

La fonction d'introduction référentielle et les moyens linguistiques dont les langues disposent pour la réaliser

Dans ce chapitre, nous introduisons notre objet de recherche et présentons les concepts théoriques pertinents pour notre analyse. Comme nous l'avons esquissé dans l'introduction, dans notre thèse nous nous intéressons aux structures linguistiques particulières qui interviennent pour ancrer un référent lorsqu'il fait son entrée dans l'univers du discours. Il s'agira donc de décrire cette fonction présentative et son lien avec le statut cognitif propre à l'entité référentielle que l'on veut justement « introduire » dans le récit.

Cela nous amènera à discuter en premier lieu des différents facteurs qui interviennent pour définir la nouveauté d'un référent : nous nous intéresserons aux catégories cognitives de l'*identifiabilité* et de l'*accessibilité* (1.1.1.1), ainsi qu'au concept de *cadre cognitif* en tant que cas particulier d'*accessibilité référentielle* (1.1.1.2) et à la notion de *proéminence thématique* qui détermine, en revanche, le statut référentiel que l'entité nouvellement introduite aura dans le discours futur (1.1.1.3).

Dans la section 1.1.2 nous décrivons comment ces considérations de caractère cognitif se reflètent formellement dans la langue : nous nous penchons sur la corrélation imparfaite entre *identifiabilité* et *détermination grammaticale* (1.1.2.1) pour discuter ensuite du phénomène de l'*anaphore associative* (1.1.2.2), étroitement lié au concept de *cadre sémantique* évoqué plus haut. Nous terminons la section avec un survol des formes utilisées en chinois pour marquer les différents degrés d'*identifiabilité* et d'*accessibilité* (1.1.2.3).

Enfin, une thèse portant sur les énoncés présentatifs ne pouvait pas se passer du concept de *théticité*, très influent dans la littérature concernée : nous traitons de l'opposition entre proposition thétique et proposition catégorique (1.1.3.1) et d'un phénomène associé aux énoncés de nature thétique, à savoir la *détopicalisation du sujet* (1.1.3.2), qui implique l'association au sujet de marques normalement réservées à l'objet, à des fins pragmatiques ; enfin, en 1.1.3.3 nous discutons de la distinction entre les énoncés thétiques centrés sur l'événement et ceux centrés sur l'entité.

La section 1.2 dans son ensemble est consacrée à la présentation des moyens formels disponibles à travers les langues pour réaliser la fonction présentative. Après des considérations de

caractère plus général, typologique, nous parlerons tout particulièrement des énoncés à sujet postverbal (1.2.1), tout d'abord en présentant les analyses formulées dans le cadre de l'hypothèse inaccusative, en montrant les avantages d'y intégrer une approche pragmatique (1.2.1.1). Nous justifierons ensuite la distinction entre l'inversion absolue et l'inversion locative (1.2.1.2), ce qui nous portera à discuter de la notion de topique scénique (1.2.1.3). Enfin, nous nous examinerons plus en détail les fonctions que l'ordre V-S peut réaliser dans le discours (1.2.1.4). La section 1.2.2 sera consacrée quant à elle aux structures présentatives biclausales, où nous introduisons d'abord la présence de l'opérateur existentiel AVOIR à travers les langues et les facteurs motivant le choix de cet opérateur (1.2.2.1) ; ensuite nous ferons une distinction entre les structures existentielles et les structures présentatives (1.2.2.2) pour terminer avec la construction relative présentative en français, qui comporte différents introducteurs (1.2.2.3).

Dans la section 1.3 nous discutons de la configuration que les structures présentatives sont censées éviter, mais qui existe tout de même – sous certaines conditions et plus ou moins fréquemment selon la langue – à savoir les configurations sujet-verbe (S-V) comportant un sujet préverbal indéfini.

Enfin, ce chapitre se conclut par un résumé des points principaux mis en évidence, ainsi que par la formulation de notre première question de recherche (section 1.4).

1.1. La fonction d'introduction référentielle

Comme indiqué en introduction à ce travail, notre thèse porte sur les événements linguistiques dans lesquels le locuteur introduit en discours un référent nouveau aux yeux de son interlocuteur. L'introduction référentielle intéresse le linguiste dans la mesure où les langues mobilisent des formes particulières pour réaliser cette fonction, dont les structures présentatives biclausales et l'ordre V-S auxquelles est consacré ce travail.

Avant de discuter des procédés utilisés dans les langues du monde pour réaliser la fonction d'introduction référentielle et de leurs caractéristiques (section 1.2), faisons un pas en arrière pour définir tout d'abord ce qu'est un référent nouveau, et voir les facteurs qui peuvent intervenir pour accroître ou bien nuancer cette « nouveauté ».

1.1.1. Le statut cognitif des référents

Certes, l'introduction référentielle – la réalisation linguistique de la première occurrence d'un référent dans le discours – est étroitement liée à l'expression de référents nouveaux. Mais qu'est-ce qu'un référent « nouveau » ? On pourrait définir, pour utiliser des mots simples, le référent nouveau par excellence comme un référent qui n'a pas été mentionné auparavant dans le discours et qui, en même temps, est entièrement inconnu aux yeux de l'interlocuteur. Ce référent sera encore plus « nouveau » s'il ne bénéficie pas d'un contexte pouvant aider à le repérer.

A partir de cette première esquisse de définition, on comprend que le niveau de nouveauté est variable, et que le statut d'une entité référentielle dans l'esprit des interlocuteurs peut être influencé par plusieurs facteurs. Tout d'abord, il convient de distinguer deux catégories cognitives, à savoir l'indentifiabilité et l'accessibilité.

1.1.1.1. Les catégories cognitives de l'indentifiabilité et de l'accessibilité

Dans la conversation, les interlocuteurs sont engagés dans une évaluation constante des suppositions opérées par leur partenaire (Clark et Marshall 1981), suppositions qui changent en continu au fur et à mesure que l'échange communicatif avance. Comme nous l'avons mentionné plus haut, au cours du processus communicatif, le locuteur influence l'image mentale du monde de son interlocuteur, car l'information se transmet au travers d'une manipulation des représentations linguistiques créées dans l'esprit des interlocuteurs durant l'échange communicatif (Clark et Wilkes-Gibbs 1986, Lambrecht 1994 : 37). Voir Clark et Wilkes-Gibbs (1986) sur la nature interactive de l'établissement des représentations linguistiques :

« The idea behind the view of reference we are taking is this: A and B must mutually accept that B has understood A's references before they let the conversation go on. Conversations proceed in an orderly way only if the common ground of the participants accumulates in an orderly way [...]. A and B must therefore establish the mutual belief that B has understood, or appears to have understood, A's current utterance before they go on to the next contribution to the conversation. They establish that belief, we argue, through an acceptance process. The two basic elements in this process are (a) a presentation and (b) an acceptance » (Clark et Wilkes-Gibbs 1986 : 9).

Les participants à la communication disposent d'une représentation mentale des référents extralinguistiques, représentation qui évolue à chaque moment de la communication. Or, le statut dans lequel un référent donné se trouve dans l'esprit des interlocuteurs est lié principalement à deux catégories cognitives : l'indentifiabilité et l'accessibilité (Lambrecht 1994 : 76).

Tout d'abord, l'identifiabilité réfère à l'estimation portée par le locuteur sur la présence ou l'absence de représentation d'un référent (ou d'un « état des choses ») dans l'esprit de son interlocuteur, il s'agit donc d'une catégorie liée à la connaissance. Un référent est dit *identifiable* quand sa représentation est partagée par le locuteur et l'interlocuteur, tandis qu'un référent est *non identifiable* lorsque sa représentation n'existe que dans l'esprit du sujet parlant (Lambrecht 1994 : 78). Si ce dernier souhaite introduire dans le discours une entité qu'il juge ne pas être encore présente dans l'esprit de l'interlocuteur, et qui n'est pas non plus rendue accessible par la situation externe au dialogue, il devra alors créer *ex nihilo* une représentation au moyen d'une description linguistique (Lambrecht 1994 : 77).

En outre, un référent non identifiable peut être *ancré* (« anchored », terme proposé par Prince 1981) lorsqu'il est lié à une autre entité référentielle du discours, typiquement au moyen d'un deuxième syntagme nominal : l'effort demandé à l'interlocuteur pour repérer l'entité référentielle désignée par le constituant nominal indéfini est bien différent selon que l'on dit *un professeur* ou *un professeur de l'Inalco*. Bien que dans les deux cas il s'agisse d'un référent non identifiable aux yeux de l'interlocuteur (ce dont témoigne dans ce cas le choix de l'article indéfini *un*), le fait de rattacher l'entité (« un professeur ») à un ensemble connu (« l'Inalco ») restreint considérablement le champ sémantico-cognitif de référence et facilite ainsi la tâche de repérage.

D'autre part, la notion d'accessibilité, en tant que phénomène cognitif, renvoie au fait que, à côté de la propriété d'être identifiable ou non, un référent donné reflète aussi différents « états d'activation » dans l'esprit des locuteurs. Un référent *activé* est au centre de l'attention à un moment donné, tandis qu'un référent *accessible* n'est pas au centre de l'attention au moment de l'énonciation, mais reste facilement repérable. A l'inverse, un référent *inactivé* se trouve uniquement dans la mémoire à long terme des sujets engagés dans l'échange communicatif (Lambrecht 1994 : 94, v. aussi la notion de *unused topic* chez Prince 1981). Ainsi, « une fois qu'une entité a revêtu le statut cognitif de référent identifiable, elle le demeure tout au long du discours, mais elle peut aussi avoir un statut plus élevé dans la hiérarchie » (Gundel *et al.* 2000). Lorsque le degré d'activation d'un référent est si fort qu'il ne nécessite pas une expression référentielle l'exprimant (anaphore zéro), on peut parler de « topique ratifié » (*ratified topic*, Lambrecht 2000b : 614).

En attribuant à un référent un certain degré d'accessibilité, fort, moyen ou faible, le locuteur réalise un jugement de type cognitif qui s'accompagne d'un choix de formes linguistiques pour le désigner. Ainsi, des opérations à la fois cognitives (hypothèse en rapport à la présence d'un référent donné dans l'esprit de l'interlocuteur) et linguistiques (sélection d'une expression référentielle adéquate) sont en jeu.

1.1.1.2. L'accessibilité : la notion de « cadre cognitif »

Un référent peut être accessible à un moment donné de l'échange linguistique dans trois situations principales. Comme nous l'avons vu, une entité est textuellement accessible lorsque suite à une sorte de « désactivation » d'un état précédent elle n'est plus active mais existe tout de même dans le monde du discours. Un référent peut également être déictiquement accessible grâce à sa présence dans le monde externe au dialogue, en référence donc à la situation communicative.

Enfin, une entité référentielle peut être accessible en vertu de son interconnexion avec un ensemble d'autres éléments. Elle peut s'insérer dans un schéma cognitif ou « cadre » (en anglais *frame*, au sens de Fillmore [1985 : 223] : « unified frameworks of knowledge, or coherent schematizations of experience »¹⁶) et être ainsi potentiellement connectée avec un autre élément de l'univers du discours plus facilement repérable. On dira que le référent dans ce cas est *accessible par inférence* (Lambrecht 1994 : 100) ; il s'agit des référents *inferrables* chez Prince (1981) : ces référents supposés inferrables de la part de son interlocuteur par déduction¹⁷.

L'idée à la base des « frames », notion empruntée à la psychologie cognitive, est que notre compréhension et perception du monde s'organise autour de schémas prototypiques et que pour certains événements ou situations, nous avons déjà une idée de toute une série de caractéristiques annexes :

« Cognitive frames are usually expressed as “slot-filler representations”, structures of interconnected roles together with constraints on the possible or likely fillers of those roles [...]. Examples of such frames are (1) the partially ordered set of events, as well as the participants in such events, that one can expect in a typical visit to a restaurant, barbershop, or hospital, (2) stages and processes in the life cycle of a human being, (3) the visual and physical properties of a cube, and (4) the organization of a human face, and countless others. » (Fillmore et Baker 2009 : 314).

On peut citer ici l'exemple du BIRTHDAY PARTY *frame*, qui fut introduit dans les travaux sur l'intelligence artificielle par Minsky (1975). Considérons l'exemple suivant :

(I.2) *Mary was invited to Jack's party. She wondered if he would like a kite.*

'Mary était invitée à la fête de Jack. Elle se demandait s'il aimerait un cerf-volant.'

¹⁶ Fillmore (1982) définit ainsi le concept de *cadre* : « by the term “frame” I have in mind any system of concepts related in such a way that to understand any of them you have to understand the whole structure in which it fits; when one of the things in such a structure is introduced into a text, or into a conversation, all of the others are automatically made available » (1982 : 111). Cf. aussi les notions de *cognitive domain* chez Langacker (1987 : 147) et d'*Idealized Cognitive Model* proposée par Lakoff (1987 : 68).

¹⁷ Lambrecht (1994 : 107) estime que cette dernière catégorie cognitive n'a pas nécessairement de corrélation directe dans la morphologie ou dans la phonologie de la langue, mais qu'elle peut avoir des manifestations au niveau de la syntaxe.

(Minsky 1986 : 261)¹⁸

Rien dans cet exemple n'évoque directement l'idée d'une « fête d'anniversaire » : c'est la contribution du sujet récepteur qui parvient à une telle association d'idées. Ce type d'évènement contient prototypiquement des « rubriques » (appelées *slots* en anglais) pour certains participants attendus, accessoires impliqués et sous-événements : *invités, roi de la fête, cadeaux, gâteau, bougies, chanson d'anniversaire* et ainsi de suite. De manière cruciale, c'est à partir des différentes rubriques évoquées dans l'exemple que l'interlocuteur déduit le cadre cognitif FETE D'ANNIVERSAIRE.

Bien sûr, ce processus d'association peut opérer dans le sens inverse. Regardons l'exemple présenté en (1.3). En ayant évoqué le cadre cognitif MARIAGE, les participants à l'échange linguistique peuvent y connecter automatiquement ses différentes rubriques. Ainsi, l'introduction du référent « mariée » peut se faire directement en tant que SN défini (*the bride*), en vertu de son appartenance au cadre MARIAGE. Nous verrons qu'une telle formulation des nouveaux référents en tant que noms grammaticalement définis relève d'un phénomène appelé « pontage inférentiel » (en anglais *bridging* ; voir la section 1.1.2.2).

(I.3) *We went to a wedding last Sunday. **The bride** wore red.*

'On a été à un mariage dimanche dernier. La mariée était habillée en rouge.'

(Poesio 1994 : 14, nt.)

S'il s'agissait d'un référent non identifiable aux yeux de l'interlocuteur, une introduction dans le monde du discours en tant que nom défini serait autrement inappropriée. Comparons avec l'exemple (I.4).

(I.4) *On a passé un bon dimanche. #?La mariée était habillée en rouge.*

L'emploi d'un déterminant défini est inadéquat dans ce cas, puisque l'élément « mariée » n'est pas connecté de manière stéréotypique à DIMANCHE, et n'est donc pas évoqué en tant qu'une de ses rubriques¹⁹.

Lambrecht (1994 : 109) résume ainsi les différents états dans lesquels un référent peut se trouver dans un moment donné du discours :

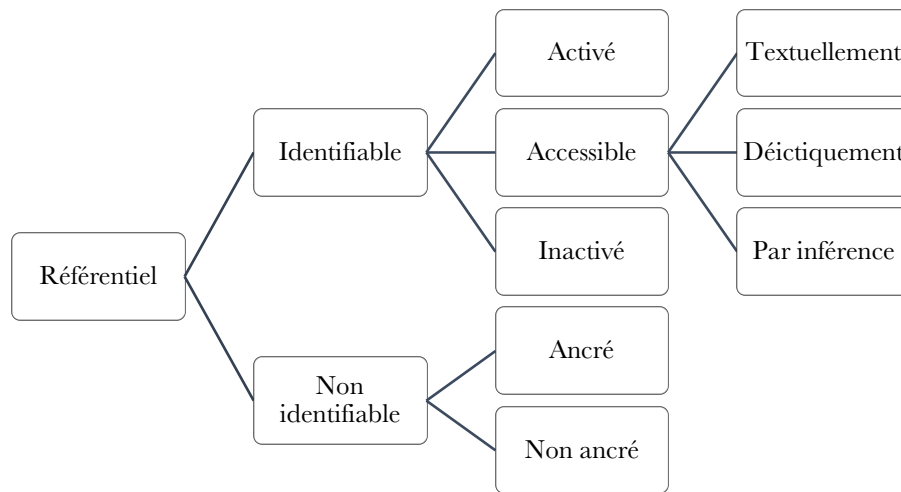
¹⁸ Cet exemple est discuté également dans Minsky (2006 : 179).

¹⁹ Peut-être, de manière quelque peu italienne, la rubrique « lasagnes » serait un bon candidat pour faire l'objet d'un pontage inférentiel dans ce contexte (puisque le DIMANCHE est typiquement associé au repas en famille, dont les lasagnes sont le plat typique, etc.) :

(i) *On a passé un bon dimanche. **Les lasagnes** étaient excellentes.*

Cet exemple nous montre aussi – en dehors de son aspect ludique – que la connexion entre un schéma et ses rubriques est en partie dépendante de la dimension culturelle du langage. Ainsi, dans le cadre de notre expérience de linguistique appliquée, nous verrons que si le « boulanger » est évoqué par le schéma de la BOULANGERIE en français, cette connexion est moins forte en chinois (chapitre VI § 6.2.2).

Figure I-1. Les différents états d'activation dans lesquels un référent donné peut se trouver selon les paramètres de l'identifiabilité et de l'accessibilité (adapté de Lambrecht 1994 : 109)



A cette classification s'ajoutent les référents qui sont accessibles en raison de leur référence unique. Un référent peut être accessible également sur la base du contexte plus large, non relatif à la situation immédiate (ex. *le président des États-Unis*) et à partir de nos connaissances générales sur le monde. Un exemple typique de ce dernier cas est « le soleil » ou « la lune », où on emploie l'article défini pour désigner un objet réputé unique.

(I.5) *It's dark. **The moon** has disappeared.*

(I.6) *I have to go. **The train** is leaving any minute.*²⁰ (Erteschik-Shir 2007 : 18)

Comme le remarque Denizot (2007), cette unicité est en quelque sorte définitoire, car « *un soleil* ou *ce soleil* sont des expressions parfaitement dicibles si l'on parle de l'univers, dans un récit de science-fiction ou dans un article de physique, par exemple ». Dans ce cas, le référent réputé unique perd sa valeur d'unicité.

Erteschik-Shir (2007 : 18) considère les référents en (I.5-6) comme ces composants du monde toujours disponibles pour être des topiques sur la base du savoir partagé entre les interlocuteurs :

²⁰ On peut toutefois considérer qu'il y a une différence entre la définitude de *la lune* et *le train*. Dans le premier cas, nous avons un référent réputé unique ; dans le deuxième l'unicité est faible et provisoire : il s'agit d'un train qui est identifiable par le locuteur. C'est à cela que réfère vraisemblablement la distinction que l'auteure évoque entre propriétés « permanentes » et « temporaires ». On peut aussi argumenter que dans (I.5-6) un cadre de référence est fourni (la NUIT et le DEPART, respectivement). Au passage, notons que (I.5) fonctionnerait plus facilement en tant qu'énoncé prononcé *out-of-the-blue* (« The moon has disappeared ») par rapport au deuxième (« The train is leaving any minute »), sûrement du fait de l'unicité plus permanente du référent qu'il inclut.

« In addition to [...] permanently available topics (stage topics, speaker and hearer), **a number of permanent and temporary fixtures of our world are also available as topics with no need for previous mention.** This type includes “the president,” “the moon,” “the train,” and whatever other world fixtures the speaker and hearer happen to have in common » (Erteschik-Shir 2007 : 18, notre emphase).

En somme, nous avons présenté ici les différents statuts qu’un référent donné peut avoir au cours de la conversation, ce qui nous a permis de constater notamment qu’un référent « nouveau » possède différents degrés d’identifiabilité – selon qu’il est ancré ou non ancré à un autre élément du discours, et d’accessibilité – en fonction du contexte linguistique, extralinguistique ou des connaissances partagées par les interlocuteurs (v. aussi Lyons 1999 : 2-7). Enfin, comme nous le verrons par la suite (section 1.1.2.3), les référents nouvellement introduits diffèrent entre eux également en rapport au statut qu’ils auront dans le discours *à suivre* – aspect qui est décrit par la notion de « proéminence thématique ».

Sur la base des considérations présentées dans cette section, nous constatons que le référent « nouveau » prototypique est un référent qui n’est ni identifiable ni ancré, tandis qu’un référent identifiable et activé à un moment donné de l’échange linguistique est déjà établi dans le discours ; *a priori* dans ce dernier cas, l’entité référentielle n’aura pas besoin d’être introduite dans le discours par le biais d’une formulation linguistique *ad hoc* (il est cependant connu que les structures présentatives du français peuvent comporter des référents accessibles – voir la § 3.7 du chapitre III).

En 1.1.2, nous discutons plus en détail des manifestations formelles propres aux différents statuts référentiels. Pour l’instant, nous nous attardons brièvement sur un autre facteur qui intervient lorsqu’un référent est introduit dans le discours, à savoir la notion de proéminence thématique.

1.1.1.3. La proéminence thématique

Nous discutons ici d’un dernier facteur qui peut influencer la manière dans laquelle un référent est introduit dans le discours. Chafe (1987, 1994) estime que le degré d’accessibilité des référents varie à la fois selon leur statut cognitif au sein de l’énoncé et selon leur saillance. Un référent facilement identifiable est, tout d’abord, présupposé connu par les participants à l’échange linguistique. Ensuite, il doit être également désigné par une expression référentielle qui permet aisément son identification. En outre, il doit acquérir un degré de saillance contextuelle élevé en situation de concurrence référentielle :

« Examined more closely, identifiability can be seen to have three components. An identifiable referent is one that is (a) assumed to be already shared, directly or indirectly by the listener, (b) verbalized in a sufficiently identifying way; and (c) contextually salient » (Chafe 1994 : 94).

C'est à cette dernière caractéristique que nous nous intéressons ici. Une entité sur laquelle est axée la conversation se trouve au centre de l'attention des interlocuteurs et se caractérise par un fort degré d'activation. Formellement, son rôle central dans le discours se manifeste par la présence d'expressions référentielles portant sur cette même entité qui sont liées entre elles par une relation de coréférence (notion de « chaîne de référence », Chastain 1975²¹).

En même temps, une entité référentielle peut jouer un rôle plus ou moins central dans l'articulation du discours. Il convient de faire une distinction entre ces entités caractérisées par une forte saillance et situées au premier plan du discours, et des entités se situant à l'arrière-plan du discours. Chafe (1994) parle en termes d'*importance* des référents et en distingue des degrés différents dans une séquence narrative :

« [W]e may speak of referents as having *primary* importance [...], *secondary* importance [...], and *trivial* importance » (Chafe 1994 : 88).

En effet, lorsqu'une nouvelle entité référentielle est introduite dans le récit, elle peut être destinée à jouer un rôle central dans le discours suivant, ou bien n'avoir qu'une incidence secondaire, marginale. C'est seulement dans le premier cas que nous avons la promotion d'une entité au rang de thème du discours.

Toujours dans le souci d'observer la réalisation linguistique de ces considérations cognitives, cet aspect est important pour le linguiste s'intéressant à la structure informationnelle dans la mesure où il peut avoir une incidence sur les moyens formels mobilisés, reflétant un degré de centralité différent vis-à-vis de la structuration narrative.

On a constaté, par exemple, que la détermination grammaticale d'un syntagme nominal, au-delà de l'opposition entre référents identifiables et non identifiables, peut dénoter une autre distinction cognitive, à savoir l'opposition entre ces référents spécifiques que l'on souhaite dans un deuxième temps élever au statut de topique du discours et ceux qui n'ont qu'un rôle accessoire. Ainsi, le « marquage local » – en l'occurrence le choix du déterminant, voir la section suivante pour les détails sur cette notion – s'utilise aussi pour mettre en avant des entités référentielles dont on veut souligner la prééminence thématique (*thematic importance*, Givón 1984). En anglais parlé on trouve, par

²¹ Voir Schnedecker et Landragin (2014) pour un approfondissement de la notion et pour plus de références à ce sujet.

exemple, un contraste entre l'article indéfini *a* et le démonstratif *this* pour marquer un référent nouvellement introduit dans le discours (Lambrecht 1994 : 83) :

- (I.7) *I met a / this guy from Heidelberg on the train*
 'J'ai rencontré un / ce gars de Heidelberg dans le train' (Lambrecht 1994 : 83)

Le locuteur aura tendance à choisir le démonstratif – entraînant l'encodage morphologiquement défini d'une entité sémantiquement indéfinie – lorsqu'il veut apporter d'autres informations à propos du référent visé dans le discours à suivre. Dans cet usage qu'on peut définir comme cataphorique, le démonstratif est appelé « indefinite *this* » (Prince 1981)²².

Dans certaines langues où le marquage de l'indétermination est réalisé par l'association du numéral et du classificateur (ex. en chinois la séquence *yí-ge* 'un-CL'), souvent le syntagme nominal quantifié se voit assigner le statut de topique pour le prochain discours (Lambrecht 1994 : 83). Ainsi, le contraste suivant a été remarqué en chinois : « an unanchored referential-unidentifiable referent which is not to become a topic (is incidental to the discourse) will often either not have the numeral plus classifier, or will have the classifier, but not the numeral » (LaPolla 1995 : 306).

Nous verrons que ce critère est crucial en chinois pour distinguer les structures présentatives comportant la mise en avant d'un référent, de ces énoncés dits « événementiels », dans lesquels l'accent repose sur l'évènement conceptualisé dans son ensemble, dont l'entité évoquée n'est qu'un simple participant nécessaire au déroulement de ce même évènement.

En effet, nos recherches de corpus concernant les structures présentatives en *yǒu* '(y) avoir' et l'ordre V-S du chinois (présentées dans le chapitre IV § 4.4.3 et dans le chapitre V § 5.2.2.1 respectivement) confirment cette hypothèse. Par exemple, dans l'extrait qui suit, le nominal nu *lǐngdǎo* 领导 'chef' introduit au sein de la structure en *yǒu* biclausale est nouveau pour le discours mais n'est pas destiné à en devenir le topique – celui-ci reste « les facteurs qui rassurent le parents » :

- (I.8) 都是愿意孩子有个正经的工作, (BJKY²³)
Dōu shì yuànyì háizi yǒu ge zhèngjīng de gōngzuò,
 tous être espérer fils AVOIR CL respectable DE travail
 ➤ 有领导管着, 有组织,
yǒu lǐngdǎo guǎn-zhe, yǒu zǔzhī,
 AVOIR chef s.occuper-DUR AVOIR structure
 这样儿家长就放心。
zhèyàngr jiāzhǎng jiù fàngxīn
 ainsi chef.de.famille alors être.tranquille

²² Un autre usage du déterminant démonstratif qui intervient pour souligner la saillance d'un référent nouvellement introduit est celui du « reminder *that* » (Gundel *et al.* 1993).

²³ Les exemples marqués par l'abréviation « BJKY » proviennent de notre recherche dans le corpus de pékinois parlé *Běijīng kǒuyǔ yǔliàokù* 北京口语语料库, et de ce fait relèvent du registre oral de la langue. Voir la section 4.6.1 du chapitre IV pour une présentation détaillée du corpus.

‘[Les parents] veulent que leurs fils aient un travail respectable, [qu’ils] aient un chef qui s’occupe [d’eux], [qu’il y] ait une structure, de cette manière les parents sont rassurés.’

En contraste, dans l’extrait en (I.9), le nom quantifié introduit par la structure en *yǒu*, à savoir *yí-ge rén* 一个人儿 ‘une personne’ est nouveau et en même temps est élevé au rang de topique dans le discours immédiat. Ce référent est ensuite marqué par un nom modifié par le déterminant démonstratif *zhè* au sein d’une unité intonative autonome (clôturée par la particule *ya*)²⁴ :

- (I.9) 我们这儿有一个人儿, (BJKY)
Wǒmen zhèr yǒu yí-ge rén,
 1PL ici AVOIR un-CL personne
 名叫杨和平。
míng jiào Yáng Héping.
 nom s.appeler Yang Heping
 这个女的呀, 是大有庄的一个难点。
Zhè-ge nǚde ya, shì Dàyǒuzhuāng de yí-ge nándiǎn.
 DEM_{prox}-CL femme SFP être Dayouzhuang DE un-CL problème
 ‘Ici chez nous on a quelqu’un qui s’appelle Yang Heping. Cette femme, hé, c’est une plaie pour le quartier.’

Nous voilà donc arrivés à la question de savoir comment les propriétés cognitives des référents se reflètent formellement dans la langue. Cet aspect est discuté en détail dans la section suivante.

1.1.2. Les réalisations linguistiques des différents statuts référentiels

Etant donné qu’une analyse se situant au niveau de la structure informationnelle ne prend en compte que les propriétés pragmatiques ayant une manifestation concrète dans la langue (aux différents niveaux de la morphosyntaxe, du lexique et de la prosodie), le linguiste travaillant dans ce cadre s’intéressera à la manière dont ces types de référents sont représentés linguistiquement.

Parmi les outils linguistiques dont le locuteur dispose pour assurer les opérations de référenciation, il convient de distinguer entre le marquage « local » et le marquage « global » (Hickmann *et al.* 1996). Par exemple, le choix d’un article défini ou indéfini en français relève du marquage local (même si sa portée va évidemment au-delà de la phrase isolée). D’autre part, le

²⁴ D’un point de vue strictement syntaxique, la structure en *yǒu* dans l’exemple (I.9) se présente dans sa variante monoclausale. Remarquons toutefois que l’énoncé *wǒmen zhèr yǒu yí-ge rén* ‘ici chez nous il y a une personne’ n’est pas complet d’un point de vue informatif sans la proposition qui suit : *míng jiào Yáng Héping* ‘qui s’appelle Yang Heping’ (lit. ‘dont le nom est Yang Heping’). En isolation, la seule interprétation possible serait ‘ici chez nous il y a une personne – et non deux, trois, aucune’ mais cela n’est pas le cas. Dans tous les cas, voir la section 4.6.4.1 du chapitre IV pour des exemples de structure présentative biclausale en *yǒu* au sens strict.

marquage global repose sur de stratégies ayant un champ d'action plus large, comme par exemple la variation de l'ordre des mots au niveau phrastique. Les structures présentatives (1.2) rentrent dans cette catégorie.

La présente section est consacrée à la description des procédés se situant au niveau du marquage local. Il s'agit d'étudier le rapport entre les diverses expressions référentielles disponibles dans une langue, et les indications qu'elles apportent sur le degré d'accessibilité de leur référent. Aux catégories cognitives de l'identifiabilité et de l'accessibilité (1.1.1.1) correspondent, donc, une série de propriétés formelles.

Tout d'abord, un référent non identifiable reçoit une proéminence prosodique, et au sein des langues pourvues de la catégorie grammaticale de la « détermination », il est typiquement encodé par un syntagme nominal indéfini. D'un point de vue de la manifestation formelle des catégories cognitives, il faut tenir compte du fait que la corrélation entre l'identifiabilité et la détermination grammaticale est forte mais pas absolue (ce sujet sera discuté dans la section suivante).

Ensuite, nous le rappelons, une fois qu'un référent donné est jugé identifiable, il se trouve nécessairement dans l'un des trois états d'activation suivants : il sera activé, inactivé ou accessible (Lambrecht 1994 : 106). Un référent inactivé, même s'il est identifiable, doit être doté de proéminence prosodique, et est typiquement exprimé par un syntagme nominal lexical défini (*ibid.* : 107), alors qu'un référent déjà activé est généralement codifié par une expression pronominale non accentuée (*ibid.* : 106). On estime que seules les entités référentielles dont la présence a déjà été établie dans le discours peuvent être encodées en tant qu'expressions anaphoriques réduites telles que les pronoms et l'anaphore zéro (Chafe 1976, Allerton 1978, Westergaard 1986 : 64) : une fois que la référence à une entité est établie dans la conversation, celle-ci devient activée dans l'esprit des interlocuteurs et peut recevoir un encodage anaphorique lexicalement atténué.

Dans leur essence, les procédés de référenciation sont orientés vers le destinataire (*recipient-designed*, Sacks et Schegloff 1979) : lorsque le locuteur choisit une forme linguistique pour désigner un référent, il doit s'assurer que celle-ci puisse permettre à son partenaire de « localiser » aisément le référent en question. Ainsi, pour assurer le succès de l'échange communicatif, les locuteurs doivent prendre nombre de décisions concernant le statut des entités discursives pour assurer leur introduction, leur maintien et leur réintroduction dans le discours (Levelt 1989). C'est ainsi que la mise en place des relations anaphoriques dans la conversation (le choix d'une forme plutôt qu'une autre pour désigner un référent donné) repose entièrement sur les hypothèses faites par le locuteur sur la manière dont l'interlocuteur pourra interpréter le référent en question.

Ariel (1988, 1990, 2001) a dédié de nombreux travaux à l'établissement d'une théorie de l'accessibilité, visant à rendre compte du degré d'accessibilité des référents au niveau du texte. Elle

identifie quatre facteurs qui permettent à l'interlocuteur d'interpréter le degré d'accessibilité d'un référent donné : la distance, la concurrence, la saillance et l'unité :

« [I]t seems that at least the following factors contribute to the assumed Accessibility status of an antecedent:

- Distance: The distance between the antecedent and the anaphor (relevant to subsequent mentions only).
- Competition: The number of competitors on the role of antecedent.
- Saliency: The antecedent being a salient referent, mainly whether it is a topic or a non-topic.
- Unity: The antecedent being within vs. without the same frame/world/point of view/segment or paragraph as the anaphor. » (Ariel 1990 : 28-29)

L'interprétation du degré d'accessibilité dépend en particulier de la distance entre l'expression référentielle et son antécédent (le concept de *distance référentielle* chez Givón 1983). En d'autres termes, plus la distance est réduite entre l'expression référentielle et l'antécédent, plus le référent est présent dans la mémoire récente de l'interlocuteur et plus son degré d'accessibilité est élevé (ce point est développé aussi par Prince 1981 : 228).

Le deuxième facteur concerne la compétition entre plusieurs référents (notion d'*interférence potentielle* chez Givón 1983). Plus un fragment de texte contient des référents potentiellement concurrents, moins le référent auquel renvoie l'expression anaphorique est accessible.

Un autre facteur concerne le degré de saillance de l'antécédent (son rôle plus ou moins important par rapport au discours). Enfin, l'unité entre l'expression anaphorique et le référent tient aussi à la position de ce dernier au niveau textuel (le facteur positionnel).

Tous ces facteurs concourent à déterminer le degré d'accessibilité d'un référent à un moment donné du discours : il est nécessaire de les considérer dans une perspective de complémentarité, car la prise en compte d'un seul facteur est insuffisante pour dégager le degré d'accessibilité propre à une expression référentielle. Voici un exemple simple de chaîne référentielle en anglais codée selon cette approche :

(I.10) *A man* [discourse new, non-accessible] *went into a store, where a woman was reading a magazine. The man* [discourse old, intermediate accessibility] *needed some food. He* [discourse old, high accessibility] *looked at some vegetables and Ø* [discourse old, highest accessibility] *took a potato.* (d'après Crosthwaite et al. 2018)

Selon Ariel (1996), l'introduction référentielle est sujette aussi au concept d'accessibilité, car d'après elle les SN démonstratifs indéfinis (comme '*this+N*' discuté plus haut) sont utilisés pour des

référents plus accessibles que les indéfinis quantifiés (*'another+N'*), lesquels à leur tour sont utilisés pour des référents plus accessibles que ceux dénotés par de simples indéfinis (*'a+N'*).

A la lumière de ces considérations, on comprend que les marques formelles ne peuvent que donner une indication approximative sur le degré d'accessibilité d'un référent donné, lequel à son tour est déterminé par plusieurs facteurs à la fois.

Dans ce qui suit, nous nous penchons plus en détail sur la corrélation « imparfaite » entre l'identifiabilité cognitive d'un référent et sa réalisation linguistique en termes de détermination grammaticale.

1.1.2.1. Identifiabilité et détermination grammaticale

Une manifestation grammaticale importante de la distinction cognitive entre référents identifiables et non identifiables prend la forme de l'opposition formelle entre noms définis et indéfinis. Dans de nombreuses langues du monde, la catégorie de l'identifiabilité est régulièrement exprimée par le contraste entre articles indéfinis et définis ou d'autres déterminants (tels que les possessifs, les démonstratifs etc.). Parmi les autres moyens morphosyntaxiques observés sont cités la présence ou l'absence d'un numéral, des particules marquant le cas, etc. (Lambrecht 1994 : 79). Au niveau proprement syntaxique, l'ordre des mots peut jouer un rôle dans l'expression de l'identifiabilité. En chinois, par exemple, l'interprétation des noms nus diffère selon leur position par rapport au verbe. L'exemple typique cité dans la littérature concernée est le suivant (Li et Thompson 1985 : 510) :

(I.11) 来了。
Lái rén le.
venir personne CRS
'Quelqu'un est arrivé.'

(I.12) 人来了。
Rén lái le.
personne venir CRS
'La/les personne.s (que nous attendions) est/sont arrivé.s.'

Cependant, même dans les langues disposant d'une opposition grammaticalisée entre définis et indéfinis, la corrélation entre la catégorie cognitive de l'identifiabilité et son expression formelle en termes de détermination grammaticale n'est pas biunivoque : certains référents non identifiables sont encodés par des syntagmes nominaux définis, et vice-versa (Lambrecht 1994 : 105). Un syntagme nominal indéfini peut aussi bien dénoter un référent spécifique, lorsqu'il est identifiable

par le locuteur mais ne l'est pas par l'interlocuteur, ou désigner un référent non spécifique, qui n'est identifiable par aucun des deux participants à l'échange linguistique (Lambrecht 1994 : 81).

Par exemple, dans (I.13) si le locuteur a un livre précis en tête, le référent est identifiable (uniquement) pour le locuteur, tandis que si le locuteur a l'intention d'acheter un livre sans savoir encore lequel, le référent n'est identifiable pour aucun des interlocuteurs, comme on le voit explicitement formulé en (I.14) :

(I.13) *Je vais acheter **un livre***

(I.14) *Je vais acheter **un livre** pour m'occuper pendant le voyage* (Creissels 2006)

A l'inverse, les définis peuvent renvoyer à un référent non identifiable par l'interlocuteur, comme dans le cas du « *this* indéfini » de l'anglais discuté plus haut, et aussi du nom modifié par « les » ici :

(I.15) *J'ai acheté un livre, mais je n'ai lu que **les premières pages*** (Creissels 2006 : 82)

En effet, nous avons vu (en 1.1.1) qu'un référent peut être accessible, et donc marqué en tant que nom défini, pour plusieurs raisons : à partir de la situation immédiate (deixis) ou du contexte plus large (nos connaissances du monde) mais aussi à partir du cotexte, de manière directe (anaphore fidèle ou infidèle) ou indirecte (anaphore associative).

Ces divers phénomènes s'expliquent encore une fois par le fait qu'une catégorie cognitive non discrète se reflète dans une catégorie formelle discrète. Les différents usages des articles défini et indéfini à travers les langues du monde renvoient à des points de saisie différents dans le *continuum* de l'identifiabilité des référents (Lambrecht 1994 : 84). Une des conséquences qui en découle est la nature polyfonctionnelle des articles et, *mutatis mutandis*, des formes linguistiques disponibles dans les langues sans articles (ex. les noms nus en chinois qui peuvent renvoyer à différents degrés d'accessibilité, voir la section 1.1.2.3).

Dans ce qui suit, nous discutons plus en détail d'un type particulier d'anaphore, à savoir l'anaphore associative que nous venons de mentionner – et dont (I.15) est d'ailleurs un exemple. En effet, il s'agit dans ce cas d'un procédé employé pour introduire les nouveaux référents dans le discours, qui semble particulièrement difficile à acquérir en L2 (chapitre II § 2.5) et que nous analyserons en rapport aux données recueillies dans le cadre de notre expérience de linguistique appliquée (chapitre VI § 6.4.1 pour les L1 et chapitre VII § 7.6.1 pour les L2).

1.1.2.2. L'anaphore associative (pontage inférentiel)

Le phénomène linguistique de l'anaphore associative est étroitement lié au concept de « cadre cognitif » discuté en 1.1.1.2. Comme nous l'avons vu au cours de ce chapitre, les nouveaux référents typiques sont encodés par un SN indéfini. En faisant ce choix de marquage local, le locuteur signale à son interlocuteur qu'il devra créer une représentation d'un référent *ex nihilo* dans le modèle mental du discours partagé (le *common ground* de Stalnaker 2002). C'est là la raison d'être des structures présentatives, qui invitent explicitement l'interlocuteur à effectuer cet effort de repérage. D'autre part, les noms définis sont typiquement employés pour maintenir la référence aux entités nouvellement introduites, pour assurer la continuité référentielle en marquant la transition entre information donnée et information nouvelle.

Cependant, les locuteurs peuvent utiliser un nom défini pour introduire un référent nouveau dans le discours si son existence ou son identité peut être présupposée par l'interlocuteur. Il s'agit alors là des formes linguistiques appelées *bridging descriptions* ou *bridging definites* (Clark 1975, Haviland et Clark 1977), termes qui renvoient au processus du « pontage inférentiel » (*bridging*) dont elles font l'objet. Hawkins (1978) utilise le terme *associative definites*. Dans la littérature linguistique en français on utilise plus souvent le terme d'*anaphore associative* (Charolles 1990, Kleiber 1993).

Les exemples typiques d'anaphores associatives comportent donc un nom défini qui renvoie à une entité nouvelle en discours « mais qui est présentée comme accessible à partir d'une entité antérieurement mentionnée, via une relation inférentielle de caractère générique ou stéréotypique, très souvent méronymique » (Charolles et Choï-Jonin 1995) :

(I.16) *Nous arrivâmes dans un village. L'église était située sur une hauteur* (Kleiber 1993)

Plusieurs motivations sous-jacentes à l'opération de pontage inférentiel ont été évoquées, qui peuvent être résumées en deux hyper-groupes : la relation lexicale avec un antécédent (ex. *gâteau ... la part*) ou la relation non lexicale qui renvoie à notre connaissance du monde ou à la structure du discours (ex. *guerre ... les survivants*) (Crosthwaite 2016).

Il est important de préciser que cet usage représente une association alternative de l'appariement forme-fonction par rapport à l'emploi typique des définis pour maintenir la référence, et que ces expressions définies n'ont pas nécessairement un antécédent dans le discours antérieur (Gundel 1996).

Par exemple, l'étude de Sanford et Garrod (1981, citée dans Crosthwaite 2014) montre que les SN définis dont la référence est dépendante du scénario (comme par ex. « le serveur » ou « le cuisinier » dans le cadre d'un RESTAURANT) ne sont pas plus longs à traiter cognitivement que les entités

mentionnées antérieurement en discours, malgré l'absence d'antécédents explicites (Crosthwaite 2014 : 459).

Il convient de souligner aussi que le phénomène du pontage inférentiel ne s'observe pas uniquement dans les langues à articles comme l'anglais ou le français, mais également dans les langues ne disposant pas de cette catégorie grammaticalisée. L'exemple suivant, adapté de Crosthwaite *et al.* (2018) montre le même énoncé en anglais, chinois et coréen²⁵.

(I.17) *A man walks into a restaurant.
The waiter gave him a menu.*

(I.18) 有一个男孩走进餐厅,
Yǒu yí-ge nánhái zǒu-jìn cāntīng,
AVOIR un-CL garçon marcher-entrer restaurant
服务员给他一份菜单。
fúwùyuán gěi tā yī-fèn càidān
serveur donner 3SG un-CL menu
'Il y a un garçon qui est rentré dans un restaurant. Le serveur lui a donné le menu
(lit : un exemplaire de menu).'

(I.19) 한 남자가 식당으로 걸어갔다.
Han namcaka siktangulo kelekassta.
un homme.NOM restaurant entrer.PAST.DEC
종업원이 그남자에게 메뉴판을 줬다.
Congepweni kunamcaeykey meynyuphanul cwesta.
serveur.NOM DEM_{dist.}homme.vers menu.ACC donner.PAST.DEC.
'Un homme est rentré dans un restaurant. Le serveur lui a donné le menu.'

Notons, au passage que le nouveau référent indéfini *yí-ge nánhái* 'un garçon' est introduit en chinois par le biais d'une structure présentative biclausale régie par *yǒu* '(y) avoir'. Ensuite, le référent du « serveur », qui fait l'objet du pontage inférentiel en raison de son association avec le cadre du RESTAURANT, est introduit par un nom modifié par l'article défini en anglais ('*the* + N'), et en tant que nom nu en chinois et en coréen.

Le pontage inférentiel peut être considéré comme la manifestation d'un phénomène plus général, que Lambrecht (1994 : 66) appelle « accommodation pragmatique », en élaborant le principe formulé par Lewis (1979)²⁶ :

²⁵ Nous avons légèrement modifié la présentation des exemples pour l'adapter au style typographique de notre thèse. Le mot-à-mot et la traduction de la phrase chinoise sont plus détaillés que dans l'article original mais nous avons gardé les gloses de l'article d'origine pour le coréen, où on n'opérait pas de séparation en morphèmes. Les catégories fonctionnelles utilisées dans cet exemple sont : NOM = nominative marker ; PAST = past-tense inflection ; DEC = declarative inflection

²⁶ Voir aussi Stalnaker (1973) et Abbott (2000).

« If at time *t* something is said that requires presupposition *P* to be acceptable, if *P* is not presupposed just before *t*, then – *ceteris paribus* and within certain limits – presupposition *P* comes into existence at *t*. » (Lewis 1979 : 340).

D'après ce principe, les présuppositions qui ne faisaient pas partie du cotexte peuvent y être associées au moment où l'énoncé est prononcé. Ainsi, le locuteur est en mesure de créer une nouvelle situation présuppositionnelle juste en utilisant une expression qui la requiert (Lambrecht 1994 : 66-67). Par exemple, les locuteurs engagés dans la conversation peuvent accepter un élément comme étant déjà établi en discours au moment même où cet élément y fait son apparition. Puisque c'est le locuteur qui « demande » à son interlocuteur de faire cet effort d'accommodation, un jugement inapproprié de sa part donnera lieu à un échec communicatif.

Avant de terminer cette section 1.1.2 sur les réalisations linguistiques des catégories cognitives de l'identifiabilité et de l'accessibilité, il convient de nous attarder un peu plus sur les formes disponibles en chinois à cet effet.

1.1.2.3. Le marquage des différents statuts référentiels en chinois

Regardons maintenant comment se manifestent linguistiquement les différents statuts référentiels en chinois. Dans cette langue il n'existe pas de catégorie grammaticale de la définitude ; cela ne veut pas dire évidemment que le chinois ne marque pas la distinction entre référents accessibles et non accessibles. En l'absence d'articles, cette fonction est réalisée localement par des procédés divers.

Sur la base de l'étude d'un corpus oral, Huang (2013 : 150) établit comme suit l'échelle d'accessibilité pour les expressions anaphoriques en chinois : Anaphore zéro > pronom > SN déterminé > nom propre > nom nu > SN possessif > SN complexe. Cette échelle se base sur la distance entre l'expression anaphorique et son antécédent. La voici dans les détails :

Tableau I-1. *L'échelle d'accessibilité pour les expressions anaphoriques en chinois d'après Huang (2013)*

- Anaphore zéro
- Pronom (*tā* 她 '3SG', *zìjǐ* 自己 'soi-même')
- SN déterminé
 - Démonstratif + nom (*zhège xuésheng* 这个学生 'cet étudiant')
 - Démonstratif + nom propre (*zhège Mǎlì* 这个玛丽 'cette Marie')

-
- Nom nu (*xuésheng* 学生 ‘étudiant’)
 - Possessif (*tāde xuésheng* 她的学生 ‘son étudiant’)
 - Nom propre (*Mǎlì* 玛丽 ‘Marie’)
 - SN complexe
 - Proposition relative + nom (*tāde nǚpéngyǒu suǒ yǎng de māo* 他的女朋友所养的猫 ‘le chat que sa copine a [chez elle]’)
 - SN possessif + nom propre (*tāde xuésheng Yáng Jiàn* 他的学生杨健 ‘son étudiant Yang Jian’)
-

Plus un référent est accessible, plus la forme linguistique sera légère, et cela jusqu’à l’anaphore zéro.

LaPolla (1995 : 305) rend compte des différents codages qu’un nom peut recevoir en chinois en considérant à la fois son degré d’identifiabilité et d’accessibilité:

Tableau I-2. *Les différents codages des référents en chinois sur la base de leur degré d’identifiabilité et d’accessibilité d’après LaPolla (1995)*

<i>Type de référent</i>	<i>Codages possibles</i>
Activé	zéro ; pronom ; nom nu ; avec déterminant démonstratif ²⁷
Accessible	pronom ; nom nu ; avec déterminant démonstratif
Inactivé	nom nu ; avec déterminant démonstratif
Non-ancré non-identifiable	nom nu ; (numéral +) classificateur
Ancré non-identifiable	SN génitif ; SN subordinatif
Non-référentiel	nom nu ; (numéral +) classificateur

Un référent activé, au centre de l’attention des interlocuteurs, est souvent encodé en chinois par un pronom ou peut rester implicite (anaphore zéro). En général, le chinois distingue entre les référents activés et inactivés en réservant l’anaphore zéro aux premiers, et entre les référents identifiables et non identifiables de par ce que ces derniers ne sont jamais représentés par un nom précédé par un déterminant démonstratif (LaPolla 1995).

A part ses usages non-référentiels, la séquence [(numéral +) classificateur] est dédiée à l’expression des référents non ancrés et non identifiables : les référents nouveaux par excellence (1.1.1.1).

²⁷ L’article d’origine parle en termes de « pronom déictique » (*deictic pronoun*) mais réfère vraisemblablement aux démonstratifs *zhè* et *nà* qui ne sont donc pas des pronoms dans leur emploi adnominal, comme semble le suggérer LaPolla (1995 : 305-307).

La forme qui présente le plus haut degré d’ambiguïté (tableau I-2) est le nom nu, lequel peut désigner un référent se trouvant dans chacun des statuts cognitifs mentionnés (à l’exception des référents non-identifiables mais ancrés). Nous citons ici un passage tiré de Hole (2012 : 61), qui résume bien les propriétés des noms nus du chinois :

« [B]are nouns constitute an especially interesting domain of investigation because they contain no functional morphemes indicating their status as given or new. Despite the lack of definiteness marking, bare nouns display definiteness effects in some syntactic environments. The first such effect concerns subjects vs. objects. Bare nouns in a non-subject (and non-topical) position are typically interpreted as indefinite (unless they denote inherently definite entities), whereas they are invariably definite in the subject position preceding the verb [...]. This contrast is particularly striking with unaccusative verbs of (dis-)appearance, which allow their sole arguments to surface either preverbally or postverbally » (Hole 2012 : 61).

Les études expérimentales de Hickmann *et al.* (1996) et de Crosthwaite (2014) montrent que la séquence [yī-CL + N] est significativement plus susceptible d’être utilisée pour encoder les nouveaux référents indéfinis non spécifiques que les noms nus le sont pour cette fonction, tandis que ces derniers sont la forme privilégiée pour renvoyer à des entités à référence définie dont les anaphores associatives (*bridging*).

L’énoncé suivant comporte le nom nu *yuèliàng*, lequel renvoie à une référence unique (« la lune ») en vertu de nos connaissances générales du monde (on peut comparer avec les exemples de la section 1.1.1.2) :

(I.20) 月亮升上来了
Yuèliàng shēng-shang-lai le
 lune surgir-monter-venir CRS
 ‘La lune s’est levée’ (Chen 2004 : 1165)

Pour l’expression d’un référent accessible depuis le cotexte on emploie généralement en chinois le nom modifié par un déterminant démonstratifs :

(I.21) 后来，嗯，这个犯罪分子从车上跑下来了 (BJKY)
hòulái, èng, zhè-ge fànzuìfēnzǐ cóng chē=shang pǎo-xia-lai-le
 après INT DEM_{prox}-CL criminel depuis voiture=sur sauter-descendre-venir-PFV
 ‘Après, euh, ce criminel est descendu de la voiture d’un saut.’

Selon Jenks (2018), le chinois fait une différence entre les définis à référence unique et les définis à référence anaphorique. Il remarque qu'en chinois les noms nus à interprétation définie sont associés à l'expression de la référence unique, tandis que les démonstratifs sont associés à l'expression de la référence anaphorique. La position sujet constituerait une exception, puisque les noms nus dans cette position peuvent avoir aussi une référence anaphorique. Plus récemment, Dayal et Jiang (à paraître) remettent partiellement en cause cette analyse et estiment que les noms nus peuvent fonctionner dans la référence anaphorique tout autant que les démonstratifs²⁸. Dans notre corpus de chinois L1, les nominaux nus en position sujet peuvent renvoyer aussi bien à une référence unique qu'anaphorique.

En dehors de l'emploi d'une structure présentative, un moyen courant d'introduire un nouveau référent non identifiable dans le discours est de le lier à un deuxième syntagme nominal ayant déjà été introduit, ou qui est plus facile à repérer (Lapolla 1995 : 306). Il s'agit dans ce cas d'un référent non identifiable mais *ancré* (dont nous avons discuté dans la section 1.1.1.1).

- (I.22) 学校的一个工人昨天出了车祸
Xuéxiào de yí-ge gōngrén zuótiān chū-le chē-huò
 école DE un-CL employé hier se.produire-PFV voiture-accident
 'Yesterday one of the workers in the school got into a car accident' (Lapolla 1995 : 306)

C'est la raison pour laquelle Chen (1986 : 16-17, cité dans LaPolla 1995) considère tous les SN génitifs ou subordinatifs comme des « définis ».

Dans la phrase suivante, les référents nouvellement introduits au moyen du verbe existentiel-possessif *yǒu* 有 '(y) avoir' (dont on discutera dans le chapitre IV) ont la forme d'un syntagme nominal indéfini quantifié (référent non identifiable, non ancré) *yí lǎo-rén* [un-CL vieille-personne] 'une personne âgée' et *liǎng-ge xiǎoháir* [deux-CL enfant] 'deux enfants' et d'un SN génitif (référent ancré) *wǒ àirén* 'mon partenaire'. Une fois introduits dans le discours, et donc activés, ils se présentent comme des nominaux nus : *àirén* 'partenaire' et *xiǎoháir* 'garçon'²⁹.

- (I.23) 家里情况也挺好的, 呵。 (BJKY)
Jiā=li qíngkuàng yě tǐng hǎode, hē.
 maison=dans situation aussi très bien INT

²⁸ « [W]e can safely take bare NPs to function as unique definites and demonstratives to function as deictic expressions. The question that remains is the status of anaphoric definites. We argue that both types of NPs can function in canonical anaphoric contexts. We have suggested looking at situations in terms of the individuals that define them as a way of capturing overlaps in distribution » (Dayal et Jiang 2020 : 22).

²⁹ On peut estimer que ces noms renvoient à des référents à la fois activés et à référence unique – liés déictiquement et implicitement à *wǒ* 'je', l'un de ces « topiques disponibles en permanence » (*permanently available topics*) chez Erteschik-Shir (2007 : 17).

有一老人, 有我爱人, 有两个小孩儿。

Yǒu yí lǎorén, yǒu wǒ àirén, yǒu liǎng-ge xiǎoháir.

AVOIR un.CL³⁰ personne.âgé AVOIR 1SG partenaire AVOIR deux-CL enfant

➤ 嗯, 爱人有工作, 小孩儿也工作啦

Nèg, àirén yǒu gōngzuò, xiǎoháir yě gōngzuò la

INT partenaire AVOIR travail enfant aussi travailler CRS.SFP

‘A la maison la situation est très bien aussi, hein. Il y a une personne âgée, il y a mon partenaire, il y a deux enfants. Mon partenaire a un travail, les enfants travaillent aussi.’

Les exemples suivants montrent le passage d’un référent du statut d’inactif non-identifiable au statut d’actif, où l’entité en question est introduite dans le discours au moyen de la structure présentative V-S (discutée dans le chapitre V) :

(I.24) 正好头里过来一老头儿, (BJKY)

Zhènghǎo tóuli guò-lai yí lǎotóur

justement devant passer-venir un.CL vieil.homme

这老头儿,

zhè lǎotóur,

DEM_{prox} vieil.homme

他站着还不动活

tā zhàn-zhe hái bù dònghuó

3SG se.tenir.début-DUR toujours NEG bouger

‘Juste à ce moment-là il y a un homme qui est venu d’en face, cet homme, il se tenait début sans bouger’

(I.25) 刚要走的时候, (BJKY)

Gāng yào zǒu deshíhòu,

juste vouloir partir quand

从那边儿过来一个妇女,

cóng nà-biānr guò-lai yí-ge fùnǚ,

depuis là.bas passer-venir un.CL femme

就问我们干什么

ø jiù wèn wǒmen gàn shénme

alors demander IPL faire quoi

‘Juste quand on était sur le point de partir, il y a une femme qui s’est rapprochée de là-bas, qui nous a demandé ce qu’on faisait’

En (I.24), on assiste à une progression thématique graduelle : le nouveau référent introduit comme non identifiable est ensuite encodé par un nom modifié par le déterminant démonstratif (marque des référents accessibles depuis le contexte) et ensuite par le pronom personnel (haut degré

³⁰ Notons que *yí* (au deuxième ton) en (I.23) mais aussi en (I.24-25) plus bas est la contraction du numéral *yī* — ‘un’ et du classificateur générique *ge* 个 en pékinois parlé, et donc crée un nom quantifié, ce qui est important à l’égard des hypothèses que nous avançons dans notre travail. Dans le même paradigme, nous avons aussi *liǎng* 俩 qui est une fusion de *liǎng* ‘deux’ et *ge*, et *sān* 仨 (*sān-ge* [trois-CL]), ces derniers étant notés à l’écrit par un caractère différent que les numéraux correspondants (sur ce sujet, nous renvoyons le lecteur aux travaux de Chirkova 2004 et Liú Xiángbó 2004).

d'activation). En (I.25) le référent nouvellement introduit reste implicite dans la proposition suivante (anaphore zéro – le plus haut degré d'activation). Rappelons l'exemple de l'anglais cité plus haut :

(I.26) *A man* [discourse new, non-accessible] *went into a store, where a woman was reading a magazine. The man* [discourse old, intermediate accessibility] *needed some food. He* [discourse old, high accessibility] *looked at some vegetables and Ø* [discourse old, highest accessibility] *took a potato.* (d'après Crosthwaite *et al.* 2018) (= I.11)

(I.27) *A man* [discourse new, non-accessible] *went into a store and Ø* [discourse old, highest accessibility] *took a potato.*

Dans le discours réel, les chaînes référentielles ne prennent pas toujours la forme linéaire exemplifiée en (I.26) ; en effet, plusieurs solutions sont possibles (telles que la forme « réduite » en [I.27], assimilable à l'exemple chinois présenté ci-dessus en [I.25])³¹.

Nous concluons ici cette section sur le statut cognitif des référents en discours et les formes linguistiques dont les langues disposent pour réaliser ces différences. Ces formes relèvent du marquage local. Avant de passer en 1.2 à la description des structures employées dans les langues pour introduire les nouveaux référents en discours – un marquage de type global –, nous consacrons en 1.1.2 une section à une notion souvent associée à la fonction d'introduction référentielle, à savoir la notion de *théticité*.

1.1.3. Jugement thétiq ue et « détopicalisation » du sujet

1.1.3.1. Proposition thétiq ue et proposition catégoriq ue

Le concept de « théticité » est de nature sémantique, inspiré par la distinction philosophique entre assertions thétiq ues et catégoriq ues³². Cette perspective prend pour point de départ la distinction cognitive entre deux types de jugement. Le jugement catégoriq ue, ou « double » (Marty 1918, cité dans Lambrecht 1987a) comprend l'acte cognitif de reconnaître un sujet et celui d'affirmer (ou de nier) ce qui est exprimé par le prédicat à propos de ce sujet. De l'autre côté, le jugement thétiq ue ou « simple » implique seulement un acte de reconnaissance ou de rejet tout court.

³¹ Voir toutefois Li et Thompson (1979) pour une discussion générale sur le choix entre l'emploi d'un pronom de troisième personne et l'anaphore zéro en chinois, et les hypothèses qui y sont avancées.

³² Pour une discussion détaillée de la notion de théticité, nous renvoyons le lecteur aux travaux de Sasse et tout particulièrement à son article de 2006 ainsi qu'aux références qui y sont citées (2006 : 257).

En d'autres termes, si dans le cas d'un jugement catégorique on dit quelque chose à propos d'une entité, le jugement thétique ne fait que « poser » (le terme est composé sur la base du grec ancien *tithēmi* 'mettre, poser') l'existence d'un certain état des choses (Sasse 2006 : 259). Kuroda (1972 : 154) le formule ainsi :

« This [Brentano-Marty's] theory assumes, unlike either traditional or modern logic, that there are two different fundamental types of judgments, the categorical and thethetic. Of these, only the former conforms to the traditional paradigm of subject-predicate, while the latter represents simply the recognition or rejection of material of a judgment. Moreover, the categorical judgment is assumed to consist of two separate acts, one the act of recognition of that which is to be made the subject, and the other, the act of affirming or denying what is expressed by the predicate about the subject. With this analysis in mind, thethetic and the categorical judgments are also called the simple and the double judgments (Einfache Urteil and Doppelurteil). » (Kuroda 1972 : 154).

Typiquement le jugement complexe se manifeste dans les phrases à structure topique-commentaire, alors que les phrases thétiques sont logiquement simples, sans présupposition. Quel est donc le lien entre une proposition à caractère thétique et la fonction d'introduction référentielle – qui nous intéresse tout particulièrement dans notre travail ? Précisément, les propositions thétiques ont une fonctionnalité discursive spécifique, car « leur raison d'être est de servir à présenter une entité, une proposition ou un état de choses en tant qu'élément d'information nouveau pour le discours » (Cornish 2008). Ainsi, à la suite des travaux de Kuroda (1972) et de Sasse (1987) cités plus haut on considère que les énoncés présentatifs ont une nature thétique, en tant qu'une instance de la classe plus large des jugements thétiques (« a special case of the more inclusive class ofthetic judgments », Gast et Haas 2011 : 132).

Ne demandant qu'une présupposition minimale, la proposition thétique se trouve souvent à la frontière des unités du discours. Elle peut signaler une rupture dans la continuité narrative et commencer de ce fait une nouvelle séquence discursive (Cornish 2008). Si l'information propositionnelle s'articule le plus fréquemment en une structure divisée entre présupposition et assertion, en contraste, il s'agit ici d'une proposition non-binaire, comportant la mise en relief de la phrase entière.

Notons que dans la conception initiale de la notion, les phrases thétiques étaient considérées comme une catégorie à part, qu'on ne peut pas analyser en termes de structure informationnelle (Lambrecht 1994 : 139). Cependant, comme l'explique Sasse (2006), le traitement de l'opposition thétique/catégorique comme un phénomène cognitif indépendant est généralement dépassé (« [t]his strong philosophical viewpoint, which regards thethetic/categorical distinction as a cognitive

phenomenon *sui generis*, reflecting two radically different types of predication, is usually rejected today. », 2006 : 259). Voir aussi Rosengren (1997) : « There is no such thing as athetic or categorical clause, but there are clauses that – because of their specific informational structure (being all-focused and all-comment) and the “right” meaning of the predicates, sometimes taking help from the cotext – unambiguously give rise to athetic or categorical reading. ».

A partir des travaux fondateurs de Halliday (1967) et de Chafe (1976), *inter alia*³³, une approche discursive-pragmatique s’est développée, selon laquelle c’est sur des catégories pragmatiques, et non logiques, qu’il faut s’appuyer pour expliquer la nature et le fonctionnement des phrases thétiques.

Une approche purement logico-sémantique ne rend pas compte du fait que l’opposition entre les deux types de jugement, catégorique et thétique, ne correspond pas nécessairement à une distinction linguistique formelle. En effet, des structures « canoniques » peuvent renvoyer à un contenu propositionnel non prédicatif. Lambrecht (1994 : 141) prend l’exemple de l’énoncé thétique de l’anglais *it is raining*, ayant la même structure formelle que tout autre énoncé incluant le pronom *it* au sens référentiel plein. Voir aussi Halliday (2004 : 259) : « In meteorological processes like *it’s raining* [...] the *it* serves the interpersonal function of Subject, liker the *there* in an ‘existential’ clause, but has no function in transitivity – if you are told that it’s raining, you cannot ask *What is?* Ant the *it* cannot be theme-predicated (we cannot say *it’s it that’s raining*) ».

Similairement, l’exemple suivant cité par Sasse (2006) est considéré logiquement comme « thétique » sans avoir aucune marque formelle spécifique (ex. la structure en *there* ou l’accentuation uniquement sur le sujet).

(I.28) NOBODY LEFT

(Ladusaw 1994)

Un terme qui semble plus approprié est celui de structure « à focus phrastique » (*sentence-focus* ; Lambrecht 2000b). Kuno (1972) décrit ce genre d’énoncé comme « description neutre » (*neutral description*) ou « sans-thème » (*themeless*). Tout en étant conscient des problématiques liées à cette appellation (comme celles que nous venons d’esquisser), Sasse (2006) choisit de garder le terme de *construction thétique*, étant donné son large emploi dans la littérature linguistique. C’est ce que nous ferons dans notre travail, en entendant par « énoncé thétique » une structure « à focus phrastique », laquelle a donc une manifestation concrète dans la langue en question.

En résumé, les phrases thétiques sont souvent considérées comme ayant une fonction présentative (Cornish 2005), car elles servent à introduire un référent nouveau dans le discours, ou à

³³ Nous renvoyons le lecteur à Lambrecht (1994 : 138) pour un récapitulatif de la littérature.

présenter un événement impliquant à son tour un référent nouveau pour le discours – ou qui est interprété de cette façon grâce au contexte (Lambrecht 2000b : 623). Comme nous le verrons, une distinction est parfois faite entre les énoncés thétiques centrés sur l'événement et ceux centrés sur l'entité. On peut utiliser l'étiquette de « event-central » (orienté vers l'évènement) quand on présente un état des choses et de « entity-central » (orienté vers l'entité) lorsque c'est une nouvelle entité référentielle qui est mise en avant (Sasse 1987). Nous en parlerons dans la section 1.1.3.3. Pour le moment, il convient de décrire l'une des propriétés les plus saillantes des énoncés thétiques, à savoir la détopicalisation du sujet.

1.1.3.2. La détopicalisation du sujet

Selon Sasse (1987) et Lambrecht et Polinsky (1997), le facteur qui compte dans la définition formelle de la catégorie des énoncés à focus phrastique (thétiques) est que l'argument qui, dans l'énoncé à structure canonique correspondante en serait le sujet (ou le « distinguished argument », v. Lambrecht 1994 : 350 note 14) est formellement marqué comme non-topique (Lambrecht 1994 : 145).

Les énoncés à nature thétique sont alors caractérisés par l'absence d'un sujet topique, car le lien entre le référent du sujet et la proposition n'est pas assuré (« the proposition expressed by the sentence is not pragmatically construed as being about the referent of this NP [noun phrase] », Lambrecht 2000b : 627). Une telle détopicalisation s'obtient à travers l'effacement des propriétés grammaticales conventionnellement associées au sujet en tant que topique de la proposition :

« Les langues adopteraient l'une ou l'autre des deux stratégies pour signaler que les sujets de propositions thétiques ne sont pas des topiques : soit les propriétés formelles habituellement associées à des arguments topicaux sont neutralisées (stratégie dite « faible ») ; soit le sujet d'une proposition thétique prend certains traits normalement associés à des objets dans une articulation topique-commentaire (les objets faisant partie, par défaut dans une telle structure, du focus) dans la stratégie dite « forte ». Ces traits relèvent de quatre catégories : *la proéminence prosodique, une position linéaire spécifique, un marquage casuel non-nominatif, et l'absence d'accord grammatical marqué sur le verbe fini.* » (Cornish 2008 : 122).

La manière la plus directe d'obtenir un effet de détopicalisation sur le sujet est donc, pour ces auteurs, celle d'associer à ce dernier les propriétés caractéristiques du constituant rhématique de la phrase. Puisque dans les énoncés ayant une structure informationnelle non marquée, l'objet représente le constituant focal par default, dans le cas d'une proposition thétique, le sujet est souvent détopicalisé en prenant des traits normalement réservés à l'objet, parmi lesquels figure par exemple la position

syntactique postverbale dans les langues SV(O) (discutée en 1.2). En faisant apparaître le SN visé au sein d'une structure biclausale on obtient le même effet. Ces formes sont décrites dans les sections 1.2.1 et 1.2.2 respectivement.

1.1.3.3. Introduction d'entités et introduction d'événements

Comme il a été dit plus haut (1.1.3.1), la distinction entre les énoncés thétiques « centrés sur l'événement » et ceux « centrés sur l'entité » vient de Sasse (1987), qui parle donc en termes de *entity-central* et *event-central thematic sentences*. Lambrecht (1988a) propose l'étiquette *event-reporting* pour designer ce dernier type. Le principe à la base d'une telle distinction est qu'à la fonction présentative propre (la mise en avant d'une entité) s'ajoute celle de « rapporter » un état des choses :

« Eventive all-focus sentences are thus uttered in a context in which something has happened: a loud noise, an expression of horror or pleasure, etc., i.e., a reaction to a contextual event. Such a context may also prompt an explicit question such as *What happened?* or *What was that?* Both questions presuppose the occurrence of an event. This event (together with the current spatio-temporal parameters) constitutes the topic of the so-called all-focus sentence. » (Erteschik-Shir 2007 : 212).

Voir aussi la « fonction interruptive » chez Sasse (2006) discutée relativement à l'ordre V-S en 1.2.1.4. Lambrecht (1988a) fait une distinction entre la fonction présentative et la fonction de rapporter un événement des structures en *il y a* du français :

(I.29) *Y'a Jean qu'a téléphoné*

(I.30) *Y'a le téléphone qui sonne !*

(Lambrecht 1988a : 136-137)

En (I.29) le nom propre *Jean*, bien que hautement accessible (les noms propres ayant par définition une référence unique³⁴) peut paraître dans une construction présentative biclausale en français (à la différence du chinois, comme nous le verrons plus loin) et est nouvellement introduit dans le discours. Une fois introduit, il est disponible pour être utilisé en tant que topique. Ainsi, la construction est un moyen de promouvoir un nouveau référent au statut de topique (*topic-promoting*, Lambrecht 1988a). En revanche, (I.30) dénote un événement dans lequel *le téléphone* n'est qu'une entité nécessairement impliquée.

³⁴ Dans les termes de Chafe (1976 : 39) : « [p]roper nouns are in a sense definite by definition, since they are direct labels for particular referents. »

Cependant, comme Lambrecht même l’admet, il n’est pas toujours aisé de différencier entre ces éléments qui sont assez « importants » pour devenir topiques de ceux qui ne le sont pas. Ainsi, il est possible de voir dans (I.29) une structure centrée sur l’événement, si *Jean* est « mentioned only as a necessary participant in the reported event, i.e. the phone call » (Lambrecht 1988a : 160)³⁵. La différence entre (I.29) et (I.30) ne serait donc pas catégorique (au sens commun du terme). De plus, comme le remarquent Karssenberget Lahousse (2018), cette distinction est problématique à mettre en place lorsqu’on analyse des occurrences authentiques de structure en *il y a ... qui* dans le cadre d’une recherche de corpus.

Or en chinois, la distinction entre les deux sous-types d’énoncé présentatif est opérationnelle, puisque – comme nous le montrerons aux chapitres IV (en ce qui concerne les structures biclausales en *yǒu*) et V (en ce qui concerne l’ordre V-S) – les énoncés présentatifs centrés sur l’entité ont tendance à comporter un nom précédé par la séquence [(numéral +) classificateur] tandis que les énoncés présentatifs centrés sur l’évènement (« événementiels ») ont tendance à comporter des noms nus (voir la discussion portant sur la proéminence thématique du SN nouvellement introduit, en 1.1.1.3). En revanche, les noms propres, et en général, les SN à référence unique, sont généralement introduits en chinois sans le recours à une structure présentative. Considérons les énoncés suivants :

(I.31) (*有)约翰打了电话
 (*Yǒu) Yuēhàn dǎ-le diànhuà
 AVOIR Jean passer-PFV appel
 ‘(*Il y a) Jean (qui) a téléphoné.’

(I.32) (有)电话在響!
 (Yǒu) diànhuà zài xiǎng!
 AVOIR téléphone PROG sonner
 ‘(Il y a) le téléphone (qui) sonne !’

En (I.31), le nom propre *Yuēhàn* – désignant un référent unique, hautement identifiable – exclut l’emploi de la structure présentative en *yǒu*. En (I.32) les deux configurations (S-V et structure en *yǒu*) sont possibles, probablement car le nom *diànhuà* ‘téléphone’ renvoie à un référent dont l’unicité est relative (comme « le train » discuté plus haut). Quoi qu’il en soit, la différence entre les énoncés centrés sur l’entité et ceux centrés sur l’évènement, distinction qui a une manifestation formelle évidente en chinois, ne se reflète pas tout à fait dans des paires de phrases françaises comme (I.29-30).

³⁵ Voir aussi Lambrecht (1988 : 149 et 2001 : 507).

Sasse (2006 : 255) reprend l'exemple de Lambrecht « il y a le téléphone qui sonne » (I.30) et présente sa traduction dans un échantillon de langues (tableau 1-3) ; les lettres majuscules notent le ou les mots sur lequel tombe l'accent phrastique.

Tableau I-3. *L'énoncé thétiqúe centré sur l'événement « Il y a le téléphone qui sonne » dans un échantillon de langues d'après Sasse (2006)*

Anglais :	<i>The PHONE's ringing.</i>
Albanien :	<i>(Po) bie TELEFONI.</i>
Basque :	<i>TELEFONOAK jo du.</i> ³⁶
Bulgare :	<i>TELEFONÁT zvánì.</i>
Français :	<i>Y'a le TELEPHONE qui SONNE.</i>
Allemand :	<i>Das TELEFON klingelt.</i>
Hongrois :	<i>Csöng a TELEFON.</i>
Italien :	<i>Squilla il TELEFONO.</i>
Japonais :	<i>DENWA ga NATTE iru yo.</i>
Grec moderne :	<i>Xtipai to TILEFONO.</i>
Polonais :	<i>TELEFON dzwoni.</i>
Serbo-croate :	<i>Zvoni TELEFON.</i>
Espagnol :	<i>Suena el TELÉFONO.</i>
Irlandais :	<i>Tá an FÓN ag ringáil.</i> ³⁷

Toutes ces configurations, à travers des procédés divers, comme l'accentuation (anglais, bulgare, polonais), le marquage morphologique (basque, japonais), la position postverbale (albanien, hongrois, italien) du sujet ou son insertion au sein d'une construction clivée (français, irlandais), manifestent qu'on s'écarte de la configuration prosodique ou (morpho-)syntaxique canonique.

Nous poursuivons la discussion engagée ici sur les structures présentatives dans les langues du monde dans la section suivante, qui est entièrement consacrée à ce sujet.

1.2. Les structures à fonction présentative : manifestations et propriétés à travers les langues

Nous nous focalisons maintenant sur les structures présentatives, à savoir les formes disponibles dans les langues pour introduire les référents nouveaux. Avant de procéder à la présentation et l'analyse des configurations syntaxiques qui interviennent pour réaliser la fonction d'introduction référentielle

³⁶ Mot-à-mot pour le basque : téléphone-ART-ERG frapper-PFV être-3SG (ART = article ; ERG = cas ergatif)

³⁷ Mot-à-mot pour l'irlandais : EXIST ART téléphone à sonner:NV (NV = nom verbal).

nominal au verbe (résultant dans l'ordre V-S) pour marquer le contraste avec l'ordre des mots canonique sujet-verbe³⁹, le français lui s'appuie (surtout) sur une structure clivée (« split structure » chez Sasse 1987 : 538), où le sujet est introduit par un marqueur existentiel ou déictique (lorsque le nom est introduit par l'opérateur existentiel AVOIR, il a le rôle d'objet). Voici quelques exemples de ce type de structure en français (tirés de Sasse 2006 : 265) :

(I.35) *J'ai le COU qui me fait MAL.*

(I.36) *Il y a un TUYAU qui fuit.*

(I.37) *Qu'est-ce qu'il y a ? – C'est MAMAN qui me bat.*

(I.38) *Voilà la SIRENE qui hurle.*

(I.39) *Nous avons une invitation ce soir, et ma FEMME qui est malade !*

Le choix d'une construction clivée est motivé par le fait que la postposition du sujet nominal au verbe est très contrainte en français oral. Comme l'inversion syntaxique de l'italien et l'inversion prosodique de l'anglais, la clivée permet de distinguer des énoncés ayant le même contenu sémantique, mais qui expriment une proposition catégorique. Les structures présentatives biclausales seront discutées dans la section 1.2.2 d'un point de vue typologique, et nous nous focaliserons ensuite sur celles du français et du chinois qui seront traitées en détail dans le chapitre III et IV respectivement.

D'autres procédés existent pour marquer les énoncés thétiques – et donc le contraste avec l'énoncé catégorique correspondant, citons les suivants (Sasse 2006 : 265-266) :

- L'incorporation du sujet dans le verbe : celui qui serait le sujet de l'énoncé catégorique correspondant est incorporé dans le verbe dans le cas d'un énoncé thétique⁴⁰ ;
- La nominalisation du verbe, qui est généralement précédé par des marqueurs existentiels ;
- L'emploi des particules, comme *wa* et *ga* en japonais (Kuroda 1972)

Bien sûr, les langues peuvent utiliser plusieurs des stratégies présentées ici pour signaler le contraste avec la structure informationnelle canonique, i.e. celle qui se déroule en présupposition-assertion. En anglais, comme il a été dit, la structure informationnelle est principalement marquée par

³⁹ L'ordre des mots en italien est peu contraint, ce qui a poussé certains auteurs à le considérer comme une « langue à inversion libre » (Burzio 1986). Une telle conception est sans doute exagérée, et des études conduites sur la base de corpus (Bernini 1995, Sornicola 1995) remettent quand même l'ordre S-V à sa place de modèle le plus fréquent en italien, et cela non seulement en ce qui concerne les verbes transitifs, mais aussi les intransitifs.

⁴⁰ Plutôt qu'une incorporation morphologique au sens strict, laquelle est restreinte à un petit nombre de langues et à des contextes précis, le procédé le plus couramment mobilisé relève d'une pseudo-incorporation, où le SN garde en partie son autonomie (Sasse 2006 : 265).

des stratégies prosodiques ; or, même si les procédés d'ordre syntaxique sont marginaux, cela ne veut pas dire qu'ils n'existent pas dans la langue, à l'inverse ils ont engendré une multitude d'études :

« At least two [syntactic] constructions can be regarded as being more or less conventionally used to encode presentative utterances in English: (i) sentences introduced by the 'existential formula' *there is/are* or *there's*, and (ii) the syntactic rearrangement rule of 'main verb inversion', often alternatively called '(locative) inversion' » (Gast et Haas 2011).

La littérature sur les phrases en *there* de l'anglais est immense, nous ne présentons ici que rapidement leurs caractéristiques. A partir de Jespersen (1924) cité en 1.2.2.1, nommons ici : Aissen (1975), Rando et Napoli (1978), Lakoff (1987, 1999), Lumsden (1988), Abbott (1993), Ward et Birner (1993), Lazard (1994), McNally (1997), Lyons (1999). On distingue généralement les phrases en *there* déictiques (I.40), existentielles (I.41) et présentatives (I.42) :

(I.40) *There is Harry* (Lakoff 2019 : 92)

(I.41) *There is a car in the backyard* (Lakoff 2019 : 92)

(I.42) *There arrived many tourists at the station.* (Bentley *et al.* 2015 : 65)

Par exemple, Lyons (1999 : 236-237) fait une distinction en termes de constructions en *there* centrales (I.41) et dérivées (I.42), ces dernières comportant un verbe autre que *be* et dont la fonction, présentative, est celle d'introduire les entités dans le discours. Lakoff (1987) *inter alia* considère la construction présentative en *there* comme une extension métaphorique de la construction déictique en *there*. Contentons-nous de remarquer qu'en anglais, l'ordre V-S requiert un élément explétif (*there*) ou locatif en position préverbale (nous discuterons de l'« inversion locative » en 1.2.1.1).

Pour le moment, revenons à notre discussion générale des procédés disponibles pour exprimer la fonction présentative à travers les langues. Pour donner au lecteur une idée d'ensemble, voici un survol typologique de la distribution de quelques structures marquées dans les langues d'Europe repris de Sasse (2006 : 268).

Dans le tableau I-4 nous avons gardé les étiquettes d'origine, à savoir : SAcc = Accentuation du sujet ; VS = ordre verbe-sujet ; XVS = inversion locative ; Split = structure clivée ; SInc = Incorporation du sujet. Rappelons que Sasse (2006) parle en termes de *phrases thétiques* qu'il entend au sens d'énoncés à structure phrastique (1.1.3.1). Nous avons ajouté la note (c) pour le hongrois dont les remarques se trouvaient dans le texte, nous avons apporté des corrections mineures au tableau d'origine.

Tableau I-4. *Distribution des structures présentatives à travers les langues d'Europe (d'après Sasse 2006)*

Langue	SAcc	VS	XVS	Split	SInc
Albanien	marginal	+	+	–	–
Basque	+	marginal	+	–	–
Bulgare	+	+	+	–	–
Catalan	+	+	+	marginal	–
Tchèque	+	marginal	+	–	–
Danois	+	–	+	–	+
Néerlandais	+	–	+	–	–
Anglais	+	–	marginal	–	–
Français	+	+ ^a	+	+	–
Allemand					
standard	+	–	+	–	–
vernaculaire	+	+	+	–	–
Hongrois	* ^c	+	+	–	* ^c
Italien	marginal	+	+	marginal	–
Latin	?	+	+	–	–
Grec moderne	marginal	+	+	–	–
Polonais	+	marginal	+	–	–
Portugais	+	+	+	marginal	–
Langues rhéto-romanes	+	marginal	+	marginal	–
Roumain	marginal	+	+	–	–
Russe	marginal	+	+	–	–
Serbo-croate	marginal	+	+	–	–
Espagnol	+	+	+	marginal	–
Turc	+	–	–	–	–
Gallois/Irlandais	–	– ^b	–	+	–

^a Seulement dans le registre écrit.

^b Dans les langues celtiques insulaires, VS est l'ordre « non marqué ». Il n'est normalement pas utilisé pour exprimer des phrases « thétiques / à focus prastique » etc.

^c L'accentuation du sujet (SAcc) en Hongrois a un statut particulier dans le système morphologique de la langue, puisque l'emplacement focal est toujours préverbal. De plus, cette forme est généralement décrite dans la littérature comme un cas d'incorporation du sujet (SInc). Voir Sasse (2006 : 268) pour les détails.

Dans l'échantillon de langues présenté dans le tableau I-4, la structure clivée représente un procédé dominant seulement en irlandais et en français, donc dans les langues d'Europe il ne s'agit pas d'un moyen privilégié pour introduire les entités dans le discours. En italien, par exemple, on trouve la forme biclausale *c'è...che* (correspondant à la tournure du français *il y a...qui*) mais son usage est en forte concurrence avec l'inversion V-S, et c'est cette dernière qui se taille la part du lion. A titre d'exemple, l'extrait oral suivant décrivant les problématiques liées au commencement des cours à l'école, contient 4 occurrences d'inversion V-S et une occurrence de structure biclausale :

(I.43) *quando circa mezza classe è a posto [comincia la lezione] poi poi dopo un po' [arrivano i ritardatari] [...] eh sempre gli stessi poi quando tutta la classe è a posto [comincia la lezione] tra parentesi si fa per dire dopo poco ci sono [c'è gente vuole **che** chiede di andare in bagno] i professori a un certo eh fino a un certo punto dicono di sì poi quando quelli che vogliono*

quando quelli che quando la gente vuole uscire diventa troppa cominciano a # a dire di no perché pensano si vada a fare giratine per la scuola # poi # ahah # poi [suona la campanella della ricreazione] e tutti escono come alla alla alla carica di centouno

Lit : ‘quand à peu près la moitié de la classe est en ordre commence le cours puis puis après un certain temps arrivent les retardataires [...] toujours les mêmes hein, puis quand toute la classe est en ordre commence le cours, entre guillemets, on dit pour dire : peu après il y a des gens qui demandent d’aller aux toilettes, les professeurs jusqu’à un certain moment disent oui puis quand ceux qui veulent.. quand ceux qui.. quand ceux qui veulent sortir sont trop nombreux, [les professeurs] commencent à à dire non car [ils] pensent qu’on va faire des promenades dans l’école, puis.. haha, puis sonne la cloche de récréation et tout le monde sort comme dans les 101 dalmatiens’ (LIP⁴¹)

En français, comme il a été dit, l’ordre V-S est très marginal dans le registre oral, toutefois, nous verrons qu’il n’est pas totalement exclu (chapitre V – section 5.3.4). Dans cette langue, le procédé le plus fréquemment exploité est celui des structures présentatives biclausales.

Dans un certain nombre de langues (ex. le danois, le néerlandais, l’anglais, l’allemand), l’ordre XVS, c’est-à-dire l’inversion locative, est plus fréquemment utilisée que l’ordre V-S (l’inversion absolue), et observé même dans des langues ne disposant pas de cette dernière stratégie, comme l’anglais. Pour le français, Sasse note que l’inversion absolue est observée juste dans le registre écrit ; or il semble bien que cette contrainte de registre s’applique aux deux types d’inversion.

En somme, nous avons discuté des principales structures disponibles dans les langues pour signaler qu’on s’écarte de la structure informative canonique, mobilisées lorsqu’on souhaite introduire une entité ou un nouvel état des choses dans le discours. Parmi les procédés présentés dans cette section, nous donnons une importance particulière à ceux se situant au niveau de la disposition syntaxique, car il s’agit de stratégies sur lesquelles le français et le chinois s’appuient fréquemment. Si la postposition du sujet nominal au verbe n’est qu’un procédé marginal en français oral, cette configuration est fréquemment utilisée en chinois, aussi bien dans le registre oral qu’écrit (avec des différences qu’on prendra le soin de préciser). En revanche, la structure biclausale, où le SN dénotant l’entité en question est introduit par l’opérateur existentiel AVOIR, représente le procédé d’introduction référentielle de loin le plus répandu dans les deux langues.

Dans les deux sections suivantes, nous situons tout d’abord ces procédés – la postposition du sujet nominal au verbe (1.2.1) et les structures présentatives biclausales (1.2.2) – dans la réflexion en linguistique générale, car on estime en effet qu’il s’agit de tournures dont la portée dépasse le cadre d’une langue spécifique. La deuxième partie de la thèse portera sur la description détaillée de ces

⁴¹ Les exemples de l’italien marqués par le sigle LIP proviennent de notre recherche dans le corpus du Lexique de fréquence de l’Italien Parlé (Corpus LIP, consultable en ligne : <http://badip.uni-graz.at/it/cerca>), constitué en 1990-1992 par un groupe de linguistes dirigé par Tullio De Mauro et fut utilisé pour rédiger, en collaboration avec la Fondation IBM Italia, le premier lexique de fréquence de l’italien oral (De Mauro, Mancini, Vedovelli, Voghera, 1993). Le corpus comprend 469 textes, composés approximativement de 490.000 mots, qui ont été recueillis dans quatre villes (Milan, Rome, Florence et Naples).

formes telles qu'elles sont employées dans chacune des langues visées dans notre étude, le français et le chinois.

1.2.1. La postposition du sujet nominal au verbe (V-S)

La prétendue « inversion » syntaxique comporte l'occurrence du sujet en position postverbale – l'emplacement normalement réservé à l'objet dans les langues à ordre dominant SV(O)⁴². Il s'agit de l'un des cas les plus manifestes de configuration à focus phrastique. Dans ce genre de structure, caractérisée par une absence d'articulation présupposition-assertion, le domaine de l'assertion s'étend à la fois sur le prédicat et sur le sujet :

« Constructions exhibiting this [V-S] word order phenomenon have been traditionally referred to as subject-verb inversion constructions. Such constructions are found in a great number of languages » (Lambrecht 2000b : 633).

Lambrecht (2000b) considère la construction existentielle ainsi que l'inversion locative comme des sous-types de constructions à focus phrastique, où le focus tombe à la fois sur l'argument et sur le prédicat :

« A subtype of SF [sentence-focus] construction is the so-called 'existential sentence' involving verbs of existence, especially the verb *be*. Another subtype, which has been the focus of much attention in recent linguistic theorizing, is the one referred to as 'Locative Inversion' » (Lambrecht 2000b : 617–18).

Comme nous l'avons dit (en 1.2), l'anglais est une langue s'appuyant principalement sur la prosodie pour marquer la structure informationnelle de l'énoncé et du discours. La postposition du sujet nominal au verbe est très contrainte et relève pour l'essentiel de l'inversion locative. En effet, un élément autre que le sujet doit être présent en position préverbale :

« English basically has a rigid SVO order. Its potential of word order variation for discourse-pragmatic purposes is severely restricted. Mechanisms of exbraciation, left-dislocation, topic fronting are present, but limited. It therefore makes extensive use of cleft constructions and intonational devices. As far as other types of discourse-pragmatically exploitable word order constellations are concerned, **VS is**

⁴² Comme l'explique Meulleman (2012 : 54), « [l']inversion du sujet nominal est un phénomène complexe qui recouvre des constructions assez différentes selon les langues ». Nous nous intéressons ici aux constructions comportant un sujet nominal postverbal qui expriment un focus phrastique. En effet, dans des langues comme l'italien et l'espagnol, l'expression d'un focus contrastif permet virtuellement à n'importe quel verbe d'apparaître au sein d'une configuration V-S.

possible to a certain extent, provided that the preverbal position is filled with an element other than the subject (i.e., it is always XVS). Except for the fossilized cases of inversion [...], there are only few types of VS constructions available. Some of them are relevant in thethetic domain. One is the well-known *there is* construction, which, in turn, is subject to lexical restrictions by virtue of its existential/deictic character. » (Sasse 2006 : 275, notre emphase).

Malgré cela, c'est sur l'anglais qu'ont débuté les études portant sur la postposition du sujet, et notamment en rapport à ce qu'on appelle parfois la « construction existentielle » (voir la section 1.2.2.2). Ainsi, les premières études sur le sujet se sont intéressées aux phrases existentielles de l'anglais, notamment aux structures en *there* et à l'inversion locative, qui étaient analysées comme une dérivation à partir d'une structure canonique. Suite à la formulation de la célèbre hypothèse inaccusative par Perlmutter (1978), la position postverbale du sujet a été associée à une classe de prédicats : les verbes inaccusatifs.

Ainsi l'italien, présentant une forte corrélation entre la susdite classe de verbes et une séquence linéaire des constituants non canonique, a très tôt attiré l'intérêt des grammairiens qui y ont consacré des nombreuses réflexions, notamment dans le cadre de la grammaire générative (citons ici Rizzi 1982 et Burzio 1986). Les tournures en V-S y sont vues comme motivées par la structure profonde du verbe, dont le sujet postverbal serait en réalité un objet (cette approche est présentée dans la section suivante). En même temps, les phrases présentatives biclausales (section 1.2.2) introduites par la formule existentielle *c'è* 'y est' ont été ignorées dans la plupart des travaux, car considérées comme des tournures relevant du registre oral et d'intérêt mineur (avec la notable exception de Berruto 1986⁴³).

D'autre part, une autre approche s'est développée qui analysait la variation de l'ordre des constituants en rapport à la structure de l'information (à partir des études de Halliday 1967 et Chafe 1976 cités plus haut). Tandis que l'on commençait à remettre en cause la pertinence de la position postverbale du sujet comme diagnostique pour identifier cette classe de verbes (voir Bentley 2006 : 363–402 pour l'italien), on a souligné l'idée d'un ordre non canonique des constituants motivé non par la valence verbale mais plutôt par des nécessités d'ordre pragmatique.

Dans la section qui suit, nous présentons l'analyse du positionnement postverbal du sujet dans le cadre de l'hypothèse inaccusative, car un grand nombre d'études ont été conduites en suivant cette approche. L'avantage d'y intégrer une approche fonctionnaliste est également discutée.

⁴³ Depuis, un certain nombre de travaux se sont intéressées aux structures en *c'è* de l'italien (Berretta 1995, De Cesare 2007, Sornicola 2010, Marzo et Crocco 2015, Karssenberget al. 2017, Karssenberget al. 2018b : 217-254). Cependant, comme le remarquent Marzo et Crocco (2015), elles sont toujours relativement peu étudiées. En effet, elles ont reçu moins d'attention que les clivées en *il y a* du français mais aussi lorsqu'on les compare aux clivées en *è* de l'italien – l'équivalent des clivées en *c'est* du français (Karssenberget al. 2018b : 219).

1.2.1.1. L'ordre V-S dans le cadre de l'hypothèse inaccusative

La postposition du sujet au verbe intransitif est l'un des comportements syntaxiques liés au phénomène dit de l'« intransitivité scindée » (*split intransitivity*). A partir de l'hypothèse sur l'inaccusativité formulée par Perlmutter (1978), on s'est rendu compte que la classe des verbes intransitifs n'est pas homogène, mais que certains d'entre eux ont un comportement particulier. Cette hypothèse y reconnaît deux groupes : les verbes *inaccusatifs* et les verbes *inergatifs*. La terminologie est motivée par l'idée que le premier type de verbes attribuerait au sujet le cas accusatif, si la langue marquait morphologiquement les cas, tandis que le sujet du deuxième aurait le cas ergatif.

Ce que l'on veut remarquer est que le seul argument d'un verbe inaccusatif ne se conduit pas à la manière du sujet d'un verbe transitif, comme on peut s'y attendre dans un alignement accusatif, et comme le fait, d'ailleurs, l'objet des verbes inergatifs. L'argument des inaccusatifs, au contraire, partage plusieurs propriétés avec l'objet direct d'un verbe transitif, ou mieux, avec le sujet grammatical du passif correspondant. Le terme de « inaccusatif », attribué à Geoffrey K. Pullum⁴⁴, a été introduit dans le débat linguistique par Perlmutter en 1978. Selon la formulation originale de l'hypothèse, le sujet des verbes inaccusatifs est un objet en structure profonde, auquel est ensuite attribué le rôle de sujet, en raison de la tendance générale des langues à ne pas laisser cette case vide. Le sujet des inergatifs, en revanche, est un sujet « plein », au niveau profond, qui apparaît en surface en tant que tel. Williams (1981) parle en termes d'emplacement des arguments du verbe : les transitifs et les inergatifs ont un argument externe, alors que le sujet des inaccusatifs, de la même manière que l'objet des transitifs, est un argument interne.

Dans ce cadre théorique, le phénomène de la postposition du sujet est interprété comme une manifestation de sa nature profonde d'objet, schématisé comme ci-dessous (PV dénote le prédicat verbal) : dans le cas des inaccusatifs, le syntagme nominal serait incorporé au sein du prédicat verbal.

Tableau I-5. Schéma sur le sujet postverbal dans l'analyse inaccusative (d'après Legendre et Sorace 2010 : 171)

Verbes inergatifs : SN [PV V]
Verbes inaccusatifs : ___[PV V SN]

Parmi les autres traits qui distinguent les inaccusatifs des inergatifs, on reconnaît la sélection de l'auxiliaire perfectif ESSE dans les temps composés pour les premiers, et de l'auxiliaire HABERE

⁴⁴ Perlmutter (1986 : 186, note 4).

pour les seconds, au sein des langues qui ont le choix⁴⁵. La sélection de l'auxiliaire HABERE est une caractéristique que les inergatifs partagent avec les verbes transitifs. L'italien, en ce sens, présente une correspondance plus nette entre la sélection de l'auxiliaire et l'inaccusativité du verbe, alors qu'en français le choix est moins révélateur car il y a de nombreux inaccusatifs qui choisissent l'auxiliaire HABERE (ex. *pâlir, rougir*, etc.). Selon Sorace (2000) il s'agit, d'ailleurs, des verbes ayant un degré d'inaccusativité plus faible.

Parallèlement à cette analyse, qui porte sur un niveau strictement syntaxique, et de façon plus ou moins complémentaire à celle-ci, une autre approche, dite sémantique, s'est développée pour expliquer le comportement particulier de certains verbes intransitifs. Il a été observé, en effet, qu'il existe des corrélations multiples entre l'appartenance d'un verbe intransitif à l'une ou l'autre des deux classes et ses propriétés sémantiques et actionnelles (Levin et Rappaport Hovav 2005, chapitre 5). D'un point de vue sémantique, l'intransitivité scindée est codée syntaxiquement, mais déterminée sémantiquement (comme Perlmutter lui-même l'avait suggéré, d'ailleurs). Les inaccusatifs n'ont alors pas nécessairement un « objet profond » comme seul argument du verbe, mais sont surtout analysés comme des verbes qui se caractérisent par un ensemble de propriétés de nature sémantique. C'est leur sémantisme qui, à son tour, en déterminerait le comportement syntaxique, comme la postposition du sujet, le choix de l'auxiliaire ESSE plutôt que HABERE, et ainsi de suite.

Dans cette approche, alors, les verbes inergatifs ont tendance à exprimer des activités intentionnelles (*travailler, marcher, parler* etc.) et ont un sujet agentif qui exerce une activité, et n'est pas affecté par le processus exprimé par le verbe. Inversement les verbes inaccusatifs décrivent généralement un changement d'état brusque et hors du contrôle du référent du sujet (*chuter, guérir, exploser, disparaître, mourir*) ou un état (*rester, être assis*). L'entité n'est pas activement la cause ou le responsable de l'action relatée par le prédicat, mais porte le rôle sémantique d'un sujet qui subit un événement (*undergoer*, Levin et Rappaport Hovav 1995 : 11). On trouve parmi ces verbes une sous-classe de verbes indiquant un changement de position à la suite d'un déplacement (*arriver, venir, entrer, descendre*). Enfin, les inaccusatifs comprennent aussi les verbes intransitifs dits pronominaux, comme *ammalarsi* 'tomber malade', *pentirsi* 'se repentir', *commuoversi* 's'émouvoir', *arrabbiarsi* 's'énervé'. Un exemple prototypique de construction V-S inaccusative est celle observée en italien avec le verbe *rompersi* 'se rompre' :

⁴⁵ A côté de la sélection de l'auxiliaire ESSE, Burzio (1986, chapitre I) qui a lancé le débat portant sur l'ordre des mots et les inaccusatifs (« ergatifs » dans sa terminologie) de l'italien, examine quatre autres paramètres syntaxiques d'inaccusativité : la phrase impersonnelle, le pronom clitique *ne* ('en'), le participe absolu et la proposition relative réduite. Ces diagnostics et leur pertinence sont discutés en détail dans Bentley (2006).

- (I.44) *Si è rotto un bicchiere!*
 se être.PRES.3SG casser.PP.M.SG un verre
 ‘Il y a un verre qui s’est cassé !’

Cette approche sémantique peut être rapprochée de celle de Dixon (1979 : 59) qui prend comme point de départ ce qu’il appelle les *universal semantic-syntactic primitives*, à savoir : S, le seul argument d’un verbe intransitif, A, l’agent d’un verbe transitif, et O, le patient d’un verbe transitif. Les langues accusatives, écrit-il, ont tendance à identifier S avec A, alors que dans les langues ergatives S converge avec O. Certaines langues ont un système mixte, de sorte qu’une partie des verbes intransitifs traitent leur argument unique à la façon de l’agent d’un verbe transitif (S_A), tandis que dans d’autres l’argument est marqué comme le patient (S_O). Parfois ce clivage peut également avoir lieu au sein du même verbe, qui sera alors à la fois inergatif ou inaccusatif, en fonction de la sémantique de chaque phrase. Les verbes *pǎo* 跑 ‘courir / s’en fuir’ et *zǒu* 走 ‘marcher / partir’ du chinois sont un bon exemple :

- (I.45) 他在走路。
Tā zài zǒu lù.
 3SG PROG marcher rue
 ‘Il est en train de marcher.’

- (I.46) 他走了。
Tā zǒu-le.
 3SG marcher-PFV
 ‘Il est parti.’

Dowty (1991 : 572) définit le « proto-agent » et le « proto-patient » comme les arguments du verbe intransitif qui réunissent, virtuellement, la totalité des propriétés typiques de leur rôle. Les attributs suivants caractérisent le proto-agent : le participant accomplit un acte de volonté (VOLONTÉ) ; il perçoit un autre participant / son propre état ou action (SENSIBILITÉ) ; il cause un événement ou un changement d’état chez un autre participant (LIEN DE CAUSALITÉ) ; il est physiquement actif (MOUVEMENT) et existe indépendamment de l’événement nommé par le prédicat (INDEPENDANCE). Le proto-patient, à l’inverse, subit un changement d’état, se trouve en position fixe par rapport au mouvement d’un autre participant, en est causalement affecté, et n’existe pas en dehors de l’événement en question. Enfin, le participant est dans ce cas « affecté de façon incrémentale » au cours de l’événement. Cela signifie que l’aboutissement d’une action télique présuppose et implique qu’il en sera entièrement affecté. Dowty (1991) utilise ces propriétés pour distinguer les deux types d’intransitifs : les verbes inaccusatifs sont alors ceux dont le sujet possède une forte concentration des propriétés typiques du proto-patient :

« The argument for which the predicate entails the greatest number of Proto-Agent properties will be lexicalized as the subject of the predicate; the argument having the greatest number of Proto-Patient entailments will be lexicalized as the direct object. » (Dowty 1991 : 576).

Dans une étude portant sur l'italien, le français, l'allemand et le néerlandais, Sorace (2000) a montré que le comportement inaccusatif se manifeste dans un *continuum* hiérarchique, allant de verbes essentiellement inaccusatifs / inergatifs, qui ont un statut syntaxique défini et prototypique, jusqu'aux verbes les plus périphériques, caractérisés par un comportement variable, et donc moins prédictible, cela aussi bien au sein de la même langue que dans une perspective translinguistique. Plus un verbe est proche de l'un des deux pôles, plus la position de son seul argument dans la structure argumentale, interne ou externe, sera certaine.

Sorace (2000) fonde sa hiérarchie sur des paramètres aspectuels, en identifiant les deux extrêmes qui sont le « changement dynamique télique » à la base de l'inaccusativité, et l'« activité non-dynamique atélique », au cœur de l'inergativité. Par définition, alors, les verbes inaccusatifs sont ceux qui expriment un déplacement (*change of location*), qui sélectionnent l'auxiliaire ESSE, qui sont associées avec le plus haut degré de télicité. A l'autre extrémité se situent les verbes qui expriment un processus contrôlé sans déplacement (*controlled processes, non motional*), sélectionnant HABERE en tant qu'auxiliaire, associés au plus haut degré d'agentivité. Entre un pôle et l'autre se situent les verbes périphériques, qui manifestent une plus grande variabilité dans le choix de l'auxiliaire (surtout en français) ou en rapport à d'autres diagnostics (par exemple, l'apparition dans les structures présentatives en *there* de l'anglais).

Un aspect important qui ressort de l'étude de Sorace est que les conditions minimales requises pour entraîner l'inaccusativité diffèrent selon les langues : en italien « l'existence d'un état » est suffisante pour déterminer un comportement inaccusatif, alors qu'en français la situation sémantique minimale nécessaire est un « changement de position ».

En somme, le comportement inaccusatif/inergatif peut varier selon le type de verbe (plus ou moins typique de sa catégorie), selon les langues (la sélection de l'auxiliaire HABERE étant plus systématique en italien qu'en français par exemple), et plus généralement selon le contexte syntaxique et sémantique. En effet, chaque diagnostic d'inaccusativité présente des nombreuses exceptions, tant au niveau intra- que interlinguistique ; il suffit de consulter la vaste littérature portant sur les « dissonances inaccusatives » (*unaccusative mismatches*, Levin et Rappaport Hovav [1995 : 281-283]), voir Levin et Rappaport 1989, Tsujimura 1989, McClure 1990, Gerds 1991, Cetnarowska 2002, Alexiaou et Schäfer 2010, *inter alia*. D'après les partisans d'une approche purement syntaxique de l'intransitivité scindée, les dissonances inaccusatives invalident l'hypothèse selon laquelle

l'intransitivité scindée est déterminée sémantiquement. A l'inverse, ces dissonances s'expliquent si on considère que l'intransitivité scindée est déterminée sémantiquement, mais que les paramètres sémantiques peuvent varier d'une langue à l'autre (Bentley 2006 : 2).

Pour revenir au paramètre de la position postverbale du sujet, on a remarqué que des verbes typiquement inaccusatifs, comme *impallidire* 'pâlis' en italien (verbe affectant son sujet, lequel n'a pas de contrôle sur l'événement ; sélection de l'auxiliaire ESSE, etc.) ne peuvent pas apparaître en V-S à focus phrastique (Pinto 1997 : 21) :

- (I.47) (Chi è impallidito? / #?Che è successo?)
 ('Qui a pâli ? / #?Qu'est ce qui s'est passé ?)
È impallidito Berlusconi.
 être.PRES.3SG pâlis.PP.M.SG Berlusconi
 'Berlusconi a pâli.'

De même, le verbe de déplacement *andare* 'aller' ne peut pas apparaître dans une structure V-S sans que le sujet postverbal reçoive une interprétation contrastive (« è innaturale creare fuori ogni contesto una frase con *andare* a soggetto postposto, perché è interpretabile solo con verbo tematico e soggetto rematico », Lonzi 1986 : 106).

A l'inverse, on a remarqué que des verbes typiquement ergatifs peuvent apparaître dans des configurations V-S normalement réservées aux inaccusatifs. C'est le cas du verbe *telefonare* en italien (Pinto 1997), à savoir un verbe d'activité, qui comporte la sélection de l'auxiliaire HABERE etc. Comparons les exemples suivants :

- (I.48) *È arrivata Maria.*
 être.PRES.3SG arriver.PP.F.SG Marie
 Lit : '(Ici) est arrivé Marie'
- (I.49) *Ha telefonato Maria*
 avoir.PRES.3SG téléphoner.PP Marie
 Lit : '(Ici / nous) a téléphoné Marie' (Bentley et Cruschina 2018 : 1-2)

C'est ainsi que plusieurs auteurs ont remis en cause la pertinence de la possibilité d'admettre un sujet postverbal en tant que diagnostique d'inaccusativité. Comme le remarquent Bentley *et al.* (2015 : 49), d'un point de vue typologique :

« Several SVO languages [...] resort to **the VS order** in a class of sentences to avoid focus initial SV order. This class **includes sentences with intransitive predicates with a state in their logical structure (so-called unaccusatives [I.48]) or with intransitive predicates which lack such a component in their logical structure but are construed as involving an overt or silent locative/goal**

argument in certain contexts (a subclass of so-called unergatives [I.49]) » (Bentley *et al.* 2015 : 49, notre emphase).

En (I.48) comme en (I.49) une référence déictique spatiale est impliquée : si dans le cas d'*arrivare* le terme du déplacement est compris dans son sémantisme, pour *telefonare* c'est le contexte qui déclenche un ancrage déictique orienté vers le locuteur.

L'aspect joue également un rôle dans l'ancrage déictique de l'énoncé V-S. En grec (I.50), l'aspect perfectif a la même fonction que le locatif préverbal dans la mesure où il améliore l'acceptabilité de la phrase, permettant au verbe inergatif d'apparaître dans une construction V-S. L'exemple de l'italien montre que le verbe *arrivare*, bien qu'inaccusatif, doit être conjugué au passé composé pour pouvoir comporter un sujet postverbal (I.51). En d'autres termes, sans le trait perfectif, ces énoncés sont ressentis comme incomplets.

- (I.50) a. ?? *dulevan kseni ergates*
worked.3PL foreign workers
- b. *dulepsan kseni ergates*
worked.PERF.3PL foreign workers
 'Des travailleurs étrangers ont travaillé (??travaillaient)'
 (Ortega-Santos 2005, cité dans Alexiou 2011)

- (I.51) a. ?? *arrivava Mario*
arriver.IMP.3SG Mario
- b. *è arrivato Mario*
être.PRES.3SG arriver.PP.M Mario
 'Il y a Mario qui est arrivé (??arrivait)' (Calabrese 1992)

Similairement, en italien des verbes comme *partire* 'partir' et *scappare* 's'enfuir' ne reçoivent au présent qu'une interprétation contrastive, une interprétation à focus phrastique demande le trait perfectif (Bentley 2006 : 379).

- (I.52) a. *scappa il gatto*
fuir.PRES.3SG le chat
 'C'est le chat qui s'échappe'
- b. *è scappato il gatto*
être.PRES.3SG fuir.PP.M le chat
 'Il y a le chat qui s'est échappé' (Bentley 2006 : 379)

Comme nous le verrons (1.2.1.1), en chinois également l'acceptabilité des énoncés V-S est étroitement liée à la sélection de la particule aspectuelle.

En règle générale, le rôle sémantique du sujet d'une phrase thétiq ue n'est pas un agent au sens strict (un « proto-agent » chez Dowty 1991 présenté plus haut). Comme le souligne Bentley (2006 :

390) à propos de V-S en italien, notons cependant que cette contrainte d'agentivité ne concerne pas le prédicat, qui peut être agentif (comme *telefonare* 'téléphoner', *bussare* 'frapper' plus haut) mais l'interprétation globale de l'énoncé⁴⁶. Dans ce type de phrases, le référent du sujet n'est pas conceptualisé comme activement impliqué dans une certaine situation, mais se présente en tant qu'une entité qui tout simplement apparaît dans la scène du discours (Lambrecht 2000b : 623-624). Comme le remarquait déjà Lonzi (1986 : 106) :

« Caratteristica della frase eventiva è di non assegnare un ruolo semantico di agente : il contenuto proposizionale è ininterpretato dal punto di vista agentivo, è la registrazione di un fenomeno. Chi enuncia una frase eventiva può solo prendere posizione sulla sua causa. »⁴⁷ (Lonzi 1986 : 106).

Ainsi, une multiplicité de facteurs entrent en jeu pour déterminer les prédicats qui permettent des structures V-S : le sémantisme lexical du verbe, sa capacité d'ancrer l'énoncé à l'« ici-et-maintenant » de l'énonciation, la monoargumentalité⁴⁸, l'aspect perfectif, etc. En fin de comptes, le sémantisme et l'Aktionsart du verbe doivent être cohérents avec l'articulation informationnelle de l'énoncé. Cela explique pourquoi les verbes pouvant apparaître dans des contextes V-S ne sont pas limités aux inaccusatifs :

« There is no doubt a relationship between the class of predicates permitting SF [sentence-focus] construal and the so-called class of 'unaccusative' [...] predicates, but the two are not coextensive [...]. The class of predicates found in SF constructions is much larger than has been assumed by most linguists who have dealt with the focus structure of such sentences. In fact, it seems to be an open class. » (Lambrecht 2000b : 623)

En somme, une approche intégrant la structure informationnelle de l'énoncé nous paraît plus apte à expliquer les phénomènes relatifs à la variation de l'ordre des constituants. Une telle approche,

⁴⁶ « I argued that the principal factor in the licensing of non-contrastive VS order with bare activity predicates is whether they can encode new events. This type of encoding depends on the possibility of a telic reading, the entailment of a topic, and the temporal or aspectual specifications of the predicate. Bare VS order is allowed with *telefonare* 'phone', *suonare* 'ring', *bussare* 'knock', and *chiamare* 'call' because, in given contexts, these verbs satisfy the mentioned criteria. Since these verbs encode agentive activities, agentivity is not the crucial factor in the licensing of non-contrastive VS order with bare predicates. Agentivity has, however, turned out to be more important in the licensing of locative inversion. » (Bentley 2006 : 390).

⁴⁷ « La particularité de la phrase événementielle est qu'elle n'attribue pas de rôle sémantique d'agent : le contenu propositionnel n'est pas interprété du point de vue agentif, il consiste dans le compte-rendu d'un phénomène. Qui prononce une phrase éventuelle ne peut prendre position que sur sa cause. » (Lonzi 1986 : 106).

⁴⁸ Comme le remarquent Bentley et Cruschina (2018 : 3, note 3), en italien l'ordre SVO basique est observé dans des configurations à focus phrastique comportant un verbe transitif :

(Che è successo?) – *I bambini (S) hanno trovato (V) una volpe (O) (in giardino).*
'(Que se passe-t-il ?) – Les enfants ont trouvé un renard (dans le jardin).'

en outre, permet de mettre en relation des structures formellement distinctes, telles que l'ordre V-S et les structures présentatives biclausales, mais qui possèdent la même fonctionnalité aux niveaux intra- et interlinguistique – avec toutes les précautions que demande une approche contrastive. Au cours de notre thèse nous utilisons les termes de *inaccusatif* et *inergatif* pour désigner une typologie de verbes, que nous entendons d'un point de vue avant tout sémantique.

Notons enfin que si des langues comme l'italien présentent un ensemble de critères morphosyntaxiques qui peuvent être étudiés pour déterminer l'inaccusativité du verbe (le choix de l'auxiliaire, la possibilité d'admettre la forme en *si* et la cliticisation avec *ne* 'en', l'accord du participe passé, et enfin l'apparition dans une phrase V-S, voir Bentley 2006 pour une analyse détaillée de ces diagnostics et de leur pertinence), en chinois, d'un autre côté, ces tests sont d'ordre syntaxique, et l'ordre des mots y est considéré comme un diagnostic-clé. Yáng (1999) en recense trois : l'apparition au sein de l'inversion absolue (V-S) ; l'apparition au sein d'une phrase météorologique ; l'alternance transitive/causative⁴⁹. Nous parlerons de l'inaccusativité discutée à propos de l'ordre V-S en chinois dans la section 5.1.2 du chapitre V.

1.2.1.2. Inversion absolue et inversion locative

On distingue généralement deux types d'inversion du sujet nominal : l'inversion absolue et l'inversion locative. L'inversion absolue (pour le français voir Lahousse [2011] ; pour l'italien voir Bernini [1995 : 55] et Sornicola [1995, 2010]) se caractérise par le fait qu'elle ne comporte pas d'élément locatif antéposé au verbe. Ainsi, ce dernier peut paraître en tête d'énoncé, comme dans l'exemple italien suivant :

(I.53) *sono arrivati due poliziotti* (LIP)
 être.PRES.3PL arriver.PP.M.PL deux policiers
 Lit : 'Sont arrivés deux policiers'

A l'inverse, dans l'inversion dite locative, le verbe est précédé d'un syntagme adverbial dénotant la référence spatio-temporelle (Marandin 2003 : 373). En anglais il s'agit d'un procédé qu'on trouve surtout dans le registre écrit :

(I.54) *In the distance appeared the spires of a town.* (Levin et Rappaport Hovav 1995: 218)

⁴⁹ Pour des contraintes de longueur nous n'entrons pas ici dans les détails des critères étudiés en chinois pour déterminer l'inaccusativité du verbe et renvoyons volontiers le lecteur à Aldridge (2015) qui fournit un récapitulatif détaillé de la littérature sur la question.

Webelhuth (2011) s'intéresse à l'inversion locative de l'anglais du point de vue de la grammaire constructionnelle de Goldberg (2006), qu'il voit comme une construction dérivée de l'inversion déictique (I.55) où s'opère un changement de perspective :

« As in the use of Deictic Inversion, the speaker wants the hearer to construct a spatial mental model in two steps in Locative Inversion. But whereas in the first plan, the hearer builds a mental model of what she herself is seeing, **in Locative inversion she builds a model of what the speaker claims an observer is seeing, where the observer needs to be inferred from the discourse context.** » (Webelhuth 2011).

(I.55) *There's Harry with his red hat on* (Lakoff 1987 : 468)

Il remarque que l'inversion locative apparaît dans des contextes discursifs précis (comme les rapports rétrospectifs de témoins oculaires, les commentaires d'événements sportifs, les descriptions d'appartements, les indications d'itinéraire, chez les guides touristiques et dans des situations narratives scéniques) qui présument toutes la présence d'un observateur.

En italien, l'inversion locative, comme l'absolue, est fréquente dans tous les registres de langue. En voici un exemple :

(I.56) *alle tre viene la estetista* (LIP)
à:les trois venir.PRES.3SG la esthéticienne
'A trois heures il y a l'esthéticienne qui vient'

Notons au passage que dans cette langue le complément de lieu peut se positionner ailleurs dans la phrase :

(I.57) *è venuto un bambinetto in negozio* (LIP)
est venir.PP.M un gamin dans magasin
'Il y a un petit garçon qui est venu au magasin'

(I.58) *proprio dieci giorni fa è venuto un signore da me* (LIP)
juste dix jours plus.tôt est venir.PP.M un monsieur chez moi
'Il y a juste dix jours un monsieur est venu chez moi'

En français, contrairement à d'autres langues romanes comme l'italien et l'espagnol, l'inversion absolue est très contrainte, surtout dans le registre oral.

(I.59) *pendant l'collège on est tous ensemble on est tous copains et arrive un moment euh "toi c'est bon euh j'peux plus marcher avec toi dans la rue"* (Français parlé, CFPP2000⁵⁰)

(I.60) *Des oiseaux passaient, plongeant parfois. On voyait d'eux un ventre duveté. Vint le moment où de nouveau elles sentirent le froid, le poids de l'eau, et que leurs bras peinaient à la jeter en arrière, leurs doigts à rester unis dans l'effort.* (Français écrit, Lahousse 2003 : 125)

(I.61) *Apparut le patron, qui essaya d'apaiser les choses* (Français écrit, Fuchs 2013)

L'inversion locative⁵¹ est également un phénomène typique du registre écrit en français. L'extrait suivant en contient deux exemples.

(I.62) *Le lendemain, ma journée commence par un entretien avec les deux secrétaires, en compagnie de Sandrine. Au mur est affiché notre nouveau logo, grand format. C'est une réalisation de l'atelier Pastel pour les 50 ans du CAES : une réussite. Sur l'écran d'ordinateur défilent de belles images.* (Fuchs 2017)

En chinois, les deux types d'inversion du sujet nominal, absolue et locative, existent, et sont fréquemment employés à l'écrit comme à l'oral. (I.63) est un exemple d'inversion absolue (*lingwài* 'en plus' étant une conjonction) et (I.64) est un exemple d'inversion locative comportant le nom locatif *wōmen xuéxiào* 'notre école' en position préverbale.

(I.63) 另外过来一小伙子 (BJKY)
Lìngwài guò-lai yí xiǎohuǒzi
en.plus passer-venir un.CL garçon
'En plus, il y a un garçon qui est arrivé'

(I.64) 那时候我们学校来了十个人 (BJKY)
Nà shíhòu wōmen xuéxiào lái-le shí-ge rén
DEM_{dist} moment IPL école venir-PFV dix-CL personne
'A ce moment-là, dix personnes sont arrivées dans notre école'

Or, une clarification s'impose à ce point de la discussion. Prenons la perspective des verbes qui présentent souvent un sujet postverbal, quel que soit le type d'inversion. A travers les langues, les verbes inaccusatifs (typiquement ceux dénotant le déplacement, l'apparition, l'existence) admettent l'inversion absolue, à laquelle toutefois on peut toujours ajouter un nom locatif préverbal (comme

⁵⁰ Les exemples du français marqués par le sigle CFPP2000 proviennent de notre recherche dans le corpus de français parlé « Discours sur la ville », le Corpus de Français Parlé Parisien des années 2000 (CFPP2000), consultable à l'adresse <http://cfpp2000.univ-paris3.fr/>. A titre indicatif, une recherche sur le verbe *arriver* nous a donné seulement deux occurrences de configuration V-S, dont celle présentée en (I.59).

⁵¹ Sur l'inversion locative en français citons, parmi les autres, les études de Marandin 1997, Cornish 2001, Fuchs 2013, et notamment la monographie de Lahousse (2011 : 63-73) citée précédemment, qui considère l'inversion absolue comme un sous-type d'inversion locative (voir le chapitre V § 5.5.1).

dans l'exemple [I.64] du chinois qui pourrait consister uniquement de la séquence V-S *lái-le shí-ge rén* lit.: 'sont arrivées dix personnes').

A l'inverse, il a été observé que généralement les verbes inergatifs (dénnotant des activités) ne peuvent pas comporter un sujet postverbal sans qu'une expression locative soit antéposée. Ces verbes apparaissent nécessairement dans une inversion locative :

- (I.65) ??(*In questo palazzo ha vissuto un poeta*
 dans ce bâtiment a vécu un poète
 'Un poète habita dans ce bâtiment' (Pinto 1997 : 68)

Les exemples suivants sont tirés d'Alexiadou (2011) et montrent l'amélioration d'acceptabilité fournie par le locatif préverbal des énoncés comportant un verbe inergatif, en espagnol (I.67), en grec (I.68) et en hébreu (I.69)⁵² :

- (I.66) a. ?? *Corren chicos*
 courent enfants
 b. ***Aqui corren chicos***
 ici courent enfants
 'Des enfants courent ??(ici)'
- (I.67) a. **Epezan pedhia*
 jouaient enfants.NOM
 b. ***Edo pezun pedjia***
 ici jouent enfants
 'Des enfants jouent *(ici)'
- (I.68) a. **Yabad gaman*
 a.travaillé jardinier
 b. ***Yabad kan gaman***
 a.travaillé ici jardinier
 'Le jardinier a travaillé *(ici)'

Notons qu'en grec l'énoncé *irthan pedhia* 'sont arrivés des enfants' est tout à fait possible, contrairement à (I.67a). Il en va de même pour sa traduction espagnole *llegaron chicos*. En d'autres termes, ces verbes (inaccusatifs) permettent une inversion absolue. Rappelons au lecteur (voir la section précédente) que la corrélation entre l'inaccusativité, que nous entendons d'un point de vue

⁵² En hébreu le locatif apparaît après le verbe. Dans les trois exemples nous avons gardé le mot-à-mot de l'article d'origine qui ne faisait pas de découpage en morphèmes.

sémantique, et l'ordre V-S (absolu donc) reflète une *tendance*, bien que forte, due au sémantisme des verbes impliqués, qui doit être compatible avec la fonction discursive portée par l'énoncé. Ainsi, si leur sémantisme le permet, des inergatifs peuvent présenter un sujet postverbal. Outre l'exemple comportant le verbe *telefonare* cité plus haut (I.49), en voici un avec le verbe *bussare* 'frapper (à la porte)' :

- (I.69) *Hanno bussato du sconosciuti.*
 avoir.3PL frapper.PP deux inconnus
 'Il y a deux inconnus qui ont frappé à la porte' (d'après Bentley *at al.* 2015 : 127)

Or en chinois, comme nous le verrons en détail plus tard (chapitre V § 5.1.2), la différence entre les deux types d'inversion se manifeste non seulement par la présence, optionnelle ou nécessaire, du locatif préverbal mais aussi par la sélection du suffixe aspectuel. Notamment, les verbes qui peuvent apparaître dans une inversion absolue sélectionnent le suffixe perfectif *-le*, comme *shēng* 生 'naître' en (I.70), tandis que les verbes qui ne peuvent qu'apparaître dans une inversion locative sélectionnent généralement le suffixe duratif *-zhe* ; c'est le cas de *zhàn* 站 'être debout' en (I.71).

- (I.70) (家里) 生了一个千金
(Jiā=li) shēng-le yí-ge qiānjīn
 maison=dans naître-PFV un-CL fille
 '(A la maison) il y a une fille qui est née' (corpus de l'auteur)

- (I.71) *(门口) 站着一个岗哨
 *(*Ménkǒu) zhàn-zhe yí-ge gǎngshào*
 entrée être.debout-DUR un-CL sentinelle
 '* (A l'entrée) une sentinelle se tenait debout' (d'après Shen 1987 : 197⁵³)

Or si les énoncés du type de (I.70) sont courants dans la langue parlée, les énoncés en V-*zhe* (I.71) sont plus typiques du registre écrit. Puisque notre thèse s'intéresse au registre oral de la langue, nous donnons une importance particulière au premier type de configuration.

En suivant Li (2014a) nous faisons une distinction entre les *constructions existentielles* (incluant des verbes d'existence, posture et localisation) et les *constructions présentatives* (incluant des verbes de déplacement ou d'apparition). En plus, pour éviter la confusion avec la structure présentative (biclausale) en *yǒu* nous appelons ce sous-type de V-S *construction présentative monoclausale*. Se référer à la section 4.1 du chapitre IV pour les détails sur la terminologie que nous employons à l'égard des structures du chinois. Enfin, dans la section 1.2.2.1 du présent chapitre nous

⁵³ Pour simplifier notre exposé, nous avons modifié l'exemple de l'article d'origine qui contenait le nom nu *gǎngshào* 岗哨 'sentinelle'. Nous avons discuté en 1.1.2.3 des différentes interprétations que le nom nu peut recevoir en chinois selon sa position dans l'énoncé.

discutons du rapport entre les structures existentielles (ex. [(SN_G +) *yǒu* + SN_F] en chinois) et les structures présentatives (ex. [(SN_G +) *yǒu* + SN_F + V]).

Tout d'abord, dans ce qui suit, il convient de préciser la valeur de l'élément spatio-temporel qui apparaît typiquement en position initiale de phrase.

1.2.1.3. Le topique scénique

Comme nous l'avons vu, l'ordre V-S est généralement considéré comme une structure à focus phrastique (ou « thétiq ue »), puisqu'elle ne demande qu'une présupposition minimale ou absente et de ce fait peut initier un nouveau segment de discours. Cependant, si elle manque d'un « topique d'à-propos » (*aboutness topic*⁵⁴), cela ne veut pas dire que ce type de phrase n'est « à propos de rien » (Krifka 2008 : 266). C'est ainsi que plusieurs auteurs se sont intéressés au statut discursif de l'élément antéposé au verbe dans l'inversion locative : faut-il le considérer comme un topique ? mais dans ce cas, XVS ne relèverait pas d'une proposition « catégorique »⁵⁵ ?

Certains auteurs considèrent l'élément spatio-temporel qui apparaît souvent en position initiale de phrase comme un topique (Chafe 1976, Erteschik-Shir 1997: 27, Nikolaeva 2001, Lahousse 2003 *inter alia*) qui spécifie le cadre de l'événement dénoté par la proposition. La notion de « cadre » est évoquée dans la terminologie de Charolles (2003) qui appelle ce type d'adverbiaux des *adverbiaux cadratifs*⁵⁶. Lahousse (2003, 2006 : 51-52, 2011) appelle l'élément préverbal dans l'inversion locative un *topique scénique* (en reprenant la notion de « stage topic » chez Erteschik-Shir 1997), à savoir un élément de cadrage qui définit la référence spatio-temporelle dans laquelle l'événement se produit (Lahousse 2011 : 51).

La notion de *topique scénique* réfère à la localisation temporelle et spatiale de l'événement, non seulement dans le cas d'un élément locatif explicitement mentionné et positionné en début d'énoncé. Déjà Gundel (1974) estimait que les énoncés thétiq ues – et donc l'inversion V-S à focus phrastique même lorsqu'un élément locatif n'est pas présent – ont toujours un topique, « la situation particulière (temps et lieu) à propos de laquelle il est asserté » (Erteschik-Shir 2007 : 16). Dans la terminologie d'Erteschik-Shir (2007 : 16-17) il s'agit des paramètres spatio-temporels, définis contextuellement, dans lesquels se déroule l'événement décrit par l'énoncé :

⁵⁴ Défini ainsi par Erteschik-Shir (2007 : 19) : « what the sentence is about; they are given; and the truth-value of the sentence is established with respect to them ». Voir aussi Gundel (1974), Reinhart (1982) et Lambrecht (1994 : 118) pour une conception du topique en termes d'« à propos ».

⁵⁵ Voir Ladusaw (1994) et Meullemans (2012 : 48).

⁵⁶ La notion de « cadre » discutée ici ne doit pas être confondue avec le concept de *cadre cognitif* (en anglais *semantic frame*) traité dans la section 1.1.1.2.

« [T]hetic sentences are viewed as having implicit “stage” topics indicating the spatio-temporal parameters of the sentence (here-and-now of the discourse). These are contextually defined. » (Erteschik-Shir 2007 : 16).

Lahousse (2011) complète le concept initial de « topique scénique » d’Erteschik-Shir (1997) en précisant qu’il s’applique, en français du moins, à tous les circonstants spatiaux ou temporels en position initiale qui ont une valeur de topique, y compris ceux qui se trouvent dans une configuration sans inversion (S-V). De plus, comme il a été dit plus haut, un topique scénique peut être explicite ou implicite. Lorsqu’il est explicite, il apparaît dans l’énoncé en tant qu’élément locatif, qui peut se trouver en position initiale, mais pas nécessairement. Lorsqu’il se trouve positionné ailleurs dans la phrase, il serait alors co-indexé avec un topique scénique vide en tête d’énoncé (les exemples [I.72-73] sont tirés et adaptés de Lahousse 2013 ; TopS signifie « topique scénique »).

(I.72) *En 1914*_{TopS}, la première guerre mondiale a éclaté.

(I.73) [*e*]_{TopS} La première guerre mondiale a éclaté [*en 1914*]_i.

Lorsque le topique scénique est implicite, il peut référer déictiquement au contexte de l’énonciation (pointant à l’« ici-et-maintenant »), comme dans (I.74), ou dépendre du cotexte, comme dans (I.75), où le point de repère spatio-temporel est spécifié par le discours :

(I.74) [*e*]_{TopS} Il pleut.
(= « Il pleut ici et maintenant »)

(I.75) *Nous sommes arrivés en Espagne.* [*e*]_{TopS} Il pleuvait.
(= « Il pleuvait en Espagne, au moment de notre arrivée »)

Contrairement à la notion de topique « classique » (le *topique d’individu* chez Lahousse 2006), la notion de *topique scénique* a l’avantage de « distinguer entre les adverbes temporels, aspectuels et locatifs, d’une part », qui peuvent paraître avec un sujet postverbal, « et les verbes modaux, verbaux et orientés vers le sujet de l’autre », qui ne permettent pas de postposer le sujet en français (Lahousse 2003). Cela permet une approche unifiée de tous ces éléments qui admettent l’inversion, en incluant les *scene-setting external topics* chez Nikolaeva (2001) « qui ne sont pas intégrés syntaxiquement dans la structure phrastique » (Lahousse 2003). De cette manière, puisque le topique scénique peut être implicite, l’inversion absolue et l’inversion locative sont traitées de manière unitaire.

Dans les cas d'inversion absolue, la présence d'un topique scénique implicite serait marquée par la présence d'un élément locatif (non situé en position initiale), ou impliqué par le sens du verbe, ou encore inférable depuis le contexte antérieur comme en (I.76) :

(I.76) [Elle ouvre la porte.]_i [**e_i**]_{TopS} Apparaît la jeune élève, âgée de 18 ans.

Le concept de *topique scénique* a été appliqué au chinois (Arslangul 2019) où il permet notamment de rendre compte de l'élément spatio-temporel qui figure en tête d'énoncé dans le cas de l'inversion XVS.

(I.77) 洞里跑出来一个兔子
dòng=li pǎo-chu-lai yí-ge tùzi
 trou=dans courir-sortir-venir un-CL lapin
 'Du trou sort un lapin.' (d'après Arslangul 2019)

En effet, dans des énoncés comme (I.77), le seul élément qui n'est pas asserté est *dòng=li* 'dans le trou' (puisque'il a été évoqué dans le cotexte) ; l'appellation de *topique scénique* reflète efficacement son aspect présupposé (« topique ») et sa fonction d'ancrage spatial (« scénique »). En plus, contrairement au « topique d'individu », manifestement l'énoncé n'est pas « à propos d'un trou » (le critère d'*aboutness* cité plus haut).

En revanche, la notion de topique scénique implicite est mobilisée par Erteschik-Shir (1997) et Lahousse (2011) pour défendre un traitement unifié de l'inversion locative et absolue, qui n'est pas pertinent en chinois, étant donné que les deux types d'inversion ont des propriétés structurelles, sémantiques et pragmatiques différentes (nous le montrerons dans la section 5.1 du chapitre V).

1.2.1.4. Fonctions réalisées par l'ordre V-S

Dans cette section nous nous appuyons sur la taxonomie détaillée proposée par Sasse (2006 : 280-290)⁵⁷. L'auteur identifie les fonctions suivantes associées à l'ordre V-S à travers son échantillon de langues examinées :

- fonction annonciatrice (annuntiative function)
- fonction introductive (introductive function)
- fonction descriptive (descriptive function)

⁵⁷ Mais voir aussi Sasse (1995) et l'étude de Sornicola (1995).

- fonction interruptive (interruptive function)
- fonction explicatrice (explanative function)
- fonction connective (connective function)

La fonction annonciatrice réfère à des énoncés employés comme des « déclarations soudaines » (*statements out of the blue*), surtout rencontrés dans le style journalistique, dont voici un exemple en serbo-croate :

(I.78) *Poginuo britanski vojnik*
 tué anglais soldat
 ‘Un soldat anglais tué’ (d’après Sasse 2006 : 281)

La fonction introductive concerne la première expression des sujets, qui intervient typiquement en tant que stratégie d’incipit discursif. C’est ce que nous appelons la fonction « présentative », comportant la mise en avant d’une nouvelle entité référentielle. Soulignons que chez Sasse (2006), cette fonction est restreinte à l’introduction des participants dans le discours.

A l’inverse, la fonction descriptive renvoie à ces énoncés qui se trouvent également au début d’un segment du discours, mais qui n’introduisent pas des participants. Leur fonction est de préparer la description de l’arrière-plan :

« There is a second type of utterance often occurring in text-initial position with first mention subjects, namely, scene-setting background descriptions. These do not introduce participants into the discourse. Moreover, in the languages examined, they are frequently not expressed in exactly the same way as participant introductions. This frequent formal difference supports the assumption that **scene-setting is an operation quite different from participant introduction.** » (Sasse 2006 : 284).

Parmi les langues qui permettent l’alternance entre les configurations S-V et V-S, un grand nombre (comme par exemple le grec, l’italien, le hongrois et le roumain) privilégient l’inversion locative plutôt que l’inversion absolue (voir la section 1.2.1.2 pour la différence) dans les séquences descriptives. Nous verrons (dans le chapitre V § 5.1.3) que le chinois se conforme à cette tendance.

La fonction explicative réfère à ces énoncés qui demandent un événement présupposé, c’est-à-dire quelque chose qui est déjà arrivé mais reste à être identifié (« a presupposed event, that is, something which has already happened but remains to be identified »). L’exemple suivant est notre traduction en italien de l’exemple allemand de Sasse (2006 : 287).

(I.79) *E poi ci fu un evento drastico nella mia vita: è nata mia sorella.*
 et puis y fut un événement drastique dans:la ma vie est née ma sœur
 ‘Et après il y eut un événement drastique dans ma vie : la naissance de ma sœur’

La fonction interruptive mérite de s'y attarder. Elle renvoie à ces énoncés, fréquents à travers les langues, qui présentent un événement soudain, inattendu :

« All authors of the VS study agreed that, in their languages, **text-internal VS clauses** (other than cases of grammaticalized inversion) **never continue statements about a topical subject commented on in the preceding sentences**. Typically, these VS clauses evoke a “sudden event” effect which we will call the interruptive function. They are usually preceded by a series of events in a topic chain which is disrupted by a sudden, unexpected new situation. **Data from all our languages converge in their preference (in some cases even obligatoriness) of VS clauses in cases of sudden, interruptive events**. The phone, the alarm clock, or the doorbell ringing, somebody knocking on the door, the door opening, the lights being turned on or off, etc., are notorious instances met over and over again. » (Sasse 2006 : 285, notre emphase).

Dans le cas d'un événement « interruptif », dont un exemple est le téléphone qui sonne de Lambrecht (1988a : 136), V-S est presque obligatoire, souligne Sasse, puisque la même proposition exprimée par une construction avec un sujet topique perdrait son caractère interruptif. C'est pourquoi certaines combinaisons V-S sont figées à travers les langues (ex. l'italien *è saltata la luce* 'la lumière s'est éteinte, il n'y a plus d'électricité').

Enfin, la fonction connective a été discutée par Ulrich (1985 : 284-302, cité dans Sasse 2006 : 294, en tant que fonction « narrative » dans sa terminologie) à propos de la polysémie de l'ordre V-S en roumain. Les énoncés V-S à fonction connective comportent un sujet qui dénote un référent accessible et donc leur fonction n'est pas celle d'introduire un personnage dans le récit mais d'assurer la continuité avec les événements qui précèdent :

« **These VS constructions are characterized by involving a given subject**, usually a main character of the story, or even the speech participants themselves in the form of explicit personal pronouns (I, you). [...] **Episode-introducing VS sentences often do not only indicate a shift in personnel and action, but are presuppositionally somewhat more intimately connected to the preceding text**. This presuppositional tie is not one of referent continuity, but a more complex one **in terms of consequences of the preceding events**. » (Sasse 2006 : 294).

Enfin, comme l'auteur l'explique (*ibid.* 288-290), toutes ces fonctions que l'ordre V-S peut réaliser se chevauchent partiellement. Par exemple, les énoncés descriptifs ont un certain nombre de caractéristiques en commun avec les énoncés interruptifs, notamment car les deux peuvent apparaître à l'intérieur ou au début d'un segment de texte. Les énoncés annonciatifs et les énoncés explicatifs, d'autre part, se situent sur un continuum. L'évènement du téléphone qui sonne, par exemple, est

souvent décrit comme un événement interruptif, mais peut aussi être exprimé en tant qu'annonce ou explication.

1.2.2. Les structures présentatives biclausales

Comme nous l'avons vu plus haut (se référer notamment au tableau I-4), la fonction présentative portée par l'ordre V-S dans certaines langues est assignée dans d'autres aux structures présentatives biclausales. Dans l'échantillon de langues examinées par Sasse (2006), les langues qui utilisent ce type de structure en tant que principal procédé pour structurer l'information sont le français et l'irlandais.

On parle de procédé *dominant* quand les deux configurations (V-S et structure biclausale) coexistent mais l'une des deux est amenée à jouer un rôle nettement plus important dans la langue en question. Comme il a été dit, en italien il existe bien une structure présentative biclausale *c'è ... che* 'il y a ... qui' (Berruto 1986, Sornicola 2010) mais c'est l'ordre V-S qui se taille la part du lion. À l'inverse, en français (parlé) l'ordre V-S n'est que marginalement employé, et son emploi reste très restreint par rapport aux structures présentatives biclausales (1.2.2.2).

En chinois, les deux procédés sont couramment employés, à l'écrit comme à l'oral. Remarquons toutefois que, étant donné les contraintes qui régissent l'emploi de V-S présentatif (notamment en ce qui concerne le sémantisme des verbes impliqués), la structure biclausale en *yǒu* a quand-même un champ d'action plus large.

Dans cette section, nous discutons donc des structures présentatives biclausales. Tout d'abord, nous nous concentrons sur les langues du monde qui ont recours à l'opérateur existentiel AVOIR (en 1.2.2.1), pour après définir le rapport entre les structures présentatives (biclausales) et les structures existentielles (monoclausales) dont elles dérivent (1.2.2.2). Enfin, nous traitons de la « construction relative présentative » en français (1.2.2.3) en tant que regroupement de configurations, pouvant présenter différents introducteurs (dont *il y a*, *c'est*, *voici*, *on voit*), ayant tous en commun la reprise pseudo-relative (*qui...*).

1.2.2.1. L'opérateur existentiel AVOIR

Les structures existentielles dans les langues romanes ont leur origine dans les verbes latins *esse*, *stare* et *habere*. Déjà en latin classique, ces trois verbes servaient à introduire un référent dans le discours : [*stare/esse* + SN_{nominatif}] et [*habere* + SN_{accusatif}]. La séquence (*il*) *y a* du français dérive du

latin *ibi habet*, comme le présentatif *hay* de l'espagnol, qui présente la particule locative uniquement à l'indicatif présent sous forme d'un *-y* enclitique fossilisé (Meulleman 2012 : 28).

Comme le remarque Creissels (2020 : 31) :

« l'espagnol, comme le français, a développé à partir du verbe latin *habere* une construction trans-possessive qui a ensuite donné naissance à une construction trans-poss-existentielle. Mais ensuite, l'espagnol a renouvelé l'expression de la prédication possessive en substituant *tener* au verbe de possession *haber*. Comme *haber* a continué à s'utiliser dans la prédication existentielle, il est devenu de ce fait un prédicatif existentiel spécialisé ».

L'italien disposait anciennement d'une forme existentielle en AVOIR (I.80), qui se retrouve à l'heure actuelle dans certains dialectes régionaux (de Calabre en [I.81] et de Pouilles [I.82]). L'italien standard contemporain emploie une formule existentielle centrée sur le verbe ETRE (*c'è*).

(I.80) *Che sotto l'acqua ha gente che sospira* (Dante, *Divina Commedia, Inferno*. 7,118⁵⁸)

(I.81) *No potimu divorziare: ave li piccinni*
 NEG pouvoir.1PL divorcer AVOIR.3SG les enfants
 'On ne peut pas divorcer : il y a les enfants.' (Bentley *at al.* 2015 : 71)

(I.82) *Guarda: ave la Maria*
 regarder.IMP.3SG AVOIR.3SG la Marie
 'Regarde : il y a (lit : la) Marie.' (Bentley *at al.* 2015 : 186)

D'autres exemples de langues disposant de l'opérateur existentiel AVOIR sont le portugais (I.83), le grec (I.84), le serbo-croate (I.85) :

(I.83) *nesta fruta há muitas sementes*
 en.ce fruit AVOIR.3SG nombreuses graines
 'Dans ce fruit il y a beaucoup de graines.' (Bentley *at al.* 2015 : 8)

(I.84) *Éxi polí kósmo sto kédro*
 AVOIR.3SG beaucoup gens in.DEF centre
 'Il y a beaucoup de gens dans le centre-ville.' (Bentley *at al.* 2015 : 14)

(I.85) *Ima nekih studenata (ovde) koji hoće samo diplomu*
 AVOIR.3SG certains étudiants.GEN ici qui vouloir.3SG seulement diplôme
 'Il y a des étudiants (ici) qui veulent seulement le diplôme' (Bentley *at al.* 2015 : 14)⁵⁹

Dans d'autres aires linguistiques, citons le Khmer (I.86) et le Hmong (I.87):

⁵⁸ Edition Petrocchi/Di Salvo ; cité dans Meulleman (2012 : 29).

⁵⁹ Les exemples (I.84) et (I.85) sont originellement de Gaeta (2013 : 490) et de Hartmann (2008 : 226) cité dans Gaeta (2013 : 491), respectivement.

(I.86) a. *Nɪw psaa nih miən mənuh craən nah*
à marché ce AVOIR personne être.très beaucoup
‘Dans ce marché il y a beaucoup de gens’

b. *Kñom miən prədap krup baep*
1SG AVOIR outil chaque type
‘J’ai tout type d’outils’

(Chappell et Creissels 2019 : 503)

(I.87) a. *Nyob hauv lawv lub vaj, muaj ib tug npua*
à dedans 3PL area jardin AVOIR un CL cochon
‘Dans leur jardin, il y a un cochon’

b. *Nws muaj peb lub ntim hauv nws lub hnab thoom*
3SG AVOIR trois CL bol intérieur 3SG CL épaule sac
‘Elle a trois bols dans son sac.’

(Chappell et Creissels 2019 : 504)⁶⁰

Sans aller jusqu’aux propos de Lyons (1967) selon lequel les phrases possessives, existentielles et locatives relèveraient d’une même structure profonde⁶¹, le lien (lexical ou étymologique selon les cas) entre l’expression de la possession et de l’existence est observable dans plusieurs langues du monde (Clark 1978). Creissels (2020 : 17-41) décrit les différentes options disponibles à travers les langues pour marquer la possession, et propose l’analyse qui suit :

« [L]es constructions possessives comportant un prédicatif de localisation / existence auprès duquel le possesseur est encodé comme la figure dans une prédication de localisation ou d’existence sont très répandues. Mais il n’est pas rare non plus de trouver dans les constructions existentielles un prédicatif qui apparaît aussi dans des constructions possessives alignées sur la construction transitive de base. Deux explications historiques sont envisageables : ou bien la construction existentielle est issue historiquement d’une construction trans-possessive (comme en français), ou bien la construction trans-possessive est issue historiquement d’une prédication de localisation / existence mais a évolué de façon à ce que l’encodage du possesseur et du possédé en vienne à s’aligner avec celui de l’agent et du patient dans la construction transitive de base. » (Creissels, 2020 : 18).

Contrairement à l’opinion diffuse dans la linguistique chinoise qui analyse le SN préverbal comme le sujet de *yǒu* aussi bien existentiel que possessif (voir p.e. Chao 1968 : 727), et *contra* Stassen (2009 : 753–54) qui classe le chinois dans les langues qui expriment la possession en s’appuyant sur une construction de topique, Chappell et Creissels (2019) considèrent que l’emploi possessif de *yǒu* est à distinguer de son usage en tant que prédicateur existentiel :

⁶⁰ Les exemples (I.86) et (I.87) sont originaires de Clark (1989 : 207).

⁶¹ Voir par exemple Heine (1997 : 89) pour une position *contra* l’assimilation des structures existentielles et possessives. Nous n’entrons pas dans les détails de la question ici et renvoyons volontiers le lecteur aux travaux de Creissels cités en bibliographie.

« [T[he possessive use of *yǒu* ‘have’ is not fundamentally an existential construction at all and [...] its possessor NP [noun phrase] acts rather as an argument of a possessive predicator [...]. Put differently, in this new analysis of ours, we neither view the possessor as a dangling topic nor as a kind of animate location in the possessive clauses with *yǒu*, thereby reclassifying it as a Have-Possessive. » (Chappell et Creissels 2019 : 487).

Yǒu est alors « un verbe transitif de possession employé impersonnellement comme prédicat existentiel », c’est-à-dire « un verbe qui a le double statut de verbe d’existence et de verbe de possession, et qui dans son emploi comme verbe de possession assigne les rôles de possesseur et de possédé à deux groupes nominaux construits comme respectivement l’agent et le patient dans la construction transitive de base » (Creissels 2020 : 23). Cet aspect sera davantage discuté dans le chapitre IV. Dans notre travail, nous glosons *yǒu* comme AVOIR, tant dans ses usages possessifs qu’existentiels (présentatifs), en raison de leur lien étymologique mais surtout pour souligner l’affinité lexicale avec la construction en *avoir* du français.⁶²

Il n’est pas sans intérêt de remarquer que chez des apprenants sinophones peu avancés (en l’occurrence d’italien L2), la formule existentielle *c’è* ‘il y a’ (lit. ‘y est’) est parfois employé avec une valeur possessive, selon l’ordre SV ([SN_{POSSESSEUR} + *c’è* + SN_{OBJET POSSEDE}] (Chini 2002) :

(I.88) CH2_IT1⁶³ *io non c’è la televisione*
 1SG NEG y.avoir la télévision
 ‘Moi il n’y a pas de télé’ (adapté de Chini 2002 : 120)

(I.89) CH2_IT1 *mio zio mia zia c’è*
 mon oncle ma tante y.avoir
 ‘Mon oncle ma tante il y a (= ont la télé)’ (adapté de Chini 2002 : 120)

Dans la mesure où cet usage relève vraisemblablement d’un calque de la langue source (Chini 2002 : 120), il suggère tout de même la perception d’une assimilation des deux emplois, possessif et existentiel, de *yǒu* 有.

Tournons-nous maintenant vers l’emploi d’AVOIR au sein des structures présentatives biclausales qui nous intéressent tout particulièrement (voir la section suivante pour le lien entre

⁶² Pour raisons d’uniformité avec le reste de notre travail nous avons changé les gloses des exemples tirés de Chappell et Creissels (2019), qui utilisent la glose EPP (‘Existential Possessive Predicator’) justement pour ne pas donner l’impression que l’une des deux constructions (existentielle ou possessive) soit nécessairement valable dans tous les cas. Les auteurs nous en excuseront.

⁶³ Rappelons-le, les exemples tirés d’études sur l’acquisition d’une L2 sont précédés par un sigle indiquant d’abord la langue cible, dans laquelle l’énoncé est produit (ici : italien L2) et la langue source de l’apprenant (ici : chinois L1).

structures existentielles et structures présentatives). Or, le verbe AVOIR⁶⁴ est employé dans les structures présentatives biclausales non seulement en tant qu'extension de son usage impersonnel dans les phrases existentielles, mais aussi en tant qu'extension désémantisée de son usage trans-possessif.

En anglais, langue où la formule existentielle comporte le verbe ETRE (*there is*), on trouve des constructions du type (I.90-91). Les exemples suivants de l'anglais américain parlé (non standard) sont tirés de Lambrecht (1988b : 333) qui appelle ces structures *presentational amalgam constructions* en raison de l'absence du pronom relatif. Voir la section 1.2.2.3 pour des exemples en anglais standard.

(I.90) *I have a friend of mine in the history department teaches two courses per semester*

(I.91) *I have a friend from Chicago's gonna meet me downstairs*

Lambrecht (1988a : 152) considère que l'usage d'AVOIR comme verbe présentatif est motivé sémantiquement par la structure thématique de ce prédicat, dont le sujet a un rôle sémantique de locatif plutôt que celui d'un agent et dont l'objet est un thème au lieu qu'un patient (Clark 1970 ; Foley et Van Valin 1984)⁶⁵. Le verbe AVOIR a donc l'avantage de permettre au sujet logique d'apparaître en position postverbale sans demander une inversion des constituants *stricto sensu* (l'ordre V-S), et de ce fait serait particulièrement exploité à des fins pragmatiques dans les langues qui ne permettent pas l'inversion des constituants, comme le français (« The use of HAVE-type predicates is therefore especially motivated in languages like modern French, which do not permit verbs in sentence-initial position », Lambrecht 1988a).

1.2.2.2. Structures existentielles et structures présentatives

Au début du XX siècle, Jespersen (1924 : 155) introduisait en linguistique le concept de « phrase existentielle », qui depuis a inspiré de nombreux travaux. Comme il l'explique lui-même, la particularité des constructions existentielles n'est pas uniquement celle, sémantique, d'asserter l'EXISTENCE d'une entité, mais au niveau de sa forme, la position postverbale du sujet (voir aussi Kuno 1971) :

⁶⁴ Ou la formule existentielle centré sur le verbe AVOIR, synchroniquement, comme *il y a* en français ou étymologiquement, comme en espagnol *hay*, par exemple.

⁶⁵ Voir aussi Keenan (1987) à qui est dû le terme « existential-*have* », et plus récemment, Bassaganyas-Bars et McNally (2019) pour une analyse de ces structures en rapport aux formes en *there*.

« Les énoncés correspondant aux phrases qui comportent en anglais *there is* ou *there are* et qui affirment ou nient l'existence de quelque chose – appelons-les *phrases existentielles* – présentent dans de nombreuses langues des caractéristiques intéressantes. Qu'ils soient introduits ou non par un mot du type de *there*, le verbe précède toujours le sujet qui ne se comporte pas véritablement comme un sujet du point de vue grammatical. » (Jespersen 1971 : 210, cité dans Meulleman 2012 : 34).

Plus précisément, comme le remarque Lambrecht (1994 : 179) il s'agit de phrases locatives, puisqu'il est toujours question de préciser l'emplacement de l'entité en question, les énoncés qui assertent strictement l'existence d'une entité étant très marginaux dans le discours commun. Ce point de vue correspond *mutatis mutandis* à la conception cognitive de l'existence comme un emplacement dans l'espace conceptuel (« location in conceptual space ») chez Lakoff (1987 : 543). Les phrases existentielles servent alors à introduire le référent dans l'univers discursif des interlocuteurs en assertant sa présence dans un lieu donné (« to introduce the NP referent into the discourse world of the interlocutors by asserting its presence in a given location », Lambrecht (1994 : 179)⁶⁶. Dans les langues que nous analysons, les structures présentatives représentent une extension des phrases existentielles (ou possessives). Creissels (2014) le formule comme suit :

« Among the languages that have grammaticalized an existential predicative construction, many also have a presentational construction in which the existential predicator is used to introduce the NP presenting the new participant, as in English *There are many students who work in supermarkets* » (Creissels 2014 : 15).

Soit les énoncés suivants du chinois – ainsi que leur traduction en français : (I.92) est généralement considérée comme une construction existentielle, tandis que (I.93) est considérée comme une construction présentative :

(I.92) 家里有三个人
Jiā=li yǒu sān-ge rén
 maison=dans AVOIR trois-CL personne
 'A la maison il y a trois personnes' (= 'On est trois chez nous')

(I.93) 有人给你打电话
yǒu rén gěi nǐ dǎ diànhuà
 AVOIR personne à 2SG passer appel
 'Il y a quelqu'un qui t'a téléphoné'
 'Someone telephoned you' (traduction d'origine) (Li et Thompson 1981 : 131)

⁶⁶ Voir Creissels (2014 : 6-8) et Sornicola (2010). Pour une discussion approfondie sur le rapport entre l'existence et la présence, nous renvoyons le lecteur à Meulleman (2012 : 69-72). Voir aussi Halliday (2004 : 256-59).

En effet, si les constructions présentatives dérivent de structures existentielles, elles sont (dans certains contextes au moins) désémantisées, au profit de leur fonction pragmatique. Cette caractéristique a été mise en avant pour le français :

(I.94) *Moi j'ai encore un formulaire que j'ai pas* (Blanche-Benveniste 1983)

L'exemple qui précède est cité dans Lambrecht (1987 : 229). D'après Lambrecht (1986, 2001) la désémantisation de *il y a* au sein des structures présentatives serait l'un des traits révélateurs de leur « non-compositionnalité », c'est-à-dire que le sens de la construction ne serait pas dérivé de la somme des parties qui le composent mais est transmis par la structure dans son ensemble. Jullien (2014) est d'avis contraire et soutient que le sémantisme des clivées en *il y a* est bien compositionnel. Voir aussi Karssenber (2018b : 266) : « the fact that *il y a* can be “frozen” in present tense, and, more generally speaking, is desemanticized, might indicate that “il y a NP” is no longer analyzed compositionally. However, this does not entail that the whole “il y a NP + cleft relative clause” pattern is non-compositional. ».

Considérons maintenant les exemples suivants en chinois. Nous constatons que les énoncés seraient contradictoires du point de vue de la logique si *yǒu* avait un sens locatif-existentiel plein :

(I.95) 家里有人去世了 (BJKY)
Jiā=li yǒu rén qùshì-le
 maison=dans AVOIR personne mourir-PFV
 ‘Dans [ma] famille il y a quelqu'un qui est décédé.’

(I.96) 有一个同学他啊在哪儿? (BJKY)
yǒu yí-ge tóngxué tā a zài nǎr?
 AVOIR un-CL camarade 3SG PAU se.trouver où
 ‘Il y a un camarade, il est passé où ?’

(I.97) 开始空中小姐说有一个乘客没到 (BJKY)
Kāishǐ kōngzhōngxiǎojiě shuō yǒu yí-ge chéngkè méi dào
 commencer hôtesse.de.l'air dire AVOIR un-CL passager NEG.AVOIR arriver
 ‘Au début l'hôtesse de l'air a dit qu'il y avait un passager qui n'était pas arrivé’

(I.98) 有些人不在了, (CCL)
Yǒu xiē rén bú zài le,
 AVOIR certains personne NEG être.là CRS
 但他的作品却还活着
dàn tā de zuòpǐn què hái huó-zhe
 mais 3SG DE œuvre en.revanche encore vivre-DUR
 ‘Il y a des gens qui ne sont plus là, mais leur œuvre continue de vivre’

Les énoncés qui précèdent (I.95-98) montrent l'incompatibilité sémantique entre *yǒu* ('avoir, y avoir, exister') et le V2, si la phrase était comprise au sens littéral⁶⁷. Ces faits semblent suggérer une désémantisation de *yǒu* au profit de sa fonction pragmatique (voir à ce sujet la notion de « existence discursive » chez Karttunen 1969).

Certes, cela ne signifie pas que les structures locatives-existentielles ne portent pas de fonction pragmatique présentative, tout au contraire⁶⁸. Remarquons toutefois que dans les structures monoclausales, l'opérateur existentiel (ex. *yǒu*) est le seul prédicat, qui agit donc nécessairement comme lien entre l'entité dénoté par le SN_F et la référence spatio-temporelle (le SN_G). En revanche, dans les structures biclausales, le noyau prédicatif est représenté par le V2 : la localisation (souvent implicite donc pointant à l'« ici-et-maintenant ») est de type pragmatique avant que sémantique (spatiale).

Une autre caractéristique bien connue propre aux structures présentatives biclausales, en quelque sorte liée au processus de désémantisation que nous venons d'illustrer, est le fait que leur contenu propositionnel peut être exprimé par un énoncé monoclausal S-V (Lambrecht 2001 : 463)⁶⁹ :

- (I.99) a. *Il y a quelqu'un / Jean qui est arrivé.*
 b. *Quelqu'un / Jean est arrivé.*
- (I.100) a. *Il y a quelqu'un qui a téléphoné.*
 b. *Quelqu'un a téléphoné.*

En chinois, la même correspondance a été relevée. Par exemple Fàn Jiyān (1985) considérait que toute structure à sujet préverbal indéfini quantifié peut commuter avec une structure en *yǒu*.

⁶⁷ De même, des énoncés tels que (ii) et (iii) montrent que l'information ajoutée par *yǒu* est superflue d'un point de vue de la structure sémantique, puisque le V2 dénote l'existence (ii) ou implique la présence (iii).

- (ii) 哪里有人存在, 哪里就有人权问题
Nǎli yǒu rén cúnzài, nǎlǐ jiù yǒu rénquán wèntí
 où AVOIR personne exister où alors AVOIR droits question
 'Partout où il y a des gens qui existent, il y a le problème des droits humains' (CCL)
- (iii) 还有一客没走了!
Hái yǒu yí kè méi zǒu le!
 encore AVOIR un.CL client NEG.AVOIR partir CRS
 'Il y a encore un client qui n'est pas parti !' (corpus de l'auteur)

⁶⁸ En effet, des auteurs comme Givón (1988) parlent en termes d'*existential presentative constructions*. Voir Karsenberg (2017) pour une étude de corpus qui montre la similarité fonctionnelle des structures monoclausales et biclausales en *il y a* du français (« existential sentences can have the same information structure articulations as *il y a* clefts [...]. In this sense, the two sentence types are closely related, which calls into question the clear-cut boundaries that are often drawn between *il y a* clefts on the one hand and existential sentences on the other »).

⁶⁹ Voir aussi Katz (2000 : 254) et Dufter (2006 : 32) à propos des clivées en général.

L'équivalence entre la structure en *yǒu* et l'énoncé monoclausale S-V comportant le même contenu propositionnel est illustrée dans les exemples qui suivent. Nous mettons un point d'interrogation puisque les énoncés à sujet préverbal indéfini quantifié sont difficilement acceptés en chinois oral, comme Zhōu et Shēn (2016 : 113) le font remarquer (nous en discuterons dans la section 1.3). Notons cependant que l'équivalence dont nous parlons entre la structure présentative biclausale et l'énoncé monoclausale S-V est d'ordre sémantique, et non pragmatique.

- (I.101) a. 有一位警察走了过来。
Yǒu yí-wei jǐngchá zǒu-le-guo-lai.
 AVOIR un-CL policier marcher-PFV-passer-venir
 'Il y a un policier qui s'est rapproché.'
- b. ?一位警察走了过来。
?Yí-wei jǐngchá zǒu-le-guo-lai.
 un-CL policier marcher-PFV-passer-venir
 'Un policier s'est rapproché.' (adapté de Zhōu et Shēn 2016)

- (I.102) a. 村里有一个孩子走丢了。
Cūn=li yǒu yí-ge háizi zǒu-diū-le.
 village=dans AVOIR un-CL enfant partir-perdre-PFV
 'Au village il y a un enfant qui a disparu.'
- b. ?村里一个孩子走丢了。
?Cūn=li yí-ge háizi zǒu-diū-le.
 village=dans un-CL enfant partir-perdre-PFV
 'Au village un enfant a disparu.' (adapté de Zhōu et Shēn 2016)

Cependant, l'équivalence propositionnelle est moins évidente dans le cas des noms nus (I.103) :

- (I.103) a. 有人给你打电话
yǒu rén gěi nǐ dǎ diànhuà
 AVOIR personne à 2SG passer appel
 'Il y a quelqu'un qui t'as appelé'
- b. #?人给你打电话
#? Rén gěi nǐ dǎ diànhuà
 personne à 2SG passer appel
 'La/les personne.s t'a/ont appelé'
- c. 某个人 / 某些人给你打电话
mǒu-ge rén / mǒu-xiē rén gěi nǐ dǎ diànhuà
 certain-CL personne certain-quelques personne à 2SG passer appel
 'Quelqu'un / certaines personnes t'a/ont appelé'

En effet, l'interprétation des noms nus en termes de définitude change selon leur position par rapport au verbe (1.1.2.3). Ainsi, *yōu* dans ce cas (I.103a) s'emploie justement pour éviter que le référent dénoté par le nom nu (en l'occurrence *rén* 人 'personne') soit interprété comme défini (ou mieux : comme pragmatiquement identifiable)⁷⁰. Une équivalence entre les deux formes [*yōu* + SN_F + V] et [SN_F + V] en termes de contenu propositionnel n'est donc pas toujours directement applicable. Cependant, avec des ajustements lexicaux, cette commutation reste effectivement possible⁷¹.

Dans notre travail, nous parlons en termes de structures présentatives en AVOIR mono- et biclausales, et donnons une importance toute particulière au deuxième type, pour les raisons susmentionnées.

1.2.2.3. La construction relative présentative

- (I.104) a. *Il était une fois une belle princesse **qui** vivait dans un vieux château*
 b. *J'ai eu mon beau-frère **qui** a fait Paris-Nice*
 c. *Y a le téléphone **qui** sonne*
 d. *Je vois le facteur **qui** arrive*
 e. *Voilà le facteur **qui** arrive* (Lambrecht 2000a : 49)

La dénomination de *construction relative présentative* (en anglais *presentational relative construction*) vient de Lambrecht (2000), qui insère la structure présentative en (*il*) *y a* dans un groupe plus large incluant un ensemble hétérogène d'introducteurs (I.105a-e)⁷². La structure est schématisée ainsi par Lambrecht (2000a) : $v[\text{pro}+V] [\text{SN}_i]_{\text{PR}}[\text{qui}_i \text{SV}]$ ⁷³. Il s'agit dans tous les cas d'une construction complexe composée d'une proposition principale à fonction présentative et d'une subordonnée relative non restrictive :

« Le verbe de la principale est soit un verbe existentiel (surtout *avoir*), soit un verbe appartenant à une classe sémantique caractérisée par la présence explicite ou implicite d'un élément de perception. [...]

⁷⁰ L'exemple (I.103a) comporte un prédicat épisodique et un nom nu (*rén* 'personne.s') à interprétation existentielle. Comme nous le montrerons plus tard (chapitre IV § 4.6.3.3), la séquence *yōu rén* [AVOIR personne] déclenche le plus souvent une lecture partitive du nom : dans ce cas *yōu* est utilisé pour éviter justement que le nom nu reçoive une *interprétation universelle*, en lui attribuant une valeur partitive (nous en parlerons dans la section 4.6.4.7 du chapitre IV).

⁷¹ Il a été noté qu'en français, au-delà des exemples prototypiques comportant un nom propre (ex. *Il y a Jean qui est arrivé* ≈ *Jean est arrivé*) les structures présentatives biclausales en *il y a* ne commutent pas toujours avec un énoncé S-V monoclausal sans que des ajustements soient nécessaires (notamment lorsqu'elles incluent un indéfini à lecture partitive). Voir à ce sujet Karssenbergh (2018b : 132) et la section 3.9 du chapitre III.

⁷² Les structures *il est SN qui...* sont limitées à certaines formes figées comme dans (I.105a).

⁷³ Pro = pronom conjoint ; PR = proposition relative ; SV = syntagme verbal.

L'élément pro- initial, qui dénote le point de repère par rapport auquel est située l'entité introduite dans le discours, a le rôle de *locatif* (entendu dans un sens large, qui inclut le rôle du siège d'une perception physique) ; le SN postverbal dénotant l'entité introduite a le rôle de *thème* ; et la relative, qui dénote une situation mettant en jeu cette entité, a le rôle de prédicat » (Lambrecht 2000a : 50).

En effet, Lambrecht (1988a : 152, 2000a : 50) associe la fonction du morphème *y* au sein de la formule *il y a* (I.105c) au siège de l'élément de perception (I.105d), en raison de leur fonction de localisation d'un référent (*referent locating function*), dénotant le point de repère par rapport auquel l'entité nouvelle introduite dans le discours est située.

Ajoutons que la structure en *c'est...qui*, laquelle est typiquement associée à l'expression du focus étroit, peut également s'employer comme structure présentative (Dufter 2008) :

(I.105) ***C'est le facteur qui arrive***

Nous incluons donc cette forme parmi les structures à fonction présentative du français – ce qui n'est pas fait par Lambrecht (2000a)⁷⁴. Ainsi, en plus de la structure en *avoir* (chapitre III), dans notre travail nous nous intéressons tout particulièrement à la construction présentative de Perception (« compte-rendu de perception » chez Lambrecht 2000a ; terme emprunté à Kleiber 1988) et à la construction présentative en *c'est*, étant donné qu'il s'agit des formulations que nous observons dans les récits de nos locuteurs (nous discuterons de ces structures, régies par des différents introducteurs, dans le chapitre V § 5.3).

Pour conclure cette section, remarquons que, s'il s'agit d'une forme fréquemment utilisée en français, la construction relative présentative existe dans d'autres langues. Sur la base d'une recherche de corpus, Duffield *et al.* (2010) rendent compte de la construction relative présentative en anglais, et remarquent qu'elle tend à comporter des verbes dénotant l'existence, la perception et la découverte (les verbes relevés dans le corpus sont : *be, get, see, hear, tell, have, find, know, look, wonder*). En voici des exemples (adaptés de Duffield *et al.* 2010) :

(I.106) ***And I know people who have been drug tested and who have not, you know, been hired by a corporation.***

(I.107) ***You get a guy down the street who comes up, uh, carrying a knife.***

⁷⁴ Mais voir Lambrecht (2010) : « in argument-focus [*i.e.* narrow focus] contexts the cleft involving the copula *être* is the typical choice in natural discourse. There are, however, certain eventive sentence-focus contexts, difficult to pin down, in which the *c'est*-cleft is used instead of the *avoir*-cleft. [...] The sentence-focus cleft involving *être* is also the construction conventionally used to introduce the characters of jokes » (2010 : 94-95). Pour une comparaison pragmatique des structures biclausales en *il y a* et celles en *c'est* voir Karssenberget Lahousse (2018).

Enfin, si une association entre les différentes structures présentatives biclausales est évidente en français pour les raisons illustrées dans cette section, elle l'est moins en chinois, en l'absence de pronoms relatifs dans cette langue. Cependant, si nous considérons ces structures sur le même plan, c'est en raison de leur fonction – présentative justement – qui ressort notamment de nos données (chapitres VI et VII), et pour permettre plus aisément une comparaison interlinguistique.

1.3. Les sujets préverbaux indéfinis

Dans cette section, nous nous intéressons à la contrainte, plus au moins forte selon les langues, qui défavorise les noms indéfinis en tant que sujets préverbaux. Depuis des travaux tel que Givón (1976) on a remarqué que dans les langues SVO les sujets préverbaux tendent à avoir des propriétés de topique : ils sont typiquement définis, référentiels et/ou génériques (Cheng et Downing 2014). Cela est bien résumé par Bentley (2011) :

« Sebbene le proprietà dei soggetti canonici possano variare da lingua a lingua, in genere le lingue naturali evitano i soggetti indefiniti non generici (eccezion fatta per stili particolari che sono irrilevanti in questo contesto perché soggetti a convenzioni che annullano i vincoli operativi nei registri spontanei), dando preferenza ai soggetti topicali e identificabili [...], altamente referenziali, agentivi e specifici »⁷⁵ (Bentley 2011 : 17).

Il s'en suit que les SN indéfinis, étant typiquement associés à l'information nouvelle, devraient être défavorisés en tant que sujets préverbaux. En effet, les échelles hiérarchiques d'accessibilité (anaphore zéro > pronoms > SN définis > SN indéfinis) dont nous avons discuté en 1.1.1.1, en dehors de donner des informations sur l'accessibilité d'une forme linguistique en contexte, nous indiquent aussi quels éléments sont plus facilement acceptés en tant que topiques.

Chafe (1994 : 83-92) estime que les sujets sont régis par un principe qu'il appelle « la contrainte du sujet léger » (« *the light subject constraint* ») : dans la conversation les sujets (préverbaux) réfèrent normalement à de l'information donnée ou accessible, tandis que l'information nouvelle apparaît plus tard dans l'énoncé. En même temps, il admet que le registre écrit n'adhère pas aussi strictement que le registre parlé à cette contrainte (Chafe 1986).

Typologiquement, on sait aussi que chaque langue est plus ou moins stricte par rapport à la légitimation d'un SN indéfini en tant que sujet : « [w]hile the cognitive constraints [preventing inaccessible referents to appear as preverbal subjects] are universal, their grammatical manifestation

⁷⁵ « Bien que les propriétés des sujets canoniques puissent varier d'une langue à l'autre, les langues naturelles évitent généralement les sujets indéfinis non génériques (à l'exception de styles particuliers qui ne sont pas pertinents dans ce contexte car ils sont soumis à des conventions qui neutralisent les contraintes à l'œuvre dans les registres spontanés), en donnant la préférence à des sujets topicaux et identifiables [...], hautement référentiels, agentifs et spécifiques ».

in individual languages is subject to typological variation » (Lambrecht 2010 : 81). Ainsi que le formule Leonetti (1998) :

« As several linguists have pointed out, the occurrence of indefinite subject NPs (in initial or preverbal position) is more restricted than the occurrence of definite ones. There seems to be a widespread typological preference for definite preverbal subjects, and some languages (Arabic, for instance) even forbid the presence of preverbal indefinite subjects. Other languages (such as English or Spanish) do not grammaticalize that universal tendency in such a strong way, but nevertheless they do impose certain constraints on indefinite subjects as well » (Leonetti 1998).

En français, diverses études ont mis en évidence des restrictions sur l'apparition des indéfinis en position de sujet préverbal (« [i]t is well known that there are restrictions on the appearance of indefinite expressions in preverbal subject position, especially in spoken French. », Karssenber 2018b : 116). En effet, les syntagmes nominaux indéfinis préverbaux n'apparaissent dans la langue parlée que dans des conditions limitées. Cela est particulièrement visible avec des partitifs en position initiale (Cappeau *et al.* 2001), comme en (I.108a-b), mais cette contrainte est également à l'œuvre lorsqu'on assigne une lecture spécifique aux énoncés (I.109a-b) :

(I.108) a. **de la neige a effacé leurs traces*
 b. **du blé a pourri sur pied* (Van de Velde 1995)

(I.109) a. *?un cheval est à l'écurie*
 b. *?un chien aboie* (Cappeau *et al.* 2001)

Cappeau (2008) montre que certains types de sujets préverbaux indéfinis sont plus fréquents dans le registre écrit (ou plus formel) que dans le parler spontané. Dans son étude il s'intéresse aux pronoms indéfinis sujet (tableau I-6) :

Tableau I-6. *Les pronoms indéfinis du français en position sujet (d'après Cappeau 2008 : 75)⁷⁶*

	Corpus parlé	Corpus écrit
<i>Certain(e)s</i>	51	80
<i>D'autres</i>	17	51
<i>Beaucoup</i>	10	33
<i>Plusieurs</i>	0	3

⁷⁶ Le tableau est tiré de Karssenber (2018b : 121), qui exclut les résultats tirés par le corpus littéraire du 19ème siècle, présentés dans l'article d'origine. Les deux corpus dont il est question dans le tableau, oral et écrit, ont la même ampleur (presque 2 millions de mots).

typologie de sujets préverbaux. Leurs résultats montrent que les sujets (préverbaux) manifestent une forte tendance à être encodés par des pronoms, étant donné que les sujets lexicaux ne constituent que le 9% des sujets totaux (tableau I-7).

Tableau I-7. *La distribution des types de sujets pour 31.021 phrases déclaratives (Francis et al. 1999 ; corpus d'anglais parlé)*

	<i>Occurrences</i>	<i>Pourcentage</i>
Sujets lexicaux	2.858	9%
Sujets pronominaux	28.163	91%

En revanche, les objets ont tendance à être lexicaux (66% des phrases transitives) :

Tableau I-8. *La distribution des types d'objets pour 7.489 phrases transitives (Francis et al. 1999 ; corpus d'anglais parlé)*

	<i>Occurrences</i>	<i>Pourcentage</i>
Objets lexicaux	4.921	66%
Objets pronominaux	2.568	34%

Considérons l'exemple suivant : le nouveau référent *a baby* est introduit par un SN indéfini (marquage local) en tant qu'objet verbal (marquage global). Puis, étant devenu accessible, son référent est encodé en tant que sujet – et après objet – pronominal :

(I.111) *My sister has a, she just had a baby. **He's** about five months old, and she was worrying about going back to work and what she was going to do with **him**.* (Francis et al. 1999)

Les auteurs s'intéressent ensuite à la forme de ces sujets et objets lexicaux, et montrent que les sujets indéfinis marqués par le déterminant *a(n)* ne constituent que le 2% des sujets lexicaux :

Tableau I-9. *La distribution des déterminants pour les sujets et les objets lexicaux (Francis et al. 1999 ; corpus d'anglais parlé)*

	<i>A/an</i>	<i>The</i>	<i>Possessif</i>	<i>Autre</i> ⁷⁸
Sujets	65 (2%)	1.070 (37%)	715 (25%)	1.008 (36%)

présentation du corpus sur le site du *The Linguistic Data Consortium* (LDC), <https://catalog.ldc.upenn.edu/LDC97S62> (consulté le 27 mai 2020).

⁷⁸ Malheureusement, dans leur article, Francis et al. (2010) ne donnent pas d'informations sur les formes morphologiques rangées sous la catégorie « autre » (tableau I-9), ce qui serait intéressant étant donné qu'il ne s'agit pas d'un pourcentage marginal (36%). Cependant, on trouve plus loin l'indication suivante : « Table 3 demonstrates the asymmetric distribution of morphological forms for subjects and objects. **In total, 62 percent** [« the » + « possessif »] **of lexical subjects are uniquely identifiable**, compared to only 23 percent for objects. The contrast between indefinite subjects and indefinite objects is also striking: Only 2 percent of subjects are indefinite compared to 29 percent of objects. **In accordance with the correlations between morphological form and givenness status described above, we conclude that subjects strongly tend to be at least uniquely identifiable** » (Francis et al. 2010, notre emphase).

Objets	1.419 (29%)	784 (16%)	346 (7%)	2.372 (48%)
--------	-------------	-----------	----------	-------------

En même temps, les sujets lexicaux relèvent principalement de deux catégories : d'une part, le défini s'utilise pour renvoyer à un référent déjà introduit dans le discours ou pour lever une ambiguïté parmi des référents possibles. Dans les autres cas, le choix d'un défini est motivé par le phénomène du pontage inférentiel : il s'agit alors de ces éléments qui sont identifiables grâce à l'activation d'un cadre sémantique. En voici un exemple :

- (I.112) *uh, actually I lived over in Europe for a couple of years, I lived in Germany and in Germany they don't have the jury system. What they do is they have, uh, three judges, basically. And you get up there and **the prosecuting attorney** presents his evidence.*
(Francis *et al.* 1999)

Nous avons discuté de ce type d'inférence référentielle en détail dans la section 1.1.2.2.

En somme, ces résultats suggèrent que les sujets lexicaux sont quand même rares par rapport aux sujet pronominaux, et qu'en particulier les sujets indéfinis (marqués par l'article) ne représentent qu'un procédé très marginal. A leur tour, les sujets lexicaux dénotent un référent déjà introduit ou bien inférable selon le procédé de l'anaphore associative. Comme Francis *et al.* (1999) le remarquent, leurs résultats sont à manier avec précaution du fait que le registre de langue y joue un rôle. En effet, dans d'autres registres linguistiques, les proportions entre sujets lexicaux et pronominaux sont altérées, voire inversées dans certains types de textes écrits.

Or, dans d'autres langues, moins « tolérantes » que l'anglais, comme le norvégien (Hasselgård 2018) les auteurs ont remarqué quand même que les SN indéfinis ne sont pas absents de leurs corpus. Cheng et Downing (2014) montrent qu'en Zulu (langue bantoue) les définis mais aussi les indéfinis forts, impliquant une présupposition d'existence, sont admis en position de sujet préverbal. En néerlandais, les indéfinis faibles (à lecture existentielle) ne peuvent apparaître que dans des constructions présentatives avec *er* mais ne sont pas tolérés en tant que sujets préverbaux ; en revanche, les indéfinis forts (à lecture spécifique ou générique) sont admis en cette position (Rullmann 1989).

En français, comme nous l'avons vu, les indéfinis ne sont pas exclus catégoriquement de cette position. Il existe en effet des conditions grammaticales qui peuvent favoriser la présence d'indéfinis en sujet. On remarquera une meilleure acceptabilité liée à la présence d'un complément dans le SN indéfini, où le complément pointe vers l'ensemble de référence par rapport auquel l'indéfini est repéré, dans le monde extralinguistique ou dans la mémoire discursive (Cappeau et Deulofeu 2001) :

- (I.113) a. ?*un ami m'a apporté ce livre*
b. *un de mes amis m'a apporté ce livre*

c. *un d'entre eux m'a apporté ce livre* (d'après Cappeau et Deulofeu 2001)

Dans ce cas, « le syntagme indéfini fonctionne comme référant à la partition d'un ensemble plus vaste préalablement identifié, ce qui l'autorise à fonctionner comme sujet » (Cappeau et Deulofeu 2001)⁷⁹. Il s'agit là des référents non identifiables mais ancrés dont nous avons parlé plus haut (en 1.1.1.1).

Rappelons qu'en chinois aussi, les nouveaux référents non identifiables sont souvent introduits au moyen d'un ancrage à un deuxième élément (comme dans [I.113b-c]) qui en restreint le champ de référence (1.1.2.3). Plus loin dans cette section nous parlerons des référents non identifiables non ancrés qui apparaissent en position de sujet préverbal en chinois.

Un autre facteur qui facilite l'apparition des indéfinis en position préverbale concerne le sémantisme du verbe. Il s'agit là de la distinction faite entre prédicats épisodiques (*stage level predicates*) et prédicats individuels (*individual level predicates*)⁸⁰.

Les prédicats épisodiques peuvent comporter un sujet préverbal indéfini puisque le prédicat indique les paramètres spatio-temporels de la proposition, et permet donc au sujet d'être ancré. Les prédicats individuels, en revanche, ne réfèrent pas à des paramètres spatio-temporels et nécessitent donc d'être ancrés par leur sujet, ce qui ne peut pas être fait par un sujet indéfini.

Comme le formule Van de Velde (2005) :

« [Le] prédicat individuel est générique au sens où il s'attribue à son sujet indépendamment du temps. Cela implique entre autres choses qu'il signifie une propriété qui n'a elle-même aucun rapport direct avec le temps mais n'en a, si elle en a, que par la médiation de son sujet. Une telle propriété ne peut donc avoir d'existence que si son sujet existe, et non l'inverse. Les propriétés dites épisodiques, particulièrement quand elles sont événementielles, ont quant à elles un rapport au temps qui les rend non seulement autonomes par rapport à leur sujet, mais susceptibles d'assurer son existence quand celle-ci n'est pas présupposée » (Van de Velde 2005 : 23).

Ces propriétés sémantiques se reflètent donc dans l'acceptabilité des sujets préverbaux indéfinis :

(I.114) *Someone has been following me* (Leonetti 1998 : 143)

(I.115) *?A student is tall* (Leonetti 1998 : 142)⁸¹

⁷⁹ Cette partition peut se faire par une complémentation dans le syntagme : *un de ces amis*, mais aussi grâce à la tournure « il y en a » (Cappeau et Deulofeu 2001). Voir le chapitre III § 3.5.1 pour une discussion de la forme *il y en a ... qui* du français.

⁸⁰ La célèbre distinction entre les deux types de prédicats a été initiée par Carlson (1977) et développé notamment par Kratzer (1995).

⁸¹ Les exemples de Leonetti (1998) sont cités dans Karssenber (2018b : 122–23), travail dont nous nous sommes inspirée plus généralement dans notre discussion des facteurs qui influencent l'acceptabilité des sujets préverbaux indéfinis, et nous y renvoyons donc volontiers le lecteur (Karssenber 2018b : 121–25).

Dans des énoncés comme (I.115), ni le prédicat ni le sujet sont en mesure de créer un contexte mutuellement approprié, et de ce fait l'énoncé résulte manquant d'un ancrage spatio-temporel (« neither the predicate nor the subject are able to create an appropriate context for each other. This causes an increase in processing effort, as the recovery of the proposition expressed is not adequately constrained », Leonetti [1998 : 158], cité dans Karssenbergh [2018b : 123]). Cependant, les sujets préverbaux indéfinis ne sont pas toujours impossibles avec un prédicat d'individu. Un moyen pour contourner l'inacceptabilité est notamment l'ancrage du SN indéfini à un deuxième élément qui en spécifie la référence (I.117) – dont nous avons parlé *supra* :

(I.116) *A student of mine is tall* (Leonetti 1998 : 156)

Karszenbergh (2018b : 120) résume les propriétés qui ont été identifiées dans la littérature concernée en tant que facteurs améliorant ou défavorisant l'acceptabilité d'un SN en tant que sujet de l'énoncé (tableau I-10).

Tableau I-10. *Les propriétés qui influencent l'acceptabilité d'un sujet (Karszenbergh 2018 : 120)*

Facteurs positifs	Facteurs négatifs
SN défini	SN indéfini
SN donné en discours	SN nouveau pour le discours
SN [+humain]	SN [-humain]
Pertinent en particulier dans le cas d'un SN indéfini :	
Prédicat épisodique (événementiel)	Prédicat d'individu (statif)
SN spécifique	SN non spécifique
SN quantifié	SN non quantifié
SN (explicitement) partitif	SN non partitif

En chinois, on a depuis longtemps remarqué que le sujet préverbal a tendance à être défini (Chao 1968 : 76, Lü 1946/1990 : 469, Li et Thompson 1981 : 20, Zhū 1982 : 96, Xu 1997) :

« The preverbal subject is interpreted as *definite* (known to both the speaker and the hearer), while the postverbal subject [...] is interpreted as *indefinite* (not known to the hearer at least » (Li et Thompson 1981 : 20).⁸²

⁸² Li et Thompson (1981) distinguent le topique et le sujet en chinois. « Nouns that are unmarked for definiteness are always interpreted as definite or generic when they are topics, however [...] an indefinite noun phrase with *yi*- 'one' or any other numeral in general does not occur as a topic » (Li et Thompson 1981 : 86).

Ainsi, si le nom nu est admis en position préverbale puisqu'il y déclenche une lecture définie (1.1.1.4), les noms précédés par la séquence [numéral + classificateur] sont évités dans ce contexte (Li et Thompson 1981 : 20 ; Zhū 1982 : 26-27) :

- (I.117) {那个人 /人/*一个人}来了
 {Nà-ge rén / rén / *yí-ge rén} lái-le
 DEM_{dist}-CL personne personne un-CL personne venir-PFV
 'Cette personne / [la/les] personne.s / *une personne est (sont) arrivée.s'

Cependant, d'autres études ont montré que, dans cette langue aussi, il ne s'agit pas d'une contrainte absolue. En effet, dans certaines phrases, le sujet peut être représenté par un nom indéfini quantifié (précédé par la séquence [numeral + classificateur]).

Dans son étude de 1985, Fàn Jiyān montrait que les sujets préverbaux indéfinis apparaissent dans certains registres de langue, en particulier dans le style journalistique (du même avis est Uchida 1989). Fàn (1985) montre aussi qu'à la plupart des structures à sujet indéfini, on peut toujours ajouter *yǒu* 有 en tête de phrase. Les structures [SN_F + V] et [*yǒu* + SN_F + V] auraient une fonctionnalité discursive comparable pour cet auteur.

Depuis, de nombreux auteurs se sont intéressés aux « phrases à sujet indéfini » (*wúding* NP *zhǔyǔ jù* 无定 NP 主语句) du chinois (Lǐ Líndìng 1986 : 2, Uchida 1989, Wáng Cànlong 2003, Liú et Zhāng 2004, Táng Cuijú 2005, Cáo Xiùlíng 2005, Zhāng Xīnhuá 2007, Xióng 2008, Zhōu et Shēn 2016). Wang Hongqi (2014) analyse la fréquence d'utilisation des phrases à sujet indéfini et constate que celles-ci sont utilisées plus fréquemment dans les styles politique et narratif. Une spécialisation au niveau du registre était identifiée également par Liú et Zhāng (2004) :

« Sur la base de l'observation de corpus de grande taille, nous avons vérifié les hypothèses des chercheurs susmentionnés. Ce type de phrase n'apparaît presque jamais dans les styles argumentatif, explicatif, dialogique, etc. On peut donc considérer la phrase à sujet préverbal indéfini comme un phénomène typique du style narratif⁸³ » (Liú et Zhāng 2004 : 97-98).

Certains auteurs ont suggéré que la configuration à sujet indéfini en chinois a une fonction descriptive, reflétant le point de vue d'un observateur. Ce type de phrase ne porte pas sur le sujet en tant qu'acteur impliqué dans une action, mais s'emploie pour décrire globalement une situation scénique (Uchida 1989, Lǚ Jíning 2004). En outre, Wáng Cànlong (2003) montre que la phrase à

⁸³ Notre traduction : « 我们观察了大量语料以后验证了上述学者的发现，在论证、说明、对话等语体里几乎不出现这种句子，无定 NP 主语句因此可以认定是叙述体里特有的现象 » (刘安春、张伯江 2004 : 97-98).

sujet indéfini a une autosuffisance syntaxique limitée, puisqu'en général elle s'utilise rarement de manière indépendante et manifeste une forte dépendance au cotexte.

Dans leur étude de 2004 que nous venons de mentionner, Liú et Zhāng s'intéressent à la fonction discursive des phrases à sujet indéfini en les comparant aux deux typologies de structures présentatives auxquelles nous nous intéressons dans notre thèse (les structures en *yǒu* et l'ordre V-S). Or il s'agit dans les trois cas de structures qui mettent en avant de l'information nouvelle en chinois. Si l'ordre V-S est ancré spatialement, la structure à sujet indéfini demande souvent un ancrage de type temporel (porté par l'adverbe *qiàhǎo* 恰好 'précisément (à ce moment-là)' dans l'exemple suivant). De plus, ils remarquent que ce type de phrase s'emploie pour marquer une rupture dans la continuité narrative.

- (I.118) 唐僧正在发愁,
Táng Sēng zhèngzài fāchóu,
 Tang Seng être.en.train.de s.inquiéter
 恰好一只小船从上游撑来。
qiàhǎo yí-zhī xiǎo chuán cóng shàngyóu chēng-lai.
 juste.à.ce.moment un-CL petit bateau depuis amont bouger.avec.un.poteau-venir
 八戒赶紧招呼摆渡。
Bājiè gǎnjǐn zhāohū bǎidù.
 Bajie précipitamment saluer ferry
 'Tang Seng était inquiet, juste à ce moment un petit bateau est venu de l'amont. Bajie se hâta de saluer le ferry. (Liú et Zhāng 2004)

En (I.119), *yí-zhī xiǎo chuán* 一只小船 'un petit bateau' est introduit dans le discours de manière accessoire, dans le cadre d'un décalage de topique. Juste après, l'attention revient aux personnages principaux de l'histoire.

L'étude de Fāng Méi (2019), qui se base sur trois types de sources (parler télévisé, roman, prose), parvient aux mêmes conclusions :

« D'un point de vue macro-discursif, le sujet indéfini [préverbal] révèle de l'information d'importance mineure, et ne dénote donc pas une entité qui s'inscrit dans la trame principale du récit. Dans ces contextes, il ne s'agit pas d'introduire des nouveaux référents topicaux car à l'inverse la fonction de la phrase dans son ensemble est celle de fournir des informations d'arrière-plan. Ainsi, bien que le nom

indéfini quantifié soit porteur d'information nouvelle, il ne possède qu'un faible degré de topicalité » (Fāng Méi 2019 : 194⁸⁴).

Zhōu et Shēn (2016 : 113) analysent les phrases à sujet indéfini du point de vue de la structure informationnelle et font remarquer que si celles-ci n'apparaissent pas souvent à l'oral c'est en raison de l'effort cognitif que leur traitement demande :

« Du point de vue de la syntaxe, [les phrases à sujet indéfini] sont conformes à la structuration typique de la phrase chinoise, et n'ont rien de particulier. Mais du point de vue de l'encodage, le nom indéfini fonctionne comme sujet, et de ce fait réunit une fonction référentielle et un rôle sémantique, ce qui présente une difficulté accrue du point de vue de son traitement cognitif. Par conséquent, ce type de phrase se rencontre rarement dans le discours oral, spontané et instantané, mais il est en revanche susceptible d'apparaître dans le registre écrit, puisqu'à l'écrit le medium de communication est différent, et le lecteur aura plus de temps pour identifier le nouveau référent, et au même temps pour interpréter l'assertion faite à son propos. » (Zhōu et Shēn 2016 : 113)⁸⁵.

Pour résumer les propriétés mises en évidence par les études discutées, les structures S-V à sujet préverbal indéfini en chinois :

- (a) relèvent du style narratif ;
- (b) sont étroitement liées au cotexte ;
- (c) ont une visée descriptive ;
- (d) signalent une interruption dans la continuité narrative ;
- (e) introduisent un événement d'arrière-plan.

Comme nous le verrons (chapitre VI § 6.4.2), dans notre corpus de chinois L1, l'ordre S-V à sujet indéfini est parfois employé pour introduire les personnages (de première plan) dans le récit, et il s'agit dans nos données d'une stratégie qui intervient notamment à l'incipit du récit. Manifestement, ces énoncés ne sont pas compatibles avec les propriétés mentionnées (b) et (d) puisqu'elles

⁸⁴ Notre traduction (libre) : « 从语篇宏观角度看，“一量名”主语是偶现信息，并非叙事主线的事物。这类情形并非引入新的具有话题性 (topicality) 的言谈对象而整句的功能在于提供背景信息。虽然“一量名”主语句引入语篇的名词是新信息的载体,但这个名词概念的话题性较弱 » (方梅 2019 : 194).

⁸⁵ Notre traduction : « 从句法上看, [无定名词主语句] 是典型的汉语句式格式, 并无异常。但是就其编码而言, 无定名词充当主语, 那么主语名词的“指称功能”和“角色功能”集中在一身, 在认知处理上有一定的难度, 因此这种句子一般很少在即时的、自然的叙事性口语中出现, 但却可以出现在书面语中, 因为在书面语中, 由于交际的媒介不同, 读者会有更充裕的时间从容地辨识所指, 同时识解 (construe) 对新引进所指的陈述。 » (周士宏、申莉 2016 : 113).

interviennent en tout début du discours (hors « contexte »), et avec (e) car l'entité qu'elles encodent est de première importance – il s'agit notamment du personnage principal de l'histoire. Nous argumenterons que, dans cet usage, le nom indéfini s'utilise justement pour initier une section de discours, ce qui est en quelque sorte assimilable à la fonction de signaler une interruption dans la continuité narrative (d) – dans les deux cas il s'agit de la « frontière » d'un épisode. Probablement, dans ce contexte précis, la propriété (e) relevée par les études est neutralisée – puisqu'il n'y a pas de concurrence entre plusieurs personnages possibles ayant des degrés différents de saillance.

Similairement, Huang (2013) remarque dans son corpus de productions orales⁸⁶ que, si les nouveaux référents non identifiables sont premièrement introduits par la structure en *yǒu*, ils apparaissent aussi parfois au sein d'une phrase prédicative canonique (S-V) :

- (I.119) 原来这个故事很简单
yuánlái zhè-ge gùshì hěn jiǎndān
à.l'origine DEM_{prox}-CL histoire très simple
就是这么描述一下
jiùshì zhème miáoshù yī-xià
être.précisément comme.ça décrire un-peu
➤ 一个工人
yí-ge gōngrén
un-CL ouvrier
在那边采巴拉
zài nàbiān cǎi bālā
à là.bas cueillir goyave
‘Alors cette histoire est très simple, on peut la décrire ainsi : un travailleur cueille des goyaves là-bas...’

Dans ce cas aussi, la structure S-V à sujet indéfini est utilisée en début de récit (les énoncés qui précèdent relevant de la dimension métalinguistique) et introduisent bien l'un des personnages principaux de l'histoire. Arslangul (2019) dans son étude sur des productions orales basées également sur un stimulus de type visuel (voir l'exemple [I.77] illustré plus haut), remarque que l'introduction des nouveaux référents en position préverbale via la construction prédicative canonique est attestée chez les locuteurs sinophones (7.9 % des énoncés comportant une introduction référentielle) mais

⁸⁶ Son corpus est basé sur un recueil de conversation spontanée ainsi que de productions élicitées sur la base des stimuli *The Pear Narratives* et *The Ghost narratives* (se référer à Huang 2013 : 82-83). Notons tout de même qu'il s'agit du chinois mandarin parlé à Taiwan, contrairement au reste des études présentées dans cette section (et dans notre thèse plus généralement), ce qui, au passage, est révélé par le choix lexical de décrire les « poires » présentes dans le stimulus *The Pear Narratives* comme des goyaves (exemple I.119), l'un des fruits les plus communs à Taiwan.

qu'elle est d'ailleurs beaucoup moins fréquente que dans les récits des locuteurs francophones (73.7 %).⁸⁷ Dans son étude, il s'agit cependant de référents d'arrière-plan par rapport au personnage principal de l'histoire, ce qui est cohérent avec les propriétés évoquées plus haut.

En somme, dans cette section nous avons vu qu'à travers les langues, les noms indéfinis, encodant des référents non identifiables, sont généralement évités en tant que sujets, puisque la position de sujet (dans les langues SVO) est typiquement associée à la fonction de topique.

Nous avons également vu que cette restriction est plus ou moins forte pour chaque langue et que le registre linguistique peut y jouer un rôle. Enfin, il a été remarqué que plusieurs facteurs interviennent pour faciliter l'apparition d'un SN indéfini en tant que sujet, notamment l'ancrage du nouveau référent à un ensemble plus large et le sémantisme verbal.

Notons particulièrement que dans les langues auxquelles nous nous intéressons dans notre travail, à savoir le français et le chinois, les noms indéfinis ne sont pas « interdits » en position de sujets, bien que les études sur ces deux langues aient relevé une forte tendance à les éviter. Cet aspect sera discuté en détail plus loin relativement à nos données de linguistique appliquée, pour le chinois et français L1 (chapitre VI § 6.4.1) et pour le chinois et français L2 (chapitre VII § 7.6.1).

1.4. Problématique et question de recherche I

Dans ce chapitre nous avons discuté de la fonction d'introduction référentielle, qui est nécessaire à tout échange communicatif, étant donné que nos conversations ne portent pas généralement sur un seul référent mais d'autres y sont introduits au fur et à mesure que nous parlons. Il a été spécifié que le référent nouveau par excellence est un référent non identifiable (inconnu par l'interlocuteur), non ancré (ne bénéficiant pas d'un lien avec une autre entité connue) et non inférable (ne bénéficiant pas d'un lien avec le contexte ou le cotexte). Puisque nous devons être en mesure de créer un discours cohérent et aisé à « décoder » par notre interlocuteur, le traitement des référents nouveaux est facilité par des structures dites « présentatives » qui sont censées l'aider dans cette tâche tout en signalant une rupture dans la continuité narrative. Parmi les structures présentatives disponibles dans les langues à cet effet, les structures biclausales introduites par un opérateur existentiel et l'ordre V-S sont citées comme des procédés syntaxiques souvent mobilisés. Deux courants d'études indépendantes se sont intéressées à ces structures en français et en chinois respectivement. Comme

⁸⁷ Notons que cette étude vise les procès spatiaux dynamiques, incluant donc un prédicat épisodique (événementiel), ce qui est considéré comme un facteur qui favorise l'emploi de l'ordre S-V dans les langues, comme il a été dit.

nous le verrons dans la deuxième partie de la thèse, si dans les deux langues on leur attribue généralement la fonction « présentative » d'introduire des nouveaux référents en discours, d'autres études ont montré la nécessité d'une analyse plus fine de cette fonction (voir p.e. la section 4.5 du chapitre IV concernant les structures en AVOIR du chinois) tandis que dans d'autres cas, notamment en rapport aux structures biclausales en AVOIR du français, on a montré qu'elles peuvent réaliser des fonctions tout à fait différentes (voir la section 3.6 du chapitre III). De plus, en ce qui concerne l'ordre V-S, on doit tenir compte d'une différence quant au registre d'usage dans les deux langues, puisque cette configuration est généralement réservée au registre écrit en français, contrairement au chinois.

Si, donc, d'un point de vue typologique on a souligné le recours systématique dans les langues à des stratégies qui permettent de contourner l'ordre des mots canonique – pour éviter à des « mauvais sujets » (Karssengerg 2018 : 115) tels que les référents non identifiables – d'apparaître en position préverbale (voir la section 1.3 du présent chapitre), les études qui s'intéressent à l'analyse d'une construction précise dévouée à cet effet rendent généralement compte d'un scénario fonctionnellement plus hétérogène.

A la lumière de ces faits, notre première question de recherche, qui se veut ouvertement exploratrice, est la suivante :

(QR1) Les structures présentatives du français et du chinois – et notamment les structures biclausales en AVOIR – relèvent-elles d'un phénomène pragmatiquement comparable ?

Cette question a été articulée en deux sous-questions :

(QR1.a) Les structures biclausales en AVOIR du chinois sont-elles multifonctionnelles, comme ce qui a été remarqué pour les formes du français ? Quelles sont leurs spécificités ?

(QR1.b) Quelle est la place des structures biclausales en AVOIR en rapport aux autres structures présentatives – notamment l'ordre verbe-sujet (V-S) – dans les deux langues (parlées) ?

Pour tenter d'y répondre, nous mobilisons des procédés divers. Tout d'abord, nous passons en revue les travaux antérieurs menés en linguistique française et en linguistique chinoise sur les structures qui nous concernent, et verrons que les deux courants n'utilisent pas une terminologie unifiée ni nécessairement les mêmes outils théoriques. Nous donnerons ainsi notre propre définition

de « structure présentative » en chinois (chapitre IV § 4.1.1). Ensuite, la discussion sera intégrée par nos recherches dans un corpus de chinois oral, qui nous permettra d'analyser d'un point de vue fonctionnel des exemples tirés de ce registre de la langue (chapitre IV § 4.6 pour les structures biclausales et chapitre V § 5.3–4 pour l'ordre V-S).

Notre analyse aura mis en évidence la multifonctionnalité propre aux structures observées, aussi bien en français qu'en chinois. En d'autres termes, une analyse en terme de « fonction présentative » n'est pas suffisante pour rendre compte de l'emploi de ces structures. Néanmoins, celle-ci demeure une fonction que les formes des deux langues ont effectivement en commun.

Dans la troisième partie de la thèse, nous opérons un changement de perspective pour viser cette fonction spécifique, dans le but d'observer comment elle est réalisée, étant donné un même contexte extralinguistique, par les locuteurs sinophones et francophones, ainsi que les apprenants des deux langues. En d'autres termes, si dans la deuxième partie de la thèse nous prenons comme point de départ les formes linguistiques (les structures présentatives), pour en conduire une analyse fonctionnelle, dans la troisième partie nous partons de la fonction (l'introduction référentielle), pour observer les formes qui y sont associées. C'est à ce point du travail que nous présenterons les données recueillies dans le cadre de notre expérience de linguistique appliquée.

Dans le chapitre VII nous étudions comment les apprenants sinophones de français et les apprenants francophones de chinois gèrent l'emploi des structures présentatives disponibles dans la langue cible en observant plus généralement comment ils introduisent les nouveaux référents dans leurs récits. L'acquisition de la dimension pragmatique de la langue, reposant sur des préférences plutôt que des choix catégoriques, constitue un enjeu pour l'apprenant. Un courant d'études s'y est consacré du point de vue des apprenants avancés, étant donné que l'appropriation de ce niveau linguistique est souvent citée comme la dernière étape à franchir avant une maîtrise complète de la langue cible. Puisque l'apprenant doit nécessairement reconsidérer les principes pragmatiques qui opèrent dans sa langue maternelle pour mettre en place ceux qui régissent la structuration du discours dans la langue cible, nombre d'études ont questionné le fait que cela soit effectivement possible. Dans le prochain chapitre nous présentons les outils théoriques sur lesquels nous nous appuyons, ainsi que les principales études réalisées dans ce domaine. Enfin, nous formulons notre deuxième question de recherche et présentons la méthodologie employée pour recueillir nos données.

Chapitre II