



HAL
open science

Conceptions lexicales et cognitives de la notion d'antécédent : une étude contrastive de l'anaphore pronominale en français et en chinois

Jiaqi Hou, Frédéric Landragin

► **To cite this version:**

Jiaqi Hou, Frédéric Landragin. Conceptions lexicales et cognitives de la notion d'antécédent : une étude contrastive de l'anaphore pronominale en français et en chinois. SHS Web of Conferences, 2020, 78, 10.1051/shsconf/20207812011 . hal-02931220

HAL Id: hal-02931220

<https://hal.science/hal-02931220>

Submitted on 5 Sep 2020

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

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

Conceptions lexicale et cognitive de la notion d'antécédent : une étude contrastive de l'anaphore pronominale en français et en chinois

Jiaqi Hou^{1*}, Frédéric Landragin^{1*}

¹ Lattice, CNRS, ENS-Université PSL, Université Sorbonne Nouvelle Paris 3
1 rue Maurice Arnoux, 92120 Montrouge, France

Résumé. Nous proposons dans cet article deux conceptions de la notion d'antécédent, telle qu'elle contribue à l'interprétation référentielle de l'anaphore pronominale. Tandis que du point de vue de la sémantique lexicale, l'antécédent lexical permet d'introduire ou de réintroduire le référent dans la conscience du destinataire, l'antécédent cognitif sert à maintenir le haut degré de saillance référentielle. En analysant des exemples et les résultats de données textuelles en français et en chinois, notre but est de montrer que malgré la différence distributionnelle des pronoms lexicaux et zéro dans les deux langues, les deux langues ont toutes besoin des deux types d'antécédent pour l'interprétation des anaphores pronominales dans un texte.

Abstract. *Lexical and cognitive approaches of antecedent: a contrastive study of pronominal anaphora in French and Chinese.* We propose in this article two conceptions of the notion of antecedent, as it contributes to the referential interpretation of pronominal anaphora. While the lexical approach of antecedent facilitates the introduction or reintroduction of the referent into the consciousness of the addressee, the cognitive antecedent is used to maintain the high degree of referential saliency. By analyzing linguistic examples and statistical results of textual data in French and Chinese, our goal is to show that, despite the distributional difference of lexical and zero pronouns in French and in Chinese, both languages require the two kinds of antecedent for the interpretation of pronominal anaphora in a text.

1 Introduction

Contrairement aux noms propres ou aux expressions nominales, les pronoms personnels n'ont pas de caractère descriptif, et leur fonction référentielle ne leur permettrait pas une existence autonome dans un discours écrit. Pour déterminer le référent des anaphores pronominales, il est nécessaire de recourir à la tâche de la résolution anaphorique qui

* Auteurs de correspondance: violainehou@gmail.com, frederic.landragin@ens.fr

consiste souvent à associer le terme pronominal à un terme antécédent. Cependant, la notion d'antécédent ne fait pas toujours l'unanimité et présente parfois un caractère complexe. D'un côté, deux niveaux d'analyses de l'antécédent peuvent être distingués, à savoir un niveau formel qui met l'accent sur la manifestation linguistique de l'antécédent par un segment du texte, et un niveau mental et cognitif qui considère l'antécédent comme la représentation mentale du référent construite par le destinataire. De l'autre côté, dans le cas de l'anaphore pronominale, les avis divergent sur la possibilité d'interpréter l'anaphore à travers sa dernière désignation et de considérer cette dernière comme antécédent.

Dans cet article, nous proposons une mise en question de la notion d'antécédent, tout en s'interrogeant sur les manières dont un antécédent peut contribuer à établir le lien entre un pronom (lexical ou zéro) et son référent. Les analyses s'appuient à la fois sur des exemples attestés ou construits et des données statistiques de l'extrait de texte *le Ventre de Paris* et de sa traduction en chinois. L'approche contrastive que nous adoptons ici nous permet de rendre compte que, malgré l'apparition plus ou moins contrainte ou fréquente des anaphores zéro dans les deux langues, nous retrouvons les mêmes façons de récupérer l'information référentielle et une distribution similaire des deux types d'antécédent dans les extraits de texte.

2 Problématique et objet de recherche

Traditionnellement, les pronoms anaphoriques sont décrits comme des « substituts abrégatifs » (Benveniste, 1966 : 256) dont la fonction est de remplacer un élément matériel dans un énoncé. On considère que pour une anaphore pronominale, la forme de son antécédent (s'il existe) doit toujours être une expression référentielle pleine (Corblin, 1985 ; Cornish, 1999 ; Mitkov, 2002 ; Jiang, 2016). Sous cette perspective traditionnelle, l'antécédent est vu comme une expression sémantiquement et référentiellement autonome (Cornish, 1999), et nous avons une relation de dépendance entre l'anaphore et son antécédent, à la différence de la relation symétrique de coréférence « entre des termes interprétables indépendamment l'un de l'autre » (Corblin, 1985 : 178). La forme de l'antécédent sert à fournir un apport du contenu sémantique permettant d'accéder au référent. Beaucoup d'analyses s'articulent autour des questions des formes de l'anaphore (pronominale ou nominale), de l'emplacement de l'anaphore et de l'antécédent (différence entre l'anaphore et la cataphore) et de la relation sémantique (synonymique, hyperonymique, etc) entre l'anaphore et son antécédent. L'accent est mis sur une relation entre deux segments textuels, le rôle d'autres segments du texte, ainsi que d'autres expressions susceptibles de coréférencer à l'anaphore et à son antécédent dans un contexte textuel plus large, semble peu prioritaire. Jiang (2016) considère que l'antécédent d'une anaphore est par défaut le GN (plein) le plus proche, même si la référence anaphorique dans un discours peut être maintenue par d'autres types d'expressions (les pronoms zéro par exemple) au milieu. Dans la mesure où plusieurs expressions en GN plein (non-pronominal) sont coréférentielles avec l'anaphore, la tâche de la résolution anaphorique est jugée comme réussie si l'une de ces expressions est identifiée comme l'antécédent (Mitkov, 2002). Dans les exemples (1) en français et (2) en chinois, l'anaphore pronominale *il* ou *tā* reçoit son interprétation de l'identité référentielle de son antécédent *Florent*. Cette interprétation est tout à fait envisageable si la phrase (1) se présente comme une phrase indépendante, ou si elle se situe au milieu co-textuel et constitue une composante parmi une suite de phrases.

(1) Alors, **Florent** se mit sur **son** séant. **Il** avait dormi, ... [*Le Ventre de Paris*, Émile Zola (extrait)]

(2) 弗洛朗一屁股坐了起来, 他睡着了, ...

Fúluòlǎng	yī	pìgu	zuò	le	qǐlái,	tā	shuǐzháo	le, ...
florent	un	fesse	s'assoier	PFV	lever-venir	3SG	dormir	PFV

« Florent se mit sur son séant. Il avait dormi, ... » [*Bāli de dùzi* 'Le Ventre de Paris', Émile Zola, traduction du français par Jīn kēngrán et luòxuéjuān (extrait)]

Dans des recherches récentes, certains linguistes (Stirling et Huddleston, 2002 ; Demol, 2011 ; Demol et Tobback, 2011) considèrent qu'il est possible de lier une anaphore (surtout pronominale) avec sa dernière mention, c'est-à-dire avec la dernière désignation du même référent et d'envisager une possibilité de considérer la dernière mention comme l'antécédent de l'anaphore. La condition préalable de ces approches est qu'on place l'anaphore et l'antécédent dans le contexte textuel où il y a plusieurs maillons dans une chaîne de référence (Stirling et Huddleston, 2002). Dans ce cas-là, l'anaphore et son antécédent sont traités comme deux maillons successifs qui sont extraits dans l'ensemble d'une chaîne de référence (s'il existe au moins trois expressions coréférentielles) et qui sont liés directement ou indirectement avec les autres maillons précédents. Selon Demol (2011), toutes les mentions (maillons) d'une même chaîne contribuent à la représentation mentale (cf. section 3) du référent : « Ce sont des traces linguistiques, en particulier la dernière mention (ou la mention précédente), qui nous informent en tant que chercheur sur les processus cognitifs et les opérations interprétatives. Notons que la dernière mention peut fort bien elle-même revêtir une forme sous-spécifiée » (*ibid.* : 33). Dans l'exemple (3), nous avons un bout de chaîne avec la succession de maillons « Florent – son – il – il – il ». Il n'est pas concevable de prétendre que l'antécédent du cinquième maillon *il* est le troisième ou le quatrième maillon *il* si on traite uniquement la phrase (3ii). Mais si on se situe au niveau textuel, la succession des mentions « Florent – son – il – il – il » sert à assurer une continuité référentielle, dans la mesure où chaque maillon constitue un ancrage pour le maillon suivant. De même en chinois (l'exemple 4), l'interprétation de la dernière anaphore zéro dans le segment de chaîne « Fúluòlǎng – tā – Ø – Ø » n'est pas indépendante de l'occurrence des deux maillons précédents.

- (3) i. Alors, **Florent** se mit sur **son** séant.
 ii. **Il** avait dormi, **il** ne sentait plus la faim ; **il** était tout hébété. [*Le Ventre de Paris*, Émile Zola (extrait)]
- (4) 弗洛朗一屁股坐了起来，他睡着了，也不感觉到饿了，显得很迟钝。
Fúluòlǎng yī pìgu zuò le qǐlái, tā shuìzháo le,
 florent un fesse s'asseoir PFV lever-venir 3SG dormir PFV
 Ø yě bù gǎnjuédao è le, Ø xiǎndé hěn chídùn.
 aussi NEG sentir faim PFV avoir.1.air très hébété
 « Florent se mit sur son séant. Il avait dormi, il ne sentait plus la faim ; il était tout hébété. »
 [*Bāli de dùzi* 'Le Ventre de Paris', Émile Zola, traduction du français par Jīn kēngrán et luòxuéjuān (extrait)]

Face à ces deux points de vue sur la notion d'antécédent, nous nous demandons si, dans le contexte textuel, à l'intérieur d'une même chaîne de référence, l'antécédent d'un pronom lexical ou zéro doit toujours être un GN plein. Si le deuxième point de vue sur l'antécédent peut paraître controversé, est-ce qu'on peut expliquer sa raison d'être si l'on se place dans une perspective cognitive ? Entre un antécédent considéré comme le GN plein le plus récent et un antécédent vu comme la dernière mention, au lieu de nier l'un au profit de l'autre, est-ce qu'on peut imaginer deux fonctionnements d'antécédent et deux modèles qui se complètent et qui permettent de récupérer différemment l'information référentielle ? En outre, s'il est communément admis que les anaphores zéro sont plus fréquemment utilisées en chinois, est-ce que la différence en termes de répartition des pronoms lexicaux et zéro peut influencer la distribution des deux fonctionnements d'antécédent dans les deux langues ?

Au cours de cet article, nous allons chercher à montrer que, pour les expressions anaphoriques pronominales dans une chaîne de référence, deux conceptions de l'antécédent pourraient être envisagées. Nous nous proposons de poser les hypothèses suivantes :

- Du point de vue de la sémantique lexicale, on aurait besoin du GN plein pour l'établissement (ou le rétablissement) référentiel des entités dans un texte.
- Du point de vue cognitif, l'antécédent cognitif permet d'assurer l'état actif du référent et de maintenir son haut degré de saillance.
- Dans une perspective contrastive, la fréquence plus importante des pronoms zéro en chinois et leur saillance extrêmement forte auraient pour résultat une plus grande proportion des antécédents cognitifs.

3 Antécédent : représentation formelle et représentation mentale

Avant de discuter nos deux conceptions d'antécédent, il nous semble nécessaire de distinguer deux niveaux d'analyse de l'antécédent, le premier est le niveau formel, et le deuxième est lié à la représentation mentale.

Traditionnellement, la notion d'antécédent est toujours liée à celle de l'anaphore, considérée comme une expression sémantiquement déficiente, « saturée » par les propriétés sémantiques et lexicales de son antécédent (Apothéloz, 1995 ; Cornish, 1999 ; Gardelle et al., 2019). Ce point de vue traditionnel de l'antécédent privilégie l'analyse de l'occurrence de l'antécédent en mettant l'accent sur la forme de l'antécédent tel qu'il existe dans le texte précédent (ou ultérieur). Pour distinguer l'aspect formel/existential de l'aspect sémantique/référentiel de l'antécédent, Cornish (1999) propose d'utiliser le terme *déclencheur d'antécédent* (en anglais *antecedent-trigger*) pour désigner l'expression linguistique qui introduit le référent de l'anaphore dans le discours, et le terme *antécédent* pour la représentation mentale et sémantique de l'antécédent. Selon l'auteur, la représentation mentale de l'antécédent se développe en même temps que le discours progresse. Les caractéristiques associées à l'antécédent peuvent connaître une accumulation ou même un changement au cours de l'évolution textuelle.

La distinction entre le niveau formel et le niveau mental de l'antécédent nous permet de rendre compte que l'importance des chaînes de référence ne se limite pas aux expressions référentielles qui en font parties. Au fur et à mesure que le texte progresse, non seulement chaque expression est enregistrée dans la chaîne, mais les informations véhiculées par les prédications sont également stockées et liées à la représentation mentale du référent. Nous avons de ce fait une sorte de réseau d'information qui s'articule autour du référent, comme le montre la figure 1. Cependant, un lecteur qui lit un texte ne retient pas toutes les expressions et toutes les informations à un moment précis. Quand un lecteur traite une phrase, certaines informations ou certains éléments d'information peuvent être activés ou réactivés, alors que d'autres aspects sont laissés au second plan. Il conviendrait de s'appuyer sur un modèle dynamique pour comprendre l'état d'attention d'un lecteur.

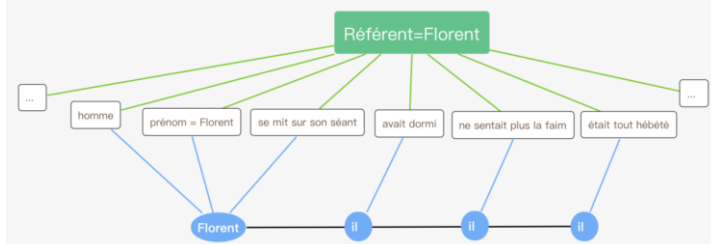


Fig. 1. La représentation mentale du référent *Florent* (dans l'exemple 3)

En se situant dans une approche plutôt mentaliste, Apothéloz (1995 : 310-311) met en question la nécessité de l'existence d'un (déclencheur d') antécédent ou d'une dénomination préalable du référent dans un texte : « le segment de texte traditionnellement nommé antécédent est considéré ici comme seulement la trace de l'opération langagière qui

a rendu accessible, dans la représentation, l'objet désigné par l'anaphorique. Mais une autre opération, non langagière (geste, mimique, regard, posture) ou, de façon générale, une autre cause (évidence perceptive) pourraient tout aussi bien avoir eu le même effet sur l'état de la représentation. » Selon l'auteur, l'identification référentielle de l'anaphore ne s'appuie pas sur un segment de texte, mais plutôt sur une représentation mentale de l'entité. Néanmoins, il est probable que l'anaphore ait une relation (coréférentielle ou non) avec sa précédente désignation.

Nos deux conceptions s'inspirent des deux dernières approches tout en y intégrant nos réflexions sur les rôles des différentes formes de surface. D'un côté, il nous semble nécessaire de distinguer le niveau référentiel/mental de l'antécédent de sa réalisation formelle dans le texte. De l'autre côté, nous considérons que, au lieu de pointer directement l'anaphore vers son antécédent référentiel, l'antécédent formel a aussi une fonction de rendre l'entité désignée par l'anaphore accessible dans la représentation mentale du destinataire. Et en fonction des formes différentes que pourraient prendre la dernière désignation, il est possible d'envisager deux conceptions de l'antécédent formel et deux modèles d'accès à l'information référentielle.

4 Conception lexicale de l'antécédent

Dans la section précédente, nous avons illustré que la représentation mentale de l'antécédent est le résultat du traitement du texte précédent par le destinataire. Dans la section 4 et la section 5, nous nous intéressons aux différentes manières d'accéder à cette représentation mentale à travers les formes différentes de l'antécédent dans le texte. Nous allons montrer, à l'aide des exemples inventés ou des données du corpus, que les rôles joués par l'antécédent en GN plein et par l'antécédent en d'autres formes possibles (en l'occurrence les pronoms) relèvent de deux fonctionnements différents. Pour la conception lexicale de l'antécédent, l'accès au référent est établi ou rétabli par le contenu lexical/sémantique de l'antécédent. En ce qui concerne la conception cognitive de l'antécédent, l'accès au référent est garanti par le maintien du statut de forte saillance du référent dans la conscience du destinataire.

4.1 L'antécédent en GN plein et l'introduction référentielle

Les analyses sur les chaînes de référence mettent souvent l'accent sur les types d'expression référentielle susceptibles d'occuper la position initiale dans une chaîne (Chastain, 1975 ; Corblin, 1985, 1987, 1995). Quand un référent est nouvellement introduit dans le discours, il faudrait fournir les éléments sémantiques nécessaires pour qu'un destinataire puisse construire ce référent dans son dossier mental. Ainsi, la mention initiale se fait très souvent sous la forme d'une expression référentielle pleine. Dans l'analyse d'Antoine et al. (2016), 92.4% de premières mentions (la première mention est le premier maillon d'une chaîne de référence) sont introduites par des GN. A travers l'exploitation des données de corpus en français et en chinois, nous obtenons les résultats (illustrés dans le tableau 1) assez similaires.

Tableau 1. Catégorie de la première mention dans les extraits de texte *le Ventre de Paris*

Catégorie de la première mention d'une chaîne (fr)	Pourcentage
GN plein	92,39%
Pronom	7,61%
Catégorie de la première mention d'une chaîne (ch)	Pourcentage
GN plein	97,09%
Pronom	2,91%

Dans les chaînes de référence, le pourcentage considérable des cas où un GN plein apparaît comme la première mention montre le rôle important d'un antécédent plein dans l'établissement référentiel. Cette procédure d'établissement est illustrée par la figure 2, dans laquelle l'expression en GN a réussi à construire un référent dans la conscience du destinataire à travers sa propriété lexico-sémantique. On peut aussi ajouter à cette représentation mentale les caractéristiques exprimées dans la phrase où le GN apparaît. Et si par la suite, une relation anaphorique est établie entre ce GN et un pronom, le pronom trouve son accès grâce à la représentation mentale établie par le GN précédent.

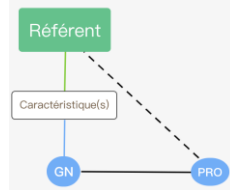


Fig. 2. L'établissement référentiel par un GN plein

Dans le cas de la première mention occupée par un pronom anaphorique, elle est plutôt perçue comme une cataphore dont l'antécédent (ou plus exactement « interprétant » [Ducrot et Todorov, 1972 : 358], ou encore « source sémantique » [Tesnière, 1959]) se trouve dans le discours suivant (Gardon, 1982 ; Tanaka, 2000). Dans la recherche de Kesik (1989), la cataphore est décrite comme « relation non structurale d'une expression indexicale (à signifiant zéro éventuellement) avec le (une partie du) contexte linguistique subséquent, telle que ce contexte permet (et est nécessaire à) l'identification du référent de l'expression indexicale » (*ibid* : 56). Reichler-Béguelin (1988) remarque un fonctionnement distinct entre les processus d'encodage et de décodage de la cataphore, dans la mesure où « le référent visé fait partie des connaissances disponibles du locuteur » (*ibid* : 38), alors que la validation en mémoire du référent visé intervient après-coup dans le processus de décodage. En comparant les textes chinois et anglais, Xu et He (2007) constatent que le phénomène de cataphore intraphrastique est moins fréquent en chinois qu'en anglais. En plus, à la place du pronom cataphorique lexical (comme *il* dans l'exemple 5 en français), une cataphore en pronom zéro est plus souvent utilisée en chinois (l'exemple 6), car l'utilisation du pronom lexical peut entraîner une interprétation non coréférentielle entre la cataphore et son interprétant. Comme dans l'exemple (7), en lisant cette phrase, un lecteur peut considérer que le pronom *tā* renvoie à un autre référent que celui du GN *wō de línjū*.

(5) Lorsqu'**il** est malade, **mon voisin** m'appelle. [Reichler-Béguelin, 1988 : 37]

(6) 生病的时候，我的邻居会打电话给我。

Ø shēngbìng de shíhòu, wǒde línjū huì
tomber.malade DE moment mon voisin être.probable.de
dǎdiànhuà gěi wǒ.
appeler à 1SG
« Lorsqu'**il** est malade, **mon voisin** m'appelle. »

(7) 他生病的时候，我的邻居会打电话给我。

Tā shēngbìng de shíhòu, wǒde línjū huì
3SG tomber.malade DE moment mon voisin être.probable.de
dǎdiànhuà gěi wǒ.
appeler à 1SG
« Lorsqu'**il** est malade, **mon voisin** m'appelle. »

Nous considérons que la cataphore n'est pas une simple anaphore anticipante et qu'elle présente sa propre spécificité. Dans l'approche cognitive et textuelle que nous adoptons, il semble que l'existence du référent de la cataphore est bien posée par la première mention, et que ce référent se voit déjà associer les caractéristiques exprimées par la phrase où la

première mention apparaît, avant que l'identité du référent soit déterminée ou précisée par l'interprétant. Il convient de noter que le fonctionnement d'accès référentiel de la cataphore nécessite une étude plus approfondie, ce qui dépasse le cadre de cet article. Ainsi, nos deux conceptions de la notion d'antécédent ne s'appliquent qu'à l'anaphore au sens étroit du terme.

4.2 La réintroduction référentielle par un antécédent en GN plein

Dans le contexte textuel (et surtout celui de narration), il se peut qu'un référent qui a déjà été construit mais qui a subi une rupture dans le développement textuel soit réintroduit dans le texte. Ce référent qui a été actif mais qui ne l'est plus dans la conscience du destinataire avant le traitement de la phrase en question se voit souvent réintroduit dans le discours par un GN plein. Comme dans l'exemple (8) en français et l'exemple (9) en chinois, les noms propres *madame François* et *Fúlǎngsuǒwǎ tàitài* sont employés pour réintroduire leur référent après deux paragraphes textuels (ou après 243 mentions différentes en français et 255 mentions différentes en chinois).

- (8) — Eh ! l'homme, là-haut ! cria brusquement **madame François**.
Et, comme il ne bougeait pas, **elle** monta, **Ø** le secoua. [*Le Ventre de Paris*, Émile Zola (extrait)]
- (9) “喂，上面那个人！”弗朗索瓦太太粗声粗气地喊道。
没有动静。她爬上车，摇着他。
“Wèi, shàngmiàn de nàge rén!” **Fúlǎngsuǒwǎ tàitài**
eh dessus DE ce homme François madame
cūshēngcūqide hǎndào.
d'une voix.bourrue crier
Méiyǒu dòngjìng. **Tā** páshàng chē, **Ø** yáo zhe tā.
NEG mouvement 3SG monter voiture secouer DUR 3SG
« — Eh ! l'homme, là-haut ! cria brusquement **madame François**.
Et, comme il ne bougeait pas, **elle** monta, **Ø** le secoua. » [*Bǎli de dùzi* 'Le Ventre de Paris',
Émile Zola, traduction du français par Jīn kēngrán et luòxuějuān (extrait)]

Dans une réintroduction référentielle, le GN permet de rétablir le lien entre un segment présent dans le texte et la représentation mentale du référent, avec éventuellement les propriétés associées qui sont encore actives dans la conscience du destinataire. Ce rôle de restauration du lien référentiel est plus souvent assumé par un GN que par un pronom, comme l'implique le tableau 2. Pour nos deux extraits de texte, nous avons mené les statistiques descriptives qui portent sur la distance entre l'expression susceptible d'être un antécédent lexical (en GN plein) ou cognitif (en pronom) et sa dernière mention (DM). Cette distance est mesurée par le nombre de mentions non coréférentielles avec l'expression en question. Les résultats montrent que les GN pleins sont généralement plus éloignés (ou possèdent plus de « compétiteurs » dans le sens d'Ariel (1990)) de la dernière mention que les pronoms. Il est vrai que certains GN pleins peuvent être eux-aussi des anaphores nominales qui ne servent pas à réintroduire un référent, et que l'état de conscience du destinataire (ou l'état de désactivation d'un référent) est difficilement mesurable, mais ces résultats semblent indiquer que les GN sont plus aptes à réintroduire un référent dans le texte.

Tableau 2. Distance du GN plein ou du pronom par rapport à la dernière mention dans les extraits de texte *le Ventre de Paris*

Distance par rapport à la dernière mention (en nb de mentions)				
Type d'expression (fr) : début – fin	Min	Médiane	Max	Moyenne

DM – GN plein	0	32	463	80,75
DM – Pronom	0	1	24	2,14
Type d'expression (ch) : début – fin	Min	Médiane	Max	Moyenne
DM – GN plein	0	19	588	67,08
DM – Pronom	0	1	28	2,10

Lors du processus de réintroduction, la représentation mentale du référent est réactivée par l'identité ou la conformité d'une ou plusieurs caractéristiques entre l'expression en GN et le référent, comme l'illustre la figure 3. Tandis que les caractéristiques du référent sont accumulées grâce aux traitements du texte précédent, celles de l'expression en GN sont associées à la dernière à travers sa propriété lexico-sémantique ou l'information véhiculée par la phrase en question. Et une fois qu'une relation de coréférence sera établie entre l'expression en GN et un pronom, ce terme se verra associé à la représentation mentale du référent qui redevient actif dans la conscience du destinataire.

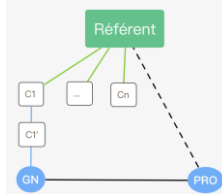


Fig. 3. La réintroduction référentielle par un GN plein

5 Conception cognitive de l'antécédent

Dans la section précédente, nous avons illustré les façons dont un antécédent en GN permet d'établir et de récupérer la représentation mentale du référent dans la conscience du destinataire. Dans cette section, nous allons d'abord expliquer les raisons pour lesquelles dans le cadre de la chaîne de référence, l'interprétation du pronom anaphorique n'est pas toujours directement liée au GN plein le plus proche. Nous montrerons ensuite comment un antécédent cognitif (en forme pronominale) qui joue un rôle de « maintenir » référentiel, permet de construire le lien entre l'anaphore et la représentation mentale du référent.

5.1 Arguments à l'appui d'une conception cognitive de l'antécédent

Afin de répondre à notre question de recherche formulée dans l'introduction et de jeter la base pour le deuxième point de notre hypothèse, trois arguments sont avancés dans cette sous-section.

Le rôle de l'antécédent cognitif dans l'interprétation du pronom anaphorique est d'abord appuyé par les exemples construits en chinois et en français. Dans (10a, b et c) comme dans (11a, b et c), les formes de surface sont presque identiques, et nous avons délibérément créé deux référents en compétition qui se voient d'abord introduire dans la première proposition par les noms propres.

Dans (10a) et (10b), le premier pronom zéro renvoie respectivement à *Xiāohóng* et à *Xiāomíng*, le deuxième pronom zéro est toujours coréférentiel avec le premier pronom zéro. Autrement dit, dans (10a) ou (10b), il s'agit d'une chaîne « GN – pronom zéro – pronom zéro » qui se construit autour du référent *Xiāohóng* ou du référent *Xiāomíng*. Quand on compare les phrases (10a, b) et (10c), si l'on considère que l'interprétation du deuxième pronom zéro dépend toujours du GN dans la première proposition, on pourrait s'imaginer que la phrase (10c) serait grammaticale, c'est-à-dire que le premier pronom pourrait renvoyer à un GN dont le référent est autre que celui du deuxième pronom. Or, cette

interprétation n'est pas possible. De ce fait, il est plus pertinent de considérer que l'interprétation référentielle du deuxième pronom anaphorique est plutôt directement liée à sa dernière mention qui tient compte de l'évolution de la saillance référentielle dans le discours.

(10) a. 小红看到了小明，从包里取出一张名片，放在了桌子上。

[Xiǎohóng] _i	kàndào	le	[Xiǎomíng] _j				
Xiaohong	voir	PFV	Xiaoming				
[Ø] _i	cóng	bāo	lǐ	qǔchū	yī	zhāng	míngpiàn,
	de	sac	dedans	retirer	un	CLF	carte.de.visite
[Ø] _i	fàng	zài	le	zhuōzi	shàng.		
	placer	à	PFV	table	sur		

« Xiaohong a vu Xiaoming, elle a sorti une carte de visite de son sac et l'a mise sur la table. »

b. 小红看到了小明，从包里取出一张名片，放在了桌子上。

[Xiǎohóng] _i	kàndào	le	[Xiǎomíng] _j				
Xiaohong	voir	PFV	Xiaoming				
[Ø] _j	cóng	bāo	lǐ	qǔchū	yī	zhāng	míngpiàn,
	de	sac	dedans	retirer	un	CLF	carte.de.visite
[Ø] _j	fàng	zài	le	zhuōzi	shàng.		
	placer	à	PFV	table	sur		

« Xiaohong a vu Xiaoming, il a sorti une carte de visite de son sac et l'a mise sur la table. »

c. 小红看到了小明，从包里取出一张名片，放在了桌子上。

[Xiǎohóng] _i	kàndào	le	[Xiǎomíng] _j				
Xiaohong	voir	PFV	Xiaoming				
[Ø] _i	cóng	bāo	lǐ	qǔchū	yī	zhāng	míngpiàn,
	de	sac	dedans	retirer	un	CLF	carte.de.visite
*[Ø] _j	fàng	zài	le	zhuōzi	shàng.		
	placer	à	PFV	table	sur		

« Xiaohong a vu Xiaoming, elle a sorti une carte de visite de son sac et il l'a mise sur la table. »

De même pour les phrases (11a, b et c) en français, nous avons un pronom zéro dans la troisième proposition de chaque phrase, sauf qu'en français, la trace du sujet zéro peut être retrouvée dans l'accord du verbe. Néanmoins, cet accord ne permet pas de distinguer entre les deux référents possibles (*Jean* et *Marie*), puisque le genre n'est pas marqué par le verbe. Dans les phrases (11a, b), nous avons une chaîne « GN – pronom lexical – pronom zéro » concernant respectivement *Jean* ou *Marie*. Au lieu de recevoir l'information référentielle directement à partir du GN plein, le troisième maillon (le pronom zéro) trouve plutôt son lien vers le référent à travers le deuxième maillon qui est réalisé par un pronom lexical dans le discours.

(11) a. [Jean]_i a rencontré [Marie]_j dans le bureau. [Il]_i a rendu un rapport, et [Ø]_j a montré les avantages et les inconvénients.

b. [Jean]_i a rencontré [Marie]_j dans le bureau. [Elle]_j a rendu un rapport, et [Ø]_j a montré les avantages et les inconvénients.

c. [Jean]_i a rencontré [Marie]_j dans le bureau. [Il]_i a rendu un rapport, et *[Ø]_j a montré les avantages et les inconvénients.

Suivant ce premier argument qui repose sur la contrainte imposée par l'antécédent cognitif exigeant une continuité référentielle, nous présentons en second lieu que, dans le cas du référent évolutif (Charolles, 2001 ; Kleiber et Riegel, 2003), l'interprétation de la mention qui apparaît à la fin d'une chaîne est plus directement et étroitement liée à sa dernière désignation en pronom lexical ou zéro. Au fur et à mesure que les événements décrits dans le texte se déroulent, les propriétés associées au référent peuvent changer ou même se trouver en contradiction avec les caractéristiques initialement construites par le GN plein.

Dans l'exemple (12), le dernier *le* ici est interprété non pas comme faisant référence à *un poulet bien actif et bien gras*, mais plutôt au poulet résultant de la préparation et de la coupe.

(12) Tuez **un poulet bien actif et bien gras**. Préparez-**le** pour le four. Coupez-**le** en quatre morceaux et faites-**le** rôtir avec du thym pendant une heure. [Kleiber et Riegel, 2003 : 126]

(13) 宰一只活鸡，将其切成四块，用料酒，酱油和耗油腌制，之后放入锅中炖三个小时。

Zǎi	yī	zhī	huó	jī,	jiāng	qí	qiè	chéng	sì
tuer	un	CLF	vivant	poulet	BA	3SG	couper	devenir	quatre
kuài,		yòng	liàojiǔ,		jiàngyóu		hé	hàoyóu	
morceaux		utiliser	vin.de.cuisine		sauce.de.soja		et	sauce.d'huître	
yānzhi	Ø,	zhīhòu	fàng	Ø	rù	guō	zhōng	dùn	
mariner		après	placer		entrer	casserole	dans	mijoter	
sān	gè	xiǎoshí.							
trois	CLF	heure							

« Tuez **un poulet bien actif**, coupez-**le** en quatre morceaux, faites-**le** mariner avec du vin de cuisine, de la sauce de soja et de la sauce d'huître, puis faites-**le** mijoter pendant trois heures dans une casserole. »

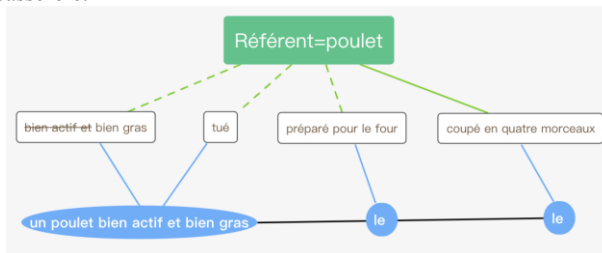


Fig. 4. La représentation mentale du référent évolutif *poulet* avant le traitement de la dernière proposition

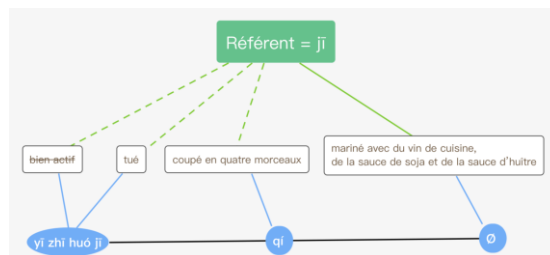


Fig. 5. La représentation mentale du référent évolutif *jī* 'poulet' avant le traitement de la dernière proposition

La figure 4 représente le résultat du changement de la représentation mentale du référent en question avant le traitement de la dernière proposition « faites-*le* rôtir avec du thym pendant une heure ». Les lignes pointillées indiquent que les propriétés anciennement associées au référent sont susceptibles d'être plus ou moins actives dans la conscience du destinataire. La propriété « bien actif » est supprimée puisque le sens de l'adjectif est en contradiction avec le sens de « tué » qui est plus récemment ajouté comme propriété du référent.

Si la présence physique des pronoms est obligatoire dans l'exemple (12) en français, dans l'exemple (13) en chinois, le référent évolutif *jī* 'poulet' est susceptible d'être désigné par les pronoms zéro, sans que cela n'entraîne une contrainte pour les deux formes zéro d'être supports des descriptions de changements. Dans cet exemple, le dernier pronom ne peut pas être conçu comme renvoyant au référent tel qu'il est présenté par le GN *yī zhī huó jī* 'un poulet vivant', mais à la représentation mentale du référent résultant des successions de modification, comme l'illustre la figure 5.

Le troisième argument a trait à la distance entre l'antécédent et l'anaphore. Dans le contexte textuel, il peut y avoir une très grande distance entre le dernier GN plein et

l'anaphore. Cette distance pourrait dépasser une distance normalement permise pour un lien direct : ce sont donc les liens intermédiaires qui maintiennent la continuité du référent dans le texte et la saillance du référent dans la conscience du lecteur, de sorte que la référence à celui-ci peut être faite à l'aide d'un pronom personnel ou d'une autre anaphore avec peu de contenu intrinsèque.

Dans la littérature, la distance (ou la récence) est considérée comme un facteur qui mesure la continuité topicale (Givón, 1983), ou qui détermine l'accessibilité (Ariel, 1990) ou la saillance (Landragin, 2004 ; Hou et Landragin, 2019, sous presse) du référent. Cette mesure se fait souvent par le nombre de phrases (Givón, 1983 ; Ariel, 1990 ; Xu, 2000) ou par le nombre de GN (Boudreau et Kittredge, 2005) ou de mentions (Poon et Domingos, 2008). Afin de savoir si l'on peut avoir une conception cognitive de l'antécédent qui sert à maintenir le lien référentiel, et si la distance entre le dernier GN plein et le pronom anaphorique peut être effectivement très longue pour que l'identité référentielle du pronom puisse être résolue à travers le dernier GN plein, nous avons effectué deux exploitations pour les anaphores pronominales dans les extraits de texte *le Ventre de Paris* en français et en chinois. Pour la première exploitation, nous avons calculé la distance en nombre de mentions entre une anaphore pronominale et sa dernière expression coréférentielle pleine. Dans la deuxième exploitation, la distance est calculée entre une anaphore pronominale et sa dernière mention, que cette dernière soit en forme nominale ou pronominale. Autrement dit, dans les deux exploitations, nous visons à mesurer la distance entre un pronom et son « antécédent ». Dans la première exploitation, l'antécédent doit toujours être un GN plein. Alors que la deuxième exploitation consiste à mesurer la distance en nombre de mentions non coréférentielles entre un pronom et son antécédent, qu'il s'agisse d'un antécédent lexical ou cognitif. Pour comparer les deux exploitations, nous résumons les résultats descriptifs dans le tableau 3.

Tableau 3. Distance du pronom par rapport au dernier GN plein ou à la dernière mention dans les extraits de texte *le Ventre de Paris*

Distance (en nb de mentions)				
Type d'expression (fr) : début – fin	Min	Médiane	Max	Moyenne
Le dernier GN plein – Pronom	0	22	123	30,27
La dernière mention – Pronom	0	1	24	2,14
Type d'expression (ch) : début – fin	Min	Médiane	Max	Moyenne
Le dernier GN plein – Pronom	0	19,5	134	30,23
La dernière mention – Pronom	0	1	28	2,10

Les résultats montrent que les distances moyennes sont beaucoup plus longues entre les pronoms et les derniers GN plein (30,27 mentions en français et 30,23 mentions en chinois) qu'entre les pronoms et les dernières mentions (2,14 mentions en français et 2,10 mentions en chinois). Un phénomène similaire peut être observé pour les distances en médiane. Entre un pronom et le dernier GN plein, on peut même avoir une distance maximale de 123 mentions en français, et de 134 mentions en chinois.

Si à chaque fois qu'un lecteur rencontre un pronom, il remonte toujours à la dernière mention pleine pour que cette dernière puisse transmettre ses propriétés sémantiques à l'anaphore pronominale, le processus sera cognitivement beaucoup plus coûteux. Il serait plus judicieux de considérer que, entre un pronom et un antécédent en GN plein, nous avons au milieu un ou plusieurs antécédents qui servent à établir les liens intermédiaires et à maintenir la saillance du référent.

5.2 L'antécédent cognitif et le maintien de forte saillance référentielle

Nous examinons maintenant le rôle joué par l'antécédent cognitif dans l'identification anaphorique du pronom dans le contexte textuel. Si l'antécédent lexical permet d'activer ou de réactiver le dossier mental du référent en question par son contenu sémantique, les référents des pronoms doivent déjà être dans la conscience de l'auditeur lors de leur mention (Chafe, 1974). Suivant Ariel (1990), l'encodage des expressions référentielles dans le texte permet de marquer le degré de l'accessibilité ou de la saillance (Landragin, 2004) du référent. Un marqueur de haute accessibilité ou de forte saillance (comme un pronom inaccentué ou un pronom zéro) permet de refléter que le référent est supposé comme suffisamment saillant dans la conscience du destinataire pour que ce dernier puisse établir le lien référentiel entre le marqueur et la représentation mentale du référent, même si ce marqueur comporte peu de contenu lexical et donne peu d'indice pour l'accès référentiel.

Si, dans le contexte textuel, nous avons une mention en forme pronominale dont le référent a déjà été identifié par le lecteur, comme le montre la figure 6 (étape 1), au moment du traitement de la phrase dans laquelle le pronom apparaît, le référent est censé avoir une forte saillance dans la mémoire du destinataire. Et si le destinataire rencontre par la suite une autre expression de forte saillance, il est plus naturellement dirigé vers le référent qui est récemment au centre de son attention. Dans ce cas-là, la nouvelle expression en forme pronominale se voit attribuer la valeur référentielle de son antécédent cognitif (figure 6, étape 2). De ce fait, dans les exemples (10c) et (11c), la rupture du maintien référentiel et l'utilisation d'un marqueur de forte saillance pour un autre référent entre l'antécédent et son anaphore rend inacceptable une représentation formelle du référent initial par un pronom.

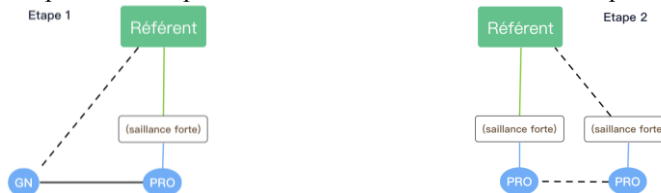


Fig. 6. Le maintien de forte saillance référentielle – étapes 1 et 2

Bien évidemment, entre deux marqueurs de forte saillance, il se peut que, avant le traitement du deuxième marqueur, un nouveau référent devienne plus saillant que le référent du premier marqueur, ce qui empêche la mise en place du mécanisme illustré dans la figure 6. Mais cela relève plutôt de la comparaison du degré de saillance entre les référents, on ne peut pas nier à ce titre l'existence du rôle joué par l'antécédent cognitif. Dans cet article, nous ne visons pas à comparer le degré de saillance des référents ou à illustrer les facteurs qui influencent la saillance des entités, mais à montrer que l'antécédent cognitif, comme l'antécédent lexical, apporte sa contribution à l'interprétation référentielle de l'anaphore pronominale.

Afin de mieux comprendre l'apport de la notion de saillance à nos deux conceptions d'antécédent, nous proposons d'illustrer par les figures 8 et 9 l'évolution du degré de saillance du référent *madame François* lors de ses quatre premières mentions dans les extraits *le Ventre de Paris* en français et en chinois (« madame François – madame François – ses – Ø » en français et « *fūlāngsuōwǎ tàitài – fūlāngsuōwǎ tàitài – tā – Ø* » en chinois). Au fil du texte, le degré de saillance d'un référent précis dans la conscience du lecteur peut évoluer. Dans les figures 8 et 9, les différentes mentions dans le texte (coréférentielle ou non) se succèdent dans le sens horizontal de gauche à droite, et le degré de saillance peut être visualisé dans le sens vertical.

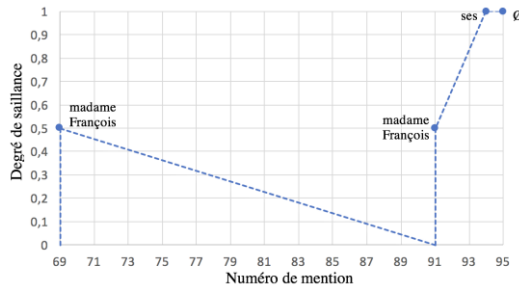


Fig. 7. L'évolution du degré de saillance du référent *madame François*

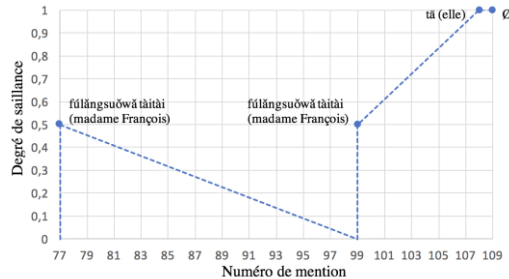


Fig. 8. L'évolution du degré de saillance du référent *fùlǎngsuǒwǎ tàitài* 'madame François'

A la fois en français et en chinois, le référent est d'abord introduit (par la mention 69 en français, ou par la mention 77 en chinois) et puis réintroduit (par la mention 91 en français, ou par la mention 99 en chinois) dans la conscience du destinataire par un GN. Et puis, la succession des pronoms reflète que la forte saillance du référent est maintenue. Alors que la mention 91 en français et la mention 99 en chinois fonctionnent comme antécédent lexical qui permet de réintroduire un référent non saillant, la mention 94 en français et la mention 108 en chinois sont des antécédents cognitifs qui marquent la forte saillance du référent et qui facilitent la mise en rapport entre les pronoms anaphoriques suivants (qui sont eux aussi les marqueurs de forte saillance référentielle) et le référent déjà saillant.

6 Discussion sur la répartition des antécédents lexicaux et cognitifs

Dans les sections précédentes, nous avons différencié deux types de manifestation linguistique de l'antécédent dans le texte qui influencent les façons dont les anaphores pronominales peuvent être résolues. Nous analysons maintenant la distribution des deux types d'antécédent dans l'extrait de texte *le Ventre de Paris* et dans sa traduction en chinois. Tout d'abord, nous observons les occurrences d'anaphores en pronom lexical ou zéro en français et en chinois. Le tableau 4 montre que parmi tous les pronoms anaphoriques qui constituent nos objets d'études, 51,38% des pronoms sont de forme zéro en chinois, alors que seulement 5,71% le sont en français, ce qui confirme le fait que l'utilisation des pronoms zéro est beaucoup plus fréquente en chinois. Selon Ariel (1990), l'antécédent des vrais pronoms zéro, comme ceux en chinois, doit être extrêmement saillant. Si les pronoms zéro sont plus fréquemment utilisés en chinois, est-ce qu'on recourt davantage aux antécédents cognitifs en chinois qui assurent déjà la forte saillance référentielle ?

Tableau 4. Répartition des anaphores dans les extraits de texte *le Ventre de Paris*

Langue	Types d'anaphore pronominale		Total
	En pronom lexical	En pronom zéro	

Français	198 (94,29%)	12 (5,71%)	210 (100%)
Chinois	123 (48,62%)	130 (51,38%)	253 (100%)

Tableau 5. Répartition des antécédents dans les extraits de texte *le Ventre de Paris*

Types d'anaphore (fr)	Types d'antécédent		
	Antécédent lexical	Antécédent cognitif	Total
Anaphore pronominale	30 (15,15%)	168 (84,85%)	198 (100%)
Anaphore zéro	4 (33,33%)	8 (66,67%)	12 (100%)
Total	34 (16,19%)	176 (83,81%)	210 (100%)
Types d'anaphore (ch)	Antécédent lexical	Antécédent cognitif	Total
Anaphore pronominale	17 (13,82%)	106 (86,18%)	123 (100%)
Anaphore zéro	33 (25,38%)	97 (74,62%)	130 (100%)
Total	50 (19,76%)	203 (80,24%)	253 (100%)

Nous résumons dans le tableau 5 la distribution des antécédents lexicaux et cognitifs selon les langues et selon les types d'anaphore. Dans les deux langues, le pourcentage des antécédents cognitifs est plus élevé que celui des antécédents lexicaux. Ceci étant dit, en lisant ces extraits de textes en français ou en chinois, le destinataire interprète plus fréquemment les référents des anaphores pronominales par le maintien du haut degré de saillance signalé par l'antécédent cognitif. Malgré le fait que le texte chinois possède plus d'anaphores zéro, parmi tous les pronoms anaphoriques, nous n'avons pas trouvé une différence significative en ce qui concerne la proportion des antécédents cognitifs en chinois (80,24%) et en français (83,81%). En outre, à la fois en français et en chinois, les pronoms zéro ne possèdent pas relativement plus d'antécédents cognitifs que les pronoms lexicaux. Cependant, comme nous avons exploité deux extraits de textes en parallèle au lieu de deux extraits comparables dans les deux langues, il se peut que la syntaxe de l'original français puisse influencer la syntaxe de la traduction. Ainsi, cette étude de caractère exploratoire pourrait se compléter avec plus de données provenant d'un corpus qui comprend à la fois des textes en parallèles et des textes comparables.

7 Remarques conclusives

Dans cet article, nous avons proposé deux conceptions de la notion d'antécédent pour les anaphores pronominales en nous plaçant dans une perspective textuelle et cognitivo-psychologique. Autrement dit, nous envisageons la relation anaphorique en prenant en considération non seulement les différentes mentions faisant parties de la même chaîne de référence, mais également les informations autour du référent, apportées par d'autres éléments du texte.

Malgré le fait que les occurrences de l'anaphore zéro soient plus contraintes et moins fréquentes en français qu'en chinois, nous n'avons pas trouvé une plus grande proportion des antécédents cognitifs en chinois, au moins dans les extraits que nous analysons. Néanmoins, les deux langues ont besoin à la fois des antécédents lexicaux et des antécédents cognitifs pour l'interprétation référentielle des anaphores pronominales. Du point de vue lexical, l'antécédent en GN plein permet de construire la représentation mentale d'un nouveau référent ou de réintroduire le référent dans la conscience du destinataire. Il joue un rôle de « déclencheur » pour (r)ouvrir le dossier mental du référent. En ce qui concerne l'antécédent cognitif, il joue un rôle de « mainteneur » qui assure l'état actif du dossier et qui maintient le haut degré de saillance du référent, de sorte qu'une autre expression qui marque une forte saillance référentielle peut facilement y être associée. Enfin, les statistiques basées sur les extraits de texte en chinois et en français indiquent que

les antécédents cognitifs auraient une place plus importante ou au moins aussi importante que les antécédents lexicaux lors du processus d'interprétation de la plupart des anaphores pronominales dans un texte. En perspective de cette étude, il serait intéressant d'analyser comment les deux types d'antécédent permettent de contribuer à l'augmentation du degré de saillance référentielle ou de maintenir la forte saillance. En d'autres termes, il s'agit de mettre en œuvre une analyse des facteurs de saillance associés aux deux types d'antécédent et de comparer éventuellement les facteurs en français avec ceux dans d'autres langues comme le chinois.

Références bibliographiques

- Antoine, J.-Y., Lefeuve, A., Schang, E. (2016). Codage en chaîne ou en première mention de la coréférence : approcher la structure des chaînes de référence par comparaison des deux annotations. in *5ème congrès mondial de linguistique française (cmlf 2016)*.
- Ariel, M. (1990). *Accessing noun-phrase antecedents*. London : Routledge.
- Benveniste, É. (1966). *Problèmes de linguistique générale*. 2 vols. Paris, France: Gallimard.
- Boudreau, S., Kittredge, R. (2005). Résolution des anaphores et détermination des chaînes de coréférences: différences entre variétés de textes. *TAL. Traitement Automatique Des Langues* 46, 41–69.
- Carden, G. (1982). Backwards anaphora in discourse context. *Journal of Linguistics* 18, 361–87.
- Chafe, W.L. (1974). Language and consciousness. *Language* 50, 111–133.
- Charolles, M. (2001). Référents évolutifs et évolution de la référence. In *Les référents évolutifs entre linguistique et philosophie*, 39–97. Klincksieck.
- Charolles, M. (2002). *La référence et les expressions référentielles en français*. Paris : Ophrys.
- Chastain, C. (1975). Reference and context. *Language, mind, and knowledge* 07, 194–269.
- Corblin, F. (1985a). Les chaînes de référence : analyse linguistique et traitement automatique. *Intellectica* 1 : 123–43.
- Corblin, F. (1985b). Remarques sur la notion d'anaphore. *Revue québécoise de linguistique* 15, 173–95.
- Corblin, F. (1995). *Les formes de reprise dans le discours : anaphores et chaînes de référence*. Rennes, France: Presses universitaires de Rennes.
- Cornish, F. (1999). *Anaphora, discourse, and understanding: evidence from English and French*. Oxford : Clarendon Press.
- Demol, A. (2010). *Les pronoms anaphoriques "il" et "celui-ci"*. Bruxelles : Editions Duculot.
- Demol, A., Tobback, E. (2011). « Je crois que c'est une histoire très inhabituelle » Le démonstratif ce dans les copulatives à prédicat nominal, un marqueur de discontinuité topicale ? *Zeitschrift Für Französische Sprache Und Literatur* 121, 113–36.
- Ducrot, O., Todorov, T. (1972). *Dictionnaire Encyclopédique Des Sciences Du Langage*. Paris: Seuil.
- Givón, T. (1983). *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Vol. 3. Typological Studies in Language. Amsterdam: John Benjamins Publishing Company.
- Hou, J., Landragin, F. (2019, sous presse). La saillance en français et en chinois: Approche multifactorielle et étude contrastive. *Linguisticae Investigationes* 42, 186–234.
- Huang, Y. (2000). *Anaphora: a cross-linguistic approach*. Oxford: Oxford University Press.
- Jiang, Y. (2016). Deixis and anaphora. in *A reference grammar of Chinese*. Cambridge, UK: Cambridge University Press.
- Kesik, M. (1989). *La cataphore*. Paris : Presses Universitaires de France.
- Kleiber, G. (1994). *Anaphores et pronoms*. Louvain-la-Neuve, Belgique : Duculot.
- Kleiber, G., Riegel, M. (2003). *Les pronoms évoluent-ils avec les référents ?* Bruxelles : De Boeck Supérieur.
- Landragin, F. (2004). Saillance physique et saillance cognitive. *Corela. Cognition, représentation, langage*, no. 2–2 (December).
- Mitkov, R. (2002). *Anaphora resolution*. New York, Etats-Unis d'Amérique : Longman.
- Muzerelle, J., Schang, E., Antoine, J.-Y., Eshkol, I., Maurel, D., Boyer, A., Nouvel, D. (2012). Annotations en chaînes de coréférences et anaphores dans un corpus de discours spontané en français. *SHS Web of Conferences* 1 : 2497–2516.
- Poon, H., Domingos, P. (2008). Joint unsupervised coreference resolution with markov logic. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 650–659.

- Pu, M.-M. (1997). Zero anaphora and grammatical relations in mandarin. In *Typological Studies in Language*, edited by T. Givón, 35:281. Amsterdam: John Benjamins Publishing Company.
- Recasens, M. (2010). Coreference: theory, annotation, resolution and evaluation. PhD Thesis, Universitat de Barcelona.
- Reichler-Béguelin, M.-J. (1988). Anaphore, cataphore et mémoire discursive. *Pratiques* 57, 15–43.
- Schnedecker, C. (1997). *Nom propre et chaînes de référence*. Paris : Klincksieck.
- Schnedecker, C., Landragin, F. (2014). Les chaînes de référence : présentation. *Langages* 195, 3–22.
- Stirling, L., Huddleston, R. (2002). Deixis and anaphora. In *The Cambridge Grammar of the English Language*. Cambridge : Cambridge University Press.
- Tanaka, I. (2000). The value of an annotated corpus in the investigation of anaphoric pronouns, with particular reference to backwards anaphora in english. PhD Thesis, Lancaster University.
- Tesnière, L. (1959). *Eléments de syntaxe structurale*. Paris : Klincksieck.
- Van Hoek, K. (1997). *Anaphora and conceptual structure*. Etats-Unis d'Amérique: University of Chicago Press.
- 陈 (Chen), 平 (Ping). 1987. 汉语零形回指的话语分析 (Analyse Discursive de l'anaphore Zéro En Chinois). *中国语文 (Zhongguo Yuwen)* 5, 363-378.
- 许 (Xu), 余龙 (Yulong). (2000). 英汉指称词语表达的不可及性 (Referring expressions and their accessibility in Chinese and English). *外语教学与研究 (Foreign Language Teaching and Research)*, no. 05, 321-328.
- 许 (Xu), 余龙 (Yulong) 贺 (He), 小聃 (Xiaodan). (2007). 英汉语下指的篇章功能和语用分析—兼谈汉语第三人称代词照应的单向性问题 (Cataphora in Chinese and English Texts: A Functional-Pragmatic Analysis). *外语教学与研究 (Foreign Language Teaching and Research)* 6, 417-423.

Liste des abréviations utilisées

1SG	première personne du singulier
3sg	troisième personne du singulier
BA	particule qui permet de poser le complément d'objet avant le verbe : <i>bǎ</i> ou <i>jiāng</i>
CLF	classificateur
DE	particule de détermination
DUR	duratif
NEG	négation
PFV	perfectif