04 : INT1	2	0	2	4	1,9
CB-PRS 03 : INT2	4	0	0	4	1,9
FAC 06 : INT2	2	0	1	3	1,4
PFC 04 : AN	3	0	0	3	1,4
FAC 03 : AN	2	0	0	2	1,0
FAC 07 : INT	1	0	0	1	0,5
FAC 05 : AN	1	0	0	1	0,5
CB-IND 02 : INT	1	0	0	1	0,5
CB-IND 04 : C	1	0	0	1	0,5
CB-PRS 04 :INT	1	0	0	1	0,5
CB-IND 07 : INT	1	0	0	1	0,5
CB : C (tous)	1	0	0	1	0,5
PFC 01 : AN	1	0	0	1	0,5
PFC 02 : AN	1	0	0	1	0,5
PFC 02 : INT1	1	0	0	1	0,5
PFC 04 : INT1	1	0	0	1	0,5

Ce tableau indique que le locuteur le plus productif d'unités de la catégorie *yeah / yes / yep* en termes de fréquence absolue est effectivement FAC : AN, ce qui s'explique par sa participation active à huit entretiens, soit 13 heures d'enregistrement. Par contre, en termes de fréquence relative, sa production d'unités cibles est comparable, et parfois même inférieure, à celle de ses interlocuteurs dans six des huit entretiens. Il en est de même pour PFC : AN, qui produit dans l'ensemble plus d'occurrences que tout autre locuteur du corpus PFC, mais seulement parce qu'elle participe à plusieurs entretiens. Il est intéressant de noter que CB : C produit nettement moins d'occurrences isolables de *yeah* (seulement une occurrence) que certains de ses interlocuteurs, ce qui s'explique sans doute par son statut de non-membre de la communauté de pratique en

question. De fait, comme ce sont justement les différences observées dans les normes et les habitudes de pratique qui ont motivé l'enquête (Beaulieu, 2012), on ne doit pas se surprendre que CB : C n'ait pas les mêmes habitudes de pratique que les participants interviewés. De plus, étant donné les attitudes des Franco-Albertains envers le français et l'anglais (*cf.* la section 3.2.1) l'enquêtrice a peut-être évité, consciemment ou inconsciemment, l'incorporation des mots anglais dans son discours afin de ne pas susciter de perceptions négatives (manque de respect envers ses participants, manque de sérieux envers la langue française, etc.) pouvant nuire à la collecte de données. Une autre explication potentielle pour la faible présence des unités cibles dans les productions de CB : C réside dans son statut de chercheuse dont l'enquête porte directement et ouvertement sur la langue française : cette locutrice aurait pu vouloir créer, consciemment ou non, un environnement clairement francophone afin d'encourager ses participants à s'exprimer en français sur la question du français.

Pour les autres locuteurs, il est également difficile de déterminer avec certitude les facteurs contribuant à leur production plus ou moins fréquente de *yeah*, *yes* ou *yep*. Par exemple, on ne peut pas dire si le sexe du locuteur a une influence, en raison du petit nombre de locuteurs étudiés et du fait qu'aucun sous-corpus n'est équilibré dans le genre. Il n'y a que 2 femmes productrices des unités cibles dans le corpus FAC, tandis que le corpus PFC ne compte que 2 hommes qui en produisent, et le corpus Beaulieu aucun. Il est difficile d'identifier l'âge comme étant un facteur contributif : la seule jeune adulte dans le corpus PFC est PFC : AN et bien qu'elle produise le plus grand nombre de *yeah* en fréquence absolue, elle n'en produit pas plus que ses interlocuteurs en termes de fréquence relative. Dans le corpus Beaulieu, tous les locuteurs qui produisent *yeah* sont d'âge moyen ou avancé, et aucun dans le corpus FAC n'est âgé de moins de 60 ans.

Il semble clair toutefois que l'exposition à l'anglais est un facteur important dans les habitudes d'usage de ces unités. La locutrice la plus productive de *yeah* dans le corpus Beaulieu indique que l'anglais est la langue dominante à la maison malgré le fait que ses enfants puissent aussi s'exprimer en français. Quant à la deuxième locutrice produisant le plus de *yeah* dans ce même corpus, elle a poursuivi des études supérieures et des certifications dans son domaine (soins infirmiers) principalement en anglais. L'attitude envers l'emploi de ressources de différents codes est aussi un facteur significatif. FAC :AN, FAC 01 : INT et FAC 08 : INT, tous les trois très productifs en ce qui concerne l'unité *yeah*, indiquent tous que non seulement le mélange des codes est normal pour eux, mais que cette manière de s'exprimer fait partie de leur identité franco-américaine. Parlant de l'alternance entre l'anglais et le français, FAC : AN dit :

AN : This is typical uh Saint John. < INT: is that right? [rire] > Saint John Valley. < INT : Yeah huh? > Uh we speak whichever (.) whichever way we want.

(FAC 08, 21:29)

FAC : 08 INT affirme qu'il pense dans les deux langues, au point où il a non seulement de la difficulté à répondre à la question « pensez-vous en anglais ou en français? », mais juge aussi qu'elle n'est pas pertinente. Quant à FAC 01 :INT, elle affirme qu'elle ne parle pas bien français parce que sa façon de le parler diverge du français standard, entre autres par le fait d'employer autant de mots en anglais. Cependant, elle valorise aussi sa façon de parler, en ce qu'elle affirme avec fierté : « je parle pas français pis-- ni anglais, je parle franglais! » (FAC 01, 15:59).

D'autres facteurs à considérer sont le style personnel et la nature de la conversation. Par exemple, on peut caractériser FAC 08 : INT par l'affirmation enthousiaste et l'écoute attentive en conversation, caractéristiques qui pourraient s'expliquer par sa profession de barbier. Étant donné son style communicatif, il n'est pas étonnant qu'il produise souvent *yeah*, et qu'il soit le locuteur le plus productif de *yes*. D'autres locuteurs, comme FAC 06 : INT et FAC 01 : INT, semblent avoir plus de difficulté à s'exprimer (ils font beaucoup de pauses et d'hésitations, cherchent souvent leurs mots, etc.). Comme la réparation et la facilitation sont des fonctions fréquentes de *yeah* (voir la section 5.3.4), cela pourrait contribuer à la haute fréquence de production de cette unité dans ces entretiens. Également, la conversation qui a lieu dans l'entretien FAC 08 comporte beaucoup de questions et des affirmations quant à la localisation des lacs, des villages, des routes, etc. par rapport à la géographie locale, et l'identification des résidents locaux et de leurs alliances familiales. Ces contextes sont donc propices à la production de *yeah* au cours de la conversation. Il en est de même dans l'entretien FAC 06.

Le tableau 5.5 présente un autre indice du style personnel des locuteurs du corpus FAC : leur propension à employer *ouais* et *oui*.⁶⁴ Le tableau présente pour chaque locuteur ayant produit au moins une occurrence des unités cibles en anglais ou en français : le nombre total d'unités cibles produites en position isolable dans chaque langue et le pourcentage que ce nombre représente du total de productions des unités cibles de chaque langue dans le même sous-corpus (soit 162 pour *yeah* / *yes* / *yep* et 55 pour *ouais* / *oui*).

Les résultats présentés dans le tableau 5.5 démontrent que les locuteurs qui produisent le plus souvent *yeah*, *yes*, ou *yep* en position isolable sont en général ceux qui produisent le plus souvent *ouais* et *oui* en position isolable. En revanche, les locuteurs qui produisent le moins les unités cibles en anglais produisent aussi moins souvent les éléments français de ce PLM. Ces résultats suggèrent que certains locuteurs ont un style personnel qui les amène à exprimer plus souvent ou plus facilement l'affirmation et l'alignement, et à participer plus activement à la collaboration interactionnelle, notamment en ce qui a trait à la réparation et à la facilitation (voir la section 5.3.4 pour une description détaillée de ces sens et ces fonctions associés à *yeah* / *yes* / *yep*). Il faut

⁶⁴ Aucune occurrence isolable définitive de *ouaip*, *ouin* ou d'autres formes semblables n'a été repérée dans le corpus d'analyse. Cependant, comme dans le cas de *yep*, la différence entre *ouaip* ou *ouin* et certaines prononciations de *ouais* est mince, et la qualité de l'enregistrement aurait pu la masquer par endroits. Le nombre réel d'occurrences de *ouaip* et de *ouin* pourrait donc être plus élevé.

se rappeler cependant que la fréquence de production des unités cibles peut être contrainte par la nature de la conversation et par la nature de la participation du locuteur en question. Une conversation qui présente plus ou moins d'occasions d'employer *yeah* – questions, négociations de tour de parole, dysfluidités à réparer, etc. – présentera sans doute aussi moins d'occasions d'employer *ouais*. De même, un locuteur qui n'intervient pas souvent ou qui ne contribue que des courtes interventions dans une conversation donnée produira nécessairement moins d'occurrences de toutes les unités cibles.

TABLEAU 5.5 : FRÉQUENCE D'OCCURRENCE DE *YEAH* / *YES* / *YEP* ET *OUAIS* / *OUI* PAR LOCUTEUR, CORPUS FAC SEULEMENT

Locuteur	# total <i>yeah</i> / <i>yes</i> / <i>yep</i>	% <i>yeah</i> / <i>yes</i> / <i>yep</i> (FAC seulement)	# total <i>ouais</i> / <i>oui</i>	% <i>ouais</i> / <i>oui</i> (FAC seulement)
FAC: AN (tous)	58	35,8	31	56,4
FAC 08: INT	34	21,0	13	23,6
FAC 01: AN	32	19,8	10	18,2
FAC 06: INT1	25	15,4	6	10,9
FAC 01: INT	19	11,7	11	20,0
FAC 03: INT1	13	8,0	1	1,8
FAC 07: AN	12	7,4	1	1,8
FAC 08: AN	6	3,7	1	1,8
FAC 06: AN	5	3,1	8	14,6
FAC 09: INT2	5	3,1	1	1,8
FAC 04: INT1	4	2,5	0	0,0
FAC 06: INT2	3	1,9	0	0,0
FAC 03: AN	2	1,2	1	1,8
FAC 07: INT	1	0,6	1	1,8
FAC 05: AN	1	0,6	0	0,0
FAC 09: an ⁶⁵	0	0,0	1	1,8

5.3.3 Tendances thématiques et cooccurrences

Aucune tendance notable n'a été observée quant à la thématique de discussion en cours lors de l'emploi de *yeah*, *yes* ou *yep*; ce résultat n'est guère étonnant étant donné la polyvalence de ces unités et le fait qu'elles sont d'usage courant dans une grande variété de contextes.

Pour des raisons similaires, aucune cooccurrence remarquable n'a été observée pour *yeah*, *yes* et *yep* de manière remarquable – à l'exception des éléments français du même PLM. De fait, dans l'ensemble du corpus d'analyse, 59 occurrences isolables des unités de la catégorie *yeah* / *yes* / *yep* ont lieu dans l'environnement rapproché d'une occurrence de *ouais* ou *oui* (isolable ou non), soit 28,2 % des 209 occurrences étudiées. Dans 21 de ces 59 cas, le même locuteur qui produit l'occurrence de *yeah*, *yes* ou *yep* produit aussi au moins une cooccurrence de *ouais* ou *oui*, par exemple :

⁶⁵ L'animatrice de FAC : 09 n'est pas la même que pour les 8 autres enregistrements du corpus FAC. Les lettres minuscules sont employées pour distinguer les deux locuteurs.

INT2 : Y avait un petit pont là au bout là XX Brook < INT1: **ouais! ouais.** > Y avait un petit pont
< INT1 : **yeah!** >

(FAC 06, 1:02:13)

Dans 42 des cas, l'interlocuteur de la personne ayant produit l'occurrence de *yeah*, *yes* ou *yep* produit au moins une occurrence de *ouais* ou *oui*, par exemple :

C : **ouais** ça fait une différence les infirmiers **oui**
INT : **yeah** le la: les personnes que tu rencontres

(CB-IND 04, 35:30)

Quatre des cas susmentionnés coïncident, en ce que le locuteur et au moins un des interlocuteurs produisent tous les deux *ouais* ou *oui* autour d'une même occurrence de *yeah*, *yes* ou *yep* :

INT : L'académie ben ça a viré en shopping center ça. < AN : **ouais** > S-- S***-- uh sa-- As*** je
pense ça porte le nom < AN : **yeah** > **ouais.**

(FAC 07, 12:27)

Il n'est pas surprenant que des éléments français et anglais de ce PLM sont souvent produits à proximité les uns des autres. Ces unités remplissent les mêmes fonctions discursives et sont employées avec les mêmes sens, quoiqu'à des proportions variables (un point qui sera discuté plus en détail à la section 5.3.4). Elles sont donc susceptibles d'être employées dans les mêmes contextes dans une interaction. De plus, comme ces unités sont relativement indépendantes sur le plan morphosyntaxique, elles peuvent facilement être employées à plusieurs endroits dans l'énoncé, dont plusieurs seront rapprochés, voire à proximité immédiate (répétition). De fait, il n'est pas rare que ces unités soient répétées dans le discours, généralement pour souligner l'affirmation ou l'accord. Ces répétitions peuvent également signaler qu'on a terminé sa contribution à la thématique en cours, et que le moment est propice pour changer le sujet. Cette répétition peut être produite par le même locuteur ou par l'interlocuteur :

AN : Garçon à *** ça?
INT : **Yes! yeah.** He lives here in Saint John.

(FAC 08, 1:04:25)

AN : J'ai eu une classe qui avait trois élèves. Ça c'est un minimum [rire] < INT1 : **ouais.** > **yeah.**
INT1 : **Yeah.** (.) Entécas, ça chan:ge ça change.

(PFC 04, 13:29)

Il semble que pour certains locuteurs du moins, cette répétition peut s'étendre à tous les éléments du PLM, qu'ils relèvent de l'anglais ou du français. Cependant, une analyse plus fine de ce phénomène, sur un plus grand échantillon, serait nécessaire afin de déterminer les circonstances et les raisons qui le favorisent.

5.3.4 Sens en usage et fonctions discursives

Comme il a déjà été mentionné au début de la section 5.3, *yeah / yes / yep* est la seule catégorie pour laquelle il y a une quantité d'occurrences isolables suffisante pour mener une analyse détaillée des sens en usage et des fonctions discursives de ces unités, et pour laquelle il est possible de comparer les résultats avec ceux d'une analyse identique des éléments français de son PLM (*ouais / oui*) qui sont aussi en position isolable. Pour ce faire, il a d'abord fallu établir les catégories d'analyse représentant les sens en usage et les fonctions discursives qui se sont révélés pertinents pour ces unités.

5.3.4.1 Établissement des catégories d'analyse

La grande majorité des études sur la signification de *yeah*, *yes* et *yep* s'inscrivent dans une approche d'analyse conversationnelle, ce qui fait en sorte que ces unités sont étudiées principalement en lien avec l'organisation de l'interaction : en tant que signaux d'écoute, en tant que signes de l'intention de prendre ou de céder la parole, et en tant qu'unités qui structurent le discours et influent sur l'interprétation des énoncés (voir par exemple Drummond et Hopper, 1993 ; Fuller, 2003 ; Gravano *et al.*, 2007 ; Lambertz, 2011). Autrement dit, ces unités sont étudiées en tant que marqueurs discursifs (ou marqueurs pragmatiques, *cf.* la section 5.5). Cela n'est pas sans raison – *yeah*, *yes* et *yep* fonctionnent effectivement comme des marqueurs pragmatiques – mais leur fréquence dans le corpus d'analyse⁶⁶ suggère que, pour les locuteurs étudiés, ce ne sont pas des marqueurs discursifs « comme les autres ».

Par contre, que ces unités soient étudiées presque exclusivement en analyse conversationnelle fait en sorte qu'il n'y a pas de cadre établi pour étudier la signification de *yeah*, *yes* et *yep* en dehors de quelques emplois spécifiques : en tant que signal d'écoute (*backchannel*), pour indiquer l'accord ou l'affirmation et pour relier deux segments de discours (énoncés, tours de parole, thématiques, etc.). Par exemple, Drummond et Hopper (1993) identifient deux emplois principaux de *yeah*, soit comme marqueur d'accord (« agreement token »), ou comme marqueur de confirmation (« acknowledgement token »). En tant que marqueur d'accord, *yeah* signifie une réponse positive aux propos de son interlocuteur. En tant que marqueur de confirmation, *yeah* indique qu'on est attentif au message de son interlocuteur. Dans cette dernière fonction, il peut être employé comme signal d'écoute (« je comprends ton message et je t'encourage à continuer ») ou comme un connecteur discursif (« je comprends ton message et je le prends comme point de départ pour mon tour de parole »). Selon

⁶⁶ Prises ensemble, *yeah*, *yes* et *yep* se trouvent en position isolable trois fois plus souvent que l'ensemble des autres marqueurs pragmatiques et 14 fois plus souvent que *so*, le deuxième marqueur pragmatique le plus fréquent.

Drummond et Hopper (1993), *yeah* peut aussi être employé de façon ambiguë, c'est-à-dire ayant les deux fonctions en même temps. Cela permettrait aux locuteurs de réinterpréter sa fonction *a posteriori* – par exemple, de se servir d'un marqueur d'accord pour ensuite prendre la parole – permettant de maintenir leurs options ouvertes pendant que l'interaction se déroule.

Gravano *et al.* (2007), travaillant sur la désambiguïsation automatique des « mots affirmatifs » comme *yeah, okay, alright*, etc., combinent les fonctions d'accord et de confirmation dans une seule catégorie (« acknowledgement/agreement : I agree with you, I believe what you say », p. 1614) mais séparent la fonction de signal d'écoute dans une catégorie à part (« backchannel : I hear you and please continue », p. 1614). Ils différencient également entre la fonction d'ouverture d'un nouveau sujet de discussion (« cue beginning ») et de fermeture du sujet de discussion en cours (« cue ending »). Gravano *et al.* (2007) voient dans la combinaison de la fonction d'accord/confirmation et soit l'ouverture, soit la fermeture du sujet de discussion une fonction de pivot (« pivot ending », « pivot beginning ») qui ressemble à l'emploi ambigu de Drummond et Hopper (1993). Gravano *et al.* (2007) ont également identifié des fonctions, très peu fréquentes avec *yeah*, de « vérification » (« check ») et « tentative de gagner du temps » (« stall »).⁶⁷

Lambertz (2011), qui s'intéresse en particulier à *yeah* et à *mm* comme signaux d'écoute, distingue entre les signaux d'encouragement (« continuers ») qui, comme les signaux d'écoute de Drummond et Hopper (1993) ou de Gravano *et al.* (2007), invitent l'interlocuteur à poursuivre son discours tout en indiquant qu'on ne cherche pas à prendre la parole; les signaux d'accord (« agreement markers ») qui indiquent ouvertement que le locuteur est en accord avec son interlocuteur; et les signaux de concordance (« alignment markers ») qui indiquent que le locuteur partage l'opinion ou le sentiment de son interlocuteur et qui peuvent ou non servir à prendre ensuite la parole, comme les marqueurs de confirmation de Drummond et Hopper (1993) et semblable au pivot d'ouverture de Gravano *et al.* (2007). Lambertz (2011) admet que la distinction entre accord et concordance est fine, et significativement, elle ne fournit pas de définition qui permettrait de clarifier cette distinction, à part le fait que « agreement seems to be more opinionated [than alignment] » (p. 16). Toutefois, une telle distinction pourrait être pertinente dans l'esprit des locuteurs : Lambertz (2011) trouve que le marqueur *mm* est employé pour signaler la concordance, mais jamais l'accord.

Les marqueurs pragmatiques, y compris *yeah*, font aussi l'objet d'études en l'acquisition des langues, en lien avec la considération grandissante que ce domaine donne à l'acquisition des compétences communicatives et pragmatiques. La plupart de ces études visent à suivre le développement des compétences pragmatiques ou à comparer l'emploi de marqueurs pragmatiques des locuteurs dans leur deuxième langue à

⁶⁷ Deux autres fonctions étudiées par Gravano *et al.* (2007), « modificateur littéral » et « retour d'une autre tâche » ne s'appliquent pas du tout à *yeah*.

celui des locuteurs natifs – il n'y a pas d'étude, à ma connaissance, qui compare l'emploi par des locuteurs bilingues des ressources dans les deux langues afin d'en comparer les différentes visées référentielles exprimées. Cependant, les études sur l'emploi des marqueurs pragmatiques par des locuteurs d'une langue seconde font état des usages de *yeah* qui pourraient être pertinents pour la présente étude. Wong (2000), par exemple, trouve que ses participants (des locuteurs du mandarin qui ont l'anglais comme langue seconde) emploient *yeah* comme un mécanisme d'auto-réparation lorsqu'ils produisent des dysfluidités ou énoncent une information erronée. De la même façon que les marqueurs *uh* ou *um* (Brennan et Schober, 2001), il semble que *yeah* permet aux participants de Wong (2000) de signaler qu'ils « annulent » ce qui a été produit avant et commencent à nouveau. *Yeah* permet aussi de gagner du temps lors d'une recherche lexicale, ce qui rappelle à la fonction « stall » identifiée par Gravano *et al.* (2007). Wong (2000) remarque de plus que cet emploi n'est pas observé chez des locuteurs natifs d'anglais. Elle suggère que les locuteurs d'anglais langue seconde, étant conscients qu'ils courent un plus grand risque que des locuteurs natifs d'être perçus comme incompetents lorsqu'ils produisent de telles « erreurs », utilisent *yeah* afin de signaler justement qu'ils sont conscients de leurs difficultés et qu'ils maîtrisent bien la situation.

Les conclusions de House (2013) vont aussi dans ce sens. Dans son étude de locuteurs dans un contexte universitaire où l'anglais est employé comme *lingua franca* entre étudiants et professeurs ayant des langues maternelles différentes, elle montre que *yeah* est employé principalement pour organiser le discours, et surtout d'une manière orientée sur soi-même, c'est-à-dire pour gagner du temps pour mieux comprendre l'énoncé de son interlocuteur, pour gagner du temps pour mieux planifier son propre énoncé ou pour signaler que le message de son interlocuteur a bien été reçu et compris (« uptaking »). House (2013) observe que ses participants utilisent *yeah* très fréquemment et dans une large gamme de fonctions. En accord avec les observations de Wong (2000), elle conclut que ces emplois sont motivés par la connotation plutôt positive de *yeah*, qui sert à modérer les multiples menaces à la face inhérentes à l'utilisation d'une langue seconde.

À la lumière de ces constats et des observations faites dans le corpus d'analyse, les catégories d'analyse ci-dessous ont été établies pour *yeah*, *yes* et *yep* (ainsi que pour *ouais* et *oui*, à des fins de comparaison). Il est à noter que l'interprétation des occurrences a été faite en consultant non seulement les transcriptions, mais les enregistrements, afin de tenir compte des éléments du contexte non transcrits comme le ton et la prosodie, le langage corporel, les mimiques, etc. qui jouent un rôle crucial dans l'interprétation (*cf.* Kerbrat-Orecchioni, 2007). Ainsi, il se peut que tous les indices qui ont mené à la classification d'une occurrence donnée ne soient pas immédiatement évidents dans les extraits fournis en exemple. Il est également à noter que les catégories d'analyse ci-dessous ne sont pas mutuellement exclusives; plusieurs catégories peuvent être attribuées à une seule et même occurrence. En particulier, les catégories d'affirmation et de concordance coïncident souvent.

- Affirmation : une occurrence fait partie de cette catégorie lorsqu'elle constitue une réponse positive (« oui, j'affirme que cela est vrai »). L'affirmation est divisée en deux sous-catégories selon que la réponse a été sollicitée par une question explicite (manifeste formellement, par la syntaxe et/ou la prosodie de l'énoncé) ou par une question implicite, comme dans l'exemple suivant :

INT1 : My mother could make thirteen different types of: patates fricassées!
 AN: [rire] Vous en avez mangé, des patates! < INT1: **yeah!** >

(FAC 03, 16:19)

- Concordance : une occurrence fait partie de cette catégorie lorsqu'elle sert à indiquer que le locuteur comprend le message ou le sentiment de son interlocuteur, qu'il sympathise ou qu'il est d'accord avec lui, sans nécessairement affirmer factuellement son énoncé ou répondre à une question. Dans l'exemple suivant, l'animateur demande à l'interviewé de confirmer une information qu'il n'avait pas précédemment, n'ayant pas séjourné au pensionnat. Ainsi, son *yeah* suivant cette confirmation n'est pas une affirmation mais bien un signe de sa compréhension du message de l'interviewé :

INT : pis le pain [*au pensionnat*] était tout le temps frais. Y faisait du pain eux autres là
 < AN : c'est vrai? > la journée ouais < AN : **yeah** >

(FAC 07, 48:57)

- Signal d'écoute : une occurrence fait partie de cette catégorie lorsqu'elle sert à indiquer que le locuteur est attentif à son interlocuteur et l'encourage à continuer de parler. Quoique cette catégorie, comme les autres, n'est pas exclusive, un signal d'écoute implique nécessairement que le locuteur n'emploie pas *yeah* comme point de départ pour son propre tour de parole.⁶⁸
- Emphase : les émotions constituent un facteur pouvant influencer sur la production de l'alternance codique. Les locuteurs bilingues jugent chacune des langues qu'ils parlent comme plus ou moins appropriée pour exprimer certains états affectifs (Dewaele, 2010). Dans la conscience collective, l'anglais est souvent reconnu comme étant « plus direct » et même « plus sévère » (voir par exemple les attitudes exprimées par les participants de Dewaele, 2010). Par ailleurs, de façon anecdotique, il a été observé qu'un nombre important des unités lexicales anglaises d'usage courant en français québécois sont employées de façon emphatique, pour exprimer un certain enthousiasme ou élan,

⁶⁸ Il est possible qu'un signal d'écoute puisse être suivi d'une prise de parole du même locuteur dans un cas bien précis : lorsque le locuteur émet un signal d'écoute auquel son interlocuteur ne réagit pas comme le locuteur l'entend, c'est-à-dire que l'interlocuteur choisit quand même de ne pas continuer son tour. Le signal d'écoute sera généralement suivi d'une pause en ce cas, après quoi le locuteur ou l'un de ses interlocuteurs pourraient tenter de relancer la conversation. Ce phénomène n'a toutefois pas été observé dans le corpus d'analyse.

parfois de façon exagérée ou sarcastique (*yes, go, nice, good job, etc.*, en plus des jurons comme *shit* ou *fuck*). La catégorie d'emphase ou d'emploi emphatique a été incluse dans l'analyse de *yeah / yes / yep* afin de voir si des observations similaires peuvent être faites au sein du corpus d'analyse. Un emploi de *yeah, yes* ou *yep* a été jugé emphatique si la prosodie était emphatique (présence de prolongements phonologiques, ton enthousiaste ou emphatique) ou s'il y avait répétition de la ou des unités cibles, ou une combinaison de ces indices :

AN : Vous mangez ben par exemple. < INT : **oh yes! oh yes.** > Uh that's what I hear, most people didn't know they were poor.

(FAC 03, 11:10)

- Réparation : suivant Wong (2000), les emplois de *yeah / yes / yep* comme marqueurs de réparation lors de la production d'un énoncé dysfluide ou factuellement incorrect ont été repérés. De plus, il a été observé dans le corpus d'analyse que *yeah*, en particulier, est souvent employé lorsque l'interlocuteur a de la difficulté à s'exprimer – notamment lorsqu'il cherche ses mots ou produit beaucoup de pauses, d'hésitations, ou de répétitions. Dans ces cas, il semble que *yeah* est employé pour indiquer à l'interlocuteur que l'énoncé a été compris malgré la difficulté, et qu'il n'a plus besoin de s'efforcer pour se corriger, mais peut « passer par-dessus » le segment qui lui pose problème et continuer avec son discours :

INT2 : Plus respectueux ça serait d'appeler Monsieur *** [nom de famille] pis après *** --[prénom] X. < C : ké. > la per-- la personne s-- dit < INT1 : **yeah** > monsi-- ***[prénom] c'est < INT1 : **yeah.** > < C : ké! > ké? Ça c'est pour partie X.

(CB-PRS 03, 14:10)

Ainsi, la catégorie de réparation est séparée en deux sous-catégories : l'auto-réparation (« self-repair ») et la facilitation (tentative d'atténuer une difficulté éprouvée par l'interlocuteur).

- Changement de sujet : suivant Drummond et Hopper (1993) et Gravano *et al.* (2007), cette catégorie d'analyse a été établie afin d'observer si les unités *yeah, yes* ou *yep* sont exploitées comme « point pivot » pour changer le sujet de la discussion, autrement dit de clore un sujet en cours, d'en ouvrir un nouveau, ou les deux. Ont été distingués les cas où le changement de sujet a été fait *post hoc*, c'est-à-dire lorsque les locuteurs ont exploité l'ambiguïté de *yeah, yes* ou *yep* en (ré)interprétant ces unités comme une clôture, notamment lorsqu'elles sont suivies d'une pause ou d'un changement de tour de parole :

INT : I did that [*faire une réservation d'hôtel*] in advance parce que c'est Memorial Day Weekend là. < AN: oh yeah yeah yeah yeah yeah yeah. > J'avais pensé là moi là là.
Yeah.

AN: **So uh when you work** uh: *** your eight-to-five here it's quite busy huh?

(FAC 08, 1:43:25)

On peut remarquer que certaines des catégories d'analyse énumérées correspondent davantage à des sens⁶⁹ des unités *yeah*, *yes* et *yep* (notamment l'affirmation), tandis que d'autres sont plutôt des fonctions de l'unité (signal d'écoute, réparation, changement de sujet), et plusieurs recouvrent à la fois un sens et une fonction (notamment la concordance et l'emphase). La combinaison de sens et de fonctions n'est pas par inadvertance : le sens en usage d'une unité donnée est établi par l'observation du contexte, et ce que l'unité fait dans le discours – comment son emploi est interprété par les locuteurs en présence et l'effet qu'il a sur l'interaction – fait partie de ce contexte (*cf.* la section 2.4.7). Envisagés de cette façon, sens et fonction sont en quelque sorte indissociables dans l'usage, et possiblement dans l'esprit des locuteurs – d'autant plus dans le cas des unités servant de marqueurs pragmatiques (*cf.* la section 5.5).

Il convient de réitérer ici ce qui a été dit à la section 2.4.7 : prendre part aux débats sur la nature de la signification des unités du lexique n'entre pas dans les objectifs de la présente étude. La mise à l'écart de ces questions n'implique nullement qu'elles ne sont pas importantes – essentielles, même – en matière de sémantique. Cependant, il n'est pas possible de confirmer ni d'infirmer aucune théorie sémantique avec les données disponibles dans la présente étude. Mieux vaut, alors, se concentrer sur ce qu'on peut observer des emplois que font les locuteurs de leurs ressources et éviter de poser des postulats ou de tirer des conclusions que les résultats ne sont pas en mesure d'appuyer. C'est toutefois dans le but d'éviter la confusion quant aux objets et aux méthodes de la présente étude que le terme *visée référentielle* a été préféré plutôt que *sens*, *signification* ou l'adjectif *sémantique* pour exprimer la totalité de ce qu'une unité peut communiquer.

La méthode d'analyse par catégorie est semblable à celle utilisée dans Vogh et Courbon (à paraître) : lorsqu'une occurrence pouvait être classée dans la catégorie en question selon les définitions établies ci-dessus, cette catégorie a été ajoutée aux annotations pour cette occurrence. Comme les catégories ne sont pas mutuellement exclusives, une occurrence a pu être classée dans plusieurs catégories si les conditions pour chacune de ces catégories étaient remplies. Le cas échéant, l'occurrence pouvait également être assignée à une sous-catégorie de la catégorie en question. Les sous-catégories de la catégorie d'affirmation sont

⁶⁹ Je ne fais pas de distinction ici entre les sens dénotatifs (comme l'affirmation) et les sens plutôt connotatifs (comme l'emphase). Selon les objectifs et l'approche de l'étude, il peut y avoir lieu de faire cette distinction; en l'occurrence, l'objectif est l'analyse de la visée référentielle, c'est-à-dire ce que les locuteurs entendent, en fonction de ce que les interlocuteurs vont comprendre de cette intention, lors de l'emploi de l'unité en question dans ce contexte précis. Dans cette perspective, la connotation fait partie de « ce que les (inter)locuteurs savent » sur l'unité et donc de ce qu'elle communique.

mutuellement exclusives (une question qui reçoit une réponse positive est soit directe ou indirecte, mais pas les deux en même temps), mais pas les sous-catégories de la catégorie de réparation. Le même processus a été répété pour les occurrences isolables de *ouais* et *oui*, dans le but de voir si les catégories identifiées s'appliquent de la même façon en français qu'en anglais.⁷⁰ La fréquence d'occurrence de chaque catégorie et sous-catégorie a également été calculée; les résultats sont présentés dans la section suivante.

5.3.4.2 Résultats et discussion

Le tableau 5.6 compare la fréquence absolue (#) et la fréquence relative (%) des occurrences de *yeah* et de *yes* pour chaque catégorie et sous-catégorie d'analyse pour l'ensemble des trois sous-corpus. La fréquence relative est calculée en fonction du nombre total d'occurrences de *yeah* (n = 190) et de *yes* (n = 15) dans le corpus d'analyse. Autrement dit, les colonnes marquées avec « % » rapportent la proportion des 190 occurrences de *yeah* ou des 15 occurrences de *yes* qui font partie de la catégorie en question. L'unité *yep* a été écartée de l'analyse en raison de sa très basse fréquence dans le corpus (n = 4).

Avant de discuter des résultats présentés dans le tableau 5.6, une note explicative est nécessaire quant aux sous-catégories présentant des chiffres identiques. Dans le cas de « question explicite » et « question implicite », comme ces sous-catégories sont mutuellement exclusives, l'identité en nombre des occurrences de *yes* est fortuite et il ne s'agit pas des mêmes occurrences. Dans le cas d'« auto-réparation » et de « facilitation », il s'agit d'un nombre égal d'occurrences de *yeah* et de *yes*, dont seulement 3 occurrences de *yeah* appartiennent aux deux sous-catégories en même temps.⁷¹ Dans le cas de « changement de sujet *post hoc* », il s'agit des mêmes occurrences précises de *yes* que dans la catégorie générale. Autrement dit, pour toutes les occurrences de *yes* qui ont donné lieu à un changement de sujet, ce changement a été fait de manière *post hoc*, après une pause ou un changement de tour de parole.

De façon générale, on remarque que l'unité *yes* est employée avec tous les sens et pour toutes les fonctions identifiés pour cette catégorie d'occurrences, à l'exception de « signal d'écoute ». De plus, comme il a déjà été mentionné, aucune occurrence de *yes* n'a mené à un changement de sujet « en continu », c'est-à-dire dans lequel le locuteur produit *yes* pour reconnaître le message de son interlocuteur et le prendre comme point de départ pour son propre tour de parole. *Yes* est donc employée comme – ou au moins interprétée comme – la clôture du sujet en cours, mais pas comme une ouverture sur le prochain sujet ou une invitation à continuer, ce qui indique que *yes* a un sens plus fermé, plus définitif, que *yeah*.

⁷⁰ Il est à rappeler ici que l'objet principal de l'analyse est les occurrences isolables en anglais et que c'est sous cet angle que les occurrences isolables en français sont examinées. Ce choix méthodologique n'implique pas pour autant que les unités lexicales du français identifiées n'auraient pas des catégories d'analyse qui leur sont propres.

⁷¹ Il s'agit des contextes où autant le locuteur que l'interlocuteur éprouvent de la difficulté à s'exprimer.

TABLEAU 5.6 : FRÉQUENCES ABSOLUE ET RELATIVE DES OCCURRENCES DE YEAH ET DE YES PAR CATÉGORIE D'ANALYSE

	affirmation		concordance		signal d'écoute		emphase		réparation		changement de sujet	
	#	%	#	%	#	%	#	%	#	%	#	%
yeah (n = 190)	156	82,1	120	63,2	27	14,2	49	25,8	48	25,3	46	24,2
yes (n = 15)	12	80,0	4	26,7	0	0,0	12	80,0	1	6,7	2	13,3
% différence (yeah – yes)		+ 2,1		+ 36,5		+ 14,2		(-54,2)		+ 18,6		+ 10,9
	question explicite		question implicite				auto-réparation		facilitation		post hoc	
	#	%	#	%			#	%	#	%	#	%
yeah (n = 190)	41	21,6	115	60,5			35	18,4	35	18,4	25	13,2
yes (n = 15)	6	40,0	6	40,0			1	6,7	1	6,7	2	13,3
différence (% yeah – % yes)		(-18,4)		+ 20,5				+ 11,7		+ 11,7		(-0,1)

89

TABLEAU 5.7 : FRÉQUENCES ABSOLUE ET RELATIVE DES OCCURRENCES DE YEAH ET DE YES PAR CATÉGORIE D'ANALYSE, CORPUS FAC SEULEMENT

	affirmation		concordance		signal d'écoute		emphase		réparation		changement de sujet	
	#	%	#	%	#	%	#	%	#	%	#	%
yeah (n = 143)	113	79,0	88	61,5	24	16,8	44	30,8	25	17,5	39	27,3
yes (n = 15)	12	80,0	4	26,7	0	0,0	12	80,0	1	6,7	2	13,3
% différence (yeah – yes)		(- 1,0)		+ 34,8		+ 16,8		(-49,2)		+ 10,8		+ 14,0
	question explicite		question implicite				auto-réparation		facilitation		post hoc	
	#	%	#	%			#	%	#	%	#	%
yeah (n = 143)	34	23,8	79	55,2			7	4,9	18	12,6	21	14,7
yes (n = 15)	6	40,0	6	40,0			0	0	1	6,7	2	13,3
différence (% yeah – % yes)		(-16,2)		+ 15,2				+ 4,9		+ 5,9		+ 1,4

On remarque également que *yes* est employée moins souvent (en fréquence relative) que *yeah* pour la plupart des catégories d'analyse, et beaucoup moins souvent pour exprimer la concordance ou pour contribuer à la réparation. Or, *yes* est employée beaucoup plus souvent que *yeah* pour indiquer l'emphase; 5 occurrences de *yes*, soit 33.3 % de toutes les occurrences de cette unité, expriment l'affirmation et l'emphase en même temps. Finalement, quoique *yeah* et *yes* soient employées en position isolable presque aussi souvent pour répondre positivement à une question, *yes* est employée beaucoup plus souvent que *yeah* en réponse à une question explicite, tandis que *yeah* est employé plus souvent pour des questions implicites. Ce résultat concorde avec l'observation de Drummond et Hopper (1993) que *yes* est plus rarement employée (et *yeah* plus fréquemment employée) comme marqueur d'accord dans des situations informelles. Ici, quoique la situation de communication soit toujours informelle, le fait de poser une question explicite a l'effet de formaliser l'échange question-réponse, et donc de solliciter une réponse qui est moins équivoque. Ce résultat va également dans le sens de House (2013), qui démontre que ses participants préfèrent *yes* à *yeah* pour exprimer un véritable accord, et emploient *yeah* dans une plus grande variété de situations. Dans l'ensemble, la comparaison des emplois de *yeah* et de *yes* suggère qu'en général, *yes* est préféré par les locuteurs étudiés pour traduire des visées référentielles semblables à *yeah*, mais moins équivoques.

Puisque les éléments français du PLM n'apparaissent en position isolable que dans un des sous-corpus, le corpus FAC, le tableau 5.7 présente seules les occurrences de *yeah* et de *yes* de ce même sous-corpus. Cette restriction permet une meilleure comparaison avec le tableau 5.8 ci-dessous et permet d'établir les tendances significatives, pour ce qui est des fréquences absolue et relative de chaque catégorie et sous-catégorie, de ce sous-corpus considéré en isolation (et donc d'identifier, s'il y a lieu, des différences par rapport aux tendances dans l'ensemble du corpus d'analyse). Dans le tableau 5.7, le nombre d'occurrences de *yeah* et de *yes* pour chaque catégorie est comparé au nombre total d'occurrences de cette même unité en position isolable dans le corpus FAC seulement ($n = 143$ pour *yeah*; $n = 15$ pour *yes*); *yep* est exclue en raison de sa fréquence extrêmement basse ($n = 4$). Il faut rappeler que la totalité des emplois de *yes* dans le corpus d'analyse provient du corpus FAC; la fréquence de *yes* pour chaque catégorie et sous-catégorie demeure donc inchangée. Toutefois, le tableau 5.7 permet déjà de constater, avant même de comparer *yeah* / *yes* / *yep* et *ouais* / *oui*, que les tendances générales observées quant à l'emploi de *yeah* et de *yes* dans l'ensemble du corpus d'analyse sont aussi présentes dans le corpus FAC en isolation.

Le tableau 5.8 présente, pour fins de comparaison, les fréquences absolue et relative d'emploi de *ouais* ($n = 33$) et de *oui* ($n = 22$) pour les mêmes catégories. Ces unités sont produites en position isolable dans le corpus FAC seulement. On peut voir que les différences observées entre *ouais* et *oui* sont semblables à celles entre *yeah* et *yes*, et que ces deux couples d'unités sont employés avec les mêmes sens et dans les mêmes fonctions discursives. Comme avec *yeah* et *yes*, *ouais* et *oui* sont employées à une fréquence (relative)

TABLEAU 5.8 : FRÉQUENCES ABSOLUE ET RELATIVE DES OCCURRENCES DE *Ouais* ET DE *Oui* PAR CATÉGORIE D'ANALYSE, CORPUS FAC SEULEMENT

	affirmation		concordance		signal d'écoute		emphase		réparation		changement de sujet	
	#	%	#	%	#	%	#	%	#	%	#	%
<i>ouais</i> (n = 33)	29	87,9	27	81,9	4	12,1	19	57,6	14	42,4	6	18,2
<i>oui</i> (n = 22)	20	90,9	10	45,5	1	4,5	11	50,0	4	12,1	6	27,3
différence (% <i>ouais</i> – % <i>oui</i>)		(-3,0)		+ 36,4		+ 7,6		+ 7,6		+30,3		(-9,1)
	question explicite		question implicite		auto-réparation		facilitation		post hoc			
	#	%	#	%	#	%	#	%	#	%		
<i>ouais</i> (n = 33)	7	21,2	22	66,7	1	3,0	13	39,4	2	6,1		
<i>oui</i> (n = 22)	9	40,9	11	50,0	0	0,0	4	18,2	2	9,1		
différence (% <i>ouais</i> – % <i>oui</i>)		(-19,7)		+ 16,7		+ 3,0		+ 21,2		(-3,0)		

TABLEAU 5.9 : FRÉQUENCES ABSOLUE ET RELATIVE DES OCCURRENCES DE LA CATÉGORIE *Yeah / Yes / Yep* ET *Ouais / Oui*, CORPUS FAC SEULEMENT

	affirmation		concordance		signal d'écoute		emphase		réparation		changement de sujet	
	#	%	#	%	#	%	#	%	#	%	#	%
<i>yeah / yes / yep</i> (n = 143)	129	79,6	93	57,4	24	14,8	57	35,2	27	16,7	42	25,9
<i>ouais / oui</i> (n = 55)	49	89,1	37	67,3	5	9,1	30	54,5	18	32,7	12	21,8
différence (% EN – % FR)		(-9,5)		(-9,9)		+ 5,7		(-19,3)		(-16,0)		+ 4,1
	question explicite		question implicite		auto-réparation		facilitation		post hoc			
	#	%	#	%	#	%	#	%	#	%		
<i>yeah / yes / yep</i> (n = 143)	42	25,9	87	53,7	7	4,3	20	12,3	24	14,8		
<i>ouais / oui</i> (n = 55)	16	29,1	33	60,0	1	1,8	17	30,9	4	7,3		
différence (% EN – % FR)		(-3,2)		(-6,3)		+ 2,5		(-18,6)		+ 7,5		

semblable dans un sens d'affirmation, alors que *oui* est employée plus souvent pour répondre à des questions explicites, et *ouais* est employée beaucoup plus fréquemment que *oui* pour exprimer la concordance et la réparation. Cependant, à la différence de *yeah* et *yes* (1) *ouais* et *oui* servent en proportion presque identique pour marquer l'emphase (2) *oui* apparaît une fois dans la fonction de signal d'écoute et (3) *oui* est plus fréquent que *ouais* pour le changement de sujet – y compris pour un changement de sujet « en continu ». Ainsi les locuteurs du FAC semblent ne pas avoir la même préférence avec *oui* qu'avec *yes* pour des emplois moins équivoques; autrement dit, les locuteurs semblent investir *oui* avec un sens plus ambivalent que celui de *yes*.

Le tableau 5.9 compare les occurrences de la catégorie *yeah / yes / yep* produites dans le corpus FAC (n = 162) aux occurrences de *ouais / oui* dans ce même sous-corpus (n = 55) quant aux fréquences absolue et relative des sens et des fonctions identifiés. Les colonnes marquées par « % » indiquent donc la proportion des 162 occurrences de *yeah, yes* ou *yep* et la proportion des 55 occurrences de *ouais* ou *oui* qui sont associées à chaque catégorie d'analyse. On voit dans ce tableau que *ouais / oui* dépasse *yeah / yes / yep* en fréquence relative dans la catégorie d'affirmation en général, et légèrement plus pour l'affirmation d'une question implicite. Les unités *ouais* et *oui* sont employées plus souvent pour exprimer la concordance, et beaucoup plus souvent pour exprimer l'emphase et pour accomplir des réparations, en particulier la facilitation. En revanche, *yeah, yes* et *yep* sont employées plus souvent comme signal d'écoute, pour faire de l'auto-réparation et pour le changement de sujet, notamment *post hoc*.

Ainsi, il n'y a pas vraiment de différence dans les préférences des locuteurs quant à l'expression du sens « canonique » que *ouais / oui* et *yeah / yes / yep* ont en commun, c'est-à-dire la réponse positive à une question explicite. Cependant, si on tient compte de la plus grande diversité de sens et de fonctions de ces unités, des tendances intéressantes apparaissent. *Ouais* et *oui* sont employées relativement plus souvent que *yeah, yes* et *yep* pour des fonctions relevant d'une plus forte (inter)subjectivité. Répondre à une question implicite, par exemple, nécessite de comprendre les normes de l'interaction et les habitudes et connaissances préalables de son interlocuteur afin de savoir interpréter son énoncé comme une question lorsque cette dernière n'est pas formellement indiquée comme telle par la prosodie ou la syntaxe, etc. La concordance implique de comprendre et de concorder avec le message ou le sentiment (ou les deux) de son interlocuteur, et la facilitation implique une certaine sympathie avec les difficultés de son interlocuteur, un désir de l'aider à les surmonter. Enfin, l'emploi emphatique nécessite une implication affective (ou au moins l'apparence d'une telle implication) dans l'interaction en cours. Pour leur part, *yeah, yes* et *yep* en position isolable participent relativement plus souvent à des fonctions de structuration du discours : en tant que signal d'écoute, qui sert entre autres à gérer les tours de parole en indiquant que le locuteur ne cherche pas à prendre la parole mais veut bien que son interlocuteur continue à parler, et pour clore un sujet de discussion (changement de sujet *post hoc*) ou en ouvrir de nouveaux (changement de sujet « en continu »).

Sans perdre de vue qu'autant les éléments anglais que français de ce PLM sont employés avec tous les sens et toutes les fonctions identifiées, il semble que globalement, les unités *ouais* et *oui* sont préférées par les locuteurs étudiés du corpus FAC pour traduire des visées référentielles qui touchent aux émotions, aux sentiments et à la relation avec l'interlocuteur – ce qui laisse en revanche des tâches plus « utilitaires » à *yeah*, *yes* et *yep*. Cette répartition pourrait être le résultat d'attitudes linguistiques, le stéréotype populaire voulant que l'anglais soit une langue plus directe (cf. Dewaele, 2010) et le français, plus romantique ou élégant. Or, il n'y a pas d'informations venant appuyer ou infirmer ces attitudes dans le corpus d'analyse, et il n'y a pas non plus, à ma connaissance, d'étude sur les perceptions des deux langues dans la communauté francophone du Maine. Il est donc difficile d'affirmer qu'il s'agit là d'une explication plausible de la répartition des préférences d'emploi des éléments de ce PLM. Cependant, on pourrait penser au contexte socio-historique des Franco-Américains du Maine (cf. la section 3.2.2.2) : historiquement, les francophones du Maine ont été victimes de violence et de discrimination sociale et économique, en partie en raison de leur langue, et les locuteurs étudiés du corpus FAC en ressentent encore les effets. Plusieurs d'entre eux ont parlé de leur expérience de l'élimination forcée du français dans les écoles, où il a été interdit non seulement comme langue d'instruction, mais comme langue parlée dans les corridors et les cours de récréation; des punitions étaient parfois administrées pour faire respecter ces règles. D'autres ont parlé de la discrimination et des moqueries auxquelles ils ont fait face en tant que francophones relativement peu scolarisés lors de leur service dans les forces armées. Beaucoup ont également parlé des difficultés sociales, psychologiques et économiques qu'ils ont vécues en raison de leur accent et de leur non-maîtrise de l'anglais. Finalement, plusieurs ont parlé du refus de la part de leurs enfants et petits-enfants (ainsi que parfois de leurs parents et grands-parents) de parler ou d'apprendre le français. Autrement dit, la langue française les réunit dans une identité franco-américaine commune, tout en les démarquant et les isolant du monde social anglo-américain plus large, et ce, même à l'intérieur de leur propre famille. Les seules autres personnes de leur entourage avec lesquelles ils partagent leur identité culturelle, leurs expériences, connaissances et pratiques, sont d'autres Franco-Américains francophones⁷². L'accès à des ressources linguistiques du français – et plus précisément la possibilité d'employer ces ressources en présupposant que ces emplois seront non seulement compris, mais *acceptés* – fait que pour ces locuteurs, ces ressources représentent, consciemment ou non, l'appartenance, la sécurité et la liberté d'expression d'être « entre nous ». Il n'est donc pas étonnant que ces locuteurs emploient des ressources linguistiques du français plutôt que de l'anglais pour traduire des visées référentielles qui relèvent de la connexion personnelle et de la cohésion sociale, et, du fait même, les renforcent.

⁷² Comme il a déjà été mentionné à la section 3.2.2.2, le fait d'avoir des compétences en français n'est pas une condition nécessaire pour s'identifier comme membre de la communauté franco-américaine.

5.4 *Right / that's right / is that right?*

Cette section présente les analyses qui ont été effectuées sur la catégorie *right / that's right / is that right?*. Elle débute comme la section précédente sur *yeah / yes / yep*, avec un aperçu de sa fréquence et sa distribution dans le corpus d'analyse (section 5.4.1); toutefois, étant donné que les occurrences de cette catégorie n'ont été produites que dans le corpus FAC, cet aperçu commence directement avec une analyse par locuteur. La section 5.4.2 traite les questions de cooccurrences et de tendances dans la thématique de discussion lors de la production de ces unités, avant de passer à l'analyse par sens en usage et par fonctions discursives dans la section 5.4.3.

5.4.1 Aperçu

L'unité *right* fait l'objet d'alternance codique en position isolable dans plusieurs formules ou constructions différentes : deux affirmatives (*right, that's right*) et une interrogative (*is that right?*). Quoique ces formules diffèrent considérablement dans leurs sens en usage et dans leurs fonctions, comme on le verra plus loin, elles ont en commun le fait de porter sur l'état épistémique du locuteur, de l'interlocuteur, ou des deux. Autrement dit, l'unité *right* fait référence dans chaque cas à la véracité, dans l'esprit du locuteur, d'une information que l'interlocuteur présente avec plus ou moins de certitude. Selon le cas, le locuteur est au courant de l'information en question et confirme qu'elle est exacte ou bien il n'est pas au courant (ou il fait comme s'il ne l'était pas) et il cherche à confirmer auprès de son interlocuteur l'exactitude de l'information telle qu'il la comprend :

AN : pis tes parents en ont eu seize, tu dis?

INT2 : **Right.** (.) I only have uh: two brothers left and two sisters.

(FAC 03, 1:01:42)

AN : pis c'était-tu pour la même affaire? [*l'hospitalisation de INT*] < INT : oui oui. Ouais. > **Is that right?**

(FAC 07, 6:25)

Le tableau 5.10 présente un survol des occurrences isolables de cette catégorie par unité et par locuteur; seuls les enregistrements où sont présentes les unités cibles sont détaillés. Une ligne supplémentaire a été ajoutée pour totaliser les productions de FAC : AN, qui participe à tous les enregistrements où les unités cibles ont été produites.

On voit que la relative prévalence de la catégorie dans le corpus d'analyse est due en grande partie aux productions de FAC : AN, et notamment à son emploi fréquent de *is that right?* qui compte pour 50 % des

occurrences dans cette catégorie. La fréquence de production élevée de FAC : AN pourrait être en partie le résultat d'habitudes d'usage qui lui sont particulières, mais elle relève nécessairement de sa participation à 12 heures d'enregistrement en tout; on verra plus loin que son rôle en tant qu'animateur de la conversation pourrait aussi être un facteur explicatif.

TABLEAU 5.10 : FRÉQUENCE DES OCCURRENCES ISOLABLES DE *RIGHT*, *THAT'S RIGHT* ET *IS THAT RIGHT?*
PAR LOCUTEUR

Entretien / Locuteur	<i>right</i>	<i>that's right</i>	<i>is that right?</i>	Total
FAC 01 : AN	2	3	10	15
FAC 07 : AN	5	2	7	14
FAC 07 : INT	1	5	1	7
FAC 03 : AN	0	0	3	3
FAC 06 : AN	0	0	2	2
FAC 06 : INT1	1	1	0	2
FAC 03 : INT2	1	0	0	1
FAC 08 : AN	0	0	1	1
FAC 08 : INT	1	0	0	1
Total	11	11	24	46
Total FAC : AN	7	5	23	35

Il est intéressant de noter que l'enregistrement FAC 08, le plus productif dans l'ensemble pour *yeah / yes / yep*, contient relativement peu de productions de *right / that's right / is that right?* en position isolable, tandis que l'enregistrement FAC 07, le plus productif dans l'ensemble pour *right / that's right / is that right?*, contient relativement peu de productions de *yeah / yes / yep*. Ces résultats suggèrent que *right / that's right / is that right?* ne remplit pas les mêmes fonctions ou ne représente pas le même style communicatif que *yeah / yes / yep*, malgré le fait que le sens « canonique » des deux catégories relèvent de l'affirmation.

5.4.2 Tendances thématiques et cooccurrences

Il a été observé dans le corpus d'analyse que les occurrences isolables de *right / that's right / is that right?* sont souvent suivies (et plus rarement précédées) d'une unité de réponse positive (*yeah, oui, ok, right*, etc.), mais jamais d'une unité de réponse négative. Une tendance a également été observé pour ces occurrences à être suivies par des marqueurs pragmatiques divers (*pis, and, mais, ben*, etc.). Le tableau 5.11 présente ces tendances de cooccurrence.

TABLEAU 5.11 : LES COOCCURRENCES FRÉQUENTES DE *RIGHT / THAT'S RIGHT / IS THAT RIGHT?*

	réponse affirmative			marqueur pragmatique							Total
	Précédant	Suivant	Total	<i>pis</i>	<i>ben</i>	<i>and / et</i>	<i>mais</i>	<i>parce que</i>	<i>alors</i>	<i>ça fait que</i>	
<i>right</i> (n = 11)	0	2	2	1	2	1	0	0	1	1	6
<i>that's right</i> (n = 11)	3	5	8	1	0	2	1	0	0	0	3 ⁷³
<i>is that right?</i> (n = 24)	2	15	17	5	1	1	1	1	0	0	9
Total (n = 46)	5	22	27	7	3	4	2	1	1	1	18

⁷³ L'une des occurrences est suivie du marqueur composite *et pis*, qui a donc été comptée à la fois sous *et* et *pis*.

Il est intéressant, mais pas tout à fait étonnant, que *right* et *that's right* aient des unités de réponse affirmative comme cooccurrence fréquente, étant donné le sens de confirmation qui leur est associé (cf. la section 5.4.3). Ces unités affirmatives servent donc à souligner ce sens, comme dans l'exemple suivant :

AN : Les jeunes qui s'en allent de la Vallée astheure < INT : ouais > c'est alentours de Portland
INT : **Ouais—oui! T'es ben avec ça ouais. That's right.**

(FAC 07, 47:39)

Similairement, il n'est pas surprenant que *is that right?* soit souvent suivi d'une réponse affirmative, étant donné qu'il correspond à une question formellement explicite (cf. la section 5.3.4.1) visant à éliciter une confirmation (quoique, comme 9 occurrences sur 24 ne suscitent pas une telle réponse, il semble qu'une réponse explicite ne soit pas toujours ressentie comme obligatoire par les locuteurs; ce point sera discuté à la section 5.4.3) En fait, la prévalence des réponses affirmatives comme cooccurrence avec la catégorie *right / that's right / is that right?* s'explique principalement par les réponses affirmatives à la question *is that right?*.

La prévalence des marqueurs pragmatiques en cooccurrence avec *right*, *that's right* et *is that right?* se comprend dès lors qu'est pris en compte le rôle des unités cibles en tant que signal d'écoute (cf. la section 5.4.3). L'une des occasions propices à la production d'un signal d'écoute dans une interaction est un « point d'articulation » dans le discours, comme la fin d'un événement dans une narration ou la fin d'un commentaire sur un sujet donné. Lorsque le locuteur reprend la parole après un signal d'écoute, il est naturel qu'il commence par un marqueur pragmatique qui lui permet d'enchaîner, d'une façon ou d'une autre, avec son discours précédent.

Quant aux thématiques, aucune thématique précise ne semble occasionner la production des unités cibles plus qu'une autre; cependant, les occurrences de *right / that's right / is that right?* semblent se concentrer autour des thématiques qu'on peut qualifier d'exceptionnelles (qui sont inhabituelles ou de quelque façon remarquables). Cette association aux situations exceptionnelles sera discutée en plus de détail à la section 5.4.3.

5.4.3 Sens en usage et fonctions discursives

Les sens et fonctions potentiels de *right*, *that's right* et *is that right?* sont très peu traités dans la littérature en général. Holmes et Chiles (2010) démontrent que *is that right?* indique que le locuteur demande une confirmation de la part de son interlocuteur plutôt qu'une information. Leurs résultats montrent que l'unité *is that right?* est employée pour contrôler le discours dans des réunions de travail, car elle peut notamment servir à résumer les propos d'un autre selon ce qu'on veut en faire ressortir, et à exiger de lui qu'il accepte formellement

cette version des choses avant de continuer la discussion (ex. « What you're saying is [...], is that right? »). Toutefois, parce que cette question n'apparaît que six fois dans leur corpus en raison de la nature de l'enquête,⁷⁴ une analyse plus poussée n'est pas possible. Banuazizi et Creswell (1999) considèrent *is that right?* comme un signal d'écoute. Ces auteures incluent également *right?*, mais cette forme interrogative n'apparaît pas en position isolable dans le corpus d'analyse. *Right* en forme déclarative fait partie des « mots affirmatifs » de Gravano *et al.* (2007), qui montrent qu'il est employé surtout comme « modificateur littéral » (« literal modifier »), c'est-à-dire pour confirmer l'exactitude d'une information, et parfois avec la fonction d'affirmation (« agreement/acknowledgement »). Cependant, il est difficile de voir la différence, dans le cas de *right*, entre un sens « littéral » et un sens affirmatif, et Gravano *et al.* (2007) ne donnent pas d'exemples ou d'explications spécifiquement pour *right*. Finalement, dans une étude sur une expression similaire en finnois, Koivisto (2013) soulève que *oh that's right* en anglais (une forme absente du corpus d'analyse) permet au locuteur de sauver la face en signalant que le locuteur vient de se souvenir d'une information qu'il aurait normalement déjà dû savoir. Selon Koivisto (2013) cette fonction dépend en particulier de la présence de *oh*, qui indiquerait un changement d'état, en l'occurrence, le rappel d'une information oubliée.

Les catégories d'analyse suivantes pour *right*, *that's right* et *is that right?* ont été établies selon cette littérature et selon des observations faites pendant le dépouillement du corpus d'analyse :

- La confirmation d'une information connue : pour *right* et *that's right*, la confirmation est le sens « canonique »; le locuteur confirme l'exactitude de l'énoncé de son interlocuteur, comme dans l'exemple suivant :

AN : pis tes parents en ont eu seize, tu dis?

INT2 : **Right.** (.) I only have uh: two brothers left and two sisters.

(FAC 03, 1:01:42)

- La recherche de confirmation / de l'accord de l'interlocuteur : *is that right?* peut signifier une véritable demande du locuteur auprès de l'interlocuteur de confirmer ou de donner son accord à l'interprétation des informations que l'interlocuteur a fournies (*cf.* Holmes et Chiles, 2010) :

AN : **Is that right huh?** Thir-- thirteen you said. [*âge à laquelle la mère de INT1 s'est mariée, mentionné plus tôt dans la conversation*] < INT1: yeah. >

(FAC 03, 21:48)

⁷⁴ Les données de Holmes et Chiles (2010) sont tirées d'une enquête sur la langue parlée au travail. La majorité des données proviennent des enregistrements des réunions d'équipe qui consistent généralement en la présentation d'informations, donc en des monologues séquentiels plutôt que des échanges dialogaux.

Cette catégorie se différencie de la catégorie précédente en ce qu'elle s'applique à la forme interrogative plutôt qu'aux formes affirmatives. Elle s'oppose à divers emplois (voir ci-dessous) qui se rapprochent davantage du signal d'écoute, c'est-à-dire dans lesquels le locuteur ne cherche pas nécessairement une réponse formelle à sa question, sait déjà que sa compréhension est correcte (cf. Banuazizi et Creswell, 1999) et emploie plutôt la formule pour signaler qu'il est attentif et participe activement à la conversation en cours.

- Le rappel d'une information déjà connue : *right* et *that's right* peuvent signaler que le locuteur vient de se rappeler d'une information qu'il était censé déjà savoir (cf. Koivisto, 2013).

AN : Tu sais-tu d'autre chose qui était ben commun itou uh pis qu'on voit on voit ben rarement ça c'est des clous!

INT : Des clous, **that's right!** Ouais.

(FAC 07, 1:00:58)

- La réaction à une information nouvelle : cette catégorie s'applique lorsque l'information était inconnue du locuteur ou semble l'être; il émet alors l'une des unités cibles en réponse. Dans l'exemple ci-dessous, on peut être relativement certain que l'information (le fait que INT a été hospitalisé pour la même maladie dont son père est décédé) est nouvelle pour AN, parce qu'il vient d'interroger INT à ce sujet :

AN : pis c'était-tu pour la même affaire? [*l'hospitalisation de INT*] < INT : oui oui. Ouais.
> **Is that right?**

(FAC 07, 6:25)

- La réaction à une situation exceptionnelle : il a été observé dans le corpus que *right*, *that's right* et *is that right?* se distinguent de *yeah*, *yes* et *yep* – qui signifient aussi l'affirmation et l'accord – en ce qu'ils s'emploient plus souvent lorsque l'interlocuteur raconte une situation qu'on peut qualifier d'exceptionnelle. Pour les fins de l'analyse, une situation exceptionnelle a été définie comme une situation inhabituelle ou remarquable (par exemple, le fait d'avoir pris la retraite avant l'âge de 60 ans), ou une situation qui aurait suscité de vives émotions (par exemple, le fait d'avoir été témoin d'une manifestation du KKK ou d'avoir été atteint de tuberculose). Lorsqu'une des unités cibles a été employée en réaction à une situation de ce genre, cette catégorie a été attribuée à l'occurrence.
- Le fait de constituer une intervention complète : cette catégorie regroupe les emplois de *right*, *that's right* ou *is that right?* qui se rapprochent de la fonction de signal d'écoute canonique, en ce que l'unité

en question ne constitue ni la clôture d'un énoncé plus long, ni l'ouverture d'un énoncé plus long de la part du locuteur, comme dans l'exemple qui suit :

INT : ça a tout grimpé par-dessus moi ça, un par de-- un par derrière l'autre [*les voitures lors d'une collision routière*] < AN : **Is that right?** > [rire] Nervous wreck the both of us!

(FAC 08, 47:14)

Cette catégorie est donc basée sur des critères plutôt formels, mais qui pourraient servir à mieux cibler la fonction discursive attribuée à l'unité par les locuteurs.

- Le fait de recevoir une réponse explicite : cette catégorie est également basée sur un critère formel, à savoir que l'interlocuteur offre ou non une réponse explicite (*oui, yeah, non, etc.*) à la question *is that right?*, encore dans le but de mieux cibler la fonction discursive attribuée à l'unité par les locuteurs.

Comme avec *yeah / yes / yep*, les catégories ci-dessus ne sont pas mutuellement exclusives, à l'exception des catégories touchant uniquement les énoncés soit affirmatifs soit interrogatifs. Le tableau 5.12 présente la fréquence par unité cible de chaque catégorie d'analyse décrite ci-dessus.

TABLEAU 5.12 : FRÉQUENCE OBSERVÉE DES CATÉGORIES D'ANALYSE DE *RIGHT*, *THAT'S RIGHT* ET *IS THAT RIGHT?*

	Confirmation	Recherche de confirmation	Rappel d'information déjà connue	Réaction à une nouvelle information	Réaction à une situation exceptionnelle	Constitue une intervention complète	Reçoit une réponse explicite
<i>right</i> (n = 11)	11	~	1	0	0	5	~
<i>that's right</i> (n = 11)	11	~	7	0	2	2	~
<i>is that right?</i> (n = 24)	~	1	~	22	23	23	16
Total (N = 46)	22	1	8	22	25	30	16

On voit que les formules affirmatives sont effectivement employées dans le corpus avec le sens « canonique » de confirmation, qui correspond à 100 % de leurs occurrences. L'unité *right* est principalement employée pour la confirmation, sans plus, tandis que sept occurrences sur 11 (63,6 %) de *that's right* sont employées pour signaler en même temps le rappel d'une information déjà connue. Il est intéressant de noter qu'aucune de ces sept occurrences ne débute par *oh* pour indiquer un changement d'état épistémique, ce qui vient à l'encontre des résultats rapportés par Koivisto (2013). Dans l'ensemble, cependant, *right* et *that's right* ne présentent rien de remarquable dans la façon dont ces unités sont employées dans le corpus d'analyse, mis à part le fait qu'elles sont préférées à *yeah, yes* ou *yep* par certains locuteurs et plus fréquentes en général que les éléments français du même PLM (voir ci-dessous).

Is that right? n'est employée qu'une seule fois comme formule de vérification de la compréhension d'une information fournie par l'interlocuteur qui nécessite une confirmation de la part de l'interlocuteur avant de continuer. En l'occurrence, il s'agit d'une tentative de relancer la conversation après une blague :

AN : se mariaient jeune dans ce temps-là?

INT1 : oh (.) y faisait frette! [rire]

AN : [rire prolongé] **Is that right huh?** Thirt- thirteen you said? < INT1: yeah >

(FAC 03, 21:48)

Ici, la conversation porte sur l'âge de la mère de INT1 au moment de son mariage, soit 13 ans, une information fournie quelques tours de parole auparavant par INT1. La blague de INT1 engendre plusieurs secondes de rires partagés, après quoi AN tente de ramener la conversation à ce sujet en demandant à INT de confirmer qu'AN avait bien compris, c'est-à-dire que INT avait bien dit que sa mère avait 13 ans.⁷⁵ En revanche, *is that right?* est employée 22 fois sur 24 (91,7 % des occurrences) en réaction à une nouvelle information que le locuteur ne connaissait pas ou fait semblant de ne pas connaître.

Is that right? n'est donc pas employée dans le corpus d'analyse pour contrôler le discours en mettant des mots dans la bouche de l'autre, comme l'ont observé Holmes et Chiles (2010). À l'exception de cette seule occurrence où l'unité est utilisée pour ramener la conversation à un sujet donné, *is that right?* est plutôt employée comme un signal d'écoute. Dans huit occurrences sur 24 (33,3 %), l'interlocuteur ne semble pas interpréter *is that right?* comme une « vraie question », en ce qu'il continue son énoncé sans y répondre explicitement, comme dans l'exemple suivant :

INT : Y est mort de cancer. [son mari] < AN : **oh is that right?** > **Mais chuis revenue par icitte**
à cause de ça moi-- que mes-- mon père était tout seul là

(FAC 01, 34:12)

De plus, 23 occurrences sur 24 de *is that right?* constituent des interventions complètes en elles-mêmes, indiquant que cette formule est employée surtout comme signal d'écoute : le locuteur n'a pas produit et ne cherche pas à produire un énoncé plus long ou complexe, mais à signaler son attention et sa participation active à l'interaction. Il est à noter que de 22 des 23 occurrences de *is that right?* constituant des interventions complètes sont produites par FAC : AN. En fait, ce locuteur utilise *is that right?* presque exclusivement comme signal d'écoute (22 occurrences de cette unité sur 23 ou 95,7 %), ce qui constitue aussi la majorité de sa contribution à la catégorie (22 occurrences sur 35 ou 62,9 %) et presque la moitié de la catégorie au complet (22 occurrences sur 46 ou 47,8 %).

La fréquence remarquable avec laquelle FAC: AN a recours à *is that right?* peut s'expliquer par son rôle en tant qu'animateur de la conversation. Quoique l'interaction ressemble davantage à une conversation qu'à une entrevue, l'objectif demeure d'enregistrer l'histoire vécue des autres membres de la communauté. Il

⁷⁵ C'est principalement la prosodie qui indique que la question de AN porte bien sur ce qu'il dit immédiatement après, et non sur ce qui vient de passer.

incombe donc à FAC : AN d'encourager ses interlocuteurs à parler le plus possible, en particulier de leurs expériences d'intérêt historique. Il doit donc gérer l'interaction de façon à mettre les interviewés à leur aise et à gagner leur confiance. À la différence de *yeah, is that right?* lui permet de signaler explicitement qu'une information fournie par son l'interviewé(e) est une nouveauté pour lui,⁷⁶ et donc d'un grand intérêt. C'est probablement pour cette raison qu'il tend à employer *is that right?* en réaction à des situations exceptionnelles racontées par ses interlocuteurs : cela lui permet de montrer une réaction plus vive et une attention encore plus soignée que *yeah*. De ce fait même, *is that right?* signale qu'il a bien compris la nature extraordinaire et significative pour l'interviewé(e) de ce qu'il ou elle vient de lui raconter et qu'il s'investit émotionnellement dans l'histoire :

INT : et pis they [*les membres du KKK*] were running around the: the orchard

AN: ok (.) {they never came to your place?}

INT: {y avaient de:s} des des des flambeaux < AN : **oh is that right?** > ouais.

(FAC 01, 18:27)

INT : pis elle [*l'enseignante*] a venu pis elle l'a fessé [*le frère d'INT*]. Pis nous autres moi pis ma sœur on se levait pour y aller grabber le bras. < AN : **oh is that right?** > Oh: yeah. And I wish she had hit him one more time!

(FAC 01, 22:06)

INT: [*parlant de comment sa mère a fait vivre ses quatre enfants après la mort de son père*] pendant tren:te-cinq ans de temps (.) elle jouait pour < AN : **oh is that right?** > des wedding receptions qu'on appelle là.

(FAC 07, 25:08)

Ainsi, en accord avec les résultats de Koivisto (2013) pour *oh that's right!*, l'emploi de *is that right?* par FAC : AN peut être interprété comme une forme de travail sur les faces en présence. Raconter un événement sensible ou remarquable qu'on a vécu représente un risque; l'interlocuteur pourrait ne pas croire à la véracité de l'événement, ou pourrait ridiculiser la personne qui se confie à lui, ou minimiser les sentiments qu'on lui révèle. En démontrant qu'il porte attention non seulement à ce que l'interviewé(e) raconte mais à la signification de cette histoire pour ce dernier, et en ayant la réaction vive escomptée, FAC : AN valide l'expérience de l'autre malgré sa nature exceptionnelle. Il protège ainsi la face de son interlocuteur, et en se révélant une personne digne de confiance, il protège aussi la sienne, tout en offrant une écoute empathique qui encourage l'interviewé(e) à continuer à parler et à collaborer avec lui dans l'interaction en cours.

⁷⁶ Que l'information soit réellement nouvelle pour lui ou qu'il le savait d'avance, c'est bien (l'apparence de) la nouveauté qu'il signale à son interlocuteur par son emploi de cette unité.

Si on compare les éléments français et anglais de ce PLM, la différence principale en est une de fréquence : les éléments français sont très peu employés. Plusieurs unités lexicales françaises pourraient faire partie d'un paradigme multicodeur avec *right*, *that's right* et *is that right?* : *c'est vrai*, *exact*, *c'est ça*, *c'est (tu) vrai?*, *pour vrai?* et même *ah oui?*, pour ne nommer que ceux-là. Dans le corpus d'analyse, on observe également *t'es ben (avec ça)*, chez FAC 07 : INT notamment, pour confirmer l'exactitude d'une information fournie par l'interlocuteur. Cependant les seules unités du français de ce paradigme qui apparaissent en position isolable sont deux occurrences de *ah ouais* ou *ah oui*, trois occurrences de *ouais hein?* ou *oui hein?* et deux occurrences de *c'est (tu) vrai?*. Avec si peu d'occurrences, il n'est pas possible de parler de tendances dans l'emploi de ces unités, mais on peut toutefois constater la préférence nette des locuteurs étudiés pour *right* / *that's right* / *is that right?* en position isolable.

Plusieurs raisons peuvent expliquer cette préférence. Étant donné que FAC : AN produit 35 sur 46 (76,1 %) des occurrences de cette catégorie à lui seul, il peut s'agir des habitudes d'usage dont les causes ne peuvent ressortir de la seule observation de ce corpus, par exemple, des raisons qui relèveraient de ses expériences à l'école, au travail, ou avec une ou des personnes en particulier de son entourage. Cela dit, on peut remarquer le nombre important d'unités et de formules du français qui peuvent, selon l'interprétation qu'on en donne, entrer dans un PLM dont la partie anglaise ne compte que trois unités ou formules. Cette asymétrie suggère qu'il n'y a pas d'élément français qui permet de traduire exactement la même visée référentielle que *right*, *that's right* ou *is that right?*, du moins dans l'usage général du français et de l'anglais. Par exemple, un locuteur pourrait trouver que *right* implique le fait de confirmer une vérité subjective (« ceci est correct selon ma connaissance des faits ») tandis que *c'est vrai* implique une vérité plus objective (« ceci est la vérité telle qu'elle existe dans le monde »), et préférer la flexibilité et le « déni plausible » du premier. Cependant, le volume de données disponibles dans le corpus d'analyse ne permet pas de trancher sur les sens que FAC : AN (ou tout autre locuteur du corpus) donne à *c'est vrai* ou d'autres éléments semblables par rapport à *right*, *that's right* ou *is that right?* et donc si une différence de sens pourrait influencer sur ses habitudes d'usage dans ce cas. Tout ce que les données permettent d'affirmer est que les locuteurs du corpus FAC, et notamment FAC : AN, se servent de la juxtaposition des codes, lorsqu'ils parlent en français, pour pouvoir employer les ressources linguistiques de l'anglais *right*, *that's right* et *is that right?* pour confirmer des informations, pour gérer la conversation et pour protéger les faces en présence – ce qui ne leur est possible que parce qu'ils ont accès à ces ressources, c'est-à-dire, parce qu'ils sont bilingues anglais-français.

5.5 Les marqueurs pragmatiques

Cette section présente les résultats d'analyse de la catégorie des marqueurs pragmatiques. La section 5.5.1 jette les bases de l'analyse en discutant de la définition de *marqueur pragmatique* et de la sélection des marqueurs individuels *so*, *anyway* et *now*, avant de présenter un aperçu de la distribution de ceux-ci par

locuteur.⁷⁷ En raison du petit nombre d'occurrences de chacun de ces marqueurs individuellement, les autres sous-sections présentent l'entièreté des occurrences repérées de *so* (5.5.2), *anyway* (5.5.3) et *now* (5.5.4) avec le contexte transcrit, les sens et fonctions discursives identifiées et d'observations qualitatives qui en découlent.

5.5.1 Aperçu

Avec un total de 69 occurrences en tout, les marqueurs pragmatiques forment la deuxième catégorie la plus fréquente d'occurrences isolables d'alternance codique vers l'anglais, après *yeah / yes / yep*. Cependant, cette catégorie est peut-être la plus difficile à définir : même si les marqueurs pragmatiques ou marqueurs discursifs font l'objet de maintes études en linguistique, il n'y a de consensus ni sur les critères qui définissent cette catégorie d'analyse, ni sur les unités qui y appartiennent ni encore sur comment la nommer (voir sur ce sujet les discussions dans Dostie, 2004 ; Fox Tree, 2010 ; Fraser, 1999). Il ne relève toutefois pas des objectifs de la présente étude de trancher sur ces questions, qui pourraient constituer une étude à part entière. Ainsi, la définition adoptée ici est large : les marqueurs pragmatiques constituent un ensemble hétérogène d'unités qui servent à relier deux ou plusieurs énoncés, idées, thématiques, etc. ensemble en structurant et en organisant le discours, ainsi contribuant à la gestion de l'interaction et même à l'entretien des relations sociales (Bolden, 2006 ; Dostie, 2004 ; Fraser, 1999). Suivant Dostie (2004) l'appellation « marqueurs pragmatiques » a été retenue afin d'englober l'ensemble des marqueurs de ce genre sans chercher à suivre une classification particulière selon le type ou la fonction des marqueurs (voir Bolden, 2006 ; Dostie, 2004 ; Fox Tree, 2010 ; Fraser, 1999 pour une vue d'ensemble de ces classifications).

L'accent mis, dans la définition de la catégorie, sur les fonctions discursives des marqueurs pragmatiques n'implique pas pour autant que ces unités n'aient pas de sens ou de signification bien à elles.⁷⁸ On n'a qu'à penser aux différences entre des discours dont les idées sont liées par *et* ou *pis* au lieu de *mais* ou *par contre* pour comprendre que les marqueurs pragmatiques ont effectivement des significations différentes. Les marqueurs pragmatiques ont toutefois en commun les fonctions de relier des éléments, de structurer le discours et de contribuer à la gestion de l'interaction; ce sont donc ces fonctions qui les réunissent en une même catégorie. Il convient cependant de rappeler ici le constat formulé dans la discussion des sens et des fonctions discursives de *yeah / yes / yep* à la section 5.3.4, à savoir que le sens en usage et la fonction en discours sont

⁷⁷ Les marqueurs *anyway* et *now* sont produits exclusivement dans le corpus FAC; seul le marqueur *so* est produit dans d'autres corpus. Pour cette raison, la fréquence et la distribution par corpus ne sont pas présentées séparément.

⁷⁸ Parallèlement, l'inclusion des unités indiquées dans le tableau 5.13 dans la catégorie de marqueurs pragmatiques n'implique pas que ces mêmes unités sont toujours des marqueurs pragmatiques lorsqu'elles apparaissent en discours. *So*, par exemple, peut aussi indiquer un degré d'intensité ou de condition suffisante, comme dans « My face was so red » ou « She was so worried that she couldn't concentrate ».

difficiles à distinguer dans l'usage, en particulier pour des marqueurs pragmatiques qui ont un rôle fonctionnel important.

Le tableau 5.13 présente les marqueurs pragmatiques de l'anglais repérés en position isolable et leur fréquence d'occurrence dans le corpus d'analyse. Étant donné l'hétérogénéité de la catégorie, j'ai décidé d'analyser les unités qui la constituent individuellement. Or, en raison du petit nombre d'occurrences de la plupart des unités de cette catégorie, seules les trois unités les plus fréquentes ont été analysées en détail : *so*, *anyway*, et *now*. À cet effet, le tableau 5.14 présente la fréquence de production de ces unités par locuteur. Des lignes supplémentaires ont été ajoutées pour donner les totaux des deux locuteurs, FAC : AN et PFC : AN, qui participent à plusieurs enregistrements; un locuteur ou un enregistrement qui n'apparaît pas dans le tableau n'a produit ou ne contient aucune occurrence des unités cibles.

TABLEAU 5.13 : FRÉQUENCE DES MARQUEURS PRAGMATIQUES DE L'ANGLAIS ISOLABLES, ANALYSE INDIVIDUELLE, TOUS LES CORPUS

Marqueur pragmatique	# d'occurrences (tout le corpus)
<i>so</i>	14
<i>anyway</i>	10
<i>now</i>	10
<i>but / yes but</i>	7
<i>y'know</i>	6
<i>too</i>	5
<i>well</i>	5
<i>see</i>	4
<i>of course</i>	3
<i>if / and if</i>	2
<i>like</i>	2
<i>because</i>	1
<i>first of all</i>	1
Total	68

TABLEAU 5.14 : FRÉQUENCE DE PRODUCTION DE *SO*, *ANYWAY* ET *NOW* PAR LOCUTEUR, TOUS LES CORPUS

Locuteur	<i>so</i>	<i>anyway</i>	<i>now</i>	Total
FAC : AN (tous)	6	2	9	17
FAC 01 : AN	3	2	7	12
FAC 01 : INT	3	6	0	9
FAC 08 : INT	1	2	0	3
FAC 07 : AN	2	0	0	2
PFC : AN (tous)	2	0	0	2
FAC 03 : AN	0	0	1	1
FAC 03 : INT1	0	0	1	1
FAC 06 : AN	1	0	0	1
FAC 06 : INT1	1	0	0	1
FAC 08 : AN	0	0	1	1
CB-PRS 02A : INT	1	0	0	1
PFC 01 : AN	1	0	0	1
PFC 04 : AN	1	0	0	1

Les marqueurs pragmatiques les plus employés dans le corpus étant quand même peu fréquents, il n'est pas possible d'établir des catégories selon les sens et les fonctions discursives de chaque unité qui permettraient de faire une analyse quantitative. L'analyse de ces unités consiste donc en des observations d'ordre plus qualitatif.

5.5.2 So

Le tableau 5.15 présente les 14 occurrences isolables de *so* avec leurs contextes immédiats, et le tableau 5.16 le seul élément français isolable du même PLM ayant été identifié dans le corpus. Pour faciliter la discussion des occurrences transcrites, chacune comporte un numéro de référence commençant par la première lettre de l'unité cible qu'elle présente. Il est à noter que, comme dans le cas de *right*, on pourrait penser à plusieurs unités ou formules en français qui pourraient faire partie d'un tel paradigme, et ce, pour tous les marqueurs pragmatiques analysés ici.⁷⁹ De fait, les marqueurs pragmatiques ont des sens relativement abstraits qui dépendent beaucoup du contexte (Fox Tree, 2010).⁸⁰ Dans le cas de la présente étude, cependant, le problème des unités lexicales française pouvant faire partie de ce PLM est largement réglé, en ce que les unités du français parmi lesquelles choisir sont déjà restreintes à celles qui apparaissent dans le corpus en position isolable. Ainsi, la seule unité lexicale française comparable qui apparaît dans le corpus d'analyse est *ça fait (que)*.

Une brève revue de la littérature sur *so* indique que le sens « canonique » de cette unité est la cause ou l'inférence (voir la revue de la littérature dans Bolden, 2006), c'est-à-dire que « *so* signals that the following segment is to be interpreted as a conclusion which follows from prior discourse » (Fraser, 1999, p. 945). Cependant, Bolden (2006) observe que le marqueur *so* est souvent employé pour lancer un nouveau sujet de discussion ou une nouvelle séquence d'énoncés – autrement dit, lorsqu'il n'y a pas de discours directement précédent auquel relier ce qui suit – ou encore est employé lorsqu'il n'y a pas de suite logique évidente entre ce qui précède et ce qui suit. S'appuyant sur l'analyse d'un corpus oral, Bolden (2006) conclut que *so* est employé dans ces contextes pour indiquer que, quoiqu'il s'agisse techniquement d'une rupture dans la conversation, le locuteur pense au sujet qui suit depuis un certain temps déjà. *So* a donc une fonction discursive de reprise d'un objectif ou d'une ligne de pensée abandonnée ou interrompue, même si cette ligne de pensée n'est pas immédiatement évidente dans le discours précédent. Bolden compare *so* avec *oh*, qui semble indiquer que le

⁷⁹ Comme il l'a déjà été signalé ailleurs dans ce chapitre (cf. notamment la section 5.1 et 5.4.3), ces résultats confirment déjà en partie l'hypothèse qui sous-tend la présente étude, à savoir que les locuteurs étudiés produisent de l'alternance codique afin d'employer des ressources dans une langue qui permettent de traduire des visées qui ne seraient pas aussi facilement exprimables dans l'autre (parce qu'elles ne sont pas lexicalisées tout à fait de la même façon).

⁸⁰ On peut faire un parallèle ici avec les prépositions, qui ont aussi des sens abstraits qui dépendent du contexte (voir à ce sujet Gruntman, 2011 ; voir aussi Duffley, 1992 sur le sens de *to* en anglais)

TABLEAU 5.15 : LES OCCURRENCES ISOLABLES DE SO EN CONTEXTE

# Ref.	Enreg.	Temps	Occurrence
S1	FAC 01	34:46	AN : What- what year- what year you talking about? INT : En quarante. AN : so : uh: dans le temps que la Guerre commençait là.
S2	FAC 01	35:56	AN : Avez-vous resté longtemps à ****? INT : uh: neuf an- neuf ans je pense presque neuf ans < AN : so (.) > it was nine years.
S3	FAC 01	41:39	INT : [parlant de son père] he-- he wrote his name-- y a été jusqu'au quatrième grade. < AN : oh ok > à l'école. (.) AN : so uh what was it like in ****?
S4	FAC 01	49:39	INT : Oui mes enfants y avait pas ça d'icitte moi quand je me suis en venue par icitte. Pis y'avaient pas parti avec moi. So uh: ben sontait tout mariés pis uh:
S5	FAC 01	54:24	INT : C'est ça qui commence ça tsé. T'as mal aux jambes puis après là ça va: c'est la hanche qui fait mal mais on sait pas ça. So! Well I brought these girls and they all went you know they all uh: (.) did their thing < AN : auditioned? > [pour une pièce de théâtre]
S6	FAC 01	1:24:26	INT : [parlant d'une réunion municipale sur l'installation des lampadaires] y savaient faire ça mais y voulaient pas le faire. (.) Pis j'ai dit on va faire ça comme faut ben la on: c'est XX de rester icitte! So j'es ai ma-- j'es ai : brassé la baratte un escousse
S7	FAC 06	1:40	AN : T'étais là souvent toi hein? [dans le magasin familial de AN] INT1 : ouais! ouais! AN : Pis tes frères. < INT1 : ouais! > L*** INT1 : L*** AN : pis R*** INT1 : pis G*** AN : G*** INT1 : G***. Yeah! AN : Et pis: uh (.) so : (.) you know INT1 : There's only one -- G-- uh: there's only one living. G***.
S8	FAC 06	43:16	AN : oh so they quit y: voulait pu < INT1 : yeah! > assurer icitte là X-- INT1: [rire] y voulait pu assurer icitte! < AN : yeah. > so . We st-- we transferred them to the other company!
S9	FAC 07	32:05	INT: c'est là que j'ai appliqué à *** < AN : o:k > j'étais à *** tu comprends? Mais j'ai été trois ans avant d'y aller tu comprends? AN : yeah. (.) < INT : ouais. > So : uh qu-- uh: (.) t'as perdu to-- ton: ton extra fifty pounds uh easily ou...?
S10	FAC 07	48:56	AN : [parlant des voisins qui sont partis vivre en ville] ça a jamais revenu ça < INT : non. non. > X y se promenaient un ti peu là < INT : ouais > pis uh. so : uh (.) on voirait pu ça INT : ouais. (.) uh: chuis parès penser, ça me passe dans la tête là
S11	FAC 08	48:44	INT : ça ça drivait pareil, the trunk was in the back seat but (.) ça drivait pareil! [rire] < AN : [rire] > So! We managed to drive that car to port like that!
S12 ⁸¹	CB-PRS 02A	8:11	INT : [parlant de l'emploi des mots familiers pour les parties du corps] ça dépendrait comment que le patient y parle de l'habitude. < C : ké! > Pis ça l'a pas dérangé du tout d-- n-- ni-- < C : ni l'un ni l'autre, ouais. > ni l'un ni l'autre. (.) < C : ké! > so ... C : ça fait que ça dérange pas < INT : non > tout ça
S13	PFC 01	10:38	INT : moi je ah j'aime ça uh travailler dans le shop. Ouais AN: ok (.) [tourne des pages] > < INT : comme uh loisir. > so . (.) Est-ce qu'il y a une raison particulière que tu vis (.) à ****? Disons.
S14	PFC 04	13:29	AN : oui [rire] j'ai eu une classe qui avait trois élèves. Ça c'est un minimum [rire] < INT1 : ouais: > yeah. INT1 : entécas ça change: ça change (.) AN : so uh (.) vous étiez dans le jardin?

⁸¹ Malgré le statut codique potentiellement ambigu de l'unité précédant immédiatement l'unité cible, il y a suffisamment d'indices disponibles pour qu'on puisse raisonnablement conclure qu'il s'agit d'un changement de code dans le déroulement linéaire de la conversation.

sujet qui suit vient tout juste à l'esprit du locuteur (cf. Koivisto, 2013, pour qui *oh* indique un changement d'état). Ainsi, selon Bolden (2006), *so* indique une attention – et une intervention – centrée plutôt sur l'autre et sur la continuité de l'interaction et de la relation, tandis que *oh* indique une attention centrée sur soi. En signalant le fait qu'on pense à l'autre depuis un certain temps déjà, mais que penser à soi vient tout juste à l'esprit, *so* (et *oh*) renforcerait les relations sociales et participerait à la protection des faces en présence.

TABLEAU 5.16 : LES OCCURRENCES DE ÇA FAIT (QUE) EN CONTEXTE

# ref.	Enreg.	Temps	Occurrence
C1	FAC 01	58:54	A N: and uh: then she started researching it, and found out that her: great grandmother avait été à l'école and was sent back whenever she learned to speak white! (.) < INT : [silencieusement : oh:!] > Ça fait que (.) y en a eu de la préjudice.
C2	FAC 01	1:07:08	INT : ça fait trente-quelques années je suis arrivée là moi. AN : ouais. Back in these parts. INT : ouais in back in these parts yeah. AN : Ça fait que : uh: you have-- that's quite a few (.) apartments here
C3 ⁸²	FAC 08	2:59	INT : pis là ça commence à diminuer. Uh: (.) people go to camp -- they go fishing, they (.) in the winter they go skidoing, w-- they take off for the weekend, they they go to *** shopping, they go to ***. Ça fait que : [pfft] (.) So I just stopped Saturdays altogether.
C4	FAC 08	10:58	INT: so they was both L (.) *** [nom de famille] and they both had three kids! (.) Ça fait que ... yeah. < AN : isn't that something! > ça fait, c'est-- c'était mêla-- c'était mêlant!
C5	FAC 08	10:58	INT: so they was both L (.) *** [nom de famille] and they both had three kids! (.) Ça fait que... yeah. < AN : isn't that something! > ça fait , c'est-- c'était mêla-- c'était mêlant!
C6	FAC 08	14:48	AN : when did they consolidate the schools into one? (.) I know in *** it happened in forty-nine. INT: oui hein? (.) no it was after that, in- (.) I think- (.) I think I was in high school at the time. See and I started high school in nineteen-sixty. Ça fait que ... AN : oh: yeah! < INT : yeah s-- > early sixties.

Cependant, dans son étude sur des locuteurs de l'anglais comme *lingua franca* étudiant ou enseignant à l'Université d'Hambourg, House (2013) révèle que ces locuteurs emploient *so* d'une manière tout à fait différente de ce que Bolden (2006) a observé. Les locuteurs non natifs de l'anglais emploieraient plutôt *so* pour gagner du temps lorsqu'ils éprouvent des difficultés dans la planification de leur énoncé. Selon House, *so* ressemble davantage dans cet emploi à *erm (um)* qu'à un marqueur discursif. Toutefois, son argumentation n'est pas tout à fait convaincante, dans la mesure où les exemples donnés ne démontrent pas clairement que les locuteurs n'emploient pas *so* pour gagner du temps justement en raison de sa connotation positive quant à l'attention portée à l'autre et à la continuité de l'interaction, tout comme cette auteure suggère qu'ils le font avec *yeah* (cf. la section 5.3.4.1).

D'un nombre aussi restreint d'occurrences à analyser, il n'est évidemment pas possible de tirer des conclusions définitives, mais certaines observations ressortent toutefois de la comparaison des emplois de *so* et de *ça fait (que)* dans le corpus FAC. Ces emplois sont résumés dans le tableau 5.17.

⁸² Malgré le statut codique potentiellement ambigu de l'unité précédant immédiatement l'unité cible, il y a suffisamment d'indices disponibles pour qu'on puisse raisonnablement conclure qu'il s'agit d'un changement de code dans le déroulement linéaire de la conversation.

On voit que le sens « canonique » d'inférence logique (y compris la causalité) est effectivement très présent. On remarque également que le marqueur *so* est souvent employé pour accomplir plusieurs fonctions en même temps (dans six occurrences sur 12), tandis que *ça fait que* remplit le plus souvent une seule fonction (cinq fois sur six). Notamment, *ça fait que* n'est jamais utilisé pour reprendre un sujet ou une ligne de pensée interrompue ou abandonnée, et seulement une fois sur six pour amorcer une rupture conversationnelle comme un changement abrupt de sujet de discussion. Or, *so* est employé avec ces fonctions quatre fois sur 12 et six fois sur 12, respectivement. En revanche, *ça fait que* est plus fréquemment employé pour exprimer une conclusion implicite. Dans ces cas, le locuteur ne cherche pas à expliciter la conclusion de son énoncé, car il présume que son interlocuteur complétera le non-dit lui-même (implicitement ou explicitement). Il est à noter que cette fonction est en lien direct avec le sens d'inférence logique, en ce qu'il repose sur l'implication forte qu'il y a une telle inférence à faire.

TABLEAU 5.17 : CATÉGORIES D'ANALYSE OBSERVÉES POUR *SO* ET *ÇA FAIT (QUE)*, CORPUS FAC SEULEMENT

# ref.	Locuteur	Inférence logique	Reprise d'un sujet précédent	Rupture conversationnelle	Conclusion implicite
S1	FAC 01 : AN	x			
S2	FAC 01 : AN	x			
S3	FAC 01 : AN		x	x	
S4	FAC 01 : INT	x			
S5	FAC 01 : INT		x	x	
S6	FAC 01 : INT	x			
S7	FAC 06 : AN	x		x	
S8	FAC 06 : INT1	x	x		
S9	FAC 07 : AN		x	x	
S10	FAC 07 : AN	x		x	
S11	FAC 08 : INT	x			
C1	FAC 01 : AN	x			
C2	FAC 01 : AN			x	
C3	FAC 08 : INT	x			
C4	FAC 08 : INT	x			x
C5	FAC 08 : INT	x			
C6	FAC 08 : INT	x			x

On observe également que FAC : AN produit huit des 18 occurrences isolables de ce PLM, et FAC 08 : INT en produit cinq. Quoique FAC : AN semble préférer employer *so* et FAC 08 : INT *ça fait que*, les deux produisent au moins une occurrence de chaque unité, ce qui suggère qu'il n'y a pas là une question d'accès lexical ou de connaissances dans l'une ou l'autre langue, donnant plus de poids à la possibilité que ce soit une stratégie d'expression. On remarque également que FAC : AN produit quatre des cinq occurrences qui ont la fonction de rupture conversationnelle. Cela est probablement relié à son rôle de facilitateur de la conversation : il lui incombe de ramener la conversation à des sujets qui permettent de documenter l'histoire de la communauté, mais il doit aussi se montrer attentif à son interlocuteur et à la continuité de l'interaction. Le fait d'amorcer une rupture avec un marqueur comme *so* lui permet de protéger les faces en présence et de maintenir une atmosphère de collaboration.

Pris dans l'ensemble, cet aperçu des emplois de ce PLM suggère que *so* est employé de manière plus « flexible » que *ça fait que* quant aux fonctions remplies dans le discours. Notamment, *ça fait que* est principalement réservé pour l'expression de l'inférence logique. Une hypothèse pour expliquer cet emploi préférentiel est que ce marqueur exprime de façon explicite une causalité logique, le rendant en quelque sorte moins abstrait ou plus transparent que *so* pour ce sens précis. Pour sa part, *so* prend des fonctions relatives à la gestion de l'interaction : il permet de signaler la fin d'une parenthèse et la reprise d'un sujet précédent, et il prévient l'interlocuteur qu'un nouveau sujet est amené, tout en signalant qu'on est attentif à la continuité de l'interaction en tant que tout cohérent. Il semble donc que, de par sa plus grande flexibilité fonctionnelle, *so* participe davantage que *ça fait que* à la gestion des relations interpersonnelles et à la protection des faces (cf. Bolden, 2006). Quoique les conclusions qu'il est possible de tirer de 18 occurrences soient limitées, ces tendances suggèrent des pistes intéressantes à explorer dans une prochaine étude de plus grande envergure sur le sujet.

5.5.3 *Anyway*

Le tableau 5.18 présente les 10 occurrences de *anyway* repérées en position isolable dans le corpus d'analyse. Pour faciliter la discussion des occurrences transcrites, un numéro de référence qui commence par les premières lettres de l'unité cible qu'elle présente. Il n'y a malheureusement pas d'occurrences isolables en français dans le corpus d'analyse qui pourraient constituer un PLM avec *anyway*, comme *de toute façon*, *en tout cas*, *néanmoins* (pour certains sens, voir ci-dessous), etc. Il y a lieu ici de discuter des sens de *anyway*, mais en raison de l'absence d'unités en français auxquelles comparer les emplois, il n'est évidemment pas possible d'effectuer une analyse différentielle.

Ferrara (1997) distingue trois sens de *anyway*, soit additif (semblable à « besides », « en plus », « d'ailleurs »), de mis à l'écart (« dismissive »; semblable à « nonetheless », « peu importe », « néanmoins ») et de reprise (« resumptive »). Ferrara considère que seulement *anyway* au sens de reprise, qui relie l'énoncé qui suit à un segment de discours antérieur suite à une digression, est un marqueur pragmatique. Dans son corpus, 89 % des occurrences de *anyway* sont employées avec ce sens de reprise. La grande majorité de ces occurrences sont produites lorsque le locuteur vient de fournir des détails supplémentaires ou de commenter son état épistémique en lien avec ce qu'il raconte, par exemple :

⁸³ Il s'agit de différencier les numéros de référence pour *anyway* (AN) ici de ceux de *astheure* (A) dans la section 5.5.4.

We were real concerned about the money. I don't know if we had credit cards back then or not.
Anyway, we stayed overnight that night.

(exemple 12c dans Ferrara, 1997, p. 355, emphase dans l'original)

TABLEAU 5.18 : LES OCCURRENCES ISOLABLES DE ANYWAY EN CONTEXTE

#	Enreg.	Temps	Occurrence
AN1	FAC 01	31:52	INT : [en parlant de la ferme de ses parents à son enfance] C'était: y avait semé des patates. (.) < AN : quand- > Pi:s uh anyway c'était rochu
AN2	FAC 01	53:05	INT : [parlant d'un individu actif dans la promotion du français, qui lui a impliqué dans la production d'une pièce de théâtre en français] Et pis um: uh je sais pas si c'est à cause de lui, mais anyway y a une année en nineteen seventy-three I think seventy-three or seventy-four
AN3	FAC 01	56:11	AN : [parlant d'un individu actif dans la promotion du français, qui lui a impliqué dans la production d'une pièce de théâtre en français] pis y se rappellerai de ça là lui? INT : ça devrait. Pis anyway parle de moi itou < AN : ouais > if you're there ouais see if he remembers me.
AN4	FAC 01	1:33:03	AN : Ouais. Ouais uh grand-mère C*** vous voulez dire là. INT : Oui grand-mère C*** oui. AN : Ouais. < INT : ouais > (.) Mais anyway uh most of those camps there on the *** they-- they were owned by people from out of town huh?
AN5	FAC 01	1:38:38	AN : Y en a gros qui aiment ça en même temps < INT : Ouais.> Mais: anyway c'était-- c'était pas cette sorte de musique là moi.
AN6	FAC 01	1:39:49	INT : [parlant de la relation entre son père et son futur mari] Mais mon père était toujours après agacer anyway ça y disait, y a-- y a une fois ça y-- sont après bâtir la grange.
AN7	FAC 01	1:40:12	INT: Pis elle [sa mère] voulait se faire venir au coin chercher de quoi au magasin-- anyway! (.) Mon père se lâchait à crier coudonc y vont penser que tu courroies avec ma fille le soir que tu vas courir avec ma bonnefemme de la journée! [rire]
AN8	FAC 01	1:57:02	INT: [parlant du maintien des toilettes] pis après ça dans l'automn-- uh le printemps y mettaient ça dedans le jardin. AN : C'était de de de l'engrais ça INT : Ben ça-- sais pas c'était de l'engrais mais y enterrait ça anyway
AN9	FAC 08	4:16	INT: nineteen: (.) eighty-five je pense c'était-- non, avan-- après ça. (.) Eighty-five. Maybe it was eighty-five. M'en rappelle pu là. Anyways! < AN : [rire] > uh : (.) j'ai oublié tout ce que je m'en rappelle, tsé!
AN10	FAC 08	1:13:34	INT: of course this was the first time we had a front-wheel drive, see, back-- I think that was in eighty-seven m'en rappelle pu là anyways! (.) so I had told her now this was a totally different driving experience

Seulement environ le quart des occurrences est produit en réponse à une digression de la part de l'interlocuteur. Autrement dit, les résultats de Ferrara (1997) montrent que le marqueur *anyway* est souvent utilisé pour garder ou pour renforcer le contrôle qu'on a sur le discours, mais l'est peu ou pas pour *prendre* le contrôle de l'interaction. À cette analyse s'ajoutent le constat de Park (2010) selon lequel *anyway* peut servir de mécanisme de clôture d'un sujet pour sortir d'une impasse dans l'interaction. Cette impasse peut être créée par une incapacité de continuer sur un sujet donné (par exemple, parce qu'aucun des interlocuteurs n'a d'autres connaissances sur le sujet), par un désaccord entre les locuteurs, par une blague ratée, etc. Dans cet emploi, *anyway* ne relie pas deux segments de discours connexes mais discontinus; il marque plutôt un passage direct entre deux segments disjoints. Park (2010) propose aussi que *anyway* peut être employé pour amorcer une conclusion abrupte avant de passer à autre chose. Cet emploi est semblable à *anyway* dans une fonction de reprise, mais au lieu de relancer et de continuer sur le sujet, il vise à le résumer et le fermer.

Le tableau 5.19 présente les emplois de *anyway* qu'on peut observer dans le corpus d'analyse. On voit qu'il y a une préférence nette pour *anyway* dans une fonction de reprise; on remarque également que les sens dismissif et de reprise (ainsi que le sens additif, qui n'apparaît toutefois pas dans le corpus) sont, comme on pouvait s'y attendre, mutuellement exclusifs, mais que le sens de reprise peut être combiné à la fonction de sortir d'une impasse dans la conversation. Cela confirme l'affirmation de Ferrara (1997) selon laquelle *anyway* en fonction de reprise est essentiellement différent⁸⁴, ainsi que l'observation de Park (2010) selon laquelle *anyway* peut prendre aussi la fonction discursive de sortir d'une impasse conversationnelle.

TABLEAU 5.19 : CATÉGORIES D'ANALYSE OBSERVÉES POUR ANYWAY

# ref.	Locuteur	Reprise	Dismissif	Sortir d'une impasse
AN1	FAC 01 : INT	x		
AN2	FAC 01 : INT	x		
AN3	FAC 01 : AN		x	
AN4	FAC 01 : AN	x		
AN5	FAC 01 : AN	x		
AN6	FAC 01 : INT		x	
AN7	FAC 01 : INT	x		
AN8	FAC 01 : INT		x	
AN9	FAC 08 : INT	x		x
AN10	FAC 08 : INT	x		x

Compte tenu d'une part de l'absence d'occurrences isolables en français qui pourraient constituer un PLM et d'autre part, de l'emploi fréquent de *anyway* dans le français québécois et dans le français de l'Ouest (attesté dans Walker, 2005), il y a lieu de se demander si *anyway* ne serait pas plutôt un emprunt. Il est toutefois impossible de répondre à la positive ou à la négative à cette question sans plus de données sur son statut au sein de la communauté francophone au Maine. On peut cependant affirmer, que cette unité soit intégrée ou non au code français dans l'esprit de ces locuteurs, qu'il y a une préférence nette pour *anyway* par rapport aux concurrents potentiels d'expression qui pourraient constituer un PLM avec lui. Ce constat, appuyé par le fait qu'il y a plusieurs unités du français possibles dans ce PLM, suggère que *anyway* traduit une ou des visées d'expression que ces unités du français ne rendent pas tout à fait.

5.5.4 Now

Le tableau 5.20 présente les occurrences isolables de *now* repérées dans le corpus d'analyse et le tableau 5.21 les occurrences isolables d'*astheure* (*là*), la seule unité lexicale en français qui apparaît en position isolable dans le corpus et qui constituent un PLM avec *now*. Pour faciliter la discussion des occurrences

⁸⁴ Si les *anyway* dismissif et additif sont ou ne sont pas des marqueurs pragmatiques dépend en large partie de la manière dont on définit les marqueurs pragmatiques. Comme ils participent à la structuration et à l'organisation du discours, il n'y a pas lieu de les exclure ici de la catégorie des marqueurs pragmatiques en l'absence d'une analyse plus fine, qui n'est bien entendu pas possible avec 3 occurrences seulement de *anyway* en fonction dismissive et aucune de *anyway* en fonction additive.

TABLEAU 5.20 : LES OCCURRENCES ISOLABLES DE NOW EN CONTEXTE

#	Enreg.	Temps	Occurrence
N1	FAC 01	29 :09	INT : Souvent je disais ça tsé < AN : ouais ouais > in my little — dans mon petit coin de- de : de vie là tsé? (.) AN : Now uh what did you write about?
N2	FAC 01	34 :28	INT : On a été à *** puis après ça on était à ***. Et puis uh : j'ai élevé la famille là. AN : Now uh you're talking—you went to ***, you and your husband.
N3	FAC 01	1 :20 :01	INT : pis dans l'automne on a ramassé des patates. (.) AN : Now uh : (.) when you went to *** là encore dites-moi quel âge vous avez là
N4	FAC 01	1 :20 :46	AN : Ben tout le monde sentait uh : < INT : les animaux > les animaux ce temps-là parce qy'y : (.) y en avait toute hein? INT : Oui. Faudrait ben. AN : Now uh (.) moi je m'en rappelle là uh : me semble qu'à *** là c'est pas—probablement c'était pareil partout, y avait deux sortes de monde deux sortes de familles.
N5	FAC 01	1 :21 :53	INT : on avait toutes des belles petites jupes puis toutes des belles petits outfit. (.) Pour aller à l'école. < AN : you had to be- > On avait pas beaucoup pis on avait deux chaque toujours. (.) AN : Now uh (.) describe uh : un journalier qui c'est que c'était ça?
N6	FAC 01	1:36:37	AN : [parlant d'un prêtre local qui est décédé récemment] Y était-tu en *** de le temps? INT : ouais! AN : Ouais. (.) Now uh (.) uh le *** Club ça ça a été populaire a plein ça.
N7	FAC 01	1 :59 :25	AN : Quessé qu'y faisait avec le restant? INT : D'argent? < AN : Ouais > N'avait pas. < C : N'avait pas. > N'avait pas! Y le donnait cinquante cennes! (.) AN : Now (.) uh — INT : Ça c'est le temps de la Dépression ça.
N8	FAC 03	2 :54	INT1 : [parlant de son père et comment il a fait vivre la famille] Y a travaillé pour uh (.) H*** (.) pis y a travaillé pour F*** (.) pis uh. AN : Now (.) tu veux-tu dire qu'y y chassait là lui c'était pour nourrir sa famille là?
N9	FAC 03	1 :01 :52	INT1 : [montrant une photographie de lui en uniforme militaire] you see? the—the patch here well (.) I was not assigned then. There's something wrong there— AN : attends moi là, montre ça moi INT1 : Now (.) after you get assigned—that's when I first graduated as a pilot
N10	FAC 08	2 :25	AN : Ben on a rentré! (.) dans le barbershop. Pis L*** nous attendait— now . (.) uh : (.) I was thinking about you L*** (.) when I was a kid uh: and there were three barbers in ***

TABLEAU 5.21 : LES OCCURRENCES DE ASTHEURE (LÀ) EN CONTEXTE

#	Enreg.	Temps	Occurrence
AS1	FAC 01	18 :10	INT : y avait un gros grocery store pis y avait du linge dedans ça pis toutes sortes d'affaires de même là. C'est vrai. AN: A—and general store. INT: And general store, yeah. AN: astheure là , avant qu'on: parte de *** là uh: the Ku Klux Klan là uh is that the only time you saw them?
AS2	FAC 01	30 :07	AN : j'ai : interviewé quelqu'un semaine passée à *** she was 88 and (.) elle m'a montré qu'a faisait des tapis pis : she's in < INT : oh ouais c'est beau hein? > she's in good shape y'know? INT: she's in good shape, that's great. AN : astheure là uh (.) uh : vous avez mové à *** when?
AS3	FAC 06	39 :30	AN: t'as gradué de H***-- H*** High School? INT1: yeah, yeah. I went four years there. AN: astheure uh uh a lot of people uh: in those days uh: y—uh didn't g—uh f—finish high school. < INT1: non. >
AS4	FAC 06	1 :03 :12	INT1: Back before E*** had it uh: F*** uh: comment qu'y s'appelle lui? < INT2: I wouldn't know > F*** F*** INT2: I'm sure I heard the name, but I didn't remember. AN: yeah. Astheure là (.) when I was a little kid là, St-Clair là was real big it was uh: popular y avait du monde là

transcrites, chacune comporte un un numéro de référence commençant par la première lettre de l'unité cible qu'elle présente. Un élément remarquable qui ressort de ce tableau est le fait que *là*, qui est produit 91 fois seul et 25 fois comme partie d'un marqueur composite⁸⁵ (cf. la section 5.1), n'a pas été inclus dans ce PLM. De fait, dans l'usage en général, *là* et *now* peuvent avoir certains sens semblables et remplir certaines des mêmes fonctions, notamment quant à la deixis du moment présent. Cependant, il semble que les emplois de *là* et *now* divergent plus qu'ils ne convergent dans l'usage général mais aussi – et surtout – dans le corpus d'analyse. Cette divergence porte notamment sur la fonction de « ponctuant » oral (Vincent, 1993) que *là* remplit beaucoup plus souvent qu'il ne prend un sens strictement déictique. Forget (1989) donne un traitement plus nuancé de cet emploi de *là* très fréquent en français québécois. L'auteure subdivise les emplois de *là* en :

- « *là* d'identification », où le locuteur « sélectionne [...] parmi un ensemble délimité, un élément à propos duquel il fera une prédication » (Forget, 1989, p. 64);
- « *là* de détachement du thème », qui « découpe les unités d'information plus larges correspondant à des propositions, de telle façon qu'elles sont en quelque sorte mises en attente avant la poursuite du message » (Forget, 1989, p. 65);
- et « *là* de renforcement de l'acte », qui « ne semble pas concerner un constituant linguistique en particulier mais plutôt l'acte dans son ensemble » qui a « retenu le choix du locuteur » (Forget, 1989, p. 66), comme dans l'exemple suivant :

Donne-lui ta raquette, là, je vais prendre le tien.

(exemple 39 dans Forget, 1989, p. 66).

Au moyen de ces différents emplois, souligne Forget, *là* accomplit plusieurs fonctions en discours : le découpage énonciatif, afin de « délimiter des unités de discours pour signaler différenciellement leur importance au destinataire, selon leur contribution au plan contextuel » (Forget, 1989, p. 67); la singularisation, c'est-à-dire « la sélection d'un objet, d'un contenu propositionnel, d'un acte pour le distinguer des autres objets, contenus ou actes possibles » (p. 71-72) et le soulignement de la pertinence de l'unité du discours ciblée par *là* et de son rôle dans l'enchaînement textuel (p. 74-78).

Ce sont en effet des emplois et des fonctions qui diffèrent significativement de l'usage qu'on pourrait imaginer de *now*, et de fait, ils sont bien présents dans les emplois de *là* qu'on observe dans le corpus d'analyse :

INT : She sent me some yarn < AN : or some uh clothing- > some yarn and uh she sent me some material to make the b- the spread uh the bedspread là. < AN : oh yeah yeah > The quilts là.

(FAC 01, 1 :17 :18)

⁸⁵ Dont *astheure là* dans le tableau 5.21.

AN : Now that road là, quessé que- quessé que le monde faisait avant qu'y a : que ce chemin-là existe?

(FAC 01, 1 :25 :36)

INT : [*parlant du moment de retour des vacanciers à un lac local*] uh : late September early October là. Nous autres on s'en venait plus de bonne heure pour aller à l'école!

(FAC 08, 13 :05)

Au plan formel également, *là* et *now* sont généralement en distribution complémentaire dans le corpus d'analyse : *now* apparaît toujours en position initiale (en début d'énoncé ou en début d'un segment de discours), tandis que *là* n'apparaît pas dans cette position, mais peut apparaître au milieu d'un segment de discours ou à la fin. Ainsi, les observations suggèrent que *là* et *now* dans le corpus d'analyse ne sont pas tout à fait comparables, et comme *là* est très complexe à analyser en soi, comme le laisse croire la discussion de Forget (1989), il a été exclu comme concurrent d'expression de *now*. Toutefois, comme il a déjà été mentionné à la section 5.1, il n'y a pas de doute que *là* est une ressource associée au français qui peut traduire des visées référentielles qui ne s'expriment pas de la même façon en anglais, d'où sa fréquence élevée en position isolable dans le corpus d'analyse.

Quant à *now*, Fraser (1990 ; 2009) le classe dans une sous-catégorie de marqueurs discursifs qu'il appelle des marqueurs d'attention, avec une fonction de focalisation de l'attention, sans toutefois en discuter plus en détail. Il en est de même pour plusieurs autres auteurs ayant travaillé sur les marqueurs discursifs ou pragmatiques, tels que Redecker (1990), qui classe *now* comme un marqueur de structure conceptuelle sans pousser davantage l'analyse. En fait, quoique l'unité *now* est identifiée dans maintes études du domaine de l'analyse de discours comme un marqueur discursif ou pragmatique par excellence, relativement peu d'entre elles se concentrent sur sa signification ou sur sa fonction. Pour cela, on doit revenir jusqu'à Schiffrin (1987), qui l'analyse comme un marqueur de progression temporelle dans le discours (par exemple, pour présenter les points d'une comparaison), et à Aijmer (1988), qui l'analyse comme un marqueur d'une quelconque disjonction, comme un changement de sujet ou pour l'émission d'un méta-commentaire. Plus récemment, Schourup (2011) avance que *now* indique plus spécifiquement une discontinuité partielle au sein d'un événement plus large conçu comme continu par le locuteur. Poussant son analyse, il suggère que *now* signale à l'interlocuteur de traiter ce qui suit comme quelque chose qui se démarque de par sa nouveauté, sans mettre complètement de côté ce qui précède. Selon Schourup (2011), c'est cette instruction procédurale encodée par *now* qui lui confère des fonctions comme marquer une disjonction, focaliser l'attention, renforcer ou créer de la cohérence discursive, signaler le commencement d'une nouvelle activité ou d'une nouvelle interaction et augmenter la perception de la pertinence de ce qui suit.

Dans le tableau 5.20, on observe effectivement que l'unité *now* est employée dans le corpus d'analyse à des points de disjonction ou de rupture conversationnelle, notamment pour changer abruptement de sujet, parfois vers un sujet qui n'a pas de lien évident avec le sujet précédent. Il est intéressant de noter qu'on observe le même emploi pour *astheure*. Le tableau 5.20 indique également qu'autant *now* que *astheure* sont employées en position isolable par un seul et même locuteur, FAC : AN, à une exception près. Ainsi, il semble ne pas y avoir de différence évidente dans le sens ou la fonction de ces deux unités qui pourrait motiver la sélection de l'une plutôt que l'autre, et on ne peut non plus attribuer la sélection de l'une ou l'autre de ces unités à une différence dans les préférences personnelles des locuteurs ou dans les connaissances lexicales dans l'une ou l'autre langue.

Encore une fois, il semble que le rôle de FAC : AN en tant qu'animateur de la conversation joue pour beaucoup. Non seulement ce locuteur est essentiellement le seul à employer ces unités en position isolable, mais il les emploie presque exclusivement pour prendre la parole après un énoncé de son interlocuteur. Autrement dit, il se sert de ces ressources pour déclarer décisivement la fin de la thématique de discussion en cours et pour contrôler lui-même la prochaine thématique qui sera lancée. Cela semble être fait pour plusieurs raisons. Pour certaines occurrences, comme N3, N6, N7, AS1 et AS2, la discussion de la thématique en cours semble arriver à une fin « naturelle », c'est-à-dire qu'aucun des locuteurs en présence ne semble avoir d'autre chose à dire sur le sujet. Il est probable que, en tant que facilitateur de la conversation, FAC : AN ressent la responsabilité particulière de relancer la conversation avec un nouveau sujet avant que la pause ne gêne l'atmosphère conviviale propice à ce genre d'interaction. Dans d'autres cas, comme N2, N4 ou AS3, l'interviewé(e) semble être ambivalent quant à un changement de sujet, c'est-à-dire qu'il ou elle serait ouvert(e) soit à continuer de parler sur le même sujet, soit de changer de sujet, en fonction de la réaction de FAC : AN. Ces interviewés laissent ainsi la responsabilité du choix entre les mains de FAC : AN, qui choisit finalement, dans les cas considérés ici, de changer de sujet. Enfin, dans certains cas, tels que N1, N8 ou AS4, il semble que FAC : AN impose un changement de sujet afin de ramener la conversation à un sujet qu'il juge plus pertinent ou plus intéressant, alors que ses interlocuteurs auraient continué de parler du sujet en cours.

Les emplois de *now* et de *astheure* que fait FAC : AN sont invariablement accompagnés de pauses et/ou d'hésitations, ce qui pourrait indiquer que ce locuteur est à la recherche d'un nouveau sujet de conversation. Toutefois, la plupart de ces pauses et hésitations sont produites juste après l'unité en question. Autrement dit, FAC : AN a déjà pris la décision de changer de sujet, ce qui suggère qu'il aurait déjà un sujet en tête qu'il souhaite introduire dans la conversation. Il semble plus probable que ces pauses et hésitations servent d'une part à renforcer l'effet de focalisation de l'attention qu'ont ces unités, et d'autre part à intervenir dans la protection des faces en présence. En tant que partenaire de conversation habile et attentif, FAC : AN doit être conscient du fait que changer abruptement le sujet de discussion est une sorte de risque : son interlocuteur

pourrait ne pas accepter ce changement, ou pourrait perdre le fil après une telle rupture. Ainsi, FAC : AN procède-t-il avec un soin particulier, intégrant du temps supplémentaire pour la compréhension à son amorce du nouveau sujet, et ancrant ladite amorce avec plusieurs signaux différents.⁸⁶ De cette façon, il signale à son ou ses interlocuteurs qu'il est conscient de traverser un moment précaire dans l'interaction et qu'il le fait en portant une attention particulière à leurs besoins. De plus, il sélectionne une unité qui, tout en signalant la nouveauté à venir, indique qu'il perçoit cette discontinuité comme ayant lieu au sein d'une continuité plus large, c'est-à-dire l'interaction dans son ensemble. Toutes ces précautions lui permettent de maintenir l'atmosphère collaborative nécessaire au bon déroulement de l'interaction alors même que l'interaction n'est pas collaborative dans les faits, car FAC : AN impose sa propre volonté quant à la direction à prendre.

Pour résumer, FAC : AN emploie tantôt *now*, tantôt *astheure* avec de toute évidence la même fonction, soit pour signaler et pour gérer l'imposition d'une rupture dans la conversation. On ne peut pas affirmer avec certitude pourquoi il sélectionne une unité plutôt que l'autre dans un contexte donné, mais le fait qu'il se sert également d'un ancrage multiple de cette rupture suggère une possibilité : il se peut que l'alternance codique elle-même soit une autre forme d'ancrage de la rupture. Autrement dit, lorsqu'il impose un changement de sujet, non seulement il emploie soit *now*, soit *astheure*, ainsi que des pauses et des hésitations, mais en fait il sélectionne entre *now* et *astheure* de façon à créer un contraste – une rupture – avec la langue employée précédemment. En ce sens, l'alternance codique serait effectivement un signal de contextualisation (« contextualisation cue ») comme l'avance Auer (1995), Alvarez-Cáccamo (1990) et Gumperz (1982), mais dans ce cas l'alternance codique ne semble pas instancier un nouveau cadre interprétatif de l'interaction selon les enjeux socio-politiques associés aux différents codes. Plutôt, l'alternance codique *en soi* semble être une ressource à laquelle les locuteurs bilingues peuvent avoir recours afin de traduire, ou de renforcer, une visée référentielle donnée.

5.6 Discussion générale

L'objectif de la présente étude a été de mettre en œuvre une méthode d'analyse des occurrences d'alternance codique qui puisse tenir compte de la visée référentielle du locuteur comme une explication possible de ces occurrences. Dans l'ensemble, les résultats démontrent que cette méthode est fructueuse. Même avec plusieurs restrictions pratiques, il a été possible d'identifier des tendances dans l'usage de certaines unités faisant l'objet d'une alternance codique, tendances qui peuvent être attribuées à ce que l'unité est employée pour exprimer. L'unité *ouais*, par exemple, lorsqu'elle est insérée dans un segment de discours en anglais, est utilisée par les locuteurs du corpus FAC pour traduire des visées référentielles qui touchent à

⁸⁶ On peut aussi y voir du temps supplémentaire pour trouver comment formuler le lancement du nouveau sujet, ce qui n'enlève rien à l'idée qu'il procède avec un soin particulier afin de protéger les faces en présence lors d'un moment délicat de l'interaction.

l'intersubjectivité et à la cohésion sociale, comme la concordance avec l'interlocuteur. En revanche, *yeah* insérée dans un segment de discours en français semble être moins souvent liée au domaine affectif. *Is that right?* est employée par un locuteur en particulier, FAC : AN, comme stratégie pour souligner une réaction vive et une certaine empathie avec son interlocuteur. Il semble que, pour ce locuteur du moins, la formule *c'est (tu) vrai?* ne traduit pas la même visée référentielle, car il ne l'emploie comme objet de l'alternance codique qu'une seule fois. *So* est de toute évidence utilisée en raison de sa flexibilité et de son degré d'abstraction, car elle apparaît lorsqu'il est nécessaire de traduire des visées référentielles pour lesquelles *ça fait que* ne semble pas être perçue comme adéquate, notamment des visées qui s'éloignent de son sens « canonique » de causalité.

Enfin, en ce qui concerne l'emploi de *now* et *astheure*, qui servent à traduire dans l'ensemble les mêmes visées référentielles, il ressort de l'analyse que l'alternance codique elle-même pourrait être une ressource linguistique. En effet, c'est grâce au fait qu'il connaît deux unités dont l'emploi est relativement équivalent dans l'usage – et qu'il peut présumer du fait que son interlocuteur connaisse également ces deux unités – que le locuteur FAC : AN a la possibilité d'employer celle qui représente un changement de langue, peu importe la langue de l'interaction au moment de la produire. Ainsi, ce locuteur peut exploiter le changement de langue en tant que tel comme une expression de rupture, pour souligner ou ancrer la rupture conversationnelle exprimée par l'unité. À la lumière de ces résultats, la présente étude semble démontrer effectivement que l'alternance codique est au moins parfois employée comme une stratégie d'expression.

Les résultats de la présente étude permettent de plus de soulever plusieurs points de discussion importants. Ils mettent en relief, premièrement, l'importance du rôle que joue le locuteur dans la conversation pour ce qui est de ses productions de l'alternance codique. On le voit avec FAC 08 : INT, qui emploie souvent *yes* et *yeah* lorsqu'il adopte un mode « d'écoute », comme on peut s'imaginer qu'il fait souvent en tant que barbier. On le voit surtout avec FAC : AN, dont le rôle d'animateur ou de facilitateur semble influencer beaucoup sur ses usages : il se sert de l'alternance codique en employant des unités qui ont des fonctions particulières dans la gestion de l'interaction et dans la protection des faces. L'analyse indique que l'unité *is that right?* lui permet de montrer qu'il a une bonne écoute et qu'il s'investit émotionnellement dans ce que son interlocuteur raconte, mettant ainsi l'interlocuteur en confiance en créant une atmosphère conviviale et cordiale, propice à la conversation informelle et plutôt personnelle. *So*, *now* et *astheure* semblent lui servir d'outils de gestion de la conversation, lui permettant d'imposer un nouveau sujet de discussion et ainsi d'exercer sa responsabilité en tant qu'intervieweur, tout en démontrant son attention aux besoins de son interlocuteur et en maintenant le caractère collaboratif de l'interaction.

L'analyse des données du corpus révèle aussi la prévalence d'unités qui sont (ou qui pourraient être interprétées comme) des marqueurs pragmatiques, tels que *yeah*, *is that right?*, *so*, *now* et *anyway*. Cette

prévalence est due en partie à la restriction de l'analyse aux unités en position isolable, car les marqueurs pragmatiques sont plus susceptibles d'apparaître dans de telles positions, en début d'énoncé ou de phrase, en incise, en parenthèse ou en commentaire méta-énonciative, ou encore comme un énoncé ou une phrase complète. On peut néanmoins considérer cette prévalence comme étant significative : environ 80 % des occurrences isolables dans le corpus d'analyse sont des insertions d'une seule unité (cf. la section 4.2.1). Or, la littérature sur l'alternance codique insertionnelle indique que ce sont les noms qui sont les plus souvent insérés (Muysken, 2000 ; Myers-Scotton, 1992b ; Pfaff, 1979 ; Poplack, 1980). Que ce soit plutôt les marqueurs pragmatiques qui prédominent dans le corpus d'analyse est un résultat pas prévu dans la littérature qui suggère qu'une des raisons pour lesquelles les locuteurs bilingues – ou du moins les locuteurs étudiés – produisent de l'alternance codique est pour se prévaloir des unités lexicales qui structurent et gèrent l'interaction, orientent l'interprétation des énoncés et jouent en plus un rôle spécifique dans la gestion des relations interpersonnelles et la protection des faces. Étant donné la nature délicate de ces tâches interactionnelles, il est tout à fait raisonnable de penser que des locuteurs bilingues prioriseraient l'emploi de ressources linguistiques qui traduisent précisément leur visée référentielle plutôt que le maintien de leur choix de langue précédent, lorsqu'ils ont confiance dans le fait qu'une telle alternance codique sera comprise et acceptée par leur(s) interlocuteur(s).

La présente étude soulève aussi la possibilité que l'alternance codique *en soi* peut être une ressource linguistique. Cela se voit le plus clairement dans la comparaison de *now* et *astheure*, dont les contextes d'emploi, les sens en usage et les fonctions discursives sont indifférenciés. Une comparaison avec les emplois de ces unités en position non isolable sera toutefois nécessaire avant de confirmer définitivement cette hypothèse, mais les résultats indiquent que c'est le contraste créé par la juxtaposition des codes que FAC : AN exploite pour attirer l'attention de son interlocuteur et pour ancrer la notion de rupture conversationnelle que ces unités permettent d'exprimer. Cela va en partie dans le sens de Gumperz (1982), Alvarez-Cáccamo (1990) et Auer (1995), selon lesquels l'alternance codique est significative en soi. La conclusion que je propose diffère toutefois de leurs observations en ce que je considère que l'alternance codique peut signifier non pas (ou non seulement) par la mise en évidence des affiliations sociales ou d'une recontextualisation du sujet de l'interaction, mais bien par la mise en évidence d'un contraste entre différents codes, contraste qui a pour effet d'attirer l'attention sur les ressources lexicales spécifiques employées (cf. Backus, 2001).

Globalement, les résultats de la présente étude mettent en relief l'importance de voir dans les locuteurs des agents pleinement impliqués dans l'emploi de leurs ressources linguistiques. En cela, la présente étude abonde dans le sens des études dans une variété de domaines en linguistique, par exemple Podesva (2007) en sociophonétique, ou Van Compernelle et Williams (2012) et LoCastro (2001) en acquisition des langues. Ce rôle d'agent n'implique pas que tout comportement langagier soit un choix conscient ou délibéré, y compris en ce qui concerne la sélection et l'emploi des unités lexicales, mais il contribue à (re)placer le locuteur au centre

de la recherche sur l'alternance codique alors même qu'une grande partie des recherches sur le sujet place l'agentivité des locuteurs en arrière-plan au profit des systèmes grammaticaux, des enjeux sociaux, politiques et interactionnels, ou bien des processus neuro-cognitifs. Bien sûr, tout un éventail de facteurs influe ou peut influencer sur les productions des locuteurs – et c'est justement pour cette raison qu'il ne faut pas négliger l'un d'entre eux, à savoir que les locuteurs peuvent employer leurs ressources linguistiques, lexicales et autres, non pas pour se conformer à un ensemble de règles ou de principes donné, mais comme ils le jugent pour atteindre leurs buts communicatifs.

Rapport-Gratuit.com

6 Conclusion

La présente étude s'est penchée sur la question de l'alternance codique comme une possible stratégie d'expression pour des locuteurs bilingues. Intuitivement, il semble évident qu'un locuteur bilingue pourrait, par exemple, sélectionner l'unité lexicale anglaise *like* dans un contexte donné parce que cette dernière lui permet de traduire plus exactement sa visée référentielle que ce que ne lui permettrait l'unité lexicale française *aimer*, qui traduit un sentiment la force duquel il pourrait ressentir comme étant plus ambiguë. Cette évidence apparente pourrait être l'une des raisons pour lesquelles si peu d'études se concentrent sur le rôle des différences de sens et de la visée référentielle dans la production de l'alternance codique. Toutefois, cet état des choses fait en sorte qu'est négligée une composante importante de ce phénomène complexe, en plus de mettre à l'écart une partie constitutive de l'expérience vécue des locuteurs bilingues, pour qui la question de lexicalisations différentes d'une langue à l'autre, qui permettent de rendre des visées référentielles similaires, mais pas tout à fait identiques, est un défi constant.⁸⁷

La présente étude se veut un point de départ pour combler cette lacune. L'approche adoptée prend appui sur la conception de la langue en tant qu'ensemble socialement défini de ressources linguistiques, c'est-à-dire des connaissances et compétences morphologiques, lexicales, syntaxiques et pragmatiques. Dans cette approche, les locuteurs bilingues sont considérés comme des agents pleinement impliqués dans l'emploi de ces ressources, ce qui leur permet d'atteindre leurs objectifs référentiels, interactionnels et sociaux. Le sens des unités est considéré non pas comme donné mais comme construit dans l'expérience des usages, et la visée référentielle que permet de traduire une unité comprend non seulement sa signification, telle qu'elle est comprise de par les expériences antérieures du locuteur, mais également ses connotations, fonctions discursives, et tout ce que le locuteur entend – ce qu'il est porté à croire que son interlocuteur va comprendre – lorsque l'unité est employée. En conséquence, la méthode d'analyse adoptée dans cette étude cherche à rapprocher, dans la mesure du possible, les critères de sélection, les définitions, les catégories d'analyse et l'intuition de l'analyste des perceptions des locuteurs.

La présente étude comporte toutefois des limites qu'il ne faut pas négliger. La plus importante, incontournable, est que ses objets d'étude principaux – à savoir les facteurs qui influencent la sélection des unités lexicales et le sens dont le locuteur investit les unités qu'il sélectionne – ne sont pas directement observables. On ne peut qu'émettre des hypothèses à partir des indices indirects, comme des tendances observées dans l'usage. C'est pour cette raison qu'un soin particulier a été apporté à ne pas présenter ces

⁸⁷ Entre autres exemples, je peux constater qu'une partie non négligeable du temps de préparation du présent mémoire a été passé à paraphraser en français des concepts lexicalisés ou lexicalisables en anglais, comme *switch* (v.) (« produire de l'alternance codique »), *switching* (« basculement tranché ») et *switches* (« occurrence d'alternance codique »).

objets non observables comme étant des faits avérés, et à tenter d'émettre les hypothèses les plus vraisemblables pour cerner les causes des phénomènes observés.

Une autre limite incontournable de l'analyse est sa subjectivité intrinsèque en raison même des phénomènes à l'étude. De grands efforts ont été déployés pour établir des critères d'analyse rigoureux et transparents en énonçant par exemple des définitions explicites des concepts-clés, en détaillant les critères de sélection et en quantifiant ce qui était quantifiable. Toutefois, il n'en demeure pas moins que toute analyse portant sur le sens des unités linguistiques comporte nécessairement une part de subjectivité. Le sens d'une unité n'existe pas en dehors de l'interprétation qu'un (inter)locuteur lui donne, rapportée à un contexte particulier. Même une mesure d'accord interjuge n'est que la moyenne d'un ensemble de subjectivités – quoiqu'en utilisant cette moyenne, on arrive à quelque chose de nettement moins subjectif dans son ensemble que l'intuition d'un analyste seul. Force est d'admettre qu'aucune étude linguistique, peu importe sa rigueur ou l'accent qu'elle met sur les données empiriques, ne peut être menée sans se fonder essentiellement sur les connaissances, l'intuition et la perception du linguiste. Déjà dans la sélection d'un objet d'étude, c'est la perception de l'analyste quant à ce qui constitue ou non un phénomène d'intérêt pour son domaine qui entre en jeu. Cette perception, ou cette intuition, participe aussi dans l'adoption de l'approche ou du cadre théorique, dans les choix méthodologiques effectués, y compris la sélection des données empiriques à étudier et l'imposition de catégories d'analyse sur le flux (et le flou) de la réalité. L'intuition est partie prenante de l'analyse de ces données et de l'interprétation des résultats. Toutefois, à une époque où l'accès à une vaste quantité de données linguistiques est ouvert – que ces données soient enregistrées, archivées, informatisées ou publiées sur des médias sociaux – le linguiste se doit de remettre constamment en question son intuition et de la confronter aux données disponibles. Mieux vaut, alors, rechercher l'objectivité tout en reconnaissant le rôle que joue l'intuition, plutôt que de prétendre à l'objectivité totale. C'est pour cette raison que l'un des objectifs secondaires de cette étude était de laisser transparaître la subjectivité qui y était inhérente, c'est-à-dire d'explicitier son rôle dans l'analyse, tout en l'ancrant le plus possible dans ce qui est empiriquement observable, par exemple en explicitant les facteurs qui influent sur la perception du statut codique.

Cette étude comporte également d'autres limites d'ordre plus pratique. La plus évidente est la petite taille de l'échantillon, qui a des effets sur les données et les résultats. Par exemple, les catégories fréquentes observées sont sans doute plus influencées qu'elles ne l'auraient été dans un corpus de plus grande taille par les usages et les sujets de discussion que favorisent quelques locuteurs très productifs de l'alternance codique, comme FAC : AN ou FAC 01 : INT. Une autre conséquence de l'utilisation d'un petit échantillon est le nombre limité d'occurrences analysées de chaque unité : l'unité examinée la plus fréquente dans le cas présent est *yeah* avec 190 occurrences. Par conséquent, aucune analyse statistique poussée n'a pu être effectuée et les sens en usage et les fonctions discursives observées pourraient être plus ou moins représentatifs des communautés

étudiées et même des habitudes d'usage réelles des locuteurs étudiés. Il n'a pas été possible non plus, dans la plupart des cas, de procéder à des analyses comparatives entre les habitudes d'usage de différents locuteurs et l'association de ces habitudes à des éléments de leur parcours sémantique. Finalement, la comparaison des contextes d'usage de différents éléments d'un même PLM peut être un outil puissant pour soulever les différences de visée référentielle ayant pu mener à la production de l'alternance codique, mais cet outil n'a pas pu être exploité de la manière espérée en raison du petit nombre d'occurrences d'alternance codique vers le français.

Ces limites sont regrettables, mais il ne faut pas oublier qu'un échantillon de petite taille est propre à l'analyse de l'oral. Un corpus oral demande beaucoup de temps et de main-d'œuvre dans la collecte de données. Il doit ensuite être dépouillé, transcrit, annoté et analysé – chacune de ces étapes étant coûteuse en temps, et aucune ne pouvant être automatisée, du moins pour l'instant. Un petit ensemble de données analysées est également le propre d'une étude exploratoire comme celle-ci. L'absence d'études antérieures sur lesquelles cette recherche aurait pu s'appuyer fait en sorte qu'une grande partie du travail a été consacré à jeter les bases de l'analyse avant même de traiter le corpus : définir les notions théoriques pertinentes (code-langue, alternance codique, bilinguisme, sens en usage, visée référentielle, etc.), établir des critères de sélection des corpus et définir les notions d'occurrence isolable et d'ambiguïté codique de façon à pouvoir les appliquer à un corpus. De plus, puisque cette étude sur corpus était exploratoire, toutes les unités en position isolable ont été traitées comme étant potentiellement pertinentes pour l'analyse. Traiter un plus grand nombre d'occurrences dans un échantillon de taille plus importante aurait été plus facile si un ensemble d'unités avait déjà été identifiées comme étant potentiellement fréquentes ou d'intérêt pour le sujet de la recherche. J'espère toutefois que la présente étude pourra servir de point d'appui pour les recherches à venir, qui pourraient par exemple reprendre ou affiner des définitions établies ici ou porter sur l'une des catégories ou l'une des unités examinées, y compris celles qui n'ont pas pu être analysées en profondeur.

Les restrictions imposées délibérément comme des choix méthodologiques constituent une autre limite de la présente étude. Ces restrictions comprennent notamment le choix de n'analyser que des occurrences isolables, l'exclusion des unités dont le statut codique est ambigu ou potentiellement ambigu (*cf.* la section 4.1.4) et l'exclusion de certaines catégories fréquentes en raison de l'ambiguïté ou de l'impossibilité de trouver des éléments de comparaison parmi les unités lexicales françaises. Ce dernier problème pourrait être résolu par une étude visant des unités spécifiques et/ou sur un plus grand ensemble de données. Les trois premières restrictions, par contre, ont été imposées dans le but d'éviter l'ambiguïté et la confusion dans l'analyse et les résultats. Puisque la présente étude est la première de ce genre, il a été jugé plus avisé de se concentrer sur ce qui pouvait réellement être fait avec les données disponibles, plutôt que de s'attarder à des questions méthodologiques et théoriques épineuses pour lesquelles il n'existe, pour le moment du moins, pas de réponse

possible. Ces choix ont effectivement permis d'atteindre les objectifs établis, soit de démontrer que la méthodologie employée fonctionne et que l'alternance codique peut servir de stratégie d'expression pour les locuteurs bilingues. Je souhaite que les recherches ultérieures puissent trouver des moyens d'éviter l'exclusion des unités comme solution méthodologique. Un premier pas dans cette direction serait par exemple d'étudier les visées référentielles associées à des éléments des PLM identifiés dans la présente étude dans tous les contextes où ils apparaissent – non pas seulement en contexte isolable – afin de voir si les différences observées sont aussi valables dans ces contextes. On pourrait également penser à une enquête qui collecterait, en plus des productions linguistiques, des informations sur les perceptions des locuteurs, leurs attitudes et leurs expériences pertinentes à l'analyse de la visée référentielle associée à des unités de différents codes. En particulier, en ce qui concerne le statut codique des unités, sont nécessaires davantage d'enquêtes lexicologiques sur des communautés de langue minoritaire, qui viseraient non seulement à documenter les anglicismes ou d'autres particularités, mais aussi à esquisser plus largement les habitudes d'usage de la communauté. Ainsi, il serait plus facile de trancher, lors d'une étude sur l'alternance codique, entre les unités d'origine extracodique qui sont diffusées et acceptées dans la communauté de pratique comme des emprunts et les unités d'origine extracodique qui suscitent un sentiment d'altérité codique marquée. Il faudrait cependant s'assurer, pour que ces enquêtes puissent servir de base de comparaison pour les études sur l'alternance codique, que la notion d'emprunt est définie selon les perceptions collectives des locuteurs du code auquel l'unité s'associe en synchronie.

Malgré les limites de l'analyse empirique, la présente étude fournit des apports théoriques et méthodologiques significatifs. Par exemple, elle contribue aux connaissances sur le fait francophone en Amérique du Nord, notamment sur le bilinguisme français-anglais en Alberta et dans le Maine. Elle constitue également une ouverture sur les pratiques langagières de locuteurs bilingues de ces deux régions, et plus généralement de locuteurs bilingues « établis » (qui ne sont pas qualifiés d'« apprenants »). Cette étude présente aussi des concepts théoriques utiles pour la sémantique, à savoir l'application des notions de la visée référentielle et de la sémantisation à un corpus de productions orales effectives. Enfin, elle propose un cadre d'analyse bien réfléchi des visées référentielles traduites au moyen de l'alternance codique qui valorise les locuteurs bilingues comme des agents dans l'emploi de leurs ressources multilingues.

Ainsi, même si la recherche sur l'alternance codique a été et continue d'être un domaine très productif, beaucoup de travail reste à faire, notamment grâce à l'ouverture de nouvelles perspectives sur le phénomène. Comme le souligne Kootstra et al. (2009) : « if we regard an individual speaker as 'someone in whom all sorts of influences on language use are expressed' (De Bot 1992: 2), it is even impossible to separate intra- and interindividual aspects of code-switching or to distinguish grammatical, social and cognitive processes in code-switching. A more comprehensive account of code-switching is therefore required in which both individual and

social variables are represented, and in which the interplay of these variables is explicitly specified » (p. 130). J'ajouterais aux aspects énumérés par ces auteurs la dimension de l'expression sémantique, et aux variables considérées les pratiques singulières de sémantisation selon les expériences et les habitudes d'usage. En mettant de l'avant l'absence relative de ces variables dans la littérature jusqu'à ce jour, et en proposant une nouvelle approche et de nouvelles méthodes pour en entreprendre l'analyse, la présente étude se veut être un premier pas vers l'intégration de la visée référentielle et des pratiques de sémantisation à l'ensemble des connaissances déjà établies sur l'alternance codique.

Il convient de souligner pour clore cette conclusion que la présente étude s'inscrit dans une lignée d'études en plein essor, dans laquelle on cherche à situer et à comprendre l'activité langagière des locuteurs bilingues et multilingues dans un monde de plus en plus globalisé, où le contact des langues et le multilinguisme sont plutôt la règle que l'exception (voir par exemple Bhatia et Ritchie, 2013 ; Blommaert, 2010 ; Nagy et Meyerhoff, 2008 ; Phillipson et Skutnabb-Kangas, 2013). Cette perspective a plusieurs implications pour la recherche, notamment, le fait de rendre nécessaire l'utilisation d'approches et de cadres théoriques pouvant rendre compte de la mobilité sans précédent entre régions géographiques et entre groupes sociaux, ainsi que des technologies de communication qui influent sur le développement et la diffusion des habitudes de pratique. Il faut, autrement dit, dépasser (ou au moins modifier) l'approche labovienne traditionnelle qui voit les groupes sociaux et les classes sociales comme étant relativement fixes, et qui postule également une certaine homogénéité et stabilité sur le plan géographique. Cette conception était plus justifiée à l'époque où Labov a élaboré sa théorie,⁸⁸ alors qu'il y avait moins de mobilité sociale et géographique (et que l'étude de la variation en linguistique était encore à ses débuts), mais elle n'a de nos jours pas la flexibilité nécessaire pour rendre compte de la singularité des parcours des locuteurs étudiés (Blommaert, 2010 ; Eckert, 2012). Ce changement implique aussi le développement de théories et de modèles du fonctionnement du langage, que ce soit sur le plan syntaxique, sémantique, cognitif, ou autre, qui peuvent rendre compte du bilinguisme et du multilinguisme non pas comme des cas d'exception nécessitant des explications et des adaptations ou des mécanismes supplémentaires, mais comme l'état ordinaire des choses. Ce travail est déjà bien amorcé – voir par exemple Ellis (2006 ; 2008), Otheguy *et al.* (2015), Hall, Cheng et Carlson (2006), Kroll *et al.* (2015) et Lowie et Verspoor (2011). Cependant, pour pouvoir développer sur ces bases, il faut un plus vaste ensemble de corpus dans lesquels on puisse observer les habitudes de pratique de locuteurs bilingues et multilingues; ces corpus doivent être récents, accessibles, et collectés dans des contextes de communication qui se rapprochent du quotidien des locuteurs. Somme toute, les comportements langagiers des locuteurs bilingues ne devraient pas être relégués à de simples outils pour confronter des théories sur l'acquisition des langues ou sur le

⁸⁸ Mais voir Silverstein (2015), qui rappelle que la diversité et même la « superdiversité » ne sont pas des phénomènes récents ou exclusifs à l'époque actuelle.

fonctionnement syntaxique ou cognitif du langage; ils devraient constituer un objet d'étude en soi, dont fait partie l'alternance codique. L'alternance codique – c'est-à dire, l'emploi de ressources associées à différentes codes reconnues comme étant des langues – doit être comprise et étudiée en tant que partie constitutive des habitudes de pratique, des compétences communicatives, et ultimement, des expériences vécues des locuteurs bilingues.

Références

- Abrahamsson, Niclas, et Kenneth Hyltenstam. 2009. « Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny ». *Language Learning*. Vol. 59, n°2, p. 249-306.
- Abutalebi, Jubin, et David W. Green. 2008. « Control mechanisms in bilingual language production: Neural evidence from language switching studies ». *Language and Cognitive Processes*. Vol. 23, n°4, p. 557-582.
- Aijmer, Karin. 1988. « “Now may we have a word on this”: The use of “now” as a discourse particle ». Dans : Kytö, M., Ihaleinen, O., Rissanen, M. (éds). *Corpus linguistics, hard and soft. The eighth International Conference on English Language Research on Computerized Corpora*. Amsterdam, Pays-Bas : Editions Rodopi B.V. p. 15-34.
- Albert, Jacob, Tony Brinkley, Yvon Labbé, et Christian Potholm. 2013. « Contemporary attitudes of Maine’s Franco-Americans ». *Franco-American Center Occasional Papers*. n°1, p. 1-124.
- Allaire, Gratien. 2014. « La francophonie de l’Ouest: pérennité, diversité et rapport à l’Autre ». Dans : Papen, R.A., Hallion, S. (éds). *À l’Ouest des Grands Lacs: Communautés francophones et variétés de français dans les Prairies et en Colombie-Britannique*. Quebec, Canada : Presses de l’Université Laval. p. 21-67.
- Allen, James P. 1974. « Franco-Americans in Maine: A geographical perspective ». *Acadiensis*. Vol. 4, n°1, p. 32-66.
- Alvarez-Cáccamo, Celso. 1998. « From “switching code” to “code-switching”: Towards a reconceptualization of communicative codes ». Dans : Auer, P. (éd.). *Code-switching in conversation: Language, interaction and identity*. Londres, R.-U. : Routledge. p. 29-48.
- . 1990. « Rethinking conversational code-switching: Codes, speech varieties, and contextualisations ». Dans : Hall, K., Koenig, J.-P., Meacham, M., Reinman, S., Sutton, L. (éds). *16th annual meeting of the Berkely Linguistics Society. General session and parasession on the legacy of Grice*. Ann Arbor, MI : Braun-Brumfield. p. 3-16.
- Anderson, John A. E., Lorinda Mak, Ahram Keyvani Chahi, et Ellen Bialystok. 2018. « The language and social background questionnaire: Assessing degree of bilingualism in a diverse population ». *Behavior Research Methods*. Vol. 50, n°1, p. 250-263.
- Angermeyer, Philipp Sebastian. 2003. « Lexical cohesion as a motivation for code-switching ». Dans : Sayahi, L. (éd.). *Selected proceedings of the first Workshop on Spanish Sociolinguistics. The first Workshop on Spanish Sociolinguistics*. Somerville, MA : Cascadilla Proceedings Project. p. 112-122.
- Anon. 2016. *attorney n.1* [En ligne]. *Oxford English Dictionary online*. Disponible sur : < <http://www.oed.com/view/Entry/12890> > (Consulté le 23 juin 2016).
- Auer, Peter. 1995. « The pragmatics of code-switching: A sequential approach ». Dans : Milroy, L., Muysken, P. (éds). *One speaker, two languages: Cross-disciplinary perspectives on code-switching*. Cambridge, R.-U. : Cambridge University Press. p. 115-135.

- . 2000. « Why should we and how can we determine the “base language” of a bilingual conversation? » *Estudios de Sociolingüística*. Vol. 1, n°1, p. 129-144.
- Backus, Ad. 2009. « Codeswitching as one piece of the puzzle of language change: The case of Turkish *yapnak* ». Dans : Ludmila, I., Winford, D., De Bot, K. (éds). *Multidisciplinary approaches to code switching*. Amsterdam, Pays-Bas : John Benjamins. p. 307-336.
- . 2001. « The role of semantic specificity in insertional codeswitching: Evidence from Dutch - Turkish ». Dans : Jacobson, R. (éd.). *Codeswitching worldwide*. Berlin, Allemagne : Mouton de Gruyter. p. 125-154.
- Baer, Doug, James Curtis, Edward Grabb, et William Johnston. 1995. « Respect for authority in Canada, the United States, Great Britain and Australia ». *Sociological Focus*. Vol. 28, n°2, p. 177-195.
- Baer, Doug, Edward Grabb, et William T. Johnston. 1990. « The values of Canadians and Americans: A critical analysis and reassessment ». *Social Forces*. Vol. 6, n°3, p. 693-713.
- Banuazizi, Atissa, et Cassandre Creswell. 1999. « Is that a real question?: Final rises, final falls and discourse function in yes-no question intonation ». Dans : Billings, S.J., Boyle, J.P., Griffith, A.M. (éds). *Papers from the 35th annual meeting of the Chicago Linguistics Society: The main session. The 35th annual meeting of the Chicago Linguistics Society*. Chicago, IL : The Chicago Linguistics Society. p. 1-14.
- Beaulieu, Suzie. 2012. [Thèse]. *Towards a sociolinguistically informed pedagogy: French for L2 nursing students in Alberta*. Edmonton, Canada : University of Alberta.
- Beaulieu, Suzie, Johnson Bresnick, Marie Duchemin, Samuel Gagnon, Stéphanie Levesque, et Leila Ranta. 2016. « Promoting sociolinguistic competence in French L2 textbooks: A case of missed opportunities? » Communication présentée à *The 35th annual Second Language Research Forum*, New York, NY.
- Beaulieu, Suzie, et Alexis Dupont Rochette. 2014. « Interventions pédagogiques visant le développement de la compétence sociolinguistique: Une exploration ». Dans : Bergeron, R., Dumais, C., Harvey, B., Nolin, R. (éds). *La didactique du français oral du primaire à l'université*. Côte St-Luc, Canada : Peisaj. p. 127-142.
- Beniak, Edouard, Stephen Carey, et Raymond Mougeon. 1984. « A sociolinguistic and ethnographic approach to Albertan French and its implications for French-as-a-first-language pedagogy ». *Canadian Modern Language Review*. Vol. 41, n°2, p. 308-314.
- Bhatia, Tej K., et William C. Ritchie. 2013. « Bilingualism and multilingualism in the global media and advertising ». Dans : Bhatia, T.K., Ritchie, W.C. (éds). *The handbook of bilingualism and multilingualism*. Chichester, R.-U. : Wiley-Blackwell. p. 565-597.
- Blanche-Benveniste, Claire. 2010. *Approches de la langue parlée en français*. Nouvelle édition 2010. Paris : Éditions Ophrys.
- Blom, Jan-Petter, et John Gumperz. 2000 (1972). « Social meaning in linguistic structures: Code-switching in Norway ». Dans : Li, W. (éd.). *The bilingualism reader*. Londres, R.-U. : Routledge. p. 111-136.
- Blommaert, Jan. 2007. « Sociolinguistic scales ». *Intercultural Pragmatics*. Vol. 4, n°1, p. 1-19.

- . 2010. *The sociolinguistics of globalization*. Cambridge, R.-U. : Cambridge University Press.
- Bolden, Galina G. 2006. « Little words that matter: Discourse markers “so” and “oh” and the doing of other-attentiveness in social interaction ». *Journal of Communication*. Vol. 56, p. 661-688.
- de Bot, Kees. 1992. « A bilingual production model: Levelt’s “Speaking” model adapted ». *Applied Linguistics*. Vol. 13, n°1, p. 1-24.
- . 2016. « Multi-competence and dynamic/complex systems ». Dans : Cook, V., Li, W. (éds). *The Cambridge handbook of linguistic multi-competence*. Cambridge, R.-U. : Cambridge University Press. p. 125-142.
- Bouchard, Chantal. 1999. *On n’emprunte qu’aux riches: La valeur sociolinguistique et symbolique des emprunts*. Montréal, Canada : Éditions Fides.
- Boulanger, Jean-Claude, et Anna Malkowska. 2008. « Itinéraires croisés des emprunts en alimentation: “Les années *Petit Robert*” ». Dans : Pruvost, J. (éd.). *Dictionnaires et mots voyageurs. Les 40 ans du Petit Robert, de Paul Robert à Alain Rey. Les Journées des dictionnaires de Cergy*. Paris, France : Éditions des Silves. p. 103-136.
- Boztepe, Erman. 2003. « Issues in code-switching: Competing theories and models ». *Working Papers in TESOL & Applied Linguistics*. Vol. 3, n°2, p. 1-27.
- Branca-Rosoff, Sonia. 1998. « Le mot comme notion hétérogène: Linguistique, histoire, discours ». Dans : Branca-Rosoff, S. (éd.). *Le mot: Analyse de discours et sciences sociales*. Aix-en-Provence, France : Presses universitaires de Provence. p. 7-39.
- Brennan, Susan, et Michael Schober. 2001. « How listeners compensate for disfluencies in spontaneous speech ». *Journal of Memory and Language*. Vol. 44, p. 274-296.
- Brennan, Susan, et Maurice Williams. 1995. « The feeling of another’s knowing: Prosody and filled pauses as cues to listeners about the metacognitive state of speakers ». *Journal of Memory and Language*. Vol. 34, p. 383-398.
- Brown, Penelope. 2005. « What does it mean to learn the meaning of words? » *Journal of the Learning Sciences*. Vol. 14, n°2, p. 293-300.
- Canagarajah, Suresh. 2011. « Codemeshing in academic writing: Identifying teachable strategies of translanguaging ». *The Modern Language Journal*. Vol. 95, n°3, p. 401-417.
- Clark, Eve. 1997. « Conceptual perspective and lexical choice in acquisition ». *Cognition*. Vol. 64, p. 1-37.
- Clark, Herbert, et Jean Fox Tree. 2002. « Using *uh* and *um* in spontaneous speaking ». *Cognition*. Vol. 84, p. 73-111.
- Clyne, Michael. 2000 (1987). « Constraints on code-switching: How universal are they? » Dans : Li, W. (éd.). *The bilingualism reader*. Londres : Routledge. p. 257-280.
- Conseil d’Europe. 2005. *Cadre européen commun de référence pour les langues: apprendre, enseigner, évaluer*. Paris, France : Didier.

- Cook, Vivian. 1992. « Evidence for multicompetence ». *Language Learning*. Vol. 42, n°4, p. 557-591.
- Cormier, Yves. 2009. *Dictionnaire du français acadien*. Montréal, Canada : Éditions Fides.
- Courbon, Bruno. 2012. « Intégration syntagmatique du sens lexical et établissement de rapports synonymiques ». Dans : Berlan, F., Berthomieu, G. (éds). *La synonymie*. Paris, France : Presses de l'université Paris-Sorbonne. p. 329-341.
- . 2015. « Pratiques sémantiques et différences interindividuelles à l'ère des corpus informatisés ». *Les Cahiers de Lexicologie*. Vol. 106, n°1, p. 91-126.
- Courbon, Bruno, et Myriam Paquet-Gauthier. 2014. « Faux amis / vrais ennemis : Réutilisations de la notion d'anglicisme dans le discours métalinguistique au Québec ». *Le discours et la langue*. Vol. 6, n°1, p. 143-173.
- Daveluy, Michelle. 2008. « Language, mobility, and (in)security: A journey through Francophone Canada ». Dans : Meyerhoff, M., Nagy, N (éds). *Social lives in language: Sociolinguistics and multilingual speech communities*. Amsterdam, Pays-Bas : John Benjamins. p. 27-42.
- Dewaele, Jean-Marc. 2010. *Emotions in multiple languages*. Basingstoke, R.-U. : Palgrave MacMillan.
- Dostie, Gaétane. 2004. *Pragmaticalisation et marqueurs discursifs: Analyse sémantique et traitement lexicographique*. Bruxelles, Belge : De Boeck Duculot.
- Drummond, Kent, et Robert Hopper. 1993. « Some uses of yeah ». *Research on Language and Social Interaction*. Vol. 26, n°2, p. 203-212.
- Duffley, Patrick J. 1992. *The English infinitive*. Londres, R.-U. : Longman.
- Durand, Jacques, Bernard Laks, et Chantal Lyche. 2002. « La phonologie du français contemporain : Usages, variétés et structure ». Dans : Pusch, C., Raible, W. (éds). *Romanistische Korpuslinguistik: Korpora und gesprochene Sprache*. Tübingen, Allemagne : Gunter Narr Verlag. p. 93-106.
- . 2009. « Le projet PFC : Une source de données primaires structurées ». Dans : Durand, J., Laks, B., Lyche, C. (éds). *Phonologie, variation et accents du français*. Paris, France : Hermès. p. 19-61.
- Eckert, Penelope. 2006. « Communities of practice ». Dans : *Encyclopedia of language and linguistics*. Boston, MA : Elsevier. p. 683-685.
- . 2012. « Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation ». *Annual Review of Anthropology*. Vol. 41, p. 87-100.
- Edwards, John. 2013. « Bilingualism and multilingualism: Some central concepts ». Dans : Bhatia, T.K., Ritchie, W.C. (éds). *The handbook of bilingualism and multilingualism*. Malden, MA : Wiley-Blackwell. p. 5-25.
- Ellis, Elizabeth. 2008. « Defining and investigating monolingualism ». *Sociolinguistic Studies*. Vol. 2, n°3, p. 311-330.
- . 2006. « Monolingualism: The unmarked case ». *Estudios de Sociolingüística*. Vol. 7, n°2, p. 173-196.

- Evans, Vyvyan. 2009. *How words mean: Lexical concepts, cognitive models, and meaning construction*. Oxford, R.-U. : Oxford University Press.
- Ferrara, Kathleen W. 1997. « Form and function of the discourse marker *anyway*: Implications for discourse analysis ». *Linguistics*. Vol. 35, p. 343-378.
- Fishman, Joshua. 2000 (1965). « Who speaks what language to whom and when? » Dans : Li, W. (éd.). *The Bilingualism Reader*. London, R.-U. : Routledge. p. 89-106.
- Flege, Jim E. 2007. « Language contact in bilingualism: Phonetic system interaction ». Dans : Cole, J., Hualde, J.I. (éds). *Laboratory phonology 9*. Berlin, Allemagne : Mouton de Gruyter. p. 353-382.
- Forget, Danielle. 1989. « Là: un marqueur de pertinence discursive ». *La pragmatique*. Vol. 18, n°1, p. 57-82.
- Fox, Cynthia A. 2007. « Franco-American voices: French in the northeastern United States today ». *The French Review*. Vol. 80, n°6, p. 1278-1292.
- Fox Tree, Jean E. 2010. « Discourse markers across speakers and settings ». *Language and Linguistics Compass*. Vol. 4/5, p. 269-281.
- Franceschini, Rita. 1998. « Code-switching and the notion of code in linguistics: Proposals for a dual focus model ». Dans : Auer, P. (éd.). *Code-switching in conversation: Language, interaction and identity*. Londres, R.-U. : Routledge. p. 51-72.
- Fraser, Bruce. 2009. « An account of discourse markers ». *International Review of Pragmatics*. Vol. 1, p. 293-320.
- . 1990. « An approach to discourse markers ». *Journal of Pragmatics*. Vol. 14, p. 383-395.
- . 1999. « What are discourse markers? » *Journal of Pragmatics*. Vol. 31, p. 931-952.
- Fraundorf, Scott, et Duane Watson. 2011. « The disfluent discourse: Effects of filled pauses on recall ». *Journal of Memory and Language*. Vol. 65, p. 161-175.
- Fuller, Janet. 2003. « The influence of speaker roles on discourse marker use ». *Journal of Pragmatics*. Vol. 3, p. 23-45.
- Gafaranga, Joseph, et Maria-Carme Torras. 2002. « Interactional otherness: Towards a redefinition of codeswitching ». *International Journal of Bilingualism*. Vol. 6, n°1, p. 1-22.
- Gardner-Chloros, Penelope. 2009. *Code-switching*. Cambridge, R.-U. : Cambridge University Press.
- . 1995. « Code-switching in community, regional and national repertoires: The myth of the discreteness of linguistic systems ». Dans : Milroy, L., Muiyksen, P. (éds). *One speaker, two languages: Cross-disciplinary perspectives on code-switching*. Cambridge, R.-U. : Cambridge University Press. p. 68-89.
- Gardner-Chloros, Penelope, et Malcolm Edwards. 2004. « Assumptions behind the grammatical approach to code-switching: When the blueprint is a red herring ». *Transactions of the Philological Society*. Vol. 102, n°1, p. 103-129.

- Granger, Sylviane. 2008. « Learner corpora ». Dans : Lüdeling, A., Kytö, M. (éds). *Corpus linguistics: An international handbook*. Berlin, Allemagne : Mouton de Gruyter. p. 259-275.
- Gravano, Agustin, Stefan Benus, Julia Hirschberg, Shira Mitchell, et Ilia Vovsha. 2007. « Classifications of discourse functions of affirmative words in spoken dialogue ». Dans : *INTERSPEECH 2007* [En ligne]. *8th annual conference of the International Speech Communication Association*. Antwerp, Belge : ISCA. p. 1613-1616. Disponible sur : < http://www.isca-speech.org/archive/interspeech_2007/i07_1613.html > (Consulté le 27 juillet 2017).
- Grosjean, François. 2010. *Bilingual: Life and reality*. Boston, MA : Harvard University Press.
- . 2013. « Bilingualism: A short introduction ». Dans : Grosjean, F., Ping, L. (éds). *The psycholinguistics of bilingualism*. Malden, MA : Wiley-Blackwell. p. 5-25.
- . 1997. « The bilingual individual ». *Interpreting*. Vol. 2, n°1-2, p. 163-187.
- Gruntman, Carleen. 2011. [Thèse]. *Towards the meaning of for: a corpus analysis*. Quebec, Canada : Université Laval.
- Gumperz, John. 1982. *Discourse strategies*. Cambridge, R.-U. : Cambridge University Press.
- Hakuta, Kenji. 1986. *Mirror of language*. New York, NY : Basic Books.
- Hall, Joan Kelly, An Cheng, et Matthew T. Carlson. 2006. « Reconceptualizing multicompetence as a theory of language knowledge ». *Applied Linguistics*. Vol. 27, n°2, p. 220-240.
- Hamers, Josiane F., et Michel H. A. Blanc. 2000. *Bilinguality and bilingualism*. 2é éd. Cambridge, R.-U. : Cambridge University Press.
- Haugen, Einar. 1956. *Bilingualism in the Americas: A bibliography and research guide*. University, AL : University of Alabama Press.
- . 1972 (1970). « The stigmata of bilingualism ». Dans : Dil, A. (éd.). *The ecology of language: Essays by Einar Haugen*. Stanford, CA : Stanford University Press. p. 307-324.
- Heller, Monica, éd. 2007. *Bilingualism: A social approach*. Basingstoke, R.-U. : Palgrave MacMillan.
- . 1995. « Code-switching and the politics of language ». Dans : Milroy, L., Muysken, P. (éds). *One speaker, two languages: Cross-disciplinary perspectives on code-switching*. Cambridge, R.-U. : Cambridge University Press. p. 158-174.
- . 2011. *Paths to post-nationalism: A critical ethnography of language and identity*. Oxford, R.-U. : Oxford University Press.
- . 1992. « The politics of codeswitching and language choice ». *Journal of Multilingual and Multicultural Development*. Vol. 13, n°1-2, p. 123-142.
- Heredia, Roberto R., et Jeanette Altarriba. 2001. « Bilingual language mixing: Why do bilinguals code-switch? » *Current Directions in Psychological Science*. Vol. 10, n°5, p. 164-168.

- Hirtle, Walter. 2007. *Lessons on the English verb: No expression without representation*. Montréal, Canada : McGill-Queen's University Press.
- Hlavac, Jim. 2011. « Hesitation and monitoring phenomena in bilingual speech: A consequence of code-switching or a strategy to facilitate its incorporation? » *Journal of Pragmatics*. Vol. 43, p. 3793-3806.
- Hockett, Charles F. 1987. *Refurbishing our foundations: Elementary linguistics from an advanced point of view*. Amsterdam, Pays-Bas : John Benjamins.
- Holmes, Janet, et Tina Chiles. 2010. « "Is that right?": Questions and questioning as control devices in the workplace ». Dans : Freed, A.F., Ehrlich, S. (éds). « *Why do you ask? »: The function of questions in institutional discourse*. New York, NY : Oxford University Press.
- Holmes, Janet, et Myriam Meyerhoff. 1999. « The community of practice: Theories and methodologies in language and gender research ». *Language in Society*. Vol. 28, p. 173-183.
- House, Juliane. 2013. « Developing pragmatic competence in English as a Lingua Franca: Using discourse markers to express (inter)subjectivity and connectivity ». *Journal of Pragmatics*. Vol. 59A, p. 57-67.
- Humbley, John. 1974. « Vers une typologie de l'emprunt linguistique ». *Les Cahiers de Lexicologie*. Vol. 25, n°2, p. 46-70.
- Husum, Russ, et R.C. Bryce. 1991. « A survey of graduates from a Saskatchewan French immersion high school ». *The Canadian Modern Language Review*. Vol. 48, n°1, p. 135-143.
- Jakobson, Roman, Gunnar Fant, et Morris Halle. 1963. *Preliminaries to speech analysis*. Cambridge, MA : M.I.T. Press.
- Jarvis, Scott, et Aneta Pavlenko. 2008. *Crosslinguistic influence in language and cognition*. New York, NY : Routledge.
- Kemp, Charlotte. 2009. « Defining multilingualism ». Dans : Aronin, L., Hufeisen, B. (éds). *The exploration of multilingualism: Development of research on L3, multilingualism, and multiple language acquisition*. p. 11-26.
- Kerbrat-Orecchioni, Catherine. 2009. « La place de l'interprétation en Analyse du Discours en Interaction ». *Signes, discours & sociétés*. n°3, p. 1-13.
- . 2007. « L'analyse du discours en interaction: Quelques principes méthodologiques ». *Limbaje si comunicare*. Vol. 9, p. 13-32.
- . 2005. *Le discours en interaction*. Paris, France : Armand Colin.
- King, Ruth, et Terry Nadasdi. 1999. « The expression of evidentiality in French-English bilingual discourse ». *Language in Society*. Vol. 28, p. 355-365.
- Kleiber, Georges. 1996. « Noms propres et noms communs: un problème de dénomination ». *La dénomination*. Vol. 414, n°4, p. 567-589.
- . 2004. « Peut-on sauver un sens de dénomination pour les noms propres? » *Functions of Language*. Vol. 11, n°1, p. 115-145.

- . 2007. « Sur le rôle cognitif des noms propres ». *Cahiers de Lexicologie*. Vol. 91, p. 151-165.
- Koch, Peter, et Wulf Oesterreicher. 2001. « Langage parlé et langage écrit ». Dans : Holtus, G., Metzeltin, M., Schmitt, C. (éds). *Lexikon der Romanistischen Linguistik*. Tübingen, Allemagne : Niemeyer. p. 584-627.
- Koivisto, Aino. 2013. « On the preference for remembering: Acknowledging an answer with Finnish *ai nii(n)* ("Oh that's right") ». *Research on Language and Social Interaction*. Vol. 46, n°3, p. 277-297.
- Kootstra, Gerrit J., Janet G. van Hell, et Tom Dijkstra. 2009. « Two speakers, one dialogue: An interactive alignment perspective on code-switching in bilingual speakers ». Dans : Isurin, L., Winford, D., de Bot, K. (éds). *Multidisciplinary approaches to code-switching*. Amsterdam, Pays-Bas : John Benjamins. p. 129-159.
- Kroll, Judith F., Susan C. Bobb, et Zofia Wodniecka. 2006. « Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech ». *Bilingualism: Language and Cognition*. Vol. 9, n°2, p. 119-135.
- Kroll, Judith F., Paola E. Dussias, Kinsey Bice, et Lauren Perrotti. 2015. « Bilingualism, mind, and brain ». *Annual Review of Linguistics*. Vol. 1, p. 377-394.
- Lamarre, Patricia. 2002. « Multilingual Montreal: Listening in on the language practices of young Montrealers ». *Canadian Ethnic Studies*. Vol. 34, n°3, p. 47-75.
- Lambertz, Kathrin. 2011. « Back-channeling: The use of *yeah* and *mm* to portray engaged listenership ». *Griffith Working Papers in Pragmatics and Intercultural Communication*. Vol. 4, n°1/2, p. 11-18.
- Langacker, Ronald W. 2008. *Cognitive Grammar: A Basic Introduction*. New York, NY : Oxford University Press.
- van Langendonck, Willy. 2007. *Theory and typology of proper names*. Berlin, Allemagne : Mouton de Gruyter.
- Lave, Jean, et Étienne Wenger. 1991. *Situated learning: Legitimate peripheral participation*. Cambridge, R.-U. : Cambridge University Press.
- Le Page, Robert, et Andrée Tabouret-Keller. 1985. *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge, R.-U. : Cambridge University Press.
- Leclercq, Pascale. 2009. « The influence of L1 French on near-native French learners of English: The case of simultaneity ». Dans : Labeau, E., Myles, F. (éds). *The advanced learner variety: The case of French*. Oxford, R.-U. : Peter Lang. p. 269-290.
- Lewis, M. Paul, Gary F. Simons, et Charles D. Fennig, éds. 2015. « The problem of language identification ». Dans : *Ethnologue: Languages of the world, 18th edition (online)* [En ligne]. Disponible sur : < <http://www.ethnologue.com/about/problem-language-identification> > (Consulté le 30 août 2015).
- Li, Wei. 2000. « Dimensions of bilingualism ». Dans : Li, W. (éd.). *The bilingualism reader*. Londres, R.-U. : Routledge. p. 1-21.
- . 2011. « Moment analysis and translanguaging space: Discursive construction of identities by multilingual Chinese youth in Britain ». *Journal of Pragmatics*. Vol. 43, p. 1222-1235.

- . 1998. « The “why” and “how” questions in the analysis of conversational code-switching ». Dans : Auer, P. (éd.). *Code-switching in conversation: Language, interaction and identity*. Londres, R.-U. : Routledge. p. 156-176.
- Liebscher, Grit, et Jennifer Dailey-O’Cain. 2005. « Learner code-switching in the content-based foreign language classroom ». *The Modern Language Journal*. Vol. 89, n°2, p. 234-257.
- Lipset, Seymore Martin. 1990. *Continental divide: The values and institutions of the United States and Canada*. New York, NY : Routledge.
- LoCastro, Virginia. 2001. « Individual differences in second language acquisition: Attitudes, learner subjectivity, and L2 pragmatic norms ». *System*. Vol. 29, p. 69-89.
- Lowie, Wander, et Marjolijn Verspoor. 2011. « The dynamics of multilingualism: Levelt’s speaking model revisited ». Dans : Schmid, M.S., Lowie, W. (éds). *Modelling bilingualism: From structure to chaos*. Amsterdam, Pays-Bas : John Benjamins. p. 267-288.
- Lüdi, Georges. 1987. « Les marques transcodiques: Regards nouveaux sur le bilinguisme ». Dans : Lüdi, G. (éd.). *Devenir bilingue – Parler bilingue*. Tübingen, Allemagne : Niemeyer. p. 1-21.
- Lüdi, Georges, et Bernard Py. 2009. « To be or not to be...a plurilingual speaker ». *International Journal of Multilingualism*. Vol. 6, n°2, p. 154-167.
- MacSwan, Jeff. 2005. « Codeswitching and generative grammar: A critique of the MLF model and some remarks on “modified minimalism” ». *Bilingualism: Language and Cognition*. Vol. 8, n°1, p. 1-22.
- . 2013. « Code-switching and grammatical theory ». Dans : Bhatia, T.K., Ritchie, W. (éds). *The handbook of bilingualism and multilingualism*. Chichester, R.-U. : Blackwell-Wiley. p. 323-350.
- Malt, Barbara C., et Eef Ameel. 2011. « The art and science of bilingual object naming ». Dans : Pavlenko, A. (éd.). *Thinking and speaking in two languages*. Bristol, R.-U. : Multilingual Matters. p. 170-197.
- McCrae, Robert R., et Antonio Terracciano. 2006. « National character and personality ». *Current Directions in Psychological Science*. Vol. 15, n°4, p. 156-161.
- Milroy, James. 2001. « Language ideologies and the consequences of standardization ». *Journal of Sociolinguistics*. Vol. 5, n°4, p. 530-555.
- Mougeon, Raymond. 1998. « French outside New Brunswick and Quebec ». Dans : Edwards, J. (éd.). *Language in Canada*. Cambridge, R.-U. : Cambridge University Press. p. 226-251.
- Mougeon, Raymond, et Edouard Beniak. 1996. « Social class and language variation in bilingual speech communities ». Dans : Guy, G.R., Feagin, C., Schiffrin, D., Baugh, J. (éds). *Towards a social science of language: Papers in honor of William Labov*. Amsterdam, Pays-Bas : John Benjamins. p. 69-99.
- Mougeon, Raymond, Terry Nadasdi, et Katherine Rehner. 2009. « Évolution de l’usage des conjonctions et locutions de conséquence par les adolescents franco-ontariens de Hawkesbury et de Pembroke (1978-2005) ». Dans : Martineau, F., Mougeon, R., Nadasdi, T., Tremblay, M. (éds). *Le français d’ici: Études linguistiques et sociolinguistiques sur la variation du français au Québec et en Ontario*. Toronto, Canada : Éditions du Gref. p. 145-184.

- Muyksen, Pieter. 1995. « Code-switching and grammatical theory ». Dans : Milroy, L., Muyksen, P. (éds). *One speaker, two languages: Cross-disciplinary perspectives on code-switching*. Cambridge, R.-U. : Cambridge University Press. p. 177-198.
- Muysken, Pieter. 2000. *Bilingual speech*. Cambridge, R.-U. : Cambridge University Press.
- Myers-Scotton, Carol. 1995. « A lexically based model of code-switching ». Dans : Milroy, L., Muyksen, P. (éds). *One speaker, two languages: Cross-disciplinary perspectives on code-switching*. Cambridge, R.-U. : Cambridge University Press. p. 233-256.
- . 1992a. « Comparing codeswitching and borrowing ». *Journal of Multilingual and Multicultural Development*. Vol. 13, n°1-2, p. 19-39.
- . 1992b. « Constructing the frame in intrasentential codeswitching ». *Multilingua*. Vol. 11, n°1, p. 101-127.
- Myers-Scotton, Carol, et Janice L. Jake. 1995. « Matching lemmas in a bilingual language production and comprehension model: Evidence from intrasentential code switching ». *Linguistics*. Vol. 33, p. 981-1024.
- Nagy, Naomi, et Myriam Meyerhoff. 2008. « Introduction: Social lives in language ». Dans : Meyerhoff, M., Nagy, N. (éds). *Social lives in language: Sociolinguistics and multilingual speech communities*. Amsterdam, Pays-Bas : John Benjamins. p. 1-15.
- Otheguy, Ricardo, Ofelia García, et Wallis Reid. 2015. « Clarifying translanguaging and deconstructing named languages: A perspective from linguistics ». *Applied Linguistics Review*. Vol. 6, n°3, p. 281-307.
- Paquet-Gauthier, Myriam. 2015. [Mémoire]. *Sens influencés de l'anglais en français au Québec: Utilisation, perception et intégration*. Québec, Canada : Université Laval.
- Paquet-Gauthier, Myriam, et Suzie Beaulieu. 2015. « Can language classrooms take the multilingual turn? » *Journal of Multilingual and Multicultural Development*. Vol. 37, n°2, p. 167-183.
- Parent, François. 2015. [Thèse]. *La définition du terme « mot » en grammaire française contemporaine : Une approche par la sémantique lexico-grammaticale*. Québec, Canada : Université Laval.
- Park, Innhwa. 2010. « Marking an impasse: The use of *anyway* as a sequence-closing device ». *Journal of Pragmatics*. Vol. 42, p. 3283-3299.
- Pavlenko, Aneta. 2011a. « Introduction: Bilingualism and thought in the 20th century ». Dans : Pavlenko, A. (éd.). *Thinking and speaking in two languages*. Bristol, R.-U. : Multilingual Matters. p. 1-28.
- , éd. 2011b. *Thinking and speaking in two languages*. Bristol, R.-U. : Multilingual Matters.
- Le Petit Robert de la langue française*. 2016. [En ligne, version 4] Dictionnaires Le Robert. Disponible sur : < <http://pr.bvdep.com/robert.asp> > (Consulté le 27 janvier 2017).
- Pfaff, Carol W. 1979. « Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English ». *Language*. Vol. 55, n°2, p. 291-318.
- Pflanz, Marie-Laure. 2014. « Emprunt lexical: Existe-t-il une typologie de la phase néologique? » *Neologica*. Vol. 8, p. 157-183.

- Phillipson, Robert, et Tove Skutnabb-Kangas. 2013. « Linguistic imperialism and endangered languages ». Dans : Bhatia, T.K., Ritchie, W.C. (éds). *The handbook of bilingualism and multilingualism*. Chichester, R.-U. : Wiley-Blackwell. p. 495-516.
- Pinette, Susan. 2002. « Franco-American studies in the footsteps of Robert G. LeBlanc ». *Québec Studies*. Vol. 33, p. 9-14.
- Podesva, Robert J. 2007. « Phonation type as a stylistic variable: The use of falsetto in constructing a persona ». *Journal of Sociolinguistics*. Vol. 11, n°4, p. 478-504.
- Poplack, Shana. 1987. « Contrasting patterns of code-switching in two communities ». Dans : Wande, E., Anward, J., Nordberg, B., Steensland, L., Thelander, M. (éds). *Aspects of multilingualism. Fourth Nordic Symposium on Bilingualism*. Uppsala, Suède : Borgströms, Motala. p. 51-77.
- . 1980. « Sometimes I start a sentence in Spanish y termino en español: Towards a typology of code-switching ». *Linguistics*. Vol. 18, p. 581-618.
- . 1989. « The care and handling of a mega-corpus: The Ottawa-Hull French project ». Dans : Fasold, R.W., Schiffrin, D. (éds). *Language change and variation*. Amsterdam, Pays-Bas : John Benjamins. p. 411-451.
- . 2012. « What does the Nonce Borrowing Hypothesis hypothesize? » *Bilingualism: Language and Cognition*. Vol. 15, n°3, p. 644-648.
- Poplack, Shana, et Marjory Meechan. 1998. « Introduction: How languages fit together in codemixing ». *International Journal of Bilingualism*. Vol. 2, n°2, p. 127-138.
- . 1995. « Patterns of language mixture: Nominal structure in Wolof-French and Fonbge-French bilingual discourse ». Dans : Milroy, L., Muysken, P. (éds). *One speaker, two languages: Cross-disciplinary perspectives on code-switching*. Cambridge, R.-U. : Cambridge University Press.
- Poplack, Shana, David Sankoff, et Christopher Miller. 1988. « The social correlates and linguistic processes of lexical borrowing and assimilation ». *Linguistics*. Vol. 28, p. 47-104.
- Rabatel, Alain. 2001. « Valeurs énonciative et représentative des "présentatifs" c'est, il y a voici/voilà: effet point de vue et argumentativité indirecte du récit ». *Revue de Sémantique et Pragmatique*. Vol. 9, p. 111-144.
- Raney, Tracey. 2010. « Quintessentially un-American? Comparing public opinion on national identity in English speaking Canada and the United States ». *International Journal of Canadian Studies*. Vol. 42, p. 105-123.
- Ravin, Yael, et Claudia Leacock. 2000. « Polysemy: an overview ». Dans : Ravin, Y., Leacock, C. (éds). *Polysemy: Theoretical and computational approaches*. Oxford, R.-U. : Oxford University Press. p. 1-29.
- Récanati, François. 1997. « La polysémie contre le fixisme ». *Langue française*. Vol. 113, p. 107-123.
- Redeker, Gisela. 1990. « Ideational and pragmatic markers of discourse structure ». *Journal of Pragmatics*. Vol. 14, p. 367-381.

- Richard, Mark Paul. 1998. « From Franco-American to American: The case of Sainte-Famille, an assimilating parish of Lewiston, Maine ». *Social History*. Vol. 31, n°61, p. 71-93.
- . 2009. « “This is not a Catholic country”: The Ku Klux Klan confronts Franco-Americans in Maine ». *New England Quarterly*. Vol. 82, n°2, p. 285-303.
- Ruhl, Charles. 1989. *On monosemy: A study in linguistic semantics*. Albany, NY : State University of New York Press.
- Sablaylorles, Jean-François. 2003. « Le sentiment néologique ». Dans : Sablaylorles, J.-F. (éd.). *L’innovation lexicale*. Paris : Honoré Champion. p. 279-295.
- Sankoff, David, et Shana Poplack. 1981. « A formal grammar for code-switching ». *Papers in Linguistics*. Vol. 14, n°1, p. 3-45.
- Schiffrin, Deborah. 1987. *Discourse markers*. Cambridge, R.-U. : Cambridge University Press.
- Schmidt, Ulrike. 2008. *Language loss and the ethnic identity of minorities* [En ligne]. European Centre for Minority Issues. Disponible sur : < http://www.ecmi.de/uploads/tx_lfpubdb/brief_18.pdf > (Consulté le 5 février 2017).
- Schourup, Lawrence. 2011. « The discourse marker *now*: a relevance-theoretic approach ». *Journal of Pragmatics*. Vol. 43, p. 2110-2129.
- Silverstein, Michael. 2015. « How language communities intersect: Is “superdiversity” an incremental or transformative condition? » *Language and Communication*. Vol. 44, p. 7-18.
- Spelke, Elizabeth S., et Sanna Tsivkin. 2001. « Language and number: A bilingual training study ». *Cognition*. Vol. 78, p. 45-88.
- St-Pierre, Adèle. 2011. [Thèse]. *Le français parlé à Jay-Livermore Falls (Maine, É-U) : Étude lexicale*. Quebec, Canada : Université Laval.
- The Task Force on Franco-Americans. 2012. *Final report of the Task Force on Franco-Americans*. State of Maine, 125th legislature. Disponible sur < <http://www.maine.gov/legis/opla/francoamerican.htm> >.
- Thomas, David R. 2006. « A general inductive approach for analyzing qualitative evaluation data ». *American Journal of Evaluation*. Vol. 27, n°2, p. 237-246.
- Tiersma, Peter. 2008. « The nature of legal language ». Dans : Gibbons J, Turell MT (éd.). *Dimensions of forensic linguistics*. Amsterdam : John Benjamins. p. 7-25.
- Toribio, Almeida Jacqueline. 2001. « On the emergence of bilingual code-switching competence ». *Bilingualism: Language and Cognition*. Vol. 4, n°3, p. 203-231.
- Tracy, Rosemarie. 2000. « Language mixing as a challenge for linguistics ». Dans : Döpke, S. (éd.). *Cross-linguistic structures in simultaneous bilingualism*. Amsterdam, Pays-Bas : John Benjamins. p. 11-36.
- Treffers-Daller, Jeanine. 2009. « Code-switching and transfer: An exploration of similarities and differences ». Dans : Bullock, B., Toribio, A.J. (éds). *The Cambridge handbook of linguistic code-switching*. Cambridge, R.-U. : Cambridge University Press. p. 58-74.

- Turpin, Danielle. 1998. « Le français, c'est le *last frontier*: The status of English-origin nouns in Acadian French ». *International Journal of Bilingualism*. Vol. 2, n°2, p. 221-233.
- United States Census Bureau. 2015. *2011-2015 American Community Survey 5-year estimates : Age by language spoken at home for the population 5 years and over* [En ligne]. Disponible sur : < factfinder.census.gov > (Consulté le 31 mars 2017).
- . 2016. « Index of questions ». Dans : *Census.gov* [En ligne]. Disponible sur : < https://www.census.gov/history/www/through_the_decades/index_of_questions/ > (Consulté le 26 février 2017).
- Usito. 2014. [En ligne] Cajolet-Laganière, H., Martel, P. et Masson, C.-E., dir. Les Éditions Delisme. Disponible sur : < <https://www.usito.com/dictio> > (Consulté le 27 janvier 2017).
- Van Compernelle, Rémi A., et Lawrence Williams. 2012. « Reconceptualizing sociolinguistic competence as mediated action: Identity, meaning-making, agency ». *The Modern Language Journal*. Vol. 96, n°2, p. 234-250.
- Vanderveen, Tamara Nicole. 2015. [Thèse]. *Life after French immersion: Expectations, motivations, and outcomes of secondary school French immersion programs in the Greater Toronto Area* [En ligne]. Toronto, Canada : York University. 268 p. Disponible sur : < <https://yorkspace.library.yorku.ca/xmlui/handle/10315/30105> > (Consulté le 22 mai 2017).
- Vaxelaire, Jean-Louis. 2009. « Lexicologie du nom propre et onomastique ». *Nouvelle revue d'onomastique*. Vol. 51, n°1, p. 301-315.
- Vincent, Diane. 1993. *Les ponctuels de la langue et autres mots du discours*. Quebec, Canada : Nuit Blanche Éditeur.
- Vogh, Kendall, et Bruno Courbon. à paraître. « "For me, it was very genetic": usage de termes issus du domaine de la génétique dans le discours en ligne de diabétiques anglophones ». Dans : *Actes de la conférence TOTH 2016. Terminologie et Ontologie: Théories et Applications*. Chambéry, France : Institut Porphyre.
- Walker, Douglas. 2012. « Albertan French phonology: French in an anglophone context ». Dans : Gess, R., Lyche, C., Meisenburg, T. (éds). *Phonological variation in French: Illustrations from three continents*. Amsterdam, Pays-Bas : John Benjamins. p. 341-368.
- . 2005. « Le français dans l'Ouest canadien ». Dans : Valdman, A., Auger, J., Piston-Hatlen, D. (éds). *Le français en Amérique du Nord: État présent*. Quebec, Canada : Presses de l'Université Laval. p. 187-205.
- Weinreich, Uriel. 1970 (1953). *Languages in contact: Findings and problems*. La Haye, Pays-Bas : Mouton.
- Wenger, Étienne. 1998. *Communities of practice: Learning, meaning, and identity*. Cambridge, R.-U. : Cambridge University Press.
- Wesche, Marjorie Bingham. 1993. « French immersion graduates at university and beyond: What difference has it made? » Dans : Alatis, J.E. (éd.). *Georgetown University round table on languages and linguistics 1992: Language, communication, and social meaning*. Washington, D.C. : Georgetown University Press. p. 208-240.

Wiley, Terrence G. 2014. « Diversity, superdiversity, and monolingual language ideology in the United States: Tolerance or intolerance? » *Review of Research in Education*. Vol. 38, p. 1-32.

Wong, Jean. 2000. « The token “yeah” in nonnative speaker English conversation ». *Research on Language and Social Interaction*. Vol. 33, n°1, p. 39-67.

Wray, Alison. 2002. *Formulaic language and the lexicon*. Cambridge, R.-U. : Cambridge University Press.

Rapport-Gratuit.com