

Institut de Psychologie
École doctorale « Neurosciences et Cognition »
Université Lumière Lyon 2

Rôle du relativiseur dans l'attachement des propositions relatives ambiguës en français

Thèse présentée et soutenue publiquement

Le vendredi 17 octobre 2008

par

Claire Delle Luche

Pour obtenir le grade de docteur de l'Université Lumière Lyon 2

(spécialité Sciences Cognitives, mention Linguistique)

Composition du jury

Prof. Harriet Jisa, Laboratoire Dynamique du Langage, Université de Lyon (directeur)

Dr. Frédérique Gayraud, MCF, Laboratoire Dynamique du Langage, Université de Lyon
(responsable scientifique)

Dr. Roger van Gompel, University of Dundee (co-directeur)

Dr. Barbara Hemforth, CR, HDR Laboratoire de Psychologie et Neurosciences, Paris 5
(rapporteur)

Prof. Edward Gibson, Department of Brain and Cognitive Sciences, MIT (rapporteur)

Prof. Michel Charolles, Université Paris 3

Dr. Joël Pynte, DR1, Laboratoire de Psychologie et Neurosciences Cognitives, Paris 5

Laboratoire Dynamique Du Langage, UMR 5596
CNRS et Université Lumière Lyon 2

REMERCIEMENTS

Ce travail a été réalisé grâce à une allocation de recherche du Ministère de l'Enseignement Supérieur et de la Recherche.

Mes premiers remerciements vont à mes directeurs de recherche, Frédérique Gayraud, Harriet Jisa et Roger van Gompel, pour m'avoir encadrée de près ou à distance, mais toujours efficacement. Présents aux moments critiques et ayant toujours une réponse pour mes interrogations de non linguiste et mes maladresses d'apprenti expérimentateur, c'est grâce à vous que ce document existe. Merci pour votre gentillesse, votre humanité et pour avoir transformé ce goût pour la recherche en une volonté d'en faire mon métier.

Je remercie Madame Barbara Hemforth et Messieurs Michel Charolles, Edward Gibson et Joël Pynte pour l'honneur qu'ils me font de faire partie de mon jury de thèse. À lire et apprécier vos travaux, je suis émue que vous ayez accepté de juger le mien.

Je remercie l'ensemble du département de psychologie de l'Université de Dundee et son directeur Monsieur Trevor Harley pour m'avoir accueillie pendant huit mois et procuré un cadre parfait pour poursuivre mes recherches. Je remercie Madame Yuki Kamide et Juhani Jarvikivi pour l'accès au laboratoire d'eye-tracking et pour m'avoir appris cette technique.

Je remercie le Laboratoire Dynamique du Langage et surtout François Pellegrino pour m'avoir accueillie parmi vous et offert un espace et des conditions de travail précieux et stimulants. Merci à Linda pour sa patience administrative, à Egidio et Christian pour leur ingéniosité informatique.

On la tient cette onde de perplexité ! Florence, Bruno, merci d'avoir si souvent trituré vos méninges pour trouver les trucs *ad hoc* pour ces expériences que j'espère non caduques, mais aussi pour que je comprenne un peu la syntaxe.

Sa gentillesse est comme son efficacité, immense. Merci à Christophe C. sans qui je n'aurais jamais pu réaliser cette étude de corpus, mais aussi pour ses autres précieux conseils.

DéDé-èLe est un labyrinthe, et merci à tous les DéDaLes ingénieux pour avoir répondu à mes multiples questions et problèmes scientifiques : Nathalie pour ses critiques gentilles quand on les voudrait brutales et pour avoir toujours le mot qu'il faut, Fanny pour des discussions intéressantes, et Mahé pour ses explications statistiquement significatives.

Je remercie aussi ceux qui m'ont accompagnée dans ma découverte de R, Messieurs Kenneth Knoblauch et Florian Jaeger. J'ai encore beaucoup de choses à apprendre et à tester.

Merci à tous mes voisins de bureau et de thèse, anciens ou futurs Thésées : Caroline pour les fous rires, Christophe d. S. pour toutes les tasses de thé, Claire, Rébecca, Delphine et Liliane pour tous ces moments d'expérimentation, de rédaction et de dégustations gourmandes.

Merci à tous ceux qui occupent les tabourets rouges de la salle de pause. Même si on ne peut pas toujours décider du fumet des autres convives, et encore moins éviter des sujets de conversation parfois animés, qu'est-ce que j'ai apprécié ces pauses à midi !

Je vous avais demandé si je pouvais faire de longues études... Papa, Maman, merci de m'avoir donné la possibilité de les faire. Voilà le résultat, à moi de faire fructifier maintenant. Un grand merci aux autres Gaston de la famille, Céline et Jean-Dominique. J'espère qu'on continuera à se faire voyager mutuellement en allant travailler dans d'autres pays. Merci à mes grands-parents pour leur présence rassurante.

J'ai mis quelques temps à comprendre la fierté qu'on peut avoir à offrir ses productions scientifiques. Parrain, Marraine, j'espère pouvoir à mon tour vous montrer un tiré-à-part, sans fautes de frappe ni de grammaire. Merci d'avoir toujours été là pour moi.

Maintenant je vais devoir travailler sur le chinois pour venir vous rendre visite. Jocelyne et Sébastien, merci d'être, merci pour votre amitié et votre joie de vivre, merci de me soutenir malgré la distance. Maintenant, bercez plutôt Corentin...

Merci à tous mes amis, merci d'être là et de ne pas avoir trop souvent posé la fatidique question « alors cette thèse, ça avance ? » Maintenant, je peux vous dire « voilà c'est fait ».

Il y a les affinités électives ou les liaisons dangereuses dans la littérature, il y a des ambiguïtés d'attachement dans la langue. Il y a des attachements qui se créent dans la vie, tissés par une Ariane invisible. Mais ces attachements ne sont pas ambigus, même s'ils sont tus. À tous ceux à qui je n'ose pas le dire : je vous aime et je suis heureuse de vous connaître.

Lab-yrinthes, DéDaLes, Arianes, je suis Icare et j'espère que la science m'emportera vers de belles découvertes.

RESUME

Parmi les ambiguïtés syntaxiques, celle liée à l'attachement de propositions relatives se distingue par la variation translinguistique et intra-linguistique des préférences d'attachement que les théories syntaxiques peinent à expliquer. Dans *je connais le père du maçon qui est amusant*, le relativiseur peut être attaché à *père* (N1) ou à *maçon* (N2). En français comme dans de nombreuses langues, la préférence va à N1, mais elle diffère pour l'anglais et parfois aussi en fonction des expérimentations à l'intérieur d'une même langue.

Hemforth, Konieczny, et Scheepers (2000) suggèrent que la spécificité de ce type d'ambiguïté réside dans le traitement anaphorique qui a lieu lors de l'attachement du relativiseur en plus du traitement syntaxique.

Selon des modèles du discours (Ariel, 1990, 2001 ; Gundel, Hedberg, & Zacharski, 1993), les expressions anaphoriques sont classées dans des hiérarchies d'accessibilité en fonction de leur forme : plus une expression est courte et peu informative, plus le référent est accessible. Je propose que les relativiseurs peuvent eux aussi être positionnés dans une hiérarchie d'accessibilité, que leur forme sert à signaler l'accessibilité de leur antécédent et qu'elle indique l'attachement à réaliser. Un relativiseur court, comme *qui* en français, signale un antécédent très accessible, N1, et son pendant *lequel* un antécédent peu accessible, N2.

La validité de cette hypothèse est évaluée en comparant les relativiseurs, *qui / lequel* où l'accessibilité signalée diffère et à *qui / auquel* où elle est contrôlée. Une série de questionnaires confirme un attachement différent pour *qui / lequel*, et non pour à *qui / auquel* : l'accessibilité indiquée par le relativiseur guide l'attachement. Une étude de corpus écrit confirme ce résultat et montre qu'en production la forme des antécédents, liée à la modification, l'animation et la fonction des antécédents dans la principale, influence leur niveau d'accessibilité. On met en évidence des attachements comparables pour à *qui* et *auquel* ; en revanche *lequel* est attaché plus souvent à un antécédent moins accessible que celui de *qui*.

Enfin, une expérience qui utilise le paradigme d'auto-présentation segmentée apporte des informations sur les difficultés d'interprétation qui peuvent survenir lors du traitement de telles ambiguïtés. La comparaison de phrases contenant *qui* ou *lequel* révèle un effet de la forme du relativiseur : dans une phrase où l'attachement est temporairement ambigu, on mesure des temps de lecture plus importants quand une relative introduite par *lequel* est ensuite désambiguïcée à N1, ce qui traduit un attachement initial à N2. Aucune préférence d'interprétation n'est mesurée pour *qui*.

Cette recherche montre que la forme du relativiseur guide l'attachement de relatives ambiguës en signalant le degré d'accessibilité de l'antécédent, aussi bien en production qu'en compréhension. La notion d'accessibilité permet d'apporter une explication nouvelle pour les préférences mises en évidence dans diverses expériences.

Mots-clés : ambiguïté syntaxique, propositions relatives, traitement anaphorique, accessibilité, corpus, *self-paced reading*, français.

ABSTRACT

Relative clause attachment ambiguities stand out among syntactic ambiguities because of their between and within language variability. In *je connais le père du maçon qui est amusant*, readers of French and many other languages interpret, e. g. attach, *père* (N1) as the relative clause antecedent, rather than *maçon* (N2). This attachment preference is not verified in English and experiments revealed language differences that theories have difficulties to account for.

Hemforth, Konieczny, and Scheepers (2000) suggest that the specificity of this type of ambiguity relies on the anaphoric binding that takes places in addition to syntactic processing when processing the relativiser.

According to discourse models (Ariel, 1990, 2001; Gundel, Hedberg, & Zacharski, 1993), anaphoric expressions are ranked on an accessibility hierarchy according to their form: the shorter and the less informative the anaphoric expression, the more accessible the referent. I propose that relativisers can also be ranked on this hierarchy and that their form signals their antecedent accessibility, thus cueing attachment. A short and low informative relativiser, like *qui* in French, signals a highly accessible antecedent, N1, while its equivalent *lequel* cues a low accessibility antecedent, N2.

The validity of this hypothesis is evaluated by comparing *qui* and *lequel* where accessibility differs, and *à qui* and *auquel* where accessibility is controlled. A series of questionnaires confirms a different attachment for *qui* and *lequel*, but not for *à qui* and *auquel*: the accessibility level signalled by the relativiser constrains the attachment. A written corpus study confirms this result and shows that the form of the antecedents, linked to their modification, animacy, and syntactic function in the main clause influences their accessibility in production. Similar attachments are evidenced for *à qui* and *auquel* but *lequel* is attached preferably to the less accessible antecedent while *qui* is attached to the more accessible.

Finally, a self-paced reading experiment investigates processing difficulties during the reading of ambiguous relative clauses with *qui* and *lequel*, and the effect of the form of the relativiser is evidenced again. In a sentence where attachment is temporarily ambiguous reading times are longer when a relative clause introduced by *lequel* is later disambiguated to N1, suggesting an initial N2 attachment. No preference is measured for *qui*.

This research shows that the form of the relativiser guides ambiguous relative clause attachment while signalling the accessibility level of the antecedent, in production as well as in comprehension. Accessibility provides a new explanation for the attachment preferences evidenced in various experiments.

Key words: syntactic ambiguity, relative clause, anaphoric processing, accessibility, corpus, self-paced reading, French.

TABLE DES MATIERES

REMERCIEMENTS	i
RESUME.....	iii
ABSTRACT	iv
INTRODUCTION GENERALE.....	7
INTRODUCTION DE LA PARTIE THEORIQUE	11
Chapitre 1 Les ambiguïtés syntaxiques	17
1. La théorie du <i>garden path</i>	17
1. 1. Le système cognitif est modulaire.....	18
1. 2. La stratégie du <i>minimal attachment</i>	18
1. 3. La stratégie du <i>late closure</i>	20
2. Un cas particulier d'ambiguïté syntaxique : l'attachement des propositions relatives.	23
2. 1. Les sources de variation	26
2. 1. 1. Les antécédents	26
2. 1. 1. 1. L'effet de la fréquence des antécédents	27
2. 1. 1. 2. La position de l'adjectif	29
2. 1. 1. 3. L'animation	30
2. 1. 2. Les relations entre les antécédents.....	33
2. 1. 3. L'effet des caractéristiques de la proposition relative	35
2. 1. 4. Les effets de la méthodologie expérimentale.....	37

2.2. Comment les modèles expliquent les variations pour les préférences d'attachement.....	39
2.2.1. Les variations dans une langue	39
2.2.1.1. L'hypothèse de De Vincenzi et Job (1993, 1995).....	40
2.2.1.2. Le principe de modifiabilité.....	40
2.2.2. Les variations translinguistiques.....	41
2.2.2.1. Le modèle de satisfaction de contraintes	41
2.2.2.2. La <i>construal hypothesis</i>	43
2.2.2.3. La <i>tuning hypothesis</i>	45
2.2.2.4. La <i>recency theory</i>	48
3. L'attachement des relatives comme résolution anaphorique (Hemforth <i>et al.</i> , 2000)..	50
Chapitre 2 Le traitement des anaphores	55
1. Givón.....	58
1.1. L'accessibilité référentielle	58
1.2. La continuité référentielle et la hiérarchie des expressions anaphoriques	59
1.3. Mise en évidence expérimentale	60
2. Les théories de l'accessibilité	61
2.1. La <i>givenness hierarchy</i>	61
2.2. L' <i>accessibility theory</i>	67
2.3. L' <i>informational load hypothesis</i>	72
2.4. Les théories de l'accessibilité et le traitement des propositions relatives	72
2.5. La <i>centering theory</i>	74
2.5.1. Les principes de la théorie	75
2.5.1.1. Les centres dans le discours	75
2.5.1.2. La première règle	78
2.5.1.3. La seconde règle.....	82
2.5.2. Les développements récents	84
2.5.2.1. Intégrer les propriétés des pronoms	84
2.5.2.2. Le parallélisme des fonctions.....	85
2.5.2.3. La théorie de la proéminence du discours.....	85
Le premier principe	86
Le deuxième principe	86
Le troisième principe.....	87
2.5.3. La <i>centering theory</i> et le traitement des propositions relatives	87
Chapitre 3 Discussion des théories.....	93
1. La saillance	93
1.1. La prosodie et l'accentuation	93
1.2. La définitude.....	95
1.3. La fréquence	96
1.4. L'ordre des mots.....	97

1. 5. Le contexte référentiel.....	99
2. Le relativiseur	101
Chapitre 4 Problématique et hypothèses	103
1. Synthèse critique des éléments présentés dans la partie théorique.....	103
2. Problématique et hypothèse théorique	109
3. L'opérationnalisation des hypothèses	111
3. 1. Objet étudié et restrictions méthodologiques	112
3. 1. 1. Restriction de l'étude à une seule langue.....	112
3. 1. 2. Le choix du français.....	112
3. 1. 3. Le type de structure étudiée	113
3. 1. 4. Les sources de variation étudiées.....	114
3. 1. 4. 1. Le relativiseur.....	114
3. 1. 4. 2. Les antécédents	115
3. 2. Les hypothèses opérationnelles	117
3. 3. Mise à l'épreuve expérimentale	120
PARTIE EXPERIMENTALE.....	122
Chapitre 1 Questionnaires	124
1. Questionnaire 1 : comparaison de <i>qui</i> et <i>lequel/laquelle</i> , interindividuel	126
1. 1. Méthodologie expérimentale	127
1. 1. 1. Participants.....	127
1. 1. 2. Matériel	127
1. 1. 3. Pré-test	128
1. 1. 4. Procédure expérimentale.....	129
1. 2. Résultats	130
1. 3. Discussion	131
2. Questionnaire 2 : comparaison de <i>auquel</i> et <i>à qui</i> , interindividuel	133
2. 1. Méthodologie expérimentale.....	134
2. 1. 1. Participants.....	134
2. 1. 2. Matériel	134
2. 1. 3. Pré-test	134
2. 1. 4. Procédure expérimentale.....	135
2. 2. Résultats	136
2. 3. Discussion	137
3. Questionnaire 3 : comparaison de <i>auquel</i> et <i>à qui</i> , intra-individuel	138
3. 1. Méthodologie expérimentale.....	139
3. 1. 1. Participants.....	139
3. 1. 2. Matériel	140
3. 1. 3. Pré-test	140

3. 1. 4. Procédure expérimentale.....	142
3. 2. Résultats	143
3. 3. Discussion	144
4. Questionnaire 4 : comparaison de <i>lequel</i> et <i>qui</i> , intra-individuel	145
4. 1. Méthodologie expérimentale.....	145
4. 1. 1. Participants.....	145
4. 1. 2. Matériel	146
4. 1. 3. Procédure expérimentale.....	146
4. 2. Résultats	147
4. 3. Discussion	147
5. Discussion générale sur les Questionnaires	148
5. 1. Aspects méthodologiques.....	149
5. 1. 1. La plausibilité du contenu sémantique.....	149
5. 1. 2. La fréquence des antécédents.....	150
5. 2. Implications théoriques	151
5. 2. 1. La question de la fréquence des relativiseurs	151
5. 2. 2. Les Hypothèses 1a et 1b	152
5. 2. 3. Les théories de l'accessibilité	153
5. 2. 4. La théorie du <i>garden path</i> et la <i>construal hypothesis</i>	154
5. 2. 5. Les modèles interactifs	154
5. 3. Conclusion.....	155
Chapitre 2 Étude de corpus	156
1. Introduction.....	157
1. 1. La structure « N1 de N2 »	158
1. 2. La restrictivité.....	158
2. Les hypothèses	159
2. 1. L'effet de la forme du relativiseur (Hypothèse 1).....	159
2. 2. Les propriétés des antécédents	160
2. 2. 1. La longueur de la proposition relative (Hypothèse 2a).....	160
2. 2. 2. La présence d'un modifieur (Hypothèse 2b)	161
2. 2. 3. L'animation (Hypothèse 2c)	162
2. 2. 4. La fonction syntaxique (Hypothèse 2e)	163
2. 3. Les interactions entre facteurs.....	165
3. Le corpus.....	165
3. 1. Corpus original et prétraitements	165
3. 2. Codage du corpus	167
3. 3. Taille du corpus	168
4. Résultats.....	168
4. 1. Ensemble des propositions relatives non restrictives.....	170
4. 2. Propositions relatives sujet.....	178

4.3.	Propositions relatives OI	183
4.4.	Propositions relatives OI – animé	188
5.	Discussion	190
5.1.	Validation des hypothèses opérationnelles.....	191
5.1.1.	Effet de la forme du relativiseur (Hypothèse 1).....	191
5.1.2.	Effet de la longueur de la proposition relative (Hypothèse 2a)	193
5.1.3.	Effet de la modification (Hypothèse 2b).....	194
5.1.4.	Effet de l’animation (Hypothèse 2c).....	195
5.1.5.	Effet de la fonction syntaxique (Hypothèse 2d)	197
5.2.	Apport pour les théories	198
5.3.	Conclusions	202
<hr/>		
Chapitre 3	Expérience avec le paradigme d’auto-présentation segmentée	204
<hr/>		
1.	Introduction.....	204
2.	Méthodologie expérimentale	211
2.1.	Participants	211
2.2.	Matériel et procédure expérimentale	212
2.2.1.	Matériel	212
2.2.1.1.	Élaboration des items expérimentaux	212
2.2.1.2.	La ponctuation.....	214
2.2.1.3.	Pré-test.....	215
2.2.1.4.	Plan expérimental.....	216
2.2.2.	Procédure expérimentale.....	217
3.	Résultats.....	219
3.1.	Les questions de compréhension.....	220
3.1.1.	Analyses des erreurs aux questions portant sur la relative	221
3.1.2.	Analyses des erreurs aux questions portant sur la principale	221
3.2.	Les temps de lecture	221
3.2.1.	Le premier nom.....	222
3.2.2.	Le « de ».....	223
3.2.3.	Le deuxième nom.....	223
3.2.4.	Le relativiseur	223
3.2.5.	Le verbe de la relative.....	223
3.2.6.	La désambiguïsation	224
3.2.7.	La région <i>spillover</i>	224
3.2.8.	Le verbe de la principale.....	225
3.2.9.	La fin de la phrase.....	225
	Post-test pour évaluer l’ordre des noms	225
4.	Discussion.....	227

Chapitre 4	Discussion générale	233
1. 1.	Récapitulatif des résultats obtenus	233
1. 2.	La validation des hypothèses opérationnelles et de l'hypothèse théorique.....	234
1. 2. 1.	La synonymie des relativiseurs en français	237
1. 3.	Apports pour les théories.....	238
1. 3. 1.	L'importance de la notion d'accessibilité dans le traitement des ambiguïtés d'attachement.....	238
1. 3. 2.	L'accessibilité comme facteur explicatif des variations translinguistiques et intra-linguistiques	241
1. 3. 3.	La fréquence.....	242
1. 3. 4.	L'accessibilité comme facteur explicatif des variations translinguistiques et intra-linguistiques	243
1. 3. 5.	L'intégration des informations syntaxiques, sémantiques ou discursives .	244
1. 3. 6.	Les modèles de la compréhension, de la production	246
1. 4.	La méthodologie expérimentale	248

CONCLUSIONS.....	251
------------------	-----

ANNEXES	252
---------------	-----

Annexe 1 – Matériel expérimental utilisé pour le Questionnaire 1 (<i>qui, lequel/laquelle, interindividuel</i>)	253
---	-----

Annexe 2 – Matériel expérimental utilisé pour le Questionnaire 2 (à <i>qui, auquel, interindividuel</i>)	254
---	-----

Annexe 3 – Matériel expérimental utilisé pour le Questionnaire 3 (à <i>qui, auquel/à laquelle, intra-individuel</i>).....	255
--	-----

Annexe 4 – Matériel expérimental utilisé pour le Questionnaire 4 (<i>qui, lequel/laquelle, intra-individuel</i>)	257
--	-----

Annexe 5 – Matériel expérimental utilisé pour l'expérience avec le paradigme d'auto-présentation segmentée.....	258
---	-----

INTRODUCTION GENERALE

Comprendre une phrase consiste à traduire un signal sonore (ou visuel) en une représentation mentale. Les phrases sont plus ou moins complexes et, parmi les sources de difficulté de traitement, les ambiguïtés sémantiques ou syntaxiques ont fait l'objet d'une littérature importante (Fuchs, 1996 ; Tanenhaus & Trueswell, 1995).

L'exemple (1) illustre un cas d'ambiguïté totale à l'écrit : le sens et les rôles syntaxiques des différents constituants sont ambigus ; aucun indice ici ne permet ni de lever l'ambiguïté ni de ne conserver qu'une seule interprétation.

(1) *Le boucher sale la tranche.*

Si *le boucher* est de toute évidence le sujet de la proposition, *sale* peut être interprété de deux façons. Si *sale* est analysé comme adjectif, le clitique *la* est l'objet du verbe *tranche*. Si *sale* est identifié comme le verbe fléchi, *la tranche* est par conséquent l'objet de ce verbe. À l'écrit, l'ambiguïté est totale puisque aucune information sémantique ou syntaxique ne permet de choisir entre les deux interprétations possibles. À l'oral, par contre, la prosodie peut lever cette ambiguïté (Pynte & Prieur, 1996 ; Schafer, Speer, Warren, & White, 2000).

L'ambiguïté est parfois temporaire. Au début de la phrase, deux interprétations sont possibles, mais un constituant ultérieur ou des informations sémantiques signalent qu'une seule interprétation est correcte. Une telle conception suppose que l'intégration des constituants de la phrase est incrémentale, un constituant après l'autre.

Une phrase comme (2) est temporairement ambiguë. *Marie* est le sujet et *repassait* le verbe de la proposition subordonnée introduite par *pendant que*. En français, l'ordre des mots canonique est sujet, verbe puis objet. *Sa jupe* est alors intégré comme l'objet de *repassait*, mais *est tombée* rend cette interprétation impossible. Il s'avère alors nécessaire de ré-analyser la phrase en intégrant *sa jupe* non pas comme l'objet de *repassait*, mais comme le sujet du verbe de la principale *est tombée*.

(2) *Pendant que Marie repassait sa jupe est tombée.*

Il existe un type d'ambiguïté lié à la relativisation, lorsque la proposition relative peut être potentiellement attachée à deux antécédents, comme dans (3). La relative (*qui a eu un accident*) a pour antécédent soit *fille*, soit *colonel*.

(3) *Le journaliste interroge la fille du colonel qui a eu un accident.* (adapté de Cuetos & Mitchell, 1988)

Des études en psycholinguistique ont montré l'existence de préférences d'attachement dans diverses langues et des théories ont été proposées pour expliquer les traitements cognitifs guidant l'interprétation, la compréhension. Ces théories ont longtemps été confrontées au problème de la variation translinguistique mesurée expérimentalement. En effet, pour des langues comme le français ou l'espagnol, dans une phrase comme (3), les locuteurs préfèrent considérer que l'antécédent de la proposition relative est le premier nom, *la fille*. Des expériences en anglais ont pu mettre en évidence que les locuteurs préféreraient attacher la proposition relative au second nom, *colonel*. L'explication la plus souvent retenue est l'application de différentes stratégies ou l'effet de certaines variables qui diffèrent selon les langues. Cependant, ces théories n'expliquent pas les différences qui peuvent être mesurées pour une même langue.

L'objectif de ce travail de thèse est de proposer une explication nouvelle des préférences d'attachement pour ce type de structure. Je propose que la forme des constituants tels que les antécédents et le relativiseur peut guider l'interprétation. La forme du relativiseur signale le degré d'accessibilité de son antécédent, elle-même signalée par la forme de ce dernier. Mais avant de présenter plus en détail les hypothèses formulées sur les facteurs qui pèsent sur la forme du relativiseur et qui ont guidé ce travail, je vais d'abord décrire le cadre théorique d'où émerge la problématique qui sous-tend ma recherche. Après une brève introduction sur ce qu'est la relativisation, je dresse une revue de la littérature sur le traitement des ambiguïtés syntaxiques, en focalisant sur celles qui expliquent les mécanismes mis en jeu lors de l'interprétation d'une phrase contenant une relative dont l'attachement est ambigu, cf. (3).

Je fais ensuite un inventaire des études sur les anaphores qui peuvent proposer un regard nouveau sur le traitement des ambiguïtés d'attachement des propositions relatives qui motive cette recherche.

Les hypothèses de travail qui seront présentées par la suite sont évaluées à l'aide de méthodes expérimentales classiques en psycholinguistique. Tout d'abord, une étude de questionnaire évalue les préférences d'attachement en français dans une tâche de compréhension. Elle est suivie d'une étude de corpus qui fournit une perspective axée sur la production langagière. Enfin, une étude utilisant le paradigme d'auto-présentation segmentée permet d'obtenir des données sur le traitement cognitif lors de la compréhension de phrases contenant des relatives dont l'attachement est ambigu.

INTRODUCTION DE LA PARTIE THEORIQUE

Ce travail porte sur le traitement d'ambiguïtés d'attachement, où une proposition relative est potentiellement attachée à plus d'un antécédent, (3). L'ensemble des théories s'accorde sur l'existence de préférences d'interprétation telles que, dans une structure « N1 de N2 », c'est le premier nom qui est considéré comme l'antécédent du relativiseur.

(3) *Le journaliste interroge la fille du colonel qui a eu un accident.* (adapté de Cuetos & Mitchell, 1988)

Avant de présenter ces théories et les expériences qui ont évalué les préférences d'attachement de façon translinguistique et en manipulant différents facteurs susceptibles d'affecter le traitement de ce type de phrase, je vais d'abord décrire, brièvement la relativisation standard en français. On s'intéressera ensuite aux théories du traitement des ambiguïtés syntaxiques, notamment à celle du *garden path*, théorie en quelque sorte historique et critiquée par d'autres courants théoriques qui ne pourront être vus de façon très détaillée ici. En effet, ces modèles se sont principalement attachés à expliquer le traitement d'autres ambiguïtés syntaxiques que celle qui fait l'objet de cette thèse. Je reviendrai sur ces théories au moment opportun.

Ces théories, souvent confrontées à la validation translinguistique des hypothèses formulées, ont pour but non seulement d'expliquer le traitement de structures ambiguës, mais aussi d'expliquer pourquoi les préférences d'interprétation varient à travers les langues. On verra que le cas des ambiguïtés d'attachement des propositions relatives constitue un cas particulier. Cependant, je montrerai qu'à une variation translinguistique s'ajoute une variation intra-linguistique, pour laquelle aucune explication très convaincante n'a pu être apportée et qui est probablement due à la variabilité du matériel utilisé dans les expériences.

Parmi ces théories, je m'attarderai sur l'hypothèse proposée par Hemforth Konieczny, et Scheepers (2000) qui semble offrir un élément de réponse intéressant pour la singularité de l'attachement ambigu des propositions relatives parmi les autres ambiguïtés syntaxiques : l'interprétation du relativiseur et l'attachement de la proposition relative consiste en un traitement syntaxique, où deux structures syntaxiques sont possibles, l'une avec un attachement au premier nom, l'autre au second. Les autres ambiguïtés syntaxiques sont de ce point de vue semblables. Par contre, il s'ajoute un traitement anaphorique du relativiseur qui fait référence à une entité du discours, son antécédent : un tel traitement est absent dans les autres ambiguïtés et l'attachement d'une proposition relative ambigu est guidé par ce traitement anaphorique, ce qui expliquerait pourquoi les principes généraux proposés pour le traitement des ambiguïtés syntaxiques (cf. théorie du *garden path*) peinent à expliquer les préférences d'attachement de relatives ambiguës.

Si l'intégration du relativiseur est guidée par ses propriétés en tant qu'anaphore, une revue des théories sur les anaphores est nécessaire. Elle montre que les différentes entités du discours peuvent être plus ou moins accessibles et que ce niveau d'accessibilité est signalé par la forme des expressions référentielles qui renvoient à ces entités du discours. Par exemple, de nombreuses études montrent qu'un pronom sert à faire référence à une entité du discours très accessible : il s'agit souvent du sujet de la proposition précédente. Comme on le verra avec l'étude d'Ariel (1999), un relativiseur peut être considéré comme une expression anaphorique, subissant les mêmes contraintes que les autres expressions référentielles : sa forme marque l'accessibilité de l'entité à laquelle il fait référence, elle signale le niveau d'accessibilité de son antécédent.

Je proposerai ensuite une approche du traitement des ambiguïtés d'attachement qui prend en compte que les relativiseurs et les antécédents sont des expressions référentielles qui signalent le niveau d'accessibilité de l'entité à laquelle ils font référence. Ainsi, la forme des

antécédents et des relativiseurs guide l'attachement de la proposition relative, ce qui peut expliquer les préférences d'attachement mesurées expérimentalement.

Mais revenons d'abord sur la relativisation¹. La stratégie standard de la relativisation en français, qui est la langue évaluée dans la partie expérimentale de ce travail, est celle qui est caractérisée par la présence d'un pronom relatif et d'un trou syntaxique (Creissels, 2006 ; Riegel, Pellat, & Rioul, 2004).

Dans l'exemple (4), la phrase est composée de deux propositions, la principale (*le reporter admet l'erreur*) et la relative (*qui attaque le sénateur*), qui est enchâssée dans la principale. Le relativiseur, *qui*, est co-référent avec son antécédent, *le reporter* et marque plusieurs aspects : c'est un modifieur du nom, son antécédent, et il marque à la fois le début de la relative et la fonction du relativisé dans la relative. *Qui* marque la fonction sujet. Le trou syntaxique, symbolisé par e_i , se trouve là où serait placé le sujet du verbe selon l'ordre canonique dans une proposition indépendante.

(4) *Le reporter qui_i e_i attaque le sénateur admet l'erreur.* (d'après King & Just, 1991)

Le français dispose de deux systèmes de relativiseurs, les *qui* et les *lequel*, dont on peut voir la distribution dans le Tableau 1. On remarque tout d'abord qu'à part la fonction sujet le système des *qui* marque l'animation de l'antécédent, ce qui les distingue des *lequel*.

L'utilisation des *lequel* est jugée plus formelle (Jones, 1996 ; Riegel *et al.*, 2004), en particulier lorsque l'antécédent est animé et que *lequel* introduit une relative sujet. Contrairement aux *qui*, ils portent le genre et le nombre de l'antécédent : ainsi *lequel* signale un antécédent masculin singulier et *laquelle* un féminin singulier.

¹ La littérature sur la compréhension des propositions relatives est très importante. Pour plus de détails, on peut se référer aux études les plus connues, comme celles de Caplan, Alpert, & Waters (1999), Gibson (1998), King & Just (1991), ou aussi celle de Sheldon (1974).

Tableau 1 : distribution des relativiseurs en français (adapté de Riegel et al., 2004)

Fonction syntaxique	Les <i>qui</i>		Les <i>lequel</i>
	Animé	Non-animé	Animé ou non-animé
Sujet	<i>qui</i>	<i>qui</i>	<i>lequel, laquelle, etc.</i>
Objet	<i>que</i>	<i>que</i>	Ø
Autres	<i>à qui</i>	<i>à quoi</i>	<i>auquel, à laquelle, etc.</i>
	prép. + <i>qui</i> <i>dont</i>	prép. + <i>quoi</i> <i>dont</i>	prép. + <i>lequel, etc.</i> <i>duquel, etc.</i>

À l'exception de la fonction objet, *lequel* ayant disparu pour cette fonction dans le français actuel (Damourette & Pichon, 1969), les deux systèmes de relativiseurs peuvent être utilisés indifféremment. Ils sont selon la norme synonymes, mis à part le codage de l'animation et des genre et nombre.

Les expériences évaluant le traitement de propositions relatives, que leur attachement soit ambigu (Baccino, De Vincenzi, & Job, 2000 ; Colonna, 2001 ; Frenck-Mestre & Pynte, 2000b ; Zagar, Pynte, & Rativeau, 1997) ou non (Holmes & O'Regan, 1981 ; Schelstraete & Degand, 1998), utilisent toujours le système des *qui*, le plus fréquent, s'intéressant d'ailleurs surtout aux fonctions sujet et objet. L'existence des deux systèmes de relativiseurs sera exploitée.

Les relativiseurs signalent aussi un autre aspect de la proposition relative, sa restrictivité (Auroux & Rosier, 1987 ; Comrie, 1981 ; Jones, 1996 ; Riegel et al., 2004 ; Tellier, 1996). Les relatives restrictives (5a) sont parfois appelées déterminatives et servent à restreindre la référence de l'antécédent (voir p. ex. Riegel et al., 2004) et à identifier une entité parmi un ensemble. Elles s'opposent aux relatives non restrictives, appelées aussi explicatives ou appositives (5b).

- (5) a. *Les enfants qui étaient en retard furent punis.*
 b. *Les enfants, qui étaient en retard, furent punis.* (d'après Galliot & Laubreaux, 1966)

Supprimer la proposition relative dans (6a), ce qui donne (6b), change le sens de la principale et la référence n'est plus à une entité spécifique mais à un ensemble général. Il s'agit d'une relative restrictive.

- (6) a. *Le roman que je viens de finir me plaît beaucoup.*
b. *Le roman me plaît beaucoup.* (d'après Riegel *et al.*, 2004)

L'antécédent réfère de façon générale à une entité qui, dans le discours, possède le statut de donnée² (*given*, voir Clark, 1975 ; Haviland & Clark, 1974), à la fois pour le locuteur³ et l'allocutaire (Givón, 1990 ; Kleiber, 1987). Aussi considère-t-on souvent qu'une relative restrictive ne peut avoir comme antécédent qu'un groupe nominal (désormais abrégé GN) avec un article défini et ne peut modifier un nom propre (Holmes, 1988 ; Jones, 1996 ; Tellier, 1996). Givón (1982) relève néanmoins de telles constructions dans les langues. On peut déjà remarquer que si l'on veut manipuler la forme de l'antécédent (p. ex. GN avec article indéfini), cela n'est guère possible avec les propositions relatives restrictives.

À ce type de propositions relatives sont opposées les relatives non restrictives, appelées aussi appositives ou explicatives.

Ces relatives fournissent des informations supplémentaires à celles déjà introduites dans la proposition principale. Elles sont en général attachées à une entité qui n'est pas nécessairement donnée, mais qui est au moins identifiable de façon unique (Jones, 1996), comme un nom propre ou un GN avec un possessif. Les propositions relatives non restrictives peuvent être supprimées sans altérer le sens de la proposition (Jones, 1996 ; Riegel *et al.*, 2004, mais voir les restrictions de Loock, 2007).

Contrairement à la proposition relative restrictive, une relative non restrictive supporte le plus souvent le test de la suppression : que ce soit en (7a) ou en (7b), le référent *ce roman* reste le même.

- (7) a. *Ce roman, que je viens de finir, me plaît beaucoup.*
b. *Ce roman me plaît beaucoup.* (d'après Riegel *et al.*, 2004)

Comme il est possible de faire référence à une entité qui n'est pas donnée, les formes de l'antécédent peuvent être plus variées, comme un nom propre, un GN avec un article indéfini (8a) ou démonstratif (7a).

- (8) a. *L'enfance est une période de la vie qui est déterminante.*
b. * *L'enfance est une période de la vie.* (*ibidem*)

² On reviendra sur la distinction *given-new*, donné-nouveau, dans le Chapitre 2 traitant des anaphores.

³ Dans le cas de l'écrit, on parlera de scripteur, mais j'utilise ici locuteur pour les deux modalités, écrite et orale.

L'exemple (8a-b) illustre que le test de la suppression n'est pas un critère toujours fiable pour identifier la non restrictivité. Un jugement pragmatique, lui, confirme qu'il s'agit bien d'une proposition relative non restrictive.

Le contenu d'une proposition relative non restrictive est le plus souvent nouveau dans le discours (Gibson, Desmet, Grodner, Watson, & Ko, 2005 ; Loock, 2007), alors que le contenu d'une proposition relative restrictive peut être considéré comme déjà connu ou donné, c'est-à-dire qu'il n'introduit pas un nouveau référent. Gibson *et al.* (2005) suggèrent que cet aspect peut expliquer les difficultés des participants dans une expérience où ils doivent lire des phrases avec des relatives appositives contenant des éléments qui font déjà partie du contexte discursif (Gibson *et al.*, 2005). Comme ils s'attendent à des éléments nouveaux, cela entraîne des difficultés de traitement plus importantes.

Enfin, notons la proposition selon laquelle l'identification de l'antécédent est réalisée à l'aide d'anaphores discursives (Sells, 1985, cité par Grodner, Gibson, & Watson, 2005), comme lors de l'identification de l'antécédent pour un pronom (D. Arnold, 2004). Cela suggère ainsi que traiter une proposition relative ne consiste pas uniquement en un traitement syntaxique, mais aussi en un traitement discursif.

D'après la norme, en anglais ou en français, les propositions relatives non restrictives sont séparées de leur proposition principale par une virgule à l'écrit (Riegel *et al.*, 2004 ; Tellier, 1996) et par une rupture prosodique à l'oral (D. Arnold, 2004 ; Holmes, 1988 ; Tao & McCarthy, 2001), bien que cette dernière ne soit pas toujours facile à analyser. Néanmoins, la ponctuation n'est pas suffisante pour déterminer le degré de restrictivité, puisque des relatives non restrictives peuvent ne pas être précédées d'une virgule (8a).

En résumé, intégrer, traiter le relativiseur et la proposition relative qu'il introduit consiste à attacher ce relativiseur à son antécédent. On a vu que cet attachement est parfois ambigu, car si plusieurs sites sont disponibles pour l'attachement.

De nombreuses théories ont tenté d'expliquer les facteurs qui régissent les préférences d'attachement, tout en s'interrogeant sur le décours temporel des différentes analyses possibles. La théorie du *garden path* apparaît comme l'une des approches majeures parmi les théories qui se sont attachées à expliquer ces préférences d'attachement (pour une revue, voir Mitchell, 1994 ; pour une revue, voir Schelstraete, 1993 ; Tanenhaus & Trueswell, 1995 ; van Gompel, Pickering, & Traxler, 2001).

Chapitre 1 Les ambiguïtés syntaxiques

Parmi les structures ambiguës qui ont pu être étudiées, c'est aux ambiguïtés d'attachement des relatives que je m'intéresse. Les ambiguïtés syntaxiques peuvent être totales ou temporaires. Celle dans (3) est totale : rien n'indique si la relative doit être attachée à *fille* ou *colonel*.

(3) *Le journaliste interroge la fille du colonel qui a eu un accident.* (adapté de Cuetos & Mitchell, 1988)

Que l'ambiguïté soit levée ou non, des expériences en compréhension, qui vont être présentées, ont mis en évidence l'existence de préférences d'attachement : si plusieurs interprétations sont possibles, le parseur en choisit une, qu'il ait ou non conscience de l'ambiguïté (il n'en n'est pas conscient, selon Frazier & Rayner, 1982). La théorie du *garden path* est une tentative d'expliquer les mécanismes qui sous-tendent ces préférences d'interprétation.

1. La théorie du *garden path*

La théorie du *garden path* (Frazier, 1985, 1987 ; Frazier & Fodor, 1978 ; Frazier & Rayner, 1982, 1987 ; Rayner, Carlson, & Frazier, 1983) a pour but d'expliquer le traitement du langage en s'appuyant sur le traitement des ambiguïtés syntaxiques et les effets *garden path*. On appelle *garden path* le fait que, confronté à une ambiguïté, le parseur fait fausse route (*garden path*) en choisissant systématiquement la même interprétation, la plus simple.

Si les constituants ultérieurs de la phrase viennent à lever l'ambiguïté et signaler le cas échéant une interprétation différente de celle choisie initialement, un processus de ré-analyse est alors mis en place (Frazier & Rayner, 1982 ; Meseguer, Carreiras, & Clifton, 2002 ; Rayner *et al.*, 1983), ce qui conduit à des temps de traitement plus importants lors de tâches de lecture.

Cette théorie repose sur trois principes. Les deux premiers sont la modularité du système cognitif et la stratégie du *minimal attachment* et sont présentés très brièvement car ils concernent moins ce travail de thèse que le troisième principe, la stratégie de *late closure*, qui explique l'attachement de propositions relatives ambiguës.

1. 1. Le système cognitif est modulaire

La théorie du *garden path* considère que le système cognitif fonctionne de façon modulaire, c'est-à-dire avec des sous-systèmes indépendants et un traitement sériel de l'information (J. A. Fodor, 1986). D'autres modèles proposent, au contraire, un traitement en parallèle de l'information, notamment le langage (Marslen-Wilson, 1987 ; Marslen-Wilson & Tyler, 1980). Ici, les modules responsables du traitement syntaxique, sémantique ou discursif sont encapsulés (Frazier, 1990a) et indépendants (Jackendoff, 1999).

L'activation de la stratégie syntaxique est automatique et réalisée en premier (F. Ferreira & Clifton, 1986 ; Frazier, 1987 ; Rayner *et al.*, 1983), avant que les stratégies basées sur la sémantiques en soient considérées. Ces dernières ne sont utilisées qu'en cas de ré-analyse.

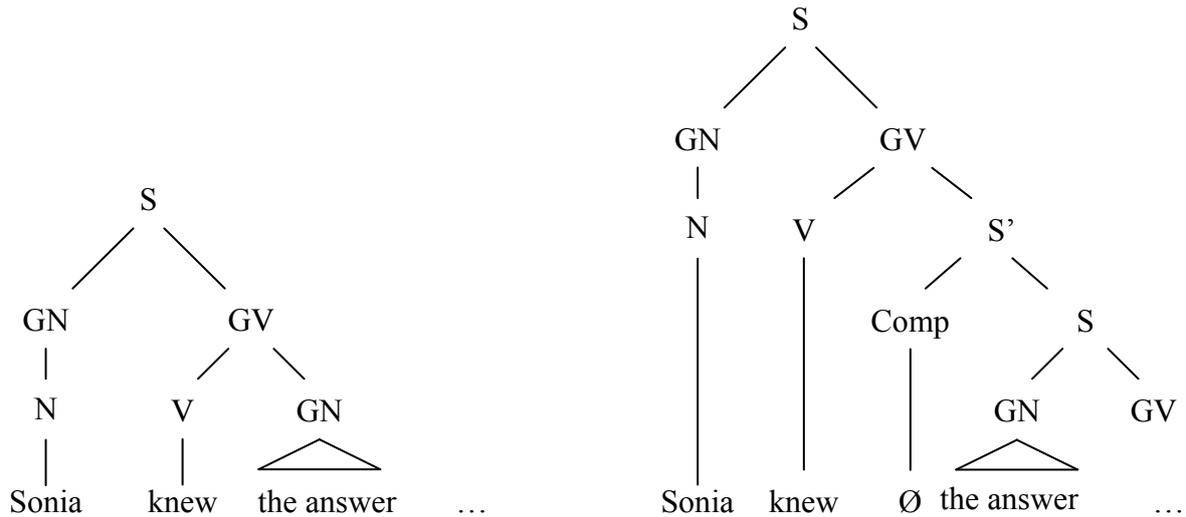
Cette conception est soutenue par une littérature importante et des données en imagerie cérébrale et en électrophysiologie (Friederici, 2002 ; Hagoort, Brown, & Groothusen, 1993 ; Neville, Mills, & Lawson, 1992 ; Osterhout & Holcomb, 1992 ; Osterhout, McLaughlin, & Bersick, 1997 ; Osterhout & Nicol, 1999b). On a pu montrer l'existence de patrons électrophysiologiques spécifiques au traitement syntaxique (Coulson, King, & Kutas, 1998 ; Friederici, Pfeifer, & Hahne, 1993) et au traitement sémantique (Hahne & Friederici, 2002 ; Kutas & Hillyard, 1984 ; Kutas & Kluender, 1994 ; Kutas & Van Petten, 1988). Des régions cérébrales spécifiques seraient aussi impliquées de façon indépendante selon l'analyse (Dapretto & Bookheimer, 1999 ; Vigliocco, 2000), syntaxique (Caplan, Alpert, & Waters, 1998 ; Friederici, Hahne, & Saddy, 2002 ; Just, Carpenter, Keller, Eddy, & Thulborn, 1996). Le mécanisme de ré-analyse qui sous-tend une interaction entre ces deux types d'informations est plus tardif, qui serait confirmé lui aussi par des données électrophysiologiques (Gunter, Friederici, & Schriefers, 2000 ; Hagoort & Brown, 2000), mais la signification du marqueur électrophysiologique invoqué reste sujette à débat (Kaan, Harris, Gibson, & Holcomb, 2000 ; Kaan & Swaab, 2003 ; Kaan, Wijnen, & Swaab, 2004 ; Palolahti, Leino, Jokela, Kopra, & Paavilainen, 2005).

1. 2. La stratégie du *minimal attachment*

La stratégie du *minimal attachment* (attachement minimal) permet de choisir entre deux interprétations qui diffèrent quant à la complexité de leurs structures syntaxiques. Une phrase comme (9) est temporairement ambiguë car, lorsqu'on intègre *the answer* à la structure syntaxique en cours de construction, ce peut être en tant qu'objet du verbe de la principale ou en tant que sujet d'une complétive.

(9) *Sonia knew the answer to the difficult problem was correct.* (Sonia savait que la réponse au problème complexe était correcte, tiré de Frazier, 1985)⁴

Ces deux analyses peuvent être représentées par deux arbres syntaxiques différents (Figure 1a-b). Interpréter *the answer* comme objet du verbe *knew* (Figure 1a) permet la construction d'un arbre syntaxique plus simple que de l'interpréter comme sujet d'une complétive (Figure 1b).



a. *Interprétation en tant qu'objet du verbe knew*

b. *Interprétation en tant que sujet d'une complétive*

Figure 1 : Représentation arborescente du début d'une phrase garden path⁵

Le principe du *minimal attachment*⁶ (Frazier, 1985 ; Frazier & Fodor, 1978) veut que « chaque item lexical (ou tout autre nœud) [soit] attaché au marqueur phrastique avec le moins de nœuds non terminaux possible et en le liant aux nœuds déjà présents » (p. 320 Frazier & Fodor, 1978). Dans (9), c'est la structure syntaxique la plus simple (Figure 1a), qui est préférée à la structure syntaxique plus complexe (Figure 1b), parce qu'elle contient le moins de nœuds syntaxiques.

⁴ Nota : les exemples traduits en français, s'ils sont répétés, ne seront plus traduits.

⁵ Liste des abréviations : S (phrase), GN (groupe nominal), GV (groupe verbal), N (nom), V (verbe), S' (proposition subordonnée), Comp (complémenteur), Ø (catégorie vide).

⁶ Pour une critique du *minimal attachment*, voir Abney (1989).

Le lecteur qui aura interprété *the answer* comme l'objet du verbe *knew* se rendra compte que son interprétation initiale est erronée en lisant *was correct* et devra ré-analyser la phrase avec une structure comme (Figure 1b), ce qui occasionne le plus souvent des temps de lecture plus importants.

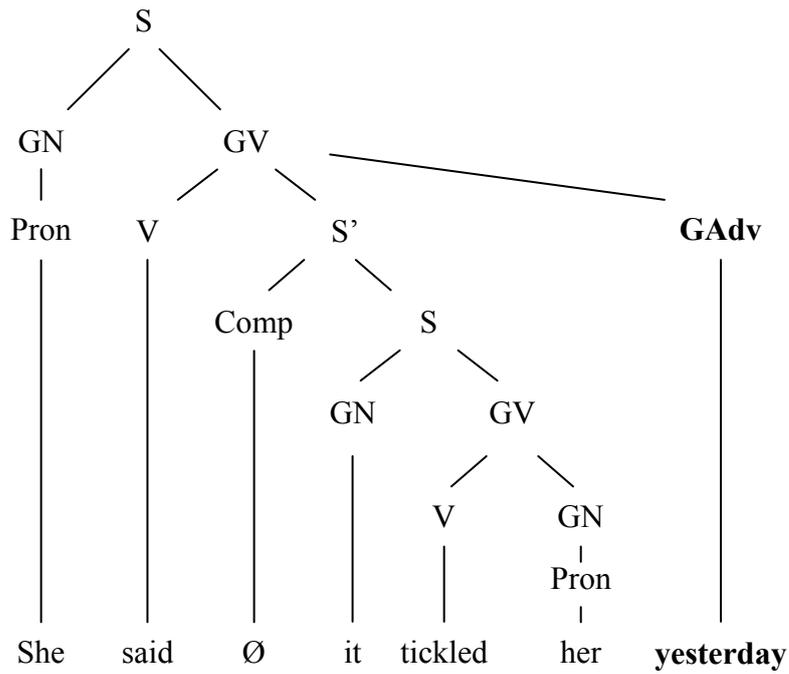
1. 3. La stratégie du *late closure*

La stratégie du *late closure* (clôture tardive) permet de choisir une interprétation lorsque deux sont possibles et que leur complexité syntaxique est comparable. Cette stratégie est proche de l'association droite (*right association*) proposée par Kimball (1973, voir aussi J. D. Fodor, 1998). Dans l'exemple (10), il est possible d'attacher l'adverbe *yesterday* au verbe de la principale (Figure 2a) ou au verbe d'une complétive (Figure 2b).

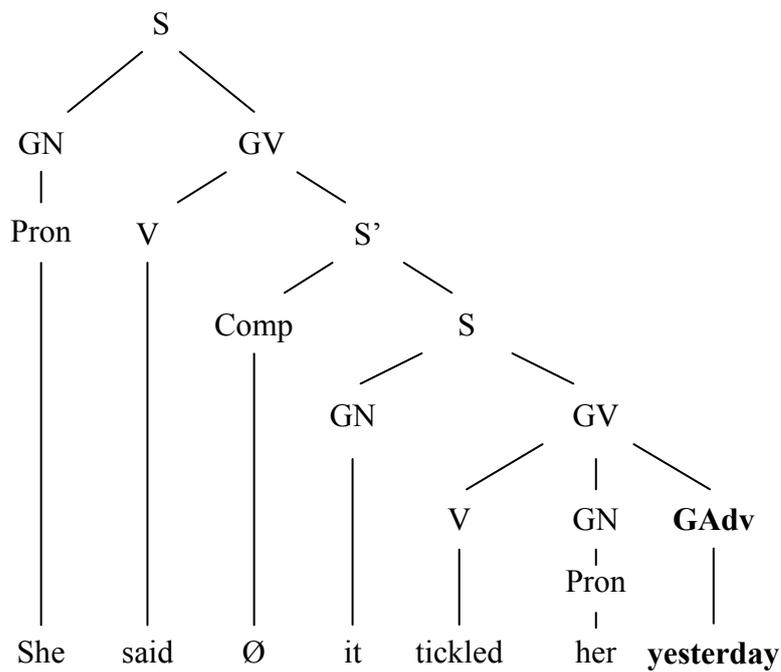
(10) *She said it tickled her yesterday.*⁷

La stratégie du *minimal attachment* ne permet pas de choisir une des interprétations disponibles, car les deux structures sont de complexité équivalente. La stratégie de *late closure* rend possible une décision d'attachement en favorisant l'interprétation la plus économique cognitivement. La stratégie consiste ainsi à attacher le constituant en cours de traitement au nœud non terminal le plus bas possible (Frazier & Fodor, 1978 ; Frazier & Rayner, 1982). Dans l'exemple présenté en (10), c'est par conséquent l'attachement de l'adverbe *yesterday* au verbe de la complétive (Figure 2b) qui est préféré, plutôt qu'au verbe de la principale (Figure 2a).

⁷ Lorsque les structures sont ambiguës, le constituant auquel les participants préférèrent attacher ces dernières est souligné. Si c'est un indice sémantique qui indique le type d'attachement à réaliser, on souligne le constituant auquel est attaché le constituant ambigu.



a. Attachement de l'adverbe au verbe de la principale



b. Attachement au verbe de la complétive

Figure 2 : Représentations arborées d'une phrase où l'attachement de l'adverbe est ambigu⁸

⁸ Liste des abréviations : Pron (pronom), GAdv (groupe adverbial), voir note 5 pour les autres abréviations.

La validité des deux stratégies, *minimal attachment* et *late closure*, a été éprouvée de façon extensive dans le cadre de la théorie du *garden path* (Clifton, 1993 ; F. Ferreira & Clifton, 1986 ; F. Ferreira & Henderson, 1990, 1991 ; Frazier & Rayner, 1982 ; Rayner *et al.*, 1983 ; Rayner & Frazier, 1987 ; Rayner, Garrod, & Perfetti, 1992). Des modèles concurrents postulent une interaction immédiate des indices syntaxiques et sémantiques. Il s'agit du modèle de satisfaction de contraintes (MacDonald, 1993, 1994 ; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993)⁹, du modèle de compétition et d'intégration (McRae, Ferretti, & Amyote, 1997 ; McRae, Spivey-Knowlton, & Tanenhaus, 1998 ; Spivey & Tanenhaus, 1998)¹⁰ et du modèle *unrestricted race* (Pickering & Traxler, 1998 ; Traxler, 2005 ; van Gompel & Pickering, 2001 ; van Gompel, Pickering, Pearson, & Liversedge, 2005 ; van Gompel, Pickering, & Traxler, 2000 ; van Gompel *et al.*, 2001).

Ces modèles ont apporté des explications et des données qui favorisent une interaction précoce (pour une revue, voir Mitchell, 1989 ; Tanenhaus, Carlson, & Trueswell, 1989) des informations syntaxiques, sémantiques (Green, 2006 ; Trueswell, Tanenhaus, & Garnsey, 1994 ; Vosse & Kempen, 2000) ou pragmatiques (Altmann & Steedman, 1988 ; Altmann, Van Nice, Garnham, & Henstra, 1998 ; Crain & Steedman, 1985 ; Farmer, Anderson, & Spivey, 2007 ; Farmer, Cargill, Hindy, Dale, & Spivey, 2007 ; Spivey-Knowlton *et al.*, 1993 ; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995).

Selon la théorie du *garden path*, l'interprétation des ambiguïtés d'attachement des relatives est guidée par la stratégie du *late closure*. Ce type de structure est présenté indépendamment des autres volontairement, car il remet en cause la théorie du *garden path* et est, on le verra, une structure problématique pour les théories formulées à ce sujet, notamment à cause de l'importante variation translinguistique et intra-linguistique qui a pu être observée expérimentalement.

⁹ Pour une mise en évidence expérimentale, voir Boland, Tanenhaus, Garnsey, & Carlson (1995), MacDonald, Just, & Carpenter (1992), MacDonald, Pearlmutter, & Seidenberg (1994), Tanenhaus, Boland, Garnsey, & Carlson (1989), Trueswell (1996), Trueswell, Sekerina, Hill, & Logrip (1999).

¹⁰ Pour une mise en évidence des prédictions du modèle basé sur les contraintes, voir Elman, Hare, & McRae (2004), Hare, Tanenhaus, & McRae (2007), Spivey-Knowlton, & Tanenhaus (1994).

2. Un cas particulier d'ambiguïté syntaxique : l'attachement des propositions relatives

L'attachement d'une proposition relative est généralement évident (il n'y a qu'un seul site d'attachement), mais il peut être ambigu lorsque plusieurs entités peuvent en être l'antécédent. On parle alors d'ambiguïté d'attachement de la proposition relative.

Par exemple, la phrase (11) contient une structure de la forme « N1 de N2 ». La proposition relative (*que tuvo el accidente*) peut être attachée soit au premier nom (abrégé désormais N1), *hija*, soit au deuxième nom (N2), *coronel*.

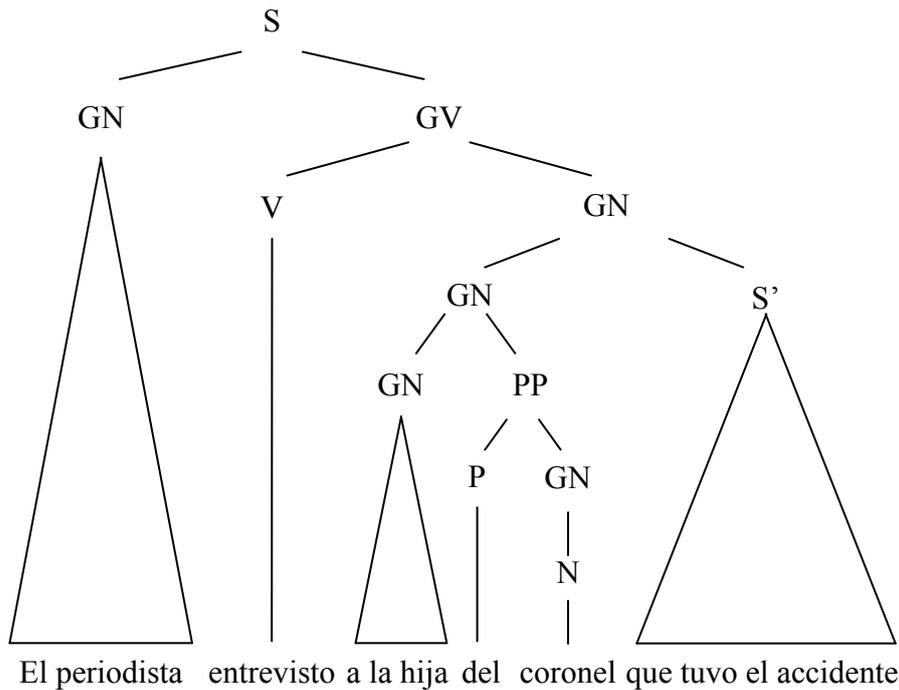
(11) *El periodista entrevistó a la hija del coronel que tuvo el accidente. (Le journaliste interroge la fille du colonel qui a eu un accident, tiré de Cuetos & Mitchell, 1988)*

Mises en regard des théories sur le traitement des ambiguïtés syntaxiques, les propositions relatives dont l'attachement est ambigu peuvent être considérées comme un bon test. La théorie du *garden path*, par exemple, résiste assez mal et c'est une théorie plus récente comme la *construal theory* qui explique mieux les préférences d'attachement qu'on observe avec ce type de structure (11).

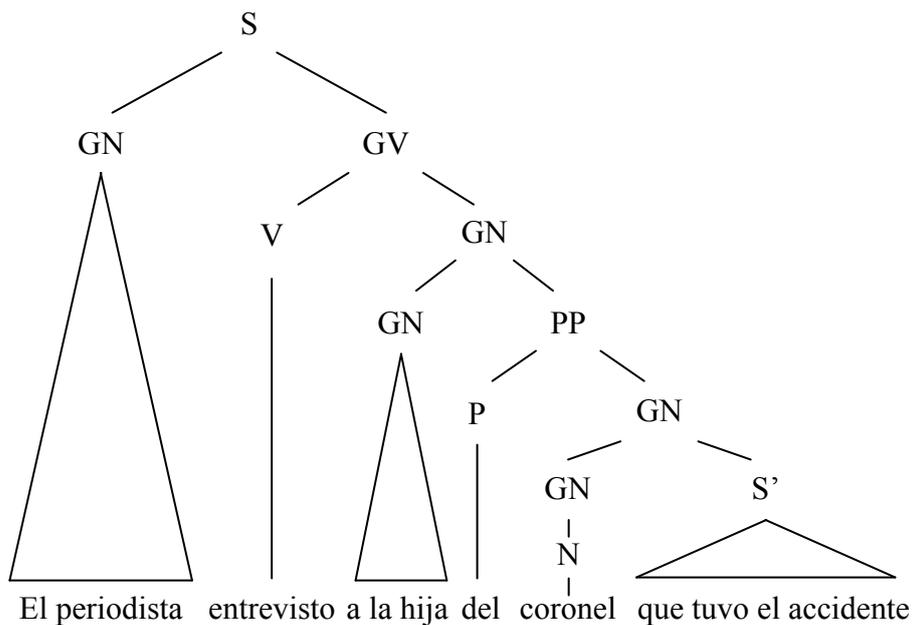
Avec cette structure, il est tout particulièrement intéressant d'observer la variation des préférences d'attachement en fonction de divers facteurs. Délibérément, je procéderai d'abord à une revue des expérimentations sur le traitement de ces ambiguïtés d'attachement, en m'intéressant aux différences translinguistiques, puis en présentant les aspects méthodologiques qui, variant d'une expérience à l'autre, pourraient contribuer aux variations entre ces expériences.

Ensuite seulement seront introduites les explications de la théorie du *garden path* et des modèles interactifs, puis d'autres modèles explicatifs plus récents, formulés dans le but d'expliquer les mécanismes sous-jacents au traitement de telles ambiguïtés.

Selon la théorie du *garden path*, la proposition relative devrait être attachée de façon préférentielle au second nom, *colonel*, en vertu de la stratégie du *late closure*. En effet, si on représente la phrase (11) sous forme arborée, on constate qu'il y a autant de nœuds pour les deux interprétations (Figure 3a-b).



a. *Attachement de la proposition relative au premier nom*



b. *Attachement de la proposition relative au second nom*

Figure 3 : Représentations arborées d'une phrase où l'attachement de la proposition relative est ambigu

La stratégie de *minimal attachment* ne peut s'appliquer et c'est par conséquent la stratégie de *late closure* qui s'applique. L'attachement se fait préférentiellement au nœud en cours de

traitement, c'est-à-dire N2 (Figure 3b). La stratégie de *late closure* prédit ainsi un attachement préférentiel vers N2 (Frazier, 1987 ; Frazier & Fodor, 1978 ; Frazier & Rayner, 1982).

Des expériences en compréhension ont confirmé, pour la langue anglaise, qu'avec une phrase de la forme de (11), les lecteurs préfèrent attacher la proposition relative à N2, que ce soit avec un questionnaire (Cuetos & Mitchell, 1988 ; Mitchell, Cuetos, & Zagar, 1990), avec une mesure des temps de lecture, ou avec la méthode de l'*eye tracking* (Carreiras & Clifton, 1999). Cependant, d'autres expériences ont montré en anglais et pour des phrases comparables une absence de différence, avec des mesures en *eye tracking* (Carreiras & Clifton, 1993 ; Traxler, Pickering, & Clifton, 1998).

Tant que l'attachement de la proposition relative reste ambigu, les temps de lecture ou les fixations oculaires mesurés ne permettent pas d'évaluer quel attachement est réalisé, vers N1 ou vers N2. Par contre, à partir du moment où l'ambiguïté est levée, que ce soit avec un indice sémantique ou grammatical comme dans (12a-b), on peut parfois observer un effet de ré-analyse. En effet, si les participants réalisent un attachement initial vers N1 et que la désambiguïstation indique un attachement vers N2, les participants doivent alors réviser leur interprétation initiale, nécessitant un traitement plus important que si la désambiguïstation concorde avec l'analyse initiale.

La réplication de telles expériences dans d'autres langues que l'anglais a montré que le lecteur attache préférentiellement la proposition relative à N1 en espagnol (Cuetos & Mitchell, 1988), ce qui n'est pas prédit par la stratégie du *late closure*. Une préférence comparable est aussi observée dans de nombreuses langues (pour une revue, voir Colonna & Pynte, 2002) comme l'espagnol (Carreiras & Clifton, 1993, 1999 ; Carreiras, Salillas, & Barber, 2004 ; Cuetos & Mitchell, 1988 ; Dussias & Sagarra, 2007 ; Gibson, Pearlmutter, & Torrens, 1999 ; Igoa, Carreiras, & Meseguer, 1998 ; Mitchell *et al.*, 1990), le néerlandais (Brybaert & Mitchell, 1996 ; Desmet, De Baecke, & Brybaert, 2002), le français (Colonna, 2001 ; Pynte & Colonna, 2000 ; Zagar *et al.*, 1997) ou l'allemand (Hemforth *et al.*, 2000). En ce qui concerne l'italien, les préférences d'attachement observées dépendent de la mesure expérimentale, questions de compréhension (N1) ou mesures de temps de lecture (N2, voir De Vincenzi & Job, 1993, 1995).

L'observation des préférences d'attachement dans ces langues met à mal la théorie du *garden path* : l'attachement, dans la majorité des langues, ne suit pas le principe de *late closure* qui devrait s'appliquer (Frazier, 1987 ; Frazier & Clifton, 1996, 1997). Cette théorie ne parvient pas à expliquer les différences translinguistiques, et d'autres théories ont proposé

des explications à cette variabilité. Avant de les présenter, je vais d'abord présenter une série de facteurs qui influencent les préférences d'attachement. En effet, d'une expérience à l'autre, d'une langue à l'autre, certains aspects du matériel testé expérimentalement changent et conduisent à des interprétations différentes. On a longtemps considéré que les différences étaient translinguistiques, mais elles peuvent tout aussi bien être liées aux différences dans le matériel utilisé.

Je vais donc revenir sur des expériences conduites sur le traitement des ambiguïtés d'attachement en insistant sur les différentes sources de variation. Les propriétés des antécédents, comme leur fréquence, leur animation, la présence d'un adjectif ou la relation sémantico-pragmatique entre les antécédents ont une influence sur les préférences d'attachement, tout comme les caractéristiques de la proposition relative (longueur, prosodie) ou la méthodologie expérimentale.

Parmi les théories, je présenterai d'abord celles qui s'attachent à expliquer l'effet de certains facteurs sur les préférences d'attachement (la proposition de De Vincenzi et al., la modifiabilité, une partie de la *construal hypothesis*), puis celles qui expliquent ce qu'on considère souvent comme étant une différence translinguistique. Il s'agit du modèle de satisfaction de contraintes, de la *construal hypothesis*, de la *tuning hypothesis* et de la *recency theory*.

2. 1. Les sources de variation

2. 1. 1. Les antécédents

Un parcours de la littérature sur les ambiguïtés d'attachement de propositions relatives montre l'importance de la nature des antécédents possibles (N1 et N2) sur les préférences d'attachement. Ainsi, les fréquences d'occurrence de N1 et N2 influencent ces préférences. Les expériences menées par Zagar *et al.* (1997) montrent qu'avec des phrases comme (12) où l'attachement est temporairement ambigu et où N1 et N2 ont des fréquences comparables, la préférence d'attachement est vers N1. Cette préférence est observée dans une tâche de complétion de phrases (incomplètes au delà de *plus* dans l'exemple (12)), les lecteurs privilégient un adjectif accordé en genre avec N1 (94 % des réponses). L'expérience suivante reprend un matériel comme (12a-b) et les mouvements oculaires réalisés pendant le traitement de la phrase sont enregistrés. Les auteurs mesurent des temps de lecture plus longs lorsque l'adjectif indique un attachement vers N2 (12b) plutôt que vers N1 (12a), à condition que les

phrases-cibles ne soient précédées d'aucun contexte. L'effet du contexte sera discuté plus loin en 2. 1. 4. Ceci témoigne d'une préférence pour un attachement vers N1.

(12) a. *Un journaliste aborda l'avocat de la chanteuse qui semblait plus confiant que de raison.*

b. *Un journaliste aborda l'avocat de la chanteuse qui semblait plus confiante que de raison.*

2. 1. 1. 1. L'effet de la fréquence des antécédents

Cette préférence pour attacher la proposition relative à N1 en français a cependant été remise en cause par les données obtenues par Baccino *et al.* (2000). En effet, les résultats montrent une préférence moindre pour N1 avec un questionnaire (55,5 %) et des temps de lecture plus courts (technique d'auto-présentation segmentée) lorsque la phrase est désambiguïsée vers N2. Dans cette étude, on mesure une préférence pour N2.

Il a été suggéré (2000) que les différences observées entre ces deux séries d'expériences (2000, 1997) sont dues à la différence de fréquence des noms choisis pour N1 et N2. En effet, la plupart des noms utilisés par Zagar *et al.* (1997) sont des noms de métiers (*avocat, chanteuse*), qui sont moins fréquents que des noms de parenté (*fils, nièce*) utilisés pour N1 par Baccino *et al.* (2000). Notons aussi que ces derniers ont employé pour leurs phrases expérimentales les mêmes noms plusieurs fois, jusqu'à trois fois, ce qui constitue un biais expérimental.

N1 et N2 diffèrent moins en fréquence dans le matériel utilisé par Zagar *et al.* (1997) que dans le matériel de Baccino *et al.* (2000), où N1 est plus fréquent. La fréquence d'occurrence des éléments de la phrase, et tout particulièrement des sites d'attachement possibles, peut avoir un effet sur le traitement de la phrase et par conséquent sur les préférences d'attachement. En effet, on peut supposer qu'une entité dont la fréquence est moindre pourra être considérée comme plus saillante (Heine, Tamm, Hofmann, Hutzler, & Jacobs, 2006 ; Schafer, Carter, Clifton, & Frazier, 1996 ; van Gompel & Majid, 2004), ce qui peut conduire à plus d'attachements vers le nom le moins fréquent, soit N2 dans les expériences de Baccino *et al.* (2000).

Pynte et Colonna (2000) ont testé cette hypothèse avec une expérience en *eye tracking* où la fréquence de N1 et N2 est manipulée. Ainsi, soit N1 est moins fréquent que N2 (13a), ou l'inverse (13b). La désambiguïsation est située au niveau de l'adjectif, marqué en genre. Dans

(13a-b), l'adjectif masculin, *dépressif*, indique un attachement à N1 et l'adjectif féminin, *dépressive*, un attachement vers N2.

- (13) a. *Il parle à l'époux de la servante qui semble dépressif(ve) depuis peu.*
 b. *Il parle au mari de la soubrette, qui semble dépressif(ve) depuis peu.*

La technique de l'*eye tracking* permet de réaliser plusieurs types de mesures (cf. 2. 1. 4), dont les premières fixations et les temps de lecture totaux. Les premières fixations, qui témoignent des préférences initiales (Rayner, 1978 ; Rayner *et al.*, 1983 ; Rayner & Clifton, 2002), montrent, au niveau de la désambiguïsation (l'adjectif), une préférence pour attacher la relative au nom le moins fréquent : les fixations sont plus courtes. Ainsi, celui-ci serait plus saillant et par conséquent l'antécédent auquel l'attachement initial est réalisé. C'est ce qui est mis en évidence par Baccino *et al.* (2000). Les temps de lecture totaux mesurés sur cette même zone révèlent aussi un effet de la fréquence : les temps de lecture sont plus courts quand N1 est le nom le moins fréquent et que l'adjectif qui lève l'ambiguïté est accordé avec cet antécédent. La fréquence lexicale influence les préférences d'attachement initiales.

Cependant, une autre façon de manipuler la fréquence lexicale consiste par exemple à présenter des noms propres. Sans vouloir tester ici l'hypothèse de la fréquence lexicale, Frenck-Mestre et Pynte (2000a) mettent en évidence en français une préférence pour N1 avec des phrases où N2 est systématiquement un nom propre. La traduction en français de l'exemple (14a) est une des phrases utilisées dans cette expérience. On peut considérer que la différence de fréquence entre N1 et N2 explique ce résultat.

Une autre étude (Frenck-Mestre & Pynte, 2000b) permet d'éprouver plus avant l'hypothèse de Pynte et Colonna (2000) au sujet de l'effet de la fréquence sur les préférences d'attachement. Dans une expérience en italien et avec la technique d'*eye tracking*, Frenck-Mestre et Pynte (2000b) reprennent des phrases (14b) utilisées auparavant par Baccino *et al.* (2000), pour fin de comparaison, mais aussi des phrases traduites du français, avec un nom propre pour N2, (14a), adaptées de Frenck-Mestre et Pynte (2000a, voir aussi Hemforth *et al.*, 2000). Quelle que soit l'origine de la phrase, N2 est considéré avoir une fréquence plus importante et, par conséquent, N1 qui a une fréquence moindre serait plus saillant. La désambiguïsation vers N1 ou N2 est réalisée par un marquage en genre.

- (14) a. *Gianni osserva il ragioniere de Caterina che sembra più pensieroso(a) del normale. (Gianni observe le comptable de Caterina qui semble plus pensif(ve) que la normale)*

b. *Patrizia conosceva la segretaria del direttore che era svenuto(a) alla festa.*
(*Patrizia connaît la secrétaire du directeur qui s'était évanoui(e) à la fête*)

Au niveau de l'adjectif qui lève l'ambiguïté, *pensieroso(a)* dans (14a), les temps de fixation sont plus importants pour des phrases comme (14a), mais il n'y a pas d'effet du type d'attachement. Il n'y a donc pas de préférence d'interprétation initiale. Par contre, au niveau des temps de lecture totaux, pour cette même région, on observe une interaction significative entre l'origine des phrases, (14a) ou (14b), et le type d'attachement, vers N1 ou vers N2. Ainsi, les temps de lecture sont plus importants pour des phrases de type (14a), où N2 est un nom propre, lorsque celles-ci sont désambiguïsées vers N2. Il y a donc une préférence pour attacher la proposition relative à N1.

Par contre, aucune préférence d'interprétation n'est observée pour les phrases comme (14b), ce qui va à l'encontre des résultats obtenus originalement avec ce matériel (Baccino *et al.*, 2000). Cette absence de préférence peut être interprétée par le fait que N1 et N2 dans les items présentés ne diffèrent pas suffisamment en fréquence (Frenck-Mestre & Pynte, 2000b). On peut aussi suggérer que le nombre d'items, 8, issus de l'expérience de Baccino *et al.* (2000) est faible comparé aux 16 items expérimentaux repris de Frenck-Mestre et Pynte (2000a), et que c'est pour cette raison qu'aucun effet du type d'attachement n'a été trouvé. Quant aux items construits pour cette expérience comme pour la version française (Frenck-Mestre & Pynte, 2000a), on observe une préférence pour N1, même si elle est moins marquée. La différence de fréquence entre N1 et N2 et la différence de saillance que cela entraîne peuvent ainsi contribuer aux variations des préférences d'attachement. Cependant, on peut aussi considérer que les noms propres, utilisés pour N2 dans des phrases comme (14a), ont un statut cognitif différent des noms communs (voir plus loin la givenness hierarchy, Gundel, Hedberg, & Zacharski, 1993). Ce statut différent est à l'origine de la différence de saillance de N1 et N2. Cette interprétation explique les résultats observés selon l'origine du matériel expérimental.

2. 1. 1. 2. La position de l'adjectif

Un des arguments avancés par Cuertos et Mitchell (1988) pour expliquer les différences de préférences entre l'anglais et l'espagnol est la position de l'adjectif dans ces langues. L'adjectif est généralement post-nominal en espagnol (*el hombre viejo, le vieil homme*) et pré-nominal en anglais (*the old man*, traduction identique). Ainsi, en espagnol, il est possible d'avoir une structure « N1 adjectif de N2 », ce qui facilite les attachements vers N1 avec les

phrases utilisées dans l'expérience. La structure « de N2 » est alors considérée comme étant un modifieur de N1, auquel la proposition relative, qui est un modifieur elle aussi, sera préférentiellement attachée (voir aussi Gibson, 1998).

Gilboy Sopena, Clifton, et Frazier (1995) testent cette hypothèse dans leur étude de questionnaires en présentant, entre autres, des phrases de la forme « N1 de N2 » suivies d'une proposition relative pour lesquelles un des noms est accompagné d'un adjectif (15b-c), ou bien aucun n'a d'adjectif (15a). D'autres propriétés sont par ailleurs évaluées, mais seront présentées plus tard. Les questionnaires sont administrés en espagnol et en anglais à des locuteurs natifs des deux langues.

(15) a. *The professor read the book of the student that was in the dining room. (Le professeur a lu le livre de l'étudiant qui était dans la salle à manger)*

b. *The professor read the new book of the student that was in the dining room. (Le professeur a lu le nouveau livre de l'étudiant qui était dans la salle à manger)*

c. *The professor read the book of the new student that was in the dining room. (Le professeur a lu le livre du nouvel étudiant qui était dans la salle à manger)*

Les résultats montrent pour l'espagnol un effet significatif de la présence de l'adjectif, mais pas en anglais. Ainsi, en espagnol, les participants sont plus enclins à réaliser un attachement vers N1 quand aucun des noms n'a d'adjectif (15a), que lorsque N1 a un adjectif (15b). On observe le moins d'attachements vers N1 lorsque N2 a un adjectif (15c). Ces résultats mettent en évidence que les participants réalisent plus d'attachements vers N2, c'est-à-dire que la proposition relative modifie le nom qui a déjà un modifieur, l'adjectif.

2. 1. 1. 3. L'animation

L'animation ou plus particulièrement le caractère humain des noms utilisés pour N1 et N2 influence aussi les préférences d'attachement. Ce type d'information suffit à contraindre l'interprétation des participants dans certains cas d'ambiguïtés syntaxiques (propositions relatives réduites, voir Trueswell *et al.*, 1994) mettent en évidence que l'animation du premier nom guide l'interprétation. On peut faire l'hypothèse que l'animation aura un effet comparable sur un autre type de structure ambiguë, l'attachement de propositions relatives. Le matériel expérimental utilisé par Cuetos et Mitchell (1988) intègre des antécédents humains ou non, qui ne sont pas séparés pour l'analyse des résultats en espagnol. Cependant, dans la version en anglais de ces items (Expérience 1B), les auteurs ont décidé d'analyser ces derniers

en deux groupes. Dans le premier, les deux noms sont humains et animés (11 items (16a)) et, dans le second, N1 est non humain et non-animé et N2 humain (14 items, (16b)).

(16) a. *Someone shot the servant of the actress who was on the balcony with her husband.* (*Quelqu'un a tué le serviteur de l'actrice qui était sur le balcon avec son mari*)

b. *Peter was looking at the book of the girl that was in the living-room watching TV.* (*Peter regardait le livre de la fille qui était dans le salon à regarder la télé*)

Les auteurs justifient cette dissociation par le fait que le relativiseur *que* en espagnol est soit *who* dans le cas où les antécédents sont humains soit *that* si un des antécédents est non humain, afin d'éviter toute désambiguïisation dès le niveau du relativiseur dans le second cas. Notons cependant que certaines variétés de l'anglais acceptent que *that* soit attaché à un antécédent animé et humain (Romaine, 1982, cité par Traxler *et al.*, 1998). Les analyses pour l'anglais révèlent une interaction significative entre ces deux groupes : il y a plus d'attachements vers N2 quand les deux antécédents sont humains (16a) et plus d'attachements vers N1 quand seul N2 est humain (16b).

La plupart des expériences présentées jusqu'à présent utilisent systématiquement deux antécédents animés et humains (Baccino *et al.*, 2000 ; Brysbaert & Mitchell, 1996 ; Carreiras & Clifton, 1993, 1999 ; Carreiras *et al.*, 2004 ; De Vincenzi & Job, 1993, 1995 ; Frenck-Mestre & Pynte, 2000a, 2000b ; Hemforth *et al.*, 2000 ; Pynte & Colonna, 2000 ; Zagar *et al.*, 1997), à l'exception de celle de Gibson *et al.* (1999) où seuls des antécédents non-animés sont utilisés et de celles qui manipulent expérimentalement l'animation (Desmet, De Baecke, Drieghe, Brysbaert, & Vonk, 2006 ; Gilboy *et al.*, 1995).

Ces dernières études mettent en évidence que l'animation influence les préférences d'attachement. Ainsi, une analyse de corpus en néerlandais (Desmet *et al.*, 2006) montre qu'il y a plus d'attachements vers N1 quand les deux antécédents sont animés, plus d'attachements vers N2 quand les deux antécédents sont non-animés et une préférence pour l'antécédent animé si les deux antécédents diffèrent quant à l'animation. À la suite de cette étude de corpus, les auteurs réalisent une expérience en *eye tracking* où deux aspects de N1 sont manipulés. N1 est animé, (17a-b) ou non-animé (17c-d), abstrait ((17a), (17c)) ou concret ((17b), (17d)). N2 est toujours un animé concret (*president*) et la désambiguïisation est située au niveau du verbe qui est soit au pluriel (*garanderen*), indiquant un attachement à N1, soit au singulier (*garandeert*), indiquant un attachement vers N2.

(17) a. *De bevolking zonder toekomstperspectieven respecteert de organisaties van de president die garanderen/garandeert dat er geen oorlog komt. (La population sans perspectives d'avenir respecte les organisations du président qui garantissent/garantit qu'il n'y aura pas la guerre)*

b. *De bevolking zonder toekomstperspectieven respecteert de raadgevers van de president die garanderen/garandeert dat er geen oorlog komt. (La population sans perspectives d'avenir respecte les conseillers du président qui garantissent/garantit qu'il n'y aura pas la guerre)*

c. *De bevolking zonder toekomstperspectieven respecteert de beslissingen van de president die garanderen/garandeert dat er geen oorlog komt. (La population sans perspectives d'avenir respecte les décisions du président qui garantissent/garantit qu'il n'y aura pas la guerre)*

d. *De bevolking zonder toekomstperspectieven respecteert de documenten van de president die garanderen/garandeert dat er geen oorlog komt. (La population sans perspectives d'avenir respecte les documents du président qui garantissent/garantit qu'il n'y aura pas la guerre)*

Les auteurs observent au niveau des temps de lecture totaux une interaction significative. En fait, ce n'est que lorsque N1 est animé et concret que les participants montrent une préférence pour réaliser un attachement vers N1 (17a). Dans les autres conditions (17b-d), c'est toujours une préférence pour N2, animé et concret, qui est mesurée. Ces résultats confirment ceux du corpus : l'attachement est réalisé de façon préférentielle vers l'antécédent animé et concret, et vers N1 quand les deux antécédents sont animés et concrets. Dans une étude de corpus en français (Kister, 2002), les fréquences d'attachement mesurées confirment la préférence pour attacher la proposition relative à l'entité animée.

Par contre, cette préférence pour un attachement à l'entité animée peut remettre en question l'interprétation de Cuetos et Mitchell (1988) concernant la préférence pour N2 en anglais quand seul N2 est humain : cet effet n'est peut-être pas dû à la stratégie du *late closure*, mais plutôt au caractère animé de N2.

L'étude de Gilboy *et al.* (1995, voir aussi Igoa *et al.*, 1998) permet aussi d'observer l'effet de l'animation sur les préférences d'attachement. Il s'agit d'une étude de questionnaires, en espagnol et en anglais. La tâche des participants consiste à indiquer le type d'attachement, vers N1 ou vers N2. Ici, non seulement l'animation des antécédents est manipulée, mais aussi

la relation entre les deux antécédents. Parmi les neuf types de relations manipulés, c'est celle où N1 est non-animé, N2 animé, N2 possédant N1 (*the book of the student, le livre de l'étudiant*) qui est particulièrement intéressante ici. En effet, les résultats de cette condition permettent une comparaison avec ceux obtenus par Desmet *et al.* (2006) puisqu'il s'agit d'un matériel comparable. Dans le cas où N1 est non-animé et N2 animé, on observe que l'attachement est réalisé de façon préférentielle vers N2, même si cette préférence est plus forte en espagnol (64 % des attachements) qu'en anglais (53 %). Ainsi, que ce soit en néerlandais, en espagnol ou en anglais, les participants préfèrent attacher la proposition relative à l'antécédent animé.

2. 1. 2. Les relations entre les antécédents

Afin d'évaluer l'effet de la structure thématique sur les préférences d'attachement, Gilboy *et al.* (1995) manipulent les relations entre N1 et N2 dans une structure « N1 de N2 », mais aussi « N1 avec N2 » (voir aussi Thornton, Gil, & MacDonald, 1998).

Les résultats des questionnaires mettent en évidence que dans le cas où les deux antécédents sont non-animés, le site d'attachement préféré est N1 pour les relations de substance (*the sweater of cotton, le sweater de coton*) et de quantité (*the glass of water, le verre d'eau*). Il n'y a pas de préférence avec une relation de possession (*the museum of the city, le musée de la ville*) ou une relation dite de représentation (*the etching of the sculpture, la gravure de la sculpture*). La proposition relative est attachée plus fréquemment à N2 avec des relations de possession inhérente (*the branch of the tree, la branche de l'arbre*). Ces résultats mettent en évidence que selon les relations entre N1 et N2, la proposition relative peut plus ou moins fréquemment être considérée être un modifieur de N2.

Enfin, la dernière relation présente non plus la préposition *of* (*de*), mais *with* (*avec* en français, *the steak with the sauce, le steak avec la sauce*). C'est ici qu'on observe la plus grande proportion d'attachements vers N2 (83 % en espagnol et 69 % en anglais). La relation, le lien entre N1 et N2 est considéré plus fort avec la préposition *de* qu'avec la préposition *avec* (pour une comparaison de ces deux prépositions, voir aussi Solomon & Pearlmuter, 2004).

Ces résultats sont de façon générale concordants avec ceux obtenus dans d'autres expériences qui utilisent une structure identique en français (Baccino *et al.*, 2000) et en italien (De Vincenzi & Job, 1993, 1995). Notons aussi que Frenck-Mestre et Pynte (2000a) observent l'influence de structures « N1 avec N2 » sur le traitement de structures « N1 de

N2 ». En effet, la préférence pour attacher la proposition relative à N2 est très forte quand elle est précédée par « N1 avec N2 ». Si les participants traitent d'abord ce type de phrases, puis les phrases avec « N1 de N2 », les attachements à N2 sont alors plus fréquents, par comparaison à la condition où ces phrases ne sont pas précédées de celles contenant « N1 avec N2 ». Cette influence des prépositions est aussi répliquée dans les études présentant conjointement ces deux types de structures, puisqu'elles mettent en évidence des attachements vers N2 plus fréquents avec une structure comme « N1 de N2 » (Baccino *et al.*, 2000 ; De Vincenzi & Job, 1993, 1995), pour des langues où on observe habituellement une préférence pour N1 (Clifton, 1993 ; Cuetos & Mitchell, 1988 ; Frenck-Mestre & Pynte, 1997 ; Zagar *et al.*, 1997).

L'effet du type de préposition s'explique par le fait qu'avec *of*, N2 est un argument de N1. Par contre, N2 n'est plus l'argument de N1 lorsqu'il est précédé de *with*, il devient un modifieur de N1. Le lien entre N1 et N2 est moins fort, ce qui facilite les attachements vers N2 (Frazier & Clifton, 1996, p. 73). Cette question de la relation entre N1 et N2 en fonction de la préposition sera détaillée plus tard dans la partie consacrée à la *construal hypothesis* (voir Chapitre 1, section 2. 2. 2. 2).

Enfin, un des arguments avancé par Frazier et Clifton (1996) pour expliquer les différences de préférence d'attachement en anglais et en espagnol (Cuetos & Mitchell, 1988) est qu'en anglais, contrairement à l'espagnol, il est possible d'exprimer le génitif d'une autre façon, avec le génitif saxon (*the colonel's daughter* au lieu de *the daughter of the colonel*), qui est d'ailleurs la forme préférée si le possesseur est une entité animée (Rosenbach, 2005). Dans ce cas, l'attachement est non ambigu (vers *daughter*). D'un point de vue pragmatique, on peut supposer qu'un locuteur respecte une des maximes de manière, « be clear » (« évitez de vous exprimer de manière obscure », d'après Grice, 1975, voir Sperber & Wilson, 1989). Par conséquent, le locuteur évite toute forme d'ambiguïté. Ainsi, il préfère le génitif saxon comme dans (18b), car l'attachement est non ambigu : N2, *daughter*, est le seul antécédent possible. Le destinataire, lui, suppose que le locuteur suit la maxime de clarté et que l'utilisation d'une phrase comme (18a) est intentionnelle, c'est-à-dire qu'il infère que le locuteur ne veut pas signaler un attachement à *daughter*, auquel cas (18b) aurait été approprié, mais un autre attachement, à *colonel*. C'est pour cette raison que les participants en anglais préfèrent attacher les propositions relatives à N2, pour des structures « N1 de N2 ».

(18) a. *The journalist interviewed the daughter of the colonel who had had the accident.* (*Le journaliste interroge la fille du colonel qui a eu un accident*, tiré de Cuetos & Mitchell, 1988)

b. *The journalist interviewed the colonel's daughter who had had the accident.*
(traduction identique, mais avec un génitif saxon)

Le génitif saxon existe aussi en néerlandais et en allemand, même si son usage est plus limité. Par exemple, il sera presque exclusivement utilisé avec un nom propre en néerlandais (*Peters dochter*, *la fille de Pierre*). Cependant, c'est une préférence pour attacher la relative à N1, contrairement à N2 en anglais, qui est observée dans ces langues avec une structure « N1 de N2 » (Brysbaert & Mitchell, 1996 ; Hemforth *et al.*, 2000).

2. 1. 3. L'effet des caractéristiques de la proposition relative

Si la forme des antécédents influence les préférences d'attachement, les caractéristiques de la proposition relative ont aussi un effet. Frazier (1990b), par exemple, suggère que des informations de niveau pragmatique peuvent entrer en jeu. Ainsi, lorsque la proposition relative contient des informations pragmatiques, dépendantes du contexte, comme la localisation dans (19b), il y a plus d'attachements vers N1 que lorsque de telles informations ne sont pas présentes, comme dans (19a). Cette hypothèse est confirmée par le questionnaire réalisé par Frazier et Clifton (1987, cité par De Vincenzi & Job, 1993, 1995) où la préférence pour N1 est la plus importante avec des phrases comme (19b) : 70 % des attachements sont vers N1, contre 59 % pour un matériel comme (19a).

(19) a. *Julie met the friend of the man who reads news on Saturday Night Live.* (*Julie a rencontré l'ami de l'homme qui présente les nouvelles sur le Saturday Night Live*)

b. *Julie met the friend of the man who was in the living room.* (*Julie a rencontré l'ami de l'homme qui était dans le salon*)

La longueur de la proposition relative, ainsi que la longueur de la région ambiguë, peuvent aussi influencer les préférences d'attachement. Se référant à son modèle originel (*sausage machine*, Frazier & Fodor, 1978), J. D. Fodor (1998) suggère que la segmentation (saucissonnage) des propositions de la phrase et l'attachement qui en découle dépendent de la taille de ces propositions. Un constituant long sera considéré comme une unité propre qui sera attachée aux autres unités de la phrase. Dans ce cas, un attachement local n'est pas favorisé, alors qu'un tel attachement le sera quand le constituant est court, puisqu'il sera plus

facilement groupé en une unité avec les constituants précédents si ceux-ci sont courts. Ainsi, plus le constituant à attacher est petit ou léger, plus il sera dans la même unité que le constituant précédent et la stratégie de *late closure* est favorisée (p. 63 Frazier & Clifton, 1996). Avec des phrases comme (20a-b) présentées dans un questionnaire, les participants montrent une préférence pour attacher les syntagmes temporels au verbe *died*. Cette préférence est guidée par la stratégie de *late closure* et d'autant plus forte que le syntagme temporel est court ((20a), *last night, tomorrow*).

(20) a. *John will explain to the kids that their grandfather died last night/tomorrow.*
(*John expliquera aux enfants que leur grand-père est mort hier soir/demain*)

b. *John will explain to the kids that their grandfather died after the operation that he had/they need.* (*John expliquera aux enfants que leur grand-père est mort après l'opération qu'il a subie/dont ils ont besoin demain/hier soir*)

De même, dans le cas d'attachements ambigus de la relative, plus la proposition relative est courte, plus elle aura tendance à être attachée à N2. *A contrario*, plus la proposition relative est longue, plus elle aura tendance à être attachée à une structure de taille proche, « N1 de N2 », et c'est alors un attachement à N1 qui est préféré (J. D. Fodor, 1998). C'est ce que montrent les résultats d'une expérience en français (Expérience 2, Pynte & Colonna, 2000) où la proposition relative est soit courte (21a), soit longue, rallongée à l'aide d'un groupe prépositionnel (21b). La désambiguïsation vers N1 ou N2 est située au niveau du verbe de la relative qui est accordé en nombre soit avec N1, soit avec N2. Les temps de lecture mesurés pour la proposition relative courte dénotent une préférence pour un attachement vers N2 alors que les mesures pour la proposition relative longue montrent une préférence pour N1.

(21) a. *Il connaît le(s) frère(s) de la (des) fille(s) qui entre(nt).*

b. *Il connaît le(s) frère(s) de la (des) fille(s) qui entre(nt) dans la cour.*

Plus qu'une segmentation contrainte par le parseur syntaxique, c'est la prosodie qui exerce en fait la plus grande contrainte. Tout d'abord, un marqueur prosodique comme l'accentuation d'un mot peut faciliter un attachement de la proposition relative vers le nom accentué. Cela a par exemple été mis en évidence dans une expérience en modalité auditive au cours de laquelle les participants écoutent des phrases (22) où N1 (*propeller*) et N2 (*plane*) sont tour à tour accentués (Schafer *et al.*, 1996). La fréquence d'attachement vers N2 augmente lorsque ce dernier est accentué.

(22) *The sun sparkled on the propeller of the plane that the mechanic was so carefully examining. (Le soleil brillait sur le moteur de l'avion que le mécanicien avait examiné si attentivement)*

Cette observation a par ailleurs été répliquée avec des phrases comme (11), répétée ci-dessous, dans plusieurs langues (Jun, 2003 ; Maynell, 1999, 2000).

(11) *El periodista entrevistó a la hija del coronel que tuvo el accidente.*

2. 1. 4. Les effets de la méthodologie expérimentale

Même dans une tâche de lecture silencieuse, la prosodie de la phrase est néanmoins présente et peut influencer le traitement et, par conséquent, l'attachement des propositions relatives (J. D. Fodor, 1998). La prosodie peut indirectement être modifiée selon la segmentation des phrases lors d'expériences utilisant la technique d'auto-présentation segmentée (voir Chapitre 3 de la partie expérimentale pour une description plus détaillée). Comme les constituants de la phrase peuvent apparaître seuls ou groupés, lire la structure « N1 de N2 » en un segment (/ N1 de N2 /) ou en deux (/ N1 / de N2 /) peut affecter les préférences d'attachement, tout du moins pendant le traitement de la phrase.

Toutes les expériences utilisant le paradigme d'auto-présentation segmentée et présentant « /N1 de N2/ » en un seul segment obtiennent une préférence pour N1 en espagnol (Carreiras & Clifton, 1993 ; Cuetos & Mitchell, 1988) ou une absence de préférence en anglais (Carreiras & Clifton, 1993). Celles qui présentent « /N1/ de N2/ » en deux segments successifs obtiennent une préférence pour N2 en français (Baccino *et al.*, 2000) ou en italien (De Vincenzi & Job, 1993, 1995).

Gilboy et Sopena (1996) ont ainsi étudié l'effet de la segmentation en réalisant une expérience en espagnol avec la technique d'auto-présentation segmentée. Les phrases contiennent des propositions relatives précédées d'une structure « N1 de N2 ». La segmentation est manipulée de façon à ce que N1 et N2 soient présentés soit dans le même segment (23a), soit dans deux segments successifs (23b). L'attachement de la proposition relative vers N1 ou N2 est signalé par l'information sémantique et par un marquage en genre.

(23) a. *La policía arrestó a la hermana del criado / que dió a luz recientemente dos gemelos. (La police a arrêté la sœur du bricoleur / qui a récemment mis au monde des jumeaux)*

b. *La policía arrestó / a la hermana / del criado / que dió a luz recientemente dos gemelos.* (traduction identique, segmentation différente)

Les analyses statistiques révèlent une interaction significative, montrant une préférence pour N1 uniquement lorsque N1 et N2 sont présentés dans le même segment (23a). Une présentation plus segmentée comme (23b) ne révèle aucune préférence initiale d'attachement.

Néanmoins, il est difficile de conclure que c'est la segmentation qui entraîne cette différence de préférence, notamment à cause de l'expérience en espagnol de Gibson *et al.* (1999) qui contredit celle de Gilboy et Sopena (1996). Dans cette expérience, « N1 de N2 » est présenté en deux segments, mais les résultats montrent cette fois une préférence pour attacher la proposition relative à N1. Qui plus est, une expérience réalisée en potentiels évoqués, où les mots sont présentés les uns à la suite des autres, confirme la préférence pour N1 en espagnol (Carreiras *et al.*, 2004).

La technique de l'*eye tracking* permet de présenter les phrases expérimentales en entier, ce qui neutralise l'influence potentielle de la segmentation sur les préférences d'attachement, en raison notamment des ruptures prosodiques que cette segmentation peut induire (Rayner, Sereno, Morris, Schmauder, & Clifton, 1989).

Les expériences utilisant l'*eye tracking* montrent de façon générale une préférence pour N1 en espagnol (Carreiras & Clifton, 1999), néerlandais (Desmet, De Baecke *et al.*, 2002) ou français (Pynte & Colonna, 2000 ; Zagar *et al.*, 1997). La seule exception demeure pour l'anglais (Carreiras & Clifton, 1999).

En ce qui concerne le traitement de la phrase, surtout lorsqu'on s'intéresse aux traitements précoces, la technique de l'*eye tracking* est parfois considérée comme la plus adaptée. En effet, certains estiment que cette technique est plus sensible aux traitements initiaux (Rayner, 1978, 1998 ; Rayner *et al.*, 1983). Cette méthode est aussi très sensible aux interruptions en cours de traitement qui sont le plus souvent causées par une difficulté supplémentaire à traiter l'information ou bien par une ré-analyse (pour une introduction, voir Rayner & Sereno, 1994).

Pour chaque région, composée d'un ou plusieurs mots, plusieurs types de mesures sont possibles. Il y a tout d'abord les premières fixations, qui correspondent au temps pendant lequel le participant reste sur un mot, qu'il atteigne une première fois. Il y a aussi les régressions, c'est-à-dire le pourcentage de saccades réalisées sur les parties antérieures de la phrase. Enfin, il y a le temps de lecture total, qui représente la somme de toutes les fixations.

Ainsi, certains auteurs (F. Ferreira & Henderson, 1993 ; Rayner *et al.*, 1983) considèrent qu'avec l'*eye tracking* on peut mesurer les traitements initiaux (premières fixations), l'effet de ré-analyse (régressions) et la difficulté globale de traitement (temps de lecture total).

Parmi les méthodologies présentées jusqu'à présent, deux types de mesures sont utilisés : les mesures *on-line* et *off-line*. Les premières sont acquises lors du traitement de l'information. Il peut s'agir de temps de lecture (auto-présentation segmentée), de fixations oculaires (*eye tracking*) ou de potentiels évoqués (EEG). Les mesures *off-line* interrogent les processus post-interprétatifs. Les erreurs aux questions de compréhension, ainsi que les réponses fournies lors de questionnaires sont deux types de mesure *off-line* qui sont fréquemment utilisées.

Aussi, les résultats observés peuvent différer selon le type de mesure réalisé. Par exemple, De Vincenzi et Job (1993, 1995) obtiennent en italien une préférence initiale (pendant la lecture) pour attacher la proposition relative à N2, mais la préférence finale est un attachement vers N1.

2. 2. Comment les modèles expliquent les variations pour les préférences d'attachement

Les modèles qui vont être présentés maintenant traitent de la variabilité qu'on peut observer dans les préférences d'attachement : variation dans les langues liée aux propriétés des constituants de la phrase, variation translinguistique liée à des différences structurelles ou autres entre les langues.

2. 2. 1. Les variations dans une langue

De nombreux facteurs varient entre les expériences et sont susceptibles d'entraîner des différences dans les préférences d'interprétation. De Vincenzi *et al.* (1993, 1995) proposent une approche pour concilier les prédictions de la théorie du *garden path* avec les données recueillies dans des expériences en compréhension. Le principe de modifiabilité proposé par Thornton *et al.* (1998 ; Thornton, MacDonald, & Gil, 1999), lui, rend compte de l'effet de la présence d'un adjectif sur les préférences d'attachement, pour une même langue.

2. 2. 1. 1. L'hypothèse de De Vincenzi et Job (1993, 1995)

Selon la théorie du *garden path*, le principe de *late closure* guide l'interprétation, l'attachement de propositions relatives ambiguës. Les expériences de la littérature invalident fréquemment cette prédiction. L'hypothèse de De Vincenzi *et al.* (1993, 1995) peut être vue comme une tentative de conserver néanmoins cette théorie qui postule que l'attachement initial est guidé par les informations syntaxiques uniquement. Les informations autres que syntaxiques, sémantiques par exemple, ne sont prises en compte que dans une deuxième étape, la ré-analyse.

Selon De Vincenzi et Job (1993, 1995), l'interprétation initiale suit le principe de *late closure*, ce qui est mesuré par les temps de lecture. Pendant le traitement de la phrase, il y a ainsi une préférence pour N2 qui est mesurée quand l'ambiguïté est levée. En italien, leurs expériences révèlent des temps de lecture plus importants quand l'ambiguïté est levée en faveur de N1. La préférence pour N1 émerge plus tard dans la compréhension de la phrase et c'est la raison pour laquelle on l'observe de façon régulière dans des études de questionnaires et surtout dans les réponses aux questions de compréhension qui suivent les phrases expérimentales (Baccino *et al.*, 2000 ; De Vincenzi & Job, 1993, 1995).

Si les expérimentations valident l'hypothèse de ces auteurs, notons cependant que ces expérimentations présentent non seulement des propositions relatives qui sont précédées d'une structure « N1 de N2 », mais aussi de « N1 avec N2 », ce qui est très rare dans la littérature (cf. Traxler *et al.*, 1998), tout du moins si on s'intéresse à des mesures *on-line*, cruciales pour la théorie du *garden path*. Il est cependant difficile de distinguer ce qui contribue aux résultats mesurés, entre la proposition des auteurs et une interaction possible entre les structures où la proposition relative est attachée.

2. 2. 1. 2. Le principe de modifiabilité

Le principe de modifiabilité de Thornton *et al.* (1998, 1999) s'inscrit dans le cadre du modèle de satisfaction de contraintes (MacDonald, 1993, 1994 ; Spivey & Tanenhaus, 1998). Très succinctement, ce modèle propose que les informations de tout niveau, lexical (Duffy, Morris, & Rayner, 1988 ; Rayner & Duffy, 1986 ; Tabor, Juliano, & Tanenhaus, 1997 ; Trueswell, 1996), syntaxique, sémantique ou pragmatique fonctionnent de façon parallèle. Lors de l'interprétation d'un mot dans une phrase, toutes les interprétations possibles sont

activées. L'ambiguïté d'interprétation est levée lorsque les analyses basées sur les différents niveaux d'information convergent vers une seule structure.

Le principe de modifiabilité affirme que la quantité de modifications qu'un antécédent subit conditionne la probabilité qu'il soit à nouveau modifié. Dans la structure « N1 de N2 », N1 est modifié par *de* N2. Une proposition relative est aussi un modifieur et, d'après Thornton *et al.* (1999), que N1 soit à la fois modifié par un adjectif et par *de* N2 a pour effet de favoriser un attachement à N2. Si, par contre, N2 est modifié par un adjectif, c'est une préférence pour N1 qui est alors attendue. L'étude de questionnaires de Gilboy *et al.* (1995, voir Chapitre 1, section 2. 1. 1. 2) permet une première estimation de l'effet de l'adjectif et de sa position. Elle confirme en partie l'influence la modifiabilité : en espagnol, il y a moins d'attachements vers N1 quand celui-ci a un adjectif, et plus d'attachement à N2 quand ce dernier en a un. Les données obtenues en anglais, par contre, ne montrent aucun effet de l'adjectif : il y a toujours un attachement préférentiel à N2. Ces résultats contredisent en partie les prédictions selon le critère de modifiabilité. Les résultats de l'Expérience 3 de Colonna (2001), conduite en français, infirment aussi l'effet de la modifiabilité : un antécédent modifié par un adjectif est le site d'attachement préféré.

2. 2. 2. Les variations translinguistiques

La *construal hypothesis* (Frazier & Clifton, 1996, 1997) est une théorie qui apporte une amélioration majeure de la théorie du *garden path*, surtout dans l'explication des préférences d'attachement de propositions relatives ambiguës. Le modèle de satisfaction de contraintes se base sur l'interaction entre traitement syntaxique et traitement pragmatique pour expliquer la variation translinguistique, alors que la *tuning hypothesis* postule que l'exposition du parseur guide les préférences d'attachement. La *recency theory*, elle, montre l'existence de deux principes dont le poids différent peut contribuer aux différences translinguistiques.

2. 2. 2. 1. Le modèle de satisfaction de contraintes

Le modèle de satisfaction de contraintes propose que c'est l'interaction de contraintes syntaxiques spécifiques et de principes pragmatiques généraux qui explique cette variation translinguistique (Thornton *et al.*, 1998 ; Thornton *et al.*, 1999). Cette interaction a lieu immédiatement et non dans la ré-analyse comme le postule la théorie du *garden path*). Les contraintes syntaxiques spécifiques aux langues sont par exemple l'ordre des mots (voir Comrie, 1981).

Reprenant l'argument de l'existence du génitif saxon en anglais émis par Cuetos et Mitchell (1988) afin d'expliquer la différence entre les résultats en anglais et en espagnol, Thornton *et al.* (1998) proposent que le fait qu'une langue ait ou non à disposition plusieurs façons d'exprimer le génitif est une première contrainte, d'ordre structurel.

En anglais, l'utilisation du génitif saxon (*the colonel's daughter*) permet des modifications ultérieures de N2, *daughter*, par exemple à l'aide d'une proposition relative. Il est par contre impossible d'ajouter ultérieurement un modifieur au possesseur, N1 (*colonel*). Avec le génitif prépositionnel (*the daughter of the colonel*), N1 et N2 peuvent être modifiés et c'est d'ailleurs la seule façon en anglais de modifier le possesseur, *colonel* (situé en N2). On devrait observer une fréquence plus importante des attachements vers N2 dans ce dernier cas en anglais. Comme en espagnol une seule façon d'exprimer le génitif est disponible, cela n'entraîne pas une préférence pour N2 aussi forte qu'en anglais, avec une structure « N1 de N2 ».

Afin de tester cette hypothèse, Thornton *et al.* (1998) proposent une tâche de complétion de phrase en anglais. Les participants doivent compléter un début de phrase comme (24a-b). Selon le modèle de satisfaction de contraintes, plus d'attachements vers N2 sont attendus pour (24a), que ce soit avec un groupe prépositionnel (commençant par *with*) ou avec une proposition relative (commençant par *who*). Dans le cas de (24b), seule une structure, « N1 près de N2 » est disponible et on devrait observer une préférence moins prononcée pour N2.

(24) a. *The assistant of the inspector with/who...* (*L'assistant de l'inspecteur avec/qui...*)

b. *The table near the window with/that...* (*La table près de la fenêtre avec/qui...*)

Les résultats obtenus confirment les attentes des auteurs. C'est donc bien la disponibilité des structures, ici l'expression de la possessivité avec le génitif, qui contraint les préférences d'interprétation. Cependant, on peut regretter que la manipulation du type de préposition (*of*, *near*) soit confondue avec la manipulation de l'animation. En effet, tous les noms sont animés pour *of* (24a) et tous sont non-animés pour *near* (24b). Par conséquent, le type de relativiseur est aussi modifié (*who* et *that*). Alors que le modèle de satisfaction de contraintes postule un effet immédiat de contraintes syntaxiques et sémantiques, l'animation est une caractéristique sémantique. On a présenté plus haut l'effet de l'animation de N1 et N2 sur les préférences d'attachement (Chapitre 1, section 2. 1. 1. 3) et on ne peut que regretter qu'animation et prépositions soient des facteurs confondus.

2. 2. 2. 2. La *construal hypothesis*

La *construal hypothesis* considère, tout comme la théorie du *garden path*, que les stratégies du *minimal attachment* et du *late closure* sont universelles, mais elles ne s'appliquent qu'aux structures primaires (Frazier, 1987 ; Frazier & Clifton, 1996, 1997). Cette théorie se pose comme une actualisation de la théorie du *garden path* qui explique de façon plus satisfaisante le traitement des ambiguïtés syntaxiques.

La préférence d'attachement vers N1 qui est observée dans de nombreuses langues, ainsi que la variabilité de ces préférences pour une même langue peuvent remettre en cause l'universalité du *late closure*. Néanmoins, Frazier et Clifton (1996) ne la rejettent pas, mais proposent que la stratégie de *late closure* ne s'applique pas de façon uniforme à toutes les structures syntaxiques, pour une même langue. L'utilisation de cette stratégie dépend du constituant ambigu, selon qu'il s'agit d'une relation primaire ou non : c'est la *construal hypothesis*.

Le principe de *construal* ne s'applique qu'à des structures syntaxiques qui ne peuvent être considérées comme primaires. Les structures primaires sont par exemple le sujet et le prédicat d'une proposition à verbe fléchi, les compléments et les constituants obligatoires. Une proposition relative ou un groupe prépositionnel sont des structures non primaires et leur interprétation suit le principe de *construal*.

Si le constituant d'une phrase peut, tout du moins temporairement, être considéré comme une structure primaire, alors ce constituant est interprété comme tel. Dans (9), *the answer* est initialement interprété comme l'objet.

(9) *Sonia knew the answer to the difficult problem was correct.*

Si un constituant peut être interprété comme impliquant deux relations primaires différentes, comme c'est le cas pour *the answer* qui peut être l'objet du verbe qui le précède ou bien le sujet de la complétive, alors les stratégies de la théorie du *garden path*, *minimal attachment* et *late closure* sont appliquées. En revanche, si un constituant peut être considéré comme deux relations non primaires différentes (cf. attachement ambigu de propositions relatives), le principe de *construal* s'applique.

Selon ce dernier, si une structure syntaxique est non primaire, alors il faut associer cette structure syntaxique au domaine de traitement thématique en cours (*current thematic processing domain*). Par exemple, si le dernier domaine de traitement thématique est une préposition, le domaine thématique en cours de traitement est constitué du groupe

prépositionnel et du nœud syntaxique que ce groupe prépositionnel domine. Si plusieurs sites d'attachement sont possibles, alors la structure syntaxique est attachée en fonction des informations structurelles ou non structurelles.

Dans leur ouvrage, Frazier et Clifton (1996) passent en revue de nombreuses structures ambiguës afin d'évaluer leur théorie, mais on ne présentera ici que la partie concernant les ambiguïtés d'attachement de propositions relatives. Le principe de *construal*, dans ce cas, consiste à « associer une proposition relative au domaine de traitement thématique en cours, à la projection maximale (étendue) du dernier assigneur de théta-rôle » (p. 31, Frazier & Clifton, 1996).

La préposition *avec*, exemple (25), est un assigneur de rôle thématique et le dernier au moment où la proposition relative est traitée. C'est donc aux constituants du groupe prépositionnel, *avec la dame*, que la proposition relative est attachée. Avec de telles phrases, on observe une forte préférence pour attacher la proposition relative à N2, ce qui a été observé en italien (De Vincenzi & Job, 1993, 1995), en français (Baccino *et al.*, 2000) ou en espagnol (Gilboy *et al.*, 1995).

(25) *Le cuisinier observait la cliente avec la dame qui s'était plainte des plats.* (tiré de Baccino *et al.*, 2000)

Par contre, la préposition *de* n'est pas un assigneur de rôle thématique. N2 est un argument de N1 et c'est le GN comportant N1 qui devient le domaine de traitement thématique en cours. Dans ce GN se trouvent N1 et N2 qui sont les deux sites d'attachement possibles de la proposition relative suivante.

Le principe de *construal* guide l'attachement, comme dans (26) où la préposition *de* n'est pas un assigneur de rôle thématique et où la proposition relative peut être attachée aux constituants du nœud dominant, N1 ou N2.

(26) *Le docteur appelait la fille de la dame qui s'était blessée à la main.* (tiré de Baccino *et al.*, 2000)

Lorsque deux sites d'attachement sont possibles, les préférences d'attachement sont influencées par trois facteurs : les relations syntaxiques et sémantiques entre N1 et N2 (voir Chapitre 1, section 2. 1. 2 et Gilboy *et al.*, 1995), le statut référentiel des sites d'attachements possibles (Frazier & Clifton, 1996 ; Gilboy *et al.*, 1995) et l'application éventuelle de maximes conversationnelles (le cas du génitif saxon, 4. 1. 2 et Sperber & Wilson, 1989).

La *construal hypothesis* conserve par ailleurs un aspect central de la théorie du *garden path* : confronté à une ambiguïté syntaxique, le parseur syntaxique choisit une interprétation unique, qui peut être ré-analysée ultérieurement si les constituants suivants infirment cette interprétation (Frazier & Clifton, 1997). Elle propose aussi que l'effet *garden path* varie selon le type de structure, primaire ou non primaire (Frazier & Clifton, 1997). Si le parseur considère qu'un constituant, dont l'interprétation est ambiguë, peut être une structure primaire ou non primaire (cf. *the answer* dans (9)), le constituant sera considéré comme une structure primaire. Dès que l'ambiguïté est levée, il peut être ré-analysé en une structure non primaire. Dans ce cas, l'effet *garden path* est important.

Si un constituant peut être interprété comme deux structures non primaires, le lecteur n'en garde qu'une en appliquant le principe de *construal* et, si une ré-analyse s'avère nécessaire, l'effet *garden path* est moindre. C'est le cas avec l'attachement de propositions relatives ambiguës.

2. 2. 2. 3. La *tuning hypothesis*

Tandis que les autres théories ne considèrent guère que l'aspect compréhension dans le traitement des ambiguïtés d'attachement, la *tuning hypothesis* (Brysbaert & Mitchell, 1996 ; Mitchell, Cuetos, Corley, & Brysbaert, 1995) accorde un rôle important à la production de ces structures ambiguës, aux attachements auxquels le parseur est exposé.

Ainsi, les préférences d'attachement de propositions relatives ambiguës sont liées à la fréquence des attachements pour des structures syntaxiques similaires, reprenant ainsi ce qui avait pu être suggéré par le modèle de satisfaction des contraintes (Thornton *et al.*, 1998).

De nombreuses mesures de fréquence sont possibles (Crocker & Keller, 2006) : lexicales (MacDonald, 1993, 1994), morphologiques (Trueswell, 1996), syntaxiques (Garnsey, Pearlmutter, Myers, & Lotocky, 1997 ; Trueswell, Tanenhaus, & Kello, 1993) ou sémantiques (Garnsey *et al.*, 1997 ; Pickering, Traxler, & Crocker, 2000). Dans une version forte de la *tuning hypothesis* (Mitchell *et al.*, 1995), seule la fréquence d'occurrence des structures syntaxiques est prise en compte, au détriment de la fréquence lexicale (MacDonald, 1994) ou d'autres facteurs linguistiques. Des études de corpus suggèrent cependant une version moins forte de la *tuning hypothesis*, incluant la fréquence des structures syntaxiques. Se pose alors la question de la granularité, de la finesse des caractéristiques dont la fréquence est prise en compte (Crocker & Keller, 2006 ; Desmet, Brysbaert, & De Baecke, 2002 ; Mitchell & Brysbaert, 1998 ; Mitchell *et al.*, 1995).

L'étude de corpus de Hindle et Rooth (1993), par exemple, met en évidence que la mesure de la co-occurrence de certains constituants, ce que les auteurs appellent l'association lexicale, permet de prédire les préférences d'attachement de groupes prépositionnels, comme (27a-b) répétées ci-dessous. On considère ici la fréquence de co-occurrence entre un verbe comme *saw*, un nom (*spy*) et une préposition (*with*). La fréquence mesurée permet à un parseur automatique d'analyser des phrases dont l'attachement du groupe prépositionnel est ambigu avec un taux de succès proche de celui de juges humains.

(27) a. *The spy saw the cop with binoculars but the cop didn't see him.* (L'espion a vu le policier avec des jumelles mais le policier ne l'a pas vu, d'après Trueswell *et al.*, 1994)

b. *The spy saw the cop with a revolver but the cop didn't see him.* (L'espion a vu le policier avec un revolver mais le policier ne l'a pas vu)

Qui plus est, dans le cas des préférences d'attachement de propositions relatives en français, la comparaison entre études en compréhension (Zagar *et al.*, 1997) et en production, *i. e.* analyse de corpus (Baltazart & Kister, 2000), montre des résultats concordants ou discordants, selon le degré de finesse des facteurs considérés dans l'analyse de corpus.

La préférence d'attachement en français, en compréhension (temps de lecture et réponse à des questions de compréhension), est généralement pour N1 (Frenck-Mestre & Pynte, 2000a ; Pynte & Colonna, 2000 ; Zagar *et al.*, 1997), à l'exception de ce qui est obtenu dans l'étude de Baccino *et al.* (2000). En production, on observe une préférence comparable. En effet, une étude de corpus confirme la préférence pour attacher une proposition relative précédée par une structure « N1 de N2 » à N1 (Baltazart & Kister, 2000 ; Kister, 1999) : sur l'ensemble des occurrences, 63 % des attachements sont réalisés vers N1. Si la définitude de N1 et N2 est prise en compte, il s'avère que cette préférence d'attachement peut varier de façon très importante. Par exemple, 100 % des attachements sont réalisés vers N1 quand ce dernier est précédé d'un article démonstratif et N2 par un indéfini (*ce résumé d'une œuvre*). La préférence peut être inverse (15 % des attachements vers N1) quand N1 est précédé par un article défini (*le, la* ou *les*) et N2 d'un indéfini (*le résumé d'une œuvre*).

Les résultats où N1 et N2 sont définis montrent que N1 n'est plus le site d'attachement préféré (40 %), ce qui est contraire à ce qui a pu être mesuré en compréhension avec des structures et constituants comparables. On peut néanmoins considérer qu'un degré de finesse aussi important ne sera pas nécessairement pris en compte par le parseur syntaxique et qu'une fréquence d'occurrence plus globale entre en jeu dans la *tuning hypothesis*. C'est ce que

proposent Mitchell *et al.* (1995) : seules les fréquences d'occurrences de structures et variables ayant des degrés relativement importants de précision, comme ici la fréquence des attachements avec une structure « N1 de N2 » (indépendamment de la définitude des articles) interviennent.

Si la fréquence d'occurrence de telles structures ou l'exposition influencent les préférences d'attachement dans d'une tâche de compréhension, on pourrait observer un effet comparable lors de situations de bilinguisme (Frenck-Mestre & Pynte, 1997) ou d'amorçage (Desmet & Declercq, 2006).

Par exemple, la fréquence des attachements dans la langue maternelle peut avoir un effet sur les préférences d'attachement dans la langue seconde, pour peu que les préférences d'attachement diffèrent d'une langue à l'autre. Un effet inverse peut aussi être possible : les préférences d'attachement dans la langue seconde influencent les préférences dans la langue maternelle. C'est ce qui est observé dans une expérience en *eye tracking* (Dussias & Sagarra, 2007), où des locuteurs natifs de l'espagnol (monolingues ou ayant une exposition limitée à l'anglais) montrent une préférence pour N1 (voir aussi Dussias, 2001). Les temps de lecture totaux sont alors plus courts quand les phrases sont désambiguïsées vers N1 par rapport à N2. Les locuteurs espagnols natifs ayant eu une longue exposition à l'anglais (vivant dans un environnement anglophone lors de l'expérience) ont un comportement inverse, avec une préférence pour N2. Les préférences d'attachement en langue seconde, l'anglais (N2), ont ainsi une influence sur le traitement de phrases dans la langue maternelle, l'espagnol (N1).

Qui plus est, l'étude de Desmet *et al.* (2006) met en évidence que, contrairement à ce que propose la *tuning hypothesis* qui considère que seule une fréquence syntaxique (et non lexicale) doit être considérée, une dimension sémantique au moins doit être prise en compte : l'animation. Ce n'est que dans ce cas que les mesures comportementales (*eye tracking*) correspondent aux fréquences observées dans des corpus. Mitchell *et al.* (1995) considèrent que les facteurs pris en compte dans l'analyse des fréquences d'occurrence ne doivent pas avoir un grain trop fin. Dans le cas des préférences d'attachement de propositions relatives, si considérer uniquement la structure « N1 de N2 » est trop flou, la trop grande précision qui consisterait à considérer la définitude, par exemple, est à éviter.

Si cette approche est intéressante, la *tuning hypothesis* ne peut cependant rendre compte ni des résultats obtenus lors d'expériences en compréhension en anglais, ni de la variabilité des résultats observés (Frazier & Clifton, 1996 ; Frenck-Mestre, 2004). De plus, la *tuning hypothesis* n'explique pas les différences de préférences d'attachement en fonction de la

préposition entre N1 et N2 (Traxler *et al.*, 1998). L'étude d'un corpus anglais montre une préférence pour N2 avec 62 % des attachements réalisés vers ce nom (Mitchell, Cuetos, & Corley, 1992, cité par Gibson, Schütze, & Salomon, 1996). Une préférence comparable devrait être observée avec des mesures comportementales. Or, si des études confirment cette attente (Carreiras & Clifton, 1999 ; Cuetos & Mitchell, 1988 ; Mitchell *et al.*, 1990), d'autres l'infirmement (Carreiras & Clifton, 1993).

De plus, l'étude de Gibson, Schütze *et al.* (1996) remet en question la validité de la *tuning hypothesis*. En effet, cette étude considère, à la fois en production (analyse de corpus) et en compréhension (questionnaire), les préférences d'attachement de propositions relatives où trois sites sont possibles (voir (28) ci-dessous). Pour les différents sites d'attachement possibles, les fréquences mesurées dans l'analyse de corpus diffèrent de celles mises en évidence dans des questionnaires de compréhension, invalidant la *tuning hypothesis*.

2. 2. 2. 4. La recency theory

L'approche envisagée par la *recency theory* (Gibson, Pearlmutter, Canseco-Gonzalez, & Hickok, 1996 ; Gibson *et al.*, 1999) ressemble à la *construal hypothesis* dans le sens où deux stratégies interprétatives sont disponibles. Selon les langues, des contraintes peuvent accorder une importance plus grande à l'une ou l'autre des stratégies, ce qui explique les variations des préférences d'attachement de propositions relatives ambiguës. La *recency theory* est un principe général qui permet d'expliquer les préférences d'attachement de propositions relatives (Gibson, Pearlmutter *et al.*, 1996), mais aussi l'interprétation d'autres structures syntaxiquement ambiguës comme l'attachement de groupes prépositionnels (Schütze & Gibson, 1999) ou la recherche de l'antécédent avec des propositions relatives distantes (Sturt, Scheepers, & Pickering, 2002).

Les deux stratégies proposées sont la préférence de récence (*recency preference*) et la proximité du prédicat (*predicate proximity*). La préférence de récence est présentée comme une variante de la stratégie du *late closure*, qui consiste à « attacher de préférence les structures d'items lexicaux en cours de traitement aux structures récemment construites »¹¹. Cette préférence favorise les attachements vers N2 pour une structure « N1 de N2 ». La

¹¹ « Preferentially attach structures for incoming lexical items to structures built more recently » (p. 26, Gibson, Pearlmutter *et al.*, 1996).

proximité du prédicat, quant à elle, favorise les attachements vers N1, car ce facteur consiste à « attacher le plus près possible de la tête d'un groupe prédicatif »¹².

La différence translinguistique des préférences d'attachement de propositions relatives, mise en évidence par l'étude de Cuetos et Mitchell (1988) en espagnol et en anglais s'expliquerait par le poids différent de la préférence de récence et de la proximité du prédicat. Les poids respectifs sont fonction par exemple de la rigidité de l'ordre des mots. L'anglais est une langue où l'ordre des mots est plus conservateur qu'en espagnol. La distance (cf. proximité du prédicat) entre un élément à intégrer, le relativiseur, et un appartenant à la structure syntaxique déjà élaborée, l'antécédent, n'est guère informative, alors que la variation dans l'ordre des mots en espagnol l'est plus. C'est pourquoi la proximité du prédicat a un poids plus important en espagnol alors qu'il s'agit de la préférence de récence en anglais. Ainsi, la proximité du prédicat conduit à plus d'attachements à N1. En anglais, c'est la préférence de récence qui domine et induit l'attachement préférentiel de la proposition relative à N2.

L'existence de ces deux facteurs a été testée au moyen d'expériences (1996) présentant des structures où trois sites d'attachement de propositions relatives sont possibles (28). L'ambiguïté est levée au niveau du verbe au singulier, forçant l'attachement à l'unique antécédent singulier, N3 dans l'exemple présenté ici.

(28) *The lamps near the paintings of the house that was damaged in the flood. (Les lampes_{N1} près des peintures_{N2} de la maison_{N3} qui a été endommagée lors de l'inondation)*

Des phrases de ce type sont présentées en espagnol et en anglais auprès de locuteurs natifs de ces langues. Le paradigme expérimental utilisé est la technique d'auto-présentation segmentée. La tâche des participants consiste à juger pour chaque nouveau mot si la phrase est grammaticale (*i. e.* correcte). C'est au niveau de l'auxiliaire *was* que l'ambiguïté est levée, car un seul des trois sites possibles d'attachement est au singulier ; dans (28) il s'agit de N3, *house*. Ce sont les phrases où l'attachement est vers N3 qui sont jugées les plus correctes. Viennent ensuite celles où il se fait vers N1, et enfin vers N2. Les temps de lecture mesurés pour chaque mot confirment cette configuration de résultats, en anglais et en espagnol.

Que les préférences d'attachement soient similaires dans ces deux langues contraste avec les résultats obtenus précédemment par Cuetos et Mitchell (1988) lorsque deux sites d'attachement sont possibles. La hiérarchie de préférences observée avec trois sites

¹² « Attach as close as possible to the head of a predicate phrase » (*ibid.* p. 41).

d'attachement est interprétée comme la mise en œuvre de la préférence de récence (préférence pour un attachement vers N3) qui prime sur la proximité du prédicat (préférence pour un attachement vers N1). L'attachement vers N2 n'étant influencé par aucun facteur, ce type d'attachement est le moins souvent réalisé.

Enfin, la différence de préférence qu'on peut observer à travers les langues est expliquée par le fait que l'importance des deux facteurs, préférence de récence et proximité du prédicat, varie d'une langue à l'autre. En anglais, l'ordre des mots est plutôt conservateur, sujet, verbe, objet. Ainsi, sujet et objet restent proches du prédicat (le verbe) et c'est pour cette raison qu'en anglais la préférence de récence prend le pas. En espagnol comme en allemand ou en néerlandais, l'ordre des mots est moins strict. Il est possible d'avoir l'ordre verbe, objet puis sujet, ce qui augmente la distance entre le sujet et le prédicat. Ici, la proximité du prédicat est le facteur le plus important Hemforth *et al.* (2000).

3. L'attachement des relatives comme résolution anaphorique (Hemforth *et al.*, 2000)

L'hypothèse qui va être maintenant présentée n'a pas pour but d'expliquer ce qui diffère d'une langue à l'autre, mais plutôt d'expliquer ce qui singularise les ambiguïtés d'attachement des propositions relatives par rapport aux autres ambiguïtés syntaxiques (Hemforth *et al.*, 2000). Pour ces dernières on retrouve des préférences similaires dans toutes les langues testées, alors qu'il y a de nombreuses variations pour l'attachement des relatives ambiguës. La spécificité de cette dernière réside dans le traitement anaphorique qui a lieu lors de l'intégration et l'attachement du relativiseur.

La *construal hypothesis*, on l'a vu dans ce Chapitre 1 en 2. 2. 2, suggère que des stratégies syntaxiques sont à l'œuvre lors du traitement d'ambiguïtés syntaxiques, mais que d'autres (p. ex. pragmatique) peuvent aussi intervenir. C'est aussi ce que postulent Hemforth *et al.* (2000) : si la stratégie de *late closure* prédit de façon correcte les préférences d'attachement dans de nombreuses structures, mais pas dans le cas d'ambiguïtés d'attachement de propositions relatives, c'est que, contrairement aux premières, ce n'est pas qu'un traitement syntaxique qui est réalisé. Il y a en plus un traitement anaphorique. Ainsi, attacher une proposition relative à son antécédent relève de deux types de processus :

- un attachement syntaxique au marqueur phrastique en cours de construction ;
- la résolution du relativiseur qui est lié anaphoriquement à une entité du discours.

Dans l'interprétation de propositions relatives ambiguës, la mise en place d'un traitement anaphorique, absent dans l'attachement d'un groupe prépositionnel, contribue aux différences qu'on a pu mesurer avec ces deux types de structures. Les auteurs ont testé cette hypothèse au moyen d'une série de questionnaires dans lesquels les participants doivent estimer l'attachement de différentes structures syntaxiques.

Pour le premier questionnaire, les items expérimentaux contiennent soit une proposition relative (29a), soit un groupe prépositionnel (29b). Le génitif est marqué grammaticalement avec le cas : dans (29a-b), *der* est l'article féminin génitif.

(29) a. *Die Tochter der Lehrerin, die aus Deutschland kam, traf John.* (La fille de l'enseignante, qui est originaire d'Allemagne, a rencontré John, tiré de Hemforth *et al.*, 2000)

b. *Die Tochter der Lehrerin aus Deutschland traf John.* (La fille de l'enseignante originaire d'Allemagne a rencontré John, *idem*)

Si l'attachement est ambigu dans les deux constructions, les participants montrent une préférence pour considérer N1 comme l'antécédent dans (29a), mais la préférence est pour N2 avec (29b). Les préférences d'attachement sont par conséquent différentes et relèvent *a priori* de mécanismes interprétatifs différents. Pour les propositions relatives, on retrouve la préférence pour attacher ces dernières à N1 (Hemforth *et al.*, 2000).

Les deuxième (estimation de l'attachement) et troisième (jugement d'acceptabilité) questionnaires testent la même hypothèse, mais N2 est ici un nom propre. L'utilisation de la préposition *von* (*de*) est alors obligatoire (*die Tochter von Mary, la fille de Mary*). De plus, N2 étant un nom propre ne peut accepter un modifieur. Cela renforce l'attachement vers N1 dans le cas d'une proposition relative comme dans (29a). Ces résultats sont similaires à ceux obtenus en français (Frenck-Mestre & Pynte, 2000a) ou en italien (Frenck-Mestre & Pynte, 2000b). Dans le cas d'un groupe prépositionnel (29b), on ne retrouve plus la préférence pour N2 obtenue dans le premier questionnaire.

Les auteurs réalisent un quatrième questionnaire où sont présentées de façon auditive les phrases du premier questionnaire, afin de neutraliser l'influence éventuelle de la ponctuation, obligatoire avec les propositions relatives en allemand, et qui entraîne une rupture prosodique (Schafer *et al.*, 1996). Cette dernière est neutralisée expérimentalement et les résultats obtenus sont comparables avec les précédents : les préférences d'attachement diffèrent pour les propositions relatives et les groupes prépositionnels. Ainsi, la différence observée dans les préférences d'attachement n'est causée ni par la ponctuation ni par la prosodie différente.

Afin de s'assurer que, dans le cas des propositions relatives, la différence d'attachement est bien due à la mise en place d'un traitement anaphorique, les auteurs réalisent un cinquième questionnaire où ils comparent des phrases avec une proposition relative (30a), un groupe prépositionnel (30b), ou une proposition adverbiale qui est soit adjacente au groupe nominal *der Student des Professors* (30c) soit distante (30d). Si c'est la présence d'un traitement anaphorique qui conduit à une préférence pour N1 avec les propositions relatives, on devrait alors observer une préférence similaire pour les phrases où un traitement anaphorique est réalisé. C'est le cas dans les phrases contenant une proposition relative (30a), mais aussi dans celles contenant une proposition adverbiale, (30c) et (30d), où un lien anaphorique entre l'antécédent (*Student* ou *Professors*) et le pronom *er* doit être réalisé. Aucun traitement anaphorique n'est réalisé dans (30b) et les préférences d'attachement devraient différer de celles qui sont réalisées dans les autres conditions.

(30) a. *Der Student des Professors, der in dem neuen Labor war, las die Ergebnisse.*
(L'étudiant du professeur, qui était dans le nouveau labo, a lu les résultats)

b. *Der Student des Professors in dem neuen Labor las die Ergebnisse.* (L'étudiant du professeur dans le nouveau labo a lu les résultats)

c. *Die Ergebnisse las der Student des Professors, als er in dem neuen Labor war.*
(L'étudiant du professeur a lu les résultats quand il était dans le nouveau labo ; l'ordre des mots est non canonique)

d. *Der Student des Professors las die Ergebnisse, als er in dem neuen Labor war.*
(traduction identique à (30c), mais l'ordre des mots ici est canonique)

La distance entre la proposition adverbiale et les antécédents possibles varie entre (30c) et (30d) avec une manipulation de l'ordre canonique. En allemand, le verbe d'une proposition principale (*las*) se trouve en deuxième position. Mis à part cette contrainte, l'ordre des autres constituants est libre, même si la succession sujet-verbe-objet est préférée. Dans (30d), l'ordre des constituants est canonique, avec *der Student des Professors* comme sujet, le verbe *las* et l'objet *die Ergebnisse*. Il ne l'est plus dans (30c), car l'objet est en tête de phrase, suivi du verbe et du sujet. Dans ce cas-là, la distance entre les sites d'attachement potentiels, *Student* et *Professors*, est minimale, alors qu'elle est plus importante dans (30d). La manipulation de l'ordre canonique permet ainsi d'évaluer l'effet de la distance entre l'antécédent et le marqueur anaphorique.

Les résultats confirment l'effet du traitement anaphorique sur les préférences d'attachement, car on observe une préférence pour N1 dans les phrases avec des propositions relatives et des propositions adverbiales, mais pas pour celles contenant des groupes prépositionnels.

Considérant qu'un traitement anaphorique a lieu lors de l'attachement de propositions relatives, les auteurs proposent que la proéminence des différentes entités du discours influence les préférences d'attachement. Les entités les plus proéminentes sont plus activées et par conséquent plus accessibles. Dès que plusieurs entités sont présentes, une interférence réciproque prend place et l'activation des entités diminue. Dans le cas des ambiguïtés d'attachement où deux sites sont possibles, cette interférence n'est pas très importante et l'activation de la première entité, N1, est suffisante pour que l'attachement y soit réalisé préférentiellement. Si trois sites d'attachement sont possibles (Gibson, Pearlmutter *et al.*, 1996 ; Hemforth *et al.*, 2000), l'activation des premières entités diminue et l'attachement ne se fait plus sur la base de l'activation des entités. Le processus syntaxique (*late closure*) prend donc le pas et c'est une préférence pour N3 qui est observée (voir la *recency theory*).

Notons cependant que ces résultats diffèrent de ceux obtenus par Traxler *et al.* (1998). En effet, dans une expérience en *eye tracking* où ils présentent des phrases similaires à (29a), en anglais, on a vu que les participants préfèrent attacher la proposition relative à N2. Quant à l'attachement du groupe prépositionnel, la préférence pour N2 est répliquée en anglais.

L'ensemble des questionnaires présentés par Hemforth *et al.* (2000) confirme que l'attachement de propositions relatives ambiguës diffère des autres types d'ambiguïtés d'attachement. Cette différence est liée à la présence d'un traitement anaphorique. Un relativiseur se comporte comme n'importe quel autre marqueur anaphorique et les mécanismes qui régissent l'attachement du relativiseur à son antécédent sont similaires à ceux mis en jeu lors du traitement des autres expressions anaphoriques.

Afin de mieux comprendre ce que cette proposition implique, il est nécessaire de passer en revue les travaux réalisés sur le traitement anaphorique. On verra que si les anaphores qui sont privilégiées dans ces études sont les GN et surtout les pronoms, il est intéressant d'élargir le champ de ces études au relativiseur, expression anaphorique rarement évoquée.

Il ressort de la littérature sur les anaphores que ces dernières sont hiérarchisées selon la proéminence dans le discours de l'entité à laquelle les anaphores font référence. Cette hiérarchie est signalée au niveau des expressions anaphoriques par leur forme, comme par exemple le caractère défini ou non.

Ensuite, je montrerai que les relativiseurs peuvent être placés sur la hiérarchie des expressions anaphoriques et que, par conséquent, ils signalent eux aussi le degré de prééminence de l'entité à laquelle ils font référence. Je propose aussi que la forme du relativiseur guide ainsi l'attachement, ce que je testerai plus tard dans la partie expérimentale.

Chapitre 2 Le traitement des anaphores

Dans le discours, la référence¹³ successive à une entité est réalisée selon des critères qui permettent de maintenir la cohérence ; celle-ci dépend à la fois de facteurs syntaxiques et sémantiques. La cohérence peut être maintenue grâce à la répétition et à la reprise des informations antérieures : c'est l'anaphore (Garrod & Sanford, 1990 ; Riegel *et al.*, 2004 ; van den Broek, 1994). Dès l'apparition d'une expression référentielle comme un groupe nominal ou un pronom, des mécanismes cognitifs sont activés afin d'établir la cohérence entre les différentes expressions référentielles du discours ou afin de les intégrer dans un modèle du discours (Sanford & Garrod, 1989 ; Tyler & Marslen-Wilson, 1982).

On a vu qu'un relativiseur se comporte comme un marqueur anaphorique (Hemforth *et al.*, 2000). Cette proposition, tout particulièrement intéressante pour expliquer les préférences d'attachement des relatives, n'a cependant jamais été abordée par la littérature sur le traitement des anaphores. Pourtant, aborder le traitement des anaphores à travers celui des propositions relatives et considérer en plus le traitement des propositions relatives en prenant en compte la dimension anaphorique du relativiseur permettrait de mieux expliquer certains points critiques de ces théories.

Comme les théories concernant le traitement des anaphores n'ont cependant pas intégré les propositions relatives dans leur objet d'étude, je reviendrai de temps à autre sur les implications de ces théories dans le traitement des propositions relatives, notamment lorsque leur attachement est ambigu.

L'étude du traitement des anaphores peut se faire de différentes façons. En intelligence artificielle, le but est d'établir des heuristiques afin de lier les différentes mentions d'un même référent (Winograd, 1972, cité par Sidner, 1981). De nombreux chercheurs en psycholinguistique se sont intéressés au traitement des anaphores (Singer, 1994, pour une revue Poesio, Stevenson, Di Eugenio, & Hitzeman, 2004, mais aussi Ferstl & von Cramon, 2001 pour une étude en IRMf), notamment à l'effet des contraintes syntaxiques et

¹³ Note sur le vocabulaire utilisé : plutôt que de parler d'antécédent pour les mentions antérieures du référent, on préférera les termes d'expression référentielle ou de référence (voir Apothéloz, 1995 ; Reboul, 1995). De plus, on appelle référent l'entité représentée dans le modèle mental du discours à laquelle il est fait référence.

sémantiques (Garnham & Oakhill, 1985 ; Garrod, Freudenthal, & Boyle, 1994 ; Gordon & Chan, 1995 ; Nicol & Swinney, 1989 ; Sturt, 2003), à l'élaboration d'inférences (van den Broek, 1994)¹⁴ afin d'établir la cohérence et créer des liens entre les éléments du discours (Klin, Guzmán, Weingartner, & Ralano, 2006 ; Levine, Guzmán, & Klin, 2000 ; Schmalhofer, McDaniel, & Keefe, 2002 ; Virtue, Haberman, Clancy, Parrish, & Beeman, 2006) ou bien aux relations entre les différentes expressions référentielles du discours (Ariel, 2001 ; Gordon, Grosz, & Gilliom, 1993 ; Grosz, Joshi, & Weinstein, 1983 ; Gundel *et al.*, 1993). C'est ce dernier aspect qui sera considéré ici.

On a vu dans l'étude de Hemforth *et al.* (2000) qu'un relativiseur est une anaphore, puisqu'il renvoie à un référent déjà présent dans le discours. Dans (31), le relativiseur *que* et le GN *le reporter* sont coréférents, dans le sens où ils font référence à la même entité, *reporter*. La relativisation est un type particulier de coréférence.

(31) *Le reporter que le sénateur attaque admet l'erreur.*

Dans ce chapitre, seules les anaphores coréférentielles seront présentées, au détriment des anaphores indirectes (Cornish, Garnham, Cowles, Fossard, & André, 2005) qui reposent plus sur le contexte que sur les références aux différentes entités du discours. Il existe plusieurs types d'anaphores coréférentielles (Garnham, Oakhill, & Cain, 1997 ; Levinson, 1987a ; Strube & Hahn, 1999) : pronominales, nominales, adverbiales, verbales et adjectivales. Les anaphores pronominales sont les plus répandues. En effet, l'emploi d'un pronom permet d'éviter la répétition du groupe nominal auquel il est fait référence (32).

(32) *Paul est sorti. Il avait trop chaud.* (tiré de Kleiber, 1994)

Les anaphores pronominales ont fait l'objet d'une littérature importante utilisant des supports variés : exemples construits (Sidner, 1979), analyses de corpus (Hobbs, 1978), productions élicitées (Stevenson, 2002), expériences en lecture (Ehrlich, 1980, 1981) ou même avec le paradigme de *visual world* (exploration visuelle, J. E. Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000 ; Karabanov, König, & Bosch, 2006). Si cette variété empêche parfois de comparer directement les résultats obtenus dans les études qui vont être présentées, on verra que ces approches peuvent être néanmoins complémentaires et sont

¹⁴ Voir aussi Frank, Koppen, Noordman et Vonk (2003) pour une modélisation de l'élaboration d'inférences dans le texte à l'aide de cartes auto-organisatrices.

nécessaires, notamment si on s'interroge aux deux versants du langage, production et compréhension.

Les anaphores adverbiales consistent à utiliser un adverbe qui réfère à une entité qui peut par exemple être un lieu (*là*) ou même une fraction d'un texte (*pareillement*). Les anaphores verbales comportent un verbe comme *faire* (*Le soleil se couche ; je vous conseille d'en faire autant*, tiré de Riegel *et al.*, 2004). Enfin, les anaphores sont adjectivales quand elles contiennent l'adjectif *tel*, afin de faire référence à la proposition antérieure.

Les anaphores nominales sont des GN contenant un article qui peut être indéfini, *a* dans (33), défini, *the* dans (33) ou autre.

(33) *They made a film of his latest novel; but after reading the book, I found the film a great disappointment. (Ils ont fait un film à partir de son dernier roman, mais après avoir lu le livre, j'ai trouvé le film très décevant, tiré de Leech, 1981)*

Dans la phrase présentée en (33), la succession des deux mentions de l'entité *film* est intéressante. La première mention est réalisée avec un article indéfini, la seconde avec un article défini. La cohérence de ce discours est établie par la succession et la forme des GN, avec un article indéfini puis défini. En effet, la forme de l'anaphore signale l'accessibilité référentielle, c'est-à-dire l'accessibilité de l'entité à laquelle il est fait référence (Givón, 1990). Dans l'exemple (34), l'emploi d'un pronom (*he*) dans (34iii) pour faire référence à l'entité *man* introduite au début du paragraphe n'est pas approprié, car la référence antérieure à l'entité *man* est trop éloignée. Une anaphore nominale comme *the man*, par contre, l'est.

(34) i. *The man finished and left to check his fish-traps.*
 ii. *The fire died out slowly. It was a cold night, the stars were out in force, no wind though, real quiet. Except for one lonely coyote.*
 iii. *After a while the man/he reappeared.*

(L'homme termina et partit contrôler ses pièges à poisson. Le feu s'éteignait lentement. La nuit était fraîche et les étoiles brillaient intensément, sans vent, tout était calme. À l'exception d'un coyote solitaire. Au bout d'un moment l'homme/il reparu, tiré de Givón, 1990)

En revanche, dans un énoncé comme (35), une anaphore nominale (*the man*) est inattendue ; par contre, une anaphore zéro (Ø) l'est.

(35) *The man finished and Ø/the man left. (L'homme termina et Ø/l'homme partit, *ibid.*)*

L'anaphore nominale est appropriée dans (34) et non dans (35) pour une raison d'accessibilité. Dans (34), la distance entre les première et seconde mentions de l'entité *man* est importante et de nombreuses entités sont successivement introduites. Lorsque l'entité *man* est reprise, la distance par rapport à la première mention est trop importante : l'accessibilité de cette entité a diminué. Il est alors nécessaire d'employer une anaphore nominale. Dans (35), la distance entre les deux mentions de l'entité *man* est suffisamment faible pour qu'une anaphore nominale soit inutile, voire inappropriée.

La notion d'accessibilité a fait l'objet de nombreux travaux sur le traitement des anaphores. Vont être présentées maintenant la théorie de Givón (1992), des théories de l'accessibilité, la *givenness hierarchy* (Gundel *et al.*, 1993 ; Gundel, Hedberg, & Zacharski, 2000), l'*accessibility hierarchy* (Ariel, 1990, 1996, 2001) et aussi l'*informational load hypothesis* (Almor, 1999). Enfin, je m'intéresserai à la *centering theory* (Gordon & Chan, 1995 ; Gordon *et al.*, 1993 ; Grosz *et al.*, 1983 ; Grosz, Joshi, & Weinstein, 1995 ; Grosz & Sidner, 1997).

1. Givón

La cohérence est caractérisée par la répétition, la récurrence de certains éléments dans le discours (Garrod & Sanford, 1990, 1994 ; Givón, 1983, 1990 ; Hobbs, 1979 ; Kintsch & van Dijk, 1978 ; Reichman, 1978 ; van den Broek, 1994). Dans le cas de la cohérence référentielle, les éléments récurrents sont les anaphores. Les corrélats de cette récurrence sont l'accessibilité mentale (*mental accessibility*) et l'effort mental (*mental effort*, Givón, 1992). L'information est soit nouvelle, non partagée par le locuteur et l'allocutaire, soit ancienne et dans ce cas partagée (Givón, 1990 ; Haviland & Clark, 1974).

1. 1. L'accessibilité référentielle

L'accessibilité référentielle des entités du discours dépend de trois sources : le contexte déictique partagé (dont dépend l'interprétation de pronoms personnels déictiques ou démonstratifs comme *cela*), la connaissance culturelle partagée (*le soleil* désigne un référent unique pour les membres de la même culture) et le contexte textuellement partagé (*i. e.* le discours précédent). En ce qui concerne le discours précédent, plusieurs facteurs doivent être pris en compte (Givón, 1982, 1983, 1990, 1992) :

- la continuité référentielle, entre la mention actuelle d'un référent et sa mention précédente (Johnson-Laird, 1980 ; Johnson-Laird & Garnham, 1980) ;
- la complexité référentielle, c'est-à-dire la compétition entre plusieurs référents, sémantiquement proches, dans le discours ;
- l'information sémantique de la proposition en cours de traitement ;
- l'information thématique du discours (p. ex. connaissances générales).

1. 2. La continuité référentielle et la hiérarchie des expressions anaphoriques

La continuité référentielle peut être mesurée de plusieurs façons : la distance référentielle (*i. e.* le nombre de propositions entre deux mentions d'un référent), le changement de référent (si deux propositions successives ont ou non le même référent comme argument) et l'interférence potentielle (*i. e.* le nombre de référents sémantiquement compatibles dans les dernières propositions). La distance référentielle induit une hiérarchie des expressions anaphoriques (Givón, 1992). À l'extrémité la plus haute se trouvent les anaphores zéros, pour lesquelles la distance avec l'anaphore précédente est la plus faible, cf. (35). À l'autre extrémité se trouvent les groupes nominaux définis répétés, comme *the man* dans (34), qui sont utilisés quand la distance avec l'anaphore précédente est la plus importante. Entre ces deux extrêmes se situent les pronoms, accentués ou non, les noms définis ou ceux qui ont un modifieur. Cette hiérarchie ressemble beaucoup aux échelles proposées par Levinson (1987b) ou par les théories sur l'accessibilité (pour une représentation de ces hiérarchies, voir Chapitre 2, sections 2. 1 et 2. 2, Ariel, 1990 ; Gundel *et al.*, 1993).

La taille morphologique des expressions référentielles induit par ailleurs une hiérarchie (Givón, 1983) du caractère marqué de ces dernières (*markedness*). En effet, les expressions référentielles comme les groupes nominaux fournissent plus d'informations et sont plus marquées que des expressions référentielles courtes.

Il ressort de la hiérarchie proposée que la forme des expressions référentielles et surtout leur taille permet d'indiquer les niveaux d'accessibilité (Apothéloz, 1995 ; Givón, 1983, 1992). Plus l'anaphore est phonologiquement courte (anaphores zéros, pronoms non accentués), plus la distance avec la mention précédente du référent est courte, c'est-à-dire que ce référent est hautement accessible. L'inverse est par ailleurs vrai : un groupe nominal avec

un article défini est phonologiquement plus long et est utilisé quand l'entité référée n'est plus accessible (34), par rapport à un pronom.

1. 3. Mise en évidence expérimentale

Fletcher (1984) met en évidence l'existence d'une telle hiérarchie en réalisant des expériences à la fois en production et en compréhension. Dans une tâche de production, les participants lisent deux phrases constituées d'une proposition chacune et faisant référence toutes les deux à une même entité. Les participants doivent ensuite reformuler ces phrases en une seule. Si la deuxième phrase est toujours identique, la continuité référentielle est manipulée dans la première (l'entité est ou non sujet de la première proposition, une seconde entité est introduite). L'auteur analyse la forme des reprises anaphoriques, un nom, un pronom ou une anaphore zéro. Les phrases produites sont proches de celles présentées dans l'expérience de compréhension, où la seconde référence à l'entité *Pete* est réalisée sous la forme du zéro (36a), d'un pronom (36b), d'un pronom accentué (36c) qui ne pouvait être produit à l'écrit dans la tâche de production, et enfin un groupe nominal (36d).

(36) a. *Pete had intended to go bowling with Sam last night but \emptyset broke his leg. (Pete avait l'intention d'aller au bowling avec Sam hier soir, mais s'est cassé la jambe)*

b. *Pete had intended to go bowling with Sam last night but he broke his leg. (Pete avait l'intention d'aller au bowling avec Sam hier soir, mais il s'est cassé la jambe)*

c. *Pete had intended to go bowling with Sam last night but HE broke his leg. (Pete avait l'intention d'aller au bowling avec Sam hier soir, mais il s'est cassé la jambe)*

d. *Pete had intended to go bowling with Sam last night but the guy broke his leg. (Pete avait l'intention d'aller au bowling avec Sam hier soir, mais le gars s'est cassé la jambe)*

Les résultats des deux expériences mettent en évidence que l'accessibilité du référent a un effet sur la forme de l'expression référentielle employée. L'anaphore zéro, très peu marquée, est utilisée pour faire référence à une entité très accessible, alors qu'un groupe nominal fait référence à une entité moins accessible, confirmant ainsi la hiérarchie proposée par Givón (1983, 1990, 1992).

Les travaux de Givón sont intéressants et la hiérarchie qu'il propose peut être retrouvée dans les modèles qui suivent. Comme le modèle de Givón, les théories de l'accessibilité, surtout la *givenness hierarchy* (Gundel *et al.*, 1993, 2000) sont basées sur la distinction *given-*

new (Clark, 1975 ; Haviland & Clark, 1974), mais la complexité des hiérarchies présentées est quelque peu augmentée.

2. Les théories de l'accessibilité

Haviland et Clark (Clark, 1975 ; Haviland & Clark, 1974) proposent la stratégie du *given-new* (donné-nouveau). Cette stratégie consiste pour le locuteur, à l'oral ou à l'écrit, à délivrer deux types d'informations qui sont soit données, celles que le locuteur estime que l'allocataire possède, soit nouvelles, celles que l'allocataire est supposé ne pas avoir. Ce dernier, de son côté, considère que le locuteur est coopératif et estime correctement les informations qu'il délivre, données ou nouvelles.

L'introduction de l'entité *film* dans le discours nécessite d'utiliser un article indéfini, cf. *a* dans (33). A partir de ce moment, le locuteur et l'allocataire considèrent cette entité comme donnée. Quand le locuteur réalise une deuxième mention à cette entité, c'est un article défini (*the*) ou bien un pronom qui sont attendus, car ils caractérisent une référence à une entité donnée dans le discours.

(33) *They made a film of his latest novel; but after reading the book, I found the film a great disappointment.*

À l'aide d'expériences en lecture, les auteurs confirment par exemple qu'il est plus facile de traiter un groupe nominal avec un article défini quand l'entité à laquelle il est fait référence est donnée. La *givenness hierarchy* proposée par Gundel *et al.* (1993, 2000) est en partie basée sur cette distinction *given-new*.

2. 1. La *givenness hierarchy*

La *givenness hierarchy* (hiérarchie de donation, dans l'article en français de Gundel *et al.*, 2000) se propose d'expliquer la forme des expressions référentielles en fonction du statut cognitif de leur référent. Cette hiérarchie repose sur la distinction *given-new* d'Haviland et Clark (Clark, 1975 ; Haviland & Clark, 1974 ; voir aussi Chafe, 1974 et Yule, 1979, 1981) : les référents présents en mémoire sont donnés et ceux pour lesquels aucune représentation en mémoire n'existe sont nécessairement nouveaux.

Selon cette hiérarchie (Gundel *et al.*, 1993), il y a six statuts différents associés à des expressions référentielles spécifiques (ici, en anglais). Le statut est :

- a. dans le focus, l'entité est représentée en mémoire et l'attention est centrée sur cette entité : un pronom est le plus souvent utilisé (*it*) ;
- b. activé, c'est-à-dire que l'entité est représentée en mémoire ; les pronoms et déterminants démonstratifs sont caractéristiques de ce statut (*that, this, this N*) ;
- c. familier, où l'entité est déjà représentée en mémoire, un déterminant démonstratif distant est alors préféré (*that N*) ;
- d. identifiable de façon unique, où une représentation unique de l'entité est réalisée lors du traitement de ce groupe nominal (*the N*) ;
- e. référentiel, qui, contrairement au niveau précédent, considère que la représentation cognitive de l'entité est élaborée au cours du traitement de la phrase (*this N indéfini (sic)*) ;
- f. type identifiable, qui permet d'identifier l'entité (*a N*).

Cette hiérarchie est illustrée dans l'exemple suivant (37), où les différentes phrases (37a-f) représentent les différents statuts cognitifs exposés ci-dessus :

- (37) *I couldn't sleep last night. (Je n'ai pas pu dormir hier soir)*
- a. *It kept me awake. (Ça m'a gardé éveillé)*
 - b. *This train/this/that kept me awake. (Ce train/ceci/cela m'a gardé éveillé)*
 - c. *That train kept me awake. (Ce train m'a gardé éveillé)*
 - d. *The train kept me awake. (Le train m'a gardé éveillé)*
 - e. *This train kept me awake. (Ce train m'a gardé éveillé)*
 - f. *A train kept me awake. (Un train m'a gardé éveillé)*

Chaque expression référentielle signale un statut cognitif donné, mais elle contient aussi tous les attributs des statuts situés plus bas dans la hiérarchie. Par exemple, une entité dans le focus est par ailleurs aussi activée, familière, etc. La hiérarchie ne fonctionne que dans un sens, puisqu'une entité familière n'est pas nécessairement dans le focus. Une fois qu'une entité a revêtu le statut cognitif d'identifiable, elle le demeure tout au long du discours (Lambrecht, 1994), mais elle peut aussi avoir un statut plus élevé dans la hiérarchie.

Si en anglais chacun des niveaux de la hiérarchie est exprimé par au moins une expression référentielle, toutes les langues n'ont pas nécessairement à leur disposition ces six niveaux. Gundel *et al.* (1993, p. 285) appliquent la *givenness hierarchy* à différentes langues comme le chinois, le japonais, le russe ou l'espagnol. De façon générale, plus la forme de l'expression référentielle est réduite (p. ex. pronom ou anaphore zéro), plus la position dans la hiérarchie est élevée et le référent dans le focus, reproduisant en cela une hiérarchie similaire à celle proposée par Givón (1992). Bien que Gundel *et al.* (1993) n'aient pas testé leur hiérarchie en français, la traduction des exemples donnés en (37) montre la pertinence d'une telle hiérarchie des expressions référentielles.

Si on reprend l'exemple proposé par Givón (1992) et repris ci-dessous, l'entité *man* est très accessible et c'est la raison pour laquelle c'est une anaphore zéro qui est préférée à une anaphore nominale.

(35) *The man finished and Ø left.*

Un tel résultat suggère par ailleurs qu'un nom signale une position plus basse dans la *givenness hierarchy*. C'est à quoi s'intéresse Mulkern (1996) : un nom complet (composé du prénom et du nom de famille en français) est utilisé pour faire référence à une entité pour laquelle l'allocutaire n'a pas nécessairement de représentation en mémoire, puisqu'aucune mention précédente à cette entité n'a été réalisée plus tôt dans le discours. Par conséquent, le statut cognitif d'un nom propre ne dépasse pas le niveau « identifiable de façon unique ». Pour un nom propre (p. ex. un prénom, un surnom ou un patronyme), on peut considérer que le statut cognitif sera plus élevé dans la hiérarchie (p. ex. familial), car l'allocutaire doit avoir une représentation de l'entité référente.

Cependant, de façon assez conservatrice les pronoms (zéros ou non) seront considérés dans le focus et les groupes nominaux avec articles définis comme ayant le statut cognitif d'identifiable de façon unique. Par contre, le statut des groupes nominaux avec un article démonstratif ou indéfini est moins clair, notamment à cause des différents types de démonstratifs disponibles dans une même langue. Ainsi, selon les langues, l'utilisation d'un article démonstratif ne sera acceptée que pour des entités familières (anglais, japonais, russe et espagnol), à l'exception du chinois, parmi les langues analysées par Gundel *et al.* (1993). La distribution de l'article indéfini, pour les langues qui en disposent¹⁵, varie aussi en fonction du statut cognitif.

¹⁵ À propos des articles indéfinis dans les langues, voir aussi Lambrecht (1994).

En anglais, Sidner (1979) propose que *this* et *this* N signalent un focus principal et que *that* N est cospécifieur d'un ancien focus, ce qui correspond aux statuts d'activé et familier dans la *givenness hierarchy*.

Si l'on prend l'exemple d'un référent qui n'est pas représenté dans la mémoire de l'allocutaire au moment où il est mentionné, ce référent reçoit alors le statut cognitif d'identifiable. Une fois que ce référent a été introduit, son statut dans la hiérarchie peut progresser. Par exemple, l'introduction de l'entité *meeting* dans (38a) nécessite l'utilisation de l'article indéfini. Dans la seconde phrase, où une deuxième référence à cette entité est réalisée (38b), c'est l'emploi d'un article défini qui est correct, signalant que cette entité est désormais identifiable de façon unique. L'emploi d'un pronom, correct lui aussi, signale un degré d'accessibilité encore plus grand : l'entité est alors considérée dans le focus.

(38) a. *I'm going to a meeting tonight. (Je vais à une réunion ce soir)*

b. *How long is it/the meeting/your meeting supposed to last? (combien de temps ça/la réunion/ta réunion va-t-elle durer ?, d'après Lambrecht, 1994)*

L'emploi d'un article défini signale d'après la *givenness hierarchy* un statut qui est au moins identifiable de façon unique. Or, on peut observer que, parfois, l'article défini est utilisé pour une entité qui est tout au plus identifiable. C'est notamment le cas pour des groupes nominaux complexes comme *the daughter of a king*, *la fille d'un roi*. Le deuxième nom, avec son article indéfini (*a king*), signifie que l'entité n'est pas identifiable dans le discours. La première entité, référée par *the daughter*, serait identifiable en vertu de son article défini, mais cela n'est possible qu'en regard de l'entité *king*. On ne peut considérer que dans le discours, l'entité *daughter* soit en effet identifiable. On parle alors d'*anchorless anchoring* (littéralement *accroche sans amarres*) ou de *boostrapping* pragmatique (Lambrecht, 1994).

Certaines entités qui peuvent être inférées à partir du discours sont référées à l'aide d'expressions anaphoriques qui ont une position élevée dans la *givenness hierarchy* alors qu'elles ne font pas partie du discours au préalable. Il s'agirait alors d'une violation de cette hiérarchie, qui peut cependant être expliquée par le phénomène d'accommodation présenté par Heim (1982, cité par Gundel *et al.*, 1993, voir aussi Poesio & Vieira, 1998). Néanmoins, les auteurs de la *givenness hierarchy* optent pour une conception plus orientée vers la performance que ne le fait Heim (Gundel *et al.*, 1993, 2000 ; Gundel, Hedberg, & Zacharski, 2001).

Les pronoms sont parmi les formes les plus réduites d'expressions référentielles et désignent un référent qui est dans le focus. Si les pronoms sont principalement utilisés pour faire référence à une entité qui fait déjà partie du discours (cf. *centering theory*), il se peut cependant que des pronoms soient utilisés en l'absence d'une mention précédente explicite de l'entité à laquelle ils font référence¹⁶, mais qui peuvent facilement être inférés (Gundel *et al.*, 1993). C'est le phénomène d'accommodation, étudié expérimentalement par Sanford, Garrod, Lucas et Henderson (1983) et illustré dans l'exemple (39). Ici, *it* fait référence à l'assemblage des deux bouts de papier.

(39) *I glued two pieces of paper together, and it flew (J'ai collé deux bouts de papier ensemble et ça s'est envolé, tiré de Gundel et al., 1993)*

L'entité désignée par un pronom comme *it* dans (39) est liée à une autre entité du discours au moyen de *bridging inferences* (inférences passerelles, voir Clark, 1975 ; Haviland & Clark, 1974). L'allocutaire interprète l'entité inférée à partir des principes de pertinence (Charolles, 1990 ; Sperber & Wilson, 1989), réalisant un effort cognitif moindre tout en choisissant une interprétation pertinente qui permette d'établir une référence pour cette nouvelle entité, en la liant à une entité déjà existante au moyen d'une *bridging inference*.

Selon la *givenness hierarchy* (Erkū & Gundel, 1987 ; Gundel *et al.*, 1993, 2000), l'emploi d'une anaphore indirecte pronominale comme (39) doit être prohibé, car elle constitue une violation de cette hiérarchie. En effet, l'emploi d'un pronom caractérise une entité du discours qui est dans le focus, ce qui ne peut être le cas avec les anaphores indirectes pronominales. Pour une anaphore indirecte, c'est une anaphore nominale qui devrait être préférée, comme dans l'exemple (40). Des études de corpus (Gundel *et al.*, 2000 ; Gundel, Hedberg, & Zacharski, 2002) montrent que la production d'anaphores indirectes pronominales reste marginale par rapport aux nominales (Erkū & Gundel, 1987).

(40) *I couldn't use the box you gave me. The bottom fell out. (Je n'ai pas pu utiliser la boîte que tu m'as donnée. Le fond est tombé, d'après Erkū & Gundel, 1987)*

Cependant, certains comme Yule (1979) postulent que des anaphores indirectes pronominales sont fréquentes dans les conversations courantes. Ce type d'anaphore fait le plus souvent référence à des entités qui sont très accessibles dans le discours (C. H. Walker &

¹⁶ Voir aussi les travaux sur les *unheralded pronouns*, pronoms inopinés, qui sont eux aussi des pronoms employés en l'absence d'un référent explicite, mais qui peuvent être inférés à partir des connaissances partagées entre le locuteur et l'allocutaire (Greene, Gerrig, McKoon, & Ratcliff, 1994 ; McKoon, Gerrig, & Greene, 1996).

Yekovitch, 1987 ; Yekovitch & Walker, 1978). La notion de focus, telle que définie par la *givenness hierarchy* (Gundel *et al.*, 1993), ne permet pas de rendre compte d'une telle observation. Sanford et Garrod (1981, cité par Cornish, 2001) proposent l'existence de deux types de focus, un focus explicite et un focus implicite, qui est inféré. Ainsi, les anaphores indirectes pronominales font référence préférentiellement à des entités qui sont dans le focus implicite et non dans le focus explicite comme les anaphores directes. Néanmoins, l'étude de Cornish *et al.* (2005) met en évidence que les participants ne peuvent interpréter correctement des pronoms qui sont dans le focus implicite, que lorsque la relation entre l'entité référée par ce pronom et celle du discours (avec laquelle une *bridging inference* est réalisée) est centrale et non périphérique. Une telle observation est assez proche des degrés de relations (entre le référent du discours et celui qui doit être inféré) proposés par Erkü et Gundel (1987).

Si on revient aux anaphores indirectes, nominales, Erkü et Gundel (1987, voir aussi Matsui, 1993 ; Wilson & Matsui, 1998) proposent que dans un fragment de discours l'entité la plus accessible sera le plus souvent considérée comme le site auquel une *bridging inference* sera liée. Dans (41), l'inférence liant *the waitress* à *a Thai restaurant* est évidente.

(41) *We went to a Thai restaurant. The waitress was from Bangkok. (Nous sommes allés dans un restaurant thaï. La serveuse était de Bangkok)*

Dans (42) par contre, *the waitress* peut être potentiellement liée à *the New York Hilton* et *the Thai restaurant*. *The New York Hilton* est néanmoins plus accessible, car cette entité est mentionnée en premier. C'est la raison pour laquelle on l'associera plus facilement à *the waitress*.

(42) *We stopped for drinks at the New York Hilton before going to the Thai restaurant. The waitress was from Bangkok. (Nous nous sommes arrêtés pour boire au New York Hilton avant d'aller au restaurant thaï. La serveuse était de Bangkok)*

Une telle interprétation va à l'encontre des connaissances générales qui pourraient intervenir, favorisant un lien entre *the Thai restaurant* et *the waitress*. Le niveau d'accessibilité des entités est ainsi plus important que les connaissances qui y sont associées. Cependant, Matsui (1995, cité par Wilson & Matsui, 1998) réalise une étude présentant un questionnaire qui met en évidence que le niveau d'accessibilité n'est pas utilisé de façon exclusive dans l'élaboration de *bridging inferences* : des facteurs pragmatiques interviennent. L'emploi de principes basés sur la théorie de la pertinence (Matsui, 1993 ; Wilson, 1998 ; Wilson & Matsui, 1998) permet d'expliquer le choix de *the New York Hilton* dans (42), par

exemple, comme entité liée à *the waitress* par le moyen d'une *bridging inference*. L'application de ces principes influence ainsi le degré d'accessibilité des entités du discours.

Dans la présentation de sa propre théorie, l'*accessibility theory*, Ariel (1990, 2001) critique la *givenness hierarchy*, tout d'abord parce que les statuts cognitifs désignent en fait plus qu'un statut à la fois. Par exemple, un groupe nominal avec l'article défini aura le statut d'identifiable de façon unique, signalant ainsi soit que la représentation de l'entité doit être récupérée en mémoire, soit que cette représentation doit être élaborée à l'instant. Ainsi, le statut d'identifiable de façon unique relève en fait de deux statuts cognitifs différents.

La seconde critique porte sur le fait que la correspondance entre une forme grammaticale et un statut cognitif n'est pas unique. Par exemple, un nom complet peut être soit identifiable de façon unique soit familier (Mulkern, 1996). De plus, l'emploi d'un pronom signale que le référent est dans le focus, mais pourrait aussi être employé pour caractériser tous les autres statuts cognitifs. On a vu aussi avec les anaphores indirectes pronominales qu'il est possible de faire référence avec un pronom à une entité qui ne fait pas partie explicitement du discours. Le modèle d'Ariel (1990, 1996, 2001), l'*accessibility theory*, propose non pas une hiérarchie des statuts cognitifs, mais une hiérarchie des expressions référentielles.

2. 2. L'*accessibility theory*

L'*accessibility theory* (théorie de l'accessibilité) développée par Ariel (1990, 1996, 2001) postule, comme pour la *givenness hierarchy*, que la forme de l'expression référentielle signale l'accessibilité en mémoire de l'entité à laquelle il est fait référence. Cette théorie s'inscrit parmi les différentes approches qui ont été vues jusqu'à présent. Tout d'abord, la hiérarchie d'accessibilité repose sur des critères aussi bien pragmatiques (Gundel *et al.*, 2000 ; Sperber & Wilson, 1989 ; M. A. Walker, Joshi, & Prince, 1998) que sémantiques (Givón, 1992 ; Hobbs, 1979 ; Levinson, 1987a). Ariel minimise l'influence des critères syntaxiques dans sa hiérarchie d'accessibilité (Ch. 6, Ariel, 1990), bien qu'elle ne réduise pas les stratégies syntaxiques comme le c-commande à une stratégie du discours ainsi que le propose van Hoek (1995, Reinhart, 1983, voir Heim, 1998).

En raison de la variabilité inter-langue de la disponibilité des différentes formes anaphoriques (p. ex. zéro, pronoms), la forme de l'expression référentielle et l'accessibilité du référent ne sont pas strictement liées. Ainsi, une langue qui n'aurait pas à sa disposition d'anaphore zéro (autorisée ou non selon la fonction grammaticale) peut cependant signaler un référent très accessible, tout comme une langue qui en disposerait.

La hiérarchie d'accessibilité proposée est la suivante, du marqueur signalant une très grande accessibilité (zéro) à celui qui signale une faible accessibilité (d'après Ariel, 2001)¹⁷ :

- zéro (T. *Mary saw John and then Ø heard Peter, Mary a vu John et puis Ø entendu Peter*) ;
- flexion du verbe (*Peter smokes and drinks a lot, Peter fume et boit beaucoup*) ;
- pronom cliticisé (*Peter saw him, Peter l'a vu*) ;
- pronom non accentué (*he knows it, il le sait*) ;
- pronom accentué (*look at me, regarde-moi*) ;
- pronom accentué et geste (*look at ME!, regarde-MOI!*) ;
- pronom démonstratif proximal (*this, ceci*) ;
- pronom démonstratif distal (T. *I want that, je veux ça*) ;
- démonstratif proximal et groupe nominal (*I see this train, je vois ce train-là*) ;
- démonstratif distal et groupe nominal (*I see that train, je vois ce train*) ;
- démonstratif proximal et modifieur (*this slow train, ce train lent*) ;
- démonstratif distal et modifieur (*that slow train, ce train lent*) ;
- prénom (T. *I saw John, j'ai vu John*) ;
- nom propre (T. *I saw Smith, j'ai vu Smith*) ;
- description courte avec article défini (T. *I saw the car, j'ai vu la voiture*) ;
- longue description avec article défini (T. *I saw the car belonging to the man down the street, j'ai vu la voiture appartenant à l'homme en bas de la rue*) ;
- nom complet (T. *I saw John Smith, j'ai vu John Smith*) ;
- nom complet et modifieur (*I saw the old John Smith living down the street, j'ai vu le vieux John Smith qui vit en bas de la rue*).

Une comparaison rapide entre l'*accessibility hierarchy* et la *givenness hierarchy* montre que, dans la première, les indéfinis sont exclus, ce qui place les groupes nominaux avec un

¹⁷ Les expressions référentielles sont soulignées dans les exemples. Si ces derniers sont tirés de Toole (1996), ils sont précédés d'un T. Les autres sont construits par moi-même à partir de la description d'Ariel (1990).

article défini à un niveau d'accessibilité bien plus bas qu'ils ne le sont sur la *givenness hierarchy* (voir aussi Chafe, 1974, 1996 ; M. A. Walker & Prince, 1996).

Notons que dans sa première version de l'*accessibility hierarchy*, Ariel (1990) place comme marqueur signalant un haut niveau d'accessibilité les trous syntaxiques, les pronoms réfléchis et les traces *wh*. Elle met aussi sur deux niveaux distincts les pronoms personnels à la première et deuxième personne d'un côté, ceux à la troisième personne de l'autre. Les premiers sont considérés comme ayant un niveau d'accessibilité intermédiaire contrairement aux pronoms à la troisième personne qui ont un niveau d'accessibilité élevée. La version plus récente de la hiérarchie (Ariel, 2001), détaillée plus haut, exclut ces formes d'expressions référentielles et cette distinction au niveau des pronoms.

Les expressions référentielles fonctionnent comme des marqueurs d'accessibilité de l'antécédent. Le niveau d'accessibilité est influencé par quatre facteurs (Ariel, 1990). Tout d'abord, plus la distance entre une expression référentielle et une anaphore croît, plus le niveau d'accessibilité diminue. La compétition est un deuxième facteur : plus il y a d'entités auxquelles l'anaphore peut faire référence, plus l'accessibilité décroît. L'unité, c'est-à-dire si les deux références sont dans le même cadre, point de vue ou paragraphe, intervient aussi dans l'accessibilité. Enfin, la saillance, entité topique ou non, joue un rôle important dans le niveau d'accessibilité du référent : elle n'est pas tant déterminée par la saillance physique de ces expressions référentielles que par la saillance de l'antécédent dans le discours. Cette saillance dépend de trois éléments, l'informativité, la rigidité et l'atténuation (Ariel, 1990, 1996, 2001).

L'informativité (Giora, 1988) désigne la quantité d'information lexicale ; par nature un pronom ou un groupe nominal complexe sont plus informatifs que le zéro. La rigidité traduit la possibilité de sélectionner un référent unique à partir d'une expression référentielle. Elle permet de distinguer les pronoms des première et deuxième personnes, qui sont plus facilement identifiables (par le locuteur et l'allocutaire) qu'un pronom à la troisième personne. La rigidité régit aussi la distinction entre les noms propres et les descriptions avec article défini. Enfin, l'atténuation traduit les différences quant à la taille phonologique, pour des expressions référentielles qui ont le même degré d'informativité. La prosodie est aussi un facteur de l'atténuation, puisque, grâce à elle, on distingue les pronoms accentués de ceux qui ne le sont pas. Et cela a un effet sur l'accessibilité (Ariel, 1988). Plus l'expression référentielle est informative, rigide et moins elle est atténuée, plus sa position sur l'*accessibility hierarchy* est faible et signale un antécédent peu accessible.

Selon leur degré d'accessibilité, les référents vont être activés à divers niveaux de la représentation mentale du discours (Kintsch & van Dijk, 1978). Ainsi, un référent très saillant, très accessible, appartient au thème discursif global ; les référents ayant un faible niveau d'accessibilité sont non thématiques et les référents ayant un niveau d'accessibilité moyen appartiennent au thème de discours local.

Ariel (2001) suggère que l'accessibilité n'est pas uniquement influencée par la distance référentielle entre deux références à une même entité, contrairement à ce que propose Givón (1992). Si la distance (O'Brien, Raney, Albrecht, & Rayner, 1997) a un effet sur l'accessibilité (Ehrlich, 1980 ; Ehrlich & Rayner, 1983) et la forme de l'expression référentielle (Ehrlich, 1981 ; Givón, 1992 ; Yule, 1981), le rôle grammatical (Halmari, 1996), la saillance thématique, mais aussi les différentes frontières du discours (phrases, paragraphes) et le nombre de références à cette entité exercent aussi une influence sur l'accessibilité des entités (Ariel, 1990 ; Toole, 1996). De plus, le caractère animé du référent (Dahl & Fraurud, 1996) a un effet sur la forme de l'expression référentielle, puisqu'un pronom à la troisième personne sera plus fréquemment employé pour une entité humaine que non humaine.

Toole (1996) réalise une analyse de corpus où l'accessibilité est calculée en fonction de la distance référentielle, de la compétition entre les référents, de la saillance de ces derniers ainsi que de l'appartenance à un même fragment de phrase. Les résultats confirment la hiérarchie proposée par Ariel (1990, 2001). En effet, on observe que plus une entité est accessible plus la forme de l'expression référentielle est réduite.

Si l'on reprend l'*accessibility hierarchy*, une entité a une position élevée à condition que plusieurs références aient été faites (Albrecht & O'Brien, 1991). Cependant, si cette entité fait partie du thème discursif global, il est possible d'utiliser une expression référentielle haute dans la hiérarchie malgré un nombre restreint de références. Il s'exerce aussi une compétition entre les différentes entités du discours (J. E. Arnold & Griffin, 2007 ; O'Brien & Albrecht, 1991) : plus il y a de référents dans le discours, plus leur niveau d'accessibilité est faible.

Les pronoms sont relativement peu informatifs et surtout courts (*i. e.* atténués) et se trouvent relativement hauts dans la hiérarchie. Les groupes nominaux sont plus longs, moins rigides et moins atténués et ont un niveau hiérarchique plus bas. Si une expression référentielle est utilisée de façon à signaler le niveau d'accessibilité correct, utiliser une expression référentielle correspondant à un niveau bien plus bas dans la hiérarchie consiste à donner trop de précisions et surtout à indiquer un changement de niveau de thème discursif

(Expérience 3, Vonk, Hustinx, & Simons, 1992). Cette surinformation fait écho à l'effet de *repeated name penalty* (cf. *centering theory*).

Kleiber (1990) propose que l'acceptabilité d'une anaphore n'est pas uniquement liée à l'accessibilité de l'entité référée. Dans les exemples (33) ou (38), on a vu que la première mention d'une entité est généralement réalisée avec un GN contenant un article indéfini. Une fois que cette entité fait partie du discours, les références ultérieures sont un pronom ou un GN avec un article défini, comme dans (43a-c).

- (43) *Un avion s'est écrasé hier à New York.*
 a. *Il transportait 100 personnes.*
 b. *L'avion transportait 100 personnes.*
 c. *Cet avion transportait 100 personnes.* (d'après Kleiber, 1990)

Les trois expressions référentielles, *il*, *l'avion* et *cet avion* sont tout à fait acceptables et signaleraient des niveaux d'accessibilité différents. Pour une phrase initiale identique, ces expressions référentielles, signalant un niveau d'accessibilité important (44a-b), ne sont plus correctes. Le niveau d'accessibilité de l'entité lors de la première référence (*un avion*) ne suffit pas expliquer le choix de l'expression référentielle dans la proposition suivante (44a-c).

- (44) *Un avion s'est écrasé hier à New York.*
 a. ? *Il relie habituellement Miami à New York.*¹⁸
 b. ? *L'avion relie habituellement Miami à New York.*
 c. *Cet avion relie habituellement Miami à New York.* (idem)

Kleiber (1990) suggère que des contraintes sémantiques ou pragmatiques interviennent en plus de l'accessibilité.

Enfin, Ariel (1990, 2001) place les pronoms et les zéros sur deux niveaux d'accessibilité distincts. Une série d'expériences en langue chinoise évalue l'effet de *repeated name penalty* (Yang, Gordon, Hendrick, & Wu, 1999), qui est observé seulement quand l'anaphore est un GN, mais pas lorsqu'il s'agit d'un pronom ou d'une anaphore zéro, qui ne diffèrent pas entre eux. Ce dernier résultat montre que ces deux expressions référentielles ne signalent pas un

¹⁸ Le « ? » signifie que la phrase n'est pas agrammaticale, mais n'est pas satisfaisante.

niveau d'accessibilité ou de proéminence différent. Les auteurs proposent qu'il s'agit d'une différence stylistique ou articulatoire.

2. 3. *L'informational load hypothesis*

L'effet de *repeated name penalty* (Gordon & Chan, 1995; Gordon *et al.*, 1993; Gordon & Searce, 1995) traduit un coût de traitement supplémentaire pour traiter un groupe nominal lorsqu'un pronom est l'expression référentielle attendue. Le coût cognitif induit est lié au niveau d'accessibilité signalé par les expressions référentielles (Ariel, 1990, 2001 ; Gundel *et al.*, 1993, 2000) : un groupe nominal indique un niveau d'accessibilité plus faible qu'un pronom. Si le niveau d'accessibilité signalé par l'expression référentielle est plus bas que celui qui est attendu, on observe un coût de traitement supplémentaire. C'est la *repeated name penalty*.

L'*informational load hypothesis* (hypothèse de la charge informationnelle, Almor, 1999 ; Almor & Nair, 2007) est basée à la fois sur les théories de l'accessibilité et les observations de Gordon *et al.* sur la compréhension, et porte sur le coût associé au traitement des anaphores. Ce coût est induit par l'intégration de la représentation sémantique de l'entité référée. Cette théorie explique pourquoi un pronom est souvent préféré lors de références successives à une même entité dans le discours.

Almor (1999, 2004 ; Almor, Kempler, MacDonald, Andersen, & Tyler, 1999) émet l'hypothèse que la difficulté engendrée par deux références successives à une même entité à l'aide d'un nom (la *repeated name penalty*) n'est pas tant liée à un surcoût de traitement pour un groupe nominal par rapport à un pronom, mais plutôt à la répétition de ce nom par rapport à la présence d'un pronom. La spécification excessive causée par la répétition du groupe nominal impose une charge de traitement sans pour autant apporter de nouvelles informations, ce qui est coûteux pour la mémoire de travail (Baddeley, 1992, 2003). L'*informational load hypothesis* prend en compte la maxime de quantité de Grice (1975, cité par Sperber & Wilson, 1989) et la typicalité des référents peut aussi affecter le coût informationnel (Rosch, 1975).

2. 4. Les théories de l'accessibilité et le traitement des propositions relatives

Les théories qui viennent d'être présentées, la *givenness hierarchy*, l'*accessibility theory* ou l'*informational load hypothesis* considèrent toutes des expressions anaphoriques qui se

trouvent dans des énoncés successifs. Elles ne discutent pas des relativiseurs ni de l'attachement des propositions relatives ambiguës, mais elles peuvent apporter un élément de réponse intéressant à propos des préférences d'attachement.

Les anaphores considérées par ces théories sont le plus souvent dans des énoncés successifs et les propositions relatives font nécessairement partie de la même phrase que l'antécédent auquel le relativiseur fait référence. Comme le relativiseur est une expression anaphorique, Ariel (1990, 1999) évalue la relation entre l'accessibilité et la forme des relativiseurs en hébreu. Dans cette langue, une relative sujet est selon la norme réalisée sans pronom ou marqueur de la relative (*i. e.* avec un trou syntaxique), mais pour les autres fonctions syntaxiques qui signalent un antécédent moins topique (objet direct et obliques, Givón, 1990), un pronom résomptif peut être utilisé. Ainsi, pour une fonction syntaxique modérément basse dans la hiérarchie (p. ex. objet direct), un trou syntaxique est généralement utilisé, mais un pronom résomptif est possible. Enfin, pour les fonctions syntaxiques les plus basses (p. ex. obliques), l'emploi d'un résomptif est obligatoire (pour une étude conjointe des deux hiérarchies d'accessibilité, voir Halmari, 1996). Avec la fonction sujet, le relativisé peut être réalisé avec un résomptif, ce qui est d'autant plus acceptable qu'il est topicalisé. La présence d'un résomptif en hébreu est aussi mieux acceptée dans les propositions relatives non restrictives plutôt que restrictives (Ariel, 1990), qui plus est quand l'antécédent est indéfini¹⁹.

Ariel (1990, 1999) propose que la stratégie de relativisation (trou syntaxique ou pronom résomptif) est fonction de l'accessibilité, de la saillance de l'antécédent, mais aussi de la fonction syntaxique du relativisé (Keenan & Comrie, 1987). L'auteur compare un score d'accessibilité avec la forme du relativiseur (trou syntaxique ou résomptif) dans un corpus d'énoncés produits spontanément. Ce score d'accessibilité est calculé à partir de la distance entre l'antécédent et le relativisé, la complexité – *i. e.* longueur de l'antécédent et compétition entre référents –, le caractère obligatoire de l'argument du relativisé et la restrictivité. Les résultats confirment de façon générale les hypothèses posées : plus l'accessibilité diminue, plus c'est un résomptif qui est utilisé. Il y a aussi un effet de la fonction syntaxique du relativisé : plus la fonction est basse dans l'*accessibility hierarchy* de Keenan et Comrie (sujet, objet direct, indirect, oblique, génitif, comparative objet Keenan & Comrie, 1987), plus

¹⁹ L'effet de la définitude pour les propositions relatives non restrictives s'oppose à la suggestion de Prince (1990), selon laquelle la définitude n'a d'effet que pour les propositions relatives restrictives.

un résomptif est fréquent. La longueur de l'antécédent, la distance entre l'antécédent et le relativiseur et la restrictivité affectent aussi la forme du relativiseur.

L'approche d'Ariel est intéressante et confirme la proposition de Hemforth *et al.* (Hemforth *et al.*, 2000) : on peut considérer les relativiseurs comme des expressions anaphoriques, signalant le niveau d'accessibilité de leur antécédent. Cependant, l'aspect optionnel ou obligatoire de l'emploi d'un pronom résomptif selon la fonction syntaxique du relativisé ne permet pas un contrôle exact d'une fonction syntaxique à l'autre.

L'application de l'*accessibility hierarchy* d'Ariel pour expliquer les préférences d'attachement permettrait d'expliquer en partie les préférences d'attachement observées en anglais par Cuetos et Mitchell (1988). Ces derniers, dans la version en anglais de leur matériel, ont utilisé *that* si N1 est non-animé et N2 animé humain et *who* si N1 et N2 sont animés humains. Dans le but d'éviter un biais du relativiseur *that* pour l'entité non-animée, entraînant une préférence pour N1, les auteurs discutent des préférences d'attachement en anglais uniquement pour les phrases avec *who*. Or, l'interaction entre les deux types de propositions relatives, avec *that* et *who*, pourrait être expliquée non pas par un biais pour l'entité non-animée, mais plutôt par le fait que *who*, plus informatif que *that*, signale un attachement à l'entité la moins accessible, en l'occurrence N2. *That*, peu informatif, signale par contre un attachement vers N1, ce qui est observé ici. Cette interprétation que je propose est fragile au vu de ces données. En effet, il y a un autre facteur de variation entre les deux types de phrases présentées : *that* désigne une proposition relative restrictive et *who* est associé de façon préférentielle à des propositions relatives non restrictives (Gibson *et al.*, 2005). Le but de cette recherche est de proposer une meilleure évaluation de l'effet de la forme du relativiseur sur les préférences d'attachement. Mais avant de présenter la problématique et les hypothèses de travail, intéressons-nous à un dernier courant théorique, la *centering theory*.

2. 5. La centering theory

La *centering theory* (Gordon & Chan, 1995 ; Gordon *et al.*, 1993 ; Grosz *et al.*, 1983, 1995 ; Grosz & Sidner, 1997 ; Poesio *et al.*, 2004 ; M. A. Walker, 1998) considère que, dans le discours, il existe deux niveaux de cohérence, un niveau global et un autre local (McKoon & Ratcliff, 1992 ; McKoon, Ratcliff, Ward, & Sproat, 1993). La cohérence globale traite des relations entre fragments relativement importants du discours, alors que la cohérence locale

traite des phrases ou des propositions et des facteurs qui permettent d'établir un lien entre ces dernières et de constituer ainsi des fragments de discours cohérent.

Cette distinction de niveaux n'est pas sans rappeler celle de Kintsch et van Dijk (1978) dans leur modèle de compréhension de texte. Ce modèle diffère cependant de celui de la *centering theory* puisqu'il s'appuie sur la sémantique du texte pour en établir la cohérence, alors que la *centering theory* repose sur la forme des expressions référentielles et les caractéristiques syntaxiques des éléments du discours.

La *centering theory*²⁰ postule l'existence de centres dans le discours. La forme grammaticale de ces centres, ainsi que leur évolution dans le discours, maintient la cohérence. Cette théorie s'attache surtout à expliquer les anaphores pronominales, forme de prédilection à l'écrit qui est traitée plus facilement dans des tâches de compréhension.

2. 5. 1. Les principes de la théorie

2. 5. 1. 1. Les centres dans le discours

La cohérence entre des propositions finies est signalée par la forme des expressions référentielles. L'utilisation d'un pronom indique que son référent est dans le focus discursif (et attentionnel Grosz *et al.*, 1995), comme le montre l'exemple suivant (45), tiré de Sidner (1981).

(45) i. *I have two dogs.*

ii. *One is a poodle; the other is a cocker spaniel.*

iii. *The poodle has some weird habits.*

iv. *He eats plastic flowers and likes to sleep in a paper bag. It's a real problem keeping him away from plastic flowers.*

v. *The cocker is pretty normal, and he's a good watch dog.*

vi. *I like having them as pets.*

(J'ai deux chiens. Un d'eux est un caniche ; l'autre est un cocker. Le caniche a des habitudes bizarres. Il mange des fleurs en plastique et aime dormir dans un sac en papier.

²⁰ Pour une revue de la *centering theory* en langue française, voir Cornish (2000).

C'est un vrai problème de le tenir éloigné des fleurs en plastique. Le cocker est plutôt normal, et il est un bon chien de garde. J'aime les avoir comme animaux de compagnie)

Il y a deux entités différentes dans ce texte, *poodle* et *cocker* (45ii), qui sont dans le focus du discours successivement²¹. Quand la première entité est dans le focus, après la mention de *the poodle* dans (45iii), ses références ultérieures sont des pronoms, *he* ou *him*, (45iv). La seconde entité est introduite par *the cocker* (45v), l'utilisation du pronom *he* est non ambiguë (45v) et y fait référence, car l'entité *cocker* est à présent dans le focus.

La *centering theory* postule l'existence de centres (*centers*) ou focus. Les centres sont définis à partir de critères syntaxiques, sémantiques, discursifs et intentionnels. Dans le cadre de ses travaux précurseurs de la *centering theory*, Sidner (1981) propose l'existence de deux focus distincts. Le premier est basé sur des critères discursifs ; c'est le focus discursif (*discourse focus*, pour une définition plus complète, voir Sidner, 1979). Le second intègre les autres critères : c'est le focus d'acteur (*actor focus*), qui est le plus souvent l'agent d'un événement.

Prenons un exemple comme (46), qui, selon Sidner (1981), illustre cette distinction entre les deux types de focus. La première phrase (46i) introduit une entité, *Jeff*, reprise dans (46ii) où une seconde entité, *Carl*, est introduite. Le pronom sujet *he* dans (46iii) fait référence à une entité masculine, *Jeff* ou *Carl*. Un locuteur natif de la langue interprète de façon non ambiguë le pronom *he* comme faisant référence à l'entité *Jeff*.

- (46) i. *I haven't seen Jeff for several days.*
 ii. *Carl thinks he's studying for his exams,*
 iii. *but I think he went to the Cape with Linda.*

(Je n'ai pas vu Jeff depuis plusieurs jours. Carl pense qu'il prépare ses examens, mais je pense qu'il est allé au Cap avec Linda, tiré de Grosz et al., 1983)

D'après Sidner (1981), le focus discursif dans (46ii) est l'entité *Jeff*, qui est réalisée par le pronom sujet *he*. Le focus d'acteur, auquel un poids plus important est accordé, est l'entité *Carl*. De cette façon, le pronom *he* dans la proposition suivante, (46iii), devrait préférentiellement faire référence au focus d'acteur. Or, dans (46iii), *he* fait référence à *Jeff* et

²¹ En ce qui concerne le traitement du discours et des anaphores et aussi l'effet du focus sur le traitement de ces dernières, voir aussi les travaux de Sanford et Garrod (Garrod *et al.*, 1994 ; Garrod & Sanford, 1977, 1990).

non à *Carl*. L'explication de Sidner ne suffit pas à justifier pourquoi le lecteur considère que *he* fait référence à *Jeff* et non à *Carl*.

La *centering theory*, elle, propose une explication plus satisfaisante. Cette théorie repose sur le postulat qu'il existe des centres dans le discours. Ils sont de deux types, *backward* et *forward looking*.

Chaque proposition contient un unique *backward-looking center* (littéralement, centre rétrospectif, abrégé Cb) et plusieurs *forward-looking centers* (centres prospectifs, Cf). Le Cb sert à créer un lien avec ce qui précède (la première phrase d'un discours en est par conséquent dépourvue) et les Cf à établir un lien avec la proposition suivante (Grosz *et al.*, 1983). Ces centres sont les différentes entités du discours, de type varié. Ils peuvent être des personnes, des objets du monde réel ou bien des fonctions ou situations. L'entité en Cb est aussi un Cf.

L'exemple (46) permet d'expliquer ce que sont les Cb et les Cf. Dans (46i) est introduit un Cf, l'entité *Jeff*. Comme il s'agit de sa première mention, l'emploi d'un nom propre plutôt qu'un pronom est attendu (Givón, 1990). L'entité *Jeff* devient aussi le Cb, car il s'agit en l'occurrence de l'unique Cf de la phrase. Dans (46ii), la phrase contient deux références à deux entités, *Carl* et *Jeff*, qui sont les Cf de cette phrase. *Jeff*, qui est l'entité prééminente, est le Cb de cette phrase. Enfin, dans (46iii), le pronom *he* fait référence au Cb de la phrase précédente, l'entité *Jeff*. Il y a deux Cf dans cette phrase, *Jeff* et *Linda*.

Contrairement à l'explication de Sidner (1981), celle de la *centering theory* parvient à expliquer pourquoi dans (46iii) *he* fait référence à *Jeff* de façon univoque, car elle en est le Cb. L'entité *Jeff* est choisie comme Cb de la façon suivante : le Cb est sélectionné parmi les Cf de la proposition qui précède, ces derniers étant ordonnés sur une échelle de prééminence (Grosz *et al.*, 1995). Le Cf qui est le plus prééminent est le centre préféré, ou Cp (Brennan, Friedman, & Pollard, 1987 ; M. A. Walker *et al.*, 1998) et constitue souvent le Cb de la proposition suivante.

Le classement des Cf est guidé par plusieurs facteurs²². La fonction grammaticale du Cf en influence la prééminence (Gordon & Chan, 1995 ; M. A. Walker, Iida, & Cote, 1994) : un référent qui a la fonction sujet est plus prééminent qu'un référent ayant la fonction d'objet ou

²² La *centering theory* a été développée dans de nombreuses versions (pour une revue, voir Poesio *et al.*, 2004). Cependant, on parlera principalement des versions fondatrices de la théorie (Brennan *et al.*, 1987 ; Gordon & Chan, 1995 ; Gordon *et al.*, 1993 ; Grosz *et al.*, 1983, 1995).

autre (Bouma & Hopp, 2006 ; McKoon *et al.*, 1993). Strube et Hahn (1999) proposent que le classement des Cf est influencé par leur entrée récente ou ancienne dans le discours. Cette approche présente des similarités avec celle de Gundel *et al.* (1993). Dans les versions plus anciennes de la *centering theory*, la fonction grammaticale est principalement considérée pour élaborer le classement des Cf (Gordon *et al.*, 1993 ; M. A. Walker *et al.*, 1994). Notons par ailleurs qu'un Cf réalisé sous la forme d'un pronom a souvent une proéminence importante parmi les Cf.

En ce qui concerne ces centres, la *centering theory* propose l'existence de deux règles. La première régit la forme grammaticale du Cb et la seconde les transitions entre les Cb successifs.

2. 5. 1. 2. La première règle

La première règle de la *centering theory*, telle qu'elle est formulée par Grosz *et al.* (1983), stipule que si le Cb d'une proposition est identique à celui de la proposition précédente, un pronom doit être utilisé. Dans (45iii-iv), on remarque que les deux références successives de la première entité (*the poodle* et *he*) sont les Cb de ces phrases. Un pronom est bien employé pour la seconde phrase (45iv)²³. Cette règle n'exclut cependant pas que des pronoms soient utilisées pour d'autres entités que le Cb d'une proposition, si ce dernier est exprimé par un pronom.

Cette règle peut aussi être violée. Dans (47i), l'entité *Terry* est introduite et constitue le Cf de la phrase. Dans (47ii), l'entité *Terry* est le Cb, réalisée avec le pronom sujet *he*, et constitue aussi le Cf de cette phrase. Dans (47iii), *Terry* est toujours le Cb (*he*) et il y a deux Cf, *he* et *Tony*. *He*, dans (47iv), fait référence à *Terry* et est encore le Cb de la phrase. Il y a deux Cf, *he* (*Terry*) and *him* (*Tony*). Dans (47v), le Cb est l'entité *Tony* à laquelle renvoie le pronom *he*. L'utilisation du pronom dans (47v) viole la première règle. En effet, *Tony* n'est pas le Cb dans (47iv) et sa référence dans (47v) ne devrait pas être réalisée par un pronom. Néanmoins, (47v) est compréhensible.

(47) i. *Terry really goofs sometimes.*

ii. *Yesterday was a beautiful day and he was excited about trying out his new sailboat.*

²³ On pourrait se demander pourquoi dans (45iii) la référence à l'entité *poodle* est réalisée avec un groupe nominal et non un pronom sujet : cela est lié au fait que cette entité n'est pas Cb dans la proposition précédente.

iii. He wanted Tony to join him on a sailing expedition.

iv. He called him at 6 am.

v. He was sick and furious at being woken up so early.

(Terry fait parfois des gaffes. Hier était un jour magnifique et il était heureux à l'idée d'étréner son nouveau bateau. Il voulait que Tony l'accompagne pour une expédition en bateau. Il l'appela à 6 heures du matin. Il était malade et furieux d'être réveillé aussi tôt, tiré de Grosz *et al.*, 1995)

Dans une version ultérieure de la *centering theory*, Grosz *et al.* (1995) formulent la première règle de façon différente : si un Cf d'une proposition est réalisé dans la proposition suivante à l'aide d'un pronom, le Cb de cette même proposition doit lui aussi être un pronom. Avec une telle définition, la première règle n'est plus violée dans l'exemple (47), alors qu'elle l'était avec la première version (Grosz *et al.*, 1983). En effet, dans (47iv), l'entité *Tony* est réalisée par *him*, qui est Cf. Dans (47v), cette même entité devenue Cb devrait normalement être réalisée par le nom *Tony*.

Les deux Cf de (47iv) sont pronominalisés et le Cb de (47v) doit donc l'être lui aussi, ce qui est le cas. Cette seconde version de la première règle paraît meilleure que la première puisqu'elle n'est pas violée dans cet exemple.

Gordon *et al.* (Gordon & Chan, 1995 ; Gordon *et al.*, 1993) proposent eux aussi une version quelque peu différente de la première règle et d'autres aspects de la *centering theory*. Ainsi, ils optent pour une formulation de la première règle encore plus forte que celles de Grosz *et al.* (1983) et (1995) : le Cb doit être réalisé sous la forme d'un pronom. La règle ainsi exprimée a été testée à l'aide d'expériences en compréhension, qui sont présentées ci-après et qui mettent en évidence l'effet de *repeated name penalty*. L'étude de corpus de Poesio *et al.* (2004) évalue les prédictions attendues selon les versions de la *centering theory*. La règle formulée en 1995 par Grosz *et al.* (1995) est vérifiée de façon plus systématique que celle qu'ils avaient proposée précédemment (Grosz *et al.*, 1983) et que celle de Gordon et Chan (1995).

Les versions fondatrices de la *centering theory* (Grosz *et al.*, 1983, 1995) proposent que la fonction grammaticale est le critère qui détermine le classement de la proéminence de Cf (cf. 3. 1. 1.). Des études plus récentes (Pearson, Stevenson, & Poesio, 2000, cité par Poesio *et al.*, 2004) ont montré que si les Cf appartiennent à la proposition principale d'une phrase, ils sont plus proéminents puisqu'ils ont plus de chance d'être repris dans la phrase suivante.

L'évaluation en compréhension de la première règle a révélé un patron de résultat appelé l'effet de *repeated name penalty* (Gordon & Chan, 1995 ; Gordon *et al.*, 1993 ; Gordon & Scarce, 1995 ; Grosz *et al.*, 1995).

Les trois premières expériences de Gordon *et al.* (1993) testent la première règle de la *centering theory*. Ces expériences sont réalisées avec la technique d'auto-présentation segmentée. Sont mesurés les temps de lecture ainsi que les performances aux questions de compréhension qui suivent chaque texte.

La première règle concerne, on l'a vu, la forme du Cb d'une phrase. Que ce soit avec la formulation de Grosz *et al.* de 1983 ou celle de 1995, la cohérence est maintenue si, dans le cas où le Cb est identique dans deux propositions consécutives, un pronom est utilisé. Gordon *et al.* (1993) proposent une version plus forte de cette première règle : le Cb d'une phrase est réalisé avec un pronom.

Afin de tester cette hypothèse, dans la première expérience, on soumet des textes présentant deux entités, *Bruno* et *Tommy* dans l'exemple (48). Tout au long du texte, l'entité introduite en premier (ici, *Bruno*) est le Cb et aussi le Cp. À partir de la seconde phrase, cette entité est réalisée soit à l'aide d'un pronom (48a-b), soit avec le nom (48c). La seconde entité, elle, est toujours introduite dans la deuxième phrase. Ses références ultérieures, de façon similaire, sont soit un pronom (48a), soit un nom (48b-c).

(48) a. *Bruno was the bully of the neighborhood. He chased Tommy all the way home from school one day. He watched him hide behind a big tree and start to cry. He yelled at him so loudly that all the neighbors came outside.*

b. *Bruno was the bully of the neighborhood. He chased Tommy all the way home from school one day. He watched Tommy hide behind a big tree and start to cry. He yelled at Tommy so loudly that all the neighbors came outside.*

c. *Bruno was the bully of the neighborhood. Bruno chased Tommy all the way home from school one day. Bruno watched Tommy hide behind a big tree and start to cry. Bruno yelled at Tommy so loudly that all the neighbors came outside.*

(traduction de a. : *Bruno était le dur du quartier. Il poursuivit Tommy sur le chemin de l'école à la maison un jour. Il le regarda se cacher derrière un grand arbre et commencer à pleurer. Il cria après lui si fort que tous les voisins sortirent*)

Les temps de lecture mesurés et les performances aux questions de compréhension qui suivaient les phrases confirment les attentes des auteurs. Il est plus difficile de traiter un texte

comme (48c) où les références successives aux deux référents sont réalisées avec un nom (voir aussi Cloitre & Bever, 1988). Ce phénomène, appelé la *repeated name penalty* (pénalité pour répétition du nom), est la différence entre une anaphore nominale et une anaphore pronominale, au niveau des temps de lecture de la phrase critique, *i. e.* la dernière phrase des discours (48a-c).

Cette *repeated name penalty* est sensible à la fonction grammaticale du Cb. En effet, les auteurs n'observent une *repeated name penalty* qu'avec l'entité remplissant la fonction sujet, l'entité *Bruno* dans les exemples présentés ici, et non avec l'entité qui remplit la fonction objet, *Tommy*. Comme l'entité *Bruno* est le Cb de cette phrase, on peut considérer que la *repeated name penalty* est spécifique au Cb, mais seulement s'il remplit la fonction sujet (Expériences 1 et 3). En effet, si le Cb a une autre fonction grammaticale, p. ex. objet indirect, mais demeure la première entité de la phrase critique (Expérience 2), l'effet de *repeated name penalty* n'est plus observé. Les auteurs concluent que le Cb est déterminé par la fonction grammaticale sujet, mais que le Cp, lui, est fixé à la fois par le rôle grammatical et la position dans la phrase (pour une évaluation en imagerie cérébrale fonctionnelle, voir Almor, Smith, Bonilha, Fridriksson, & Rorden, 2007).

L'étude d'Arnold (2001) évalue l'effet du rôle thématique sur la proéminence des entités du discours et montre un effet du but : quand les participants traitent une phrase où le sujet du verbe est la source et l'objet le but (*Lisa gave the leftover pie to Brendan, Lisa_{SOURCE} a donné le reste de tarte à Brendan_{BUT}*) ou l'inverse (*Marguerite caught a cold from Eduardo, Marguerite_{BUT} a attrapé le rhume par Eduardo_{SOURCE}*) et qu'ils doivent continuer le texte, le sujet de la phrase produite est plus souvent le sujet de l'énoncé précédent lorsque son rôle était le but. Ainsi, la fonction grammaticale, mais aussi le rôle thématique, influencent la proéminence des entités du discours : les participants font référence en premier lieu au Cb de la proposition précédente avec un pronom, quand il remplit la fonction sujet et quand il est le but. L'effet du rôle thématique est par ailleurs confirmé par deux expériences en production (Stevenson, Crawley, & Kleinman, 1994) qui manipulent le rôle thématique selon le type de verbe et selon la causalité. Par exemple, le sujet d'un verbe comme *prendre* est le but, mais le sujet de *passer* (*John passe le livre*) est la source. La forme pronominale signale la proéminence du référent (Stevenson, 2002 ; Stevenson *et al.*, 1994), elle est privilégiée pour faire référence au but. La causalité est marquée par les connecteurs *so* et *because* (*donc* et *parce que* en français) ; elle signale l'importance du rôle thématique des référents qui les précèdent (Ehrlich, 1980 ; McDonald & MacWhinney, 1995 ; Stevenson, Knott, Oberlander,

& McDonald, 2000). Par exemple, dans *John irritated Bill because...*, le connecteur *because* induit un agent et le pronom qui le suit fait le plus souvent référence au sujet de la principale, *John*. Une anaphore pronominale suivant *so* est utilisée pour faire référence à l'objet de la principale (*Bill* dans *John irritated Bill so...*). Des facteurs thématiques guident ainsi la préférence pour utiliser une anaphore pronominale au lieu de nominale qui signale le haut degré de proéminence de l'entité référée.

2. 5. 1. 3. La seconde règle

Dans l'établissement des Cb de chaque proposition, trois types de transition de Cb sont possibles d'une proposition à l'autre : le centre est conservé (*center continuation*), retenu (*center retaining*) ou modifié (*center shifting*).

Si le centre est conservé, c'est que les Cb de deux propositions successives font référence à la même entité, qui est le Cp, c'est-à-dire le centre préféré, dans la première proposition. Par « centre retenu », Grosz *et al.* (1995) entendent que pour une proposition donnée, le Cb fait référence à la même entité que le Cb de la proposition précédente ; cependant, l'entité n'est plus le Cp. Enfin, quand le centre est modifié, c'est que les Cb de deux propositions successives ne font pas référence à la même entité.

Ces trois types de transition permettent d'établir le degré de cohérence au sein du discours et c'est l'ordre de succession des transitions qui est à la base de la cohérence globale. L'ordre de ces transitions est régi par la seconde règle de la *centering theory* (Grosz *et al.*, 1995). Il est le suivant : centre conservé, retenu puis modifié. L'exemple (49) tiré de Grosz *et al.* (1995) illustre bien cet enchaînement. Si ce fragment de texte est cohérent pour le lecteur, c'est parce que l'ordre des transitions est respecté. *John* est introduit dans (49i), puis devient le Cb de (49ii) et (49iii), réalisé avec le pronom *he*. Comme l'entité *John* continue à être le Cp, le centre est conservé. Un second référent, *Mike*, est introduit dans (49iii) ; il y a alors deux Cf, *he* (i. e. *John*) et *Mike*. *He* est le Cp, c'est-à-dire le Cf le plus proéminent. Dans (49iv), le Cb est toujours *John*, réalisé par le pronom *him*, et il y a deux Cf, *Mike* et *him*. L'entité *Mike* devient le Cp, car elle remplit la fonction sujet. Le centre, dans (49iv), est retenu. Enfin, dans (49v), *he* fait référence non plus à *John*, mais à *Mike*. Le Cb est différent de celui de la phrase précédente ; le centre a été modifié.

- (49) i. *John has been having a lot of trouble arranging his vacation.*
 ii. *He cannot find anyone to take over his responsibilities.*

iii. *He called up Mike yesterday to work out a plan.*

iv. *Mike has annoyed him a lot recently.*

v. *He called John at 5am on Friday last week.*

(*John a eu beaucoup de difficultés à arranger ses congés. Il ne peut trouver personne pour prendre ses responsabilités. Il a appelé Mike hier afin de mettre en place une solution. Mike lui a causé beaucoup de problèmes récemment. Il a appelé John à 5 heures du matin vendredi de la semaine dernière*)

Dans des développements plus récents (Brennan *et al.*, 1987 ; Poesio *et al.*, 2004 ; M. A. Walker *et al.*, 1998 ; M. A. Walker & Prince, 1996), un classement différent des transitions a été présenté. D'une proposition à l'autre, deux facteurs peuvent varier. Tout d'abord, soit les Cb sont identiques pour deux propositions successives, soit ils font référence à deux entités distinctes ; c'est la cohésion selon Kibble (2001). Ensuite, soit le Cb de la phrase correspond au Cp de la proposition précédente, soit il n'est pas le Cf le plus proéminent, ce qui correspond à la saillance selon Kibble (2001).

En combinant ces facteurs, le centre est toujours conservé, retenu ou modifié. En effet, si le centre est conservé, les deux Cb successifs font référence à la même entité et le Cb de la seconde phrase est le Cp de la précédente. Quand le centre est retenu, on a de nouveau deux Cb identiques, mais le Cb de la seconde phrase n'est plus le Cp de la précédente. Enfin, si les deux Cb successifs ne font pas référence à la même entité, deux types de centres modifiés sont possibles. Le premier dénote une modification en douceur (*smooth shift*), quand le Cp de la première phrase est le Cb de la deuxième et le second Cb montre une modification brutale (*rough shift*) si tel n'est pas le cas.

La validité de ces deux règles et des aménagements proposés après la première formulation de la *centering theory* (Grosz *et al.*, 1983) a été éprouvée avec des études en production et en compréhension. Brennan (1995) réalise une étude d'élicitation où chaque participant décrit à un partenaire la vidéo d'un match de basket. L'auteur établit au préalable le degré de proéminence des entités, *i. e.* des joueurs, en fonction de leur rôle dans une action à un moment donné, puis il analyse la forme des expressions référentielles produites et hiérarchise les Cf et Cb des énoncés. Le pronom est utilisé de façon systématique quand l'entité référée est au cœur d'une action et la forme des anaphores en fonction des transitions suit ce qui est prédit par la seconde règle.

L'étude d'un corpus en italien (Di Eugenio, 1998) confirme par ailleurs que c'est l'expression référentielle la plus réduite, le zéro, qui est privilégiée par rapport à un pronom pour exprimer le sujet d'un verbe si l'entité référée est le Cb de la phrase.

2. 5. 2. Les développements récents

Outre l'approche de Beaver (2004) qui propose de reformuler les principes de fonctionnement de la *centering theory* avec le formalisme de la théorie de l'optimalité et qui ne sera pas présentée ici, il convient d'aborder la question des propriétés des pronoms dans le traitement des anaphores. C'est ce que suggère Smyth (Chambers & Smyth, 1998 ; Smyth, 1994) lorsqu'il relève l'effet du parallélisme des fonctions dans la compréhension des expressions référentielles, observé par ailleurs dans une expérience avec la technique d'auto-présentation, phrase par phrase (Chambers & Smyth, 1998) ou mot par mot (Wolf, Gibson, & Desmet, 2004).

Enfin, Gordon *et al.* (Gordon & Hendrick, 1997, 1998 ; Gordon, Hendrick, Ledoux, & Yang, 1999 ; Ledoux, Gordon, Camblin, & Swaab, 2007) proposent une sorte d'actualisation de la *centering theory* avec la théorie de la proéminence du discours, qui intègre aussi les apports d'autres approches théoriques.

2. 5. 2. 1. Intégrer les propriétés des pronoms

On l'a vu, le classement des Cf dépend de plusieurs facteurs comme la fonction grammaticale, le rôle thématique ou bien le caractère ancien ou nouveau. Il a aussi été suggéré que l'ordre d'apparition des Cf dans la proposition influence leur classement. Dans les langues où l'ordre des mots est sujet-verbe-objet, la fonction grammaticale se trouve le plus souvent confondue avec la position du mot dans la proposition.

Dans (49v), l'interprétation du pronom *he* est potentiellement ambiguë puisque le pronom peut faire référence aux deux entités, *Terry* et *Tony*, tous deux masculins singuliers. En vertu de la première règle (Grosz *et al.*, 1995), *he* est interprété comme faisant référence à *Tony*. Cependant, si les deux entités avaient été d'un genre différent, cette interprétation n'aurait pas été ambiguë. C'est pour cette raison que Kehler (1993, 1997) propose que le coût d'une modification de centre sera modulé par les propriétés du pronom comme le genre, le nombre ou le caractère réfléchi. Le pronom fera préférentiellement référence au Cp de la proposition précédente, pour peu que les contraintes d'accord ou de coréférence soient respectées.

2. 5. 2. 2. Le parallélisme des fonctions

Selon Gordon et Chan (1995), on préfère considérer que l'entité en position de sujet sera le Cb de la proposition suivante si cette entité a déjà été mentionnée. Selon la version de la *centering theory* de Gordon *et al.* (Gordon & Chan, 1995 ; Gordon *et al.*, 1993 ; Gordon & Scearce, 1995) l'effet de la position sujet subsume le parallélisme des fonctions (Sheldon, 1974). Crawley, Stevenson et Kleiman (1990) confirment dans leurs expérimentations la prééminence d'une stratégie où le pronom désigne l'entité sujet de la proposition précédente par rapport à une stratégie où le parallélisme des fonctions est pris en compte. Cette observation est cependant limitée au cas où le pronom est ambigu, car il peut potentiellement désigner deux antécédents de genre identique. Si le pronom n'est pas ambigu, aucun effet des deux stratégies n'est observé.

Smyth (1994) quant à lui critique la méthodologie expérimentale utilisée dans cette dernière étude : les phrases cibles ne sont pas entièrement parallèles avec les phrases qui les précèdent, ce qui remet en question l'absence de parallélisme des fonctions. Quand celui-ci est contrôlé dans le matériel expérimental (Chambers & Smyth, 1998 ; Smyth, 1994), une anaphore pronominale est plus souvent produite par les participants qu'une anaphore nominale si les deux références successives remplissent la même fonction syntaxique (voir aussi Streb, Rösler, & Hennighausen, 1999).

2. 5. 2. 3. La théorie de la proéminence du discours

Dans des travaux plus récents, Gordon *et al.* (Gordon & Hendrick, 1997, 1998 ; Gordon *et al.*, 1999 ; Ledoux *et al.*, 2007) présentent une théorie de la coréférence qui intègre plusieurs approches²⁴ : celle de la *centering theory*, celle de la théorie du liage de Chomsky (voir aussi Reinhart, 1983, cité par Heim, 1998) ou encore celle de Kamp et Reyle (1993, cité par Gordon & Hendrick, 1998). Cette théorie de la proéminence du discours (*discourse prominence theory*, Gordon & Hendrick, 1998) se veut une approche intégrée du discours (voir aussi Cornish, 2000), puisqu'elle prend en compte les apports des théories syntaxiques comme la *centering theory* (sur les relations entre propositions) ou la théorie du liage (sur les relations à l'intérieur des propositions), ainsi que les apports d'une théorie sémantique, celle de Kamp & Reyle (1993).

²⁴ Voir aussi Sturt (2003) pour une évaluation de la théorie du liage et du focus discursif lors de la compréhension de phrases complexes.

La théorie de la proéminence du discours repose sur trois principes qui vont être abordés maintenant.

Le premier principe

Le premier principe traite de la forme de l'expression référentielle dans les références successives aux entités du discours. Ici, la référence et la coréférence sont établies à travers trois règles (Gordon & Hendrick, 1997). La première est qu'un nom est utilisé pour introduire une nouvelle entité dans le discours. La seconde porte sur l'usage du pronom, qui est obligatoire pour faire référence à un antécédent déjà présent dans le discours. Cela est mis en évidence avec l'effet de *repeated name penalty* (voir aussi Ledoux *et al.*, 2007). Enfin, la troisième règle, celle de la construction finale, est déclenchée si deux entités distinctes du discours sont introduites avec un nom identique.

Le deuxième principe

Si la forme de l'expression référentielle sert à la coréférence, la structure syntaxique détermine la proéminence des entités du discours (Gordon *et al.*, 1999). Cela est en partie lié à l'aspect séquentiel du langage. C'est le deuxième principe de la théorie de la proéminence du discours. Il traduit l'effet de la fonction grammaticale sur le classement des Cf dans la *centering theory* (Gordon & Chan, 1995 ; M. A. Walker *et al.*, 1994). Cette règle est comparable à ce qui a été avancé par d'autres auteurs, comme par exemple la stratégie gauche-droite²⁵ (Tetreault, 2001) ou le parallélisme des fonctions introduit plus haut (Chambers & Smyth, 1998 ; Smyth, 1994). La proéminence d'une entité du discours facilite le traitement des références ultérieures qui y sont faites, notamment sous la forme de pronoms. Selon Gordon *et al.* (1999), ce sont plus les fonctions syntaxiques, plus que l'ordre d'apparition des constituants dans la phrase, qui ont un effet sur la coréférence.

Revenons brièvement sur la théorie du liage. Le liage est une relation de co-indexation entre deux éléments, dont l'un commande l'autre. Il intervient tout particulièrement pour les expressions anaphoriques, qui sont soumises au principe A. Ce principe veut qu'une anaphore ait un antécédent qui le c-commande et qui appartienne au même GN ou à la même proposition. La c-commande (voir Reinhart, 1981, cité par Gordon & Hendrick, 1997) postule qu'un élément en c-commande un autre si, et seulement si, le nœud syntaxique qui est au-

²⁵ La stratégie gauche-droite de Tetreault (2001) est un algorithme intégrant la version de Grosz *et al.* (1995) qui considère un fonctionnement incrémental pour l'intégration des constituants.

dessus du premier élément contient aussi le second élément. De fait, le premier élément est proéminent. Les expériences de Gordon et Hendrick (1997) ont mis en évidence que l'effet de *repeated name penalty* est plus important quand la première référence c-commande la seconde, parce qu'elle est plus proéminente. Cet effet est tout particulièrement observé si le nom est répété. La proéminence de la première entité rend une anaphore pronominale préférable à une anaphore nominale.

Le troisième principe

Enfin, le troisième principe possède un caractère plus sémantique et traite de l'établissement du sens dans le discours. Chaque proposition s'ajoute dans le modèle du discours qui est construit par le locuteur et son auditeur. Cette conception est similaire à celle qui est développée par Gernsbacher (1989 ; Gernsbacher & Hargreaves, 1988), bien que les prédictions des deux modèles sur la forme des expressions anaphoriques soient opposées. En effet, Gernsbacher (1989 ; Gernsbacher & Hargreaves, 1988) suggère que la répétition d'un référent par un nom permet de renforcer sa représentation au niveau du modèle mental du discours en cours de construction. Avec l'effet de *repeated name penalty* (Gordon & Chan, 1995 ; Gordon *et al.*, 1993 ; Gordon & Hendrick, 1997, 1998 ; Gordon *et al.*, 1999 ; Gordon & Scearce, 1995), Gordon et ces coauteurs postulent que la répétition d'un nom à la place d'un pronom pour faire référence à la même entité proéminente est coûteux cognitivement.

Le caractère incrémental induit par l'élaboration du modèle du discours empêche une interprétation immédiate des pronoms, ce qui rend possible l'établissement de liens cataphoriques.

2. 5. 3. La *centering theory* et le traitement des propositions relatives

La *centering theory* propose une explication de la forme des expressions anaphoriques d'un énoncé à l'autre, plus que d'une proposition à l'autre. C'est tout particulièrement le cas dans l'approche de Brennan *et al.* (1987) puisque l'algorithme utilisé pour identifier les centres et les transitions ne permet pas de rendre compte de l'interprétation de (50b). En effet, d'après cet algorithme, *he* devrait faire référence à *John* qui serait le Cp dans (50a) et par conséquent le Cb dans (50b). Or, les locuteurs interprètent *he* comme faisant référence à *Bill*.

- (50) a. *John criticized Bill,*
 b. *so he tried to correct the fault.*

(*John a critiqué Bill, alors il a essayé de corriger l'erreur*, d'après Stevenson *et al.*, 2000)

Cette préférence d'interprétation s'explique, selon Stevenson *et al.* (2000), par le fait que le verbe induit un focus sur l'entité *Bill* et que le connecteur *so* porte aussi le focus sur les conséquences de la proposition précédente, sur l'entité *Bill*. L'interprétation est alors basée plus sur des critères sémantiques et pragmatiques que structurels comme avec l'approche de Brennan *et al.* (1987). La stratégie d'interprétation proposée suggère par ailleurs que le classement des Cf n'est pas actualisé à la fin de la phrase, mais plutôt après chaque constituant, de façon dynamique. Le pronom est interprété, dès sa lecture, comme faisant référence à l'entité dans le focus.

Si la *centering theory* a pour objectif d'expliquer l'interprétation et l'utilisation des pronoms personnels dans le discours, peut-elle aussi expliquer l'interprétation, *i. e.* l'attachement, d'un pronom relatif, d'un relativiseur ? Si l'on utilise une perspective dynamique, constituant par constituant, comme celle qui est proposée par Stevenson *et al.* (2000), la *centering theory* peut expliquer que, dans de nombreuses langues, on préfère attacher une relative ambiguë à N1. Dans la structure « N1 de N2 », c'est N1 qui est plus dans le focus et le relativiseur sera plus souvent interprété comme faisant référence à la même entité.

Miltsakaki (2004a, 2004b, voir aussi Kameyama, 1998) propose une approche quelque peu différente : les anaphores interphrases subissent des contraintes structurelles, prédites par la *centering theory* et les anaphores au sein d'une même phrase sont régies par des contraintes sémantiques ou grammaticales, comme le suggèrent Stevenson *et al.* (2000). Les propositions subordonnées sont considérées comme faisant partie de la même unité discursive que la proposition principale dont elles dépendent. Séparer les propositions subordonnées des principales, (51b) et (51c), conduit à des transitions brutales de Cb (*rough shift*), suggérant un discours discontinu avec un niveau de cohérence faible.

- (51) a. *John had a terrible headache.* (Cf : *John, headache*)
 b. *When the meeting was over,* (Cf : *meeting*, pas de Cb)
 c. *he rushed to the pharmacy.* (Cf : *he*, pas de Cb)

(*John avait un mal de tête affreux. Quand la réunion fut finie, il courut à la pharmacie*, d'après Miltsakaki, 2004b)

A contrario, considérer que les deux propositions (51b) et (51c) constituent une unité discursive revient à avoir pour cette proposition deux Cf, *the meeting* et *he*, mais aussi un Cb, l'entité *John* dans (51a) qui est réalisée à l'aide du pronom *he*. Le centre est alors conservé entre les deux propositions, ce qui est plus cohérent.

Pour des phrases contenant des propositions relatives dont l'attachement est ambigu, le relativiseur, pronom relatif, fera référence à l'entité la plus haute dans la hiérarchie des Cf de cet énoncé. Ainsi, pour (11), le Cf le plus haut dans la hiérarchie est *hija* et non *coronel*, aussi *que* est-il interprété comme faisant référence à l'entité *hija*.

(11) *El periodista entrevistó a la hija del coronel que tuvo el accidente.*

Miltsakaki (2004a) suggère aussi que le choix d'utiliser une relative plutôt qu'une autre proposition principale signifierait des entités à faible saillance dans le discours et réduirait la complexité de l'inférence nécessaire au traitement d'une telle proposition.

Les travaux de Warren et Gibson (2002) sont tout particulièrement intéressants pour l'effet de la forme des expressions référentielles dans une relative, bien qu'ils n'étudient pas à proprement parler des références successives à une entité du discours comme la *centering theory*. Par exemple, dans l'Expérience 4, les participants lisent des phrases qui contiennent une relative enchâssée au centre, (52) avec la technique d'auto-présentation ; des phrases avec une complétive sont aussi testées. Les auteurs émettent l'hypothèse qu'une expression référentielle comme un pronom de la deuxième personne (*you*) fait référence à une entité présente par défaut dans le discours, alors qu'un GN comme *the boy* nécessite d'y intégrer une entité nouvelle, ce qui est plus coûteux.

(52) *The woman who you/the boy had accidentally pushed off the sidewalk got upset and decided to report the incident to the policeman standing nearby. (La femme que tu/le garçon as/a poussé hors du trottoir s'est énervée et a décidé de signaler l'incident au policier non loin, Warren & Gibson, 2002)*

Les pourcentages de réponses correctes aux questions de compréhension qui suivent ces phrases confirment l'hypothèse des auteurs, tout comme les temps de lecture mesurés sur les régions suivant l'expression référentielle (*accidentally* et *pushed off*). L'ensemble des expériences présentées dans cet article (Warren & Gibson, 2002) s'accorde avec la *centering theory* pour reconnaître la complexité du traitement d'un nom par rapport à un pronom, mais aussi un prénom ou un pronom à la première personne. Une telle observation est aussi suggérée par la *givenness hierarchy*, bien que ces anaphores ne reprennent pas une entité déjà présente dans le discours.

Si on veut évaluer les prédictions attendues selon le parallélisme des fonctions en regard avec les résultats obtenus expérimentalement sur les préférences d'attachement de relatives, force est de constater que, dans les phrases présentées aux participants, N1 a le plus souvent une autre fonction que celle de sujet dans la principale, alors que la proposition relative est toujours sujet. De plus, le parallélisme des fonctions ne peut expliquer la différence d'attachement de la relative lorsqu'elle est précédée par une structure « N1 de N2 » ou « N1 avec N2 » (Baccino *et al.*, 2000 ; De Vincenzi & Job, 1993, 1995).

Parce qu'elle considère les relations des expressions référentielles entre des énoncés et à l'intérieur d'un même énoncé, la théorie de la proéminence du discours permet d'expliquer l'interprétation des propositions relatives dont l'attachement est ambigu.

Cependant, le principe A et celui de c-commande, qui sont utilisés par Gordon et Hendrick (1997) pour expliquer la forme des anaphores, ne permettent pas d'expliquer la présence de préférences d'attachement dans le cas des propositions relatives comme (11).

Par contre, les expériences sur les GN complexes (Gordon & Hendrick, 1998 ; Gordon *et al.*, 1999 ; Ledoux *et al.*, 2007) retiennent l'attention parce qu'elles étudient le traitement d'anaphores, dans des phrases qui comportent un génitif saxon. Pour des énoncés comme (53), Gordon *et al.* (Expérience 5, technique d'auto-présentation, 1999) manipulent la seconde proposition (53b), qui est ou une subordonnée ou une proposition indépendante qu'il n'est pas nécessaire de représenter ici. Le pronom dans (53b) fait référence soit à l'entité possédante, *Bill*, soit à l'entité possédée, *aunt*. Les temps de lecture mesurés au niveau de la phrase critique (53b) montrent qu'un pronom est traité plus rapidement s'il fait référence à l'entité possédée. L'anaphore pronominale fait référence de façon préférentielle à l'entité proéminente.

- (53) a. *Bill's aunt owns a lake house*
 b. *where he/she likes to go swimming.*
 c. *It's nice to live beside a lake.*

(*La tante de Bill possède une maison au bord d'un lac où il/elle aime aller nager.
 C'est agréable de vivre au bord d'un lac*)

Dans le cas des relatives précédées d'un GN avec un génitif saxon, l'interprétation et l'attachement du relativiseur est non ambiguë, et va toujours vers le possédé (Frazier & Clifton, 1996). Une structure « N1 de N2 » est un GN complexe au même titre que celles qui ont été étudiées par Gordon *et al.* (Gordon & Hendrick, 1997, 1998 ; Gordon *et al.*, 1999 ;

Ledoux *et al.*, 2007), « N1 avec N2 » ou le génitif saxon. L'hypothèse de la proéminence des entités pour expliquer la préférence du pronom pour faire référence à une entité permettrait d'expliquer de façon satisfaisante la préférence d'attachement pour N1 qu'on observe dans de nombreuses langues : N1 est plus proéminent que le complément du nom N2.

Dans cette perspective, l'attachement de la relative est guidé par la proéminence des entités du discours. L'étude des préférences d'attachement de propositions relatives apparaît ainsi comme un bon test des prédictions de la théorie de la proéminence du discours, ne serait-ce que parce qu'elle a pour objet d'expliquer les relations anaphoriques à l'intérieur d'une phrase.

Chapitre 3 Discussion des théories

Dans les chapitres précédents, les théories ont été introduites successivement, mais des recoupements s'avèrent nécessaires et pertinents. Tout comme des théories du traitement syntaxique intègrent des aspects sémantiques, pragmatiques, les théories sur le traitement des anaphores, des théories du discours, prennent aussi en compte des aspects syntaxiques comme la fonction grammaticale.

Dans ce court chapitre, je propose de discuter des différentes approches en me focalisant sur des aspects qui paraissent importants. La question de la saillance des entités du discours a par exemple été posée à nombreuses reprises et peut être observée de différentes façons. Une section sera ensuite consacrée au statut des relativiseurs.

1. La saillance

On l'a vu, la saillance des antécédents peut influencer les préférences d'attachement. La saillance des entités du discours est déterminée par de nombreux facteurs : le type de mot, la fréquence d'occurrence, la longueur ou l'informativité. Cette section ne peut prétendre à l'exhaustivité et seuls certains aspects vont être abordés, comme la prosodie, la définitude, la fréquence d'occurrence, l'ordre des mots ou le contexte référentiel.

1. 1. La prosodie et l'accentuation

Les théories du discours s'accordent à considérer que l'accentuation permet de signaler que l'entité à laquelle il est fait référence est nouvelle (Chafe, 1974) et peu accessible (Ariel, 1990, 2001 ; Givón, 1983, 1992). À cet aspect s'ajoute aussi la longueur phonologique (Givón, 1983) : plus une expression référentielle est longue, plus elle a de chances de signaler un référent peu accessible (Ariel, 1990, 2001).

Conformément à la théorie de la proéminence du discours (Gordon & Hendrick, 1997) ou à la *centering theory* (Gordon *et al.*, 1993 ; Grosz *et al.*, 1995), dans des énoncés successifs, le Cb, le plus souvent réalisé sous la forme d'un pronom, fait référence à l'entité la plus saillante. De plus, si on considère qu'un relativiseur est une expression anaphorique

(Hemforth *et al.*, 2000), une proposition relative est attachée de façon préférentielle à l'entité qui est la plus saillante, et qui est ici accentuée.

L'étude de Schafer *et al.* (1996) confirme cette hypothèse, puisqu'effectivement, les participants attachent plus souvent la relative à celui des antécédents potentiels qui est accentué (voir Chapitre 12. 1. 4). La prosodie permet ainsi de lever l'ambiguïté, comme c'est aussi le cas pour d'autres ambiguïtés syntaxiques (Beach, 1991).

Fodor (1998) propose que la prosodie influence aussi la compréhension de phrases écrites, où le lecteur établit des limites prosodiques sur lesquelles il fonde son interprétation syntaxique (voir aussi Carlson, Clifton, & Frazier, 2001 ; Kjelgaard & Speer, 1999). Les variations translinguistiques au niveau des limites prosodiques, notamment entre propositions principale et relative, pourraient expliquer les différences de préférences d'attachement. Premièrement, différentes structures syntaxiques peuvent être utilisées pour mettre en exergue des informations dans le discours. Ensuite le locuteur peut recourir à des constructions qui varient d'une langue à l'autre. En français par exemple, les propositions clivées sont plus fréquentes qu'en anglais (Holmes, 1995).

De plus, la longueur des pauses entre les différentes propositions du discours peut aussi signaler des attachements différents. La réalisation de ces pauses vides ou remplies varie en français (Holmes, 1995) et en anglais (Holmes, 1988, 1995) : les locuteurs du français produisent moins de pauses après un mot de liaison entre deux propositions (comme une conjonction ou un relativiseur) que des locuteurs de l'anglais.

La segmentation dans les paradigmes d'auto-présentation pourrait aussi induire une prosodie différente et par conséquent des préférences d'attachement différentes (Gilboy & Sopena, 1996). Cependant, il subsiste des variations dans les résultats obtenus pour un même type de segmentation, d'une expérience à l'autre.

La *centering theory*, dans ses premières versions (Gordon & Chan, 1995 ; Gordon *et al.*, 1993 ; Grosz *et al.*, 1983, 1995), propose que la hiérarchie des Cf (qui est une échelle de saillance) est actualisée à la fin de chaque énoncé. Des versions ultérieures suggèrent même que ce classement est actualisé après chaque mot (Brennan *et al.*, 1987 ; Chambers & Smyth, 1998). Accentuer une expression référentielle signifie que l'entité référée est peu accessible, mais, en même temps, si elle peut attirer la proposition relative (Schafer *et al.*, 1996), cela laisse supposer que cette entité est devenue la plus saillante, le Cb, en quelque sorte. Une telle hypothèse n'est pas envisagée par la *centering theory* qui, étant évaluée principalement avec

des mesures lors de tâches de lecture, n'inclut ni la prosodie ni l'accentuation dans l'élaboration de la hiérarchie des Cf.

1. 2. La définitude

Le degré de définitude²⁶ (présence dans le GN d'un déterminant indéfini, démonstratif ou défini) signale le statut de l'entité référée. Selon Givón (1982, 1990), l'emploi d'un article indéfini dans un GN signale que l'entité référée est nouvelle (cf. Clark, 1975 ; Haviland & Clark, 1974) et qu'il n'est pas nécessaire de la chercher dans le modèle du discours actuel.

De nombreux auteurs (pour une revue voir Epstein, 2002 ; Lambrecht, 1994 ; Reboul, 1997) se sont intéressés au statut référentiel d'un GN avec un indéfini : il est admis qu'un tel GN ne peut faire référence à une entité déjà présente dans le discours (voir les travaux de Russell, présentés par Reboul, 2001), ou bien que l'entité est inconnue du locuteur ou de l'allocutaire (Reboul, 1997).

Pourtant, d'autres chercheurs proposent qu'un GN fait référence à une entité, même avec un indéfini (Donnellan, 1966). Une proposition relative restrictive doit normalement être attachée à une expression référentielle, son antécédent, qui fait référence à une entité déjà présente dans le modèle du discours. De fait, une proposition relative ne peut pas être attachée à un GN avec un article indéfini, puisque par définition ce dernier fait référence à une entité qui ne fait pas partie du système référentiel du discours. De telles constructions existent pourtant et les énoncés produits alors sont compréhensibles, comme (54).

(54) *A man I had never seen before came into my office and...* (Un homme que je n'avais jamais vu auparavant entra dans mon bureau et..., d'après Givón, 1982)

Une étude en production (Holmes, 1995) montre que des locuteurs de l'anglais et du français produisent des phrases où l'antécédent d'une proposition relative restrictive est un GN avec un article indéfini, dans des proportions non négligeables (respectivement 38 % et 41 % des propositions relatives). L'étude de corpus de Baltazart et Kister (2000) montre aussi que pour des structures « N1 de N2 » en français, pour « N1 défini de N2 indéfini » (*le résumé d'une œuvre*), 85 % des attachements sont réalisés vers N2. Les locuteurs du français produisent ainsi des propositions relatives dont l'antécédent est un GN avec un article indéfini. Ce type de GN, selon Givón (1982), fait bien référence à une entité qui est dans l'arrière-plan discursif (*discourse background*).

La *givenness hierarchy* (Gundel *et al.*, 1993, 2000) considère qu'un GN avec article indéfini est référentiel et signale que le statut cognitif du référent est de type identifiable, le plus bas dans la hiérarchie. Si ce référent ne fait pas partie du discours avant d'être mentionné, il en fait partie dès son introduction par cette expression référentielle. Une telle conception est comparable avec celle de Hawkins (1984). Un GN avec un article défini signale un référent dont le statut cognitif est plus élevé dans la hiérarchie ; l'entité est ici donnée (*given*), alors qu'elle était nouvelle dans le cas d'un GN avec article indéfini.

Epstein (2002) propose par ailleurs que ni la familiarité ni la qualité d'identifiable de façon unique ne sont nécessaires pour expliquer l'emploi de l'article défini. Ce sont la proéminence du référent et le point de vue non canonique employé dans le discours lors de l'introduction de ce référent qui importent. L'article défini ou indéfini signale en fait le niveau d'accessibilité de ce référent. Ce niveau est faible dans le cas d'un indéfini.

Pourtant, l'*accessibility hierarchy* (Ariel, 1990, 2001), qui a justement pour objectif de classer les expressions référentielles selon le niveau d'accessibilité qu'elles signalent, omet d'intégrer le cas des GN avec article indéfini. Cette catégorie est cependant prise en compte dans l'évaluation de l'*accessibility theory* faite par Toole (1996), qui place ce type de GN au plus bas dans la hiérarchie. L'analyse de corpus réalisée montre qu'en effet, un GN avec un article indéfini signale un faible niveau d'accessibilité du référent.

1. 3. La fréquence

Certaines études traitant des préférences d'attachement de propositions relatives ont évalué l'effet de la fréquence de N1 et N2 et ont montré que les lecteurs attachent préférentiellement la proposition relative au nom dont la fréquence est moindre (Pynte & Colonna, 2000) et pour lequel la saillance est comparativement plus importante (Heine *et al.*, 2006 ; van Gompel & Majid, 2004).

Les résultats obtenus par Frenck-Mestre et Pynte (2000b), s'ils montrent un effet de la fréquence, peuvent néanmoins être questionnés en regard des théories de l'accessibilité. Dans cette étude sont présentées des phrases comme (14a-b) où soit N1 est un nom et N2 un prénom (14a), soit N1 et N2 sont des noms (14b). Cette expérience est présentée dans le Chapitre 1, section 2.1.1.1. Les auteurs étudient l'effet de la fréquence relative de N1 et N2 sur les préférences d'attachement en italien.

²⁶ Pour une présentation en français de cet aspect du langage, voir Creissels (2006).

(14) a. *Gianni osserva il ragionere de Caterina che sembra più pensieroso(a) del normale.* (matériel adapté de Frenck-Mestre & Pynte, 2000a)

b. *Patrizia conosceva la segretaria del direttore che era svenuto(a) alla festa.* (matériel adapté de Baccino *et al.*, 2000)

Que ce soit selon la *givenness hierarchy* ou l'*accessibility hierarchy*, un statut cognitif ou un niveau d'accessibilité plus faible est accordé à un nom propre par rapport à un nom commun. De fait, dans le matériel de Frenck-Mestre et Pynte (2000a, 2000b), non seulement la fréquence varie entre N1, un nom commun, et N2, un prénom, mais aussi le niveau d'accessibilité. De plus, d'après les résultats obtenus, les participants attacheraient la proposition relative à l'expression référentielle la moins fréquente et pour laquelle le niveau d'accessibilité est plus faible (d'après l'*accessibility hierarchy* d'Ariel, 2001).

Si l'on se réfère à la *centering theory* et à la théorie de la proéminence du discours, une anaphore telle qu'un relativiseur, qui est assimilé à un pronom, devrait faire référence à l'entité la plus haute dans la hiérarchie des Cf, la plus proéminente du discours. La préférence pour N1 obtenue par Frenck-Mestre et Pynte (2000a, 2000b) suggérerait que la hiérarchie des Cf est établie plus en fonction de la fonction grammaticale (en (14), N1 est objet direct et N2 génitif, complément du nom) et de la fréquence qu'en fonction de la forme de l'expression référentielle et de l'accessibilité qu'elle signale (prénom ou nom commun).

Lorsque la fréquence de N1 et N2 est manipulée (Pynte & Colonna, 2000) et que N1 et N2 sont de même type (signalant ainsi un niveau d'accessibilité équivalent), l'attachement est réalisé préférentiellement vers le nom le moins fréquent, dont la saillance est renforcée. L'hypothèse de la saillance émise par Pynte et Colonna (2000) est par ailleurs confirmée dans une étude en potentiels évoqués (Heine *et al.*, 2006), où la fréquence de l'antécédent est manipulée, ce qui a un effet sur les potentiels évoqués mesurés au niveau de l'anaphore pronominale.

1. 4. L'ordre des mots

Quand l'ordre des mots n'est pas trop strict, il est possible de mettre en exergue une entité en la plaçant en début de phrase ; elle est alors plus saillante. Par exemple en français (Riegel *et al.*, 2004), le recours à des propositions clivées (*c'est à Paul que je parle*) ou des dislocations (*les italiens, je les aime pas*) place le constituant à mettre en valeur en début de phrase. En anglais, où l'ordre des mots est strict, la voix passive place le patient en position de

sujet, au début de la proposition (*the bird was chased, l'oiseau a été poursuivi*, voir Holmes, 1995).

Placer un constituant en début de phrase permet de le situer dans le focus, mais augmente aussi la distance entre cette référence et l'anaphore ultérieure et aurait selon Ariel (1990) un effet sur l'accessibilité de l'entité. Dans les expériences de Gordon et Chan (1995), où la voix est manipulée, les résultats montrent que la *repeated name penalty* n'est pas influencée par la position dans la phrase de l'entité référée (*i. e.* la distance), mais plutôt par la fonction grammaticale de cette entité. La *repeated name penalty* est plus importante quand la première référence a la fonction de sujet dans la première phrase.

Le critère de distance proposé par Ariel (1990) ne suffit pas pour expliquer l'interprétation des anaphores pronominales dans des énoncés comme ceux qui sont étudiés par les tenants de la *centering theory*.

Selon la *recency theory* (Gibson, Pearlmutter *et al.*, 1996 ; Gibson *et al.*, 1999), l'ordre des mots dans les langues guide les stratégies d'interprétation utilisées lors du traitement d'ambiguïtés syntaxiques, en particulier les ambiguïtés d'attachement des relatives. Cette théorie repose sur deux principes, la préférence de récence et la proximité du prédicat. Les différences translinguistiques entre les préférences d'attachement de propositions relatives sont expliquées par une différence du poids accordé selon les langues à ces deux principes, cela entrant en relation avec l'ordre des mots propre à ces langues.

En anglais, les locuteurs se basent principalement sur l'ordre des mots conservateur pour comprendre des phrases, alors qu'en italien par exemple, l'interprétation est fondée sur les informations sémantiques, l'ordre des mots étant plus libre (E. Bates, Devescovi, & D'Amico, 1999; E. Bates, McNew, MacWhinney, Devescovi, & Smith, 1982).

En anglais, et dans le cas où l'attachement de la proposition relative est ambigu, c'est la préférence de récence qui a plus d'importance, alors que pour une langue dont l'ordre des mots est plus permissif, en italien ou en espagnol, la proximité du prédicat prédomine (Gibson, Pearlmutter *et al.*, 1996).

Notons aussi que l'attachement préférentiel de la proposition relative à N2 en anglais (Carreiras & Clifton, 1999; Cuetos & Mitchell, 1988) n'est peut-être pas lié à la distance minimale entre les références à la même entité (N2 et le relativiseur), mais plutôt à la stratégie syntaxique qu'est la préférence de récence.

Pour conclure, les travaux présentés dans cette section suggèrent que l'ordre des mots et la distance que ce dernier induit n'influenceraient pas directement les préférences d'attachement,

mais favoriseraient selon les langues l'utilisation d'une stratégie ou d'une autre selon la *recency theory*.

1. 5. Le contexte référentiel

Des expériences ont montré un effet du nombre de référents sur l'interprétation de phrases syntaxiquement ambiguës (Altmann & Steedman, 1988 ; Altmann *et al.*, 1998 ; Spivey-Knowlton *et al.*, 1993 ; Spivey & Tanenhaus, 1998 ; Tanenhaus *et al.*, 1995) : un groupe prépositionnel est plus facilement interprété comme modifieur d'un nom, une relative réduite, plutôt que par exemple la destination dans (55a), si le nom peut faire référence à deux entités dans le contexte référentiel (deux pommes). Ainsi, si un nom peut faire référence à plus d'une entité, l'interprétation d'une proposition relative comme modifieur est plus fréquente que s'il n'y a qu'une entité. Le contexte référentiel a donc une influence sur l'interprétation d'une proposition relative réduite.

(55) a. *Put the apple on the towel in the box. (Mets la pomme (qui est) sur la serviette dans la boîte)*

b. *Put the apple that's on the towel in the box. (Mets la pomme qui est sur la serviette dans la boîte, Tanenhaus *et al.*, 1995)*

D'après Ariel (1990, 2001), le niveau d'accessibilité d'une entité du discours dépend de la compétition entre référents, du nombre de référents dont une expression référentielle peut être l'anaphore.

Les travaux de Strohner, Sichelschmidt, Duwe et Kessler (2000) ont pour but d'évaluer plus précisément l'interaction entre les éléments du discours et les informations verbales. Dans l'Expérience 1, les participants se voient présenter sur papier des ensembles de trois objets (pyramides et cubes) en dessous desquels se trouve une série de trois instructions (p. ex. *marquez un cube sur le côté gauche*). Les deux premières instructions servent à mettre le focus sur deux objets et la troisième est la tâche cible. Dans cette dernière instruction, l'objet est désigné par un article défini ou indéfini. Un article défini est incorrect quand deux entités peuvent être désignées par cette expression (*le cube* quand deux cubes sont présents). Il y a alors une ambiguïté référentielle (ainsi qu'une compétition entre ces entités). De façon similaire, l'emploi d'un article indéfini (*un cube*) est incorrect quand seul un élément peut être désigné par cette expression. Les autres variables expérimentales sont l'ordre dans lequel les éléments sont dans le focus et la consigne donnée aux participants, à qui on demande soit

d'être tolérants, même si l'expression référentielle dans la tâche cible n'est pas correcte (cf. définitude) ou bien au contraire de signaler une instruction correcte. Les résultats montrent que les participants prennent en compte la définitude du GN, choisissant indifféremment un objet quand il est désigné par un article indéfini, d'autant plus qu'ils ont reçu la consigne d'être tolérant vis-à-vis de l'exactitude de l'instruction.

Les travaux sur le traitement des anaphores ont étudié de façon assez étendue les ambiguïtés référentielles, notamment avec la résolution d'anaphores pronominales. Quand deux entités diffèrent en genre et que l'anaphore pronominale est marquée en genre, l'interprétation du pronom est non ambiguë. C'est le cas dans (56), où *he* fait référence sans ambiguïté à l'entité *Craig*.

(56) *Fiona complimented Craig, and he congratulated James. (Fiona a fait des compliments à Craig et il a félicité James, tiré de Wolf et al., 2004)*

Cependant, l'interprétation d'un pronom peut être ambiguë si, par exemple, les deux référents sont du même genre que le pronom (57).

(57) *Sophie helped Cathy to prepare the food and Simon chatted to her. (d'après Crawley et al., 1990)*

Les stratégies mises en place selon la *centering theory* ou la théorie du parallélisme des fonctions ont déjà été présentées : un pronom fait référence au sujet de la proposition précédente pour la première et le pronom fait référence à l'entité qui est dans une position parallèle dans la proposition précédente pour la seconde. Si on se réfère à la *givenness hierarchy* ou à l'*accessibility theory*, l'emploi d'un pronom signifierait une entité hautement accessible : dans la proposition précédente, il s'agirait du sujet, *Sophie* dans (57).

Des expériences utilisant la technique des potentiels évoqués ont révélé la présence de marqueurs électrophysiologiques différents en fonction du contexte référentiel. Quand une anaphore est référentiellement ambiguë, *i. e.* peut référer à deux entités, on observe une négativité prolongée dans les régions frontales pour une anaphore nominale (Nieuwland, Otten, & Van Berkum, 2007 ; Van Berkum, Brown, & Hagoort, 1999a ; Van Berkum, Brown, Hagoort, & Szwiterlood, 2003 ; Van Berkum, Koornneef, Otten, & Nieuwland, 2007) ou pronominale (Nieuwland & Van Berkum, 2006). Quand une anaphore ne fait pas référence à une entité du discours, *i. e.* diffère en genre, c'est une P600 qui est observée (Nieuwland & Van Berkum, 2006 ; Osterhout & Mobley, 1995).

L'expérience en IRMf de Nieuwland, Petersson et Van Berkum (2007) met en évidence que le contexte référentiel a un effet sur le traitement des anaphores. En effet, des activations cérébrales dissociées sont observées lorsqu'une anaphore pronominale est ambiguë, cohérente (*i. e.* ne réfère qu'à une entité) ou incorrecte (l'anaphore n'est pas du même genre que les antécédents), suggérant des traitements cognitifs différents. En cas d'ambiguïté référentielle, les participants évaluent tous les candidats. Quand l'anaphore pronominale diffère en genre par rapport aux antécédents potentiels, ce sont les régions associées au traitement morphosyntaxique qui sont activées (erreur d'accord) ou bien celles associées à l'élaboration d'inférences (Gina R. Kuperberg, Lakshmanan, Caplan, & Holcomb, 2006).

Enfin, Greene, McKoon et Ratcliff (1992) suggèrent que l'entité à laquelle une anaphore pronominale fait référence n'est identifiée correctement de façon automatique que si le focus discursif ne contient qu'un seul référent correct. Si tel n'est pas le cas, le pronom est attaché aux entités possibles, à moins que plus de temps de traitement ne soit disponible pour engager des stratégies particulières (non automatiques).

Le contexte référentiel et la compétition entre entités du discours qui peut exister conduisent à des différences au niveau des traitements cognitifs engagés qui peuvent être influencés par des facteurs discursifs (nombre de référents), sémantiques (forme de l'expression référentielle) ou même syntaxiques (marquage en genre).

2. Le relativiseur

Un relativiseur est un marqueur anaphorique (Hemforth *et al.*, 2000). Contrairement aux autres expressions référentielles comme un nom ou un pronom, un relativiseur fait nécessairement référence à une entité de la proposition précédente, la proposition principale.

Il est difficile d'appliquer directement les prédictions de la *centering theory* (Grosz *et al.*, 1983, 1995) au traitement d'un relativiseur, car cette dernière s'attache tout particulièrement à expliquer les relations anaphoriques entre énoncés ou phrases, mais aussi parce qu'elle formule des règles sur la pronominalisation. Néanmoins, cette théorie est intéressante puisqu'elle explique les difficultés de traitement rencontrées si l'anaphore est plus informative, d'après la cohérence, ici la hiérarchie des Cf et le Cb ; c'est l'effet de *repeated name penalty*. Celui-ci est observé dans les expériences en compréhension (Gordon & Chan, 1995 ; Gordon *et al.*, 1993) et résulte de la différence de difficultés entre le traitement d'une anaphore nominale et d'une anaphore pronominale, mais ne peut guère être utilisé dans le cas des propositions relatives.

Si l'on émet l'hypothèse qu'un relativiseur se comporte comme un pronom (Hemforth, Konieczny, & Scheepers, 1996, cité par Frazier & Clifton, 1997), il fait alors référence à l'entité la plus proéminente du discours. Cette hypothèse permettrait d'expliquer la préférence d'attachement pour N1 quand la proposition relative est précédée d'une structure « N1 de N2 » dans des langues comme le français.

Un autre facteur pourrait aussi influencer la préférence pour N1. En effet, le degré d'enchâssement plus important pour N2 rend ce dernier moins accessible (Cornish, 2000) et favorise un attachement vers N1, plus saillant. Une telle hypothèse s'oppose à la notion de modifiabilité proposée par Thornton *et al.* (1999), selon laquelle N1, modifié par N2, a moins de chance d'être à nouveau modifié, par la proposition relative cette fois.

Chapitre 4 Problématique et hypothèses

Cette introduction théorique avait pour but de présenter chacune des théories. Je propose ici une synthèse critique de cette littérature, à l'issue de laquelle une problématique émerge : est-ce que la forme du relativiseur peut guider l'attachement d'une proposition relative, par le degré d'accessibilité signalé ? A cette question, plusieurs hypothèses sont formulées.

1. Synthèse critique des éléments présentés dans la partie théorique

Dans le Chapitre 1, des théories sur le traitement d'ambiguïtés syntaxiques, théorie du *garden path*, modèles de satisfaction de contraintes, de compétition et d'intégration ou *unrestricted race* ont été présentées. Elles expliquent les stratégies interprétatives mises en place lors de l'interprétation de ces ambiguïtés. L'ambiguïté liée à l'attachement de propositions relatives (18) se distingue des autres car elle remet en cause la validité de la théorie du *garden path* et la variabilité translinguistique et intra-linguistique qui a été mise en évidence expérimentalement.

(18) *The journalist interviewed the daughter of the colonel who had had the accident.*

En effet, la théorie du *garden path* prévoit un attachement à N2 en vertu de la stratégie du *late closure*, mais si une telle préférence est mesurée en anglais (Carreiras & Clifton, 1999 ; Cuetos & Mitchell, 1988), des expériences ne répliquent pas cette préférence, que ce soit en anglais (Carreiras & Clifton, 1993 ; Traxler *et al.*, 1998) ou dans d'autres langues comme l'espagnol (Cuetos & Mitchell, 1988), le français (Zagar *et al.*, 1997) ou le néerlandais (Desmet *et al.*, 2006).

Des théories ont suggéré l'intervention de différents facteurs pour expliquer ces variations translinguistiques. Par exemple, la *construal hypothesis* (voir Chapitre GARDEN PATH Frazier & Clifton, 1996) avance l'effet de contraintes pragmatiques et conversationnelles : en anglais, l'emploi du génitif saxon (*the colonel's daughter*) rend l'attachement non ambigu, à *daughter*, et par l'emploi de la forme « N1 de N2 », le parseur interprète qu'un attachement autre est signalé, à *colonel*. La *recency theory* (Gibson, Pearlmutter *et al.*, 1996 ; Gibson *et al.*, 1999) postule l'existence de deux stratégies, conduisant à un attachement à N1 ou à N2,

dont le poids varie d'une langue à l'autre en fonction de la rigidité de l'ordre des mots. La *tuning hypothesis* (Brysbaert & Mitchell, 1996 ; Mitchell *et al.*, 1995), elle, explique que les préférences d'attachement en compréhension sont le miroir des ambiguïtés syntaxiques auxquelles le parseur est exposé (en production).

Les mises à l'épreuve expérimentales ont révélé une variation translinguistique, mais aussi des résultats différents pour une même langue. En anglais, certaines expériences ont révélé une préférence pour N2 (Carreiras & Clifton, 1999 ; Cuetos & Mitchell, 1988) alors que d'autres ne mettent en évidence aucune préférence (Carreiras & Clifton, 1993 ; Traxler *et al.*, 1998). Les changements dans les procédures expérimentales, ainsi que certains aspects du matériel étudié, peuvent contribuer à l'explication des différences. Par exemple, on a pu mesurer en français une préférence pour N1 dans des phrases comme (3) avec une tâche de complétion de phrases ou une expérience en *eye tracking* (Zagar *et al.*, 1997), mais c'est une préférence pour N2 qui est mesurée avec les temps de lecture (Baccino *et al.*, 2000).

La fréquence d'occurrence relative de N1 et N2, manipulée expérimentalement ou variant de façon non contrôlée (Baccino *et al.*, 2000 ; Frenck-Mestre & Pynte, 2000b ; Pynte & Colonna, 2000), mais aussi l'animation (Cuetos & Mitchell, 1988 ; Desmet *et al.*, 2006 ; Gilboy *et al.*, 1995) sont à l'origine de différences d'attachement.

La variabilité qui existe au niveau du matériel utilisé dans ces expérimentations peut contribuer de façon importante aux résultats observés. Ainsi, on pourrait se demander si la variabilité translinguistique est liée à des préférences d'attachement différentes selon les langues, ou bien à des configurations expérimentales différentes selon les expériences. Cependant, des expériences telles que celles de Cuetos et Mitchell (1988) ou Gilboy *et al.* (1995) contrôlent le matériel d'une langue à l'autre. Cuetos et Mitchell (1988) réalisent deux questionnaires, en espagnol et en anglais. Les phrases en anglais sont traduites de l'espagnol, mais un facteur de variation supplémentaire est ajouté. En effet, si toutes les phrases en espagnol contiennent le relativiseur *que*, indépendamment de l'animation de N1 et N2, il n'en va pas de même pour les phrases en anglais. C'est *that* qui est employé quand N1 est non-animé et N2 animé humain et *who* quand N1 et N2 sont animés. Selon Cuetos et Mitchell (1988), la présence de *that* entraîne un biais pour attacher la relative à l'entité non-animée, N1. C'est la raison pour laquelle une telle préférence est alors mesurée, alors que les résultats avec *who* confirment la validité de la stratégie du *late closure* prédite par la théorie du *garden path*. C'est d'ailleurs sur cette version que les auteurs concluent sur les préférences d'attachement en anglais. En fait, selon les variétés de l'anglais, *that* peut avoir comme

antécédent une entité humaine (Romaine, 1982, cité par Traxler *et al.*, 1998). C'est par exemple le cas dans certaines phrases du questionnaire de Gilboy *et al.* (1995) où, quand N1 est non-animé, N2 humain et la proposition relative introduite par *that*, on observe 53 % d'attachements vers N2. La préférence n'est certes pas très importante, mais on ne retrouve pas le biais pour l'entité non-animée avancé par Cuetos et Mitchell (1988) ; d'autres facteurs pourraient être à l'origine de cette préférence d'attachement.

La fréquence de N1 et N2 a un effet sur les préférences d'attachement, tout comme le type de nom employé (nom propre, de métier ou de parenté, voir Frenck-Mestre & Pynte, 2000a). Pour les noms propres, la restrictivité est une source de variabilité supplémentaire, puisqu'on ne peut attacher à un nom propre qu'une proposition relative non restrictive. Les théories sur le traitement des anaphores qui ont été présentées dans le Chapitre 2, *informational load hypothesis*, *accessibility* et *givenness hierarchy*, montrent que ces facteurs, ainsi que la définitude (Gundel *et al.*, 1993 ; Holmes, 1995), ou le rôle grammatical (théorie de la proéminence du discours et la *centering theory*) ont un effet sur la référence successive aux entités du discours, sur leur accessibilité. On peut supposer que ces facteurs peuvent influencer les préférences d'attachement des relatives. Par ailleurs, notons que l'étude de Baltazart et Kister (2000) montre un effet de la définitude des antécédents disponibles sur les préférences d'attachement de propositions relatives.

Un relativiseur est une expression anaphorique : il renvoie à une entité du discours appelé antécédent. Si un traitement syntaxique est mis en place, il y a aussi un traitement anaphorique qui lie l'expression référentielle, c'est-à-dire le relativiseur, à une entité du discours (Hemforth *et al.*, 2000). Les facteurs susceptibles d'influencer le traitement anaphorique peuvent être alors appliqués aux cas d'ambiguïtés syntaxiques qui nous intéressent. Les théories sur le traitement des anaphores et les résultats des expériences sur les préférences d'attachement doivent être analysés de concert. Outre les théories syntaxiques, les théories sur le traitement des anaphores pourraient expliquer les variations de préférences d'attachement.

Les expériences testant les prédictions de la *centering theory* ou celles de la théorie de la proéminence du discours montrent qu'une anaphore pronominale est préférée (cf. *repeated name penalty*) en général lorsqu'elle réfère à l'entité proéminente. En ce qui concerne l'établissement de la hiérarchie des expressions référentielles, ces études ont montré un effet de la fonction syntaxique (Gordon & Chan, 1995 ; Gordon *et al.*, 1993 ; M. A. Walker *et al.*, 1994), du caractère ancien de l'entité la plus proéminente (Strube & Hahn, 1999), ou du rôle

thématique (J. E. Arnold, 2001). La différence de saillance, de proéminence entre les entités du discours en influence la reprise anaphorique, sous la forme d'un pronom ou d'une anaphore nominale. Ceci pourrait en partie expliquer les variations observées dans les études sur les préférences d'attachement de propositions relatives ambiguës.

La saillance des antécédents apparaît comme un facteur qui contribue de façon importante aux préférences d'attachement ; elle varie par exemple selon la fréquence de N1 et N2 (Frenck-Mestre & Pynte, 2000b ; Pynte & Colonna, 2000), leur fonction syntaxique, mais aussi selon l'animation. Il y a plus d'attachements vers l'entité la plus saillante. N1 remplissant la fonction objet, sa proéminence (ou saillance) est plus importante que celle de N2, génitif, et il sera plus fréquemment repris par un pronom. Dans le cas qui nous intéresse, on s'attend à ce que se soit avec un relativiseur. L'animation d'un antécédent en fait une entité discursive proéminente dans le modèle du discours : cela peut expliquer la préférence pour attacher une relative à une telle entité (Desmet *et al.*, 2006).

Les travaux d'Ariel (1999) confortent l'idée selon laquelle un relativiseur est un marqueur anaphorique puisque l'emploi d'une forme de relativiseur, anaphore zéro ou pronom résomptif, est influencé par des facteurs qui guident la forme des expressions référentielles dans le discours. De fait, un relativiseur peut être placé sur une hiérarchie d'accessibilité comme celle de Gundel *et al.* (1993) ou celle d'Ariel (1990, 2001). Un relativiseur, comme un pronom, se situe haut dans la hiérarchie et signale un antécédent très accessible. Dans l'étude de corpus d'Ariel (1999), une anaphore zéro, peu informative et atténuée au maximum, sert à faire référence à un antécédent accessible alors qu'un pronom résomptif est utilisé dans les cas où l'antécédent est moins accessible.

Dans le prolongement de l'hypothèse de Hemforth *et al.* (2000), et en me fondant sur des théories du traitement des anaphores, je propose que la forme du relativiseur sert à signaler l'accessibilité de l'antécédent et que cela peut guider le choix du site d'attachement d'une proposition relative, à N1 ou à N2, selon leur niveau d'accessibilité.

Comme la forme d'une anaphore, celle du relativiseur indique l'accessibilité de l'entité à laquelle il fait référence et par conséquent le type d'attachement à réaliser. Un relativiseur qui signale un antécédent hautement accessible sera de préférence attaché à N1, plus accessible que N2. En revanche, quand le relativiseur signale un antécédent moins accessible, il sera plus fréquemment attaché à N2. Je présenterai plus loin en quoi cette hypothèse permet d'expliquer les préférences d'attachement, après avoir repris certains aspects des théories sur le traitement des anaphores nécessaires pour comprendre cette hypothèse.

En effet, la forme d'une expression référentielle signale le niveau d'accessibilité de cette dernière, assimilable à une saillance des entités. Selon la *givenness hierarchy*, la forme de l'expression référentielle traduit le statut cognitif de l'entité à laquelle il est fait référence et, selon l'*accessibility hierarchy*, il s'agit de l'accessibilité de cette entité.

Une anaphore pronominale signale le plus souvent une entité déjà présente dans le discours. Elle peut faire référence à une entité de façon non ambiguë (à l'entité la plus saillante notamment), mais un pronom peut parfois faire référence à plusieurs entités. Dans ce cas, le marquage grammatical, par exemple en genre et nombre, n'est pas nécessairement suffisant au niveau du pronom pour pouvoir l'attacher au bon antécédent.

C'est le cas dans l'exemple (57) où le pronom *her*, féminin singulier, peut faire référence à *Sophie* ou *Cathy*.

(57) *Sophie helped Cathy to prepare the food and Simon chatted to her.*

Pour identifier le référent de ce pronom, des stratégies interprétatives sont mises en place (p. ex. hiérarchie des entités du discours, parallélisme des fonctions, proéminence). Si on se réfère à la hiérarchie proposée par Ariel (1990, 2001), la forme du pronom à l'oral, accentuée ou non, pourrait indiquer un niveau d'accessibilité plus ou moins important de l'entité à laquelle ce pronom fait référence et permettre ainsi de lever l'ambiguïté d'interprétation. Je propose qu'il en va de même pour les relativiseurs : leur forme indique à quel antécédent un relativiseur doit être attaché, grâce au niveau d'accessibilité qu'il signale.

La forme du relativiseur et son influence sur les préférences d'attachement peuvent expliquer les variations d'interprétation pour une même langue, mais aussi entre les langues. Avant de présenter plus en détails la problématique qui émerge ici, ainsi que les hypothèses qui en découlent, revenons sur l'étude d'Ariel (1999) et sur la proposition de Frazier et Clifton (1997).

Ariel (1999) étudie la distribution des relativiseurs en hébreu (anaphore zéro ou pronom résomptif). Elle montre qu'une anaphore plus informative, le pronom résomptif, se situe plus bas sur la hiérarchie d'accessibilité (Ariel, 1990, 2001) et est utilisée de façon préférentielle quand l'accessibilité de l'antécédent dans la proposition principale diminue. Cependant, on pourrait considérer que la distribution inégale de ces deux relativiseurs, anaphore zéro (fonction sujet) et pronom résomptif (autres fonctions), rend leur comparaison difficile. Si, par contre, on considère que l'accessibilité signalée par les relativiseurs a été grammaticalisée en hébreu, cela explique pourquoi selon la norme c'est la forme la plus réduite (anaphore zéro)

qui est obligatoire quand l'antécédent est hautement accessible, alors que c'est une forme moins réduite (pronom résomptif) qui est nécessaire quand l'accessibilité de l'antécédent est moindre.

Une évaluation plus satisfaisante de la forme du relativiseur et de l'accessibilité signalée par ce dernier est possible si l'on compare deux systèmes de relativiseurs dont la distribution est égale selon la grammaire normative, c'est-à-dire qu'ils peuvent être utilisés indifféremment.

Frazier et Clifton (1997) font un rapprochement entre l'interprétation des pronoms dans le discours et les relativiseurs. L'interprétation d'un pronom, c'est-à-dire l'identification de l'entité à laquelle il fait référence, est basée sur la proéminence des entités du discours. De même, l'interprétation d'un pronom relatif, qui ressemble plus à un pronom qu'un complémenteur, est guidée par les critères de l'interprétation d'un pronom, c'est-à-dire par la proéminence des entités du discours.

Les auteurs émettent l'idée que les différences translinguistiques de préférences d'attachement peuvent être expliquées par la forme du relativiseur. Les langues qui utilisent des pronoms relatifs (p. ex. allemand) sont plus susceptibles d'avoir recours à la proéminence des entités du discours dans l'attachement des propositions relatives, d'où une préférence pour N1. L'anaphore, *i. e.* le pronom relatif, est interprétée de la même façon qu'un pronom. *A contrario*, un relativiseur qui a perdu son statut pronominal est moins sensible à l'effet de la proéminence. La forme du relativiseur aurait un effet sur l'attachement des propositions relatives.

Prenons une langue comme l'allemand, qui a pour relativiseurs des pronoms relatifs, marqués quant à la fonction syntaxique du relativisé dans la proposition relative, mais aussi au niveau du genre et du nombre de l'antécédent. L'interprétation et l'attribution de l'antécédent de ce pronom relatif seraient ainsi sujettes à l'effet de la proéminence du discours. Comme un pronom fait de préférence référence à l'entité proéminente, la plus accessible, c'est donc une préférence pour N1 qui est observée quand une proposition relative est précédée par une structure « N1 de N2 » (Hemforth *et al.*, 2000). Si cette explication proposée par Frazier et Clifton (1997) paraît intéressante, est-elle aussi correcte pour les données observées en anglais ?

Les phrases contenant *who* dans la version anglaise du questionnaire présenté par Cuetos et Mitchell (1988) sont caractérisées par une préférence pour N2, comme pour les phrases utilisées par Carreiras et Clifton (1999). Ce relativiseur étant un pronom relatif, son

attachement est guidé par la proéminence des entités du discours et devrait alors être attaché à l'antécédent le plus accessible, N1. Or, tel n'est pas le cas.

De façon assez remarquable, c'est pour les phrases contenant *that*, qui est un complémenteur et non un pronom relatif (Seppänen, 1999 ; Stahlke, 1976), qu'on peut observer une préférence pour N1 dans les données de Cuetos et Mitchell (1988), ou bien une absence de préférence (Traxler *et al.*, 1998). L'attachement dans ce cas devrait être guidé non pas par la proéminence des entités du discours, mais par le principe de *late closure* qui devrait conduire à un attachement à N2. Ces résultats, dans leur ensemble, sont contraires à ce qui est proposé par Frazier et Clifton (1997). Afin d'expliquer ces résultats, Cuetos et Mitchell (1988, voir aussi Bayley, 1999 et Guy & Bayley, 1995) proposent l'existence d'un biais pour attacher *that* à un antécédent non-animé (ici N1).

On ne peut cependant rejeter la suggestion de Frazier et Clifton (1997) sur la base unique de ces données en anglais. En ce qui concerne la relativisation et la forme du relativiseur employé dans les variétés de l'anglais, il existe en plus de nombreuses variations (Bayley, 1999 ; Kortmann, Burrigge, Mesthrie, Schneider, & Upton, 2004) qui ne permettent pas de jugement tranché sur les préférences d'attachement en fonction du relativiseur. On peut aussi considérer qu'il existe deux stratégies de relativisation différentes selon que le relativiseur est de la forme *wh-* ou *that* (Seppänen, 1999).

La question de l'effet de la forme du relativiseur sur les préférences d'attachement reste entière. Dans la partie suivante, je vais proposer une autre interprétation des différences translinguistiques et intra-linguistiques, pour laquelle je formulerai une série d'hypothèses qui seront testées expérimentalement.

2. Problématique et hypothèse théorique

Depuis de nombreuses années, des théories ont étudié le traitement des ambiguïtés syntaxiques, se heurtant à une variabilité translinguistique pour un cas particulier : l'attachement des propositions relatives. Diverses explications ont été proposées, mais n'ont pas nécessairement pris en compte, dans la validation expérimentale de leurs hypothèses, des aspects linguistiques qui peuvent affecter les préférences d'attachement.

Dans la section précédente on a vu que la variabilité du matériel utilisé dans les expériences en compréhension pouvait contribuer à la variation des préférences d'attachement, translinguistique et intra-linguistique. Le lien entre N1 et N2, leur fréquence ou leur modifiabilité (parmi d'autres aspects) influencent les préférences d'attachement. On

peut trouver une explication de la variabilité des préférences d'attachement dans les études sur le traitement des anaphores. En effet, ces dernières ont mis en évidence l'importance de divers facteurs sur l'accessibilité des entités du discours, ce qui a un effet sur leur reprise possible dans la suite du discours.

L'attachement de la proposition relative à N1 ou N2 serait ainsi influencé par la différence d'accessibilité entre les deux entités auxquelles N1 et N2 font référence. Cependant, cette interprétation est à l'heure actuelle plus un questionnement qu'une affirmation.

Un autre aspect pourrait contribuer aux variations au sein d'une même langue et permettre une meilleure compréhension de la variation translinguistique : il s'agit de la forme du relativiseur (voir aussi à ce sujet Desmet, Brysbaert *et al.*, 2002). En effet, si l'on reprend les études réalisées en anglais, on observe une différence dans l'attachement selon la forme du relativiseur utilisé. Cette dernière pourrait influencer les préférences d'attachement. On mesure en effet une préférence pour N2 quand la proposition relative est introduite par *who* (une partie du matériel de Cuetos & Mitchell, 1988, mais aussi Carreiras & Clifton, 1999). Par contre, une préférence pour N1 (l'autre partie du matériel de Cuetos & Mitchell, 1988) ou une absence de préférence (Traxler *et al.*, 1998) est mesurée avec une proposition relative introduite par *that*. On pourrait se demander si la forme du relativiseur, plus qu'un biais de l'animation (tel qu'il est proposé par Cuetos & Mitchell, 1988), ne serait pas liée à un niveau d'accessibilité signalé, comme c'est le cas pour les autres expressions anaphoriques. Les résultats de ces expériences ne permettent pas d'affirmer cette idée.

En effet, les hiérarchies d'accessibilité ont montré que la forme d'une expression référentielle signale un niveau d'accessibilité de l'entité à laquelle elle fait référence. De la même façon, on pourrait considérer que la forme du relativiseur, qui est aussi une expression anaphorique, signale le niveau d'accessibilité de son antécédent.

Problématique Ainsi, et de façon plus générale, on peut se poser la question de savoir si les préférences interprétatives sont guidées par les propriétés des constituants de la phrase : l'attachement des propositions relatives est influencé par les propriétés, c'est-à-dire l'accessibilité signalée par le relativiseur et ses antécédents potentiels. C'est le questionnement qui est à l'origine de cette recherche.

Hypothèse théorique : D'après les théories de l'accessibilité, le niveau d'accessibilité de toute entité du discours est signalé par la forme de l'expression référentielle utilisée pour y faire référence. Plus le niveau d'accessibilité est élevé, plus la probabilité pour cette entité d'être reprise dans la suite du discours est importante. Par conséquent, la forme d'une anaphore est d'autant moins réduite et plus informative que cette anaphore fait référence à une entité peu accessible.

À la question soulevée dans la problématique, je propose l'hypothèse théorique suivante : la forme du relativiseur, qui est un marqueur anaphorique, signale le niveau d'accessibilité de son antécédent. Un relativiseur qui est moins atténué et plus informatif se situe à un niveau plus bas dans la hiérarchie d'accessibilité qu'un relativiseur qui est plus atténué et moins informatif et signale un antécédent peu accessible. Dans le cas des ambiguïtés d'attachement des propositions relatives, la forme du relativiseur guide l'attachement, à N1 ou à N2 pour une structure de la forme « N1 de N2 », selon le niveau d'accessibilité de ces antécédents.

Cette hypothèse n'est pas nouvelle dans l'absolu, mais son application dans le cadre des ambiguïtés d'attachement des propositions relatives est innovante et permet de rendre compte des variations translinguistiques et intra-linguistiques observées dans la littérature. De cette hypothèse théorique en découlent d'autres à propos de l'effet de la forme du relativiseur et des antécédents auxquels la proposition relative peut être attachée.

Cependant, avant de présenter les hypothèses opérationnelles qui vont permettre de tester cette hypothèse théorique, je vais préciser l'objet de mon étude.

3. L'opérationnalisation des hypothèses

L'hypothèse théorique est appliquée à une structure particulière où l'attachement d'une proposition relative est ambigu car deux sites d'attachement, deux antécédents sont possibles, comme dans (18).

(18) *The journalist interviewed the daughter of the colonel who had had the accident.*

Les théories de l'accessibilité postulent que la forme d'une expression référentielle signale le degré d'accessibilité de l'entité à laquelle il est fait référence. Dans le cas des relatives, on pourrait proposer que la forme des relativiseurs et des antécédents potentiels signale le degré d'accessibilité des référents et qu'elle peut par conséquent guider l'attachement de la relative. Dans ce qui suit, j'étudierai séparément de ces deux types. Mais avant cela, je vais revenir sur

des restrictions méthodologiques qui conditionnent l'élaboration des hypothèses opérationnelles, le choix d'une langue, le français et le type de proposition relative, les non restrictives.

3. 1. Objet étudié et restrictions méthodologiques

3. 1. 1. Restriction de l'étude à une seule langue

Le premier choix qui s'impose est celui d'une langue donnée, disposant de plusieurs stratégies de relativisation : le français. Cette restriction pourrait mettre ce travail en porte-à-faux par rapport à la critique qui a pu être apportée sur les travaux traitant des ambiguïtés d'attachement. La validation translinguistique permet de conforter l'argument de la portée universelle des hypothèses proposées, ce que l'observation d'une langue unique ne peut pas.

Cependant, on a vu dans les travaux de Cuetos et Mitchell (1988) qu'il est difficile d'adapter de façon fidèle le matériel d'une langue à l'autre. Se focaliser sur une seule langue qui justement dispose de plusieurs stratégies de relativisation rend le contrôle des conditions plus facile que de façon translinguistique.

De plus, que des langues comme le français, l'anglais, l'hébreu ou l'espagnol disposent de plusieurs stratégies de relativisation conforte l'idée qu'une utilisation de ces stratégies n'est pas une exception dans les langues du monde.

Plusieurs choix méthodologiques sont nécessaires : une langue, un type de proposition relative et la limitation de l'observation à quelques aspects des antécédents.

3. 1. 2. Le choix du français

Mon choix s'est porté sur le français, qui dispose de deux systèmes de relativiseurs et permet de tester l'hypothèse posée. Ce sont des relativiseurs comme *qui* ou *que* qui sont le plus souvent employés. C'est pourquoi ils sont aussi utilisés dans les expérimentations sur le traitement des propositions relatives en français (Baccino *et al.*, 2000 ; Frenck-Mestre & Pynte, 2000b ; Zagar *et al.*, 1997).

L'autre système, celui des *lequel*, est considéré plus formel (Jones, 1996 ; Riegel *et al.*, 2004). Ces relativiseurs sont plus informatifs que les *qui* : ils sont marqués au niveau du genre (*lequel, laquelle*) et du nombre (*lesquels, lesquelles*), ce qui pourrait aider à identifier l'antécédent si plusieurs sont possibles et qu'il y a une ambiguïté d'attachement. Si l'on se réfère aux hiérarchies d'accessibilité, on peut considérer que des relativiseurs comme *lequel*

sont situés plus bas dans la hiérarchie que des relativiseurs comme *qui*, parce qu'ils sont plus informatifs et marqués en genre et en nombre.

De façon générale, on peut considérer que, pour une fonction syntaxique donnée, les *qui* sont plus courts et moins informatifs que les *lequel*, à l'exception de *à qui* et *auquel* qui ont la même taille phonologique, et de la fonction objet où aucun relativiseur comme *lequel* n'est disponible). On peut ainsi considérer que, placés sur une hiérarchie d'accessibilité, les *lequel* sont situés plus bas que les *qui* et signalent ainsi une entité moins accessible. Le choix du français semble donc pertinent pour tester les hypothèses qui vont être maintenant proposées.

3. 1. 3. Le type de structure étudiée

Comme ce travail a pour but d'estimer les préférences d'interprétation pour une proposition relative dont l'attachement est ambigu, c'est une phrase semblable à (58) qui est retenue.

(58) *Je connais le père du maçon, qui est amusant.*

Si l'on s'intéresse à l'accessibilité des entités du discours (Ariel, 1990, 2001 ; Gundel *et al.*, 1993), on peut considérer que N1, dans (58) *père* est plus accessible que N2, *maçon* parce qu'il est la tête du GN, N2 étant à un niveau plus profond de la structure syntaxique.

Comme on l'a vu dans l'Introduction, il existe deux types de propositions relatives : les restrictives et les non restrictives (Jones, 1996 ; Riegel *et al.*, 2004). Tout comme certains considèrent qu'en anglais les propositions relatives avec *that* sont nécessairement restrictives (Gibson *et al.*, 2005), en français les propositions relatives introduites par des *lequel* sont non restrictives. Le degré de restrictivité des relatives introduites par des *qui* peut être déterminé par la ponctuation (Riegel *et al.*, 2004 ; Tellier, 1996), mais aussi par des critères sémantico-pragmatiques. La restrictivité est contrôlée pour comparer ces deux formes de relativiseurs : seules les propositions relatives non restrictives sont retenues.

Choisir ce type de relative a un avantage supplémentaire pour l'évaluation de l'effet de l'accessibilité signalée par la forme des antécédents. On a vu qu'une proposition relative restrictive ne peut, selon la norme, être attachée à un nom propre (Frenck-Mestre & Pynte, 2000b ; Jones, 1996, ou la critique de Baccino *et al.*, 2000). De tels attachements sont possibles avec une proposition relative non restrictive, ce qui nous permettra d'étudier des niveaux d'accessibilité relativement fins pour les antécédents (voir la hiérarchie proposée par Gundel *et al.*, 1993).

3. 1. 4. Les sources de variation étudiées

L'aspect central de ce travail est l'évaluation de l'effet de la forme du relativiseur sur les préférences d'attachement. Je propose, en relation avec les théories de l'accessibilité, que le relativiseur signale l'accessibilité de l'antécédent. Ce sera le premier aspect étudié.

De plus, comme on l'a vu dans la littérature, divers aspects des antécédents peuvent influencer les préférences d'attachement. Je propose que le niveau d'accessibilité des antécédents signalé par la forme de l'expression référentielle peut varier en fonction de plusieurs facteurs : la modification d'un nom par un modifieur, la compétition entre antécédents, mais aussi leur animation, leur forme ou encore leur fonction syntaxique. On peut aussi proposer que la longueur de la relative sert à signaler l'accessibilité de l'antécédent : une proposition relative longue fournit plus d'informations et une entité récemment introduite dans le discours est plus susceptible d'être l'objet de ces informations.

3. 1. 4. 1. Le relativiseur

Au niveau du relativiseur, je comparerai les deux formes de relativiseurs pour une même fonction syntaxique, afin que la seule source de variation soit la forme du relativiseur. L'analyse portera sur la fonction sujet (*qui* d'un côté, *lequel* et *laquelle* de l'autre) et sur la fonction objet indirect²⁷ (*à qui*, *auquel* et *à laquelle*). Comme il s'agit de l'étude des préférences d'attachement de propositions relatives, ce sont des phrases comme (58a-b), avec une structure de la forme « N1 de N2 », qui sont étudiées.

(58) a. *Je connais le père du maçon, qui est amusant.*

b. *Je connais le père du maçon, lequel est amusant.*

Entre *qui* et *lequel/laquelle*, deux aspects changent : l'**atténuation** (ou la taille phonologique), l'**informativité** (marquage en genre et nombre pour *lequel/laquelle*) et potentiellement la rigidité (grâce au marquage en genre et nombre, voir Chapitre 2 et Ariel, 2001). *Lequel* et *laquelle* sont moins atténués, plus informatifs et potentiellement plus rigides que *qui*, car ils sont plus longs (deux syllabes vs une seule), marqués en genre et en nombre et peuvent potentiellement identifier un référent unique (cf. rigidité), si le genre et le nombre marqué correspondent à un seul référent. *Lequel* et *laquelle* se situent alors à un niveau plus

²⁷ Désormais abrégée OI.

bas de la hiérarchie d'accessibilité d'Ariel (1990, 1999, 2001) et signalent une entité du discours moins accessible que *qui*.

En ce qui concerne la fonction OI, avec *à qui* et *auquell/à laquelle*, il y a toujours une différence d'informativité due au marquage en genre et en nombre pour *auquel* et *à laquelle*, mais *à qui* fournit aussi une information supplémentaire par rapport à *auquel* et *à laquelle* : il marque le caractère forcément animé de son antécédent, alors que *auquel* et *à laquelle* peuvent être attachés à un antécédent animé ou non. L'observation d'une différence entre ces deux formes de relativiseurs indique que le marquage en genre et nombre ou le marquage de l'animation situent les relativiseurs sur deux niveaux différents de la hiérarchie et que cela entraîne des attachements différents.

Le caractère animé des entités du discours joue un rôle important aussi bien en compréhension que dans la production du langage (Bock & Levelt, 1994 ; Branigan, Pickering, & Tanaka, 2008; G. R. Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, 2006). D'autres études montrent aussi que le marquage en genre et en nombre contraint de façon rapide l'interprétation (J. E. Arnold, Eisenband *et al.*, 2000 ; Van Berkum, Brown, & Hagoort, 1999b). Je propose que ces deux types de marquage sont informatifs, mais qu'*a priori* ils situent les deux relativiseurs, *à qui* et *auquel*, à un même niveau de la hiérarchie d'accessibilité. Il n'y aura pas de différence d'attachement signalée.

Le niveau d'atténuation entre *auquel* et *à qui* est comparable, puisque les deux relativiseurs font deux syllabes, alors que *à laquelle* est moins atténué que *qui*. Si l'atténuation est un facteur déterminant dans l'accessibilité signalée par les relativiseurs, une différence d'atténuation devrait conduire à des préférences d'attachement différentes et l'on pourrait alors s'attendre à une différence entre *à qui* et *à laquelle*. Si, au contraire, la différence d'atténuation n'est pas suffisante pour signaler un niveau d'accessibilité et un attachement différents, on devrait alors observer des résultats comparables entre *à qui* et *auquel* et entre *à qui* et *à laquelle*.

3. 1. 4. 2. Les antécédents

À la suite d'études sur le traitement des anaphores, différents facteurs ont été proposés pour expliquer le niveau d'accessibilité ou de saillance des référents du discours.

Il y a tout d'abord la **distance** (Ariel, 1990, 2001 ; Givón, 1983, 1990) entre les deux références successives, qui fait diminuer le niveau d'accessibilité de l'entité référée, mais aussi l'**ordre des mots** et la possibilité d'utiliser des dislocations (Givón, 1983, 1990).

Étant donné la structure utilisée, ces deux aspects ne varient pas²⁸. L'ordre des deux antécédents est strict et la proposition relative suit N2, sauf si un adjectif modifiant N1 ou N2 est intercalé. N2 reste toujours plus proche de la proposition relative que N1.

La **compétition** (Ariel, 1990 ; J. E. Arnold & Griffin, 2007 ; O'Brien & Albrecht, 1991 ; Toole, 1996) entre les entités du discours a été avancée comme un facteur influençant l'accessibilité. Plus il y a de référents dans le discours, plus il y a de compétition et plus leur niveau d'accessibilité est bas. Pour la structure qui nous concerne, les relatives, les référents en compétition sont les entités désignées par N1 et N2 qui sont les deux sites où la proposition relative peut être attachée. Le relativiseur peut faire référence aux deux entités référées par N1 et N2. Comme seule la structure de la forme « N1 de N2 » est étudiée, la compétition est constante et s'effectue toujours entre ces deux entités du discours.

On pourrait aussi ajouter que la saillance, signalée par la prosodie et l'accentuation (Ariel, 1990, 2001 ; Givón, 1983), affecte les préférences d'attachement. Ceci a en effet été mis en évidence par Schafer *et al.* (1996). Cependant, comme ce travail est basé sur l'écrit, de telles informations ne sont pas disponibles et il n'est pas possible d'étudier l'effet de la saillance. Notons que, d'après Fodor (1998), le lecteur établit des frontières prosodiques dans des tâches de compréhension, notamment écrites (Gilboy & Sopena, 1996), sur lesquelles il se fonde pour interpréter la structure en cours d'analyse (Kjelgaard & Speer, 1999).

D'autres facteurs sont susceptibles d'affecter le niveau d'accessibilité de N1 et N2 et vont être évalués par la suite. Il s'agit tout d'abord de l'**animation** (Dahl & Fraurud, 1996 ; Givón, 1990). On a vu dans l'étude de Desmet *et al.* (2006) que l'animation influence les préférences d'attachement, entraînant plus d'attachements vers l'entité animée humaine. Je propose que l'effet de l'animation est lié au niveau d'accessibilité des antécédents : il y a plus d'attachements vers l'entité humaine animée que vers l'entité non-animée, car la première, par son animation même, est plus accessible, plus proéminente que la seconde (pour une discussion sur la saillance et l'animation, voir Bock & Irwin, 1980 ; Bock & Levelt, 1994 ; Branigan *et al.*, 2008).

La **forme de l'expression référentielle** est un aspect important dans les hiérarchies d'accessibilité (Ariel, 1990 ; Givón, 1992 ; Gundel *et al.*, 1993) et peut par conséquent affecter les préférences d'attachement. Un GN avec un article indéfini signale un référent

²⁸ Comme la distance et l'ordre des mots ne varient pas, aucune hypothèse n'est formulée à ce sujet. Il en va de même pour la compétition.

moins accessible qu'un GN avec un article défini, qui fait déjà partie du discours. Entre ces deux niveaux, Ariel considère qu'un prénom signale une entité du discours plus accessible qu'un GN avec l'article défini (Ariel, 2001), ce qui contredirait les résultats de Frenck-Mestre et Pynte (2000a, 2000b), bien que la question de la restrictivité, soulevée par Baccino *et al.* (2000), ne permette pas de conclure de façon certaine sur les résultats alors obtenus (voir Chapitre 1, section 2.1.1.1).

La **fonction syntaxique** a aussi un effet sur l'accessibilité ou la proéminence des entités du discours. La *centering theory* propose qu'une anaphore pronominale, qui fait référence à l'entité la plus proéminente à un moment donné, est préférée quand elle réfère au GN qui remplit la fonction sujet dans la proposition précédente. Si l'on reprend la hiérarchie proposée par Givón (1990), la fonction sujet est plus topique, plus accessible que la fonction objet direct, indirect ou oblique. Notons cependant que dans la présentation de sa théorie, Ariel (1990, 2001) minimise l'importance de la fonction syntaxique (pour une mise en évidence expérimentale, voir Halmari, 1996).

Les facteurs qui viennent d'être présentés vont avoir une influence sur l'accessibilité signalée par la forme des expressions référentielles, relativiseurs comme antécédents. Je formule à ce sujet une série d'hypothèses opérationnelles qui constituent un test de l'hypothèse théorique selon laquelle, rappelons-le, plus une anaphore fait référence à une entité peu accessible, moins la forme de cette anaphore est réduite et plus elle est informative.

3. 2. Les hypothèses opérationnelles

Dans la section précédente, j'ai rappelé trois aspects qui d'après Ariel (1990, 2001) varient en fonction du niveau d'accessibilité signalé par l'expression référentielle : atténuation, informativité et rigidité. Comme le dernier ne varie pas dans la structure étudiée, je pose une hypothèse sur l'effet de l'atténuation et celui de l'informativité, à propos de la forme du relativiseur.

Hypothèse 1a, effet de l'atténuation : quand deux types de relativiseurs sont disponibles, par exemple *qui* et *lequel*, le relativiseur le moins atténué (*lequel*) sera attaché de façon plus importante à l'entité du discours, moins accessible que le relativiseur qui est plus atténué (*qui*). Avec le type de structure étudié ici, je m'attends à ce que l'emploi de *lequel* au détriment de *qui* entraîne plus d'attachements vers N1 dans (58b) que *qui* dans (58a), car l'entité référée par N2 est moins accessible que celle référée par N1.

(58) a. *Je connais le père du maçon, qui est amusant.*

b. *Je connais le père du maçon, lequel est amusant.*

Hypothèse 1b, effet de l’informativité : de la même façon, l’emploi d’un relativiseur plus informatif, *lequel*, signale un attachement à l’entité moins accessible, toujours N2 dans (58).

Ces deux hypothèses seront évaluées en comparant des propositions relatives sujet, introduites par *qui* ou *lequel*, ce qui nous procurera un test de l’Hypothèse 1. Le relativiseur *lequel* est à la fois moins atténué et plus informatif que *qui*, ce qui laisse prédire un effet cumulatif de ces deux facteurs et devrait conduire à une différence des préférences d’attachement.

Cependant, la comparaison entre *qui* et *lequel* ne permet pas de distinguer les deux volets de l’Hypothèse 1 sur l’effet de la forme du relativiseur, l’Hypothèse 1a et l’Hypothèse 1b, puisque atténuation et informativité varient de concert. La comparaison des relativiseurs pour les relatives OI est alors intéressante. En effet, le degré d’atténuation est différent entre *à qui* et *auquel*, mais non entre *à qui* et *à laquelle*. Si l’Hypothèse 1a est correcte, on devrait observer une différence dans les préférences d’attachement entre *à qui* et *à laquelle*, mais non entre *à qui* et *auquel*, avec plus d’attachements à N2 pour *à laquelle* qu’avec *à qui*. L’étude des relatives OI permettra aussi d’évaluer l’Hypothèse 1b. D’un côté *à qui* est plus informatif que *auquel* et *à laquelle* parce qu’il marque l’animation de l’antécédent et, de l’autre côté, ces derniers sont plus informatifs car marqués en genre et en nombre. De fait, tous ces relativiseurs auraient un niveau d’informativité équivalent. Si c’est la différence d’informativité qui contribue à la différence attendue entre *qui* et *lequel*, que le degré d’informativité soit équivalent entre *à qui*, *auquel* et *à laquelle* ne devrait pas mener à des attachements différents.

En ce qui concerne les antécédents, j’ai sélectionné cinq facteurs qui vont influencer le niveau d’accessibilité signalé, ce qui affecte l’attachement de la proposition relative. Je propose une hypothèse pour chacun.

Hypothèse 2a, effet de la longueur de la proposition relative : Fodor (1998) propose que la longueur de la proposition relative a un effet sur son attachement (cf. *sausage machine*, Frazier & Fodor, 1978). Une proposition relative longue est attachée préférentiellement à N1, une courte à N2.

Si on prend le critère de l'accessibilité comme guidant l'attachement des relatives, on peut poser une hypothèse différente. Une proposition relative longue fournit plus d'informations sur son antécédent qu'une courte. Une entité du discours moins accessible, sur laquelle on dispose de moins d'informations, est alors plus susceptible de recevoir des informations nouvelles qu'une entité très accessible. Selon l'Hypothèse 2a, on s'attend alors à plus d'attachements à l'antécédent moins accessible, N2, avec une relative longue qu'avec une courte.

Hypothèse 2b, effet de la présence d'un modifieur : on l'a vu, le principe de modifiabilité (Thornton *et al.*, 1999) suggère que si N1 est modifié, par un adjectif et par « de N2 », il n'est plus disponible pour être le site d'attachement de la relative. Colonna (2001) invalide cette prédiction en mettant en évidence que la présence d'un adjectif attire en quelque sorte l'attachement au nom qu'il modifie. Selon la hiérarchie d'accessibilité d'Ariel (1990), un GN avec un modifieur (une description selon les termes de l'auteur) est situé plus bas qu'un GN sans modifieur. L'Hypothèse 2a propose que la présence d'un adjectif signale un antécédent moins accessible. Un relativiseur qui signale un antécédent moins accessible, *lequel* vs *qui*, sera attaché plus fréquemment à un antécédent modifié.

Hypothèse 2c, effet de l'animation : une entité humaine peut être réalisée de différentes façons : nom complet, propre, prénom, GN avec article défini, pronom, etc. Pour deux entités, N1 et N2, humaines et animées, de la même forme (GN avec article défini), c'est N1 qui est proéminent. Par conséquent, l'emploi d'un relativiseur qui signale une entité moins accessible fait référence de façon préférentielle à l'entité non-animée, pour peu que les relativiseurs disponibles ne restreignent pas l'animation de l'antécédent. Cette hypothèse ne peut donc être formulée pour la comparaison entre *à qui* et *auquel/à laquelle*, car l'attachement avec *à qui* est obligatoirement à un antécédent animé. Par contre, on s'attend à plus d'attachements vers l'entité non-animée avec *lequel*, quand N1 et N2 diffèrent quant à l'animation.

Hypothèse 2d, effet de la fonction syntaxique : je suggère que plus la fonction syntaxique de N1 est basse dans la hiérarchie suggérée par Givón (1990), plus l'accessibilité de N1 diminue (cf. Ariel, 1999). On peut émettre l'hypothèse que plus la fonction syntaxique de N1 est basse, plus un relativiseur qui signale un niveau d'accessibilité bas sera attaché à N1. De façon inverse, plus la fonction syntaxique de N1 est élevée dans la hiérarchie, plus un relativiseur signalant un niveau d'accessibilité élevé y sera attaché.

Un relativiseur signalant un niveau d'accessibilité bas (p. ex. *lequel*) sera préférentiellement attaché à N2 quand N1 a une fonction syntaxique élevée (sujet), mais plus

d'attachements vers N1 sont attendus quand ce dernier a une fonction syntaxique plus basse (objet direct, indirect et oblique).

Ces hypothèses vont être testées dans une série d'expériences, une étude de questionnaires, une analyse de corpus écrit et une expérience avec le paradigme d'auto-présentation qui vont maintenant être proposées. L'ensemble des hypothèses peut être résumé dans le Tableau 2. Notons que pour les différents aspects de l'Hypothèse 2, c'est la différence d'attachements entre *qui* et *lequel* qui nous intéresse, plus que la préférence pour *qui*.

Tableau 2 : Récapitulatif des hypothèses formulées

		Effets attendus
Hypothèse 1	a. Atténuation	Moins le relativiseur est atténué, plus il signale un attachement à un antécédent moins accessible, N2, et moins il signale un attachement à un antécédent accessible
	b. Informativité	Plus le relativiseur est informatif, plus il signale un attachement à un antécédent moins accessible, N2
Hypothèse 2	a. Longueur	Il y a plus d'attachements à l'antécédent moins accessible, N2, quand la proposition relative est longue ; une relative courte sera attachée plus fréquemment à un antécédent hautement accessible.
	b. Modifieur	La présence d'un adjectif signale un antécédent peu accessible. Il y a plus d'attachements à cet antécédent quand la relative est introduite par un relativiseur qui signale un niveau d'accessibilité faible.
	c. Animation	Un antécédent animé est hautement accessible et un relativiseur signalant un niveau d'accessibilité élevé y sera attaché plus fréquemment. Un antécédent non-animé, moins accessible, sera plus fréquemment le site d'attachement d'un relativiseur signalant un niveau d'accessibilité plus bas.
	d. Fonction syntaxique	Un relativiseur signalant un degré d'accessibilité élevé sera attaché préférentiellement à N1 quand celui-ci remplit une fonction syntaxique élevée (sujet), alors qu'un relativiseur plus bas dans la hiérarchie est attaché à un antécédent dont la fonction syntaxique est basse (objet direct, indirect ou oblique pour N1, génitif pour N2).

3. 3. Mise à l'épreuve expérimentale

Lorsqu'on s'intéresse au traitement d'ambiguïtés, on a souvent recours à des études de questionnaires. Cette méthode expérimentale a l'avantage d'évaluer les préférences de locuteurs et de donner une première idée des mécanismes mis en jeu lors du traitement de phrases ambiguës. C'est pourquoi je commencerai par une série de questionnaires qui testent l'Hypothèse 1, avec la comparaison de *qui* et *lequel* d'une part, de *à qui* et *auquel* d'autre part.

Si les questionnaires procurent un indice sur les préférences d'attachement, je propose ensuite une étude de corpus qui évalue l'effet de la forme du relativiseur sur les préférences d'attachement en production. Cette analyse de corpus est intéressante dans l'évaluation de la *tuning hypothesis* qui postule que les attachements en production influencent les préférences d'attachement en compréhension. Elle permet aussi de vérifier la validité des hypothèses posées en relation avec les théories de l'accessibilité principalement formulées dans le domaine de la production. Il est alors possible d'évaluer les prédictions de l'Hypothèse 2 et de ses différents volets, d'estimer non seulement si la forme du relativiseur influence les préférences d'attachement, mais aussi si les attachements sont guidés par l'accessibilité des antécédents.

Enfin, je propose une étude en auto-présentation segmentée, qui évalue plus précisément les mécanismes cognitifs dans la compréhension de propositions relatives dont l'attachement est temporairement ambigu. La mesure des temps de lecture et les différences mesurées sont autant d'indices des préférences d'attachement et des difficultés de traitement rencontrées pendant la compréhension.

PARTIE EXPERIMENTALE

Lorsqu'on s'intéresse au traitement d'ambiguïtés, on a souvent recours à des études de questionnaires. Ils évaluent de façon explicite les préférences d'interprétation des locuteurs et fournissent ici un premier aperçu des mécanismes mis en jeu dans le traitement des ambiguïtés d'attachement. Si on obtient des résultats significatifs avec cette méthode d'investigation explicite, il est probable qu'une mesure implicite, comme la mesure des temps de lecture par mot, révèle des différences de comportement selon le type de phrase présentée. Aussi je commencerai par une série de questionnaires, qui testent les deux volets de l'Hypothèse 1, avec la comparaison de *qui* et *lequel* d'une part, de *à qui* et *auquel* d'autre part.

Si les questionnaires procurent un indice sur les préférences d'attachement, je propose ensuite une étude de corpus qui évalue l'effet de la forme du relativiseur sur les préférences d'attachement en production. Cette étude est un apport intéressant à la discussion sur l'utilité de données statistiques en production langagière (Desmet, Brysbaert *et al.*, 2002 ; Gibson & Schütze, 1999 ; Mitchell *et al.*, 1995 ; Rayner & Clifton, 2002) et du lien entre les fréquences d'occurrence en production et les mesures en compréhension. L'étude de corpus, mise en relation avec une étude comportementale en compréhension, permettra aussi d'évaluer la *tuning hypothesis*, qui avance que les attachements en production influencent les préférences

d'attachement en compréhension. Elle permet aussi de vérifier la validité des hypothèses posées en relation avec les théories de l'accessibilité principalement formulées dans le domaine de la production. Il est alors possible d'évaluer les prédictions de l'Hypothèse 2 et de ses différents volets, d'estimer non seulement si la forme du relativiseur influence les préférences d'attachement, mais aussi si les attachements sont guidés par l'accessibilité des antécédents.

Enfin, je propose une étude en auto-présentation segmentée, qui évalue plus précisément les mécanismes cognitifs dans la compréhension de propositions relatives dont l'attachement est temporairement ambigu. La mesure des temps de lecture et les différences mesurées sont autant d'indices des préférences d'attachement et des difficultés de traitement rencontrés pendant la compréhension. Il sera alors possible d'évaluer si la forme du relativiseur guide l'attachement initial des lecteurs et de localiser les endroits de la phrase où des difficultés de traitement peuvent être rencontrées.

Chapitre 1 Questionnaires

La méthode des questionnaires est fréquemment utilisée pour évaluer les préférences d'interprétation de structures ambiguës (Cuetos & Mitchell, 1988 ; Gilboy *et al.*, 1995 ; Hemforth *et al.*, 2000 ; Traxler *et al.*, 1998 ; van Gompel *et al.*, 2001). La série de questionnaires qui est proposée ici teste les deux volets de la première hypothèse opérationnelle (Hypothèse 1a et 1b), selon laquelle la forme du relativiseur influence les préférences d'attachement : on s'attend à plus d'attachements à l'antécédent moins accessible avec *lequel* qu'avec *qui*, car *lequel* signale précisément un antécédent moins accessible. Dans un GN complexe de la forme « N1 de N2 », on considère que N1 est plus accessible que N2 car il est à la tête du GN et N2 est le complément du nom. Il devrait y avoir plus d'attachements à N2 avec *lequel* plutôt qu'avec *qui*. Des phrases comme (59a-b) sont présentées dans le Questionnaire 1.

(59) a. *Je connais le père du maçon, qui est amusant.*

b. *Je connais le père du maçon, lequel est amusant.*

Les résultats de ces questionnaires procurent une estimation de la préférence d'attachement qui est ici explicite : en choisissant leur réponse, les participants signalent quel type d'attachement ils réalisent.

Les premiers questionnaires utilisent une procédure interindividuelle : les participants ne traitent qu'un seul type de relativiseur, afin d'éviter un biais ou une influence mutuelle des relativiseurs (Questionnaire 1 et Questionnaire 2). Selon la grammaire, les deux systèmes de relativiseurs en français sont synonymes et les participants pourraient décider d'une interprétation, d'un attachement, et l'appliquer à toutes les phrases expérimentales, quelle que soit la forme du relativiseur. Ces questionnaires sont répliqués avec une procédure intra-individuelle, où les participants traitent des phrases contenant les deux relativiseurs (Questionnaire 3 et Questionnaire 4). L'avantage de cette procédure intra-individuelle est que la différence statistique mesurée entre les deux formes de relativiseurs ne peut être attribuée à la différence entre les deux groupes de sujets, ce qu'on peut reprocher dans le cadre d'une procédure interindividuelle.

Comme cela a été présenté dans les hypothèses de travail, la comparaison des relativiseurs pour la fonction sujet et OI est intéressante pour tester l'Hypothèse 1. Entre *qui* et *lequel*, il y a plusieurs types de variations. Il y a tout d'abord l'atténuation de l'expression référentielle (une syllabe pour *qui*, deux pour *lequel*), mais aussi l'informativité ou la rigidité. Ces facteurs sont utilisés pour évaluer la position de l'expression référentielle sur l'*accessibility hierarchy* proposée par Ariel (1990) et son effet sur les préférences d'attachement.

Dans la comparaison entre *qui* et *lequel* (Questionnaire 1 et Questionnaire 4), deux facteurs, atténuation et informativité, varient de façon équivalente : ils signalent une position plus basse pour *lequel*, moins atténué et plus informatif. La comparaison entre *à qui* et *auquel* (Questionnaire 2) neutralise l'atténuation et l'informativité. Celle-ci varie selon deux aspects, soit le marquage en genre et en nombre (*auquel* est alors plus informatif), soit le marquage de l'animation (*à qui* est alors plus informatif). Si, comme je le propose, ces marquages respectifs rendent les relativiseurs aussi informatifs l'un que l'autre, cela ne devrait pas affecter les préférences d'attachement. Enfin, *à qui* et *à laquelle* (Questionnaire 3) ont un niveau d'informativité équivalent, mais leur niveau d'atténuation est différent, car *à laquelle* est plus long. Comme dans le Questionnaire 3, on manipule les deux formes de relativiseurs ainsi que le genre : *à qui*, *auquel* et *à laquelle*. On peut alors évaluer si l'atténuation et l'informativité ont un effet sur les préférences d'attachement, ou si un seul facteur contribue aux préférences d'attachement.

Dans le contexte précis des phrases étudiées, la rigidité est neutralisée. La rigidité d'une expression référentielle est, selon Ariel (1990), la capacité de celle-ci à sélectionner de façon unique l'entité référée lorsque plusieurs sont en compétition. *Qui* est peu rigide puisqu'il ne permet pas de sélectionner un antécédent et *lequel*, alors que N1 et N2 sont volontairement identiques en genre et nombre, ne distingue pas l'antécédent et est dans ce cas tout aussi peu rigide que *qui*.

Ces questionnaires, dans leur ensemble, constituent une première évaluation en compréhension des facteurs influençant la position des relativiseurs sur une hiérarchie d'accessibilité et de leur capacité à signaler un antécédent plus ou moins accessible.

1. Questionnaire 1 : comparaison de *qui* et *lequel/laquelle*, interindividuel

Dans le Questionnaire 1, les participants lisent des phrases comme (59a-b) répétées plus bas. Elles contiennent une proposition relative non restrictive introduite par *qui* ou *lequel/laquelle* qui est précédée d'une structure « N1 de N2 ». L'attachement de la relative est ambigu, car le contenu de la phrase n'indique pas le site d'attachement, N1 ou N2.

N1 et N2 sont identiques en genre et nombre, sinon le marquage de *lequel/laquelle* signalerait l'antécédent et l'attachement ne serait plus ambigu. Des études ont montré que ce type d'information est pris en compte rapidement dans l'interprétation des pronoms (Caramazza, Grober, Garvey, & Yates, 1977 ; Ehrlich, 1980 ; Garnham & Oakhill, 1985). N1 et N2 désignent des entités humaines, pour éviter un biais possible de l'animation (Bock & Levelt, 1994 ; Desmet *et al.*, 2006). Un pré-test, présenté plus bas, assure que le contenu de la proposition relative est sémantiquement neutre.

En premier lieu, le Questionnaire 1 constitue une réplication des expériences réalisées en français sur les préférences d'attachement. Conformément à ce qui a pu être mesuré avec des mesures des mouvements oculaires (Zagar *et al.*, 1997) ou avec des questions de compréhension (Baccino *et al.*, 2000), on s'attend à une forte préférence pour N1 avec *qui*.

De plus, ce questionnaire teste les deux hypothèses qui ont été formulées sur l'atténuation (Hypothèse 1a) et l'informativité (Hypothèse 1b). On a vu dans la présentation des hypothèses que *qui* et *lequel/laquelle* diffèrent au niveau de l'atténuation et de l'informativité, ce qui peut affecter les préférences d'attachement. *Lequel* et *laquelle* sont moins atténués que *qui* et signalent un niveau d'accessibilité plus faible. Par conséquent, une proposition relative introduite par *lequel* ou *laquelle* entraîne plus d'attachements vers N2, entité la moins accessible des deux antécédents possibles. Comme on mesure la fréquence d'attachement, on s'attend à plus d'attachements à N2 avec *lequel/laquelle* qu'avec *qui* (Hypothèse 1a).

Le niveau d'accessibilité signalé par *lequel/laquelle* est plus bas que *qui* si on prend en considération le critère de l'informativité. Ainsi, on s'attend à plus d'attachements à N2 avec une proposition relative introduite par *lequel/laquelle* qu'à une relative introduite par *qui* (Hypothèse 1b).

D'après la proposition de Frazier et Clifton (1997), si le relativiseur est un pronom relatif, son interprétation est guidée par la proéminence des entités du discours. En français *qui* et

lequel sont des pronoms relatifs et on pourrait s'attendre à une préférence pour attacher la proposition relative à N1, car il s'agit de l'antécédent proéminent. Il ne devrait pas y avoir de différence de traitement selon que la relative est introduite par *qui* ou *lequel*.

Les autres modèles du traitement des ambiguïtés syntaxiques ne posent aucune hypothèse quant à l'effet de la forme du relativiseur sur les préférences d'attachement, bien que les modèles interactifs n'excluent pas dans leur formulation que de telles informations peuvent affecter l'interprétation de telles phrases. D'après la théorie du *garden path*, par contre, il ne devrait pas y avoir d'effet de la forme du relativiseur, avec toujours un attachement à N2 (cf. stratégie du *late closure*).

Les théories sur le traitement des anaphores (*centering theory* et *informational load hypothesis*) suggèrent que plus l'anaphore est réduite, p. ex. anaphore pronominale, plus le référent est proéminent, ce qui est conforme avec les prédictions de l'Hypothèse 1.

1. 1. Méthodologie expérimentale

1. 1. 1. Participants

Soixante étudiants en deuxième ou troisième année de Licence de sciences du langage de l'Université Lumière Lyon 2 ont librement participé à cette expérience lors d'un de leurs cours. Seuls les questionnaires des participants de langue maternelle française sont pris en compte.

1. 1. 2. Matériel

Un ensemble de 32 items expérimentaux est construit (voir Annexe 1), sélectionné après un pré-test évaluant la plausibilité sémantique du contenu des propositions relatives. Les phrases expérimentales comprennent le sujet et le verbe d'une proposition principale, suivis de l'objet, un GN complexe de la forme « N1 de N2 ». Ce GN est suivi d'une proposition relative dont le caractère non restrictif est signalé par la virgule précédant le relativiseur. L'attachement est ambigu, car aucune information sémantique n'en favorise un en particulier, vers N1 ou N2 (voir le pré-test ci-après). Les antécédents sont toujours identiques en genre et en nombre.

La proposition relative est introduite par *qui* (59a) ou par *lequel* (59b) pour les items avec N1 et N2 masculins, et *qui* et *laquelle* si les antécédents sont féminins.

- (59) a. *Je connais le père du maçon, qui est amusant.*
b. *Je connais le père du maçon, lequel est amusant.*

N1 et N2 sont des noms communs désignant des entités humaines identiques en genre et nombre. Un tiers des items expérimentaux contient des noms féminins. N1 et N2 sont contrôlés de façon à ce qu'ils aient une longueur (*i. e.* nombre de lettres) et un nombre de syllabes comparables. Ces informations sont obtenues à partir de la base de données Lexique 3 (New, Pallier, Ferrand, & Matos, 2001).

1. 1. 3. Pré-test

Le but du pré-test est de sélectionner les meilleurs items expérimentaux, ceux pour lesquels le contenu de la proposition relative est plausible aussi bien pour N1 que pour N2, afin de s'assurer que, si une préférence d'attachement est mesurée dans le Questionnaire 1, elle n'est pas induite par un biais sémantique pour attacher la proposition relative à N1 ou N2.

Vingt participants, issus de la même population que ceux du Questionnaire 1, mais n'y ayant pas participé, ont pour tâche d'évaluer la plausibilité du matériel. Ils doivent juger des énoncés simples dérivés des propositions relatives des items expérimentaux. Pour l'item (59a-b) on dérive une proposition compatible avec un attachement vers N1 (*le père est amusant*) et une avec un attachement vers N2 (*le maçon est amusant*). Des 40 items expérimentaux construits au départ on obtient 80 propositions qui sont toutes présentées aux participants dans un ordre aléatoire contrôlé. Tout d'abord, les deux propositions dérivées du même item expérimental sont réparties dans les première et deuxième moitiés du pré-test, avec au moins douze propositions intercalées. Ensuite, les propositions sont distribuées de façon à ce qu'il y ait autant de propositions compatibles avec un attachement à N1 ou à N2 dans les deux moitiés du corpus. Deux listes sont réalisées pour contrebalancer l'ordre des propositions : si dans la Liste 1 c'est la proposition compatible avec un attachement à N1 qui est présentée pour un item donné, dans la Liste 2, c'est la proposition compatible avec un attachement à N2 que l'on trouve au même endroit dans la liste.

Les participants estiment la plausibilité de ces 80 propositions à l'aide d'une échelle à sept degrés allant de 1 à 7. Plus la valeur estimée est élevée, plus la phrase est plausible. La durée de passation du pré-test est d'environ un quart d'heure. Cinq phrases avec une structure

identique à celle des propositions sont présentées au début du pré-test pour familiariser les participants avec l'utilisation de l'échelle.

Les 32 meilleurs items sont ensuite sélectionnés pour le Questionnaire 1 (Tableau 3, Annexe 1). Un test de Student ne révèle aucune différence significative entre les propositions compatibles avec N1 ou N2 [$t(27) = 1,27$; *n. s.* ; $t(47) = 1,47$; *n. s.*]²⁹. On peut considérer que, pour les 32 items expérimentaux retenus, le contenu de la proposition relative est plausible de façon équivalente pour N1 et N2. La longueur [$t(31) = 0,71$; *n. s.*] et le nombre de syllabes [$t(31) = 0,65$; *n. s.*] pour N1 et N2 sont contrôlés.

Tableau 3 : caractéristiques du matériel expérimental
(Questionnaire 1, 32 items retenus)

	Score de plausibilité	Longueur	Nb. syllabes
N1	5,83 ($\sigma = 1,68$)	7,25 ($\sigma = 7,09$)	2,25 ($\sigma = 0,80$)
N2	5,95 ($\sigma = 1,68$)	7,00 ($\sigma = 1,63$)	2,16 ($\sigma = 0,57$)

1. 1. 4. Procédure expérimentale

Un essai (Figure 4) consiste à lire une phrase comme (59a-b) suivie de deux propositions situées à droite de la phrase qui sont identiques à celles du pré-test, l'une congruente avec un attachement vers N1 (*le père est amusant*), l'autre vers N2 (*le maçon est amusant*).

64 - J'accompagne le père du maçon, qui est amusant.	Le père est amusant.	<input type="checkbox"/>
	Le maçon est amusant.	<input type="checkbox"/>

Figure 4 : exemple d'essai pour le Questionnaire 1

Pour ce questionnaire de type « papier-crayon », les participants ont pour tâche de lire une phrase et les propositions correspondantes, puis de cocher celle des propositions qui leur semble la « plus correcte », de façon spontanée. Un tel paradigme expérimental est fréquemment utilisé dans les études sur les préférences d'attachement (Cuetos & Mitchell, 1988 ; Gilbois *et al.*, 1995 ; Hemforth *et al.*, 2000 ; Traxler *et al.*, 1998 ; van Gompel *et al.*, 2001).

²⁹ *n. s.* : non significatif.

Cependant, afin de maximiser les chances d'obtenir des résultats et d'éviter la mise en place de stratégies (cf. présentation de *de* et *avec*, Baccino *et al.*, 2000 ; De Vincenzi & Job, 1993, 1995), le type de relativiseur varie de façon interindividuelle. La moitié des participants traite la version avec *qui* et l'autre moitié des participants reçoit la version avec *lequel* et *laquelle*, afin d'éviter la mise en place de stratégies, en vertu de la synonymie des relativiseurs selon la norme. Les participants sont moins enclins à conserver la même interprétation quelle que soit la forme du relativiseur.

À ces items expérimentaux 100 remplisseurs de structures syntaxiques et de longueurs variées sont ajoutés. Ces phrases ne contiennent pas de proposition relative, mais souvent plusieurs noms ou plusieurs actions pour un même personnage afin de présenter une complexité d'interprétation.

Ce Questionnaire manipule une variable indépendante, le Relativiseur, à deux modalités (*qui*, *lequel/laquelle*) et 32 items sont présentés dans chaque condition, de façon interindividuelle. La variable dépendante est la réponse du participant, c'est-à-dire un attachement à N1 ou à N2. Quatre listes expérimentales sont élaborées pour contrebalancer le Relativiseur et l'ordre de présentation des propositions (attachement vers N1 ou N2 proposé en premier). Un ordre de présentation pour l'ensemble des listes est élaboré, où chaque item expérimental apparaît au même endroit parmi les remplisseurs, dans chacune des listes ainsi construites. Les passations sont collectives et leur durée totale n'excède pas 20 minutes.

1. 2. Résultats

En cochant une des deux propositions, les participants indiquent le type d'attachement réalisé. Dans ce questionnaire, la variable dépendante est la fréquence des attachements vers N1. Elle est analysée dans les tests statistiques et transformée ici en pourcentage pour plus de lisibilité.

Comme un participant n'avait pas fourni de réponse pour tous les items, un participant est au hasard retiré de chacune des autres listes, afin d'équilibrer ces listes. Les analyses sont conduites sur 56 participants, 28 ayant traité *qui* et autant *lequel/laquelle*. La procédure est interindividuelle.

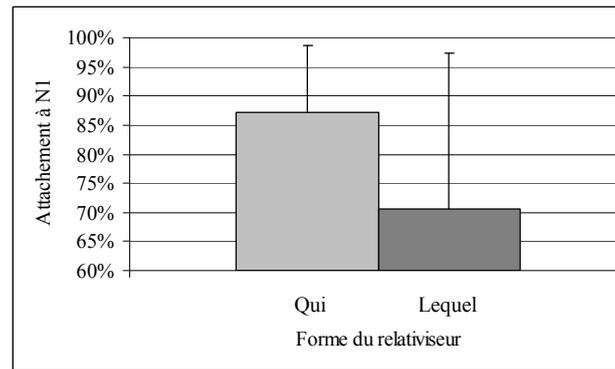


Figure 5 : Pourcentage d'attachements à N1 en fonction du Relativiseur (qui, lequel, procédure interindividuelle)

La préférence d'attachement globale des participants (Figure 5) favorise N1 comme site d'attachement de la proposition relative (78,9 %). Que ce soit pour les propositions relatives introduites par *qui* ou par *lequel/laquelle*, on observe une forte préférence pour attacher la relative à N1. En effet, pour les relatives avec *qui*, le pourcentage moyen d'attachements vers N1 est de 87,2 % ($\sigma = 11,6$) et diffère du hasard [$t1(27) = 17,70$; $p < 0,0001$; $t2(31) = 6,77$; $p < 0,0001$]. Il en va de même pour les propositions relatives avec *lequel/laquelle* où 70,6 % ($\sigma = 26,7$) des attachements sont à N1 et diffèrent du hasard [$t1(27) = 4,40$; $p < 0,0001$; $t2(31) = 27,06$; $p < 0,0001$].

Cette préférence d'attachement semble néanmoins plus forte avec *qui*, ce qui est confirmé par les analyses statistiques réalisées sur la fréquence des attachements vers N1, puisqu'on trouve une différence significative entre *qui* et *lequel*, à la fois par sujet [$t1(54) = 3,00$; $p = 0,004$] et par item [$t2(31) = 7,33$; $p < 0,0001$]. Il y a plus d'attachements à N2 pour les phrases avec *lequel/laquelle*.

1. 3. Discussion

La préférence globale d'attachement à N1 observée dans le Questionnaire 1 confirme les résultats obtenus en français, même si cette préférence avec *qui* (87,2 %) est plus faible que dans la tâche de complétion (sans contexte) de Zagar *et al.* (1997).

Avec *lequel* on retrouve une forte fréquence d'attachement à N1, bien que cette préférence soit significativement moins importante : il y a plus d'attachements vers N2 avec *lequel/laquelle* qu'avec *qui*. Ce résultat montre un effet de l'atténuation (Hypothèse 1a) et de l'informativité (Hypothèse 1b) : *lequel* et *laquelle* sont moins atténués et plus informatifs que

qui, ce qui les positionne à un niveau inférieur dans la hiérarchie d'accessibilité, signalant un attachement à l'antécédent moins accessible, N2.

La préférence globale de N1 mesurée pour les propositions relatives est conforme à ce qu'on pouvait attendre selon la *construal hypothesis* : *qui* et *lequel* sont des pronoms relatifs et leur attachement est par conséquent guidé par la proéminence de N1 et N2. Comme N1 est proéminent, ce dernier est donc le site d'attachement préféré. L'effet de la forme du relativiseur n'était pas prévue.

Enfin, la préférence pour N1 infirme ce que prédisait la théorie du *garden path* : la stratégie du *late closure* ne guide pas la préférence d'attachement pour les phrases testées ici. L'effet du Relativiseur sur les préférences d'attachement suggère que les théories interactives gagneraient à considérer ce type d'information dans les processus de traitement.

Le Questionnaire 1 montre que, pour une fonction relativisée donnée, ici la fonction sujet, si les deux relativiseurs disponibles diffèrent quant au degré d'atténuation (Hypothèse 1a) et d'informativité (Hypothèse 1b), le relativiseur le moins atténué et le plus informatif signale un antécédent moins accessible. Cependant, la comparaison entre *qui* et *lequel/laquelle* ne distingue pas quel facteur, atténuation ou informativité, influence les préférences d'attachement.

On pourrait proposer une interprétation alternative de ces résultats. On l'a dit plus haut, employer *lequel* pour faire référence à un antécédent humain fait partie du registre formel. La fréquence d'utilisation de ce relativiseur est plus faible que celle de *qui* (voir l'analyse de corpus, Chapitre 1). La différence dans les préférences d'attachement mesurée pour le Questionnaire 1 pourrait alors être liée au fait que *lequel* signale par sa fréquence moindre un attachement à l'antécédent qui est moins accessible et n'est généralement pas le site d'attachement préféré.

Afin d'évaluer cette hypothèse, je propose de comparer les relativiseurs à *qui* et *auquel*. La différence de fréquence est moindre et inversée : *auquel* est plus fréquent que à *qui*. Le Questionnaire 2 permet non seulement de tester l'Hypothèse 1, mais aussi l'effet possible de la fréquence sur l'attachement des relatives ambiguës.

2. Questionnaire 2 : comparaison de *auquel* et *à qui*, interindividuel

Le Questionnaire 2 évalue trois facteurs qui ont pu contribuer à la différence mesurée entre *qui* et *lequel/laquelle* dans le Questionnaire 1. Il s'agit de l'atténuation (Hypothèse 1a), de l'informativité (Hypothèse 1b) et de la fréquence.

En manipulant des propositions relatives OI (*à qui*, *auquel*), deux sources de variation sont contrôlées, l'atténuation et l'informativité. Longs de deux syllabes, *à qui* et *auquel* ont un niveau d'atténuation équivalent (cela n'est valable que pour le cas où l'antécédent est masculin singulier, cf. *à laquelle*).

Le marquage en genre et nombre de *auquel* rend ce relativiseur plus informatif que *à qui*. Par contre, comme *à qui* marque l'animation de son antécédent, contrairement à *auquel*, il est de ce point de vue plus informatif ; ces deux types d'informations paraissent tout aussi importants et on peut considérer que *auquel* et *à qui* ont un niveau d'informativité équivalent.

Si l'atténuation (Hypothèse 1a) et l'informativité (Hypothèse 1b) contribuent à signaler un niveau d'accessibilité et un attachement différents, qu'elles soient ici contrôlées ne devrait pas conduire à une différence dans les préférences d'attachement.

Quant à la fréquence, la différence entre *à qui* et *auquel* paraît moindre qu'entre *qui* et *lequel*. Dans l'étude de corpus présentée plus tard (Chapitre 1), on mesure même une tendance inverse : il y a plus d'occurrences où une proposition relative précédée d'un GN « N1 de N2 » contient *auquel* vs *à qui*, même si on restreint les premières aux cas où le site d'attachement est animé, afin d'avoir une même condition d'animation. Si la fréquence d'occurrence des relativiseurs affecte les préférences d'attachement dans le Questionnaire 1 et entraîne plus d'attachements à N2 pour le relativiseur le moins fréquent, on devrait alors observer plus d'attachements à N2 pour les relatives introduites par *à qui*.

On s'attend à des résultats différents selon l'atténuation et l'informativité (pas de différence d'accessibilité et d'attachement signalée) ou la fréquence (plus d'attachements à N2 pour *à qui*).

Selon la *construal hypothesis*, comme pour *qui* et *lequel/laquelle* (Questionnaire 1), l'attachement est réalisé de façon préférentielle à N1 parce que *auquel* et *à qui* sont des pronoms relatifs. Aucune différence n'est attendue selon la forme des relativiseurs.

2. 1. Méthodologie expérimentale

2. 1. 1. Participants

Les 48 participants ayant répondu à ce questionnaire sont issus de la même population étudiante que pour les expérimentations précédentes, auxquelles ils n'ont pas participé. Tous sont de langue maternelle française.

2. 1. 2. Matériel

Un ensemble de 32 items expérimentaux comme (60a-b) est constitué (Annexe 2), avec des phrases contenant une relative OI introduite par *à qui* (60a) ou *auquel* (60b). Le relativisé est le récipient du verbe de la relative ; il s'agit donc d'un vrai datif. La proposition principale contient le sujet (pronom singulier, première ou deuxième personne), le verbe et l'objet qui est de la forme « N1 de N2 ». La proposition relative non restrictive OI, précédée d'une virgule, contient le relativiseur suivi du sujet de la relative (entité humaine), du verbe et d'un objet direct (toujours non-animé). Les deux antécédents possibles de la relative sont masculins singuliers pour n'utiliser que *à qui* et *auquel* et sont contrôlés pour le nombre de lettres et surtout de syllabes. La plausibilité du contenu de la proposition relative est contrôlée de façon à ce qu'il n'y ait pas de biais sémantique pour attacher la relative à N1 ou N2 (cf. pré-test). Ces items sont sélectionnés à l'issue d'un pré-test évaluant 40 items.

La forme de ces items est similaire à celle des items du Questionnaire 1, mais la fonction relativisée est le OI et N1 et N2 sont masculins singuliers, contrôlés pour la longueur et le nombre de syllabes.

- (60) a. *Tu rencontres l'admirateur du chanteur, à qui le copain prête de l'argent.*
b. *Tu rencontres l'admirateur du chanteur, auquel le copain prête de l'argent.*

La liste des items expérimentaux conservés pour le Questionnaire 2 se trouve dans l'Annexe 2.

2. 1. 3. Pré-test

Comme pour le pré-test du Questionnaire 1, ce pré-test sert à évaluer la plausibilité du contenu de la proposition relative et à s'assurer de l'absence d'un biais vers N1 ou N2.

Vingt participants provenant de la même population étudiante ont jugé la plausibilité de propositions dérivées des items expérimentaux. Pour chaque item comme (60a-b), on

construit deux propositions à partir de la relative qui sont compatibles avec un attachement soit vers N1 (*le copain prête de l'argent à l'admirateur*) soit vers N2 (*le copain prête de l'argent au chanteur*). On obtient alors 80 propositions dont l'ordre de présentation est contrôlé : les deux propositions issues d'un même item sont réparties dans les deux moitiés du pré-test et séparées par au moins 12 propositions. Les propositions sont distribuées dans les deux moitiés de façon à ce qu'il y ait autant de propositions compatibles avec un attachement à N1 qu'à N2. Deux listes expérimentales sont constituées pour contrebalancer l'ordre de présentation des propositions.

Les participants évaluent la plausibilité de ces propositions à l'aide d'une échelle allant de 1 à 7 (du moins au plus plausible). Ils traitent toutes les propositions précédées de cinq énoncés d'entraînement, de structure comparable. La passation dure environ 15 minutes.

Les 32 items sélectionnés pour le Questionnaire 2 sont ceux pour lesquels les scores de plausibilité obtenus sont proches entre une proposition compatible avec un attachement à N1 et une compatible avec un attachement à N2 (voir Tableau 4). Le test de Student conduit sur ces items ne montre pas de différence significative de plausibilité [$t(19) = 0,49$; *n. s.* ; $t(31) = 0,27$; *n. s.*].

Pour les items conservés, N1 et N2 ne diffèrent pas statistiquement, que ce soit pour la longueur [$t(31) = 0$; *n. s.*] ou le nombre de syllabes [$t(31) = 0,23$; *n. s.*].

Tableau 4 : caractéristiques du matériel expérimental
(Questionnaire 2, 32 items retenus)

	Score de plausibilité	Longueur	Nb. syllabes
N1	4,87 ($\sigma = 0,86$)	7,22 ($\sigma = 2,32$)	2,28 ($\sigma = 0,81$)
N2	4,91 ($\sigma = 0,97$)	7,22 ($\sigma = 1,95$)	2,25 ($\sigma = 0,62$)

2. 1. 4. Procédure expérimentale

La procédure employée est identique à celle du Questionnaire 1. Il s'agit d'un questionnaire de type « papier-crayon » : les participants lisent sur la gauche des feuilles les phrases (items expérimentaux et remplisseurs) suivies sur la droite de deux propositions. Ils cochent la proposition qui leur semble la « plus correcte ». Les remplisseurs sont identiques à ceux du Questionnaire 1.

La variable indépendante, manipulée de façon interindividuelle, est le Relativiseur, variable à deux modalités (*à qui, auquel*). Les 32 items expérimentaux sont répartis parmi 100

remplisseurs identiques à ceux du Questionnaire 1. Un ordre aléatoire est généré pour ces items et conservé pour tous les participants et pour les quatre listes construites. Chaque proposition issue d'un même item figure au même endroit dans les listes. Ces dernières permettent de présenter les deux formes de relativiseurs et de contrebalancer l'ordre des propositions (compatible avec N1 ou N2 en premier). Les passations sont collectives et durent environ 25 minutes.

2. 2. Résultats

Tous les participants ayant répondu à l'intégralité du questionnaire, leurs réponses sont conservées dans l'analyse statistique. Comme pour le Questionnaire 1, la proposition choisie pour chaque phrase est transformée en fréquence d'attachement à N1 sur l'ensemble des 32 items. C'est cette fréquence qui est analysée statistiquement, mais présentée ici sous forme de pourcentage des attachements à N1, comme sur la Figure 6.

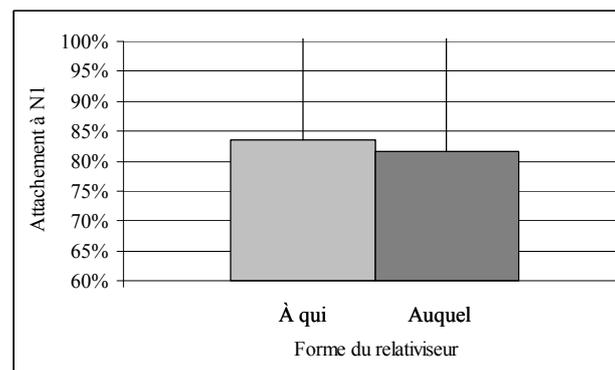


Figure 6 : Pourcentage d'attachements à N1 en fonction du Relativiseur (à qui, auquel, procédure interindividuelle)

On observe une préférence importante pour attacher la relative à N1 (85,3 %), quel que soit le Relativiseur. La préférence pour N1 diffère du hasard, pour *à qui* [$t(11) = 8,28$; $p < 0,0001$; $t(31) = 11,0$; $p < 0,0001$] et pour *auquel* [$t(11) = 4,06$; $p = 0,002$; $t(31) = 16,03$; $p < 0,0001$].

Le pourcentage d'attachements à N1 ne varie pas de façon significative en fonction de la forme du relativiseur [$ts < 1$]. Il y a ainsi une préférence pour N1 comparable entre des phrases avec *à qui* (87,8 %) et celles avec *auquel* (82,8 %).

2. 3. Discussion

Les résultats montrent une préférence pour attacher la proposition relative à N1, indépendamment du type de relativiseur. L'absence de différence observée entre *à qui* et *auquel* indique que, lorsque les relativiseurs sont contrôlés aussi bien au niveau de l'atténuation (Hypothèse 1a) que de l'informativité (Hypothèse 1b), cela ne conduit pas à une différence d'attachement.

Si, pour une fonction relativisée donnée, les relativeurs ont une position comparable sur la hiérarchie d'accessibilité, ils signalent alors le même attachement à l'antécédent le plus accessible.

La différence de fréquence d'occurrence entre *à qui* et *auquel* ne suffit pas à signaler un attachement différent : bien que *à qui* soit plus rare, il n'est pas attaché plus souvent à N2 que *auquel*. L'hypothèse alternative d'un effet de la fréquence sur les préférences d'attachement proposée dans le Questionnaire 1 n'est pas vérifiée ici. La différence de fréquence ne contribue pas, tout du moins seule, à signaler un attachement autre.

Les résultats obtenus sont en partie conformes à ce que la *construal hypothesis* prédit avec la préférence globale pour N1. L'absence de différence entre les deux formes de relativiseurs n'était cependant pas attendue selon cette théorie. Notons aussi que la préférence pour N1 infirme l'hypothèse de l'utilisation de la stratégie de *late closure* (cf. théorie du *garden path*) selon laquelle le parseur réalise un attachement à N2.

Dans le Questionnaire 1 (*qui, lequell/laquelle*) et le Questionnaire 2 (*à qui, auquel*), on observe une préférence pour attacher la proposition relative à N1, comme dans la plupart des expériences réalisées en langue française (Colonna, 2001 ; Pynte & Colonna, 2000 ; Zagar *et al.*, 1997). La préférence pour N1 semble encore plus marquée pour les propositions relatives OI (85,3 %, Questionnaire 2) que pour les relatives sujet (78,9 %, Questionnaire 1). L'élaboration des frontières prosodiques pendant la lecture des phrases expérimentales pourrait aussi expliquer la très forte préférence pour N1.

Des études ont montré que la longueur de la structure dont l'attachement est ambigu peut influencer l'attachement réalisé (voir p. ex. Colonna, 2001). Plus la structure ambiguë est longue, moins un attachement local est favorisé, que ce soit avec l'attachement d'un groupe prépositionnel (Frazier & Clifton, 1996 ; Petrone, Colonna, Hemforth, d'Imperio, & Pynte, 2005 ; Thornton, MacDonald, & Arnold, 2000) ou celui d'une proposition relative (Pynte & Colonna, 2000). Ceci s'explique selon Fodor (1998) par une sorte d'équilibrage prosodique :

la structure ambiguë est attachée à la tête d'une structure de taille comparable. Dans le Questionnaire 2, la taille de la proposition relative est importante et c'est pour cela qu'une préférence pour N1 est observée. Dans le Questionnaire 1, par contre, la plus petite taille de la proposition relative favorise le nombre des attachements à N2.

En résumé, le Questionnaire 2 montre que lorsque l'atténuation et l'informativité sont contrôlées entre deux relativiseurs, il n'y a pas de différence d'attachement. Les résultats de ce présent questionnaire confortent l'idée selon laquelle, dans le Questionnaire 1 qui compare *qui* et *lequel/laquelle*, c'est bien une modification de l'atténuation et de l'informativité, plus que de fréquence, qui contribue à la différence d'attachement mesurée. Afin de tester plus précisément l'effet de la forme du relativiseur et celui de l'atténuation (Hypothèse 1a) et de l'informativité (Hypothèse 1b), il faut comparer deux relativiseurs où soit l'atténuation, soit l'informativité varie.

En français, il n'y a que pour la fonction sujet que le niveau d'informativité des deux relativiseurs est différent, puisque c'est le seul cas où s'il y a toujours un marquage en genre et nombre pour un relativiseur (*lequel, laquelle*), l'autre, *qui*, ne distingue pas l'animation de son antécédent. Par contre, les deux relativiseurs disponibles diffèrent souvent au niveau de l'atténuation, à l'exception de *à qui* et *auquel/auxquels*. En comparant *à qui* et son équivalent féminin *à laquelle*, on observe les préférences d'attachement pour deux relativiseurs dont le degré d'informativité est contrôlé, mais où le degré d'atténuation varie. Le relativiseur *à laquelle* étant plus long, moins atténué, si c'est la différence d'atténuation qui situe les relativiseurs à des positions différentes sur la hiérarchie d'accessibilité, on devrait observer plus d'attachements à N2 avec *à laquelle* qu'avec *à qui*. C'est ce que le Questionnaire 3 évalue.

3. Questionnaire 3 : comparaison de *auquel* et *à qui*, intra-individuel

Le but de ces questionnaires est de tester l'hypothèse selon laquelle la forme du relativiseur signale le niveau d'accessibilité de son antécédent et guide l'attachement à réaliser. D'après la théorie de l'accessibilité d'Ariel (Ariel, 1990, 2001), les expressions référentielles se situent sur les divers niveaux de la hiérarchie d'accessibilité en fonction de leur degré d'atténuation et d'informativité. Dans le Questionnaire 1, les relativiseurs évalués sont *qui* et *lequel/laquelle* et se distinguent sur ces deux aspects. La différence d'attachement alors mesurée confirme

que ces relativiseurs signalent un attachement autre, mais leur fréquence d'occurrence est susceptible de contribuer à ce résultat. Le Questionnaire 2, lui, évalue deux relativiseurs pour la fonction OI, *à qui* et *auquel* : l'atténuation et l'informativité sont contrôlées, mais la fréquence d'occurrence des relativiseurs diverge. L'absence de différence dans les préférences d'attachement suggère que la fréquence ne contribue pas à elle seule à un attachement différent.

Dans le Questionnaire 3, ce sont des propositions relatives OI qui sont étudiées, avec *à qui* et *auquel*, comme dans le Questionnaire 2, mais des relatives avec des antécédents féminins sont ajoutées, pour pouvoir comparer *à qui* et *à laquelle*. Entre ces deux relativiseurs, le niveau d'informativité est contrôlé, mais l'atténuation et la fréquence d'occurrence varient. Comme la différence de fréquence n'a pas conduit à une modification de l'attachement dans le Questionnaire 2, on peut supposer qu'il en sera de même dans le Questionnaire 3. Si le degré d'atténuation signale un niveau d'accessibilité différent, on devrait alors observer une différence d'attachement entre *à qui* et *à laquelle*. Le relativiseur *à laquelle* étant plus long que *à qui*, on s'attend à plus d'attachements à N2.

Ainsi, l'observation d'une variation de l'attachement des relatives avec *à qui* et de celles avec *auquel* et *à laquelle* montrera un effet de la fréquence d'occurrence de ces relativiseurs. Si la seule différence statistique est observée entre *à qui* et *à laquelle*, on pourra considérer que c'est la différence d'atténuation qui signale un attachement spécifique.

Contrairement au Questionnaire 1 et au Questionnaire 2, le Relativiseur varie de façon intra-individuelle : les participants traitent aussi bien les phrases avec *à qui* que celles avec *auquel/à laquelle*. Cette procédure permet de s'assurer que si une différence selon le Relativiseur est mesurée, elle n'est pas due à un effet de groupe, ce qu'on pourrait suspecter avec une procédure interindividuelle.

3. 1. Méthodologie expérimentale

3. 1. 1. Participants

Quarante étudiants en deuxième année de sciences du langage, tout comme pour les premières expériences, ont participé à ce questionnaire. Tous sont de langue maternelle française et aucun n'avait participé aux expériences précédentes.

3. 1. 2. Matériel

Comme précédemment, les 32 items expérimentaux retenus pour le questionnaire (Annexe 3) sont issus d'un pré-test de 40 items. La structure des phrases est semblable à celle des phrases du Questionnaire 2 : une proposition principale avec un GN de la forme « N1 de N2 » suivi d'une virgule, d'une proposition relative non restrictive contenant le relativiseur, le sujet (toujours humain), le verbe et l'objet de la relative (toujours non-animé). Afin d'éviter toute désambiguïisation par le marquage en genre et nombre avec *auquel* et *à laquelle*, les deux antécédents sont toujours singuliers et du même genre. Le contenu de la proposition relative est contrôlé de façon à ce qu'il soit neutre (cf. pré-test). Plus précisément, il s'agit de relatives datif, où l'antécédent est le récipient du verbe de la relative.

Chaque item expérimental a quatre versions, ce qui permet de manipuler la forme du relativiseur, variable à deux modalités (*à qui*, (61a) et (61c), *auquel/à laquelle*, (61b) et (61d)), mais aussi le genre, variable à deux modalités (masculin dans (61a-b) et féminin dans (61c-d)).

- (61) a. *Je n'oublie pas le grand-père du marié, à qui le témoin apporte le dessert.*
b. *Je n'oublie pas le grand-père du marié, auquel le témoin apporte le dessert.*
c. *Je n'oublie pas la grand-mère de la mariée, à qui le témoin apporte le dessert.*
d. *Je n'oublie pas la grand-mère de la mariée, à laquelle le témoin apporte le dessert.*

Pour la version féminisée des items, N1 et N2 sont dérivés de la version masculine soit en féminisant le nom (article *la* et *-e* final si nécessaire) soit en utilisant l'équivalent féminin (p. ex. *grand-père* et *grand-mère*). N1 et N2 sont contrôlés de façon à avoir une longueur et un nombre de syllabes comparables. Il est fait en sorte que les versions masculine et féminine soient proches sur ces aspects.

3. 1. 3. Pré-test

Ce pré-test est conçu de la même façon que les précédents et 32 participants, n'ayant pas participé aux autres expérimentations, évaluent la plausibilité d'énoncés à l'aide d'une échelle graduée de 1 à 7. Une valeur élevée signale un fort degré de plausibilité.

À partir des items expérimentaux des propositions sont construites. Elles sont compatibles avec un attachement à N1 (*le témoin apporte le dessert au grand-père, le témoin apporte le*

dessert à la grand-mère) ou à N2 (*le témoin apporte le dessert au marié, le témoin apporte le dessert à la mariée*). Pour chaque item, on dérive quatre propositions, pour obtenir 160 propositions à partir des 40 items initialement construits.

Le pré-test est comparable aux précédents. Par contre, pour limiter la longueur de ce pré-test et éviter la répétition des débuts de phrases (p. ex *le témoin apporte le dessert*), des listes sont élaborées pour ne présenter que les propositions avec un nom masculin (*le témoin apporte le dessert au grand-père*) ou avec un nom féminin (*le témoin apporte le dessert à la grand-mère*). Chaque participant traite 80 propositions.

Huit listes sont construites pour ainsi répartir les propositions selon le genre (masculin, féminin) et l'ordre de présentation (présentation d'une proposition contenant N1 ou N2). Un ordre de présentation est généré et reproduit sur l'ensemble des listes, de telle sorte que les deux versions de chaque item, masculines ou féminines, soient réparties dans les deux moitiés de la liste et séparées d'au moins douze autres propositions. Chaque proposition issue d'un même item est présentée au même endroit dans chacune des listes. Chaque liste est précédée de cinq phrases de structure comparable afin de familiariser le participant à l'utilisation de l'échelle. La passation est collective et dure environ 15 minutes.

Les 32 items retenus pour le Questionnaire 3 sont ceux dont la différence de longueur et du nombre de syllabes entre les noms N1 et N2 est la plus faible et où la plausibilité est la plus proche entre les versions masculines et féminines. Les scores de plausibilité, ainsi que la longueur et le nombre de syllabes moyens sont reportés dans le Tableau 5.

Tableau 5 : caractéristiques du matériel expérimental
(Questionnaire 3, 32 items retenus)

		Score de plausibilité	Longueur	Nb. syllabes
Masculin	N1	5,15 ($\sigma = 0,78$)	8,39 ($\sigma = 2,69$)	2,55 ($\sigma =$)
	N2	5,29 ($\sigma = 0,76$)	7,81 ($\sigma = 1,96$)	2,44 ($\sigma = 0,80$)
Féminin	N1	5,55 ($\sigma = 0,71$)	9,06 ($\sigma = 2,62$)	2,52 ($\sigma = 1,03$)
	N2	5,51 ($\sigma = 0,74$)	8,63 ($\sigma = 1,80$)	2,47 ($\sigma = 0,76$)

Les tests statistiques conduits sur ces items montrent une différence significative entre les deux noms [$F1(1, 31) = 13,21$; $p = 0,001$; $F2(1, 31) = 4,41$; $p = 0,04$] : N2 reste plus plausible que N1. Il n'y a, en revanche, pas de différence significative entre les versions masculines et féminines [$F_s < 1$] ni d'interaction [$F1(1, 31) = 3,31$; $p = 0,08$; $F2(1, 31) = 2,39$; *n. s.*].

S'il n'est pas possible d'éviter un biais de plausibilité en faveur de N2, ce dernier favorise de toute façon l'attachement le moins fréquent. Ce biais, s'il affectait les préférences d'attachement, entraîne plus d'attachements à N2. Afin de vérifier si N1 et N2 sont contrôlés au niveau de la longueur du mot et du nombre de syllabes, des analyses de la variance (ANOVA) sont réalisées.

Les noms féminins sont plus longs que leurs pendants masculins [$F(1, 31) = 59,54$; $p < 0,00001$], mais cela indépendamment de la position [$F(1, 31) = 1,11$; *n. s.*] puisqu'il n'y a pas d'interaction significative [$F < 1$]. Cette variation est en partie liée à la présence de morphèmes marquant le féminin. Le nombre de syllabes est par ailleurs équilibré : il n'y a d'effet ni du genre [$F(1, 31) = 1,00$; *n. s.*], ni de la position [$F < 1$], ni d'interaction [$F < 1$].

3. 1. 4. Procédure expérimentale

Le pré-test a permis de choisir 32 items expérimentaux répartis dans huit listes expérimentales. Un essai consiste à lire une phrase suivie à droite de deux propositions, compatible avec un attachement à N1 ou à N2 (cf. proposition du pré-test). Le participant choisit l'une des deux propositions.

Chaque item a quatre versions, voir (61a-d), et n'apparaît qu'une fois dans chaque liste. Les huit listes construites permettent de contrebalancer la forme du relativiseur (*à qui, auquel/à laquelle*), le genre (masculin, féminin), mais aussi l'ordre de présentation des propositions (attachement à N1 ou à N2 en premier). Les variables indépendantes, Relativiseur et Genre, sont manipulées de façon intra-individuelle et les quatre conditions expérimentales sont réparties dans chaque liste pour qu'il y ait autant de phrases de chaque condition.

Ces items sont répartis parmi un ensemble de remplisseurs, chaque version d'un même item apparaissant au même endroit dans les listes. Les remplisseurs sont ceux qui ont été utilisés pour les questionnaires précédents (100), avec en plus 20 de structures syntaxiques comparables. On a aussi ajouté 20 phrases dont la structure syntaxique est semblable à celle des phrases expérimentales, comme (62) où le contenu sémantique signale un attachement à N2.

(62) *Tu attends le patient du chirurgien, à qui l'infirmier tend le scalpel.*

Ces remplisseurs contiennent une ambiguïté syntaxique puisque la proposition relative non restrictive est précédée d'une structure « N1 de N2 ». Cette ambiguïté est temporaire, car le

contenu sémantique de la proposition relative indique l'attachement attendu, N2 dans (62). La moitié de ces remplisseurs indique un attachement à N1, l'autre moitié à N2. Le genre des antécédents est contrebalancé. Ces remplisseurs sont inclus dans le Questionnaire 3 pour signaler la possibilité d'attacher la relative aussi bien à N1 qu'à N2 et éviter que les participants ne décident d'une interprétation unique indépendamment du relativiseur.

Les participants se voient assigner une liste de façon aléatoire et la passation dure environ 25 minutes.

3. 2. Résultats

Pour pouvoir évaluer si les résultats observés diffèrent du hasard, un facteur, qu'on appelle Groupe, est ajouté. Il est à deux modalités, avec soit les participants, soit le hasard. Une ANOVA ayant pour facteurs le Relativiseur, le Genre et le Groupe est réalisée.

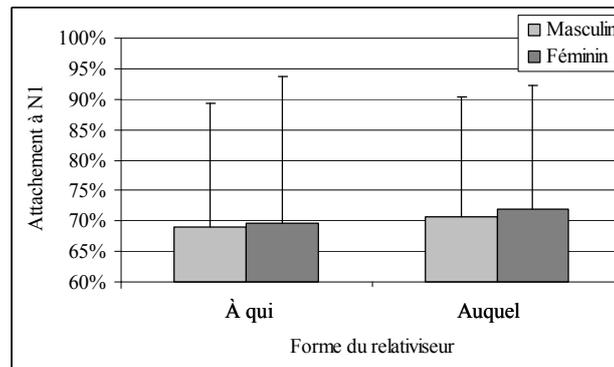


Figure 7 : Pourcentage moyen d'attachements à N1 en fonction du Relativiseur (à qui, auquel) et du Genre (masculin, féminin), procédure intra-individuelle

On obtient un effet très important du Groupe : les participants montrent une préférence importante pour attacher la proposition relative à N1 (70,6%) et cette préférence est différente du hasard [$F_1(1, 39) = 59,17 ; p < 0,0001 ; F_2(1, 31) = 16,21 ; p = 0,0003$].

Aucune différence significative n'est révélée par les ANOVAs. Il n'y a ni effet du relativiseur [$F_s < 1$], ni effet du genre [$F_s < 1$] ou interaction double ou triple entre les facteurs [$F_s < 1$].

3. 3. Discussion

Les résultats confirment la préférence pour N1, conformément aux études dans la littérature et aux questionnaires précédents, mais elle est moindre ici. Ceci pourrait être expliqué par le biais de plausibilité mesuré dans le pré-test : il favorise N2 comme site d'attachement.

Par contre, on n'observe ni un effet du Relativiseur, ni du Genre ni d'interaction. Tout d'abord, notons que le biais de plausibilité du matériel ne peut expliquer aucune des différences, ou plutôt de l'absence de différence, selon ces facteurs. Deux aspects varient entre les relativiseurs, l'atténuation (contrôlée entre *à qui* et *auquel*, différente entre *à qui* et *à laquelle*) et la fréquence de ces relativiseurs (moindre pour *à qui*).

L'absence d'effet du Relativiseur tendrait à montrer que la fréquence de ce dernier ne contribue pas à révéler un attachement différent, mais aussi que le niveau d'informativité marqué par les relativiseurs est équivalent. Par conséquent, ils se situent à un niveau proche dans la hiérarchie et suggèrent un niveau d'accessibilité et un attachement comparables.

Qu'on n'ait observé ni d'effet du Genre ni d'interaction, mais une analogie entre *à qui* et *à laquelle*, montre que la différence d'atténuation ne conduit pas non plus à des attachements différents.

Si on revient aux facteurs qui peuvent contribuer à la différence d'attachement mesurée dans le Questionnaire 1, le Questionnaire 3 et le Questionnaire 2 permettent de rejeter l'hypothèse selon laquelle c'est la différence de fréquence entre *qui* et *lequel/laquelle* qui guide l'attachement. Quand les relativiseurs sont contrôlés au niveau de l'atténuation (Questionnaire 2 et Questionnaire 3), on n'observe aucune différence et, quand seule l'atténuation varie entre deux relativiseurs (*à qui* et *à laquelle*, Questionnaire 3), il n'y a pas non plus de différence.

C'est l'informativité qui semble le facteur déterminant pour situer les relativiseurs à des positions différentes sur la hiérarchie d'accessibilité. Parmi *qui* et *lequel/laquelle*, *qui* est moins informatif et se situe plus haut, signalant un attachement à une entité plus accessible que *lequel/laquelle*. C'est ce qui est mesuré. Mais lorsque l'informativité est contrôlée, comme entre *à qui*, *auquel* et *à laquelle*, aucune différence n'est mesurée dans les préférences d'attachement. L'informativité serait l'indice le plus pertinent dans la hiérarchie d'accessibilité.

Que les deux formes de relativiseurs aient été présentées de façon intra-individuelle peut avoir contribué à favoriser plus d'attachements à N2, à cause d'un biais de synonymie, où les participants réaliseraient le même attachement en vertu du sens identique selon la norme des deux systèmes de relativiseurs. Cependant, les résultats obtenus entre *à qui* et *à laquelle* dans le Questionnaire 2, pour lequel la manipulation du type de relativiseur est réalisée de façon interindividuelle, sont répliqués dans le Questionnaire 3. De fait, on pourrait aussi estimer que la différence observée entre *qui* et *lequel/laquelle* dans le Questionnaire 1 traduit un effet de groupe, les participants qui avaient lu les phrases contenant *lequel* ou *laquelle* ayant une préférence pour attacher la proposition relative à N2 plus importante. Afin d'estimer si la procédure expérimentale a pu contribuer aux résultats sur l'ensemble des questionnaires, le Questionnaire 4 sert de réplique du Questionnaire 1, à la différence près que le type de relativiseur varie cette fois de façon intra-individuelle.

4. Questionnaire 4 : comparaison de *lequel* et *qui*, intra-individuel

Le Questionnaire 4 a pour objectif de répliquer le Questionnaire 1, mais avec la présentation intra-individuelle des deux types de relativiseurs, afin de vérifier que la différence observée entre *qui* et *lequel/laquelle* n'est pas due à un effet de groupe.

Le Questionnaire 4 sert aussi à évaluer les hypothèses qui ont été formulées à propos de l'effet de la forme du relativiseur. Selon les Hypothèses 1a et 1b, on attend une fréquence plus importante des attachements à N2 avec *lequel/laquelle* qu'avec *qui*, car ils sont moins atténués et plus informatifs et signalent par conséquent un attachement à l'entité la moins accessible, N2.

4. 1. Méthodologie expérimentale

4. 1. 1. Participants

Quarante étudiants provenant de la même population que pour les expériences précédentes sans y avoir participé, ont pris part à ce questionnaire. Tous sont de langue maternelle française.

4. 1. 2. Matériel

Les 32 items expérimentaux (59a-b) sont identiques à ceux utilisés dans le Questionnaire 1 (s’y reporter pour une description du matériel ; Annexe 1). Cette fois cependant, la variable indépendante (i.e. le type de relativiseur) varie de façon intra-individuelle : les participants lisent aussi bien des phrases contenant *qui* que *lequel/laquelle*. Le contenu de la proposition relative et la plausibilité sont contrôlés (cf. pré-test du Questionnaire 1).

(59) a. *Je connais le père du maçon, qui est amusant.*

b. *Je connais le père du maçon, lequel est amusant.*

4. 1. 3. Procédure expérimentale

Le Questionnaire 4, de type « papier-crayon », consiste à lire des phrases à la droite desquelles deux propositions sont proposées. Les participants choisissent celle qu’ils estiment « la plus correcte », de façon spontanée.

Les 32 items sont présentés parmi 140 remplisseurs, 120 issus du Questionnaire 3 et 20 de structure syntaxique comparable aux items expérimentaux (59a-b) pour signaler que N1 et N2 sont des sites d’attachement possibles. L’ambiguïté d’attachement est levée avec l’information sémantique de la proposition relative : à N1 dans (*Tu rencontres le grand-père du nourrisson, qui est à la retraite*) et à N2 dans (*Je reconnais le cuisinier du prêtre, qui célèbre la messe*).

La moitié de ces remplisseurs contient une proposition relative introduite par *qui* et l’autre par *lequel* ou *laquelle*. Le biais sémantique signale dans la moitié des cas un attachement à N1 et dans l’autre moitié à N2.

La variable indépendante est le Relativiseur, variable à deux modalités, *qui* et *lequel/laquelle*, manipulée de façon intra-individuelle. Chaque participant traite seize items de chaque modalité. Quatre listes expérimentales sont construites pour répartir les items en fonction du Relativiseur et de l’ordre de présentation des propositions (attachement à N1 présenté en premier ou en second).

Les versions d’un même item expérimental apparaissent au même endroit parmi les remplisseurs dans chacune des listes. Comme pour les questionnaires précédents, la passation est collective et dure environ 20 minutes.

4. 2. Résultats

L'analyse est réalisée sur les items expérimentaux, ceux dont l'attachement est ambigu. Pour les remplisseurs avec une relative, les participants choisissent l'attachement signalé par le contenu de la relative.

La fréquence moyenne des attachements à N1 est très importante, puisque sur l'ensemble des items expérimentaux on observe une moyenne de 72,9 %. La préférence pour N1 diffère du hasard, aussi bien pour la condition avec *qui* [$t1(39) = 6,96$; $p < 0,0001$; $t2(31) = 11,92$; $p < 0,0001$] que celle avec *lequel/laquelle* [$t1(39) = 7,12$; $p < 0,0001$; $t2(31) = 8,57$; $p < 0,0001$].

Comme cela est illustré dans la Figure 8, il y a plus d'attachements à N1 avec les phrases contenant *qui* qu'avec celles contenant *lequel/laquelle*.

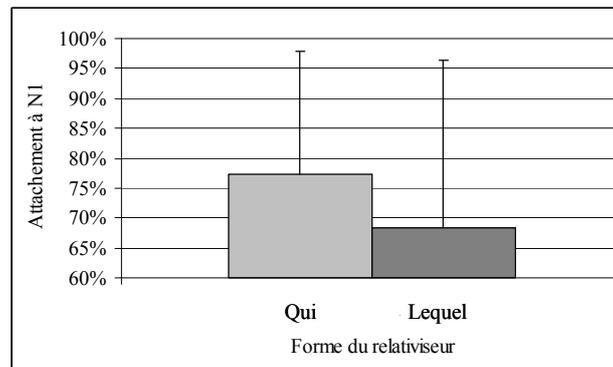


Figure 8 : Pourcentage moyen d'attachements vers N1 en fonction du Relativiseur (*qui*, *lequel*, procédure intra-individuelle)

Cette différence est significative, par sujet [$t1(39) = 2,24$; $p = 0,031$] et par item [$t2(31) = 3,92$; $p < 0,0001$].

4. 3. Discussion

On retrouve dans le Questionnaire 4 une préférence globale importante pour attacher une proposition relative à N1 en français, mais aussi plus d'attachements à N2 lorsque la relative est introduite avec *lequel/laquelle*.

On pourrait considérer que le Questionnaire 4 valide l'Hypothèse 1a sur l'effet de l'atténuation et l'Hypothèse 1b sur l'effet de l'informativité : quand le relativiseur signale un niveau d'accessibilité plus faible, par son degré d'atténuation et d'informativité, il y a plus

d'attachements à l'antécédent le moins accessible. Cependant, le résultat obtenu pour le Questionnaire 3 indique que la variation d'atténuation ne suffit pas. Seule la différence d'informativité permet de situer les relativiseurs à des niveaux différents d'accessibilité. Ces relativiseurs signalent alors un attachement autre. *Lequel/laquelle* étant plus informatifs et plus bas dans la hiérarchie d'accessibilité que *qui*, ils signalent un attachement à l'antécédent moins accessible. C'est pourquoi on observe plus d'attachements à N2 avec *lequel/laquelle*.

Le Questionnaire 4 reproduit le résultat obtenu dans le Questionnaire 1, tout en répliquant la préférence pour N1 avec *qui* mise en évidence dans d'autres expérimentations en français (Pynte & Colonna, 2000 ; Zagar *et al.*, 1997). Cette préférence pour N1 est en accord aussi avec les mesures *off line* de Baccino *et al.* (Baccino *et al.*, 2000).

On peut conclure de ce Questionnaire 4 que pour une même fonction syntaxique, ici la fonction sujet, un relativiseur situé plus bas sur une hiérarchie d'accessibilité (Ariel, 1990, 2001) signale un attachement vers l'entité la moins accessible, quand la proposition relative peut être attachée à plusieurs antécédents. Le degré d'informativité de ces relativiseurs permet de les positionner sur la hiérarchie d'accessibilité.

5. Discussion générale sur les Questionnaires

Cette série de questionnaires vérifie la préférence en français pour attacher une proposition relative à N1 quand cette dernière est précédée d'une structure « N1 de N2 », répliquant les résultats observés dans des expériences précédentes (Baccino *et al.*, 2000 ; Colonna, 2001 ; Frenck-Mestre & Pynte, 2000a ; Zagar *et al.*, 1997).

Ces questionnaires permettent aussi de tester l'effet de la forme du relativiseur sur les préférences d'attachement. Selon Ariel (1990, 2001), la forme des expressions référentielles signale l'accessibilité de l'entité dans le modèle du discours. De la même façon, un relativiseur peut signaler le degré d'accessibilité de son antécédent.

Deux facteurs susceptibles d'affecter le niveau dans la hiérarchie d'accessibilité sont manipulés ici : l'atténuation (Hypothèse 1a) et l'informativité (Hypothèse 1b). Dans les Questionnaires 1 et 4, la comparaison de *qui* et *lequel/laquelle* permet de faire varier ces deux facteurs. *Lequel* et *laquelle* sont à la fois moins atténués et plus informatifs que *qui*, d'où une position plus basse dans la hiérarchie d'accessibilité. Ils signalent alors un antécédent moins accessible que *qui*, et plus d'attachements vers N2, moins accessible que N1, sont réalisés. Cet effet est significatif aussi bien quand les participants ne traitent qu'un type de relativiseur

(procédure interindividuelle, Questionnaire 1) que lorsqu'ils traitent les deux (procédure intra-individuelle, Questionnaire 4).

Le Questionnaire 2 neutralise l'atténuation et l'informativité avec la comparaison de *à qui* et *auquel* et, si la préférence globale pour N1 est conservée, *auquel* n'entraîne pas plus d'attachements vers N2. On peut en fait considérer que *à qui* et *auquel* ont un degré d'informativité et d'atténuation comparables et signalent des degrés équivalents d'accessibilité.

Enfin, le Questionnaire 3 présente des propositions relatives introduites par *à qui* (féminin et masculin), *à laquelle* et *auquel*. L'atténuation, neutralisée entre *à qui* et *auquel*, ou non neutralisée entre *à qui* et *à laquelle*. Le niveau d'informativité est maintenu constant. La préférence pour N1 est vérifiée et l'absence de différence selon l'atténuation (*i. e.* plus d'attachements vers N2) suggère que le degré d'atténuation ne suffit pas à signaler un degré d'accessibilité différent.

5. 1. Aspects méthodologiques

Dans cette série de questionnaires, plusieurs aspects du matériel expérimental sont contrôlés, et en particulier la plausibilité du contenu sémantique, mais aussi la fréquence des antécédents qui peuvent affecter l'attachement réalisé.

5. 1. 1. La plausibilité du contenu sémantique

Dans ces questionnaires, les participants réalisent un jugement explicite de l'attachement de la proposition relative dont le contenu doit être neutre : elle doit pouvoir être attachée aussi bien à N1 qu'à N2.

Des pré-tests ont été construits afin de s'assurer de cette neutralité. Dans les pré-tests, les participants, en estimant le degré de plausibilité de propositions compatibles avec un attachement à N1 ou à N2, indiquent quels contenus sont sémantiquement neutres, c'est-à-dire si N1 et N2 ont un degré de plausibilité comparable. À l'issue des pré-tests, on peut considérer que les propositions relatives sont sémantiquement ambiguës pour les Questionnaire 1, Questionnaire 2 ou Questionnaire 4, puisque qu'il n'y a pas de différence significative entre les valeurs estimées pour des propositions compatibles avec N1 ou N2. Pour le Questionnaire 3, par contre, il subsiste une différence de plausibilité : une proposition avec N2 est jugée dans l'ensemble plus plausible qu'une proposition avec N1. Si cette différence de plausibilité avait affecté les préférences d'attachement, ce biais aurait alors

conduit à une préférence plus marquée pour N2, quelle que soit la forme du relativiseur. Le pourcentage global d'attachement à N1 est en effet plus faible dans le Questionnaire 3 (70,6 %) que dans le Questionnaire 2 (85,3 %), il est possible que ce ne soit pas tant la différence de plausibilité que la manipulation de la forme du relativiseur qui ait entraîné plus d'attachements vers N2.

Dans le Questionnaire 2, la forme du relativiseur est une variable indépendante qui est manipulée de façon interindividuelle, alors qu'elle varie de façon intra-individuelle dans le Questionnaire 3. La co-présence des deux formes de relativiseurs a pu contribuer à des attachements à N2 plus fréquents ; c'est ce qu'on observe aussi entre le Questionnaire 1 (78,9 % d'attachements à N1) et le Questionnaire 4 (72,9 %), où la seule variation est la manipulation, inter- ou intra-individuelle. Qu'on observe plus d'attachements à N2 dans le Questionnaire 3 que dans le Questionnaire 2, et davantage d'attachements à N2 dans les Questionnaires 4 et 1 — dans une moindre mesure — pourrait être expliqué par le fait que la présence de *auquel*, *à laquelle*, *lequel* et *laquelle* incite à plus d'attachements à N2 qu'avec *qui* et *à qui*.

On peut considérer que la préférence pour N1, moindre dans le Questionnaire 3 que dans le Questionnaire 2, est plus le fait de la manipulation de la forme du relativiseur que de la différence de plausibilité.

5. 1. 2. La fréquence des antécédents

Dans la partie théorique, on a pu voir que la fréquence de N1 et N2 influence les préférences d'attachement : les attachements sont plus nombreux à l'antécédent le moins fréquent car il est plus saillant (Pynte & Colonna, 2000). À l'exception de l'étude de Pynte et Colonna (2000), les autres travaux ne contrôlent pas la fréquence d'occurrence des noms utilisés pour N1 et N2, qui pourrait être plus ou moins importante selon la nature des noms (de parenté ou de métier). Dans les questionnaires présentés ici, la fréquence est évaluée, mais on obtient une différence marginalement significative. N1 serait un peu plus fréquent que N2 et cela est en partie lié à l'utilisation de noms de parenté en N1 (*filis*, *papa*), qui sont plus fréquents que des noms de métier pour N2 (*maçon*). Cette différence de fréquence est bien moindre que celle du matériel de Pynte et Colonna (2000) où des mots très peu fréquents sont utilisés (p. ex. *soubrette*). Néanmoins, si la fréquence des antécédents avait affecté les préférences d'attachement, elle aurait alors conduit à une préférence pour N2. Elle n'explique pas les préférences obtenues.

En contrôlant la fréquence relative des antécédents dans les questionnaires, on s'assure que l'observation d'une préférence globale pour N1 n'est pas due à une saillance supérieure de ce dernier du fait de la fréquence. Qui plus est, dans les Questionnaires 2 et 3, où la fréquence de N1 et N2 présente une différence marginalement significative, N1 est plus fréquent et c'est N2 qui serait alors le nom le plus saillant et le site d'attachement préféré. Tel n'est pas l'effet mesuré, car la différence n'est pas significative.

5. 2. Implications théoriques

5. 2. 1. La question de la fréquence des relativiseurs

La différence de fréquence entre les relativiseurs comme *qui* et ceux comme *lequel* peut aussi expliquer les préférences d'attachement observées. L'emploi d'un relativiseur comme *lequel* ou *laquelle* est bien moins fréquent que celui de *qui* (voir l'étude de corpus dans le Chapitre 1), notamment lorsque l'antécédent est une entité humaine. Si on reprend la hiérarchie d'accessibilité d'Ariel (2001), la fréquence des expressions référentielles est un corrélat de l'atténuation, de l'informativité et de la rigidité. Une expression référentielle qui a une position élevée dans la hiérarchie, p. ex. un pronom, a aussi une fréquence d'occurrence élevée alors qu'une plus basse est moins fréquente, p. ex. un nom commun.

Une explication alternative de la différence d'attachement entre *qui* et *lequel/laquelle* était que la fréquence moindre de *lequel* et *laquelle* signale un attachement à l'antécédent qui est le site d'attachement le moins fréquent, N2. La différence de fréquence entre *à qui* d'un côté et *auquel* et *à laquelle* de l'autre est inversée (cf. étude de corpus, Chapitre 1) : *à qui* est moins fréquent, même quand on le compare à *auquel* avec un antécédent animé. L'absence de différence d'attachement tend à montrer que la fréquence d'occurrence des relativiseurs ne contribue pas aux préférences d'attachement mesurées entre *qui* et *lequel/laquelle*.

Peut-on dire que seule la différence de fréquence entre les relativiseurs contribue à la différence d'attachement réalisé ? Dans un sens, oui, puisque l'emploi de *lequel* est bien moins fréquent que celui de *qui*. C'est justement l'utilisation d'une expression référentielle plus rare qui est interprétée comme signalant une interprétation qui diffère de l'interprétation habituelle. *Lequel* signale un attachement différent, à N2 au lieu de N1. Cependant, la fréquence n'explique pas à elle seule les stratégies d'interprétation, puisqu'il n'y a pas de différence d'attachement pour les relatives OI en dépit de la variation de fréquence entre les formes de relativiseurs.

Qui plus est, si une proposition relative introduite par *lequel* est nécessairement non restrictive, une proposition relative introduite par *qui* est généralement restrictive, sauf si le relativiseur est précédé d'une virgule. Comme une proposition relative non restrictive introduite par *qui* est bien moins fréquente que son homologue restrictif, une préférence pour N2 plus marquée aurait pu être observée avec *qui*, mais ce n'est pas le cas.

5. 2. 2. Les Hypothèses 1a et 1b

Les Questionnaires 1 et 4 comparent *qui* et *lequel/laquelle*. Entre ces deux relativiseurs, atténuation et informativité varient. La différence dans les préférences d'attachement peut être liée soit à l'effet des deux facteurs, soit à l'effet d'un seul. La comparaison des deux relativiseurs ne permet pas de distinguer la contribution de ces deux facteurs aux préférences d'attachement.

Les comparaisons de *à qui* et *auquel* (Questionnaire 2) et de *à qui*, *auquel* et *à laquelle* (Questionnaire 3) permettent d'évaluer l'effet de l'atténuation alors que l'informativité est contrôlée. Aucune différence d'attachement n'est observée : la préférence pour N1 est aussi forte pour tous les relativiseurs. La différence d'atténuation entre *à qui* et *à laquelle* ne signale pas d'attachement différent. Il est néanmoins possible que l'atténuation ait un effet pour les relatives sujet, mais il n'a pas été possible de le mettre en évidence ici.

L'absence de différence dans le Questionnaire 2, qui compare *à qui* et *auquel*, peut ainsi être due à la neutralisation de l'atténuation, mais aussi à des niveaux d'informativité équivalents. Si l'absence de différence significative pouvait alors être attendue dans ce questionnaire, une interaction entre type de relativiseur et genre aurait pu être observée dans le Questionnaire 3. L'ajout de *à laquelle* dans le Questionnaire 3 permet en effet de manipuler le degré d'atténuation entre *à qui* et *à laquelle*. Aucun effet principal ni aucune interaction n'atteignent le seuil de significativité : il y a une préférence pour N1 quelle que soit la forme du relativiseur. Cette absence de différence est au moins en partie prévisible : si les relativiseurs (*à qui* et *auquel*) ne diffèrent ni pour l'atténuation, ni pour l'informativité, ils signalent un niveau d'accessibilité comparable, c'est-à-dire un attachement identique. Par contre, s'ils diffèrent quant à l'atténuation (*à qui* et *à laquelle*), l'absence de différence de préférence tend à montrer que ces deux relativiseurs ne signalent pas un attachement différent : l'atténuation à elle seule n'est pas suffisante pour montrer un niveau d'accessibilité différent. Si l'Hypothèse 1a n'est pas validée, on peut proposer que dans les comparaisons entre *qui* et *lequel/laquelle*, c'est l'informativité plus que l'atténuation qui contribue à signaler

un niveau d'accessibilité différent des antécédents et qui conduit à plus d'attachements à N2 avec *lequel/laquelle*. L'Hypothèse 1b est validée.

5. 2. 3. Les théories de l'accessibilité

Les Hypothèses 1a et 1b sont basées sur les théories de l'accessibilité, en particulier celle d'Ariel (1990, 2001) qui a l'avantage d'être plus détaillée que celle de Gundel *et al.* (Gundel *et al.*, 1993) et de proposer des critères d'accessibilité applicables aux relativiseurs. Selon Ariel, la position des expressions référentielles dans la hiérarchie et l'accessibilité signalée sont fonction de la distance et de la rigidité, qui ne varient pas ici, mais aussi de l'atténuation et de l'informativité, qui sont soit manipulées soit contrôlées.

Les résultats des questionnaires montrent que la position des expressions référentielles sur la hiérarchie d'accessibilité varie en fonction de critères assez fins. Que *lequel* et *laquelle* soient marqués en genre et nombre contrairement à *qui* (et aussi le fait qu'ils soient plus longs) les positionne plus bas dans la hiérarchie. Cette différence de position suffit à signaler un degré d'accessibilité distinct et plus d'attachements à N2, entité moins accessible que N1 qui, à la tête du GN, domine N2.

Par contre, quand les relativiseurs ont un degré d'informativité et d'atténuation comparable, comme c'est le cas pour *à qui* et *auquel* (Questionnaire 2), le niveau d'accessibilité signalé est équivalent et n'entraîne pas de différence dans les préférences d'attachement. Le Questionnaire 3 montre par ailleurs que l'atténuation seule ne suffit pas à distinguer la position des relativiseurs dans la hiérarchie d'accessibilité et le niveau d'accessibilité signalé est comparable. La préférence d'attachement à N1 ne semble pas affectée par la forme du relativiseur. Cela suggère que l'atténuation seule ne contribue pas à la différence d'attachement pour *qui* et *lequel/laquelle* (Questionnaires 1 et 4). En revanche, le degré d'informativité peut expliquer les différences d'attachement.

Ces Questionnaires montrent que les relativiseurs occupent des positions différentes sur la hiérarchie d'accessibilité et signalent des attachements variés. Les résultats sont compatibles avec ceux obtenus par Ariel (1999) dans son étude de corpus où elle montre que dans des propositions relatives non ambiguës, une anaphore zéro signale un antécédent hautement accessible et qu'un résomptif signale lui un antécédent moins accessible. Cependant, cette étude de questionnaires et celle d'Ariel (1999) évaluent la compréhension pour la première et la production pour la seconde. Les Questionnaires ont mis en évidence que la forme du

relativiseur sert à signaler l'accessibilité de son antécédent, ce qui affecte la préférence d'attachement.

5. 2. 4. La théorie du *garden path* et la *construal hypothesis*

D'après la théorie du *garden path* (Frazier, 1985 ; Frazier & Fodor, 1978 ; Frazier & Rayner, 1982, 1987 ; Rayner *et al.*, 1983), la stratégie du *late closure* guide l'interprétation de propositions relatives précédées d'une structure « N1 de N2 » : une préférence pour N2 est attendue. Sur l'ensemble des questionnaires, on mesure toujours une préférence pour N1, mais la puissance de cette préférence est influencée par la forme du relativiseur. On peut dire que les données recueillies invalident l'hypothèse d'une préférence pour N2 proposée par la *garden path theory*.

Selon la *construal hypothesis* (Frazier & Clifton, 1996, 1997), l'attachement d'une proposition relative est possible aussi bien vers N1 que vers N2 dans une structure « N1 de N2 » ; les relations sémantiques entre N1 et N2 peuvent influencer l'attachement réalisé. Ces relations sont maintenues constantes à travers les Questionnaires. À propos de la différence d'interprétation en fonction de la forme du relativiseur, Frazier et Clifton (1997) proposent que plus le relativiseur ressemble à un pronom, plus il se comporte comme tel : la proéminence des entités du discours influence l'attachement. Plus d'attachements à N1, entité proéminente, sont alors attendus, aussi bien pour les relatives avec *qui* que celles avec *lequel*. Or, les Questionnaires 1 et 4 montrent bien qu'il y a une préférence globale pour N1, mais que des attachements à N2 sont plus fréquents avec *lequel/laquelle*. Ces résultats invalident en partie la prédiction de Frazier et Clifton (1997).

5. 2. 5. Les modèles interactifs

D'après les modèles interactifs, comme le modèle de satisfaction de contraintes (MacDonald, 1993, 1994 ; Spivey-Knowlton *et al.*, 1993) ou le modèle de compétition et d'intégration (McRae *et al.*, 1998 ; Taraban & McClelland, 1988 ; Trueswell *et al.*, 1993), les informations syntaxiques, sémantiques ou discursives interagissent immédiatement et guident l'interprétation. Parmi les phrases qui avaient été présentées dans les questionnaires, les pré-tests permettaient de s'assurer qu'aucun biais sémantique n'indiquait un attachement vers N1 ou N2. Les préférences d'attachement ne sont pas influencées par ces informations.

Si les informations discursives sont prises en compte, la proéminence des entités du discours et l'accessibilité signalée par les expressions référentielles peuvent intervenir. Ainsi,

ces théories prédisent une préférence pour N1, ce qui est observé. La forme du relativiseur, notamment l'informativité signalée par le marquage en genre et en nombre, serait prise en compte et, de façon comparable à ce qui est proposé par la théorie de l'accessibilité, un relativiseur plus informatif signalerait un attachement à l'entité moins accessible, N2.

Cependant, les modèles basés sur les contraintes considéreraient plutôt que c'est la fréquence des relativiseurs qui guide l'attachement, plutôt que l'accessibilité signalée. On a pu voir que la fréquence des relativiseurs n'explique pas les préférences d'attachement mesurées ici. L'accessibilité signalée par les relativiseurs, en revanche, explique mieux les attachements réalisés. Les modèles basés sur les contraintes pourraient alors tout à fait intégrer les principes des théories de l'accessibilité.

Selon le modèle *unrestricted race* (Pickering & Traxler, 1998 ; Traxler, 2005 ; van Gompel & Pickering, 2001), si aucune information ne favorise une interprétation par rapport à l'autre (attachement à N1 ou N2), le lecteur prend sa décision au hasard. Or, les préférences observées dans les questionnaires diffèrent du hasard : différentes contraintes guident non seulement l'interprétation, mais aussi la vitesse à laquelle un attachement à N1 ou à N2 est réalisé. C'est l'attachement établi le plus rapidement qui est conservé. Les résultats des questionnaires suggèrent qu'à la fois la proéminence des antécédents (N1 est proéminent) et la forme des relativiseurs ont un effet sur l'interprétation adoptée.

Qu'il y ait une préférence globale pour N1 indique que la proéminence des antécédents est prise en compte. De même, la différence d'accessibilité signalée par la forme des relativiseurs guide l'interprétation.

5. 3. Conclusion

La forme du relativiseur sert ainsi à signaler le degré d'accessibilité de l'antécédent. Un relativiseur moins atténué et plus informatif signale ainsi un attachement à une entité moins accessible. L'emploi de *lequel* ou *laquelle* signale un attachement différent d'avec *qui*, entraînant plus d'attachements vers N2.

La technique des questionnaires est relativement peu riche en informations sur le traitement cognitif de ces ambiguïtés d'attachement. Elle reste néanmoins un outil intéressant pour une première exploration des préférences d'interprétation de locuteurs natifs (Cuetos & Mitchell, 1988 ; Gilboy *et al.*, 1995 ; Hemforth *et al.*, 2000 ; Traxler *et al.*, 1998 ; van Gompel *et al.*, 2001) : on a pu montrer que les préférences d'attachement diffèrent selon que la proposition relative est introduite par *qui* ou par *lequel/laquelle*.

Chapitre 2 Étude de corpus

Les théories sur le traitement des ambiguïtés syntaxiques ont majoritairement été éprouvées à l'aide d'expériences en compréhension, en lecture de phrases. Mais les études sur le traitement des anaphores sont basées principalement sur l'observation en production des formes anaphoriques. Les prédictions des théories sur les anaphores sont aussi évaluées en compréhension (notamment la *centering theory*).

Des travaux sur l'attachement des propositions relatives ambiguës ont cependant pris en considération les productions de telles structures dans les langues. La *tuning hypothesis* (Mitchell, 1994 ; Mitchell *et al.*, 1995) affirme que les préférences d'attachement en compréhension sont basées sur la fréquence de ces attachements en production (Brysbaert & Mitchell, 1996 ; Dussias & Sagarra, 2007). Cependant, d'autres études, où les données en production (analyse de corpus) et en compréhension (expériences en temps de lecture) sont comparées, tendent à infirmer la *tuning hypothesis*, tout du moins dans sa version à gros grain (Desmet, Brysbaert *et al.*, 2002 ; Desmet *et al.*, 2006).

Cette étude de corpus se présente comme un test de la *tuning hypothesis*, mais il est aussi intéressant d'évaluer les prédictions de cette théorie sur l'effet de la forme du relativiseur. *A priori*, il ne devrait pas y avoir de différence d'interprétation selon que la proposition relative est introduite par *qui* ou *lequel*, par exemple. Cette hypothèse est motivée par le fait que, quand bien même il y aurait une différence au niveau des fréquences d'attachements à N1 et N2 selon le type de relativiseur, ce degré d'analyse serait trop fin pour influencer les préférences d'attachement en compréhension.

Les hypothèses formulées dans le cadre de ce travail peuvent aussi être évaluées au cours de cette analyse de corpus, dans une dimension axée sur la production de structures syntaxiquement ambiguës et dans le prolongement des théories sur l'accessibilité. La conjugaison d'études en compréhension (les questionnaires) et en production (l'analyse de corpus) apportera une contribution intéressante sur la question des mécanismes impliqués dans ces deux aspects, indépendants selon certains chercheurs (Gibson & Schütze, 1999). Des facteurs tels que la proximité du prédicat (Gibson, Pearlmutter *et al.*, 1996) ou le traitement anaphorique qui a lieu lors de l'attachement de la relative (Hemforth *et al.*, 2000) expliqueraient que les stratégies d'attachement en compréhension ne pourraient pas être

appliquées au niveau de la production (Desmet, Brysbaert *et al.*, 2002 ; Gibson & Schütze, 1999).

Or, les Hypothèses 1 et 2 reposent justement sur la prise en compte d'informations discursives fournies par les anaphores, les relativiseurs et les antécédents. Si les mécanismes en production et en compréhension du langage sont indépendants, il se peut que les informations fournies par ces anaphores affectent différemment la compréhension et la production. Si on considère qu'au moins certains mécanismes cognitifs sont communs, on peut alors s'attendre à un effet de la forme du relativiseur qui soit comparable à ce qui a été mesuré en compréhension, dans les études de Questionnaires (Chapitre 1).

L'étude de corpus analyse des textes écrits issus du journal « Le Monde ». Il faut garder à l'esprit qu'il s'agit de textes écrits et préparés, ce qui ne donne qu'un aperçu de l'ensemble des productions langagières (Biber, 1986 ; Biber, Conrad, & Reppen, 1998), mais qui est cependant riche d'informations.

Afin de rester proche des structures étudiées dans les Questionnaires 1 à 4, des propositions relatives à attachement ambigu sont analysées : la fonction relativisée y est sujet (*qui, lequel*) ou OI (*à qui, auquel*) et la relative est non restrictive. Le site d'attachement appartient à un GN complexe de la forme « N1 de N2 » où *de* N2 est un génitif, le complément du nom, ce qui restreint l'analyse à des structures comparables à celles des Questionnaires.

1. Introduction

L'étude de Questionnaires présentée dans le Chapitre 1 a permis une première évaluation de l'Hypothèse 1 : les relativiseurs, s'ils diffèrent assez quant au niveau de l'accessibilité signalée (*qui* et *lequel*), conduisent à un attachement différent. *Lequel* est plus informatif (car marqué en genre et en nombre) et moins atténué (plus long) que *qui* et se situe plus bas sur une hiérarchie d'accessibilité telle que celle proposée par Ariel (1990, 2001), marquant ainsi un attachement à une entité, un antécédent, moins accessible dans le discours. Au contraire, quand les relativiseurs ont un degré d'informativité comparable, et quand l'atténuation est contrôlée, ils signalent un attachement comparable (*à qui, auquel* et *à laquelle*).

Les études en production et l'analyse de corpus en particulier permettent au chercheur d'avoir un regard sur la façon dont le locuteur utilise sa propre langue (Biber *et al.*, 1998), surtout quand les études sur la compréhension et la production du langage sont plus souvent

isolées qu'analysées de concert (MacDonald, 1999). En réalisant une étude de corpus, on pourra estimer les contraintes qui s'exercent en production.

Le but premier de cette étude de corpus est d'évaluer l'Hypothèse 1, cette fois en production. L'emploi d'un relativiseur comme *lequel* au lieu de *qui* signale un attachement à une entité moins accessible si plusieurs sites d'attachement sont possibles. Seront comparés dans cette étude de corpus *qui* et *lequel*, mais aussi *à qui* et *auquel*, ce qui permet d'évaluer l'Hypothèse 1 avec les mêmes relativiseurs que dans l'étude de questionnaires.

Le second intérêt de cette étude de corpus est de tester les différents versants de l'Hypothèse 2, avec les effets attendus de : la longueur de la relative (Hypothèse 2a), la présence d'un modifieur (Hypothèse 2b), l'animation (Hypothèse 2c) et la fonction syntaxique (Hypothèse 2d). Je vais introduire ces hypothèses et préciser les effets attendus, tout en revenant sur les apports de la littérature propices à étayer ces hypothèses. Mais avant cela, je vais revenir sur un aspect capital de l'analyse : la structure « N1 de N2 ». La question de la restrictivité sera abordée ensuite.

1. 1. La structure « N1 de N2 »

L'étude de questionnaires de Gilboy *et al.* (1995) a montré l'importance des relations entre N1 et N2 sur les préférences d'attachement : pour une relation dite de substance (*le pull de coton*) ou de quantité (*le verre d'eau*), on mesure une très forte préférence pour N1. Quand *de* est un génitif (*le parent de l'enfant*, *le musée de la ville*, etc.), une préférence pour N1 est mesurée pour des relations professionnelles (*l'assistant de l'inspecteur*), par exemple, mais pour une relation de possession (*le livre de l'étudiant*), on observe une préférence pour N2. Cette étude révèle le statut particulier des relations de substance et de quantité : les relations entre les antécédents, N1 et N2, seront contrôlées (voir dans ce Chapitre dans la section 3. 2).

1. 2. La restrictivité

Dans l'étude de Questionnaires, des propositions relatives non restrictives sont étudiées, car une relative avec un relativiseur comme *lequel* ou *auquel* est nécessairement non restrictive. Afin de pouvoir les comparer, les relatives avec *qui* et *à qui* étaient aussi non restrictives. La présence d'une virgule en début de relative (et à la fin si la relative est enchâssée) permet de signaler la non restrictivité de la phrase.

La restrictivité de la proposition relative est susceptible d'influencer les préférences d'attachement. Par exemple, le contenu d'une proposition relative non restrictive apporte des

informations nouvelles (Gibson *et al.*, 2005 ; Loock, 2007). Une proposition relative non restrictive a généralement pour antécédent une entité du discours qui a le statut de donnée, ce qui lui confère un niveau d'accessibilité élevé.

Une proposition relative restrictive, par contre, sert à restreindre le référent parmi un ensemble, ce qui signifierait un niveau d'accessibilité plus faible, car plusieurs entités sont en compétition dans le discours (cf. Ariel, 2001).

La restrictivité affecte le traitement de relatives ayant un seul antécédent (Gibson *et al.*, 2005) et peut par conséquent avoir un effet sur l'attachement quand plusieurs sites d'attachements sont disponibles. Néanmoins, pour exclure la possibilité que l'observation de différences dans les préférences d'attachement soit liée à la probabilité d'un relativiseur utilisé pour des relatives restrictives et non restrictives, les analyses se limitent aux relatives non restrictives.

2. Les hypothèses

2. 1. L'effet de la forme du relativiseur (Hypothèse 1)

L'Hypothèse 1 postule que la forme du relativiseur permet de le positionner sur la hiérarchie d'accessibilité (Ariel, 1988, 2001) et signale le degré d'accessibilité de l'antécédent auquel la proposition relative est attachée. Ce degré d'accessibilité dépend de l'atténuation (Hypothèse 1a) et de l'informativité (Hypothèse 1b).

Pour deux relativiseurs disponibles, celui qui est le moins atténué et le plus informatif se situe plus bas sur la hiérarchie d'accessibilité et signale un attachement à un antécédent moins accessible. Pour une structure « N1 de N2 », N1 est considéré plus accessible que N2 parce qu'il est la tête du GN.

Si on compare *qui* et *lequel*, les relativiseurs utilisés pour la fonction sujet, *lequel* est moins atténué (plus long d'une syllabe) et plus informatif (car marqué en genre et en nombre) que *qui* et signifierait un attachement à N2. Pour les relatives OI, *à qui* et *auquel* sont disponibles. Leur degré d'atténuation est contrôlé (sauf avec *à laquelle*), tout comme leur degré d'informativité : *à qui* est plus informatif, car il marque l'animation de son antécédent et *auquel* l'est aussi car il est marqué en genre et en nombre. L'étude de questionnaires (Chapitre 1) suggère que l'absence de différence dans les préférences d'attachement avec *à qui* et *auquel* est liée à un degré contrôlé d'informativité et que l'atténuation, à elle seule, ne contribue pas à signaler un attachement différent.

Dans le corpus, on s'attend, au niveau des fréquences d'attachement à des résultats comparables à ceux obtenus dans les questionnaires. Une proposition relative introduite par *lequel* sera plus fréquemment attachée à N2 qu'une proposition relative introduite par *qui*, car *lequel* signale un antécédent moins accessible. Pour les relatives OI, à *qui* et *auquel* signalent un niveau d'accessibilité équivalent et on devrait observer des fréquences d'attachement comparables.

2. 2. Les propriétés des antécédents

Les relativiseurs sont des expressions référentielles dont la forme signale l'accessibilité de leur antécédent. Les antécédents, N1 et N2, sont aussi des expressions référentielles : leur forme indique le niveau d'accessibilité dans le discours de l'entité à laquelle ils font référence. La présence d'un modifieur (Hypothèse 2b) est un critère utilisé pour positionner les expressions référentielles sur la hiérarchie d'accessibilité d'Ariel (1988, 2001), tout comme le degré de définitude et le type de GN (Hypothèse 2d). L'animation (Hypothèse 2c) ainsi que la fonction syntaxique (Hypothèse 2e) sont elles aussi susceptibles d'influencer le degré d'accessibilité. Observons tout d'abord comment la longueur de la proposition relative peut guider l'attachement (Hypothèse 2a).

2. 2. 1. La longueur de la proposition relative (Hypothèse 2a)

L'effet de la longueur de la relative en compréhension incite, selon Fodor (1998), à attacher une relative longue à une structure de taille comparable, comme « N1 de N2 », ce qui entraîne une préférence pour N1. Une proposition relative courte est aussi attachée à un antécédent de taille comparable, N2.

Relativement à la production de telles structures, les préférences d'attachement selon la taille de la relative sont évaluées dans deux études de corpus, une en français (Toussenet et Abeillé, 2001, cité par Colonna, 2001) et l'autre en néerlandais (De Baecke, Brysbaert & Desmet, 1999, cité par Colonna, 2001). La première analyse porte sur 292 phrases contenant des propositions relatives dont l'attachement est ambigu qui sont précédées d'une structure « N1 prép. N2 ». Les résultats confirment l'effet de la taille sur les préférences d'attachement : plus la relative est courte, plus elle est attachée à N2. Cette observation n'est cependant pas confirmée par l'étude de De Baecke *et al.* (1999) où un attachement préférentiel à N2 est mesuré quelle que soit la longueur de la relative.

On pourrait néanmoins poser une autre hypothèse quant à l'effet sur l'attachement de la longueur de la relative. En effet, une proposition relative longue peut *a priori* apporter plus d'informations nouvelles qu'une courte. L'entité du discours qui n'est pas dans le focus, ou peu accessible, est plus susceptible de recevoir ces informations supplémentaires. Par conséquent, une proposition relative longue sera attachée de façon préférentielle à l'entité la moins accessible, N2.

Une proposition relative courte sera attachée préférentiellement à N2 d'après l'hypothèse de Fodor. S'il s'agit de la préférence d'attachement par défaut, avec un relativiseur comme *qui*, un relativiseur plus bas sur la hiérarchie d'accessibilité signale un tel attachement, sans entraîner de différence. On s'attend à une préférence pour N2 avec une proposition relative courte, indépendamment de la forme du relativiseur.

2. 2. 2. La présence d'un modifieur (Hypothèse 2b)

On l'a vu dans la partie théorique, la modifiabilité est un facteur avancé par Thornton *et al.* (1999) pour expliquer les préférences d'attachement dans des tâches de compréhension (voir Chapitre 1 de la partie théorique, section 2.2.1.2) : un antécédent est moins susceptible d'être le site d'attachement d'une proposition relative s'il est déjà modifié par un adjectif. La validation expérimentale n'est guère tranchée, puisque les données acquises en espagnol valident l'hypothèse, au contraire de celles acquises en anglais (Thornton *et al.*, 1999) ou en français (Colonna, 2001). Colonna (2001) suggère une autre explication basée sur le modèle DLT de Gibson (dependency locality theory, 1998, 2000). Ce modèle postule l'existence de deux coûts de traitement, celui d'intégration et celui de distance. Ici, attacher la relative, c'est établir un lien entre le relativiseur et son antécédent. Colonna (2001) propose que la présence d'un adjectif pour N2 augmente la distance entre N1 et le relativiseur, mais pas entre N2 et le relativiseur (mais cela n'est vrai que lorsque l'adjectif est antéposé). Le parseur préfère choisir une interprétation où le coût de traitement est moindre (cf. limitations de la charge en mémoire de travail), c'est pourquoi il préfère attacher la proposition relative à N2 plutôt qu'à N1, quand N2 est modifié. Ainsi, Colonna (2001) propose que l'effet de la modification de N1 n'a pas d'incidence sur les préférences d'attachement (pas d'augmentation de la distance), alors que la modification de N2 entraîne plus d'attachements à cet antécédent (distance augmentée entre N1 et le relativiseur). La mise à l'épreuve expérimentale présentée par Colonna (2001) valide cette hypothèse.

La notion d'accessibilité procure une autre explication de l'effet de l'adjectif sur les préférences d'attachement. D'après la hiérarchie d'accessibilité d'Ariel (1990, 2001), un GN avec un long modifieur (description courte selon l'appellation d'Ariel) sert à faire référence à une entité dont le niveau d'accessibilité est plus faible qu'un GN plus court de la forme « dét. N ». Si un relativiseur signale un attachement à une entité peu accessible, comme *lequel*, alors il devrait y avoir plus d'attachements vers l'entité modifiée par un adjectif qu'avec un relativiseur comme *qui*.

Les prédictions formulées pour l'effet de la taille de la proposition relative et celui de l'adjectif permettront de tester l'Hypothèse 2a.

2. 2. 3. L'animation (Hypothèse 2c)

La plupart des études en compréhension ont utilisé pour N1 et N2 des référents animés humains. L'animation affecte les préférences d'attachement, aussi bien en compréhension qu'en production. Des études de corpus en néerlandais ont montré que l'attachement le plus fréquent est à N2, ce qui contraste avec la préférence pour N1 observée dans des tâches de compréhension (Desmet, Brysbaert *et al.*, 2002 ; Desmet *et al.*, 2006 ; Kister, 2002). Desmet *et al.* (Desmet, Brysbaert *et al.*, 2002 ; Desmet *et al.*, 2006) suggèrent que, dans les corpus, la préférence pour N2 est liée à la plus grande fréquence de structures « N1 de N2 » où N2 est animé humain et où il est le site d'attachement. L'animation des antécédents guide les préférences d'attachement. De plus, si on considère dans le corpus les cas où N1 et N2 sont humains, comme dans les expériences en compréhension, on mesure alors plus d'attachements à N1, ce qui est conforme à la préférence pour attacher une relative à cet antécédent (Brysbaert & Mitchell, 1996 ; Desmet, De Baecke *et al.*, 2002).

L'animation est une caractéristique linguistique (ou extralinguistique selon Comrie, 1989) qui exerce une forte influence sur d'autres aspects de la langue, comme les fonctions grammaticales (Bock & Levelt, 1994 ; Dahl & Fraurud, 1996) ou l'ordre des mots (Rosenbach, 2005, 2008). Bock et Warren (1985) suggèrent que ces observations sont en partie liées à l'accessibilité conceptuelle plus élevée des entités animées par rapport à celles d'entités non-animées (voir aussi Branigan *et al.*, 2008). L'accessibilité conceptuelle influence ainsi l'ordre des mots. Cela a des conséquences sur l'ordre des constituants selon le statut *given-new* (Bock & Irwin, 1980), le niveau de saillance des référents ou l'animation (Prat-Sala & Branigan, 2000).

Le statut *given-new* et le niveau de saillance des référents du discours sont fondamentaux pour la *givenness hierarchy* de Gundel *et al.* (1993) et l'*accessibility hierarchy* d'Ariel (cf. effet de la saillance, 2001). On peut ainsi émettre l'hypothèse qu'un antécédent animé a un niveau d'accessibilité plus important qu'un antécédent non-animé. C'est la raison pour laquelle on observe un attachement préférentiel pour attacher une proposition relative à l'entité animée.

On l'a vu, certains relativiseurs signalent l'animation de leur antécédent. C'est le cas des relativiseurs ayant pour base *qui* et une préposition, comme *à qui*, qui ne peuvent être attachés qu'à un antécédent animé (et humain). Les autres relativiseurs, *qui*, *lequel* et ceux ayant pour base *lequel* et une préposition (p. ex. *auquel*) peuvent faire référence aussi bien à un antécédent animé que non-animé.

Il serait intéressant de voir, quand deux relativiseurs n'indiquent pas l'animation de leur antécédent, si le degré d'accessibilité qu'ils signalent guide néanmoins l'attachement. Si *lequel* signale un attachement à l'antécédent moins accessible que *qui*, il devrait alors être attaché à un antécédent non-animé plus fréquemment que *qui*.

Si on compare *à qui* et *auquel*, il est possible de faire référence de façon non ambiguë à un antécédent animé en utilisant *à qui*. Cet attachement reste cependant ambigu quand N1 et N2 sont animés. *À qui* étant marqué pour l'animation, il sert à signaler un attachement à un antécédent moins accessible. Comme il ne peut renvoyer qu'à un antécédent animé, il sera utilisé de façon préférentielle quand cet antécédent est moins accessible, en N2 par exemple. On s'attend ainsi à plus d'attachements à N2 quand l'antécédent est animé et la relative introduite par *à qui* plutôt que *auquel*.

2. 2. 4. La fonction syntaxique (Hypothèse 2e)

Dans les études en compréhension (voir p. ex. Carreiras & Clifton, 1993, 1999 ; Cuetos & Mitchell, 1988 ; Zagar *et al.*, 1997), la fonction syntaxique de N1 dans la principale est le plus souvent sujet ou objet. Ariel (1999) a montré, dans une étude de corpus sur des relatives non ambiguës en hébreu, l'emploi plus fréquent d'un pronom résomptif par rapport à une anaphore zéro, dès lors que l'antécédent de la proposition relative a une fonction syntaxique autre que sujet.

Givón (1990) propose une hiérarchie de topicalité pour la fonction syntaxique : la fonction sujet est la plus haute et signale un référent topique, l'objet est le topique secondaire et enfin l'objet indirect et les obliques sont non topiques et les moins accessibles.

J'ai soulevé un problème à propos de l'étude d'Ariel (1999) : l'emploi d'un pronom résomptif est selon la norme incorrect quand la fonction sujet est relativisée, et une anaphore zéro est incorrecte pour les fonctions basses. Il est néanmoins possible que la distribution de ces relativiseurs traduit en fait une grammaticalisation, en quelque sorte, de l'accessibilité qu'ils signalent.

Si on élargit la hiérarchie de Keenan et Comrie (1987) aux fonctions syntaxiques dans la proposition principale, la fonction sujet est la plus accessible, suivie de la fonction objet direct, de l'objet indirect, etc. On peut considérer que plus N1 a une fonction syntaxique basse, moins il est accessible. Par conséquent, on devrait observer une fréquence plus importante de propositions relatives introduites par *lequel* quand l'antécédent, N1, a une fonction syntaxique basse.

Dans son étude de corpus, Kister (2002) évalue l'effet du rôle syntaxique de N1 sur les préférences d'attachement, pour des propositions relatives en *qui*. La fréquence d'attachement à N1 est la plus importante (74 %) lorsque N1 remplit la fonction sujet et diminue si la fonction syntaxique baisse (objet direct 55 %, objet indirect 60 %, autres 45 %). Ces résultats suggèrent que les propositions relatives sont attachées de façon préférentielle à des antécédents qui remplissent un rôle grammatical plus saillant.

On devrait observer une interaction entre la forme du relativiseur et la fonction syntaxique de N1. Considérons tout d'abord les relatives avec *qui*. Ce relativiseur signale un attachement à un antécédent hautement accessible. Si N1 remplit une fonction syntaxique élevée dans la hiérarchie, il est plus accessible que N2. On s'attend alors à un attachement à N1. Quand la fonction syntaxique de N1 baisse, son niveau d'accessibilité baisse comparativement à N2, qui devient désormais un site d'attachement plus probable. Les relatives introduites par *lequel* sont attachées préférentiellement à un antécédent peu accessible. C'est pourquoi on s'attend à peu de relatives avec *lequel* pour des phrases où N1 est sujet, qui plus est avec un attachement à N1. Les relatives avec *lequel* sont attachées de façon préférentielles à l'antécédent moins accessible, N2 quand N1 remplit une fonction syntaxique élevée, mais cette préférence devient moindre si N1 a une fonction syntaxique basse, car la différence d'accessibilité entre N1 et N2 est réduite.

Pour les relatives avec *à qui* et *auquel*, on s'attend à des préférences équivalentes, puisqu'ils signalent un même niveau d'accessibilité. Les préférences devraient ressembler à celles observées dans l'étude de Kister (2002).

2. 3. Les interactions entre facteurs

J'ai tenté dans les points précédents de séparer les effets des différents facteurs. Ces derniers peuvent interagir. Par exemple, si on reprend le caractère *given-new* qui caractérise les entités du discours et qui est important pour l'accessibilité des référents du discours, on constate qu'il est tout aussi influent sur la fonction syntaxique des constituants de la phrase. En effet, les informations déjà connues, *given*, sont le plus souvent situées en début de phrase, remplissant la fonction sujet (pour les langues de type sujet-verbe-objet comme le français). Les entités nouvelles, *new*, sont introduites plus tard, notamment en position d'objet (Branigan *et al.*, 2008 ; Temperley, 2007). Elles sont aussi souvent plus longues que les entités données. Temperley (2007) analyse le lien entre la longueur d'un GN avec un article indéfini (*i. e.* nouveau dans le discours) et sa fonction syntaxique (sujet ou objet), afin d'estimer si la longueur de ce GN est guidée plus par le caractère nouveau ou par la fonction syntaxique, sujet ou objet. C'est pour la fonction objet plutôt que sujet que le GN est le plus long, quel que soit son degré de nouveauté dans le discours. La longueur de l'expression référentielle est plus contrainte par la fonction syntaxique que par la distinction *given-new*.

Analyser toutes les interactions entre ces facteurs entraînerait une explosion de la complexité de l'analyse et par conséquent une forte diminution du nombre d'occurrences dans chacune des sous-catégories. Néanmoins, l'analyse statistique, les modèles linéaires mixtes, révèlent des interactions significatives entre ces facteurs.

3. Le corpus

3. 1. Corpus original et prétraitements

Le matériel qui constitue le corpus est issu d'articles publiés dans le journal « Le Monde » de 1987 à 1996. Ce journal, dont le registre est formel, est constitué de productions écrites planifiées.

Au départ, les structures sont extraites d'une portion du corpus en choisissant les mois de janvier sur les dix années couvertes, ce qui représente environ 2 millions de mots. Cependant, comme il y avait trop peu de phrases contenant une structure « N1 de N2 » suivie d'un relativiseur autre que *qui*, le corpus a été augmenté à deux reprises, en incluant les mois de juillet et d'octobre.

Le corpus est dans un premier temps nettoyé pour supprimer le codage des articles ou des erreurs de ponctuation. Ensuite, le corpus est étiqueté à l'aide de TreeTagger (Schmid, 1994), qui permet après segmentation des mots du corpus d'en déterminer la catégorie et qui en calcule le lemme. La catégorie grammaticale d'un mot est déterminée à partir d'un arbre de décision binaire qui calcule la probabilité que ce mot appartienne à telle ou telle catégorie en fonction des étiquettes précédentes. Pour un mot donné, la probabilité qu'il soit un nom est importante s'il est précédé d'un adjectif, lui-même précédé d'un article. C'est la catégorie avec la plus grande probabilité qui est conservée pour la suite de l'étiquetage. Son taux de succès est de l'ordre de 96 % (Schmid, 1994).

C'est à partir des entrées lexicales étiquetées que sont extraites les catégories syntaxiques de ces entrées, à l'aide de Lexique 3 (New *et al.*, 2001). Cette base de données comporte 135 000 mots du français auxquels sont associées leurs catégories grammaticales, utilisées ici, ainsi que d'autres informations comme le nombre de syllabes, la représentation orthographique ou phonémique. C'est avec Lexique 3 qu'un programme classe les phrases contenant les différents relativiseurs en fonction des mots précédant ou suivant les relativiseurs. Le programme sélectionne les phrases contenant *qui*, *lequel* (mais aussi *laquelle*, *lesquels*, *lesquelles*), *à qui* et *auquel* (avec *à laquelle*, *auxquels* et *auxquelles*).

Si on veut extraire de ce corpus des phrases contenant des propositions relatives, il faut distinguer les relativiseurs (*qui*, *que*, etc.) des pronoms interrogatifs (63). Aussi une liste de contraintes ou cas particuliers est entrée à la main pour faciliter la sélection des phrases cibles contenant des propositions relatives. Par exemple, pour éviter de retenir des complétives (64), le relativiseur *que* ne doit pas être directement précédé d'un verbe. La liste des prépositions qui peuvent précéder les relativiseurs est aussi encodée (*à*, *pour*, *de...*). Seules les relatives sujet (*qui*, *lequel*) et OI (*à qui*, *auquel*) sont analysées par la suite.

(63) *Lequel a tiré le premier ?*

(64) *Le PS affirme que c'est faux.*

À l'issue de l'étiquetage avec TreeTagger, un programme utilise Lexique 3 ainsi qu'une liste de contraintes pour extraire les phrases à propositions relatives. Les phrases rejetées car codées comme interrogatives, complétives, etc. sont ensuite vérifiées manuellement, afin de s'assurer de l'efficacité de l'extraction. Chaque phrase classée comme « avec relative » est vérifiée manuellement. Les phrases rejetées ne contiennent en effet pas de relatives, mais bien

des interrogatives, directes ou non, ainsi que des complétives. Il y a aussi peu de phrases considérées à tort comme contenant des relatives.

Ensuite, on procède à une sélection manuelle des phrases, parmi les phrases contenant *qui*, *lequel*, *à qui* ou *auquel*, où on choisit les propositions relatives dont le site d'attachement fait partie d'un GN « N1 de N2 ». Seules les phrases où « de N2 » est un vrai génitif sont conservées pour l'analyse ultérieure. Cette sélection est réalisée pour les relativiseurs *qui*, *lequel*, *à qui* et *auquel* (et leurs pendants féminin et pluriel). Enfin, différents aspects sont encodés, manuellement.

3. 2. Codage du corpus

Tout d'abord, en ce qui concerne les antécédents N1 et N2, sont codées l'animation (animé, non-animé) et la présence ou non d'un modifieur (adjectif, adverbe). Pour un GN complexe contenant un *de N* non génitif, comme dans *l'homme de confiance d'un chevalier*, je considère que *l'homme* est N1 modifié par *de confiance* et que *un chevalier* est N2.

Sont exclues les phrases où l'antécédent de la relative appartient à une structure « N1 de N2 », mais où il s'agit de noms composés comme *pomme de terre* ou bien de relations de substance (le pull de coton, cf. Gilboy *et al.*, 1995) ou de quantité (*le verre d'eau*), mais aussi les cas où N1 est un quantificateur (*le groupe de gamins*, aussi exclu dans l'analyse de Kister, 2002) ou quand le GN contient une date (*la loi de 1957*).

Comme l'analyse porte sur les relatives non restrictives, il faut contrôler cet aspect dans les phrases retenues. Les relatives précédées d'une marque de ponctuation sont considérées comme non restrictives, tout comme celles qui peuvent être supprimées sans gêner la compréhension.

En ce qui concerne la forme des expressions référentielles (p. ex. définitude), les antécédents sont codés comme défini, indéfini, démonstratif ou possessif selon l'article qui les précède et les noms propres sont codés comme tels. Un nom de pays, un acronyme désignant une organisation ou un groupe sont codés comme des noms propres et non-animés.

La fonction syntaxique de N1 est codée en partie d'après la classification de Givón (ch. 20, 1990), avec trois catégories : sujet, objet direct et « OI et obliques ». J'ajoute une autre catégorie que j'appelle « sans fonction ». Quand la structure « N1 de N2 » est dans une incise (65), s'il s'agit d'une phrase sans verbe (66) ou bien d'une apposition, j'assigne à N1 la catégorie « sans fonction ».

(65) *Ceux-ci envoient alors trois véhicules blindés et des camions militaires remplis de manifestants vers le siège d'Ostankino - principale source d'information des Russes, à laquelle l'opposition n'avait plus accès.*

(66) *Deux entretiens de Philip Roth, qui, dans l'un, s'interroge sur lui-même, dans l'autre, s'adresse à Primo Levi, chimiste, industriel, romancier, qui vient de mourir.*

L'attachement de la proposition relative est identifié par l'accord en genre et nombre entre l'antécédent et le relativiseur (*lequel, auquel*) et le verbe de la relative (*qui, à qui*). Le caractère animé de *à qui* permet aussi d'établir l'attachement si un seul antécédent est animé. Sinon, le contenu sémantique de la phrase suffit à établir l'attachement et aucun cas ambigu n'est répertorié. La ponctuation au début de la relative ainsi que la longueur de cette dernière sont aussi annotées.

3. 3. Taille du corpus

Les propositions relatives considérées dans cette étude de corpus sont des relatives non restrictives précédées d'une structure « N1 de N2 ». Les fonctions relativisées sont la fonction sujet (*qui, lequel*) et la fonction OI (*à qui, auquel*). 1937 phrases sont analysées : on trouve parmi elles 1162 occurrences avec *qui*, 149 avec *lequel* (et ses déclinaisons), 69 avec *à qui* et enfin 557 avec *auquel* (et ses déclinaisons).

La taille de ce corpus est de l'ordre de grandeur de ceux qui sont publiés et qui vont de 32 occurrences pour des ambiguïtés avec trois sites d'attachement possibles (Gibson & Schütze, 1999) à 2544 pour celles présentées dans l'étude de Desmet et Gibson (2003).

4. Résultats

La majorité des études de corpus (contrairement à la linguistique computationnelle) ne réalise pas de test statistique pour évaluer la différence dans les préférences d'attachement ; seules les fréquences moyennes d'occurrence sont répertoriées (Baltazart & Kister, 2000 ; Desmet, Brysbaert *et al.*, 2002 ; Desmet *et al.*, 2006 ; Gibson & Schütze, 1999 ; Kister, 1999, 2002), mais on peut noter toutefois que Spivey-Knowlton et Sedivy (1995) font une analyse de χ^2 pour calculer une interaction dans le cas d'ambiguïtés d'attachement de groupes prépositionnels. Étant donné qu'il y a souvent plus de deux modalités pour les facteurs observés, une mesure de χ^2 , est impossible. Une analyse de la variance ne peut non plus être

réalisées puisqu'il s'agit de variables catégorielles et que seules des variables continues peuvent être testées.

Je propose une analyse des données quelque peu différente des études de corpus que l'on peut retrouver dans la littérature, puisque j'étaye la présentation des fréquences d'attachement d'une analyse statistique avec les modèles linéaires mixtes (Baayen, à paraître ; D. M. Bates & DebRoy, 2004 ; Harville & Mee, 1984 ; Henderson, 1975). Ils permettent d'évaluer la contribution statistiquement significative de ces facteurs sur les attachements de propositions relatives et sont adaptés à la fois pour des variables dépendantes continues et catégorielles.

Un modèle linéaire mixte est une combinaison d'effets fixes, les prédicteurs, et d'effets aléatoires, ici les références des textes d'où sont issues les phrases du corpus. La procédure consiste à adapter les paramètres du modèle pour obtenir un ajustement optimal de ces derniers et expliquer les données, en se basant sur le maximum de vraisemblance. Cette analyse est réalisée à l'aide du logiciel R qui utilise l'approximation de Laplace pour maximiser le log de vraisemblance.

La procédure d'analyse est la suivante (Demberg & Keller, à paraître ; Jaeger, 2008) : tout d'abord, en plus de la variable dépendante (l'attachement, à N1 ou à N2) et de la variable aléatoire (la référence de texte), les prédicteurs sont ajoutés au modèle, ainsi que toutes les doubles interactions. Les prédicteurs sont : le Type de relative (les *qui*, *qui* et *à qui*, et les *lequel*³⁰, *lequel* et *auquel*), la Fonction du relativisé (sujet ou OI), la longueur de la relative, la présence d'un modifieur de N1 ou de N2, l'Animation (animé, non-animé), ainsi que la Fonction syntaxique de N1 (sujet, objet direct, « OI et obliques » et « sans fonction »). A chaque pas de l'analyse, les résultats renvoient une liste des effets fixes (effets principaux et interactions) auxquels sont associées une valeur de z et une valeur de p . L'interaction la moins significative est retirée du modèle qui est à nouveau testé. La procédure est incrémentale et le modèle le plus vraisemblant, celui qui explique le mieux les données, est celui où il n'y a plus que des effets significatifs, pour les effets principaux et les interactions. Les effets principaux des prédicteurs où la valeur de z est non significative ne sont conservés que si leur interaction avec un autre prédicteur est significative.

Trois modèles sont réalisés selon cette procédure : le premier traite toutes les propositions relatives non restrictives, les deuxième et troisième modèles comparent les propositions relatives sujet d'un côté (*qui*, *lequel*) et OI de l'autre (*à qui*, *auquel*). Le premier modèle, le

³⁰ Toutes les déclinaisons de *lequel* et *auquel* sont bien sûr prises en compte.

modèle global, fournit des informations sur les préférences d'attachement en général, ce qui procure une évaluation de l'Hypothèse 2 et des différents facteurs indépendamment de la forme du relativiseur. Les interactions de la Forme du relativiseur avec les autres facteurs est néanmoins intéressante. On verra qu'il est important de séparer les relatives sujet des relatives OI, parce que les préférences d'attachement diffèrent, comme c'est le cas dans l'étude de questionnaires. Tous les prédicteurs (énumérés plus haut) et leurs interactions doubles sont introduits dans la première itération des trois modèles. La variable dépendante est l'attachement, à N1 ou à N2.

Un quatrième modèle est généré, et ne prend en compte que les relatives OI dont le site d'attachement est animé (relative OI – animé). Comme les données sont plus réduites, il est impossible d'utiliser la procédure décrite plus haut pour estimer le meilleur modèle. En fait, il est possible d'aboutir au meilleur modèle en utilisant la procédure inverse, en ajoutant incrémentalement les prédicteurs, en ne conservant que ceux qui sont significatifs. Cette méthode est couramment utilisée pour les régressions classiques.

4. 1. Ensemble des propositions relatives non restrictives

Pour les propositions relatives non restrictives, les tableaux contenant les fréquences d'occurrence ne sont pas rapportés, afin d'éviter la répétition de ces tableaux lors de la présentation des résultats pour les relatives sujet et OI. On peut retrouver les fréquences en compilant les fréquences obtenues pour les deux types de relatives. La direction des effets observés est indiquée par la valence du coefficient associé à chaque prédicteur dans le Tableau 6, en fonction des niveaux de ces prédicteurs.

Le Tableau 6 rapporte les coefficients et les niveaux de significativité pour le meilleur modèle obtenu. Conventionnellement, l'interaction entre deux prédicteurs est symbolisée avec « : » ; seuls les effets significatifs sont discutés.

Considérons d'abord les effets principaux significatifs seuls (cf. Demberg & Keller, à paraître). La valence des coefficients indique la direction de l'effet : si elle est positive, il y a plus d'attachements vers N2 pour le deuxième niveau du prédicteur (p. e. pas d'adjectif pour N1 dans le Tableau 6). Par opposition, une valence négative, comme par exemple pour la présence d'un adjectif pour N2, signifie qu'il y a moins d'attachements à N2 pour le deuxième niveau du prédicteur, c'est-à-dire « sans adjectif ».

Tableau 6 : Coefficients et niveaux de significativité pour un modèle de l'attachement de propositions relatives non restrictives contenant les différents prédicteurs et leurs interactions

Effets fixes	Coefficient	<i>p</i>	Significativité
(Résidu)	-1,60	< 0,0001	***
Fonction de la relative (OI, sujet)	-0,25	<i>n. s.</i>	<i>n. s.</i>
Forme du relativiseur (les <i>qui</i> les <i>lequel</i>)	1,90	< 0,0001	***
Longueur de la relative	0,01	<i>n. s.</i>	<i>n. s.</i>
Adjectif pour N1 (adjectif, pas d'adjectif)	0,49	< 0,0001	***
Adjectif pour N2 (adjectif, pas d'adjectif)	-0,56	< 0,0001	***
Animation de N1 (animé, non-animé)	1,74	< 0,0001	***
Animation de N2 (animé, non-animé)	-0,58	0,03	*
Fonction syntaxique de N1 (sujet, objet direct)	1,77	< 0,0001	***
Fonction syntaxique de N1 (sujet, OI et obliques)	0,98	0,004	**
Fonction syntaxique de N1 (sujet, sans fonction)	2,59	< 0,0001	***
Fonction de la relative : Forme du relativiseur	-1,50	0,003	**
Fonction de la relative : Animation de N1	1,58	0,0004	***
Fonction de la relative : Animation de N2	0,51	0,04	*
Forme du relativiseur : Animation de N1	-2,11	< 0,0001	***
Longueur de la relative : Animation de N2	-0,07	0,002	**
Animation de N1 : F. synt. de N1 (sujet, objet direct)	-0,64	0,2	<i>n. s.</i>
Animation de N1 : F. synt.de N1 (sujet, OI et obliques)	0,16	<i>n. s.</i>	<i>n. s.</i>
Animation de N1 : F. synt.de N1 (sujet, sans fonction)	-2,01	0,0002	***

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$; *n. s.* : non significatif ; F. synt. : Fonction syntaxique

Dans le modèle linéaire mixte généré (Tableau 6), la Forme du relativiseur est un bon prédicteur. Comme on peut le voir dans le Tableau 7, il y a plus d'attachements à N2 quand la relative est introduite par *lequel* ou *auquel* (et leurs déclinaisons), 58 %, que par *qui* ou *à qui*, 55 %. Les relativiseurs comme *lequel* et *auquel* et leurs déclinaisons sont dans l'ensemble moins atténués (la paire *auquel* et *à qui* est le seul cas où l'atténuation est contrôlée) et plus informatifs (c'est le cas pour *lequel* par rapport à *qui*) et signalent un attachement à l'antécédent le moins accessible.

Tableau 7 : Fréquence globale d'attachements à N1 et N2 selon la forme du relativiseur (relatives non restrictives)

F. relativiseur	Attachement à N1		Attachement à N2		Total
	Nb.	%	Nb.	%	
Qui - A qui	551	45	680	55	1231
Lequel - Auquel	300	42	406	58	706
Total	851	44	1086	56	1937

F. relativiseur : forme du relativiseur ; Nb. : nombre d'occurrences

Notons que, sur l'ensemble des propositions relatives non restrictives, il y a plus d'attachements à N2 qu'à N1 (56 % vs 44 %). Il y a, en production, une préférence pour un attachement à N2.

En ce qui concerne la présence d'un adjectif ou autre modifieur, on obtient un effet significatif de la Modification de N1 et N2 (Tableau 6). Les fréquences d'occurrences obtenues (Tableau 12) montrent plus d'attachements à N1 quand il est modifié (52 %) que s'il ne l'est pas (41 %). Par contre, si N2 est modifié (62 %), il y a plus d'attachements à cet antécédent que s'il ne l'est pas (52 %).

Tableau 8 : Fréquence globale d'attachements à N1 et N2 selon la présence ou l'absence d'un adjectif (relatives non restrictives)

F. relativiseur	N1	N2				Total
		Adjectif		Sans adjectif		
		Att. N1	Att. N2	Att. N1	Att. N2	
Qui – A qui	Adjectif	46 (15)	64 (21)	104 (35)	85 (28)	299
	Sans adjectif	143 (15)	222 (24)	258 (28)	309 (33)	932
Lequel - Auquel	Adjectif	39 (20)	31 (16)	64 (34)	57 (30)	191
	Sans adjectif	65 (13)	155 (30)	132 (26)	163 (32)	515
	Total	293 (15)	472 (24)	558 (29)	614 (32)	1937

Le pourcentage d'attachements est indiqué entre parenthèses ; F. relativiseur : forme du relativiseur ; Att. : attachement

Ces résultats sont incompatibles avec la modifiabilité (Thornton *et al.*, 1999) puisqu'ils vont à l'opposé des prédictions proposées par ces auteurs, mais ils confirment, au moins en partie, l'hypothèse de Colonna (2001). En effet, on se serait attendu à une préférence pour N2

quand ce dernier est modifié, ce qui est le cas, mais pas à un effet de la modification de N1. Ces résultats ne valident pas non plus l'Hypothèse 2b sur l'effet de la modification.

L'Animation est un bon prédicteur du modèle. Dans le Tableau 9, on voit qu'il y a plus d'attachements à N2 quand N1 est non-animé (60 %, contre 35 %), mais aussi quand N2 est animé (66 %, non-animé : 47 %). De façon générale, un antécédent animé est le site d'attachement le plus fréquent, ce qui confirme l'Hypothèse 2c.

Tableau 9 : Fréquence globale d'attachements à N1 et N2 selon l'animation des antécédents (relatives non restrictives)

F. relativiseur	N1	N2				Total
		Animé		Non-animé		
		Att. N1	Att. N2	Att. N1	Att. N2	
Qui – A qui	Animé	77 (28)	40 (14)	130 (47)	31 (11)	278
	Non-animé	110 (12)	391 (41)	234 (25)	218 (23)	953
Lequel - Auquel	Animé	15 (14)	22 (20)	44 (40)	30 (27)	111
	Non-animé	54 (9)	121 (20)	187 (31)	233 (39)	595
	Total	256 (13)	574 (30)	595 (31)	512 (26)	1937

Le pourcentage d'attachements est indiqué entre parenthèses ; F. relativiseur : forme du relativiseur ; Att. : attachement

Pour l'évaluation de l'effet de la Fonction syntaxique de N1, le prédicteur ayant plus de deux niveaux, les fonctions objet direct, autres (*i. e.* objet indirect et oblique) et sans fonction sont comparées à la fonction sujet, choisie comme pivot parce qu'elle est susceptible de différer le plus des autres (cf. accessibilité et hiérarchie des fonctions syntaxiques). Les coefficients positifs obtenus pour les trois comparaisons montrent qu'il y a plus d'attachements à N2 pour une fonction autre que la fonction sujet (objet direct, autres et sans fonction). Les fréquences d'occurrences sont reportées dans le Tableau 10.

Tableau 10 : Fréquence globale d'attachements à N1 et N2 selon la fonction syntaxique de N1 (relatives non restrictives)

	Fonction syntaxique de N1								Total
	Sans fonction		Sujet		Objet direct		OI et obliques		
F. relativiseur	Att. N1	Att. N2	Att. N1	Att. N2	Att. N1	Att. N2	Att. N1	Att. N2	
Qui – A qui	34 (3)	45 (4)	182 (15)	68 (6)	116 (9)	220 (18)	219 (18)	347 (28)	1231
Lequel - Auquel	16 (2)	22 (3)	46 (7)	27 (4)	80 (11)	119 (17)	158 (22)	238 (34)	706
Total	50 (3)	67 (3)	228 (12)	95 (5)	196 (10)	339 (18)	377 (19)	585 (30)	1937

Le pourcentage d'attachements est indiqué entre parenthèses ; F. relativiseur : forme du relativiseur ; Att. : attachement ; OI : objet indirect

On trouve 29 % d'attachements à N2 quand N1 est sujet, 79 % d'attachements à N2 s'il ne remplit pas de fonction, 63 % s'il est l'objet direct et 55 % s'il est OI ou oblique. Ces résultats confirment que lorsque N1 est très accessible, c'est-à-dire sujet, il est alors le site d'attachement préféré. Pour une fonction syntaxique plus basse (Givón, 1990), il n'est guère plus accessible que N2 qui devient alors le site d'attachement préféré. L'Hypothèse 2e est validée, telle qu'elle est formulée indépendamment de la forme des relativiseurs.

Que la fonction relativisée (relative sujet ou OI) ne soit pas un prédicteur du modèle suggère que les préférences observées le sont indépendamment d'elle.

Pour l'interprétation des interactions, il faut à la fois prendre en compte la valence de l'interaction, mais aussi celle des effets principaux significatifs. La représentation graphique de l'interaction est aussi utile. L'interaction entre la Fonction de la relative et la Forme du relativiseur est significative. Sur le graphe d'interaction représenté sur la Figure 9, on peut voir les pourcentages d'attachement à N2. Il y a plus d'attachements à N2 pour la fonction sujet quand la relative est introduite par *lequel*, par comparaison à une relative introduite par *qui*. Pour la fonction OI, c'est l'inverse qui est mesuré, avec plus d'attachements à N2 quand la relative est introduite par *à qui* plutôt que *auquel*.

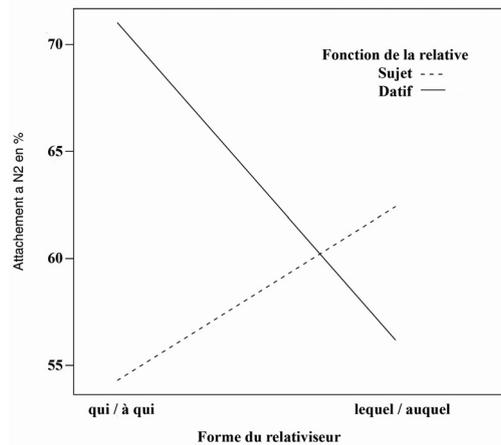
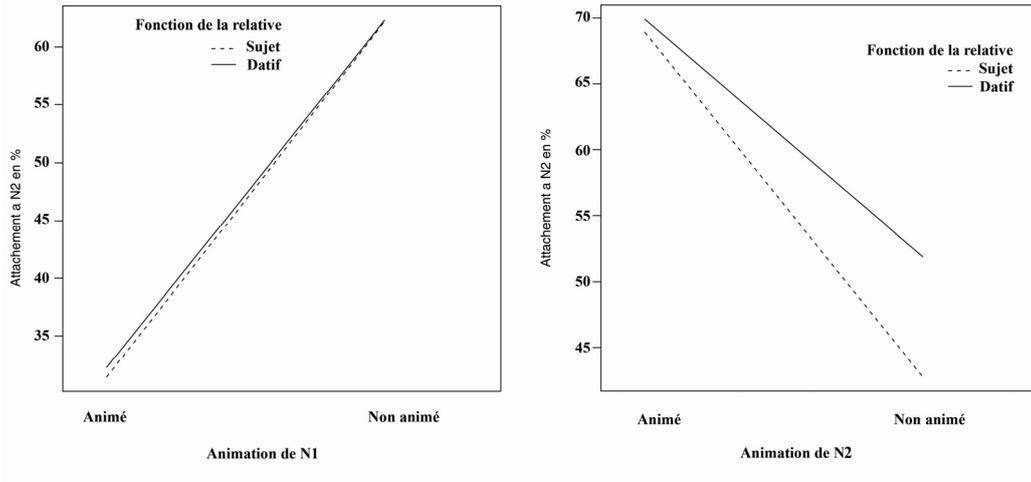


Figure 9 : Graphe d'interaction entre la fonction de la relative et la forme du relativiseur (relatives non restrictives)

Bien que, selon que N1 est animé ou non, la différence soit proche entre les relatives sujet et les relatives OI, l'interaction entre la Fonction de la proposition relative et l'Animation de N1 est significative (Figure 10a). Quand N1 est animé il y a une différence entre les relatives sujet (31 % d'attachements à N2) et les relatives OI (32 %). Lorsque N1 est non animé, on observe une préférence de N2 comparable pour les deux types de relatives (62 %).

L'interaction entre la Fonction de la relative et l'Animation de N2 (Figure 10b) est elle aussi significative. On observe une préférence pour N2 proche avec N2 animé, pour les relatives sujet (69 %) et les relatives OI (70 %), mais que N2 soit non-animé conduit à plus d'attachements à cet antécédent pour les relatives OI (52 %) que pour les relatives sujet (43 %).



a.

b.

Figure 10 : Graphe d'interaction entre la fonction de la relative et l'animation de N1 (a) et de l'animation de N2 (b) (relatives non restrictives)

Je reviendrai par la suite sur l'effet de l'Animation des antécédents en fonction du type de relative, lorsque les relatives, sujet et OI, seront présentées de façon isolée.

L'interaction entre la Forme du relativiseur et l'Animation de N1 est elle aussi significative (Figure 11) : la fréquence d'attachements est plus affectée par l'animation de N1 pour les relatives introduites par les *qui* (*qui*, *à qui*) que celles introduites par les *lequel* (*lequel*, *auquel*). Le pourcentage d'attachements à N2 est plus faible avec les *qui* si N1 est animé. Cet effet est peut-être dû à la restriction à un antécédent animé pour les relatives avec *à qui*, alors que les autres relativiseurs ne distinguent pas l'animation. Je reviendrai plus tard sur cet aspect.

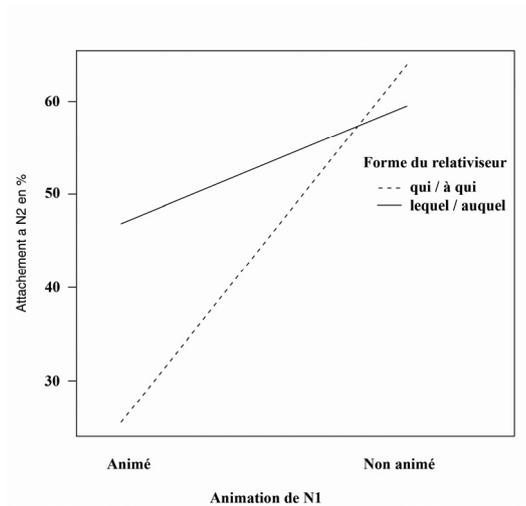


Figure 11 : Graphe d'interaction entre la forme du relativiseur et l'animation de N1 (relatives non restrictives)

L'interaction entre l'Animation de N2 et la Longueur de la relative est significative. Cette interaction est calculée en prenant la Longueur de la relative qui est une variable continue. Afin de rendre ce résultat plus lisible, les occurrences sont rassemblées dans trois catégories de Longueur dans le Tableau 11, en fonction de l'Animation de N2. Ces catégories sont établies d'après celles de Toussanel et Abeillé (2001, cité par Colonna, 2001), avec des propositions relatives courtes (≤ 5 mots), longues (≥ 10 mots) ou intermédiaires.

Tableau 11 : Fréquence globale d'attachements à N1 et N2 selon la longueur de la proposition relative et l'animation de N2

	Longueur de la relative						Total
	Courte		Moyenne		Longue		
N2	Att. N1	Att. N2	Att. N1	Att. N2	Att. N1	Att. N2	
Animé	44 (5)	88 (11)	93 (11)	236 (28)	119 (14)	250 (30)	830
Non-animé	92 (8)	107 (10)	248 (22)	237 (21)	255 (23)	168 (15)	1107
Total	136 (7)	195 (10)	341 (18)	473 (24)	374 (19)	418 (22)	1937

Le pourcentage d'attachements est indiqué entre parenthèses ; F. relativiseur : forme du relativiseur ; Att. : attachement

Quand la relative est courte et que N2 est animé, on mesure une forte préférence pour N2, avec 67 % des attachements. Cette préférence demeure importante même quand la relative est longue (68 %). Mais, quand N2 est non-animé, la préférence pour N2 est plus faible si la relative est courte (54 %) et décroît lorsque la taille de la relative augmente (40 %).

Enfin, la dernière interaction significative est entre l'Animation de N1 et la Fonction syntaxique de N1 (ici, la comparaison entre la fonction sujet et l'absence de fonction), car elle ne l'est pas entre la fonction sujet et les fonctions objet direct ou autres.

Il est intéressant de voir que pour toutes les fonctions syntaxiques, il y a plus d'attachements à N2 avec N1 non-animé, sauf quand N1 est sans fonction. Dans ce cas précis, N1 est très peu accessible et son caractère animé ne suffit pas à le rendre plus accessible que N2, site d'attachement préféré (car plus accessible). Par contre, quand N1 ne remplit pas de fonction syntaxique, il y a une préférence pour N1 quand ce dernier est non-animé.

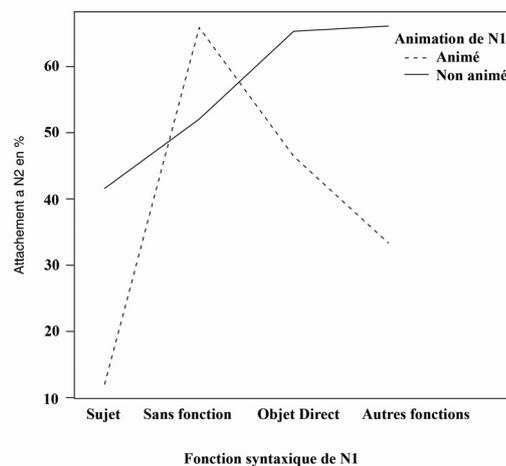


Figure 12 : Graphe d'interaction entre l'animation et la fonction de N1

Ce résultat confirme l'Hypothèse 2d, selon laquelle N1 est le site d'attachement préféré s'il est hautement accessible et remplit une fonction élevée (Givón, 1990), d'autant plus qu'il est animé (Bock & Irwin, 1980). Quand N1 remplit une fonction basse, il est peu accessible, qu'il soit animé ou non et il n'est plus le site d'attachement préféré.

4. 2. Propositions relatives sujet

L'effet significatif de l'interaction entre la Fonction de la proposition relative et la Forme du relativiseur, mais aussi avec d'autres prédicteurs comme l'Animation de N1 et N2, indique qu'une analyse séparant les relatives sujet des relatives OI renseignerait sur les préférences d'attachement et les facteurs qui influencent ces dernières.

Une analyse de modèle linéaire mixte est conduite sur les propositions relatives sujet, comparant ainsi les relatives introduites par *qui* et par *lequel* (et ses déclinaisons). Les effets

fixes sont reportés dans le Tableau 12. Comme précédemment, la présentation des résultats débute par les effets principaux significatifs, puis les interactions significatives, accompagnée des tableaux avec les fréquences d'occurrences.

Tableau 12 : Coefficients et niveaux de significativité pour un modèle de l'attachement de propositions relatives non restrictives sujet contenant les différents prédicteurs et leurs interactions

Effets fixes	Coefficient	<i>p</i>	Significativité
(Résidu)	-1,20	< 0,0001	***
Forme du relativiseur (<i>qui, lequel</i>)	3,72	0,002	**
Adjectif pour N1 (adjectif, pas d'adjectif)	0,44	0,002	**
Adjectif pour N2 (adjectif, pas d'adjectif)	-0,60	< 0,0001	***
Animation de N1 (animé, non-animé)	1,38	< 0,0001	***
Animation de N2 (animé, non-animé)	-1,29	< 0,0001	***
Fonction syntaxique de N1 (sujet, objet direct)	1,35	< 0,0001	***
Fonction syntaxique de N1 (sujet, OI et obliques)	1,18	< 0,0001	***
Fonction syntaxique de N1 (sujet, sans fonction)	1,32	< 0,0001	***
Forme du relativiseur : Animation de N1	-1,32	0,003	**
Forme du relativiseur : F. synt. de N1 (sujet, objet direct)	-2,46	0,05	*
Forme du relativiseur : F. synt. de N1 (sujet, OI et obliques)	-2,52	0,04	*
Forme du relativiseur : F. synt. de N1 (sujet, sans fonction)	-1,71	0,2	<i>n. s.</i>

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$; *n. s.* : non significatif ; F. synt. : Fonction syntaxique

L'effet de la Forme du relativiseur est significatif : il y a plus d'attachements à N2 quand une proposition relative sujet est introduite par *lequel* plutôt que par *qui* (Tableau 13).

Tableau 13 : Fréquence globale d'attachements à N1 et N2 selon la forme du relativiseur (comparaison de *qui* et *lequel*)

F.relativiseur	Attachement à N1		Attachement à N2		Total
	Nb.	%	Nb.	%	
Qui	531	46	631	54	1162
Lequel	56	38	93	62	149
Total	587	45	724	55	1311

F. relativiseur : forme du relativiseur ; *Nb.* : nombre d'occurrences

Ce résultat confirme l'Hypothèse 1 : la forme du relativiseur signale l'accessibilité de son antécédent. Le relativiseur *lequel*, à la fois moins atténué (Hypothèse 1a) et plus informatif

(Hypothèse 1b) que *qui*, est attaché plus fréquemment à un antécédent moins accessible, N2, car ce dernier est dominé par N1, tête du GN.

Il y a aussi un effet de la Modification de N1 par un adjectif, entraînant plus d'attachements à N2 quand N1 n'est pas modifié (67 %), plutôt que modifié (59 %), comme on peut le voir dans le Tableau 14.

Tableau 14 : Fréquence globale des attachements à N1 et N2 selon la présence ou l'absence d'un adjectif (comparaison entre qui et lequel)

F. relativiseur	N1	N2				Total
		Adjectif		Sans adjectif		
		Att. N1	Att. N2	Att. N1	Att. N2	
Qui	Adjectif	44 (15)	60 (21)	103 (36)	82 (28)	289 (22)
	Sans adjectif	136 (16)	205 (23)	248 (28)	284 (33)	873 (67)
Lequel	Adjectif	8 (19)	5 (12)	10 (24)	19 (45)	42 (3)
	Sans adjectif	11 (10)	35 (33)	27 (25)	34 (32)	107 (8)
	Total	199 (15)	305 (23)	388 (30)	419 (32)	1311

Le pourcentage d'attachements est indiqué entre parenthèses ; F. relativiseur : forme du relativiseur ; Att. : attachement

La Modification de N2 (Tableau 14) affecte aussi de façon significative les préférences d'attachement : il y a plus d'attachements à cet antécédent quand il est modifié (61 %, non modifié : 52 %).

L'Animation de N1 ou de N2 est un bon prédicteur du modèle. Quand N1 est animé (Tableau 15, c'est à cet antécédent que la relative est le plus souvent attachée (58 % des attachements à N1), alors que ce n'est pas le site d'attachement préféré quand il est non-animé (29 %). Quand N2 est animé, il est lui aussi le site d'attachement préféré (69 % des attachements sont à N2), alors que non-animé, il ne l'est plus (43 %). Ces résultats confirment une fois de plus qu'un antécédent animé, plus accessible, est le site d'attachement le plus fréquent (Hypothèse 2c).

Tableau 15 : Fréquence globale des attachements à N1 et N2 selon l'animation des antécédents (comparaison entre qui et lequel)

F. relativiseur	N1	N2				Total
		Animé		Non-animé		
		Att. N1	Att. N2	Att. N1	Att. N2	
Qui	Animé	71 (28)	37 (15)	116 (45)	31 (12)	255
	Non-animé	110 (12)	345 (38)	234 (26)	218 (24)	907
Lequel	Animé	2 (5)	13 (32)	14 (34)	12 (29)	41
	Non-animé	11 (10)	35 (32)	29 (31)	33 (31)	108
	Total	194 (15)	430 (33)	393 (22)	294 (22)	1311

Le pourcentage d'attachements est indiqué entre parenthèses ; F. relativiseur : forme du relativiseur ; Att. : attachement

La Fonction syntaxique de N1 est un bon prédicteur (voir Tableau 12 et Tableau 16) : la fonction sujet diffère significativement des autres, car la fréquence des attachements à N1 est plus importante pour cette fonction (71 %) que pour les autres (sans fonction, 17 % : objet direct, 36 % ; autres, 50 %).

Tableau 16 : Fréquence globale d'attachements à N1 et N2 pour des propositions relatives introduites par qui et lequel en fonction de la fonction syntaxique de N1

F. relativiseur	Fonction syntaxique de N1								Total
	Sans fonction		Sujet		Objet direct		OI et obliques		
	Att. N1	Att. N2	Att. N1	Att. N2	Att. N1	Att. N2	Att. N1	Att. N2	
Qui	32 (3)	42 (4)	171 (15)	66 (6)	114 (10)	205 (18)	214 (18)	318 (27)	1162
Lequel	2 (1)	7 (4)	1 (1)	3 (2)	16 (11)	28 (19)	37 (25)	55 (37)	149
Total	34 (3)	49 (3)	172 (13)	69 (5)	130 (10)	233 (18)	251 (19)	373 (28)	1311

Le pourcentage d'attachements est indiqué entre parenthèses ; F. relativiseur : forme du relativiseur ; Att. : attachement ; OI : objet indirect

Quand N1 remplit la fonction sujet, il est très accessible et constitue le site d'attachement préféré (Hypothèse 2e). Si sa fonction syntaxique est plus basse, son niveau d'accessibilité est plus proche de celui de N2, alors site d'attachement préféré.

Pour les relatives sujet, deux interactions font partie du modèle linéaire mixte (Tableau 12). Il s'agit de l'interaction entre la Forme du relativiseur et l'Animation de N1 (Figure 13), et de celle entre la forme du relativiseur et la fonction syntaxique de N1 (Figure 14).

Pour la première interaction (Figure 13), on observe qu'une relative introduite par *qui* est attachée de façon plus importante à N1 quand ce dernier est animé (66 %), plutôt que non-animé (24 %).

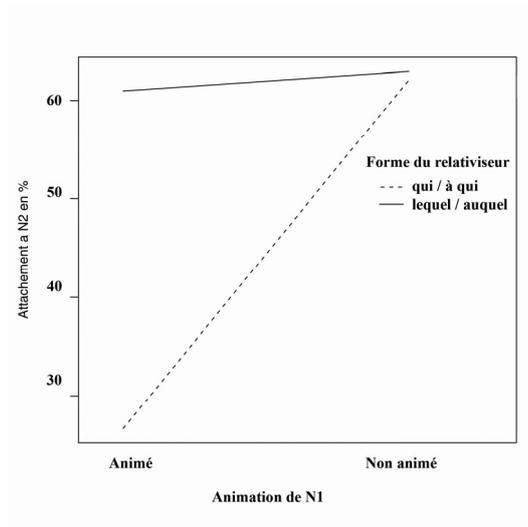


Figure 13 : Graphe d'interaction entre la forme du relativiseur et l'animation de N1 (relatives sujet)

Pour les relatives avec *lequel*, par contre, on mesure toujours une préférence pour N2, plus importante quand N1 est non-animé (87 %) que lorsqu'il est animé (76 %). Ces résultats valident l'Hypothèse 2c qui était proposée : un relativiseur comme *lequel* signale un antécédent peu accessible, c'est pourquoi N2 est alors le site d'attachement préféré, même si l'animation affecte cette préférence, mais sans que N1 devienne le site préféré pour autant. Avec une relative introduite par *qui*, c'est une préférence pour N1 qui est attendue et plus marquée encore lorsque N1 est animé. Un antécédent animé, qui plus est N1, est une entité hautement accessible.

Il y a une deuxième interaction significative (Figure 14), entre la Forme du relativiseur et la Fonction syntaxique de N1, pour la comparaison entre d'une part les fonctions sujet et objet direct et d'autre part les fonctions sujet et autres (Tableau 16).

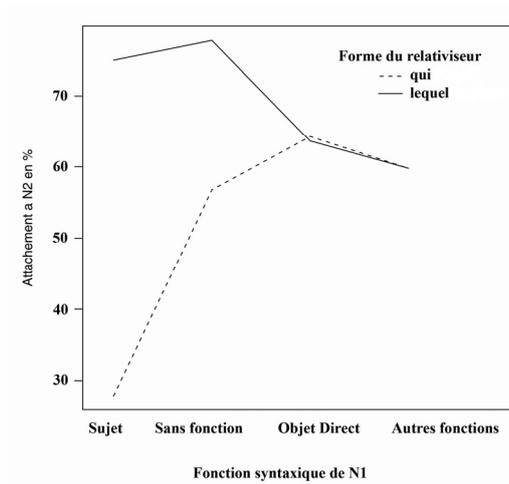


Figure 14 : Graphe d'interaction entre la forme du relativiseur et la fonction syntaxique de N1 (relatives sujet)

Si la relative est introduite par *qui*, on mesure 72 % d'attachements à N1 si ce dernier est sujet. Cette préférence est plus importante que lorsque N1 est l'objet direct de la phrase (36 %) ou OI et oblique (40 %). N1 n'est le site d'attachement préféré que s'il remplit une fonction hautement accessible, la fonction sujet.

Pour une relative avec *lequel*, que la fonction syntaxique soit élevée ne conduit pas à une préférence pour N1, puisqu'on mesure peu d'attachements à N1 quand il est sujet (25 %), objet direct (36 %) ou OI et obliques (40 %).

La forme du relativiseur signale le degré d'accessibilité de son antécédent. Quand l'antécédent est hautement accessible et que le relativiseur signale un antécédent accessible, la relative est attachée de façon préférentielle à cet antécédent. C'est ce qu'on observe pour les relatives avec *qui* où N1 est sujet. Par contre, quand le relativiseur signale un attachement à un antécédent peu accessible, *lequel*, il est attaché plus fréquemment à un antécédent peu accessible, N2, que N1 soit très accessible (sujet) ou moins accessible (objet direct et autres).

4. 3. Propositions relatives OI

Dans l'étude de questionnaires (Chapitre 1), les antécédents sont humains afin de pouvoir comparer à *qui* et *auquel*. Le relativiseur à *qui* ne peut renvoyer qu'à un antécédent animé alors que *auquel* peut être attaché à un antécédent quelle que soit son animation. Avant de présenter le modèle linéaire mixte obtenu pour les relatives OI ayant un antécédent animé

(voir en 4. 4), c'est celui qui a été obtenu sur l'ensemble des relatives OI qui est reporté maintenant.

Dans le Tableau 17, on retrouve l'ensemble des effets fixes, effets principaux et interactions d'ordre 2 significatifs retenus pour le meilleur modèle des relatives OI. Dans les tableaux contenant les fréquences d'occurrence, les relatives introduites par *auquel* sont séparées, selon que leur site d'attachement est animé ou non. Cette séparation permet de discuter la distribution des relatives avec *auquel* selon l'animation de l'antécédent et surtout de comparer les fréquences obtenues avec *à qui* et *auquel*, quand tous les deux sont attachés à un antécédent animé (cf. 4. 4).

Tableau 17 : Coefficients et niveaux de significativité pour un modèle de l'attachement de propositions relatives non restrictives OI contenant les différents prédicteurs et leurs interactions

Effets fixes	Coefficient	<i>p</i>	Significativité
(Résidu)	1,22	0,2	<i>n. s.</i>
Forme du relativiseur (<i>à qui</i> , <i>auquel</i>)	-1,00	0,01	*
Longueur de la relative	-0,23	0,01	*
Adjectif pour N1 (adjectif, pas d'adjectif)	0,64	0,002	**
Adjectif pour N2 (adjectif, pas d'adjectif)	-0,47	0,01	*
Animation de N1 (animé, non-animé)	-1,37	0,07	<i>n. s.</i>
Animation de N2 (animé, non-animé)	0,59	0,2	<i>n. s.</i>
Fonction syntaxique de N1 (sujet, objet direct)	1,18	0,0002	***
Fonction syntaxique de N1 (sujet, OI et obliques)	1,32	< 0,0001	***
Fonction syntaxique de N1 (sujet, sans fonction)	1,41	0,003	**
Longueur de la relative : Animation de N1	0,33	0,0005	***
Longueur de la relative : Animation de N2	-0,16	0,002	**

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$; *n. s.* : non significatif

On obtient un effet significatif de la Forme du relativiseur. Comme on peut le voir dans le Tableau 18, les relatives avec *à qui* sont attachées préférentiellement à N2 (71 %), alors que cette préférence est moindre pour les relatives avec *auquel* (56 %). Si ces résultats semblent contradictoires avec ce qui a pu être mis en évidence dans les Questionnaires 2 et 3, où les préférences d'attachement sont comparables selon que la relative est introduite par *à qui* ou *auquel*, c'est plutôt le modèle où seules les phrases avec un attachement à un antécédent animé sont comptabilisées qui fournit un test de l'Hypothèse 1 pour les relativiseurs *à qui* et *auquel*.

Tableau 18 : Fréquence globale d'attachements à N1 et N2 selon la forme du relativiseur (comparaison de à qui et auquel)

F. relativiseur	Attachement à N1		Attachement à N2		Total
	Nb.	%	Nb.	%	
À qui	20	29	49	71	69
Auquel (antécédent animé)	43	31	95	69	138
Auquel (antécédent non-animé)	201	48	218	52	419
Total	264	42	362	58	626

F. relativiseur : forme du relativiseur ; Nb. : nombre d'occurrences

Il y a aussi un effet de la Longueur de la relative (regroupée en trois catégories dans le Tableau 19) : plus la relative est longue, plus il y a d'attachement à N1 (longue : 47 % ; moyenne : 40 % ; courte : 39 %, d'après les fréquences totales). Ce résultat est en conformité avec ce que prédit Fodor (1998) : une relative longue est attachée de façon préférentielle à un constituant de taille comparable.

Tableau 19 : Fréquence globale d'attachements à N1 et N2 selon la longueur de la proposition relative (comparaison de à qui et auquel)

N1	Longueur de la relative						Total
	Courte		Moyenne		Longue		
	Att. N1	Att. N2	Att. N1	Att. N2	Att. N1	Att. N2	
À qui	0	2	8	22	12	25	69
Auquel (antécédent animé)	5	23	18	39	20	33	138
Auquel (antécédent non-animé)	52	64	84	103	65	51	419
Total	57	89	110	164	97	109	626

Le pourcentage d'attachements est indiqué entre parenthèses ; F. relativiseur : forme du relativiseur ; Att. : attachement

La présence d'un adjectif pour N1 et N2 est un prédicteur significatif du modèle (Tableau 17), comme on peut aussi le voir dans le Tableau 20 qui présente les fréquences d'occurrence des relatives en fonction de la modification des antécédents.

Tableau 20 : Fréquence globale d'attachements à N1 et N2 selon la présence ou l'absence d'un adjectif (comparaison entre à qui et auquel)

F. relativiseur	N1	N2				Total
		Adjectif		Sans adjectif		
		Att. N1	Att. N2	Att. N1	Att. N2	
À qui	Adjectif	2 (20)	4 (40)	1 (10)	3 (30)	10
	Sans adjectif	7 (12)	17 (29)	10 (17)	25 (42)	59
Auquel (antécédent animé)	Adjectif	7 (20)	9 (26)	9 (26)	10 (29)	35
	Sans adjectif	8 (8)	33 (32)	19 (18)	43 (42)	103
Auquel (antécédent non-animé)	Adjectif	24 (21)	17 (15)	45 (39)	28 (25)	114
	Sans adjectif	46 (15)	87 (29)	86 (28)	86 (28)	305
	Total	94 (15)	167 (27)	175 (28)	195 (31)	626

Le pourcentage d'attachements est indiqué entre parenthèses ; F. relativiseur : forme du relativiseur ; Att. : attachement

De façon générale, pour une relative OI, il y a plus d'attachements à N2 si N1 est non modifié (62 %) que modifié (45 %), mais aussi quand N2 est modifié (64 %, non modifié : 53 %). Comme pour les relatives sujet et le modèle global, ces résultats infirment l'hypothèse de la modifiabilité (Thornton *et al.*, 1999), favorisant celle de Colonna (2001). Qu'un antécédent soit le site d'attachement préféré quand il est modifié contredit aussi l'Hypothèse 2b qui proposait qu'un GN sans modifieur se situerait plus haut dans la hiérarchie d'accessibilité (Ariel, 1990, 2001 ; Gundel *et al.*, 1993) qu'un GN avec modifieur et qu'il devrait donc être le site d'attachement préféré.

La Fonction syntaxique de N1 est aussi un bon prédicteur, la fonction sujet (68 % des attachements à N1) se détache significativement des fonctions objet direct (38 %), OI et obliques (37 %) et des cas où N1 ne remplit pas de fonction syntaxique (47 %). Ces résultats sont reportés dans le Tableau 21.

Tableau 21 : Fréquence globale d'attachements à N1 et N2 pour des propositions relatives introduites par à qui et auquel, en fonction de la fonction syntaxique de N1

	Fonction syntaxique de N1								Total
	Sans fonction		Sujet		Objet direct		OI et obliques		
	Att. N1	Att. N2	Att. N1	Att. N2	Att. N1	Att. N2	Att. N1	Att. N2	
F. relativiseur	2	3	11	2	2	15	5	29	69
A qui	(3)	(4)	(16)	(3)	(3)	(22)	(7)	(42)	
Auquel (antécédent animé)	5	3	13	13	4	29	21	50	138
	(4)	(2)	(9)	(9)	(3)	(21)	(15)	(36)	
Auquel (antécédent non-animé)	9	12	32	11	60	62	100	133	419
	(2)	(3)	(8)	(8)	(14)	(15)	(24)	(32)	
Total	16	18	56	26	66	106	126	212	626
	(3)	(3)	(9)	(9)	(11)	(17)	(20)	(34)	

Le pourcentage d'attachements est indiqué entre parenthèses ; F. relativiseur : forme du relativiseur ; Att. : attachement ; OI : objet indirect

Quand N1 remplit une fonction syntaxique élevée comme la fonction sujet, il est très accessible, ce qui en fait le site d'attachement préféré. Par contre, quand il remplit une fonction syntaxique plus basse, son niveau d'accessibilité baisse et il n'est plus le site d'attachement préféré. Ce résultat est conforme à ce que prévoit l'Hypothèse 2e.

Enfin, les derniers prédicteurs du modèle sont deux interactions, entre la Longueur de la relative et l'Animation, de N1 (Tableau 22) et de N2 (Tableau 23).

Tableau 22 : Fréquence globale des attachements à N1 et N2 selon l'animation de N1 et la longueur de la relative (relatives OI)

N1	Longueur de la relative						Total
	Courte		Moyenne		Longue		
	Att. N1	Att. N2	Att. N1	Att. N2	Att. N1	Att. N2	
Animé	5 (5)	12 (13)	26 (28)	13 (14)	32 (34)	5 (5)	93
Non-animé	52 (10)	77 (14)	84 (16)	151 (28)	65 (12)	104 (20)	533
Total	57 (9)	89 (14)	110 (18)	164 (26)	97 (15)	109 (17)	626

Le pourcentage d'attachements est indiqué entre parenthèses ; Att. : attachement

L'interaction entre la Longueur de la relative et l'Animation de N1, révèle que lorsque N1 est animé, le pourcentage d'attachements à N1 augmente avec la longueur de la relative

(courte : 29 % ; moyenne : 67 % ; longue : 86 %). Quand N1 est non-animé, par contre, la préférence d'attachement à N1 n'est pas influencée par la longueur de la relative (courte : 40 % ; moyenne : 36 % ; longue : 38 %).

Si on s'intéresse maintenant à l'interaction entre la Longueur de la relative et l'Animation de N2, cf. Tableau 23, on observe que, lorsque N2 est animé, cet antécédent est le site d'attachement préféré, et que cette préférence devient plus importante à mesure que la taille de la relative augmente (courte : 63 % ; moyenne : 69 % ; longue : 75 %). Le résultat observé quand N2 est non-animé va dans la direction opposée, puisqu'il y a moins d'attachements à N2 si la taille de la relative augmente (courte : 60 % ; moyenne : 56 % ; longue : 40 %).

Tableau 23 : Fréquence globale des attachements à N1 et N2 selon l'animation de N2 et la longueur de la relative

	Longueur de la relative						Total
	Courte		Moyenne		Longue		
N2	Att. N1	Att. N2	Att. N1	Att. N2	Att. N1	Att. N2	
Animé	15 (7)	25 (12)	28 (14)	61 (30)	19 (9)	58 (28)	206
Non-animé	42 (10)	64 (15)	82 (20)	103 (25)	78 (19)	51 (12)	420
Total	57 (9)	89 (14)	110 (18)	164 (26)	97 (15)	109 (17)	626

Le pourcentage d'attachements est indiqué entre parenthèses ; Att. : attachement

Aucune hypothèse n'a été formulée précédemment sur l'interaction entre l'Animation des antécédents et la Longueur de la relative. On peut considérer à partir des résultats que ces deux facteurs s'additionnent dans l'établissement du niveau d'accessibilité de l'antécédent, pour peu que l'on modifie l'hypothèse sur la relation entre la longueur de la relative et l'accessibilité de l'antécédent. Une relative longue est attachée de façon préférentielle à un antécédent très accessible, N1 (par rapport à une relative courte), mais cette préférence est d'autant plus forte que cet antécédent est accessible, quand il est animé. De façon converse, N2 n'est pas le site d'attachement préféré quand la relative est longue, et encore moins si le caractère non-animé de cet antécédent signale un faible niveau d'accessibilité.

4. 4. Propositions relatives OI – animé

Si on veut comparer les données en production avec celles en compréhension (p. ex. les questionnaires, Chapitre 1) et surtout si on applique la suggestion de Desmet *et al.* (Desmet, Brysbaert *et al.*, 2002 ; Desmet *et al.*, 2006) d'affiner les données observées dans les corpus

avec les facteurs manipulés dans les expériences en compréhension, il est nécessaire de réaliser une analyse sur les propositions relatives OI dont le site d'attachement est animé, comme c'est le cas pour les phrases testées en compréhension.

Le Tableau 24 présente les prédicteurs significatifs (marginalelement significatif pour l'effet de la Modification de N1). Notons tout d'abord que la Forme du relativiseur n'est pas un prédicteur retenu pour ce modèle, ce qui suggère que la forme du relativiseur ne contribue pas à signaler un attachement différent. Cette première observation tendrait à confirmer ce qui a pu être mesuré dans les Questionnaires 2 et 3 : *à qui* et *auquel* signalent un degré d'accessibilité et un attachement comparables, quand ils sont attachés à un antécédent animé.

Tableau 24 : Coefficients et niveaux de significativité pour un modèle de l'attachement de propositions relatives non restrictives OI dont l'antécédent est animé

Effets fixes	Coefficient	<i>p</i>	Significativité
(Résidu)	-1,01	0,02	*
Adjectif pour N1 (adjectif, pas d'adjectif)	0,74	0,06	<i>n. s.</i>
Fonction syntaxique de N1 (sujet, objet direct)	2,52	< 0,0001	***
Fonction syntaxique de N1 (sujet, OI et obliques)	1,53	0,0002	***
Fonction syntaxique de N1 (sujet, sans fonction)	0,29	<i>n. s.</i>	<i>n. s.</i>

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$; *n. s.* : non significatif

L'effet de la Modification de N1 est marginalelement significatif. Si on se reporte au Tableau 20 et aux relatives avec *à qui* et *auquel* (antécédent animé), on voit qu'il y a plus d'attachements à N1 quand celui-ci est modifié par un adjectif (42 %) que s'il n'est pas modifié (27 %). L'accessibilité de N1 est plus importante quand il est modifié.

L'effet de la Fonction syntaxique de N1 est significatif si on compare la fonction sujet avec la fonction objet direct, mais aussi OI et obliques (voir aussi le Tableau 21). La préférence pour N1 est plus importante quand N1 est sujet (62 %) qu'objet direct (12 %) ou qu'OI et obliques (30 %). La différence entre la fonction sujet pour N1 et les cas où N1 ne remplit pas de fonction syntaxique (54 %) n'est pas significative.

Ces résultats montrent que, indépendamment de la forme du relativiseur et si le site d'attachement est animé, l'attachement à N1 est important lorsqu'il est très accessible, c'est-à-dire quand il remplit la fonction syntaxique la plus élevée, la fonction sujet.

5. Discussion

La discussion des résultats de l'étude de corpus est organisée de la façon suivante : tout d'abord, je propose un récapitulatif des résultats obtenus et les mets en regard avec les prédictions et les hypothèses posées dans ce travail de thèse. Ces résultats sont aussi discutés par rapport aux prédictions d'autres théories et aux autres études de corpus de la littérature sur les ambiguïtés d'attachement des relatives.

Mais avant cela, revenons sur le modèle global et sur la notion de restrictivité. Kister (Baltazart & Kister, 2000 ; Kister, 2002) ne contrôle pas la restrictivité des relatives qu'elle analyse et c'est aussi le cas dans d'autres études de corpus (p. ex. Desmet, Brysbaert *et al.*, 2002). Des études en production orale (Holmes, 1988, 1995 ; Tao & McCarthy, 2001) ou en compréhension (Gibson *et al.*, 2005) ont mis en évidence des traitements différents selon la restrictivité. Les travaux de Loock (2003, 2007) portent justement sur 450 relatives non restrictives en anglais, mais dont l'attachement est non ambigu puisqu'un seul antécédent est possible. Si la restrictivité n'est pas testée ici, il est intéressant de confronter les résultats obtenus dans des études où la restrictivité est contrôlée ou non (Loock, 2003, 2007).

Si l'on revient au modèle linéaire mixte généré pour l'ensemble des relatives non restrictives, aucune hypothèse n'est formulée pour les propositions relatives en général. Les résultats sont cependant intéressants. On remarque tout d'abord qu'il y a une majorité d'attachements à N2, sur l'ensemble des relatives non restrictives (56 %), mais aussi pour chaque type de relativiseur. La préférence pour N2 contraste avec celle pour N1 mesurée dans le corpus de Kister (Baltazart & Kister, 2000 ; Kister, 2002) qui étudie l'attachement de relatives avec *qui* précédées de « N1 de N2 », sans toutefois contrôler la restrictivité. La différence d'attachement entre ces deux corpus peut s'expliquer par le fait que Kister comptabilise des noms composés comme *pomme de terre*, qui bien que structure « N1 de N2 » sont ici rejetés car « de N2 » n'est pas un génitif. Pour ce type de structure, la proposition relative est exclusivement attachée à N1, ce qui contribue à la préférence globale de cet antécédent. Les pourcentages d'attachement sont par ailleurs comparables entre l'étude Kister et celle qui est présentée ici.

Les préférences d'attachement varient selon certains facteurs comme la modification, l'animation, la fonction syntaxique ou la longueur de la relative.

Le modèle global et les tableaux contenant les fréquences d'occurrences révèlent que la modification des antécédents affecte l'attachement : quand un antécédent est modifié, par

exemple par un adjectif, il est le site d'attachement préféré. Sinon, c'est un attachement à N2 qui est mesuré. On observe aussi un effet de l'animation, car l'attachement à N1 et à N2 est plus important quand ces derniers sont animés que non-animés. La fonction syntaxique de N1 affecte les attachements, entraînant plus d'attachements à N1 quand ce dernier remplit une fonction syntaxique élevée, la fonction sujet, par comparaison aux autres fonctions syntaxiques (sans fonction, objet direct, objet indirect et obliques) ; on mesure alors plus d'attachements à N2. Dans l'interaction entre l'animation et la fonction syntaxique de N1, le cas où N1 ne remplit pas de fonction syntaxique se détache des autres, puisque dans ce cas l'attachement à N1 est plus fréquent lorsque N1 est non-animé qu'animé. Pour les autres fonctions syntaxiques, il y a plus d'attachements à N1 quand il est animé.

Plutôt que de discuter des interactions entre la forme du relativiseur et les autres facteurs dans le modèle global, intéressons-nous aux modèles pour les relatives sujet et OI. L'analyse séparée de ces deux fonctions est tout à fait pertinente puisque d'une part l'analyse du modèle global et de l'autre les interactions entre les facteurs testés et la forme du relativiseur témoignent de différences de traitement selon la fonction relativisée. C'est en comparant *qui/lequel* et à *qui/auquel* qu'on peut évaluer l'Hypothèse 1 sur l'effet de la forme du relativiseur. De plus, la variation des résultats entre les relatives sujet et OI sont importantes et méritent une analyse séparée.

5. 1. Validation des hypothèses opérationnelles

5. 1. 1. Effet de la forme du relativiseur (Hypothèse 1)

Au niveau des relatives sujet, on mesure une différence selon la forme du relativiseur : il y a plus d'attachements à N2 avec *lequel*³¹ (et ses déclinaisons, *laquelle*, *lesquels* et *lesquelles*) qu'avec *qui*. Ce résultat valide l'Hypothèse 1, selon laquelle un relativiseur comme *lequel* est moins atténué et plus informatif que *qui* et par conséquent situé à un niveau plus bas de la hiérarchie d'accessibilité. *Lequel* signale un attachement à un antécédent moins accessible. N2 est le moins accessible des deux sites d'attachements possibles, parce qu'il est le modifieur de N1, et c'est bien à cet antécédent qu'on mesure plus d'attachements avec *lequel*.

La comparaison entre ces deux formes de relativiseurs ne permet pas de vérifier si c'est l'atténuation (Hypothèse 1a) ou l'informativité (Hypothèse 1b), ou bien les deux, qui

contribuent à la différence mesurée. En effet, d'après l'Hypothèse 1a on a comptabilisé plus d'attachements à N2 avec *lequel* qu'avec *qui*, parce que ce dernier est plus long, moins atténué. L'Hypothèse 1b propose que c'est le degré d'informativité plus important de *lequel* qui signale un niveau d'accessibilité plus faible et un attachement à N2.

Quant à l'atténuation, toutes les déclinaisons de *lequel* (*laquelle*, *lesquels* et *lesquelles*) sont moins atténuées que *qui*. D'après l'Hypothèse 1a, *lequel* est situé plus bas dans la hiérarchie et signale un antécédent peu accessible. Il y a en effet plus d'attachements à N2 avec *lequel* qu'avec *qui*. Dans les modèles linéaires mixtes et contrairement aux études de Questionnaires (cf. Chapitre 1), la différence d'atténuation entre *à qui* et *auquel* et ses déclinaisons (*à laquelle*, *auxquels* et *auxquelles*) n'est pas contrôlée. Il est difficile de discuter alors de l'effet de l'atténuation pour les relatives OI.

Ces dernières offrent une autre évaluation de l'Hypothèse 1b, complétant celle des relatives sujet. *A qui* signale l'animation de son antécédent et il est de ce point de vue plus informatif que *auquel*. Pour les relatives OI en général, on peut considérer que *à qui* est plus informatif que *auquel* et, si l'Hypothèse 1b est vraie, les attachements à N2 sont plus fréquents avec *à qui*. Les résultats confirment cette hypothèse. Lorsqu'on prend en compte l'information d'animation de *à qui* et que l'on restreint l'analyse des relatives avec *auquel* à celles qui ont comme site d'attachement un animé, on peut contrôler la différence d'informativité de *à qui* et *auquel*. Dans ce cas, le modèle linéaire mixte ne conserve pas la forme du relativiseur en tant que prédicteur. Ces résultats suggèrent que lorsque le degré d'informativité varie entre deux relativiseurs pour une fonction syntaxique donnée (*qui-lequel*, *à qui-auquel* tout antécédent), l'attachement signalé est différent. Cette différence d'accessibilité conduit à attacher plus souvent la relative à l'antécédent peu accessible lorsque le relativiseur est plus informatif. Sinon, un niveau d'informativité équivalent (*à qui-auquel* antécédent animé) ne conduit pas à une différence d'attachement. C'est la différence d'informativité entre *qui* et *lequel* qui influence leur position dans la hiérarchie d'accessibilité. L'Hypothèse 1b est validée.

On peut émettre l'hypothèse que la fréquence d'occurrence de ces types de relativiseurs et celle de phrases contenant des relativiseurs précédés d'une structure « N1 de N2 » peut guider l'attachement. Un relativiseur et une construction peu fréquente seraient caractérisés par un attachement à l'antécédent moins accessible. Un relativiseur comme *lequel* est plus rare que

³¹ Que ce soit pour *lequel* ou *auquel*, leur mention inclut systématiquement leurs déclinaisons, sauf précision explicite.

qui et on mesure bien une préférence pour N2 plus importante. Le relativiseur à *qui* est moins fréquent et attaché plus souvent à N2 que *auquel*. Cependant, si on restreint les relatives avec *auquel* à celles dont le site d'attachement est animé, à *qui* reste deux fois moins fréquent, mais il n'y a pas de différence significative dans les attachements réalisés. La fréquence d'occurrence seule ne permet pas d'expliquer les résultats obtenus.

L'Hypothèse 2 et ses différents volets postulent que la forme des antécédents signale leur degré d'accessibilité et que cela influence l'attachement du relativiseur dont la forme, elle aussi, signale le degré d'accessibilité de l'antécédent. Plusieurs facteurs sont étudiés, la longueur de la proposition relative, la modification et l'animation des antécédents ainsi que la fonction syntaxique de N1.

5. 1. 2. Effet de la longueur de la proposition relative (Hypothèse 2a)

Évaluer l'effet de la longueur de la proposition relative sur l'attachement teste de façon indirecte le niveau d'accessibilité signalée de l'antécédent. Considérant que l'allocutaire a reçu vraisemblablement peu d'informations sur une entité nouvelle dans le discours, ce qui le rend peu accessible, une proposition relative longue fournit plus d'informations qu'une courte et sera attachée à un antécédent faisant référence à une entité relativement nouvelle. On s'attendait en effet à plus d'attachements à N2 lorsque la taille de la relative est importante (Hypothèse 2a).

Les résultats escomptés sont à l'inverse de ceux attendus selon Fodor (J. D. Fodor, 1998) pour qui un constituant long sera attaché préférentiellement à un constituant plus long, c'est-à-dire à « N1 de N2 » (cf. *sausage machine*, Frazier & Fodor, 1978). C'est ce patron de résultats qui est observé dans l'étude de corpus, avec des attachements plus nombreux à N2 quand la relative est courte et à N1 lorsqu'elle est importante. Ceci invalide l'Hypothèse 2a, mais il est notable que l'effet de la longueur de la relative n'est significatif que pour les relatives OI et l'interaction avec l'animation de N2 (toutes les relatives non restrictives et les relatives OI) ou de N1 (les relatives OI). L'interaction entre la longueur de la relative et l'animation de N2 suggère un effet quelque peu antagoniste de ces facteurs. Si N2 est animé, il est très accessible et la fréquence d'attachement à cet antécédent n'est pas affectée par la longueur de la relative. Lorsque N2 est non-animé, il est peu accessible et constitue d'autant moins souvent le site d'attachement de la relative que la taille de cette dernière est grande.

Par contre, l'interaction entre la longueur de la relative et l'animation de N1 traduit un effet cumulatif des facteurs. N1 animé est très accessible et l'attachement est sensible à la quantité

d'informations fournies par la relative : N1 animé est plus susceptible d'être le site d'attachement quand la relative est longue. Si N1 est non-animé, il est peu accessible et la longueur de la relative, comme la quantité d'informations, n'affectent pas l'attachement.

Ces résultats montrent qu'un antécédent animé est très accessible et voit son niveau d'accessibilité maintenu, voire augmenté, par rapport à un antécédent moins accessible avec la quantité d'informations contenues dans la proposition relative. Ce n'était pas prédit par l'Hypothèse 2a. Néanmoins, certains auteurs (cf. Brennan *et al.*, 1987 ; Chambers & Smyth, 1998) postulent que le niveau d'accessibilité de celle-ci est évalué (cf. Brennan *et al.*, 1987 ; Chambers & Smyth, 1998) à chaque référence à une entité. Par conséquent, une proposition relative longue permet de signaler l'accessibilité relative de l'antécédent auquel il est attaché. La longueur de la relative est un marqueur d'accessibilité.

Notons que les résultats obtenus pour les relatives sujet sont en conformité avec ceux de De Baecke *et al.* (1999, cité par Colonna, 2001) où la préférence d'attachement N2 n'est pas influencée par la longueur de la relative, contrairement à ce qu'observent Toussenet et Abeillé (2001, *ibidem*) dans leur étude de corpus en français qui comporte plus d'attachements à N1 quand la relative est longue.

5. 1. 3. Effet de la modification (Hypothèse 2b)

Selon les théories de l'accessibilité, une expression référentielle avec un modifieur signale un niveau d'accessibilité plus faible qu'une expression sans modifieur (Ariel, 1990 ; Gundel *et al.*, 1993). À partir de ce constat, l'Hypothèse 2b propose qu'un antécédent sans modifieur est plus accessible, ce qui en fait le site d'attachement préféré, par comparaison avec un antécédent avec modifieur.

Que ce soit pour le modèle global ou ceux pour les relatives sujet et pour les relatives OI, un antécédent modifié est plus souvent le site d'attachement que s'il n'est pas modifié. Ce premier résultat tend à montrer que la modification d'un antécédent le rend plus saillant, ce qui en fait le site d'attachement préféré.

L'absence d'interaction entre la modification de N1 ou de N2 avec la forme du relativiseur (notamment *qui* et *lequel*) suggère qu'il s'agit de marqueurs d'accessibilité indépendants. Un nom avec un modifieur est préféré comme site d'attachement, mais cette préférence n'est pas affectée par l'accessibilité signalée par la forme du relativiseur.

Les résultats ne confirment pas l'Hypothèse 2b et ils infirment l'hypothèse de la modifiabilité de Thornton *et al.* (1999). Contrairement à ce que cette hypothèse prédit, un

nom déjà modifié reste disponible pour une nouvelle modification apportée par la proposition relative. Que N1 soit modifié par un adjectif et « de N2 » aurait du le rendre non disponible pour que la relative y soit attachée. Or, il y a plus d'attachements à N1 quand il est modifié. Les préférences d'attachement mesurées dans ce corpus sont en partie en accord avec les prédictions de Colonna (2001 ; Pynte & Colonna, 2000) et les résultats dans des questionnaires (Gilboy *et al.*, 1995) ou une expérience en compréhension (Colonna, 2001). On mesure plus d'attachements à N2 quand celui-ci est modifié, car le coût d'intégration de la relative à N2 est moindre que celui de son attachement à N1. En revanche, la modification de N1 ne doit pas affecter l'attachement, parce que la présence d'un modifieur, *a priori* prénominal, n'augmente pas la distance entre N1 et le relativiseur. C'est pourtant le cas ici et les propositions de Colonna ne sont qu'en partie vérifiées.

5. 1. 4. Effet de l'animation (Hypothèse 2c)

Une entité animée est considérée comme plus saillante dans le discours, plus accessible qu'une entité non-animée (Bock & Irwin, 1980 ; Bock & Levelt, 1994 ; Branigan *et al.*, 2008 ; Dahl & Fraurud, 1996 ; Stevenson, 2002). L'Hypothèse 2c prévoit qu'on jugera préférable d'attacher une proposition relative à l'antécédent animé, mais aussi qu'il y aura un effet de la forme du relativiseur capable d'interagir avec l'animation.

Si on considère pour le modèle global et celui pour les relatives sujet l'effet de l'animation de N1 et de N2, il est significatif, avec plus d'attachement à l'antécédent, N1 ou N2, quand il est animé plutôt que non-animé. L'animation rend l'antécédent plus saillant que s'il est non-animé.

Dans son étude où il analyse un corpus écrit de relatives non restrictives et non ambiguës, Loock (2007) met en évidence un effet de l'animation : il y a une majorité de relatives introduites par *who*, qui marque le caractère animé de son antécédent, tandis que *which*, utilisé pour un antécédent non-animé, est moins fréquent. Les propositions relatives non restrictives sont attachées à un antécédent animé, très accessible.

Pour évaluer l'effet de l'animation dans le corpus présenté, il est plus pertinent d'étudier les relatives sujet. Les relativiseurs à *qui* et *auquel* ne sont pas directement comparables dans la mesure où seul le premier marque le caractère animé de son antécédent. La différence constatée au niveau de l'animation entre ces deux relativiseurs est évidemment liée à ce marquage. D'ailleurs, les modèles des relatives OI, avec un antécédent animé pour *auquel* ou sans restriction d'animation, ne conservent pas l'animation de N1 et N2 comme prédicteurs.

L'animation n'influence pas l'attachement des relatives OI. Les deux types de relativiseurs, *à qui* et *auquel*, signalent un niveau d'accessibilité et un attachement équivalents.

Contrairement aux relativiseurs pour la fonction OI, *qui* et *lequel* peuvent être attachés aussi bien à un antécédent animé qu'à un antécédent non-animé. Une différence d'attachement liée à l'animation avec les relatives sujet peut être interprétée comme associée à l'accessibilité signalée de *qui* et *lequel*, ce dont on ne peut s'assurer avec *à qui* et *auquel*.

Sur l'ensemble des propositions relatives non restrictives, le modèle global montre qu'il y a un effet de l'animation de N1, c'est-à-dire plus d'attachements à N2 quand N1 est non-animé, ainsi qu'une interaction entre ce prédicteur et la forme du relativiseur. Avec les relativiseurs *qui* et *à qui*, il y a significativement plus d'attachements à N1 si ce dernier est animé que non-animé. Les relativiseurs *lequel* et *auquel* sont attachés en majorité à N2 quel que soit le degré d'animation de N1. L'interaction indique que l'attachement d'un relativiseur comme *qui* et *à qui* est guidé par la haute accessibilité de N1 liée à son caractère animé. Avec *lequel* et *auquel*, par contre, l'attachement majoritaire est à N2, l'antécédent le moins accessible. Qu'on retrouve cette interaction pour les relatives sujet et non pour les relatives OI suggère que c'est bien la différence d'attachement, signalée par *qui* et *lequel*, qui contribue à l'interaction dans le modèle global.

L'effet de l'animation de N2 pour le modèle global et les relatives sujet révèle que si N2 est animé, il est plus accessible que s'il ne l'est pas. Il est alors le site d'attachement majoritaire. L'animation est ainsi un marqueur pertinent d'accessibilité : il interagit avec l'accessibilité signalée par les relativiseurs, uniquement selon l'animation de N1. *Qui* et *lequel* se situent à des niveaux différents de la hiérarchie d'accessibilité et *lequel*, qui signale un attachement à un antécédent peu accessible, est attaché à N2, peu accessible (N2 est le complément du nom et N1 la tête du GN), même si l'accessibilité relative de N1 change. Ce dernier reste plus accessible quelle que soit son animation. *Qui*, par contre, est d'autant plus attaché à l'antécédent plus accessible, N1, que ce dernier signale une accessibilité élevée (*i. e.* animé). Dans l'ensemble, l'Hypothèse 2c est validée.

Les résultats pour les relatives sujet, indépendamment de la forme du relativiseur, sont conformes avec ceux des études de corpus où l'animation est évaluée (Desmet, Brysbaert *et al.*, 2002 ; Desmet *et al.*, 2006 ; Kister, 2002) et qui ont montré une préférence d'attachement pour l'antécédent animé.

5. 1. 5. Effet de la fonction syntaxique (Hypothèse 2d)

Givón (ch. 20, 1990) propose une hiérarchie de topicalité selon la fonction syntaxique des constituants. Celui qui remplit la fonction sujet est dans le topique principal, alors que le constituant qui remplit la fonction objet direct est dans le focus secondaire. Pour les autres fonctions syntaxiques, l'entité est non topique. Si N1 remplit la fonction sujet, il est plus accessible que N2 qui est génitif, complément du nom. En revanche, si la fonction syntaxique de N1 est plus basse, son niveau de topicalité, mais aussi d'accessibilité, est moindre. L'Hypothèse 2d prédisait que si N1 remplit une fonction syntaxique élevée, la fonction sujet, il est plus accessible. Si les relativiseurs signalent des niveaux différents d'accessibilité, comme c'est le cas entre *qui* et *lequel*, celui qui signale un antécédent peu accessible (*lequel*) est attaché préférentiellement à N2 qui est toujours moins accessible. *Qui* signale un antécédent hautement accessible et est d'autant plus attaché à N1, plus accessible que N2, que la fonction syntaxique de N1 est élevée.

Dans cette analyse de corpus, la catégorie « sans fonction » est ajoutée à celles de Givón (1990). Ces occurrences ne peuvent être intégrées à la catégorie « autres » de Givón ou la catégorie « OI et obliques » utilisée ici puisque, justement, N1 ne remplit pas de fonction syntaxique. Comme dans ce cas « N1 de N2 » est le plus souvent en incise, on peut supposer que le degré d'accessibilité de N1 est faible.

Que ce soit pour le modèle global, les relatives sujet et OI, on mesure un effet de la fonction syntaxique de N1 : quand N1 est sujet, il est le site d'attachement préféré, mais quand sa fonction est plus basse (objet direct, OI et obliques), N2 est préféré. Ces résultats valident l'Hypothèse 2d. Pour le modèle global, on mesure une très forte préférence pour N1 quand il est animé et sans fonction. Ceci est lié à la structure des phrases du corpus où N1 est dans une apposition (67) ou dans une incise. Placer l'information en tête de phrase sert à signaler une entité très proéminente qui est souvent animée. Quand cette entité fait partie d'une structure « N1 de N2 », l'attachement préféré est à N1 qui est très proéminent (par l'apposition et l'animation).

(67) *Témoin de ces changements, Catherine Thieck, directrice de la Galerie de France, à qui nous avons demandé comment elle avait travaillé avec les artistes qu'elle expose.*

On obtient aussi une interaction significative entre la forme du relativiseur et la fonction syntaxique de N1 pour le modèle global et les relatives sujet. Ce sont ces dernières qui contribuent à l'interaction dans le modèle global.

Pour les relatives sujet, il y a une forte préférence pour N1 quand il remplit la fonction sujet avec le relativiseur *qui*. Cette préférence diminue quand la topicalité et l'accessibilité de N1 baisse. *Lequel*, par contre, est attaché majoritairement à N2 quelle que soit la fonction syntaxique de N1. L'accessibilité signalée par la fonction syntaxique de N1 et la forme du relativiseur guident l'attachement réalisé. *Qui* signale un antécédent hautement accessible et est attaché à N1 quand sa fonction signale un tel niveau d'accessibilité. *Lequel* signale un antécédent peu accessible, N2, indépendamment de la fonction syntaxique de N1.

Comme *à qui* et *auquel* signalent un niveau d'accessibilité équivalent, les relatives introduites par ces relativiseurs sont attachées majoritairement à N1 quand il remplit la fonction sujet, mais cette préférence baisse quand la fonction syntaxique de N1 signale une accessibilité plus faible (objet direct, indirect, obliques et sans fonction). Les attachements ne sont pas différents selon la forme du relativiseur (*à qui*, *auquel*). On peut considérer que l'Hypothèse 2d est validée.

5. 2. Apport pour les théories

On l'a vu plus haut, les résultats sur l'effet de l'adjectif infirment ce qui est prédit par le principe de modifiabilité de Thornton *et al.* (1999), issu des modèles basés sur les contraintes. Ce modèle prédit que si un antécédent est déjà modifié, soit par un adjectif pour N1 et N2, soit par « de N2 » pour N1 seulement, il n'est pas susceptible d'être à nouveau modifié par la relative. Or, les résultats de ce corpus et d'autres études (Colonna, 2001 ; Gilboy *et al.*, 1995) infirment ces prédictions.

Les résultats de l'analyse de corpus, en particulier l'effet de la forme du relativiseur, sont intéressants dans l'évaluation de la *tuning hypothesis* (Mitchell *et al.*, 1995). Selon cette théorie, les préférences d'attachement en compréhension sont guidées par la fréquence des structures et des attachements auxquels le parseur est confronté et qui peuvent être observés dans des corpus, en production. La *tuning hypothesis* avance l'hypothèse d'un grain d'analyse élevé : le parseur se base sur les fréquences pour ces structures en général, par exemple sur l'ensemble des relatives dont l'attachement est vers N1 ou N2 dans une structure « N1 de N2 ». La forme du relativiseur n'aurait alors aucun effet, ni aucun autre paramètre syntaxique (p. ex. la fonction de N1) ou sémantique (p. ex. l'animation).

L'observation d'une différence dans les préférences d'attachement en compréhension selon que les propositions relatives sont introduites par *qui* ou *lequel* et une fréquence d'attachement à N2 plus importante pour *lequel* infirme la prédiction d'un grain grossier de l'analyse : la forme du relativiseur a un effet sur l'attachement.

Qui plus est, en ce qui concerne la fréquence globale des attachements en production, on a mesuré une préférence plus marquée pour N2, alors que dans les études en compréhension, on observe une forte préférence pour N1, que ce soit dans l'étude de Questionnaires présentée plus haut ou dans la littérature sur la compréhension de telles ambiguïtés syntaxiques en français (Colonna, 2001 ; Frenck-Mestre & Pynte, 2000a ; Pynte & Colonna, 2000 ; Zagar *et al.*, 1997). Ces résultats sont incompatibles avec les prédictions de la *tuning hypothesis* (grain d'analyse grossier), ce que les résultats d'études en néerlandais suggèrent aussi (Desmet, Brysbaert *et al.*, 2002 ; Desmet *et al.*, 2006).

Il est cependant possible de trouver une préférence similaire en production et en compréhension si on prend un degré d'analyse plus fin, qui tient compte par exemple l'animation des antécédents. On a vu que le trait d'animation affecte de façon importante les attachements observés dans le corpus. Pour les données en néerlandais (Desmet, Brysbaert *et al.*, 2002 ; Desmet *et al.*, 2006) ou en français (ce corpus), quand on restreint l'observation aux cas où N1 et N2 sont animés, on trouve une préférence pour N1 (66 % dans ce corpus). Une préférence similaire est mesurée dans des tâches en compréhension. Ces résultats confortent l'hypothèse de l'utilisation d'un grain d'analyse mixte (Desmet *et al.*, 2006), à la fois grossier, c'est-à-dire les structures syntaxiques considérées, et fin, avec les informations lexicales comme l'animation (voir p. ex. les modèles de Jurafsky, 1996 ; Tabor *et al.*, 1997). Ce modèle paraît plus satisfaisant que l'approche à gros grain prônée par la *tuning hypothesis*.

Les résultats de ce corpus apportent des données intéressantes pour les théories de l'accessibilité (Ariel, 1990, 2001 ; Gundel *et al.*, 1993) : la forme du relativiseur permet de signaler le degré d'accessibilité de l'antécédent. Un relativiseur situé plus bas sur la hiérarchie d'accessibilité (*lequel*, par comparaison avec *qui*) est attaché plus fréquemment à l'antécédent moins accessible dans une structure « N1 de N2 » que son pendant situé plus haut dans la hiérarchie (*qui*). Avec une structure « N1 de N2 », les attachements à N2 sont plus fréquents avec le relativiseur *lequel* qu'avec *qui* ; N2 est en effet moins accessible que N1 qui est la tête du GN. La comparaison entre les attachements avec *à qui* et *auquel* montre une différence significative, le relativiseur *à qui* signalerait un niveau d'accessibilité plus faible que le relativiseur *auquel*, puisqu'il y a plus d'attachements à N2 pour *à qui*. Mais cela n'est vrai que

si l'animation de l'antécédent pour *auquel* n'est pas contrôlée et si le marquage de l'animation différencie les relativiseurs. Cependant, le modèle où l'animation du site d'attachement est contrôlé (*à qui* et *auquel* avec antécédent animé) ne révèle pas d'interaction significative avec d'autres facteurs : ces facteurs influencent de façon comparable les attachements pour ces deux formes de relativiseurs. L'ensemble de ces résultats, comme ceux de l'étude de corpus d'Ariel (1999), confirme que les relativiseurs, comme les autres types d'anaphores, fonctionnent comme des marqueurs d'accessibilité.

L'attachement et le niveau d'accessibilité signalés par les relativiseurs est sensible au niveau d'accessibilité relatif des antécédents disponibles, N1 et N2. La fonction syntaxique de N1 influence l'accessibilité de ce dernier : quand il remplit la fonction sujet, il est alors hautement accessible. C'est la raison pour laquelle l'emploi de *qui* est favorisé, car ce relativiseur signale un degré d'accessibilité plus important que *lequel*. Une entité animée est plus proéminente qu'une entité non-animée et par conséquent elle est plus accessible dans le discours, devenant le site d'attachement le plus fréquent. Cependant, pour les relatives avec *lequel*, signalant un antécédent moins accessible, l'attachement de ces relatives n'est pas sensible à l'animation de N1 qui reste malgré tout l'antécédent plus accessible. On observe toujours plus d'attachements à N2 avec *lequel*, indépendamment de l'animation de N1, confirmant que l'attachement de *lequel* est bien l'antécédent le moins accessible.

L'interaction significative, pour les relatives sujet, entre la forme du relativiseur et l'animation de N1 montre que pour les relatives avec *lequel*, le degré d'animation de N1 n'influence guère l'attachement, plus fréquent à N2, alors que pour les relatives avec *qui*, il y a un très fort taux d'attachement à N1 quand celui-ci est animé. Le relativiseur *qui*, qui signale un antécédent plus accessible, est d'autant plus souvent attaché à N1 que ce dernier est animé et donc très accessible.

On pourrait se demander pourquoi l'animation de N1 influence l'attachement avec *qui* et non avec *lequel*. Que pour les autres fonctions relativisées, *auquel*, *duquel*, etc. le relativiseur marque l'animation de l'antécédent peut favoriser un attachement comparable, à l'antécédent animé, pour *qui*. Des critères sémantiques peuvent aussi intervenir : une entité animée est le plus souvent agent (cf. expériences sur l'effet des informations sémantiques dans l'interprétation d'ambiguïtés, Clifton *et al.*, 2003 ; Trueswell *et al.*, 1994) et elle est par conséquent plus saillante, plus accessible. C'est pourquoi *qui* serait préféré. Néanmoins, la proportion d'occurrences où N1 est animé avec une relative introduite par *qui* est proche (22 %) de celle qui est introduite par une relative contenant *lequel* (28 %). Cette interprétation

ne peut être retenue : que N1 soit animé ne permet pas de prédire l'emploi de *qui* au détriment de *lequel*, puisque ces relativiseurs sont utilisés dans des proportions similaires.

Il est intéressant de remarquer que, pour les relatives OI, on obtient un effet de la forme du relativiseur, mais que ce facteur n'interagit pas avec les autres facteurs. L'attachement des relatives est guidé par la forme du relativiseur. *À qui* est attaché plus fréquemment à N2 (cf. effet significatif dans le modèle des relatives OI), car la configuration où N1 est non-animé et N2 animé est plus fréquente avec ce relativiseur (67 % des occurrences avec *à qui*) qu'avec *auquel* (tout antécédent, 23 %). Par contre, si on contrôle l'animation de l'antécédent avec *auquel*, une proportion comparable de configurations N1 non-animé et N2 animé est observée (62 %). L'absence d'interaction suggère que l'attachement avec *à qui* et *auquel* ne diffère pas selon le niveau d'accessibilité signalé par la forme des antécédents. S'il y a davantage d'attachements à N2 avec *à qui*, le sens de l'effet des différents facteurs est identique pour *à qui* et *auquel*.

C'est pour la paire *qui-lequel* qu'on observe des interactions entre la forme des relativiseurs et les facteurs susceptibles d'agir sur le niveau d'accessibilité signalée des antécédents, suggérant des attachements différents quand les relativiseurs signalent eux aussi un niveau d'accessibilité plus ou moins important.

En ce qui concerne l'effet de la modification des antécédents sur l'attachement et leur niveau d'accessibilité, les résultats sont contraires à ce qu'on aurait attendu selon la hiérarchie d'accessibilité proposée par Ariel (1990, 2001) : un nom modifié signifierait un antécédent moins accessible. Or, pour N1 ou pour N2, qu'un antécédent soit modifié y entraîne plus d'attachements, même quand le relativiseur signale un antécédent moins accessible (*lequel*). La présence d'un modifieur n'est pas pertinente pour signaler un antécédent moins accessible.

Plus une proposition relative est longue, plus elle fournit d'informations nouvelles sur son antécédent. On s'attendait alors à ce qu'une proposition relative plus longue soit attachée plus fréquemment à un antécédent moins accessible, N2 dans « N1 de N2 », d'autant plus que le relativiseur signale un degré d'accessibilité moindre. Cette prédiction s'oppose à celle de Fodor (1998) selon laquelle une proposition relative est attachée à un constituant de taille comparable, N1 quand la proposition relative est longue. On observe dans le corpus un effet de la longueur de la relative pour les relatives OI qui contribuent ainsi à l'interaction avec la longueur de la relative révélée pour le modèle global.

Pour les relatives sujet, aucune des hypothèses n'est vérifiée puisqu'il n'y a pas d'effet significatif de la longueur de relative. Pour les relatives OI, il y a plus d'attachements à N1

quand la relative est longue et plus d'attachements à N2 quand la relative est courte, confirmant l'hypothèse de Fodor (1998).

En conclusion, les données de ce corpus confirment la première hypothèse : deux relativiseurs comme *qui* et *lequel* signalent un niveau d'accessibilité différent qui sert d'indice pour désigner le site d'attachement, *lequel* indiquant un antécédent moins accessible. Les attachements avec les relativiseurs à *qui* et *auquel* sont globalement comparables, pour peu qu'on contrôle l'animation de l'antécédent et par conséquent l'informativité de à *qui* et *auquel*. Dans ce cas, l'absence de différence d'attachement conforte l'idée selon laquelle ces deux relativiseurs signalent un degré d'accessibilité équivalent.

L'accessibilité relative de N1 et N2 peut être modifiée par leur animation, mais aussi par la fonction syntaxique de N1, et on constate qu'il y a plus d'attachements à l'antécédent moins accessible avec *lequel* qu'avec *qui*. Les autres facteurs étudiés ne semblent pas affecter le degré d'accessibilité signalé par les antécédents, ou tout du moins n'entraînent pas d'attachement différent selon le degré d'accessibilité signalé par les relativiseurs.

5. 3. Conclusions

La plupart des hypothèses de travail qui ont été proposées sont validées. *Qui* et *lequel* signalent un niveau d'accessibilité différent car leur degré d'atténuation et d'informativité les positionne à des niveaux différents de la hiérarchie d'accessibilité. La forme du relativiseur permet de signaler par exemple un attachement à N2 avec *lequel*, ce qui entraîne plus d'attachement à cet antécédent qu'à *qui*. Quand l'animation de l'antécédent de *auquel* est contrôlée, à *qui* et *auquel* signalent un niveau d'accessibilité équivalent et un attachement comparable.

L'animation des antécédents, la modification et la fonction syntaxique de N1 interagissent avec la forme des relativiseurs sujet, *qui* et *lequel*, confirmant que la forme des antécédents et celle des relativiseurs signalent le degré d'accessibilité des référents du discours, guidant l'attachement des propositions relatives.

Cette étude de corpus est originale par rapport à la majorité des études de corpus publiées en psychologie cognitive ou en psycholinguistique, notamment sur la production de propositions relatives ambiguës (voir p. ex. Desmet, Brysbaert *et al.*, 2002 ; Desmet *et al.*, 2006 ; Kister, 2002) ou d'autres ambiguïtés syntaxiques (Gibson & Schütze, 1999), qui ne présentent que rarement des évaluations statistiques. On peut noter l'exception des travaux de Demberg et Keller (à paraître) et ceux d'Arnold, Wasow, Losongco, et Ginstrom (2000), par

exemple, qui ont utilisé les modèles linéaires mixtes pour éprouver la validité des facteurs observés.

Les modèles linéaires mixtes complètent de façon intéressante une étude de corpus où les fréquences sont reportées, ajoutant des données statistiques à propos des variables ou des prédicteurs qui peuvent expliquer les données du corpus.

L'intérêt particulier de cette étude de corpus vient de ce qu'elle procure des informations sur le versant « production » des structures syntaxiques étudiées dans cette thèse, confirmant un effet de la forme du relativiseur sur l'attachement réalisé. Elle souligne aussi la différence d'attachement entre des données en compréhension, l'étude de Questionnaires (cf. Chapitre 1), et des données en production. Cette différence est liée au degré d'analyse et aux critères pris en compte, comme l'animation des antécédents (cf. *tuning hypothesis*, Desmet, Brysbaert *et al.*, 2002).

Je propose d'évaluer si la forme du relativiseur influence les préférences d'interprétation lors de la résolution d'ambiguïtés d'attachement à l'aide du paradigme d'auto-présentation segmentée. Cette technique procure des informations *on-line*, pendant la compréhension et la lecture de phrases.

Chapitre 3 Expérience avec le paradigme d'auto-présentation segmentée

1. Introduction

Le but de l'étude de Questionnaires présentés dans le Chapitre 1 était d'évaluer l'hypothèse selon laquelle la forme du relativiseur signale l'attachement à réaliser par l'accessibilité de l'antécédent qu'il signale. Pour une fonction syntaxique donnée, les relativiseurs peuvent varier selon l'atténuation (Hypothèse 1a) ou l'informativité (Hypothèse 1b, voir Problématique et hypothèses). Ces hypothèses reposent sur les théories de l'accessibilité, telles que celle de Gundel *et al.* (1993) ou celle d'Ariel (1990, 2001). Cette dernière postule que la forme d'une anaphore permet de signaler le degré d'accessibilité de l'entité référée. Appliqué au relativiseur, qui est un marqueur anaphorique (Hemforth *et al.*, 2000), cela implique que la forme du relativiseur signale le degré d'accessibilité de l'entité référée, son antécédent et par conséquent qu'elle guide l'attachement.

Les résultats obtenus confirment l'existence d'une préférence pour attacher une proposition relative ambiguë à N1, conformément à ce qui a été mesuré dans d'autres expériences en français (Baccino *et al.*, 2000 ; Zagar *et al.*, 1997) ou dans d'autres langues (voir Introduction théorique, Chapitre 1, section 2). Les Questionnaires 1 et 4 ont mis en évidence l'existence d'une préférence d'attachement différente selon que la proposition relative est introduite par *lequel* ou par *qui*. *Lequel* est moins atténué et plus informatif que *qui* et se situe plus bas dans la hiérarchie d'accessibilité. Il signale ainsi un attachement à un antécédent moins accessible, N2 pour une structure « N1 de N2 ». Pour des relatives avec *lequel*, on mesure plus d'attachements à N2. La comparaison des préférences d'attachement entre les relativiseurs *à qui* et *auquel* (atténuation et informativité comparables, Questionnaires 2 et 3) et entre *à qui* et *à laquelle* (informativité comparable, atténuation moindre pour *à laquelle*, Questionnaire 3), par contre, n'a révélé aucune différence.

Ces résultats suggèrent que de l'atténuation (Hypothèse 1a) et de l'informativité (Hypothèse 1b), c'est cette dernière qui influence le plus les préférences d'attachement,

puisque la différence d'interprétation est observée seulement quand le niveau d'informativité diffère (*qui, lequel*). La forme du relativiseur signale le type d'attachement à réaliser.

Il était intéressant d'étudier l'effet de la forme du relativiseur en fonction des variations dans l'accessibilité des antécédents possibles, N1 et N2. L'étude de corpus (voir Chapitre 1) a ainsi évalué les préférences d'attachement en fonction de la saillance (Hypothèse 2a), de l'animation (Hypothèse 2b) et de la fonction syntaxique (Hypothèse 2d), facteurs qui influencent le degré d'accessibilité signalé par les anaphores (Ariel, 1990 ; Gundel *et al.*, 1993). Les résultats montrent une préférence pour attacher la proposition relative, non pas à N1 comme cela est mesuré dans des tâches de compréhension, mais à N2. Ces résultats infirment les prédictions de la *tuning hypothesis* (Mitchell, 1994 ; Mitchell *et al.*, 1995), tout du moins dans sa version à grain d'analyse élevé : elle propose que les préférences d'attachement en compréhension sont guidées par les préférences d'attachement auxquelles le parseur est confronté et qui peuvent être mesurées dans des productions de telles structures. Les différences obtenues dans les questionnaires et le corpus sont conformes avec ce qui est observé en néerlandais (Desmet, Brysbaert *et al.*, 2002 ; Desmet *et al.*, 2006) et avec une version de la *tuning hypothesis* où une analyse plus fine intègre par exemple le caractère animé des antécédents. Dans ce cas, on observe une préférence pour N1 dans le corpus, comme dans les tâches en compréhension. Les résultats obtenus valident nos hypothèses : la différence d'accessibilité signalée par *qui* et *lequel* conduit à plus d'attachements à l'antécédent moins accessible.

Ces résultats sont intéressants, mais si la technique des questionnaires est utile pour avoir une première estimation des préférences d'attachement, elle a néanmoins l'inconvénient de ne révéler que l'interprétation finale. En effet, au moment de répondre, les participants ont fini de traiter la phrase et c'est leur interprétation finale qu'ils signifient.

L'expérience qui est présentée ici a l'avantage de tester de façon plus approfondie ce qui a été évalué avec ces Questionnaires. Dans cette expérience, le paradigme d'auto-présentation est utilisé et permet d'enregistrer mot par mot les temps de lecture (ou temps de réaction, abrégés TR). Ces mesures *on-line*, pendant le processus de compréhension, procurent des informations sur les éventuelles difficultés de traitement que le lecteur peut rencontrer à certains endroits dans une phrase. Je reviendrai sur les caractéristiques de ce paradigme dans la section dédiée à la procédure expérimentale (2. 2. 1. 4).

Contrairement aux phrases des Questionnaires pour lesquelles l'attachement devait rester ambigu afin d'évaluer les préférences d'attachement, le matériel expérimental contient une

proposition relative dont l'attachement n'est que temporairement ambigu. Les temps de lecture, de traitement, mesurés au niveau de la région où l'ambiguïté est levée, révèlent si l'information procurée dans cette région est prise en compte. Des temps de lecture plus importants d'une condition à l'autre, dans cette région et les suivantes, peuvent être interprétés comme mettant en évidence un phénomène de réinterprétation, impliquant que dans les régions antérieures le parseur avait choisi une interprétation, infirmée par la désambiguïsation.

Dans l'exemple suivant (68a-d), l'attachement de la relative est ambigu jusqu'à l'objet de la relative (*son service*), qui signale un attachement à N1, (68a) et (68c), ou N2, (68b) et (68d). C'est la région de désambiguïsation.

(68) a. *La serveuse de la cliente, qui termine son service dans moins d'une heure, renverse l'assiette de légumes.*

b. *La cliente de la serveuse, qui termine son service dans moins d'une heure, renverse l'assiette de légumes.*

c. *La serveuse de la cliente, laquelle termine son service dans moins d'une heure, renverse l'assiette de légumes.*

d. *La cliente de la serveuse, laquelle termine son service dans moins d'une heure, renverse l'assiette de légumes.*

Cette expérience complète l'étude de l'effet de la forme du relativiseur en fournissant des informations sur la compréhension de ces phrases, pendant l'intégration des mots qui la composent, testant l'Hypothèse 1a. D'après cette hypothèse, on s'attend à ce que les lecteurs, confrontés à une ambiguïté d'attachement, décident d'un attachement initial. Cet attachement est guidé par la forme du relativiseur : quand la proposition relative est introduite par *lequel*, le niveau d'accessibilité par ce relativiseur indique un attachement à N2. Avec une proposition relative introduite par *qui*, c'est un attachement initial à N1 qui devrait être réalisé. Cela implique que l'effet de la forme du relativiseur est immédiat, et se produit dès l'intégration du relativiseur. On ne peut exclure un effet de la fréquence du relativiseur : *lequel* est beaucoup moins fréquent que *qui* et cela peut entraîner des temps de lecture plus importants lors de son traitement. Après le relativiseur, au niveau du verbe de la proposition relative, aucune différence entre les conditions expérimentales ne devrait être mesurée. Les participants ont choisi une interprétation et aucune information du verbe ne vient modifier cette interprétation.

L'objet de la proposition relative lève l'ambiguïté d'attachement en signalant, par son contenu sémantique, que l'attachement est à N1 ou N2. Considérons les phrases introduites par *lequel* : si l'interprétation initiale est un attachement à N2, que l'ambiguïté soit levée à N1 infirme cette interprétation, qui doit être révisée. Cette révision est coûteuse et entraîne des temps de lecture plus importants, au niveau de la région de désambiguïsation et des régions suivantes, si la réinterprétation est difficile. On prévoit ainsi une différence entre les phrases (68a) et (68b) à partir de la région de désambiguïsation, avec des temps de lecture plus importants pour (68b). En fait, on s'attend à une interaction entre les deux variables manipulées, le type de Relativiseur (*qui, lequel*) et l'Attachement (vers N1 ou vers N2).

Pour la fin de la phrase, à partir du verbe de la principale, il ne devrait pas y avoir de différence dans les temps de lecture, car les phrases sont identiques quelle que soit la condition expérimentale. Les difficultés liées à la levée de l'ambiguïté sont *a priori* résolues et cela n'affecte plus les temps de lecture.

L'Hypothèse 1 postule un effet de la forme du relativiseur et les résultats attendus sont compatibles avec ce qu'on pourrait proposer en se basant sur l'*informational load hypothesis* développée par Almor (1999). Les relativiseurs sont des anaphores qui apportent peu d'informations nouvelles ; leur charge informationnelle est par conséquent faible et leur traitement peu coûteux. On pourrait cependant émettre l'hypothèse d'une charge différente selon la forme du relativiseur. *Lequel* est marqué en genre et en nombre, ce qui le rend plus informatif que *qui*. Ce surcroît d'information aurait une justification fonctionnelle : *lequel* signalerait un antécédent moins saillant. Dans le contexte spécifique des phrases expérimentales, les deux antécédents sont identiques en genre et en nombre et on pourrait considérer que l'information procurée par *lequel* devient redondante et entraîne un coût informationnel plus important. Une telle hypothèse serait envisageable à condition que l'*informational load hypothesis* situe les différentes expressions anaphoriques et leur charge informationnelle en fonction du contexte, ce qui n'est pas le cas. Il semble plus conforme de considérer que l'*informational load hypothesis* prédit que *lequel* sera attaché à l'antécédent le moins saillant et *qui* au plus saillant.

La plupart des théories sur le traitement d'ambiguïtés syntaxiques ont élaboré des hypothèses sur l'interprétation et le traitement de phrases telles que celles qui vont être manipulées. Cependant, rares sont celles qui considèrent que la forme du relativiseur peut avoir un effet ; c'est le cas de la *construal hypothesis* (Frazier & Clifton, 1996, 1997). D'autres théories comme la *recency theory* pourraient néanmoins intégrer la forme du

relativiseur comme un facteur capable d'affecter les préférences d'interprétation. Les modèles basés sur les contraintes prédiraient aussi une différence de traitement, fondée sur la fréquence des relativiseurs. Selon la théorie du *garden path* et la *tuning hypothesis*, par contre, il ne devrait pas y avoir d'effet. Je vais maintenant revenir sur ces différentes théories.

D'après la *construal hypothesis*, une proposition relative est une structure non primaire et le principe de *construal* s'applique. Le site d'attachement de la relative est le domaine thématique en cours, c'est-à-dire le GN complexe « N1 de N2 ». Comme il y a deux noms, l'attachement, à l'un des deux, est guidé par des informations sémantico-pragmatiques, p. ex. les relations entre N1 et N2. Pour les phrases de cette expérience, le lien entre N1 et N2 est de type relationnel (liens de parenté, *le père de l'enfant*) ou fonctionnel (l'assistant de l'inspecteur, d'après l'appellation de Gilboy *et al.*, 1995). D'après Frazier et Clifton (1996), ce type de relation entre N1 et N2 est caractérisé par une préférence pour N1.

Frazier et Clifton (1997) proposent, dans le cadre de la *construal hypothesis*, que la forme du relativiseur peut aussi influencer les préférences d'attachement. *Lequel*, qui ressemble plus à un pronom par son marquage en genre et nombre, est alors sensible aux mêmes facteurs qui guident l'interprétation des pronoms, comme la proéminence des entités du discours. Dans la structure « N1 de N2 », N1 est l'entité proéminente et par conséquent on s'attend à une préférence plus marquée pour attacher la proposition relative à N1, quel que soit le pronom relatif.

La *recency theory* (Gibson, Pearlmutter *et al.*, 1996 ; Gibson *et al.*, 1999) repose sur deux principes, la proximité du prédicat et la préférence de récence. Selon le premier, un attachement à N1 est attendu alors que le second prédit celui à N2. Le poids de ces deux principes peut varier selon les langues, ce qui explique les variations des préférences d'attachement. La rigidité de l'ordre des constituants de la phrase est avancée pour expliquer le poids de ces principes. Gibson, Pearlmutter *et al.* (1996) ont proposé que le poids de la proximité du prédicat est déterminé par la distance moyenne entre un prédicat et ses arguments. Pour une langue où l'ordre des mots est plutôt libre, comme en espagnol ou en français, la distance entre un prédicat (p. ex. un verbe) et ses arguments (p. ex. le sujet, l'objet) a une distance moyenne plus importante, qui est liée à l'ordre des mots moins rigide. En français, c'est la proximité du prédicat qui a un poids plus important et c'est pourquoi une préférence de N1 est attendue dans les phrases manipulées dans cette expérience. Bien que la *recency theory* ne postule aucun effet de la forme du relativiseur sur le poids de ces deux

principes et les préférences d'attachement, elle n'exclut pas cependant qu'un tel facteur puisse intervenir dans l'interprétation de propositions relatives dont l'attachement est ambigu.

Bien que le modèle *unrestricted race* soit sous-spécifié quant à l'effet du relativiseur, il pourrait tout à fait l'intégrer. En effet, ce modèle considère que, pendant l'interprétation des mots de la phrase, les informations syntaxiques, sémantiques ou discursives sont prises en compte : la forme du relativiseur pourrait alors avoir un effet. Signalant un antécédent moins accessible, *lequel* signale un attachement à N2. Comme le traitement est incrémental, il y a un effet immédiat de la forme du relativiseur. Au niveau de la région de désambiguïsation, et pour les régions ultérieures si la difficulté de ré-interprétation se prolonge, on s'attendrait alors à une interaction. Avec *qui*, il y a un attachement initial pour N1, ce qui entraîne une ré-analyse et des temps de lecture plus importants quand l'attachement est en fait à N2. Pour *lequel*, c'est l'inverse qui est attendu, avec des temps de lecture plus faibles quand l'attachement final est à N2.

Les modèles interactifs, comme le modèle de satisfaction de contraintes ou le modèle de compétition et d'intégration, sont basés sur le traitement des ambiguïtés lexicales. Lors de l'intégration d'un mot dans une phrase, les différentes interprétations lexico-sémantiques de ce mot sont activées. Dans le cas des relativiseurs, on pourrait alors s'attendre à une différence de traitement liée à la différence de fréquence de *qui* et *lequel*. Par sa fréquence moindre, *lequel* signale un attachement différent de *qui* vers N2 alors que *qui* est attaché à N1. Il y aurait ainsi un effet du relativiseur. Selon ces modèles interactifs, face à une ambiguïté d'attachement, toutes les interprétations restent activées jusqu'à ce qu'une seule soit possible (modèle de satisfaction de contraintes) ou atteigne un seuil d'activation critique (modèle de compétition et d'intégration). Si la fréquence du relativiseur est suffisante pour restreindre l'interprétation, on ne devrait pas observer de différence de traitement tant que la phrase est ambiguë. Quand l'ambiguïté est levée, par contre, l'information sémantique confirme ou non l'interprétation initialement retenue. C'est donc au niveau de la région de désambiguïsation qu'on attend une différence de traitement. En revanche, on ne devrait pas observer d'effet prolongé de la désambiguïsation : dès la levée de l'ambiguïté, l'interprétation est éventuellement modifiée avant de passer au mot suivant.

Quant aux modèles interactifs, si on reprend la proposition de Thornton *et al.* (1999) sur la modifiabilité, on devrait s'attendre à plus d'attachements à N2, car N1 est déjà modifié par « de N2 » et serait moins susceptible de l'être à nouveau par la proposition relative.

Cependant, aussi bien les résultats de Colonna (2001) que ceux de l'étude de corpus (effet de la modification, Chapitre 2) ont infirmé cette prédiction.

Enfin, deux théories excluent que des informations portées par le relativiseur interviennent dans le traitement de structures ambiguës et qu'elles guident l'attachement. D'après la *tuning hypothesis*, les préférences d'interprétation de structures ambiguës, telles que celle étudiée ici, sont guidées par les attachements auxquels le locuteur est confronté dans son environnement. C'est l'attachement le plus fréquent dans la langue, en production, qui est retenu lors du traitement d'une structure similaire. Mitchell *et al.* (1995) proposent que c'est une estimation grossière de la structure ambiguë en production qui est prise en compte, en excluant les informations lexicales des constituants de cette structure. Comme l'animation des antécédents (voir Desmet *et al.*, 2006), la forme du relativiseur ne devrait ainsi pas avoir d'effet sur les préférences d'attachement. Si on reprend les mesures de Kister (2002) sur les préférences d'attachement dans un corpus écrit pour des propositions relatives introduites précédées d'une structure « N1 de N2 », c'est une préférence de l'ordre de 57 % en faveur de N1 qui est relevée. L'analyse de corpus présentée dans le Chapitre 1, qui restreint le choix des « N1 de N2 » à de vrais génitifs, révèle cependant des attachements plus fréquents à N2, ce qui tranche avec les mesures réalisées en compréhension en français (voir l'étude de Questionnaires du Chapitre 1 et l'expérience de Zagar *et al.*, 1997).

Ainsi, selon la version à gros grain de Mitchell *et al.* (1995), on escompte une préférence pour N1, mais pas un effet de la forme du relativiseur. Aucune interaction n'est attendue à partir de la région de désambiguïsation.

La théorie du *garden path*, elle aussi, exclut un effet précoce de la forme du relativiseur. Confronté à l'ambiguïté d'attachement du relativiseur à N1 ou N2, le principe de *late closure* est appliqué : l'attachement initial est à N2. Comme l'interprétation repose sur des critères syntaxiques et que l'information procurée par la forme du relativiseur n'est pas un critère syntaxique, il n'y a pas d'effet de la forme du relativiseur sur l'interprétation initiale. La levée de l'ambiguïté au niveau de la région de désambiguïsation peut entraîner une ré-analyse de l'interprétation initiale (i. e. attachement à N2) si l'attachement signalé est à N1. Des informations sémantiques ou d'un autre type peuvent alors être prises en compte et un effet tardif de la forme du relativiseur peut être envisagé. *Qui* signalant un attachement à un antécédent très accessible, sa présence faciliterait la ré-analyse. En résumé, on s'attend, selon la théorie du *garden path*, à un attachement initial à N2 dès l'apparition du relativiseur, dont la forme n'a pas d'effet. Aucune différence au niveau du relativiseur ou du verbe de la relative

n'est attendue. Avec la levée de l'ambiguïté, des temps de lecture plus importants devraient être mesurés quand le contenu de l'objet de la proposition relative signale un attachement à N1, pour *qui* et pour *lequel*. Si la forme du relativiseur peut avoir un effet, la difficulté de ré-analyse devrait être plus faible avec *qui*. Rappelons à ce propos qu'aussi bien l'expérience en français de Zagar *et al.* (1997) que celles en espagnol, en allemand ou en néerlandais infirment l'hypothèse d'un attachement initial à N2, puisque c'est une préférence pour N1 qui est mesurée.

De Vincenzi *et al.* (1993, 1995) ont suggéré une sorte de compromis entre les prédictions de la théorie du *garden path*, une préférence pour N2, et la préférence pour N1 mesurée expérimentalement. Ils proposent que l'interprétation finale est guidée par le principe de *late closure* et que la préférence pour N1 n'émerge que tardivement. Des données en français (Baccino *et al.*, 2000) et en italien (De Vincenzi & Job, 1993, 1995) confirmeraient cette hypothèse.

Si l'expérience d'auto-présentation ne permet pas de distinguer un modèle par rapport aux autres, elle rend toujours possible l'évaluation des modèles selon qu'ils peuvent ou non considérer un effet de la forme du relativiseur, ce qui est prédit par l'Hypothèse 1.

2. Méthodologie expérimentale

2. 1. Participants

Les 48 participants, dont 38 femmes et d'âge moyen 22,4 ans, sont tous des étudiants en deuxième année de Licence de sciences du langage à l'Université Lumière Lyon 2. Tous sont de langue maternelle française, ont une vue normale ou corrigée et n'ont pas d'antécédents connus de dyslexie.

Cette expérience est réalisée au cours d'une session d'une heure pendant laquelle les participants font deux autres expériences ne portant pas sur le langage. Leur participation est dédommagée de 7,50 €.

2. 2. Matériel et procédure expérimentale

2. 2. 1. Matériel

2. 2. 1. 1. Élaboration des items expérimentaux

Le matériel expérimental est au départ constitué de 50 items en quatre versions (68a-d) où sont manipulées deux variables, le type de Relativiseur (*qui* ou *lequel*) et la Désambiguïsation de la proposition relative soit vers N1, (68a) et (68c), soit vers N2, (68b) et (68d). Les 40 meilleurs items sont conservés à l'issue de pré-tests (voir plus bas).

(68) a. *La serveuse de la cliente, qui termine son service dans moins d'une heure, renverse l'assiette de légumes.*

b. *La cliente de la serveuse, qui termine son service dans moins d'une heure, renverse l'assiette de légumes.*

c. *La serveuse de la cliente, laquelle termine son service dans moins d'une heure, renverse l'assiette de légumes.*

d. *La cliente de la serveuse, laquelle termine son service dans moins d'une heure, renverse l'assiette de légumes.*

Ces phrases sont construites de la façon suivante : elles débutent avec le sujet de la proposition principale, un GN complexe de la forme « N1 de N2 », comme par exemple *la cliente de la serveuse* dans (68a). La proposition relative non restrictive, toujours encadrée par une virgule est une relative sujet, constituée du relativiseur, d'un verbe, de l'objet et d'un groupe prépositionnel.

Afin de conserver l'ambiguïté d'attachement jusqu'à la région de désambiguïsation, il est nécessaire que les deux noms, N1 et N2, soient identiques en genre et en nombre, sinon l'attachement serait désambiguïté dès le traitement de *lequel* ou *laquelle* (Brybaert & Mitchell, 1996). Le marquage en genre et en nombre avec ces derniers contraint par ailleurs la forme de désambiguïsation. En effet, c'est une désambiguïsation par le genre de l'adjectif qu'utilisent Zagar *et al.* (1997), avec N1 et N2 de genres différents, ce qui n'est pas possible ici. La désambiguïsation sémantique est alors la seule possibilité. Afin de s'assurer de l'effet de cette désambiguïsation un pré-test est réalisé (voir dans ce Chapitre la section 2. 2. 1. 3).

Les deux noms, N1 et N2 sont contrôlés de façon à ce qu'ils aient des fréquences d'occurrence comparables, mais aussi une longueur et un nombre de syllabes proches. Comme les positions des deux noms sont inversées afin de spécifier l'attachement, les GN constitués doivent être réversibles : *la serveuse de la cliente* doit être aussi acceptable que *la cliente de la serveuse*. Le genre de ces substantifs est marqué et non ambigu, sauf pour quelques uns dont le genre par défaut est le même que l'autre nom.

L'objet de la relative (*son service*) est systématiquement composé d'un nom précédé d'un article, dont le contenu sémantique lève l'ambiguïté d'attachement (cf. Carreiras & Clifton, 1993, 1999). Il s'agit de la région de désambiguïsation. L'objet de la relative est identique pour toutes les conditions expérimentales, ce qui permet de s'assurer que si le temps de lecture diffère au niveau de cette région, il n'est pas dû à une différence de contenu, de longueur ou de fréquence. Ce choix méthodologique conditionne par conséquent l'élaboration des GN « N1 de N2 », où la position des noms doit être inversée selon qu'un attachement à N1 ou à N2 est signalé. Dans (68), c'est *la serveuse* qui est l'agent le plus probable de *termine son service* (cf. 2. 2. 1. 3, pré-test de plausibilité) ; sa position en N1 conduit à signaler un attachement à N1 lorsque l'objet de la relative est atteint. Notons aussi que la région ambiguë est courte, ce qui devrait faciliter la ré-analyse de l'attachement (Christianson, Hollingworth, Halliwell, & Ferreira, 2001).

La relative se poursuit avec un groupe prépositionnel de cinq mots. Cette région dite *spillover* permet de mesurer éventuellement la prolongation sur les constituants suivants d'un effet de l'attachement situé au niveau de la région de désambiguïsation. On appelle cela un effet de *spillover* ; cet effet s'observe assez fréquemment dans des expériences en *eye tracking* (Hoeks, Hendriks, Vonk, Brown, & Hagoort, 2006 ; van Gompel *et al.*, 2001).

Puis viennent le verbe de la principale, son objet (de la forme article et nom) et enfin un groupe prépositionnel ou un modifieur de l'objet de la principale. Le contenu de la principale (verbe, objet et fin) est conçu de façon à ce qu'il soit plausible pour les deux noms (*la serveuse* et *la cliente*).

Pour chacun de ces items, on formule des questions de compréhension, où la question porte soit sur le contenu de la proposition relative (*c'est la serveuse qui termine son service ?*) soit sur le contenu de la proposition principale (*c'est la serveuse qui renverse l'assiette ?*).

Un ensemble de 100 remplisseurs est construit et constitué de phrases de longueur et de complexité variables, mais sans aucune proposition relative. Afin de diminuer les effets

d'attente, certains remplisseurs débutent par une structure « N1 de N2 », avec des noms animés ou non.

2. 2. 1. 2. La ponctuation

On a vu au début de l'introduction que, selon la norme, les propositions relatives non restrictives doivent être précédées par une virgule. La ponctuation est en quelque sorte une trace visuelle de la prosodie et des études ont confirmé la congruence entre ponctuation à l'écrit et frontières prosodiques à l'oral pour des structures contenant une ambiguïté syntaxique (Speer, Kjelgaard, & Dobroth, 1996 ; Steinhauer, 2003 ; Steinhauer & Friederici, 2001). La présence de la virgule sert à signaler le caractère non restrictif des relatives.

Dans des tâches de compréhension, notamment dans le cas d'ambiguïtés syntaxiques, on a pu voir que la ponctuation est un indice pertinent qui est pris en considération dans l'interprétation initiale (Schafer *et al.*, 1996) et qui facilite le traitement de ces structures ambiguës (Clifton, 1993 ; Cohen, Douaire, & Elsabbagh, 2001 ; Pickering & Traxler, 1998 ; Rayner, Kambe, & Duffy, 2000), en particulier si la virgule indique l'interprétation non préférée dans des phrases de type *garden path* (Hill & Murray, 1999). Ici, la ponctuation ne suffit pas à lever l'ambiguïté, mais elle devrait être un indice pertinent, tout du moins pour interpréter une proposition relative introduite par *qui* comme étant une relative non restrictive.

Dans l'étude d'un important corpus de mesures en *eye tracking* récoltées en français et en anglais, Pynte et Kennedy (2007) ont montré que la présence de ponctuation ne gêne pas en tant que telle le traitement de la phrase, mais que les mots suivant une virgule sont souvent omis (*skipped*), principalement car il s'agit souvent d'un mot de fonction. Cependant, la technique d'auto-présentation rend impossible d'omettre le traitement du relativiseur, contrairement à la technique de l'*eye tracking* où la phrase est présentée d'un seul tenant. L'observation d'une différence des temps de lecture au niveau des relativiseurs montrera une difficulté d'intégration plus qu'une différence du statut syntaxique de ces deux constituants, tous deux des mots de fonction.

De plus, la présence de la virgule avant le relativiseur ne devrait pas être trop coûteuse, si on se réfère aux résultats de Hirovani *et al.* (2006) : les temps de lecture (*eye tracking*) sont plus longs après une virgule, sauf quand celle-ci est à la tête d'une proposition relative non restrictive (comparée à une relative restrictive, sans virgule). Selon les auteurs, ce résultat indique qu'une relative non restrictive signale un contenu moins important, comme le suggère

aussi Givón (1990). Si tel était le cas, on pourrait s'attendre à plus d'erreurs pour les questions de compréhension qui portent sur la proposition relative plutôt que sur la principale.

Enfin, si on se réfère à l'étude de Clifton, Carlson et Frazier (2002) étudient l'effet de la prosodie dans l'attachement de propositions relatives comme (18) et montrent qu'il y a plus d'attachements vers N1 (*daughter*) quand la frontière prosodique à la tête de la proposition relative est plus importante (indiquant une proposition relative non restrictive) que la frontière prosodique précédente, placée après *daughter*.

(18) *The journalist interviewed the daughter of the colonel who had had the accident.*

La virgule au début de la proposition relative dans les phrases expérimentales, si elle affecte les préférences d'attachement, pourrait tout au plus favoriser un attachement à N1.

2. 2. 1. 3. Pré-test

Comme l'attachement est signalé par le contenu sémantique de la région de désambiguïsation, il convient de s'assurer que le contenu de la relative est plausible pour un seul des deux antécédents possibles et qu'il suffit à signaler l'attachement final.

Pour chaque item expérimental des propositions comme (69a-b) sont dérivées, soit 100 propositions. Le pré-test est un questionnaire de type « papier-crayon » où la tâche consiste à évaluer le degré de plausibilité de chaque proposition, à l'aide d'une échelle graduée de sept points : plus la valeur est élevée, plus la proposition est estimée plausible.

(69) a. *La serveuse termine son service.*

b. *La cliente termine son service.*

Afin d'augmenter au maximum la distance entre les deux propositions dérivées d'un même item, ces dernières sont réparties dans les deux moitiés du questionnaire, avec au moins dix propositions issues d'autres items qui les séparent, ceci pour éviter que les participants ne comparent directement les deux propositions. Il est fait en sorte d'équilibrer les propositions entre les deux moitiés du questionnaire, afin qu'il y ait autant de phrases très plausibles que peu plausibles. Ce questionnaire est réalisé en deux listes, pour contre-balancer l'ordre des propositions. Vingt participants issus de la même population que ceux qui ont participé à l'expérience réalisent ce questionnaire. La passation est collective et dure quinze minutes.

Les quarante meilleurs items expérimentaux sont sélectionnés à partir des valeurs de plausibilité accordées, en réalisant un compromis entre la plus grande valeur de plausibilité et la plus grande différence entre les deux propositions.

Les caractéristiques de ces quarante items sont présentées dans le Tableau 25. Un test de Student évalue le degré de significativité des facteurs. Le biais sémantique est très significatif [$t_1(19) = 37,99$; $p < 0,0001$; $t_2(49) = 33,09$; $p < 0,0001$]. On peut ainsi considérer que le contenu sémantique de la proposition relative, au niveau de la région de désambiguïsation, biaise suffisamment l'interprétation et lève l'ambiguïté d'attachement.

Tableau 25 : Caractéristiques du matériel expérimental (40 items retenus)

	Score de plausibilité	Fréquence	Longueur	Nb. syllabes
N plausible	6,46 ($\sigma = 0,52$)	23,85 ($\sigma = 61,06$)	7,90 ($\sigma = 1,66$)	2,40 ($\sigma = 0,50$)
N peu plausible	2,19 ($\sigma = 0,72$)	17,38 ($\sigma = 29,21$)	7,90 ($\sigma = 2,01$)	2,40 ($\sigma = 0,63$)

Les paires de noms ne diffèrent pas significativement du point de vue de leur fréquence d'occurrence, de leur longueur et du nombre de syllabes [$t < 1$ pour les trois comparaisons]. Notons aussi que parmi les items retenus, N1 et N2 sont féminins pour 8 d'entre eux.

2. 2. 1. 4. Plan expérimental

Trois variables indépendantes sont manipulées et varient de façon intra-individuelle. Le Relativiseur est une variable à deux modalités : *qui* ou *lequel/laquelle* (selon le genre de N1 et N2). La seconde variable est l'Attachement, signalé par la région de désambiguïsation, soit à N1, soit à N2. Enfin, le type de Question de compréhension, portant sur la proposition relative ou la proposition principale, constitue la troisième variable, elle aussi à deux modalités.

Les variables dépendantes sont les temps de lecture par région et l'exactitude de la réponse aux questions de compréhension.

Le plan expérimental est le suivant : $P_{48} * I_5 < R_2 * A_2 * Q_2 >$ (P : participants, I : items, 5 par condition ; R : Relativiseur, A : Attachement et Q : Question de compréhension).

Huit listes expérimentales sont construites, où 5 items de chacune des conditions expérimentales sont répartis ; de plus chaque item n'est présent qu'une seule fois dans chaque liste. Les items expérimentaux sont répartis parmi 100 remplisseurs et chaque item apparaît à la même position dans les listes. Les trois variables expérimentales sont contre-balançées sur ces huit listes qui sont attribuées aux participants de façon aléatoire.

2. 2. 2. Procédure expérimentale

Les participants sont installés dans une salle équipée de 8 ordinateurs isolés les uns des autres par de grands panneaux. Ces ordinateurs sont de type PC équipés du logiciel DMDX (Forster & Forster, 2002) et les réponses sont données en utilisant une manette de jeu Logitech Wingman. Tout d'abord, les participants donnent leur consentement écrit et quelques informations telles que leur âge, leur main dominante, s'ils ont des problèmes de vision, d'audition ou de dyslexie.

Les instructions concernant l'expérience sont délivrées de façon orale puis écrite, sur l'écran de l'ordinateur. Les participants ont pour tâche de lire les phrases région par région en cliquant sur un bouton de la manette, en prenant soin d'avoir un rythme de lecture le plus naturel possible. Ils gardent la manette en main, un index reposant sur le bouton droit (pour répondre « oui » aux questions de compréhension) ou gauche (réponses « non »). La consigne pour les questions de compréhension est d'être le plus rapide et le plus juste possible.

La procédure expérimentale utilisée est celle d'auto-présentation segmentée ou *moving window non cumulative self-paced reading*. Ce paradigme est couramment utilisé dans les études en compréhension écrite (F. Ferreira & Henderson, 1990 ; Just, Carpenter, & Wooley, 1982 ; Mitchell & Green, 1978) car il permet de mesurer les temps de lecture pour chaque région, puisque c'est le lecteur lui-même qui fait apparaître les régions en actionnant un bouton. Une illustration du paradigme est présentée Figure 15.

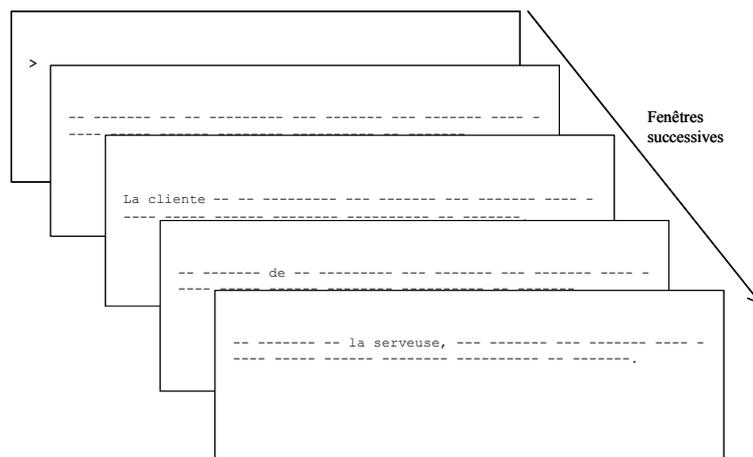


Figure 15 : Le paradigme d'auto-présentation segmentée

Le participant fait apparaître les fenêtres successives en appuyant sur un bouton de la manette. Il est alors possible d'enregistrer les temps de lecture pour chaque fenêtre, avec une

précision de l'ordre de la milliseconde. Les caractères de la phrases sont symbolisés par des tirets (-). La procédure est dite non cumulative, car une fois qu'une région est présentée, elle est remplacée à nouveau par les tirets.

Une contrainte est appliquée dans la présentation des contenus des phrases : un nom apparaît systématiquement en même temps que son article. Cela permet d'afficher la région de désambiguïsation, pour tous les items, en une seule fenêtre, mais aussi N1 et N2. A l'exception de cette contrainte, chaque région est composée d'un seul mot.

Les phrases expérimentales étant trop longues pour n'occuper qu'une seule ligne à l'écran, il est nécessaire d'insérer une rupture de ligne. Cette dernière se trouve toujours dans la région *spillover* qui est le groupe prépositionnel suivant l'objet de la relative (i. e. la région de désambiguïsation), entre le troisième et le quatrième mot, que ces noms soient ou non regroupés en une fenêtre « déterminant et nom ».

Cette technique expérimentale est tout à fait satisfaisante pour mesurer les éventuels effets de ré-analyse auxquels on s'attend. À propos des techniques expérimentales utilisées pour évaluer le traitement d'ambiguïtés syntaxiques, Frazier et Clifton (1996) suggèrent que la technique d'auto-présentation convient très bien pour un matériel expérimental si l'ambiguïté temporaire est résolue dans la phrase, ici au niveau de la région de désambiguïsation, et qu'aucune autre information contextuelle ne contraint l'interprétation initiale. La mesure des temps de lecture peut alors révéler des interruptions de traitement (*disruptions*) lorsque l'information, ici sémantique, signale une interprétation incongrue avec les préférences structurelles de traitement (voir aussi F. Ferreira & Clifton, 1986 ; F. Ferreira & Henderson, 1990, 1991).

Par contre, si le contexte est manipulé et précède la levée de l'ambiguïté, Frazier et Clifton (1996) estiment qu'une technique dont le niveau d'analyse temporel est très fin s'avère nécessaire. L'*eye tracking* est alors utilisé.

Revenons un moment sur la présentation des fenêtres successives, la segmentation de la phrase, et leur effet potentiel sur les préférences d'attachement. Gilboy et Sopena (1996) l'ont évalué dans une expérience en auto-présentation. Si le GN « N1 de N2 » est présenté en une seule fenêtre, il y a une préférence pour attacher la proposition relative à N1 (voir aussi Carreiras & Clifton, 1993 ; Cuetos & Mitchell, 1988 ; Gilboy *et al.*, 1995). Cette préférence est plus faible si les deux noms, N1 et N2, sont présentés dans des fenêtres différentes. C'est ce qui est mesuré aussi dans les expériences de Carreiras et Clifton (1999). La segmentation choisie pour cette expérience pourrait ainsi favoriser une préférence d'attachement à N2.

Un essai se déroule de la façon suivante. Tout d'abord, le symbole « > » signifie le début de la phrase ; le participant fait apparaître la première fenêtre en appuyant sur un bouton de la manette. Par pressions successives, la phrase défile et les temps de réponse sont mesurés pour chaque fenêtre. Après la dernière fenêtre, une question de compréhension apparaît, aussi bien pour les items expérimentaux que pour les remplisseurs. Le participant répond par « oui » ou « non » en appuyant sur le bouton correspondant. L'exactitude de la réponse ainsi que les temps de réponse sont mesurés et l'essai suivant est prêt.

Les 140 essais sont répartis en trois blocs : 50 essais pour les deux premiers blocs et 40 pour le dernier. Ces blocs sont séparés par des pauses dont la longueur est déterminée par le participant. L'expérience débute avec 6 essais d'entraînement afin de se familiariser avec le paradigme d'auto-présentation. L'expérience dure 25 à 30 minutes.

3. Résultats

Les analyses portent sur les erreurs aux questions de compréhension ainsi que sur les temps de réponse (TR) pour chaque région. L'analyse des temps de réponse aux questions de compréhension n'ayant révélé aucun effet statistiquement significatif, cette analyse n'est pas présentée.

Les analyses de la variance (ANOVAs) sont conduites par sujet ($F1$) et par item ($F2$) et trois facteurs varient de façon intra-individuelle pour les questions de compréhension : la Question (sur la relative ou la principale), le Relativiseur (*qui, lequel/laquelle*) et l'Attachement (à N1 ou N2). Le facteur Question est non pertinent pour les temps de lecture, aussi n'est-il pas inclus dans l'analyse.

Afin de limiter la variabilité due à la répartition aléatoire des items et de leurs conditions expérimentales à travers les listes (Pollatsek & Well, 1995; van Gompel *et al.*, 2005), la Liste sujet est ajoutée en tant que variable interindividuelle dans l'analyse par sujet et la Liste des items dans l'analyse par item. Les valeurs de F pour les Listes et les interactions avec les autres variables ne sont pas informatives en elles-mêmes et ne sont pas reportées dans les résultats.

3. 1. Les questions de compréhension

Le pourcentage d'erreurs au niveau des remplisseurs confirme que les participants lisent les phrases dans le but de les comprendre, puisqu'ils font peu d'erreurs, 13 % en moyenne.

L'ANOVA qui croise la Question (sur la relative ou la principale), l'Attachement et le Relativiseur révèle un effet significatif de la Question [$F1(1, 40) = 21,39$; $p < 0,001$; $F2(1, 32) = 19,00$; $p < 0,001$] et un effet de l'Attachement [$F1(1, 40) = 14,13$; $p = 0,001$; $F2(1, 32) = 11,20$; $p = 0,002$]. Il n'y a pas d'effet principal du Relativiseur [$F_s < 1$].

Les phrases expérimentales, elles aussi, sont globalement bien comprises puisqu'il n'y a sur l'ensemble des participants que 17 % d'erreurs. Les pourcentages d'erreurs pour chaque condition expérimentale sont représentés sur la Figure 16.

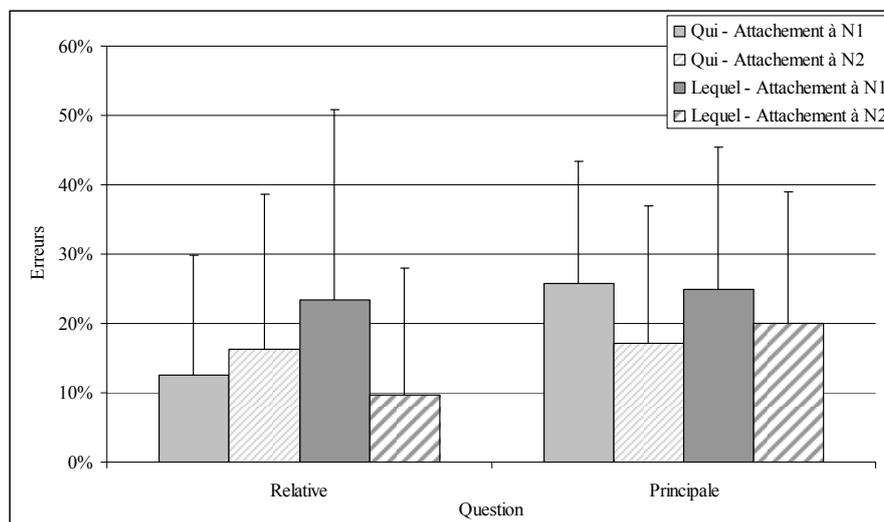


Figure 16 : Pourcentage d'erreurs pour les questions portant sur la relative ou sur la principale, en fonction du Relativiseur et de l'Attachement

Les questions portant sur la proposition principale sont plus difficiles (23 %) que les questions qui concernent la proposition relative (14 %). Les phrases dont la relative est attachée à N2 sont plus faciles que les phrases avec un attachement à N1. Il y a en effet moins d'erreurs (N2 : 16 % ; N1 : 22 %).

Les interactions doubles ne sont pas significatives, que ce soit entre le Relativiseur et l'Attachement [$F1(1, 40) = 1,68$; $p = 0,20$; $F2(1, 32) = 2,10$; $p = 0,16$], entre la Question et l'Attachement [$F_s < 1$] ou entre la Question et le Relativiseur [$F_s < 1$].

Si les interactions doubles ne sont pas statistiquement significatives, l'interaction triple, elle, l'est [$F1(1, 40) = 5,65 ; p = 0,02 ; F2(1, 32) = 5,31 ; p = 0,03$]. Afin de déterminer plus précisément la nature de cette interaction, des ANOVAs sont conduites sur les erreurs aux questions de compréhension portant sur la relative ou la principale séparément.

3. 1. 1. Analyses des erreurs aux questions portant sur la relative

Ces ANOVAs montrent un effet significatif de l'Attachement [$F1(1, 40) = 6,60 ; p = 0,01 ; F2(1, 32) = 6,86 ; p = 0,01$], mais pas du Relativiseur [$F_s < 1$]. L'interaction est significative [$F1(1, 40) = 4,71 ; p = 0,04 ; F2(1, 32) = 16,94 ; p = 0,0003$].

Des analyses des effets simples sont réalisées pour évaluer l'effet de l'Attachement en fonction du Relativiseur. On observe que l'Attachement, à N1 ou à N2, n'a d'effet que pour les relatives introduites par *lequel*, puisqu'on mesure une différence significative [$F1 = 6,98 ; p = 0,01 ; F2 = 25,12 ; p = 0,0001$], absente avec les relatives introduites par *qui* [$F_s < 1$]. Quand la question de compréhension porte sur la relative, les phrases contenant *lequel* et désambiguïsées vers N2 sont mieux comprises, car les participants font moins d'erreurs.

Pour l'effet du Relativiseur selon que les phrases sont désambiguïsées vers N1 ou N2, on mesure une différence entre *qui* et *lequel* pour les relatives attachées à N1, par item [$F2 = 9,44 ; p = 0,004$], mais pas pour l'analyse par sujet [$F1 = 2,62 ; p = 0,11$]. Un résultat comparable est mesuré pour les questions sur des relatives attachées à N2, avec une différence significative par item [$F2 = 7,56 ; p = 0,01$] et non par sujet [$F1 = 2,10 ; p = 0,12$]. Ainsi, pour une relative attachée à N1, les phrases avec *qui* entraînent moins d'erreurs qu'avec *lequel*. L'inverse est observé pour les relatives attachées à N2.

3. 1. 2. Analyses des erreurs aux questions portant sur la principale

Les ANOVAs sur les questions portant sur la principale montrent qu'il y a un effet de l'Attachement [$F1(1, 40) = 7,93 ; p = 0,007 ; F2(1, 32) = 4,29 ; p = 0,05$], mais pas d'effet du Relativiseur [$F_s < 1$] ni d'interaction [$F_s < 1$].

Pour ce type de questions, les participants font moins d'erreurs pour les relatives désambiguïsées vers N2.

3. 2. Les temps de lecture

Pour l'analyse des temps de lecture, les différents constituants des phrases expérimentales sont regroupés en régions d'analyse. Les constituants de la fin de la relative, après la

désambiguïsation, sont ainsi regroupés en une seule région, la région *spillover*, et la fin de la principale subit le même traitement. Ces régions sont illustrées dans l'exemple (70) ci-dessous.

(70) *La serveuse*_{N1} / *de*_{de} / *la cliente*_{N2} / *qui*_{relativiseur} / *termine*_{verbe de la relative} / *son service*_{désambiguïsation} / *dans moins d'une heure*_{spillover} / *renverse*_{verbe de la principale} / *l'assiette de légumes*_{fin} /

Les variables intra-individuelles sont le Relativiseur (*qui, lequel*) et l'Attachement (N1, N2) et le facteur Liste est une variable interindividuelle. Les effets de la Liste ne sont pas reportés car ils ne sont pas informatifs. Les temps de lecture moyens sont représentés dans le Tableau 26. En ce qui concerne la région *spillover*, les régions 3 et 4 sont rassemblées car tous les items n'ont pas nécessairement une quatrième région (cf. rupture de phrase). Les derniers mots de la phrase (Fin) sont eux aussi rassemblés.

Tableau 26 : Temps de lecture moyen (ms) par région d'analyse, en fonction du relativiseur et du type d'attachement

Mot	Qui		Lequel	
	Att. N1	Att. N2	Att. N1	Att. N2
N1	637 (404)	639 (481)	675 (539)	657 (447)
de	359 (176)	380 (248)	376 (213)	372 (217)
N2	572 (596)	585 (592)	571 (523)	612 (668)
Rel.	581 (450)	523 (407)	674 (556)	600 (494)
Verbe de la relative	436 (354)	426 (361)	470 (386)	469 (397)
Désambiguïsation	535 (487)	517 (567)	511 (371)	515 (499)
Spillover 1	443 (144)	430 (138)	493 (198)	436 (136)
Spillover 2	456 (148)	454 (139)	483 (159)	417 (136)
Spillover 3 et 4	778 (330)	702 (292)	825 (418)	706 (291)
Verbe de la principale	475 (316)	483 (303)	495 (339)	508 (386)
Fin	1545 (924)	1528 (921)	1639 (1067)	1560 (964)

3. 2. 1. Le premier nom

Pour cette première région, les TR sont très proches et aucune différence significative n'est observée au niveau du Relativiseur [$F1(1, 40) = 2,36$; $p = 0,13$; $F2(1, 32) = 1,58$; $p = 0,22$], de l'Attachement [$F_s < 1$] ni de l'interaction de ces deux facteurs [$F_s < 1$].

3. 2. 2. Le « de »

Pour *de*, on observe des temps de lecture proches d'une condition à l'autre, ce qui est confirmé par l'ANOVA pour laquelle aucune différence statistique significative n'est mise en évidence, que ce soit pour le Relativiseur [$F_s < 1$], l'Attachement [$F_s < 1$] ou une interaction entre ces deux facteurs [$F_1(1, 40) = 1,95 ; p = 0,17 ; F_2(1, 32) = 1,12 ; p = 0,30$].

3. 2. 3. Le deuxième nom

Là encore, les temps de lecture restent proches, car les phrases sont jusqu'à présent de structure et de contenu équivalents d'une condition à l'autre. Aucune différence significative n'est observée. Il n'y a effet ni du Relativiseur [$F_s < 1$], ni de l'Attachement [$F_1 < 1 ; F_2(1, 32) = 1,27 ; p = 0,27$] et il n'y a pas d'interaction [$F_s < 1$].

3. 2. 4. Le relativiseur

C'est à partir du relativiseur que les phrases diffèrent, puisqu'il s'agit soit de *qui* soit de *lequel/laquelle*.

Les ANOVAs mettent en évidence un effet du Relativiseur [$F_1(1, 40) = 12,80 ; p = 0,001 ; F_2(1, 32) = 18,53 ; p < 0,0001$] : *lequel* est plus long à traiter que *qui* (respectivement, 638 ms ; $\sigma = 372$ et 551 ms ; $\sigma = 302$). Il y a aussi un effet de l'Attachement [$F_1(1, 40) = 18,71 ; p < 0,001 ; F_2(1, 32) = 10,45 ; p = 0,003$], avec des temps de lecture plus courts quand le site d'attachement est N2, tel que cela est signalé plus tard dans la phrase. Cette différence n'est pas liée à la levée de l'ambiguïté ni à l'attachement final qui n'est signalé que plus tard. Elle pourrait être due à une différence selon l'ordre des noms dans « N1 de N2 », *la cliente de la serveuse* ou *la serveuse de la cliente*. Je reviendrai sur cet aspect lors de la discussion. L'interaction n'est pas significative [$F_s < 1$].

3. 2. 5. Le verbe de la relative

Au niveau du verbe de la relative, il semble qu'on retrouve un effet *spillover* du type de relativiseur. La différence selon le Relativiseur est en effet statistiquement significative [$F_1(1, 40) = 6,83 ; p = 0,01 ; F_2(1, 32) = 5,80 ; p = 0,02$] : les temps de lecture sont plus longs quand le verbe suit *lequel* (470 ms ; $\sigma = 264$) plutôt que *qui* (431 ms ; $\sigma = 242$).

Les autres comparaisons n'atteignent pas le seuil de significativité. C'est le cas pour l'Attachement [$F_s < 1$] : les temps de lecture sont comparables que la phrase soit

ultérieurement désambiguïsée vers N1 ou N2 (453 ms ; $\sigma = 248$ et 448 ms ; $\sigma = 258$, respectivement). L'interaction n'est pas non plus significative [$F_s < 1$].

3. 2. 6. La désambiguïsation

L'ANOVA conduite sur la région de désambiguïsation montre l'absence d'un effet du Relativiseur [$F_1 < 1$; $F_2(1, 32) = 1,10$; $p = 0,30$]. Les temps de lecture sont similaires que la relative soit introduite par *qui* (526 ms ; $\sigma = 351$) ou *lequel* (512 ms ; $\sigma = 301$). Il n'y a pas non plus d'effet de l'Attachement [$F_s < 1$], avec un temps moyen de 522 ms ($\sigma = 307$) pour les phrases désambiguïsées vers N1 et 516 ms ($\sigma = 345$) pour celles désambiguïsées vers N2. L'interaction n'atteint pas non plus le seuil de significativité [$F_s < 1$].

3. 2. 7. La région *spillover*

La région *spillover* suit celle où l'ambiguïté est levée. S'il faut du temps pour intégrer l'information sémantique, des effets des conditions expérimentales peuvent être mesurés dans cette région qui est analysée mot par mot.

Pour le premier mot qui suit la région de désambiguïsation, l'ANOVA révèle un effet du Relativiseur qui est marginalement significatif par sujet [$F_1(1, 40) = 3,90$; $p = 0,06$] et significatif par item [$F_2(1, 32) = 4,18$; $p = 0,05$] : les temps de lecture sont plus importants pour les relatives avec *lequel* (464 ms ; $\sigma = 172$) que celles avec *qui* (437 ms ; $\sigma = 141$). L'effet de l'Attachement est marginalement significatif par sujet [$F_1(1, 40) = 3,47$; $p = 0,07$] et significatif par item [$F_2(1, 32) = 5,28$; $p = 0,03$]. Que l'attachement soit vers N1 cause plus de difficultés que pour un attachement à N2 (N1 : 467 ms ; $\sigma = 174$; N2 : 433 ms ; $\sigma = 137$). L'interaction n'est pas significative [$F_1(1, 40) = 2,04$; $p = 0,16$; $F_2(1, 32) = 1,62$; $p = 0,21$].

Au niveau du second mot du *spillover*, il n'y a pas de différence significative selon la forme du relativiseur [$F_s < 1$]. Les temps de lecture sont comparables (*qui* : 455 ms ; $\sigma = 142$; *lequel* : 449 ms ; $\sigma = 151$). On obtient par contre un effet significatif de l'Attachement [$F_1(1, 40) = 8,44$; $p = 0,006$; $F_2(1, 32) = 8,16$; $p = 0,008$], où une désambiguïsation vers N2 (435 ms ; $\sigma = 138$) entraîne des temps de lecture plus courts qu'une désambiguïsation vers N1 (469 ms ; $\sigma = 154$). L'interaction entre ces deux facteurs est marginalement significative par sujet [$F_1(1, 40) = 3,52$; $p = 0,07$] et significative par item [$F_2(1, 32) = 6,67$; $p = 0,01$].

L'analyse des effets simples pour le deuxième mot porte sur l'effet de l'Attachement en fonction de la forme du relativiseur. C'est l'effet de l'Attachement pour les phrases avec

lequel qui contribue à cette interaction, puisqu'on obtient une différence significative pour ce relativiseur [$F1(1, 40) = 9,78$; $p = 0,003$; $F2(1, 32) = 15,19$; $p = 0,0005$], mais pas pour *qui* [$F1(1, 40) = 1,98$; $p = 0,17$; $F2 < 1$]. Les temps de lecture moyens sont de 417 ms ($\sigma = 136$) pour une relative introduite par *lequel* et avec un attachement à N2, de 482 ms ($\sigma = 159$) avec un attachement à N1 (cf. Tableau 26). Pour les phrases avec *qui*, les temps de lecture sont de 456 ms ($\sigma = 148$) avec un attachement à N1 et de 454 ms ($\sigma = 139$) avec un attachement à N2.

Au niveau du troisième mot (et éventuellement quatrième) de la région spillover, on retrouve l'absence d'effet du Relativiseur [$F1(1, 40) = 1,16$; $p = 0,29$; $F2 < 1$] ; les temps de lecture ne diffèrent pas entre les relatives avec *qui* (740 ms ; $\sigma = 312$) et celles avec *lequel* (766 ms ; $\sigma = 363$). L'effet de l'Attachement est significatif [$F1(1, 40) = 10,96$; $p = 0,002$; $F2(1, 32) = 19,49$; $p = 0,0001$] : les temps de lecture sont plus importants quand l'Attachement est à N1 (802 ms ; $\sigma = 376$) plutôt qu'à N2 (704 ms ; $\sigma = 290$). L'interaction n'est pas significative [$F_s < 1$].

3. 2. 8. Le verbe de la principale

Les ANOVAs conduites sur le verbe de la principale ne révèlent aucun effet significatif. Il n'y a pas de différence selon le Relativiseur [$F1(1, 40) = 1,38$; $p = 0,25$; $F2(1, 32) = 2,61$; $p = 0,12$] : les temps de lecture mesurés pour les phrases avec *qui* sont proches de ceux mesurés pour celles avec *lequel* (respectivement, 479 ms ; $\sigma = 223$ et 502 ms ; $\sigma = 222$).

L'effet de l'Attachement n'est pas non plus significatif [$F_s < 1$]. Les temps de lecture sont comparables quel que soit l'attachement signalé (N1 : 485 ms ; $\sigma = 225$ et N2 : 496 ms ; $\sigma = 220$). Il n'y a pas non plus d'interaction significative [$F_s < 1$].

3. 2. 9. La fin de la phrase

Cette dernière région est identique dans toutes les conditions expérimentales et, si un effet est observé, il s'agirait d'un effet de fin de phrase (*wrap up* en anglais).

Il n'y a pas non plus d'effet du Relativiseur [$F1(1, 40) = 2,65$; $p = 0,11$; $F2(1, 32) = 1,60$; $p = 0,22$] ni de l'Attachement [$F1(1, 40) = 2,99$; $p = 0,09$; $F2(1, 32) = 1,05$; $p = 0,31$], ni d'interaction [$F_s < 1$].

Post-test pour évaluer l'ordre des noms

Conserver un contenu identique au niveau de la région de désambiguïsation quel que soit l'attachement signalé obligeait à inverser l'ordre des noms dans le GN « N1 de N2 ». Lors de

l'élaboration du matériel, il a été fait en sorte que les paires de noms soient réversibles, que dans (68a-d) *la cliente de la serveuse* soit aussi bien que *la serveuse de la cliente*. L'effet significatif de l'Attachement sur le relativiseur n'est pas dû à la manipulation du contenu sémantique de la relative qui n'apparaît que plus tard. Il est possible qu'un ordre des noms soit préféré à l'autre. Le post-test évalue la possibilité d'un biais de l'ordre des noms.

Dix participants, tous de langue maternelle française et membres du Laboratoire Dynamique du Langage, ont réalisé un questionnaire de type « papier-crayon ». La tâche consiste à évaluer les deux ordres à l'aide d'une échelle à cinq gradients placée entre les deux propositions (*la cliente de la serveuse est là* ou *la serveuse de la cliente est là*), comme illustré dans la Figure 17. Cocher la plus faible valeur signifie que la proposition à gauche est meilleure et cocher la plus haute valeur signifie que la droite est meilleure.

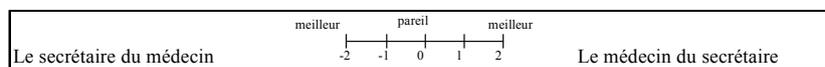


Figure 17 : Exemple d'un essai dans le pré-test évaluant l'ordre des noms

L'ordre des propositions (N plus plausible en N1 ou N2, à gauche ou à droite) est contre-balançé sur l'ensemble des essais, 50 au total (i. e. testant toutes les propositions issues des items construits originellement). Ces essais sont précédés de propositions d'entraînement afin de familiariser les participants avec l'utilisation de l'échelle. Le questionnaire nécessite environ 10 minutes.

Les résultats montrent que sur l'ensemble des items il y a une préférence pour l'ordre de présentation où N2 est le nom qui est biaisé sémantiquement dans les items expérimentaux. La valeur moyenne est en effet de -0,62 ($\sigma = 1,29$).

Il y aurait ainsi une préférence pour un ordre de N1 et N2 comme dans (68b), qui peut expliquer pourquoi les participants ont plus de difficultés à intégrer le relativiseur quand la phrase est ultérieurement désambiguïsée à N1 : ce n'est pas un ordre favorisé. Néanmoins, ce biais du matériel disparaît au delà du relativiseur et ne devrait pas affecter outre mesure les TR.

4. Discussion

Le but de cette expérience est d'évaluer les difficultés rencontrées par les participants lors du traitement de phrases contenant une proposition relative dont l'attachement est temporairement ambigu. Deux antécédents sont possibles, N1 et N2, dans une structure « N1 de N2 » et l'ambiguïté n'est levée qu'au niveau de l'objet de la relative.

L'hypothèse principale concerne l'effet de la forme du relativiseur. Une expression anaphorique, par exemple un relativiseur, signale le degré d'accessibilité de l'entité à laquelle elle fait référence formellement dans le discours. Plus une expression anaphorique est courte (i. e. atténuée) et moins elle est informative (p. ex. un pronom), plus cette entité est accessible dans le discours. Dans le cas d'un relativiseur, je propose que *lequel*, relativiseur moins atténué et plus informatif que *qui*, signale un attachement à un antécédent moins accessible. Dans une structure « N1 de N2 », c'est N2 l'antécédent le moins accessible, car N1 est la tête du prédicat, mais aussi le sujet de la principale (cf. effet de la fonction syntaxique, Chapitre 1) et N2 est le modifieur de N1. C'est pourquoi on s'attend à une préférence pour N2 plus importante avec *lequel* plutôt qu'avec *qui*.

Regardons d'abord l'effet de la forme du relativiseur, indépendamment de l'attachement final. On mesure une différence entre les phrases avec *qui* ou avec *lequel* dès la présentation du relativiseur : traiter *lequel* est plus long que traiter *qui*. On peut interpréter cette différence de deux façons. La différence de fréquence entre les deux relativiseurs peut tout à fait contribuer aux effets observés (cf. *informational load hypothesis*, modèle de satisfaction de contraintes et de compétition et d'intégration). *Lequel* est un relativiseur moins fréquent que *qui* et le lecteur peut être confronté à une difficulté plus importante en traitant ce premier relativiseur. Cet effet de la fréquence devrait conduire par exemple à une meilleure intégration de la structure en cours de traitement et faciliter le traitement de la suite de la phrase (van Gompel & Majid, 2004). Néanmoins, les temps de lecture plus importants pour certains mots ultérieurs (pour les phrases avec *lequel*) infirment tout du moins en partie l'effet de la fréquence. Selon Fodor cependant (p. 289, 1998), la fréquence n'est pas suffisante pour contribuer à une sollicitation plus importante des ressources en mémoire.

Cette difficulté au niveau du relativiseur peut aussi être liée aux informations supplémentaires que fournit *lequel* par rapport à *qui* : le marquage en genre et en nombre, mais aussi la restrictivité, puisque *lequel* ne signale que des propositions relatives non restrictives, contrairement à *qui*. Néanmoins, on peut supposer que les lecteurs ont pris en

compte la présence de la virgule après N2 et, au moment où ils traitent *qui*, ils intègrent la non restrictivité pour la relative avec *qui*. Intégrer le relativiseur, c'est entre autres choses établir un lien entre ce relativiseur et son antécédent, c'est-à-dire attacher la relative. Les expériences relevées dans la littérature (p. ex. Carreiras & Clifton, 1993, 1999; p. ex. Cuetos & Mitchell, 1988), notamment celle en français (Baccino *et al.*, 2000 ; Frenck-Mestre & Pynte, 2000a ; Zagar *et al.*, 1997), ont reproduit de façon extensive une préférence pour N1 dans de nombreuses langues. Un attachement initial à N1 semble majoritaire. Lors de l'intégration du relativiseur, l'accessibilité signalée par celui-ci peut guider l'attachement initial, avec un attachement à N1 pour *qui* et un attachement à N2 pour *lequel*. L'attachement à N2 étant moins fréquent, il peut nécessiter plus de temps aux participants, d'où un temps de lecture plus long pour *lequel*.

L'effet de la forme du relativiseur se prolonge sur le mot suivant, le verbe de la relative. On retrouve un effet du relativiseur au niveau du premier mot qui suit la désambiguïsation, avec plus de difficultés quand la relative est introduite par *lequel*. Quant aux erreurs aux questions de compréhension, qui traduisent les mécanismes post-interprétatifs, il n'y a pas d'effet du relativiseur. Les phrases avec *qui* ou avec *lequel* sont aussi bien comprises et la difficulté de traitement due à la forme du relativiseur est résolue. À l'issue de la présentation des phrases, les participants les ont aussi bien comprises, quelle que soit la forme du relativiseur. Il s'agirait ainsi d'une difficulté locale, pendant l'interprétation et la compréhension des phrases.

La variable Attachement manipule le contenu sémantique de la relative, de façon à indiquer un attachement à N1 ou à N2. On observe une différence au niveau du relativiseur : les phrases qui sont ultérieurement désambiguïsées vers N1 sont plus longues à traiter. Cela ne peut être dû à l'information sémantique qui signale l'attachement, car elle est postérieure, mais c'est peut-être dû à un biais expérimental.

La manipulation de l'attachement est liée à l'ordre de présentation des antécédents N1 et N2. Si on reprend les exemples (68a-d) répétés ci-dessous, *la serveuse de la cliente* est désambiguïsée vers N1, exemple (68a c) et *la cliente de la serveuse* vers N2, (68b d). Un biais de l'ordre de présentation pourrait tout à fait contribuer à cet effet de l'Attachement. Le post-test confirme dans l'ensemble une préférence pour un ordre, celui où l'attachement est signalé à N2. Ce biais expérimental est cependant temporaire, puisqu'il disparaît pour le mot suivant.

(68) a. c. *La serveuse de la cliente, qui/laquelle termine son service dans moins d'une heure, renverse l'assiette de légumes.*

b. d. *La cliente de la serveuse, qui/laquelle termine son service dans moins d'une heure, renverse l'assiette de légumes.*

La levée de l'ambiguïté au niveau de l'objet de la relative n'entraîne pas d'effet immédiat de l'Attachement. Cet effet n'est statistiquement significatif que dans la région *spillover*, pour tous les mots. Que l'attachement soit à N1 entraîne des temps de lecture plus importants. Cette difficulté est résolue à partir du verbe de la proposition principale jusqu'à la fin de la phrase.

L'attachement final rend néanmoins les phrases avec attachement à N1 plus difficiles que celles désambiguïsées vers N2. En effet, on retrouve un effet de l'Attachement sur les questions de compréhension : il suggère que les participants ont moins bien compris les phrases avec un attachement à N1, qui peuvent être ainsi considérées comme plus difficiles. On pourrait alors envisager que dans cette expérience, les participants montrent une préférence pour un attachement à N2.

C'est la présence d'interactions entre les variables Relativiseur et Attachement ainsi que l'analyse des effets simples qui sont les plus informatives. On mesure de telles interactions pour des questions de compréhension sur la relative, mais aussi, et de façon marginale avec l'analyse par sujet, au niveau du deuxième mot de la région *spillover*. Dans les cas énumérés, on mesure des performances comparables pour les relatives introduites par *qui*, que l'attachement signalé soit à N1 ou à N2. Pour les relatives avec *lequel*, par contre, les participants ont plus de difficultés à traiter un attachement à N1 qu'à N2. On mesure alors des temps de lecture plus importants (région *spillover*, deuxième mot) et plus d'erreurs aux questions de compréhension. Ce sont donc les phrases avec *lequel* qui contribuent à l'effet de l'Attachement. La préférence pour N2 est liée à la présence du relativiseur *lequel*, qui signale bien un attachement à un antécédent peu accessible, N2.

Dans l'ensemble, ces résultats sont conformes à l'Hypothèse 1 : un relativiseur comme *lequel* signale un antécédent peu accessible et c'est pourquoi il est attaché de façon préférentielle à N2. Quand l'ambiguïté est levée et qu'un attachement à N1 est signalé par le contenu sémantique, cela entraîne des difficultés pour le traitement de phrases. L'attachement initial est à N2 et changer cette interprétation, ré-analyser la phrase, est coûteux.

Pour les phrases avec *qui*, on se serait attendu à une préférence pour N1, avec des temps de lecture plus courts quand l'ambiguïté est levée vers cet antécédent. L'absence de différence selon l'attachement pour les phrases avec *qui* suggère qu'il n'y a pas de préférence d'interprétation initiale. On pourrait interpréter que cette absence de différence signale le défaut de ré-analyse, selon la théorie du *garden path* (Frazier, 1987 ; Frazier & Rayner, 1982, 1987 ; Rayner *et al.*, 1983). D'après le modèle *unrestricted race* (van Gompel *et al.*, 2000), cela signifie que les participants interprètent alternativement la relative en *qui* comme attachée au hasard à N1 ou à N2. La révision possible qui a lieu dans la région *spillover* ne conduit pas à une différence significative parce que, s'il y a parfois révision, elle est contrebalancée par les cas où l'analyse initiale est correcte. On pourrait aussi interpréter l'absence d'effet de l'Attachement pour les relatives en *qui* par le fait que les participants activent les deux interprétations possibles et les conservent jusqu'à la levée de l'ambiguïté. C'est ce que propose le modèle de compétition et d'intégration (McRae *et al.*, 1997 ; McRae *et al.*, 1998). S'il n'est pas vraiment possible de départager ces interprétations, on peut néanmoins considérer que le relativiseur *qui*, dans cette expérience, ne signale pas un attachement en particulier.

Si on met en regard ces résultats avec ceux qui ont été obtenus dans la littérature sur le traitement des propositions relatives (non ambiguës), on observe que les participants commettent plus d'erreurs quand les questions portent sur la principale que lorsque les questions portent sur la relative. Ce résultat est conforme avec les données de la littérature (Delle Luche, 2004, 2006 ; Gibson, 1998, 2000 ; Gibson *et al.*, 2005 ; Holmes & O'Regan, 1981 ; King & Just, 1991) : la quantité d'intégration et le coût de traitement sont plus importants au niveau de la principale qu'à celui de la relative sujet (Gibson, 1998).

Selon la *construal theory* (Frazier & Clifton, 1997), que les relatives en français aient pour relativiseurs un pronom relatif implique que la proéminence des entités du discours guide l'attachement. *Qui* et *lequel* sont des pronoms relatifs et devraient être attachés préférentiellement à N1, de façon comparable. Or, on mesure dans cette expérience, aussi bien pour les mesures *on-line* (temps de lecture) que *off-line* (questions de compréhension) une préférence pour N2 avec *lequel*, ainsi qu'un effet de la forme du relativiseur. Ces résultats infirment les prédictions de la *construal theory*, mais aussi celle de la théorie du *garden path*. Cette dernière prévoit d'abord une préférence initiale pour N2 (stratégie de *late closure*) mais surtout l'absence d'un effet de la forme du relativiseur, puisque seules les informations

syntaxiques sont prises en compte. Si on observe bien une préférence initiale pour N2, cette préférence est liée à la forme du relativiseur.

Enfin, revenons sur les aspects méthodologiques et sur le biais du matériel qui peut être envisagé. Le type de désambiguïsation est contraint par la forme des relativiseurs, et surtout par le marquage en genre et en nombre de *lequel*. En effet, le type de désambiguïsation le plus souvent utilisé est le marquage en genre d'un adjectif, quand N1 et N2 diffèrent à ce niveau (Baccino *et al.*, 2000 ; Carreiras *et al.*, 2004 ; De Vincenzi & Job, 1993, 1995 ; Frenck-Mestre & Pynte, 2000a, 2000b ; Zagar *et al.*, 1997). Cela n'était pas possible avec le matériel utilisé. D'autres chercheurs ont utilisé une désambiguïsation sémantique (Carreiras & Clifton, 1993, 1999), mais les informations sémantiques sont souvent basées sur des prototypes sexués (p. ex. *a accouché* ne peut s'appliquer qu'à une femme). Afin d'avoir une région de désambiguïsation identique que l'attachement soit signalé à N1 ou à N2, ce qui était souhaitable, la position des antécédents est inversée selon l'attachement. Les paires d'antécédents ont été élaborées de façon à être acceptables dans les deux positions, mais il n'a pas été possible d'éviter un biais de l'ordre des antécédents (cf. post-test). Les mesures de TR semblent montrer que l'effet de ce biais est temporaire et résolu au delà du relativiseur, mais on ne peut écarter la possibilité que ce biais ait une influence sur les traitements ultérieurs dans la phrase. Il serait intéressant de réaliser une expérience comparable où cette fois l'ordre des antécédents serait fixe, mais où le contenu sémantique de la région de désambiguïsation changerait selon le type d'attachement. On pourrait alors présenter des phrases comme (71a-b), où *whisky* signale un attachement à N1 (71a) et *biberon* un attachement à N2 (71b).

(71) a. Le papa du bébé, qui/lequel réclame son whisky à la fin de la journée, supporte le bruit assez difficilement.

b. Le papa du bébé, qui/lequel réclame son biberon à la fin de la journée, supporte le bruit assez difficilement.

Il faudra néanmoins contrôler le matériel pour que les constituants utilisés pour la désambiguïsation diffèrent le moins possible, tant au niveau de la fréquence que de la longueur ou du nombre de syllabes, mais aussi de la plausibilité : que celui qui signale un attachement N1 soit aussi plausible avec N1 et non plausible avec N2 que celui qui indique un attachement à N2 (plausible avec N2 et non plausible avec N1).

Chapitre 4 Discussion générale

L'objectif de ce travail de thèse est d'évaluer, en français, l'attachement des relatives ambiguës et le traitement anaphorique mis en place. J'ai proposé que les expressions référentielles, relativiseurs et antécédents, signalent par leur forme le niveau d'accessibilité de l'entité à laquelle ils font référence. Trois méthodes expérimentales ont été utilisées pour éprouver mes hypothèses : une étude de corpus, une étude de questionnaires et une expérience avec le paradigme d'auto-présentation segmentée.

Avant de débattre de la validation ou de l'invalidation des hypothèses et de l'apport de ce travail pour les autres théories explicatives sur le traitement des ambiguïtés syntaxiques et le traitement anaphorique, je vais récapituler l'ensemble des résultats obtenus jusque là.

1. 1. Récapitulatif des résultats obtenus

L'effet de la forme du relativiseur a été évalué à plusieurs reprises. L'étude de corpus montre un attachement différent quand la forme des relativiseurs signale un niveau d'accessibilité différent. Le relativiseur *lequel* et ses déclinaisons signalent un antécédent moins accessible que *qui*, parce qu'ils sont plus longs et marqués en genre et nombre. De façon générale, *lequel* est attaché plus fréquemment à N2 que *qui*. Cet effet est répliqué dans des tâches de compréhension, dans une étude de questionnaires (Questionnaire 1 et 4) et une expérience d'auto-présentation. Pour celle-ci, la mesure des temps de lecture révèle que les participants ont plus de difficultés à traiter une phrase avec *lequel* quand l'ambiguïté est levée vers N2. La comparaison entre *à qui* et *auquel*, dont le niveau d'accessibilité est similaire quand on contrôle l'animation de l'antécédent, ne révèle pas de différence d'attachement (corpus et Questionnaires 2 et 3).

L'étude de corpus évalue l'effet de la forme des antécédents et de leur niveau d'accessibilité sur l'attachement, alors que les Questionnaires et la tâche de lecture contrôlent forme et niveau. L'analyse à l'aide de modèles linéaires mixtes et les fréquences d'occurrence mettent en évidence que l'animation, la modification des antécédents, ainsi que la fonction syntaxique de N1 influencent non seulement l'attachement de façon générale, mais que l'attachement diffère selon la forme du relativiseur (entre *qui* et *lequel*, et non entre *à qui* et

auquel). Si l'on mesure une préférence globale pour attacher la relative à l'antécédent soit à N1 ou N2 quand il est plus accessible, c'est-à-dire animé ou modifié, soit à N1 quand il remplit la fonction sujet ou que la relative est longue, cette préférence interagit aussi avec la forme du relativiseur. L'étude des relatives sujet (*qui, lequel*) montre que *lequel*, qui signale un attachement à un antécédent moins accessible, est attaché plus souvent à l'antécédent moins accessible (N2 quand N1 est sujet ou remplit une autre fonction) et que *qui*, attaché de préférence à un antécédent très accessible, est en effet attaché plus fréquemment à un tel antécédent (N1 animé ou sujet de la principale).

1. 2. La validation des hypothèses opérationnelles et de l'hypothèse théorique

Le point de départ pour l'élaboration de la problématique est la proposition de Hemforth *et al.* (2000) : la spécificité de l'ambiguïté d'attachement des relatives par rapport aux autres ambiguïtés syntaxiques (relatives réduites et attachement de groupes prépositionnels) réside dans le fait que le constituant dont l'attachement est ambigu, *i. e.* le relativiseur, fait l'objet non seulement d'un attachement syntaxique, mais aussi de l'élaboration d'un lien anaphorique avec son antécédent.

Un relativiseur est une expression anaphorique et les facteurs régissant le traitement des anaphores et de la cohérence doivent nécessairement s'y appliquer. La forme d'une expression référentielle signale le degré d'accessibilité dans le discours de l'entité à laquelle elle fait référence (Givón, 1990, 1992). Les expressions référentielles peuvent aussi être placées sur des hiérarchies d'accessibilité (Ariel, 1990, 2001 ; Gundel *et al.*, 1993) : plus l'expression référentielle est courte et peu informative (p. ex. pronom), plus elle signale un référent accessible, alors qu'une expression référentielle longue (p. ex. GN avec modifieur) signale un référent peu accessible.

Dans une structure de la forme « N1 de N2 », N1 est considéré comme le référent le plus accessible, parce qu'il est la tête du GN et que N2 est le complément du nom. Un relativiseur, expression référentielle courte, fera ainsi préférentiellement référence à l'antécédent le plus accessible, N1. C'est pourquoi dans la littérature une préférence pour N1 est souvent observée, pour de nombreuses langues. Néanmoins, des expériences ont mis en évidence des variations translinguistiques (préférence d'attachement à N2) ou intra-linguistiques qui pourraient être expliquées par la notion d'accessibilité. Par exemple, la forme des antécédents affecte les préférences d'attachement (nom propre ou nom commun, Frenck-Mestre & Pynte,

2000b) ; un nom commun (généralement N1) est d'autant plus le site d'attachement choisi que N2 est un nom propre. La notion d'accessibilité apporte une explication intéressante : un nom propre signale un référent moins accessible qu'un nom commun. Comme le relativiseur est attaché préférentiellement à un antécédent accessible, il est attaché à N1 qui est un nom commun. La forme des antécédents signifierait ainsi l'accessibilité de leur référent, ce qui affecte les préférences d'attachement.

De même, la forme du relativiseur peut signaler l'accessibilité de l'entité référée et cela guide son attachement. Pour Ariel (1990, 2001), l'accessibilité varie selon l'atténuation, l'informativité et la rigidité. Le français dispose de deux systèmes de relativiseurs, les *qui* et les *lequel*, et je propose que, pour une même fonction syntaxique, l'emploi d'un relativiseur au détriment de l'autre peut signaler un niveau d'accessibilité différent, si leur forme les positionne à des rangs différents de la hiérarchie. *Lequel* étant moins atténué que *qui*, parce que plus long et plus informatif par son marquage en genre et nombre, il signale un antécédent moins accessible. L'ensemble des expérimentations valide cette hypothèse : il y a plus d'attachements à N2 avec *lequel* qu'avec *qui*. On observe une fréquence plus importante d'attachements à N2 dans les Questionnaires 1 et 4, mais aussi moins d'erreurs aux questions de compréhension sur des phrases avec *lequel* et dont l'attachement final est à N2 dans l'expérience d'auto-présentation segmentée. Les mesures des temps de lecture montrent l'absence de préférence d'interprétation avec *qui*, mais un attachement initial à N2 pour les relatives avec *lequel*, puisqu'une désambiguïsation vers N1 conduit à des TR plus longs qu'une désambiguïsation à N2. Dans l'ensemble, on peut considérer que la différence au niveau de la forme des relativiseurs *qui* et *lequel* sert à signaler pour le second un attachement à un antécédent moins accessible, N2.

La comparaison entre *à qui* et *auquel* (et *à laquelle* pour le Questionnaire 3) permet de contrôler le degré d'informativité avec le marquage de l'animation pour *à qui*, du genre et du nombre pour *auquel*, mais aussi le cas échéant le degré d'atténuation (seul *à laquelle* est plus long que *à qui* pour la fonction OI). Les résultats, tant pour les Questionnaires 2 et 3 que pour l'étude de corpus (relatives OI avec antécédent animé), ne révèlent pas de différence d'attachement pour *à qui* et *auquel*. Comme ils ont des degrés d'atténuation et d'informativité comparables, ils ont une position équivalente dans la hiérarchie d'accessibilité et signalent un attachement identique.

Les préférences d'attachement sont influencées par l'accessibilité signalée par la forme des antécédents. L'étude de corpus a mis en évidence que l'animation (animé, non-animé) et la

modification (présence ou absence d'un modifieur) des antécédents d'une part, et la fonction syntaxique de N1 (sujet, objet direct et indirect, obliques et sans fonction) de l'autre, affectent l'accessibilité signalée de l'antécédent. De façon générale, s'il y a une ambiguïté d'attachement, une relative est d'autant plus souvent attachée à l'un des référents qu'il est plus accessible. Si l'antécédent est animé ou modifié, il est très saillant et constitue le site d'attachement préféré. Quand N1 est sujet, fonction syntaxique où l'entité est dans le focus principal (Givón, 1990), il y a plus d'attachements à cet antécédent qu'à N2. La présence d'interactions entre ces facteurs et la forme du relativiseur est remarquable. D'abord, il n'y a pas d'interaction entre la forme du relativiseur et ces facteurs pour les relatives OI (*à qui, auquel*), suggérant que l'attachement est affecté de façon comparable pour les relatives avec *à qui* et *auquel*. Les interactions au niveau des relatives sujet, entre la forme du relativiseur (*qui, lequel*), l'animation et la fonction syntaxique de N1 sont intéressantes, puisqu'elles révèlent un effet différent de l'animation et de la fonction syntaxique selon le relativiseur. Une relative avec *qui* est sensible à la haute accessibilité de N1, quand il est sujet de la principale ou qu'il est animé. L'attachement à N1 est alors très fréquent. Quand N1 est moins accessible, non-animé ou qu'il remplit une autre fonction, la préférence pour attacher la relative à cet antécédent est moindre. Le relativiseur *lequel* semble moins sensible à la très haute accessibilité de N1. Même quand ce dernier est sujet ou animé, il y a plus d'attachements à N2 avec *lequel*, qui est attaché majoritairement à l'antécédent peu accessible.

Certaines de mes hypothèses sur l'attachement préférentiel à l'antécédent le plus accessible ne sont pas validées. D'après les études sur la cohérence et le caractère *given-new*, je m'attendais à ce qu'une relative longue, qui fournit plus d'informations, soit attachée de façon préférentielle à l'antécédent le moins accessible, N2, car il est relativement nouveau dans le discours : de nouvelles informations, fournies par la relative, sont plus susceptibles de le concerner. Les résultats montrent que plus la relative est longue, plus elle est attachée à N1. La modification des antécédents signale selon les théories de l'accessibilité (cf. Ariel, 1990, 2001) un référent moins accessible et les relatives dans l'ensemble y sont d'autant plus attachées. Si ce résultat ne valide pas l'hypothèse de la modification, il indique que celle-ci augmente le degré de saillance de l'antécédent.

Enfin, les hypothèses sur l'animation des antécédents et la fonction syntaxique de N1 sont validées. Un antécédent animé est proéminent dans le discours et constitue le site d'attachement préféré d'une relative par rapport à un antécédent non-animé. Quand N1 est

sujet de la principale, il est dans le focus principal et donc très proéminent. C'est pourquoi il est le plus souvent l'antécédent.

Dans l'ensemble, ce travail montre que la forme des relativiseurs, ainsi que celle des antécédents permettent de signaler le niveau d'accessibilité des entités du discours auxquelles ils font référence. Cela exerce une influence sur les préférences d'attachement, un relativiseur situé plus bas dans la hiérarchie d'accessibilité signale un attachement à un antécédent moins accessible. Un relativiseur plus réduit signale un antécédent hautement accessible, ce qui confirme ce que la *centering theory* (Brennan *et al.*, 1987; Gordon *et al.*, 1993 ; Grosz *et al.*, 1995) a montré de façon extensive : une anaphore pronominale, dont la forme est réduite, est préférée à une anaphore moins réduite quand l'entité du discours est très accessible.

1. 2. 1. La synonymie des relativiseurs en français

Dans cet exposé, les relativiseurs comparés, *qui* et *lequel* d'un côté, à *qui* et *auquel* de l'autre, sont synonymes. Rien dans la grammaire normative n'indique une différence de sens, si ce n'est l'emploi plus formel de *lequel* — et dans une moindre mesure *auquel* — quand l'antécédent est animé (Riegel *et al.*, 2004). La seule différence entre ces deux types de relativiseurs concerne le marquage de l'animation : pour les fonctions autres que sujet, dès qu'il y a une préposition, les relativiseurs comme *à qui* ne peuvent être utilisés que pour les entités animées. Les relativiseurs comme *auquel* peuvent être utilisés indépendamment de l'animation.

Si ce marquage de l'animation, et non celui de l'accessibilité, revêt un sens pour le parseur, il devrait affecter les préférences d'attachement quand les relativiseurs se distinguent à ce niveau : un antécédent animé est très accessible et un relativiseur qui signale un antécédent animé serait privilégié. La différence d'animation entre *à qui* et *auquel* affecterait alors les préférences d'interprétation. *A contrario*, *qui* et *lequel* qui ne marquent pas l'animation devraient être interprétés de façon comparable. Les résultats obtenus dans ce travail (Questionnaires et corpus) montrent un patron inverse, avec une différence d'attachement pour *qui* et *lequel* et aucune différence pour *à qui* et *auquel* dont l'antécédent est toujours animé. On peut donc considérer que les deux formes de relativiseurs sont bien synonymes, que ce n'est pas cet aspect qui contribue aux résultats mesurés, mais plutôt la notion d'accessibilité signalée.

1. 3. Apports pour les théories

1. 3. 1. L'importance de la notion d'accessibilité dans le traitement des ambiguïtés d'attachement

La différence de traitement entre les relativiseurs *qui* et *lequel*, mesurée dans les trois études menées ici, montre que la présence d'un relativiseur comme *lequel* entraîne plus d'attachements à l'antécédent le moins accessible qu'un relativiseur comme *qui*. L'étude de corpus montre une fréquence plus importante des attachements à N2 avec *lequel* qu'avec *qui*. Les Questionnaires 1 et 4 mettent en évidence un résultat comparable, avec des procédures interindividuelle et intra-individuelle. L'expérience avec le paradigme d'auto-présentation montre, pour les relatives avec *lequel*, des temps de lecture plus importants après la levée de l'ambiguïté à N1, traduisant que les participants ont au préalable interprété *lequel* comme signalant un attachement à N2. Pour les relatives avec *qui*, les questionnaires montrent une préférence pour N1 (l'attachement reste ambigu) et l'expérience en auto-présentation ne montre aucun effet de l'attachement ; il n'y a pas de préférence d'attachement ni de difficulté à traiter l'un ou l'autre des attachements signalés. Ce résultat est confirmé par l'absence de différence au niveau des questions de compréhension, pour les relatives avec *qui*. Ces résultats, dans leur ensemble, suggèrent que *lequel* est situé plus bas que *qui* dans la hiérarchie d'accessibilité (Ariel, 1990, 2001) et qu'il signale un niveau d'accessibilité et un attachement à l'antécédent le moins accessible. Comme *lequel* est à la fois moins atténué et plus informatif que *qui*, ces deux facteurs, atténuation et informativité, pourraient contribuer à la différence d'accessibilité signalée.

La comparaison entre *à qui* et *auquel/à laquelle* permet d'évaluer l'effet de ces facteurs en contrôlant leur niveau d'informativité. *À qui* est plus informatif que *auquel* parce qu'il marque le caractère animé de son antécédent. Ce dernier, par contre, est plus informatif car il renseigne le genre et le nombre de son antécédent. Comme l'animation et le marquage en genre et en nombre sont des caractéristiques importantes pour le traitement de matériel linguistique (J. E. Arnold, Eisenband *et al.*, 2000 ; Bock & Levelt, 1994 ; Branigan *et al.*, 2008 ; G. R. Kuperberg *et al.*, 2006 ; Van Berkum *et al.*, 1999b), on a considéré que le niveau d'informativité de ces deux relativiseurs est contrôlé. L'atténuation, en revanche, peut être contrôlée entre *à qui* et *auquel* (deux syllabes) ou varier entre *à qui* et *à laquelle*. Le Questionnaire 3 (*à qui*, *auquel* et *à laquelle*) montre que la différence d'atténuation n'entraîne pas d'attachement différent. Ainsi la différence d'informativité entre *qui* et *lequel* expliquerait

les résultats alors obtenus. Les relativiseurs fonctionnent bien comme des marqueurs d'accessibilité et signalent l'attachement à réaliser.

La plupart des théories sur l'accessibilité (*accessibility theory* et *givenness hierarchy*) et sur le traitement anaphorique en général (*centering theory* et ses développements ultérieurs et l'*informational load hypothesis*) considèrent les expressions référentielles d'un énoncé à l'autre. Les résultats de ces études montrent que le traitement anaphorique dans un même énoncé (propositions principale et relative) est lui aussi guidé par la forme des expressions anaphoriques, augmentant le champ d'application de ces théories.

L'étude de questionnaires et celle avec le paradigme d'auto-présentation contribuent aussi à une première évaluation de la validité en compréhension des théories de l'accessibilité. Ces dernières expliquent la forme des expressions référentielles dans les productions langagières, contribuant à la cohérence. En accord avec le principe général de la *tuning hypothesis*, selon quoi les préférences en compréhension sont guidées par l'exposition du parseur, en production, à certaines structures et certaines interprétations, les résultats obtenus ici montrent que l'accessibilité signalée par les expressions référentielles, relativiseurs et antécédents, contraint l'interprétation en compréhension et qu'elle signale l'attachement tel que l'envisage le locuteur. D'un point de vue de la pertinence et des principes de Grice (voir Sperber & Wilson, 1989), cela confirme que le locuteur utilise ce qui est à sa disposition, les deux systèmes de relativiseurs, pour indiquer à son allocutaire l'attachement à réaliser. On a vu ici que la forme du relativiseur sert à guider l'attachement dans le cas de structures ambiguës, mais comme l'attachement dépend de l'accessibilité des entités du discours, rien n'exclut que la forme du relativiseur soit aussi contrainte par l'accessibilité de l'antécédent quand l'attachement est non ambigu. Ainsi, *qui* est le relativiseur privilégié pour faire référence à un antécédent très accessible, sujet de la principale ou entité animée, tandis que *lequel* est utilisé pour un antécédent moins accessible.

L'étude de corpus a permis de tester quelques aspects utilisés dans l'élaboration des hiérarchies d'accessibilité, comme la présence d'un adjectif et par conséquent la longueur de l'expression référentielle (cf. Ariel, 1990). D'autres aspects susceptibles d'influencer l'accessibilité des antécédents sont étudiés : l'animation et la fonction syntaxique de N1. La longueur de la proposition relative peut aussi contribuer à signaler le niveau d'accessibilité de l'antécédent. Les résultats montrent que l'animation et la fonction syntaxique contribuent à signaler le degré d'accessibilité des antécédents, guidant l'attachement de la proposition relative, et qu'en plus ces facteurs interagissent avec la forme des relativiseurs, pour les

relatives sujet. Lorsque le niveau d'accessibilité est bas (antécédent non animé, fonction syntaxique basse), il est le site d'attachement préféré pour un relativiseur signalant un faible niveau d'accessibilité, *lequel* par comparaison avec *qui*. Même quand l'accessibilité signalée par N1 est importante (animé, sujet de la principale), *lequel* reste attaché plus souvent à l'antécédent moins accessible, alors que *qui* est dans ce cas attaché de façon très importante à N1.

La longueur de la relative n'a d'effet que pour les relatives OI, et sa direction est opposée à ce qui était prédit selon l'accessibilité : plus la relative est longue, plus elle est attachée à l'antécédent plus accessible.

La modification des antécédents, par exemple avec un adjectif, conduit à des résultats opposés à ceux qu'on attendrait selon les théories de l'accessibilité. Ajouter de l'information à un antécédent en fait le site d'attachement préféré et il n'y a pas de différence entre les relatives avec *qui* et *lequel*, alors que selon les hiérarchies d'accessibilité on se serait attendu à moins d'attachements à un antécédent modifié pour *qui*. La modification n'est pas un facteur pertinent pour le degré d'accessibilité signalé des antécédents et sert à signaler la spécificité de l'expression référentielle. Par contre, la modification augmente le niveau de saillance de cette référence par rapport à un GN sans modifieur, à cause de la taille phonologique plus importante. Peut-être faut-il distinguer saillance et accessibilité ? La saillance semble liée à la prosodie ou tout du moins à la longueur des constituants, la taille du GN, avec ou sans modifieur, ou celle de la relative. Quand le GN ou la relative sont longs, ces expressions ont une saillance plus importante qu'un GN non modifié ou une relative courte. Ce degré de saillance plus élevé attire l'attachement. Notons que ces deux facteurs, la modification et la longueur de la relative, n'interagissent pas avec la forme des relativiseurs pour la fonction sujet, alors que l'effet de l'animation et celui de la fonction syntaxique de N1 sont modulés selon la forme du relativiseur, *qui* ou *lequel*. On pourrait envisager que la longueur de la relative et la modification des antécédents sont des facteurs signalant la saillance et que l'animation et la fonction syntaxique sont des facteurs signalant l'accessibilité.

Ainsi, l'accessibilité signalée des antécédents et des relativiseurs guide l'attachement de propositions relatives lorsqu'il est ambigu. Si un relativiseur signale un antécédent peu accessible et que l'antécédent est lui aussi peu accessible, alors ce dernier est le site d'attachement préféré. En revanche, si le relativiseur signale un antécédent hautement accessible et que la forme de l'antécédent signale un niveau d'accessibilité élevé, qu'il soit animé ou que la fonction syntaxique soit élevée, ce dernier est le site d'attachement préféré.

1. 3. 2. L'accessibilité comme facteur explicatif des variations translinguistiques et intra-linguistiques

Le relativiseur fonctionne comme une autre anaphore : sa forme signale le degré d'accessibilité de l'entité à laquelle il fait référence, son antécédent. L'effet de l'accessibilité signalée par les expressions référentielles que sont les relativiseurs et les antécédents pourrait apporter une explication nouvelle et intéressante pour les variations translinguistiques et intra-linguistiques relevées dans la littérature.

Tout d'abord, les études en anglais qui mettent en évidence une préférence pour attacher une relative ambiguë à N2 présentent en général le relativiseur *who* (Carreiras & Clifton, 1999 ; Mitchell *et al.*, 1990) alors que l'utilisation de *that* ne révèle aucune préférence d'attachement (Traxler *et al.*, 1998). Notons aussi que Cuetos et Mitchell (1988) mesurent une préférence pour N2 avec *who* et une préférence pour N1 avec *that*. Ces résultats ressemblent à ceux que j'ai obtenus et suggèrent que le degré d'accessibilité des relativiseurs *that* et *who* sert à signaler un attachement différent. Néanmoins, la possibilité d'un biais de l'animation avec *that* (Cuetos & Mitchell, 1988) ne peut permettre de conclure de façon certaine.

On a vu que Baccino *et al.* (2000) ont mesuré une préférence pour N2 avec des phrases comme (14a) où N1 et N2 sont des noms communs et que Frenck-Mestre et Pynte (2000a, 2000b) observent une préférence pour N1 avec des phrases comme (14a). La différence de fréquence entre les noms dans les deux types de phrases est invoquée pour expliquer la différence alors obtenue (Pynte & Colonna, 2000).

(14) a. *Gianni osserva il ragioniere de Caterina che sembra più pensieroso(a) del normale.*

L'accessibilité des antécédents propose une explication alternative à propos de la forte préférence d'attachement à N1 pour (14a) : un nom propre signale un antécédent moins accessible qu'un nom commun. Comme une relative est attachée de façon préférentielle à l'antécédent le plus accessible, c'est un attachement à N1 qui est alors observé.

Les différences translinguistiques sont expliquées par Frazier et Fodor (1997) par le fait que des langues qui disposent du pronom relatif utilisent la proéminence des entités du discours pour établir l'attachement. Néanmoins, leur explication ne vaut pas pour les données en anglais : *who* est un pronom relatif et on observe avec ce relativiseur un attachement préférentiel à N2. Par contre, si on considère qu'une langue qui possède plusieurs systèmes de relativiseurs attache de façon préférentielle la proposition relative à l'antécédent le plus

accessible avec un relativiseur qui signale un tel niveau d'accessibilité, il sera alors plus atténué et moins informatif (*that* en anglais, *qui* en français) que l'autre système de relativiseur (*who* en anglais, *lequel* en français). Que les expériences en compréhension en anglais utilisent généralement *who* au détriment de *that* peut ainsi expliquer la préférence pour N2 qui est mise en évidence. En français, le relativiseur *qui* est toujours utilisé (Colonna, 2001 ; Frenck-Mestre & Pynte, 2000b ; Pynte & Colonna, 2000 ; Zagar *et al.*, 1997), à l'exception du travail présenté ici. Quand *qui* est utilisé, on observe un attachement à N1, l'antécédent le plus accessible et si *lequel* est utilisé, il y a une préférence pour N2. Les langues qui ne disposent que d'un système de relativiseur vont attacher la relative à l'antécédent le plus accessible, N1. C'est le cas par exemple en allemand (Hemforth *et al.*, 2000).

1. 3. 3. La fréquence

On pourrait se demander si la différence de fréquence d'occurrence entre *qui* et *lequel*, plus importante qu'entre *à qui* et *auquel*, contribue à la différence des préférences d'attachement. La fréquence est un corrélat des hiérarchies d'accessibilité : les expressions référentielles situées le plus haut dans la hiérarchie, que ce soit une anaphore zéro ou un pronom, sont plus fréquentes que des expressions référentielles plus basses dans la hiérarchie, comme un nom propre. En ce qui concerne les relativiseurs, *lequel* est plus bas dans la hiérarchie et aussi moins fréquent que *qui*. Les relativiseurs pour la fonction OI, *à qui* et *auquel*, présentent une différence de fréquence moindre et inversée : *auquel* est plus fréquent que *à qui*, même lorsqu'on ne prend en compte que les cas où l'antécédent est animé (cf. marquage de l'animation de *à qui*). Dans les Questionnaires 2 et 3, où l'animation est contrôlée, et dans le modèle élaboré pour les relatives OI avec un antécédent animé, on ne mesure pas de différence d'attachement selon les relativiseurs. Or, quand on compare les relatives avec *à qui* et *auquel* (Questionnaire 2), seule la fréquence varie et non l'atténuation ou l'informativité. La différence de fréquence ne suffit pas à entraîner un attachement différent. D'après ces résultats et ceux obtenus pour *qui* et *lequel*, on peut favoriser l'idée que c'est la différence d'informativité entre *qui* et *lequel* qui signale un attachement différent. Quand l'atténuation et l'informativité sont contrôlées et que la fréquence varie, il n'y a pas de différence (Questionnaire 2, modèle linéaire mixte pour les relatives OI avec antécédent animé). Quand seule l'informativité est contrôlée et que l'atténuation et la fréquence sont plus importantes (*à laquelle* vs. *à qui*, Questionnaire 3), il n'y a pas non plus de différence d'attachement. Par contre, quand l'informativité, l'atténuation et la fréquence diffèrent (*qui*, *lequel*, corpus,

Questionnaires 1 et 4 et expérience d'auto-présentation), cela conduit à des attachements différents. L'informativité apparaît comme le marqueur d'accessibilité le plus important. Il y a bien dans le corpus une différence selon que la relative est introduite par *à qui* ou *auquel* (tout antécédent), mais cette différence semble plus liée à la fréquence plus importante de constructions « N1 de N2 » où N1 est non-animé et N2 animé avec *à qui* par comparaison avec *auquel*. Cette différence n'est pas liée à l'accessibilité.

1. 3. 4. L'accessibilité comme facteur explicatif des variations translinguistiques et intra-linguistiques

Selon la *tuning hypothesis* (Mitchell *et al.*, 1995), un autre type de fréquence intervient dans les préférences d'attachement : ces dernières, dans un processus de compréhension, reflètent les fréquences d'attachement en production. Seules les fréquences grossières sont prises en considération. Ici, il s'agirait des fréquences d'attachement pour une proposition relative précédée d'une structure « N1 de N2 », indépendamment de la forme du relativiseur ou des aspects syntaxiques ou sémantiques des antécédents.

D'après l'étude de corpus, aussi bien pour l'ensemble des relatives non restrictives que pour les relatives sujet ou OI, on mesure une fréquence plus importante des attachements à N2 (56 % sur toutes les occurrences analysées). Si la fréquence d'attachement contraint l'interprétation des propositions relatives ambiguës, un attachement à N2 est alors attendu dans les études en compréhension, pour les Questionnaires et pour l'expérience avec le paradigme d'auto-présentation segmentée. Cette préférence pour N2 n'est pas confirmée. On obtient une forte préférence pour N1 sur l'ensemble des Questionnaires, ce qui ne confirme pas les prédictions de la *tuning hypothesis*. Les erreurs aux questions de compréhension pour l'expérience d'auto-présentation révèlent un effet de l'Attachement avec plus d'erreurs commises pour des phrases où le contenu sémantique signale un attachement à N1. L'analyse des effets simples montre que ce sont les relatives avec *lequel* qui contribuent à ce résultat, puisque les relatives avec *qui* sont bien traitées indépendamment de l'attachement, alors que celles avec *lequel* sont plus difficiles quand l'attachement est à N1, traduisant une préférence pour N2. Les mesures des temps de lecture sont cohérentes avec les résultats pour les questions de compréhension. On ne peut pas dire que les résultats généraux soient supportés par une conception à gros grain de la *tuning hypothesis*.

La prise en compte d'informations lexicales comme l'animation (Desmet, Brysbaert *et al.*, 2002 ; Desmet *et al.*, 2006), qui suppose un modèle à grain d'analyse fin ou mixte (Desmet,

Brysbaert *et al.*, 2002), et le calcul des fréquences d'attachement pour des structures proches de celles présentées expérimentalement, permettent de trouver un accord entre les fréquences d'attachement dans les corpus et les préférences d'attachement en compréhension. Desmet *et al.* ont montré que si on prend en compte les fréquences d'attachement avec une configuration d'animation comparable avec le matériel utilisé dans les tâches de compréhension (p. ex. N1 et N2 animés), les attachements réalisés en compréhension sont bien le miroir des fréquences dans le corpus, confirmant la préférence pour N1. C'est le cas ici : quand N1 et N2 sont animés, N1 est le site d'attachement le plus fréquent dans le corpus et les questionnaires mesurent la préférence de N1. L'effet du relativiseur, congruent en production et en compréhension puisqu'il y a plus d'attachements à N2 quand le relativiseur signale un antécédent moins accessible, suggère que les fréquences considérées pour le modèle de *tuning* mixte doivent prendre en compte la forme des relativiseurs comme critère, en plus de l'animation. D'autres facteurs sont aussi envisageables (Gibson & Schütze, 1999), comme la fonction syntaxique ou la modification. En effet, cette étude de corpus a montré qu'il y a plus d'attachement à un antécédent modifié et des expériences en compréhension (Colonna, 2001 ; Gilboy *et al.*, 1995 ; Pynte & Colonna, 2001) ont mis en évidence une préférence pour attacher la relative à l'antécédent modifié. Il s'agit d'un autre critère pertinent.

1. 3. 5. L'intégration des informations syntaxiques, sémantiques ou discursives

La théorie du *garden path* postule l'existence de deux stratégies d'interprétation. La première est purement syntaxique et la seconde, qui intègre des informations sémantiques ou autres, n'intervient que dans un second temps, en cas d'échec de la première stratégie, dans un processus de ré-analyse (Clifton, 1993 ; F. Ferreira & Clifton, 1986 ; F. Ferreira & Henderson, 1990 ; Frazier & Rayner, 1982, 1987 ; Rayner *et al.*, 1983). L'expérience avec le paradigme d'auto-présentation est la plus appropriée pour évaluer les prédictions de cette théorie puisque la mesure des temps de lecture mot à mot permet d'estimer une éventuelle ré-analyse si l'interprétation initiale est différente de celle signalée à partir de la levée de l'ambiguïté. Des temps de lecture plus importants sont attendus.

Dans le cas des structures étudiées ici, l'analyse initiale est réalisée selon le principe de *late closure* et conduit à un attachement à N2 dès l'apparition du relativiseur. Par ailleurs, la forme du relativiseur n'affecte pas cette interprétation initiale. Quand l'ambiguïté d'attachement est levée, à partir de l'objet de la proposition relative et des régions *spillover*, et

que N2 est le site d'attachement, ceci est congruent avec l'analyse initiale selon le *late closure* et le participant peut continuer à traiter la phrase sans en changer l'interprétation. Par contre, si l'analyse initiale est un attachement à N2 et que l'attachement signalé est à N1, il est nécessaire de ré-analyser la structure choisie initialement. Cela entraîne des temps de traitement plus importants à partir du moment où l'ambiguïté est levée.

Qu'on ait obtenu un effet immédiat de la forme du relativiseur invalide les prédictions de la théorie du *garden path*, d'autant plus que le relativiseur a un effet précoce sur l'interprétation de la phrase : une phrase contenant *qui* est traitée plus rapidement quand l'attachement est à N1 et l'inverse est mesuré pour les phrases contenant *lequel*. La forme du relativiseur est un indice de l'attachement à réaliser.

Des théories sur le traitement syntaxique qui ont été présentées dans l'introduction, seule la *construal hypothesis* mentionne l'effet éventuel de la forme du relativiseur sur l'interprétation de ces structures syntaxiquement ambiguës (Frazier & Clifton, 1997). Pour les langues possédant un vrai pronom relatif, la saillance des antécédents peut guider l'interprétation, comme c'est le cas pour l'interprétation des pronoms (cf. Hemforth *et al.*, 2000). En français, les relativiseurs *qui* et *lequel* sont considérés comme des pronoms relatifs (Jones, 1996 ; Riegel *et al.*, 2004) et c'est pourquoi on obtient une préférence d'attachement de la relative à l'antécédent le plus proéminent, N1 dans les études de compréhension (Colonna, 2001 ; Colonna & Pynte, 2002 ; Zagar *et al.*, 1997). Ici, notons que les résultats de l'étude de corpus montrent une préférence générale pour N2 et que l'expérience d'auto-présentation confirme cette préférence.

De plus, l'effet de la forme du relativiseur n'est pas prévu ni expliqué par la *construal hypothesis*. Cette théorie n'est qu'en partie vérifiée : elle prévoit une préférence pour N1, ce qu'on obtient dans les Questionnaires, mais pas une différence entre *qui* et *lequel*.

Les modèles interactifs, c'est-à-dire les modèles de satisfaction de contraintes et de compétition et d'intégration, le modèle *unrestricted race* et la *recency theory* ne formulent pas d'hypothèse spécifique à propos de l'effet de la forme du relativiseur sur les préférences d'attachement, mais ils n'excluent pas non plus que les informations, notamment d'accessibilité, procurées par les relativiseurs interviennent dans le traitement de ces ambiguïtés d'attachement. Les travaux menés ici montrent que des informations de niveau discursif, le niveau d'accessibilité, sont prises en compte dans l'interprétation de ces phrases dès le niveau du relativiseur. Ces modèles pourraient ainsi intégrer cet aspect.

1. 3. 6. Les modèles de la compréhension, de la production

Les théories de l'accessibilité et du traitement anaphorique sont souvent élaborées sur la base d'exemples construits ou extraits de productions écrites ou orales (Ariel, 1990 ; Givón, 1990 ; Gundel *et al.*, 1993). Leur évaluation et leur validation sont réalisées avec des études en production, des études de corpus (Dahl & Fraurud, 1996 ; Hobbs, 1978 ; Poesio *et al.*, 2004 ; Toole, 1996), avec le paradigme de *visual world* (J. E. Arnold, Eisenband *et al.*, 2000), avec des tâches de production écrite (Fletcher, 1984 ; Stevenson *et al.*, 1994) ou orale (Brennan, 1995). Des études en compréhension sont aussi réalisées (Almor, 1999 ; Fletcher, 1984), en particulier dans l'évaluation de la *centering theory* (Gordon & Chan, 1995 ; Gordon *et al.*, 1993).

Les théories sur le traitement des ambiguïtés syntaxiques se sont principalement attachées à expliquer le versant relatif à la compréhension du langage (Bever, 1970 ; Frazier, 1987 ; Frazier & Fodor, 1978). Des données en production sont parfois analysées afin d'évaluer si les facteurs mis en évidence en compréhension ont aussi un effet sur la production, et inversement. C'est le cas par exemple avec la *tuning hypothesis* et ses mises à l'épreuve, où l'on évalue si les fréquences d'attachement en production expliquent les préférences d'attachement en compréhension (voir Desmet, Brysbaert *et al.*, 2002 ; Desmet *et al.*, 2006). Des données en production, par exemple une tâche de complétion, sont parfois utilisées pour évaluer les prédictions du modèle basé sur les contraintes (Thornton & MacDonald, 2003), du modèle de compétition et d'intégration (Hare *et al.*, 2007) ou du modèle concurrent (*meaning through syntax*, McKoon & Ratcliff, 2003), avant de les tester en compréhension.

Certains chercheurs considèrent que des mécanismes qui sous-tendent la compréhension et la production sont communs (Thornton & MacDonald, 2003), par l'accord entre les données en production et celles en compréhension (Branigan, Pickering, & Cleland, 1999 ; Pearlmutter, Garnsey, & Bock, 1999). D'autres proposent que des facteurs linguistiques s'appliquent de façon différente en compréhension et en production (Gibson & Schütze, 1999), car les données acquises ne sont pas toujours comparables et que des contraintes spécifiques peuvent s'appliquer à une modalité seulement. Des contraintes imposées par les limites cognitives de la mémoire de travail (Gibson, 1998, 2000) affectent la production de phrases et leur compréhension, favorisant un attachement local (Frazier, 1987, 1995 ; Gibson, Pearlmutter *et al.*, 1996 ; Gibson *et al.*, 1999 ; Gibson & Schütze, 1999). Gibson et Schütze (1999) proposent qu'un attachement à N1 est plus difficile à produire qu'un attachement à N2 à cause du déclin des informations plus anciennes dans la mémoire de travail. C'est pourquoi

on mesure plus d'attachements à N2 dans cette étude de corpus et dans d'autres (Desmet, Brysbaert *et al.*, 2002 ; Desmet *et al.*, 2006).

En revanche, des stratégies rendent possible un attachement distant en compréhension (Gibson & Schütze, 1999) : la proximité du prédicat, stratégie proposée dans la *recency theory* (Gibson, Pearlmutter *et al.*, 1996 ; Gibson *et al.*, 1999), ou bien l'élaboration d'un lien anaphorique (Hemforth *et al.*, 2000) ne s'appliqueraient qu'à la compréhension ou en tout cas n'affectent pas la production. Ceci expliquerait par exemple la différence qu'on observe pour les ambiguïtés d'attachement des relatives entre les données de corpus et celles qui sont acquises en compréhension. Desmet *et al.* (2002) relèvent cependant dans une étude de corpus un attachement compatible avec la proximité du prédicat (attachement à N1 pour une proposition relative précédée d'une structure « N1 de N2 » où les deux antécédents sont animés). Ces auteurs suggèrent aussi que ce n'est peut-être pas l'application de ces stratégies qui contribue à la différence entre production et compréhension, mais plutôt les contraintes liées à l'étude de corpus, telles qu'une taille réduite et la modalité écrite (pour une discussion voir Biber, 1986), et l'élaboration du matériel expérimental, avec le contrôle de la fréquence ou de la longueur des constituants.

Si des contraintes diffèrent entre production et compréhension, il faut tenir compte du fait que ces deux aspects impliquent deux perspectives, celle du locuteur et celle de l'allocutaire (Givón, 1992). Le locuteur a son propre modèle interne du discours et il doit aussi envisager le modèle interne de l'allocutaire, afin d'employer des expressions référentielles cohérentes. Qu'on ait observé un effet de la forme du relativiseur à la fois en compréhension et en production suggérerait que le locuteur est coopératif. Il utilise plus souvent un relativiseur signalant un faible niveau d'accessibilité quand l'attachement est à un antécédent moins accessible. Les données présentées ici ne permettent guère d'estimer de façon certaine si le locuteur est ou non coopératif, mais il serait intéressant d'étudier cet aspect.

Cependant, la plupart des expérimentations évaluent les contraintes mises en jeu du point de vue de l'allocutaire, que ce soit à propos du traitement syntaxique (Frazier & Fodor, 1978 ; Gibson, 1998, 2000) ou du traitement anaphorique (Ariel, 1990 ; Reboul, 1989). Certains travaux questionnent ces deux perspectives en s'intéressant à celle du locuteur (J. E. Arnold, Wasow, Asudeh, & Alrenga, 2004 ; J. E. Arnold, Wasow *et al.*, 2000). Arnold *et al.* (2004, 2000) étudie par exemple la production de phrases et comment des critères tels que la nouveauté de l'information et la complexité de l'expression référentielle influencent l'ordre

des constituants de la phrase, dans une étude de corpus et lors d'une tâche de production élicitée.

Il serait intéressant d'évaluer les hypothèses formulées dans le cadre de cette thèse à l'aide d'une expérience d'élicitation, comme celle de Brennan (1995) ou celle d'Arnold *et al.* (2000), bien qu'il semble difficile de concevoir une tâche où les participants produiraient une phrase contenant une proposition relative et les deux formes de relativiseurs possibles.

1. 4. La méthodologie expérimentale

Dans cette recherche, plusieurs techniques ont été utilisées : des tâches en compréhension, avec les questionnaires et le paradigme d'auto-présentation segmentée et une étude en production avec le corpus. Ces techniques présentent des avantages et des inconvénients.

J'ai voulu évaluer mes hypothèses quant à l'effet de l'accessibilité des expressions référentielles sur le traitement des ambiguïtés syntaxiques en production, avec une étude de corpus, afin de tester si les contraintes valables en compréhension s'appliquent aussi à la production. Cette étude de corpus présente l'intérêt premier de procurer une vision sur la production de phrases contenant une ambiguïté d'attachement, mais aussi sur la forme du relativiseur utilisé. À cette étude de corpus sont ajoutées des études en compréhension, avec des questionnaires et une tâche avec le paradigme d'auto-présentation segmentée. D'aucuns considèrent qu'études de corpus et expérimentales se complètent mutuellement, mais qu'elles ont des avantages et des inconvénients propres.

Si on veut réaliser une étude en production qui soit écologique, il serait intéressant de s'inspirer de la méthodologie utilisée par Brennan (1995), où des productions orales spontanées sont élicitées. Cependant, la comparaison entre *qui* et *lequel* pourrait être difficile, puisque *lequel* est d'un registre plus formel, moins susceptible d'être produit à l'oral qu'à l'écrit. Une tâche de production orale permettrait d'aborder la question de la prosodie (cf. Schafer *et al.*, 1996) et d'estimer si les contours prosodiques sont différents selon la forme du relativiseur et l'attachement, à N1 ou à N2.

Les résultats obtenus dans la comparaison entre *qui* et *lequel* (corpus) sont globalement conformes à ceux d'Ariel (1999) qui compare les pronoms résomptifs et les anaphores zéros. Il serait intéressant de réaliser des études translinguistiques sur d'autres langues qui disposent de plusieurs formes de relativiseurs, ce qui permettrait d'évaluer plus avant les hypothèses sur l'effet de la forme du relativiseur.

Parmi les formes de relativiseurs, ce sont deux types de pronoms relatifs qui ont été étudiés ici : *qui* et *à qui* d'un côté, *lequel* et *auquel* de l'autre. En anglais par exemple, on peut utiliser un pronom relatif, ce qui est le cas dans la majorité des expériences sur les ambiguïtés d'attachement qui utilisent *who* (Carreiras & Clifton, 1993, 1999 ; Cuetos & Mitchell, 1988 ; Mitchell *et al.*, 1990), mais il est aussi possible d'utiliser un complémenteur (*that*, cf. Traxler *et al.*, 1998) ou bien d'omettre ce complémenteur. L'emploi du complémenteur, et le cas échéant de l'auxiliaire, est parfois optionnel (Temperley, 2003) : il pourrait dans ce cas lever l'ambiguïté d'interprétation, comme dans des phrases de type *garden path* (72).

(72) a. *The horse raced past the barn fell.* (*Le cheval qui a été conduit devant la grange est tombé*, tiré de Frazier, 1985)

b. *The horse that was raced past the barn fell.* (traduction identique)

Cependant, des études en production montrent que le complémenteur n'est pas utilisé plus fréquemment quand il permet d'éviter des cas d'ambiguïtés (V. S. Ferreira & Dell, 2000), mais plutôt quand le contenu ultérieur de la phrase est complexe à produire ou moins accessible (V. S. Ferreira, 2003, à paraître ; V. S. Ferreira & Dell, 2000).

On pourrait proposer que la forme de l'expression référentielle en anglais sert à signaler le degré d'accessibilité et à guider l'attachement des propositions relatives : l'omission du relativiseur s'apparente à une anaphore zéro, signalant un niveau d'accessibilité très élevé (Ariel, 2001). L'utilisation de *that* au lieu de son omission sert à signaler un antécédent moins accessible, mais plus accessible cependant que *who*, qui serait situé au plus bas dans la hiérarchie, car il est plus informatif (il marque l'animation de l'antécédent et la fonction relativisée). Des données de corpus semblent conforter l'hypothèse d'un effet de l'accessibilité de l'antécédent, pour des propositions relatives non ambiguës. Ainsi, Jaeger et Wasow (à paraître) relèvent dans une étude de corpus sur les propositions relatives autres que sujet, où le relativiseur peut être omis, que le complémenteur *that* est utilisé plus fréquemment quand l'antécédent est nouveau dans le discours, mais aussi lorsque cet antécédent a un degré de définitude plus faible, p. ex. un GN avec un article indéfini. Certains suggèrent que la production de *that* au lieu de son omission permet au locuteur de gagner du temps et de préparer la suite de la proposition relative (Race & MacDonald, 2003). La comparaison de la distribution en anglais des pronoms relatifs (mots *wh-*) et du complémenteur *that*, mais aussi l'observation des effets de la suppression dans cette étude (Jaeger & Wasow, à paraître) montrent pour des propositions relatives restrictives un biais de l'animation. En effet, l'antécédent est plus souvent non-animé qu'animé avec *that*. L'effet de la fonction syntaxique

relativisée est plus difficile à interpréter : il est tout d'abord exceptionnel d'omettre le relativiseur quand la fonction sujet est relativisée et s'il s'agit de la fonction génitif, c'est un pronom relatif qui est alors obligatoire. Les résultats montrent que le complémenteur *that* est employé plus fréquemment qu'un pronom relatif pour des relatives sujet ou objet.

Il serait intéressant de réaliser une étude en compréhension qui étudie l'effet de la suppression du relativiseur sur les préférences d'attachement. Afin d'éviter le biais d'animation qui est parfois avancé (Bayley, 1999 ; Cuetos & Mitchell, 1988 ; Guy & Bayley, 1995), les antécédents N1 et N2 doivent être non-animés, comme dans (73a-d). Les phrases sont désambiguïsées vers N1, (73a) et (73c), ou vers N2, (73b) et (73d), le relativiseur étant soit présent, (73a-b), soit omis, (73c-d).

(73) a. *Peter is watching the film of the ceremony that I made yesterday.* (*Peter regarde le film de la cérémonie que j'ai fait hier*)

b. *Peter is watching the film of the ceremony that I organised yesterday.* (*Peter regarde le film de la cérémonie que j'ai organisée hier*)

c. *Peter is watching the film of the ceremony I made yesterday.* (*Peter regarde le film de la cérémonie que j'ai fait hier*)

d. *Peter is watching the film of the ceremony I organised yesterday.* (*Peter regarde le film de la cérémonie que j'ai organisée hier*)

Si la présence du complémenteur *that* sert à signaler un antécédent moins accessible que son absence, on devrait alors observer une préférence pour N2 dans le premier cas et pour N1 dans le second. Ainsi, la comparaison des temps de lecture entre les phrases (73a) et (73b) devraient montrer des temps de traitement plus importants quand l'attachement est désambiguïsé vers N1, alors que le contraire serait attendu entre (73c) et (73d).

CONCLUSIONS

Cette recherche présente une perspective innovante puisqu'elle intègre la notion d'accessibilité dans l'explication des préférences d'attachement de propositions relatives ambiguës. La forme du relativiseur et des antécédents signale le niveau d'accessibilité des entités du discours et guide l'attachement, aussi bien en production qu'en compréhension.

La méthodologie expérimentale utilisée pour l'analyse de corpus se distingue des études généralement publiées sur la production de ce type de phrases en ce qu'elle évalue l'effet de différents facteurs à l'aide de tests statistiques, les modèles linéaires mixtes.

La combinaison d'une étude en production et d'autres en compréhension permet d'évaluer si les facteurs qui affectent la compréhension du langage ont aussi un effet sur la production de structures semblables. Enfin, intégrer la notion d'accessibilité dans l'étude des préférences d'attachement confirme non seulement l'hypothèse selon laquelle ce type d'attachement est spécifique à cause du traitement anaphorique (Hemforth *et al.*, 2000), mais cela ouvre une nouvelle perspective explicative sur les variations translinguistiques et intra-linguistiques qui ont pu être mises en évidence.

Cette nouvelle approche doit être élaborée de façon plus approfondie, par exemple avec une mise à l'épreuve de ces hypothèses dans d'autres langues comme l'anglais, qui dispose de plusieurs stratégies de relativisation. D'autres langues peuvent s'avérer intéressantes grâce aux différences de marquage des systèmes de relativiseurs dans les langues du monde.

ANNEXES

Annexe 1 – Matériel expérimental utilisé pour le Questionnaire 1 (*qui, lequel/ laquelle, interindividuel*)

- 1 - Tu ne revois plus l'employé du notaire, qui/lequel vient de Provence.
- 2 - Je reconnais le banquier du préfet, qui/lequel aime jardiner.
- 3 - Tu oublies le neveu de l'avare, qui/lequel est endormi.
- 4 - Tu attends la concierge de la locataire, qui/laquelle parle espagnol.
- 5 - Tu rencontres l'admirateur du chanteur, qui/lequel est heureux.
- 6 - Tu revois l'associée de la fleuriste, qui/laquelle est malhonnête.
- 7 - J'affectionne l'apprenti du bottier, qui/lequel aime travailler.
- 8 - Tu cherches le frère du pêcheur, qui/lequel est hypocrite.
- 9 - Tu n'oublies pas l'assistant du chercheur, qui/lequel est attentif.
- 10 - Tu écoutes le fils du juge, qui/lequel habite Lyon.
- 11 - Je ne comprends plus la marraine de la fermière, qui/laquelle arrive maintenant.
- 12 - Tu comprends l'adversaire du coureur, qui/lequel sue beaucoup.
- 13 - Je revois le parrain du bambin, qui/lequel chante très bien.
- 14 - J'aide le patron du secrétaire, qui/lequel rit très fort.
- 15 - Je n'aide pas le cousin du boxeur, qui/lequel adore voyager.
- 16 - Je sous estime l'agent de l'acteur, qui/lequel est ruiné.
- 17 - J'accompagne le père du maçon, qui/lequel est amusant.
- 18 - Tu n'attends pas l'ennemi du chevalier, qui/lequel est fatigué.
- 19 - J'écoute le supérieur de l'ouvrier, qui/lequel va manger.
- 20 - Tu connais l'avocat de l'accusé, qui/lequel est chauve.
- 21 - Je cherche le maître de l'élève, qui/lequel est enrhumé.
- 22 - J'attends l'employeur du couturier, qui/lequel est âgé.
- 23 - Tu vois le chauffeur du comte, qui/lequel est grand-père.
- 24 - Tu aides la mère de la pianiste, qui/laquelle est radieuse.
- 25 - Tu reconnais la confidente de la commère, qui/laquelle est malade.
- 26 - Tu sous estimes le voisin du soldat, qui/lequel part en vacances.
- 27 - Tu affectionnes la nièce de la danseuse, qui/laquelle arrive demain.
- 28 - Je rencontre le valet du baron, qui/lequel est voûté.
- 29 - Tu accompagnes l'entraîneur de l'athlète, qui/lequel est bronzé.
- 30 - Tu ne comprends plus l'adjoint du maire, qui/lequel déménage bientôt.
- 31 - Tu n'aides pas le traducteur de l'écrivain, qui/lequel habite en centre-ville.
- 32 - Tu appelles le fournisseur du commerçant, qui/lequel veut partir.

Annexe 2 – Matériel expérimental utilisé pour le Questionnaire 2 (à *qui, auquel*, interindividuel)

- 1 - Tu rencontres l'admirateur du chanteur, à qui/auquel le copain prête de l'argent.
- 2 - Tu appelles le fournisseur du commerçant, à qui/auquel l'huissier notifie la dette.
- 3 - Je ne revois plus le successeur du directeur, à qui/auquel le comptable présente les factures.
- 4 - Je ne comprends plus le collègue de l'étudiant, à qui/auquel la bibliothécaire apporte le livre.
- 5 - J'affectionne l'apprenti du bottier, à qui/auquel le passant demande le chemin.
- 6 - J'écoute le supérieur de l'ouvrier, à qui/auquel l'enfant chante une chanson.
- 7 - Tu comprends l'adversaire du coureur, à qui/auquel l'arbitre demande de partir.
- 8 - Je revois le parrain du bambin, à qui/auquel la maman donne les clés.
- 9 - Je n'oublie pas l'ami du curé, à qui/auquel la cuisinière apporte le repas.
- 10 - Tu connais l'avocat de l'accusé, à qui/auquel le greffier téléphone le verdict.
- 11 - J'aide le patron du secrétaire, à qui/auquel l'épouse suggère la retraite.
- 12 - Je rencontre le valet du baron, à qui/auquel la bonne fait confiance.
- 13 - Tu cherches le frère du pêcheur, à qui/auquel la voisine confie le chaton.
- 14 - Tu n'oublies pas l'assistant du chercheur, à qui/auquel le chef refuse les congés.
- 15 - Je connais le gardien du prisonnier, à qui/auquel le policier conseille la prudence.
- 16 - Tu affectionnes le biographe du danseur, à qui/auquel la femme dit adieu.
- 17 - J'attends l'employeur du couturier, à qui/auquel le mannequin impose le silence.
- 18 - Je reconnais le banquier du préfet, à qui/auquel le mendiant ouvre la porte.
- 19 - Tu sous estimes le voisin du soldat, à qui/auquel le parieur interdit de tricher.
- 20 - Tu revois le concierge du locataire, à qui/auquel la propriétaire défend de fumer.
- 21 - Je vois le guide du touriste, à qui/auquel le docteur prescrit des antiviraux.
- 22 - Je sous estime l'agent de l'acteur, à qui/auquel la comédienne pardonne le retard.
- 23 - J'accompagne le père du maçon, à qui/auquel la mère écrit une lettre.
- 24 - Je comprends l'oncle du moine, à qui/auquel l'hôtelier souhaite la bienvenue.
- 25 - Tu ne revois plus l'employé du notaire, à qui/auquel la banque fait crédit.
- 26 - Tu aides le client de l'antiquaire, à qui/auquel le livreur fait appel.
- 27 - Tu reconnais l'associé du fleuriste, à qui/auquel le boucher doit de l'argent.
- 28 - J'oublie le correspondant du journaliste, à qui/auquel l'informateur confie le secret.
- 29 - Tu oublies le neveu de l'avare, à qui/auquel l'homme fait allusion.
- 30 - Tu vois le chauffeur du comte, à qui/auquel le jardinier porte des fleurs.
- 31 - Je n'aide pas le cousin du boxeur, à qui/auquel la fille promet de revenir.
- 32 - Tu écoutes le fils du juge, à qui/auquel la nourrice annonce une bonne nouvelle.

Annexe 3 – Matériel expérimental utilisé pour le Questionnaire 3 (à *qui, auquel/à laquelle*, intra-individuel)

- 1 - Je n'oublie pas le filleul du pédiatre, à qui/auquel le livreur apporte la pizza.
Je n'oublie pas la filleule de la pédiatre, à qui/à laquelle le livreur apporte la pizza.
- 2 - Tu écoutes le fils du juge, à qui/auquel la nourrice annonce une bonne nouvelle.
Tu écoutes la fille de la juge, à qui/à laquelle la nourrice annonce une bonne nouvelle.
- 3 - Tu ne revois plus le commanditaire du styliste, à qui/auquel la banque fait crédit.
Tu ne revois plus la commanditaire de la styliste, à qui/à laquelle la banque fait crédit.
- 4 - Tu cherches le frère du vendeur, à qui/auquel le vétérinaire confie le chaton.
Tu cherches la sœur de la vendeuse, à qui/à laquelle le vétérinaire confie le chaton.
- 5 - Je rencontre le valet du baron, à qui/auquel le bottier fait confiance.
Je rencontre la servante de la baronne, à qui/à laquelle le bottier fait confiance.
- 6 - Tu oublies le neveu du traducteur, à qui/auquel l'homme fait allusion.
Tu oublies la nièce de la traductrice, à qui/à laquelle l'homme fait allusion.
- 7 - Tu sous estimes le voisin de l'infirmier, à qui/auquel le restaurateur offre le dessert.
Tu sous estimes la voisine de l'infirmière, à qui/à laquelle le restaurateur offre le dessert.
- 8 - J'aide le patron du secrétaire, à qui/auquel le confident suggère la retraite.
J'aide la patronne de la secrétaire, à qui/à laquelle le confident suggère la retraite.
- 9 - Tu connais l'avocat du criminel, à qui/auquel le greffier téléphone le verdict.
Tu connais l'avocate de la criminelle, à qui/à laquelle le greffier téléphone le verdict.
- 10 - Je revois le parrain du bambin, à qui/auquel la maman donne les clés.
Je revois la marraine de la gamine, à qui/à laquelle la maman donne les clés.
- 11 - Je reconnais le banquier du libraire, à qui/auquel le mendiant ouvre la porte.
Je reconnais la banquière de la libraire, à qui/à laquelle le mendiant ouvre la porte.
- 12 - Tu appelles le scénariste du réalisateur, à qui/auquel l'huissier notifie la dette.
Tu appelles la scénariste de la réalisatrice, à qui/à laquelle l'huissier notifie la dette.
- 13 - J'affectionne le client du couturier, à qui/auquel le passant demande le chemin.
J'affectionne la cliente de la couturière, à qui/à laquelle le passant demande le chemin.
- 14 - Tu vois le chauffeur du duc, à qui/auquel le jardinier porte des fleurs.
Tu vois la bonne de la duchesse, à qui/à laquelle le jardinier porte des fleurs.
- 15 - J'attends le psychologue du malade, à qui/auquel le surveillant décrit le jardin.
J'attends la psychologue de la malade, à qui/à laquelle le surveillant décrit le jardin.
- 16 - Tu n'oublies pas l'assistant du chercheur, à qui/auquel le chef refuse les congés.
Tu n'oublies pas l'assistante de la chercheuse, à qui/à laquelle le chef refuse les congés.
- 17 - J'écoute le supérieur de l'ouvrier, à qui/auquel l'enfant chante une chanson.
J'écoute la supérieure de l'ouvrière, à qui/à laquelle l'enfant chante une chanson.
- 18 - Je ne revois plus le remplaçant du directeur, à qui/auquel le comptable présente les factures.
Je ne revois plus la remplaçante de la directrice, à qui/à laquelle le comptable présente les factures.
- 19 - Tu comprends le masseur du gymnaste, à qui/auquel l'arbitre demande de partir.
Tu comprends la masseuse de la gymnaste, à qui/à laquelle l'arbitre demande de partir.

- 20 - Tu rencontres l'admirateur du chanteur, à qui/auquel le copain prête de l'argent.
Tu rencontres l'admiratrice de la chanteuse, à qui/à laquelle le copain prête de l'argent.
- 21 - J'oublie le correspondant du journaliste, à qui/auquel l'informateur confie le secret.
J'oublie la correspondante de la journaliste, à qui/à laquelle l'informateur confie le secret.
- 22 - Je sous estime le maquilleur de l'acteur, à qui/auquel la comédienne pardonne le retard.
Je sous estime la maquilleuse de l'actrice, à qui/à laquelle la comédienne pardonne le retard.
- 23 - Je connais le gardien du détenu, à qui/auquel le policier conseille la prudence.
Je connais la gardienne de la détenue, à qui/à laquelle le policier conseille la prudence.
- 24 - Tu revois le concierge du locataire, à qui/auquel la propriétaire défend de fumer.
Tu revois la concierge de la locataire, à qui/à laquelle la propriétaire défend de fumer.
- 25 - Je ne comprends plus le collègue de l'étudiant, à qui/auquel la bibliothécaire apporte le livre.
Je ne comprends plus la collègue de l'étudiante, à qui/à laquelle la bibliothécaire apporte le livre.
- 26 - Tu reconnais le collaborateur du fleuriste, à qui/auquel le boucher doit de l'argent.
Tu reconnais la collaboratrice de la fleuriste, à qui/à laquelle le boucher doit de l'argent.
- 27 - Tu affectionnes le biographe du danseur, à qui/auquel le facteur dit bonjour.
Tu affectionnes la biographe de la danseuse, à qui/à laquelle le facteur dit bonjour.
- 28 - Je vois le guide du touriste, à qui/auquel le docteur prescrit des antiviraux.
Je vois la guide de la touriste, à qui/à laquelle le docteur prescrit des antiviraux.
- 29 - Je n'oublie pas le grand-père du marié, à qui/auquel le témoin apporte le dessert.
Je n'oublie pas la grand-mère de la mariée, à qui/à laquelle le témoin apporte le dessert.
- 30 - Tu ne revois plus le réceptionniste de l'hôtelier, à qui/auquel la blanchisseuse livre le linge.
Tu ne revois plus la réceptionniste de l'hôtesse, à qui/à laquelle la blanchisseuse livre le linge.
- 31 - Je comprends le préparateur du perchiste, à qui/auquel le sponsor offre du matériel.
Je comprends la préparatrice de la perchiste, à qui/à laquelle le sponsor offre du matériel.
- 32 - Tu sous estimes le dentiste de l'agriculteur, à qui/auquel le garçon tient la porte.
Tu sous estimes la dentiste de l'agricultrice, à qui/à laquelle le garçon tient la porte.

Annexe 4 – Matériel expérimental utilisé pour le Questionnaire 4 (*qui, le quel/laquelle, intra-individuel*)

- 1 - Tu ne revois plus l'employé du notaire, qui/lequel vient de Provence.
- 2 - Je reconnais le banquier du préfet, qui/lequel aime jardiner.
- 3 - Tu oublies le neveu de l'avare, qui/lequel est endormi.
- 4 - Tu attends la concierge de la locataire, qui/laquelle parle espagnol.
- 5 - Tu rencontres l'admirateur du chanteur, qui/lequel est heureux.
- 6 - Tu revois l'associée de la fleuriste, qui/laquelle est malhonnête.
- 7 - J'affectionne l'apprenti du bottier, qui/lequel aime travailler.
- 8 - Tu cherches le frère du pêcheur, qui/lequel est hypocrite.
- 9 - Tu n'oublies pas l'assistant du chercheur, qui/lequel est attentif.
- 10 - Tu écoutes le fils du juge, qui/lequel habite Lyon.
- 11 - Je ne comprends plus la marraine de la fermière, qui/laquelle arrive maintenant.
- 12 - Tu comprends l'adversaire du coureur, qui/lequel sue beaucoup.
- 13 - Je revois le parrain du bambin, qui/lequel chante très bien.
- 14 - J'aide le patron du secrétaire, qui/lequel rit très fort.
- 15 - Je n'aide pas le cousin du boxeur, qui/lequel adore voyager.
- 16 - Je sous estime l'agent de l'acteur, qui/lequel est ruiné.
- 17 - J'accompagne le père du maçon, qui/lequel est amusant.
- 18 - Tu n'attends pas l'ennemi du chevalier, qui/lequel est fatigué.
- 19 - J'écoute le supérieur de l'ouvrier, qui/lequel va manger.
- 20 - Tu connais l'avocat de l'accusé, qui/lequel est chauve.
- 21 - Je cherche le maître de l'élève, qui/lequel est enrhumé.
- 22 - J'attends l'employeur du couturier, qui/lequel est âgé.
- 23 - Tu vois le chauffeur du comte, qui/lequel est grand-père.
- 24 - Tu aides la mère de la pianiste, qui/laquelle est radieuse.
- 25 - Tu reconnais la confidente de la commère, qui/laquelle est malade.
- 26 - Tu sous estimes le voisin du soldat, qui/lequel part en vacances.
- 27 - Tu affectionnes la nièce de la danseuse, qui/laquelle arrive demain.
- 28 - Je rencontre le valet du baron, qui/lequel est voûté.
- 29 - Tu accompagnes l'entraîneur de l'athlète, qui/lequel est bronzé.
- 30 - Tu ne comprends plus l'adjoint du maire, qui/lequel déménage bientôt.
- 31 - Tu n'aides pas le traducteur de l'écrivain, qui/lequel habite en centre-ville.
- 32 - Tu appelles le fournisseur du commerçant, qui/lequel veut partir.

Annexe 5 – Matériel expérimental utilisé pour l'expérience avec le paradigme d'auto-présentation segmentée

Les phrases sont présentées dans les versions « attachement à N1 ». Il suffit d'inverser les deux premiers noms pour obtenir les versions « attachement à N2 ».

- 1 - Le plombier de l'architecte, qui/lequel installe les toilettes entre les deux chambres, prévoit d'aller courir demain.
- 2 - L'écolière de la maîtresse, qui/laquelle recommence sa punition jusqu'à la grande récréation, préfère le bleu au rouge.
- 3 - L'enfant du parrain, qui/lequel fait ses dents sur son jouet préféré, imite le bruit du train.
- 4 - Le garagiste du curé, qui/lequel change le moteur malgré son bon état, sursaute à chaque feu d'artifice.
- 5 - La serveuse de la cliente, qui/laquelle termine son service dans moins d'une heure, renverse l'assiette de légumes.
- 6 - L'infirmier du patient, qui/lequel prend la température avec un nouveau thermomètre, aperçoit un lapin courant dehors.
- 7 - La gamine de la maman, qui/laquelle finit son biberon avant son jeu favori, préfère les carottes aux haricots.
- 8 - Le geôlier du détenu, qui/lequel planifie sa ronde pour la longue veille, doit laver son linge demain.
- 9 - Le sponsor de l'athlète, qui/lequel fournit l'équipement pendant toute la compétition, demande le classement d'hier.
- 10 - Le seigneur du domestique, qui/lequel déclare la guerre contre son vieil ennemi, éteint la vieille bougie fondue.
- 11 - La voleuse de la victime, qui/laquelle cache son butin grâce à son complice, conserve son sang-froid malgré tout.
- 12 - Le facteur du libraire, qui/lequel livre les colis dès le petit matin, organise sa fête d'anniversaire.
- 13 - Le locataire du propriétaire, qui/lequel paie le loyer grâce aux allocations familiales, loge dans un bel appartement.
- 14 - Le bambin du pédiatre, qui/lequel jette son doudou près de la porte, feuillette encore le même livre.
- 15 - Le romancier du lecteur, qui/lequel accepte le Goncourt avec une modestie simulée, repart avec le gros livre.
- 16 - L'esclave du tyran, qui/lequel traîne ses chaînes avec beaucoup de difficultés, referme la porte après lui.
- 17 - Le docteur du chauffeur, qui/lequel adapte le traitement après la forte crise, siffle son air favori continuellement.
- 18 - Le banquier du maçon, qui/lequel ferme son guichet avec un sourire narquois, déjeune toujours au restaurant.
- 19 - Le champion du coach, qui/lequel tombe du podium pendant l'émouvante cérémonie, veut prendre sa retraite bientôt.

-
- 20 - Le boucher du livreur, qui/lequel prend son hachoir sur son présentoir neuf, apprécie les dimanches soir tranquille.
- 21 - Le comptable du dentiste, qui/lequel établit le budget pour l'année prochaine, peut facilement garer sa voiture.
- 22 - La concierge de l'épicière, qui/laquelle distribue le courrier malgré sa jambe fatiguée, emprunte de l'argent régulièrement.
- 23 - Le charpentier du trésorier, qui/lequel manipule son rabot malgré un mauvais aiguisage, remarque la fenêtre mal fermée.
- 24 - Le jardinier du laitier, qui/lequel répare la tondeuse avant le premier rendez-vous, circule dans sa camionnette blanche.
- 25 - Le dyslexique de l'orthophoniste, qui/lequel réussit ses exercices malgré un stress immense, habite loin du cabinet.
- 26 - Le bébé du papa, qui/lequel réclame du lait après une longue sieste, supporte le bruit assez difficilement.
- 27 - Le couturier du conseiller, qui/lequel présente sa collection devant un public intéressé, explique le fil conducteur suivi.
- 28 - Le prisonnier du gardien, qui/lequel retourne en cellule avec son courrier hebdomadaire, trouve les romans policiers débiles.
- 29 - L'étudiant du tuteur, qui/lequel passe ses examens après les vacances d'hiver, habite hors de la ville.
- 30 - Le formateur de l'apprenti, qui/lequel transmet son savoir avec une passion communicative, remplit la cafetière d'eau.
- 31 - Le coiffeur du patron, qui/lequel refait le shampoing sans sa brosse ronde, évoque son enfance très heureuse.
- 32 - Le boulanger du confesseur, qui/lequel retourne le pain d'un geste assuré, amuse les enfants en grimaçant.
- 33 - Le colonel du soldat, qui/lequel donne des ordres d'une voix autoritaire, met son costume à nettoyer.
- 34 - Le candidat du partisan, qui/lequel remporte les élections avec une large avance, envoie son courrier en retard.
- 35 - La fleuriste de la caissière, qui/laquelle enveloppe le bouquet dans un beau papier, surveille l'arrivée du camion.
- 36 - La surveillante de la collégienne, qui compte ses élèves après la sortie extrascolaire, enfile son manteau en laine.
- 37 - Le médecin du secrétaire, qui rédige l'ordonnance avant la prochaine réunion, arrive en retard le mercredi.
- 38 - Le psychiatre de l'autiste, qui/lequel reçoit en consultation pendant toute la journée, contemple la bibliothèque bien.
- 39 - L'avocat de l'employeur, qui/lequel figole sa plaidoirie jusqu'à l'audience suivante, redescend les escaliers en courant.
- 40 - La malade de l'infirmière, qui/laquelle avale son remède entre ses deux repas, élève ses enfants toute seule.

REFERENCES BIBLIOGRAPHIQUES

- Abney, S. P. (1989). A computational model of human parsing. *Journal of Psycholinguistic Research*, 18, 129-144.
- Albrecht, J. E. & O'Brien, E. J. (1991). Effects of centrality on retrieval of text-based concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 932-939.
- Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review*, 106, 748-765.
- Almor, A. (2004). A computational investigation of reference in production and comprehension. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use: bridging the language-as-product and language-as-action traditions*. Cambridge, MA: MIT Press.
- Almor, A., Kempler, D., MacDonald, M. C., Andersen, E. S., & Tyler, L. K. (1999). Why do Alzheimer patients have difficulty with pronouns? Working memory, semantics, and reference in comprehension and production in Alzheimer's disease. *Brain and Language*, 67, 202-227.
- Almor, A. & Nair, V. A. (2007). The form of referential expressions in discourse. *Language and Linguistics Compass*, 1, 84-99.
- Almor, A., Smith, D., V., Bonilha, L., Fridriksson, J., & Rorden, C. (2007). What is in a name? Spatial brain circuits are used to track discourse reference. *NeuroReport*, 18, 1215-1219.
- Altmann, G. T. M. & Steedman, M. (1988). Interactions with context during human sentence processing. *Cognition*, 30, 191-238.
- Altmann, G. T. M., Van Nice, K. Y., Garnham, A., & Henstra, J.-A. (1998). Late Closure in context. *Journal of Memory and Language*, 38, 459-484.
- Apothéloz, D. (1995). *Rôle et fonctionnement de l'anaphore dans la dynamique textuelle*. Genève: Droz.
- Ariel, M. (1988). Retrieving propositions from context: Why and how. *Journal of Pragmatics*, 12, 567-600.
- Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Croom Helm.
- Ariel, M. (1996). Referring expressions and the +/- coreference distinction. In T. Fretheim & J. K. Gundel (Eds.), *Reference and referent accessibility* (pp. 13-35). Amsterdam: John Benjamins.
- Ariel, M. (1999). Cognitive universals and linguistic conventions: The case of resumptive pronouns. *Studies In Language*, 23, 217-269.
- Ariel, M. (2001). Accessibility theory: An overview. In T. J. M. Sanders (Ed.), *Text representation. Linguistic and psycholinguistic aspects*. (pp. 29-87). Philadelphia, PA, USA: John Benjamins.
- Arnold, D. (2004). *Non-restrictive relative clauses in construction based HPSG*. Paper presented at the HPSG04 Conference, Stanford.
- Arnold, J. E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes*, 31, 137-162.

- Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition*, 76, B13-B26.
- Arnold, J. E. & Griffin, Z. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, 56, 521-536.
- Arnold, J. E., Wasow, T., Asudeh, A., & Alrenga, P. (2004). Avoiding attachment ambiguities: The role of constituent ordering. *Journal of Memory and Language*, 51, 55-70.
- Arnold, J. E., Wasow, T., Losongco, A., & Ginstrom, R. (2000). Heaviness vs. newness: The effect of structural complexity and discourse status on constituent ordering. *Language*, 76, 28-55.
- Auroux, S. & Rosier, I. (1987). Les sources historiques de la conception des deux types de relatives. *Langages*, 88, 9-29.
- Baayen, R. H. (à paraître). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baccino, T., De Vincenzi, M., & Job, R. (2000). Cross-linguistic studies of the late closure strategy: French and Italian. In M. De Vincenzi & V. Lombardo (Eds.), *Cross-linguistic perspectives on language processing* (pp. 89-118): Kluwer Academics.
- Baddeley, A. D. (1992). Working memory. *Science*, 255, 556-559.
- Baddeley, A. D. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, 36, 189-208.
- Baltazart, D. & Kister, L. (2000). Is it possible to predetermine a referent included in a French N de N structure? In S. Botley (Ed.), *Corpus-based and computational approaches do discourse anaphora*. Philadelphia, PA: John Benjamins.
- Bates, D. M. & DebRoy, S. (2004). Linear mixed models and penalized least squares. *Journal of Multivariate Analysis*, 91, 1-17.
- Bates, E., Devescovi, A., & D'Amico, S. (1999). Processing complex sentences: A cross-linguistic study. *Language and Cognitive Processes*, 14, 69-123.
- Bates, E., McNew, S., MacWhinney, B., Devescovi, A., & Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition*, 11, 245-299.
- Bayley, R. (1999). Relativization strategies in Mexican-American English. *American Speech*, 74, 115-139.
- Beach, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language*, 30, 644-663.
- Beaver, D. (2004). The optimization of discourse anaphora. *Linguistics and Philosophy*, 27, 3-56.
- Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279-362). New York: Wiley.
- Biber, D. (1986). Spoken and written textual dimensions in English: Resolving the contradictory findings. *Language*, 62.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Bock, K. & Irwin, D. E. (1980). Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior*, 19, 467-484.

- Bock, K. & Levelt, W. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945-984). Sand Diego, CA: Academic Press.
- Bock, K. & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21, 47-67.
- Boland, J., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language*, 34, 774-806.
- Bouma, G. & Hopp, H. (2006, 7-11 August). *Effects of word order and grammatical function on pronoun resolution in German*. Paper presented at the Ambiguity in Anaphora Workshop, ESSLI 2006, Malaga, Spain.
- Branigan, H. P., Pickering, M. J., & Cleland, A. A. (1999). Syntactic priming in written production: Evidence for rapid decay. *Psychonomic Bulletin & Review*, 6, 635-640.
- Branigan, H. P., Pickering, M. J., & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua*, 118, 172-189.
- Brennan, S. E. (1995). Centering attention in discourse. *Language and Cognitive Processes*, 10, 137-167.
- Brennan, S. E., Friedman, M. W., & Pollard, C. J. (1987). *A centering approach to pronouns*. Paper presented at the 25th Annual Meeting of the ACL, Stanford, CA.
- Brysbaert, M. & Mitchell, D. C. (1996). Modifier attachment in sentence parsing: Evidence from Dutch. *The Quarterly Journal of Experimental Psychology*, 49A, 664-695.
- Caplan, D., Alpert, N., & Waters, G. (1998). Effects of syntactic structure and propositional number on patterns of regional cerebral blood flow. *Journal of Cognitive Neuroscience*, 10, 541-552.
- Caplan, D., Alpert, N., & Waters, G. (1999). PET studies of syntactic processing with auditory sentence presentation. *NeuroImage*, 9, 343-351.
- Caramazza, A., Grober, E., Garvey, C., & Yates, J. (1977). Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behavior*, 16, 601-609.
- Carlson, K., Clifton, C., Jr., & Frazier, L. (2001). Prosodic boundaries in adjunct attachment. *Journal of Memory and Language*, 45, 58-81.
- Carreiras, M. & Clifton, C., Jr. (1993). Relative clause interpretation preferences in Spanish and English. *Language and Speech*, 36, 353-372.
- Carreiras, M. & Clifton, C., Jr. (1999). Another word on parsing relative clauses: Eyetracking evidence from Spanish and English. *Memory and Cognition*, 27, 826-833.
- Carreiras, M., Salillas, E., & Barber, H. (2004). Event-related potentials elicited during parsing of ambiguous relative clauses in Spanish. *Cognitive Brain Research*, 20, 98-105.
- Chafe, W. L. (1974). Language and consciousness. *Language*, 50, 111-133.
- Chafe, W. L. (1996). Inferring identifiability and accessibility. In T. Fretheim & J. K. Gundel (Eds.), *Reference and referent accessibility* (pp. 37-46). Amsterdam: John Benjamins.
- Chambers, C. G. & Smyth, R. (1998). Structural parallelism and discourse coherence: A test of Centering Theory. *Journal of Memory and Language*, 39, 593-608.
- Charolles, M. (1990). Coût, surcoût et pertinence. *Cahiers de linguistique française*, 11, 49-75.
- Christianson, K., Hollingworth, A., Halliwell, J., & Ferreira, F. (2001). Thematic roles assigned along the garden-path linger. *Cognitive Psychology*, 42, 368-407.

- Clark, H. H. (1975). Bridging. In R. C. Schank & B. L. Nash-Webber (Eds.), *Theoretical issues in natural language processing*. New York: Association for Computing Machinery.
- Clifton, C., Jr. (1993). Thematic roles in sentence parsing. *Canadian Journal of Experimental Psychology*, 47, 222-246.
- Clifton, C., Jr., Carlson, K., & Frazier, L. (2002). Informative prosodic boundaries. *Language and Speech*, 45, 87-114.
- Clifton, C., Jr., Traxler, M. J., Mohamed, M. T., Williams, R. S., Morris, R. K., & Rayner, K. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory and Language*, 49, 317-334.
- Cloitre, M. & Bever, T. G. (1988). Linguistic anaphors, levels of representation, and discourse. *Language and Cognitive Processes*, 3, 293-322.
- Cohen, H., Douaire, J., & Elsabbagh, M. (2001). The role of prosody in discourse prosody. *Brain and Cognition*, 46, 73-82.
- Colonna, S. (2001). *Facteurs influençant la levée des ambiguïtés syntaxiques*. Université Aix-Marseille I, Aix-en-Provence.
- Colonna, S. & Pynte, J. (2002). La levée des ambiguïtés syntaxiques : apport des recherches inter-langues. *L'Année Psychologique*, 102, 151-187.
- Comrie, B. (1981). Relative clauses. In *Language universals and linguistic typology*. Oxford: Basil Blackwell.
- Cornish, F. (2000). L'accessibilité cognitive des référents, le Centrage d'attention et la structuration du discours : une vue d'ensemble. *Verbum*, 22, 7-30.
- Cornish, F. (2001). L'anaphore pronominale indirecte : une question de focus. In W. de Mulder, C. Vet & C. Vetters (Eds.), *L'anaphore pronominale et nominale : études pragmatiques* (pp. 1-25). New York - Amsterdam: Rodopi.
- Cornish, F., Garnham, A., Cowles, H. W., Fossard, M., & André, V. (2005). Indirect anaphora in English and French: A cross-linguistic study of pronoun resolution. *Journal of Memory and Language*, 52, 363-376.
- Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, 13, 21-58.
- Crain, S. & Steedman, M. (1985). On not being led up the garden path: the use of context by the psychological parser. In D. R. Dowty, L. Karttunen & A. M. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives*. Cambridge: Cambridge University Press.
- Crawley, R. A., Stevenson, R., & Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research*, 19, 24-264.
- Creissels, D. (2006). *Syntaxe générale, une introduction typologique*. Paris: Hermès.
- Crocker, M. W. & Keller, F. (2006). Probabilistic grammars as models of gradience in language processing. In G. Fanselow, C. Féry, R. Vogel & M. Schlesewsky (Eds.), *Gradience in grammar: Generative perspectives*. Oxford: Oxford University Press.
- Cuetos, F. & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition*, 30, 73-105.
- Dahl, Ö. & Fraurud, K. (1996). Animacy in grammar and discourse. In T. Fretheim & J. K. Gundel (Eds.), *Reference and referent accessibility* (pp. 47-64). Amsterdam: John Benjamins.
- Damourette, J. & Pichon, E. (1969). *Des mots à la pensée. Essai de grammaire de la langue française 1911-1934* (Vol. 4). Paris: Editions d'Artrey.

- Dapretto, M. & Bookheimer, S. Y. (1999). Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron*, 24, 427-432.
- De Vincenzi, M. & Job, R. (1993). Some observations on the universality of the late-closure strategy. *Journal of Psycholinguistic Research*, 22, 189-206.
- De Vincenzi, M. & Job, R. (1995). An investigation of Late Closure: The role of syntax, thematic structure, and pragmatics in initial and final interpretation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1303-1321.
- Delle Luche, C. (2004). *Contributions des indices syntaxiques et sémantiques dans la compréhension de phrases*. Unpublished Master Thesis, Université Lyon 2 Lumière, Lyon.
- Delle Luche, C. (2006, 4-7 juillet). *Syntaxe et sémantique dans la compréhension de propositions relatives en français*. Paper presented at the Colloque international des étudiants chercheurs en didactique des langues et en linguistique, Grenoble.
- Demberg, V. & Keller, F. (à paraître). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*.
- Desmet, T., Brysbaert, M., & De Baecke, C. (2002). The correspondance between sentence production and corpus frequencies in modifier attachment. *The Quarterly Journal of Experimental Psychology*, 55A, 879-896.
- Desmet, T., De Baecke, C., & Brysbaert, M. (2002). The influence of referential discourse context on modifier attachment in Dutch. *Memory and Cognition*, 30, 150-157.
- Desmet, T., De Baecke, C., Drieghe, D., Brysbaert, M., & Vonk, W. (2006). Relative clause attachment in Dutch: On-line comprehension corresponds to corpus frequencies when lexical variables are taken into account. *Language and Cognitive Processes*, 21, 453-485.
- Desmet, T. & Declercq, M. (2006). Cross-linguistic priming of syntactic hierarchical configuration information. *Journal of Memory and Language*, 54, 610-632.
- Desmet, T. & Gibson, E. (2003). Disambiguation preferences and corpus frequencies in noun phrase conjunction. *Journal of Memory and Language*, 49, 353-374.
- Di Eugenio, B. (1998). Centering in Italian. In M. A. Walker, A. K. Joshi & E. F. Prince (Eds.), *Centering in discourse* (pp. 115-137). Oxford: Oxford University Press.
- Donnellan, K. S. (1966). Reference and definite descriptions. *The Philosophical Review*, 75, 281-304.
- Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, 27, 429-446.
- Dussias, P. E. (2001). Sentence parsing in fluent Spanish-English bilinguals. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 159-176). Cambridge, MA: Blackwell.
- Dussias, P. E. & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish-English bilinguals. *Bilingualism: Language and Cognition*, 10, 101-116.
- Ehrlich, K. (1980). Comprehension of pronouns. *The Quarterly Journal of Experimental Psychology*, 32, 247-255.
- Ehrlich, K. (1981). *Search and inference strategies in pronoun resolution: An experimental study*. Paper presented at the 19th annual meeting on Association for Computational Linguistics.
- Ehrlich, K. & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior*, 22, 75-87.

Elman, J. L., Hare, M., & McRae, K. (2004). Cues, constraints, and competition in sentence processing. In M. Tomasello & D. I. Slobin (Eds.), *Beyond nature-nurture: Essays in honor of Elizabeth Bates*. Mahwah, NJ: Lawrence Erlbaum Associates.

Epstein, R. (2002). The definite article, accessibility, and the construction of discourse referents. *Cognitive Linguistics*, 12, 333-378.

Erkù, F. & Gundel, J. K. (1987). The pragmatics of indirect anaphors. In J. Verschueren & M. Bertucelli-Papi (Eds.), *The pragmatic perspective: Selected papers from the 1985 International Pragmatics Conference*. Amsterdam, PA: John Benjamins.

Farmer, T. A., Anderson, S. E., & Spivey, M. J. (2007). Gradiency and visual context in syntactic garden-paths. *Journal of Memory and Language*, 57, 570-595.

Farmer, T. A., Cargill, S. A., Hindy, N. C., Dale, R., & Spivey, M. J. (2007). Tracking the continuity of language comprehension: Computer mouse trajectories suggest parallel syntactic processing. *Cognitive Science*, 31, 889-909.

Ferreira, F. & Clifton, C., Jr. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25, 348-368.

Ferreira, F. & Henderson, J. M. (1990). Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 555-568.

Ferreira, F. & Henderson, J. M. (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language*, 30, 725-745.

Ferreira, F. & Henderson, J. M. (1993). Reading processes during syntactic analysis and reanalysis. *Canadian Journal of Experimental Psychology*, 47, 247-275.

Ferreira, V. S. (2003). The persistence of optional complementizer production: Why saying "that" is not saying "that" at all. *Journal of Memory and Language*, 48, 379-398.

Ferreira, V. S. (à paraître). Ambiguity, accessibility, and a division of labor for communicative success. *Psychology of Learning and Motivation*.

Ferreira, V. S. & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, 40, 296-340.

Ferstl, E. C. & von Cramon, Y. D. (2001). The role of coherence and cohesion in text comprehension: an event-related fMRI study. *Cognitive Brain Research*, 11, 325-340.

Fletcher, C. R. (1984). Markedness and topic continuity in discourse processing. *Journal of Verbal Learning and Verbal Behavior*, 23, 487-493.

Fodor, J. A. (1986). *La modularité de l'esprit : essai sur la psychologie des facultés*. Paris: Editions de Minuit.

Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research*, 27, 285-319.

Forster, K. I. & Forster, J. C. (2002). DMDX version 2.9.01. Retrieved 12/2000, from <http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm>

Frank, S., Koppen, M., Noordman, L. G., & Vonk, W. (2003). Modeling knowledge-based inferences in story comprehension. *Cognitive Science*, 27, 875-910.

Frazier, L. (1985). Syntactic complexity. In D. R. Dowty, L. Karttunen & A. M. Zwicky (Eds.), *Natural language parsing. Psychological, computational, and theoretical perspectives* (pp. 129-189). Cambridge: Cambridge University Press.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *The psychology of reading. Attention and performance XII* (pp. 559-586). Hove: Lawrence Erlbaum.

- Frazier, L. (1990a). Exploring the architecture of the language-processing system. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing*. Cambridge: MIT press.
- Frazier, L. (1990b). Parsing modifiers: Special purpose routines in the human sentence processing mechanism? In D. A. Balota, G. B. Flores d'Arcais & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 303-330). Hillsdale: Lawrence Erlbaum.
- Frazier, L. (1995). Constraint satisfaction as a theory of sentence processing. *Journal of Psycholinguistic Research*, 24, 437-468.
- Frazier, L. & Clifton, C., Jr. (1996). *Construal*. Cambridge, MA: MIT Press.
- Frazier, L. & Clifton, C., Jr. (1997). Construal: Overview, motivation, and some new evidence. *Journal of Psycholinguistic Research*, 26, 277-295.
- Frazier, L. & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6, 291-325.
- Frazier, L. & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178-210.
- Frazier, L. & Rayner, K. (1987). Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. *Journal of Memory and Language*, 26, 505-526.
- Frenck-Mestre, C. (2004). Ambiguities and anomalies: What can eye-movements and event-related potentials reveal about second language sentence processing? In J. Kroll & A. M. B. de Groot (Eds.), *Handbook of Bilingualism* (pp. 268-284). Amsterdam: Elsevier.
- Frenck-Mestre, C. & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *The Quarterly Journal of Experimental Psychology*, 50A, 119-148.
- Frenck-Mestre, C. & Pynte, J. (2000a). Resolving syntactic ambiguities: Cross-linguistic differences? In M. De Vincenzi & V. Lombardo (Eds.), *Cross-linguistic perspectives on language processing*: Kluwer Academics.
- Frenck-Mestre, C. & Pynte, J. (2000b). 'Romancing' syntactic ambiguity: Why the French and the Italians don't see eye to eye. In A. Kennedy, R. Radach & D. Heller (Eds.), *Perceptual processes in reading*: Elsevier.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *TRENDS in Cognitive Sciences*, 6, 78-84.
- Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research*, 31, 45-63.
- Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, 1, 183-192.
- Fuchs, C. (1996). *Les ambiguïtés du français*. Paris: Ophrys.
- Galliot, M. & Laubreaux, R. (1966). *Le français langue vivante*. Paris: Privat-Didier.
- Garnham, A. & Oakhill, J. (1985). On-line resolution of anaphoric pronouns: Effects of inference making and verb semantics. *British Journal of Psychology*, 76, 385-393.
- Garnham, A., Oakhill, J., & Cain, K. (1997). The interpretation of anaphoric noun phrases: Time course, and effect of overspecificity. *The Quarterly Journal of Experimental Psychology*, 50A, 149-162.

- Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, 37, 58-93.
- Garrod, S., Freudenthal, D., & Boyle, E. (1994). The role of different types of anaphor in the on-line resolution of sentences in discourse. *Journal of Memory and Language*, 33, 39-68.
- Garrod, S. & Sanford, A. (1977). Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior*, 16, 77-90.
- Garrod, S. & Sanford, A. (1990). Referential processes in reading: Focusing on roles and individuals. In D. A. Balota, G. B. Flores d'Arcais & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 465-485). Hillsdale: Lawrence Erlbaum.
- Garrod, S. & Sanford, A. (1994). Resolving sentences in a discourse context. How discourse representation affects language understanding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 675-698). San Diego, CA: Academic Press.
- Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition*, 32, 99-156.
- Gernsbacher, M. A. & Hargreaves, D. J. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language*, 27, 699-717.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1-76.
- Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz (Ed.), *Image, language, brain: Papers from the first Mind Articulation Project symposium* (pp. 95-126). Cambridge, MA: MIT Press.
- Gibson, E., Desmet, T., Grodner, D., Watson, D., & Ko, K. (2005). Reading relative clauses in English. *Cognitive Linguistics*, 16, 313-353.
- Gibson, E., Pearlmutter, N. J., Canseco-Gonzalez, E., & Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition*, 59, 23-59.
- Gibson, E., Pearlmutter, N. J., & Torrens, V. (1999). Recency and lexical preferences in Spanish. *Memory and Cognition*, 27, 603-611.
- Gibson, E. & Schütze, C. T. (1999). Disambiguation preferences in noun phrase conjunction do not mirror corpus frequency. *Journal of Memory and Language*, 40, 263-279.
- Gibson, E., Schütze, C. T., & Salomon, A. (1996). The relationship between the frequency and the processing complexity of linguistic structure. *Journal of Psycholinguistic Research*, 25, 61-92.
- Gilboy, E. & Sopena, J.-M. (1996). Segmentation effects in the processing of complex NPs with relative clauses. In M. Carreiras, G.-A. J. E. & N. Sebastián-Gallés (Eds.), *Language processing in Spanish* (pp. 191-206). Mahwah, NJ: Lawrence Erlbaum.
- Gilboy, E., Sopena, J.-M., Clifton, C., Jr., & Frazier, L. (1995). Argument structure and association preferences in Spanish and English complex NPs. *Cognition*, 54, 131-167.
- Giora, R. (1988). On the informativeness requirement. *Journal of Pragmatics*, 12, 547-565.
- Givón, T. (1982). Logic vs. pragmatics, with human language as the referee: Toward an empirically viable epistemology. *Journal of Pragmatics*, 6, 81-133.
- Givón, T. (1983). Topic continuity in discourse: An introduction. In T. Givón (Ed.), *Topic continuity in discourse: A quantitative cross-language study* (pp. 5-41). Amsterdam: John Benjamins.

- Givón, T. (1990). *Syntax: A functional typological introduction* (Vol. 2). Amsterdam, Ph: John Benjamins.
- Givón, T. (1992). The grammar of referential coherence as mental processing instructions. *Linguistics*, 30, 5-55.
- Gordon, P. C. & Chan, D. (1995). Pronouns, passives, and discourse coherence. *Journal of Memory and Language*, 34, 216-231.
- Gordon, P. C., Grosz, B., & Gilliom, L. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 3, 311-347.
- Gordon, P. C. & Hendrick, R. (1997). Intuitive knowledge of linguistic co-reference. *Cognition*, 62, 325-370.
- Gordon, P. C. & Hendrick, R. (1998). The representation and processing of coreference in discourse. *Cognitive Science*, 22, 389-424.
- Gordon, P. C., Hendrick, R., Ledoux, K., & Yang, C. L. (1999). Processing of reference and the structure of language: An analysis of complex noun phrases. *Language and Cognitive Processes*, 14, 353-379.
- Gordon, P. C. & Searce, K. A. (1995). Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory and Cognition*, 23, 313-323.
- Green. (2006). Absence of real evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, 55, 1-17.
- Greene, S. B., Gerrig, R. J., McKoon, G., & Ratcliff, R. (1994). Unheralded pronouns and management by common ground. *Journal of Memory and Language*, 33, 511-526.
- Greene, S. B., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 266-283.
- Grodner, D., Gibson, E., & Watson, D. (2005). The influence of contextual contrast on syntactic processing: Evidence for strong interaction in sentence comprehension. *Cognition*, 95, 275-296.
- Grosz, B., Joshi, A. K., & Weinstein, S. (1983). *Providing a unified account of definite noun phrases in discourse*. Paper presented at the 21st Annual Meeting of the Association for Computational Linguistics, Cambridge, MA.
- Grosz, B., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics*, 21, 203-225.
- Grosz, B. & Sidner, C. L. (1997). Lost intuitions and forgotten intentions. In M. A. Walker, A. K. Joshi & E. F. Prince (Eds.), *Centering in discourse*. Oxford: Oxford University Press.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69, 274-307.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (2000). Statut cognitif et forme des anaphoriques indirects. *Verbum*, 22, 79-102.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (2001). Definite descriptions and cognitive status in English: Why accommodation is unnecessary. *English Language and Linguistics*, 5, 273-295.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (2002, September 18-20). *Pronouns without explicit antecedents: How do we know when a pronoun is referential?* Paper presented at the DAARC-4 (the fourth Discourse Anaphora and Anaphora Resolution Colloquium), Lisbon, Portugal.

- Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, *12*, 556-568.
- Guy, G. R. & Bayley, R. (1995). On the choice of relative pronouns in English. *American Speech*, *70*(148-162).
- Hagoort, P. & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: The P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia*, *38*, 1531-1549.
- Hagoort, P., Brown, C. M., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8*, 439-483.
- Hahne, A. & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERP's. *Cognitive Brain Research*, *13*, 339-356.
- Halmari, H. (1996). On accessibility and coreference. In T. Fretheim & J. K. Gundel (Eds.), *Reference and referent accessibility* (pp. 155-178). Amsterdam: John Benjamins.
- Hare, M., Tanenhaus, M. K., & McRae, K. (2007). Understanding and producing the reduced relative construction: Evidence from ratings, editing and corpora. *Journal of Memory and Language*, *56*, 410-435.
- Harville, D. A. & Mee, R. W. (1984). A mixed-model procedure for analyzing ordered categorical data. *Biometrics*, *40*, 393-408.
- Haviland, S. E. & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, *13*, 512-521.
- Hawkins, J. A. (1984). A note on referent identifiability and co-presence. *Journal of Pragmatics*, *8*, 649-659.
- Heim, I. (1998). Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach. In U. Sauerland & O. Percus (Eds.), *The interpretive tract: Working paper in syntax and semantics* (Vol. 25, pp. 205-246): MIT Working Paper in Linguistics.
- Heine, A., Tamm, S., Hofmann, M., Hutzler, F., & Jacobs, A. M. (2006). Does the frequency of the antecedent noun affect the resolution of pronominal anaphors? *Neuroscience Letters*, 7-12.
- Hemforth, B., Konieczny, L., & Scheepers, C. (2000). Syntactic attachment and anaphor resolution: The two sides of relative clause attachment. In M. W. Crocker, M. J. Pickering & C. Clifton, Jr. (Eds.), *Architectures and mechanisms for language processing* (pp. 259-282). Cambridge, UK: Cambridge University Press.
- Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics*, *31*, 423-447.
- Hill, R. L. & Murray, W. S. (1999). Commas and spaces: Effects of punctuation on eye movements and sentence parsing. In A. Kennedy, R. Radach, D. Heller & J. Pynte (Eds.), *Reading as a perceptual process*. Oxford: Elsevier.
- Hindle, D. & Rooth, M. (1993). Structural ambiguity and lexical relations. *Computational Linguistics*, *19*, 103-120.
- Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language*, *54*, 425-443.
- Hobbs, J. R. (1978). Resolving pronoun references. *Lingua*, *44*, 311-338.
- Hobbs, J. R. (1979). Coherence and coreference. *Cognitive Science*, *3*, 67-90.

- Hoeks, J. C. J., Hendriks, P., Vonk, W., Brown, C. M., & Hagoort, P. (2006). Processing the noun phrase versus sentence coordination ambiguity: Thematic information does not completely eliminate processing difficulty. *The Quarterly Journal of Experimental Psychology*, *59*, 1581-1599.
- Holmes, V. M. (1988). Hesitations and sentence planning. *Language and Cognitive Processes*, *3*, 323-361.
- Holmes, V. M. (1995). A crosslinguistic comparison of the production of utterances in discourse. *Cognition*, *54*, 169-207.
- Holmes, V. M. & O'Regan, J. K. (1981). Eye fixation patterns during the reading of relative-clause sentences. *Journal of Verbal Learning and Verbal Behavior*, *20*, 417-430.
- Igoa, J. M., Carreiras, M., & Meseguer, E. (1998). A study on late closure in Spanish: Principle-grounded vs. frequency-based accounts of attachment preferences. *The Quarterly Journal of Experimental Psychology*, *51A*, 561-592.
- Jackendoff, R. (1999). Parallel constraint-based generative theories of language. *TRENDS in Cognitive Sciences*, *3*, 393-400.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformational or not) and towards logit mixed models. *Journal of Memory and Language*.
- Jaeger, T. F. & Wasow, T. (à paraître). *Processing as a source of accessibility effects on variation*. Paper presented at the 31st Annual Meeting of the Berkeley Linguistic Society.
- Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, *4*, 71-115.
- Johnson-Laird, P. N. & Garnham, A. (1980). Descriptions and discourse models. *Linguistics and Philosophy*, *3*, 371-393.
- Jones, M. A. (1996). *Foundations of French syntax*. Cambridge: Cambridge University Press.
- Jun, S.-A. (2003). Prosodic phrasing and attachment preferences. *Journal of Psycholinguistic Research*, *32*.
- Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, *20*, 137-194.
- Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science*, *274*, 114-116.
- Just, M. A., Carpenter, P. A., & Wooley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*, 228-238.
- Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*, 159-201.
- Kaan, E. & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience*, *15*, 98-110.
- Kaan, E., Wijnen, F., & Swaab, T. Y. (2004). Gapping: Electrophysiological evidence for immediate processing of "missing" verbs in sentence comprehension. *Brain and Language*, *89*, 584-592.
- Kameyama, M. (1998). Intrasentential centering: A case study. In M. A. Walker, A. K. Joshi & E. F. Prince (Eds.), *Centering in discourse* (pp. 89-114). Oxford: Oxford University Press.
- Karabanov, A., König, P., & Bosch, P. (2006). *Eye-tracking evidence for online processes in the comprehension of referential expressions*. Paper presented at the International Conference on Linguistic Evidence.

- Keenan, E. L. & Comrie, B. (1987). Noun phrase accessibility and Universal Grammar. In E. L. Keenan (Ed.), *Universal Grammar: 15 essays* (pp. 3-45). London: Croom Helm.
- Kehler, A. (1993). *Intrasentential constraints on intersentential anaphora in Centering Theory*. Paper presented at the Workshop on Centering Theory in Naturally Occurring Discourse, University of Pennsylvania.
- Kehler, A. (1997). Current theories of Centering for pronoun interpretation: A critical evaluation. *Computational Linguistics*, 23, 467-475.
- Kibble, R. (2001). A reformulation of Rule 2 of Centering Theory. *Computational Linguistics*, 27, 579-587.
- King, J. W. & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30, 580-602.
- Kintsch, W. & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363-394.
- Kister, L. (1999). *Identification des chaînes de référence dans les systèmes automatiques : le cas des anaphores dans les N de N*. Paper presented at the VEXTAL'99, Venezia, San Servolo.
- Kister, L. (2002, 24-27 juin). *Relatifs et référents inclus dans un SN : des paramètres pour présélectionner la saisie*. Paper presented at the TALN 2002, Nancy.
- Kjelgaard, M. M. & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, 40, 153-194.
- Kleiber, G. (1987). Relatives restrictives/relatives appositives : Dépassement(s) autorisé(s). *Langages*, 88, 41-63.
- Kleiber, G. (1990). Marqueurs référentiels et processus interprétatifs : pour une approche "plus sémantique". *Cahiers de linguistique française*, 11, 241-258.
- Kleiber, G. (1994). *Anaphores et pronoms*. Louvain-la-Neuve: Duculot.
- Klin, C. M., Guzmán, A. E., Weingartner, K. M., & Ralano, A. S. (2006). When anaphor resolution fails: Partial encoding of anaphoric inferences. *Journal of Memory and Language*, 54, 131-143.
- Kortmann, B., Burridge, K., Mesthrie, R., Schneider, E. W., & Upton, C. (Eds.). (2004). *A handbook of varieties of English* (Vol. 2: Morphology and Syntax). Berlin: Mouton de Gruyter.
- Kuperberg, G. R., Kreher, D. A., Sitnikova, T., Caplan, D. N., & Holcomb, P. (2006). The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. *Brain and Language*.
- Kuperberg, G. R., Lakshmanan, B. M., Caplan, D. N., & Holcomb, P. J. (2006). Making sense of discourse: An fMRI study of causal inferencing across sentences. *NeuroImage*, 33, 343-361.
- Kutas, M. & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307, 161-163.
- Kutas, M. & Kluender, R. (1994). What is who violating? A reconsideration of linguistic violations in light of event-related brain potentials. In H. T. Heinze, G. R. Mangun & T. F. Münte (Eds.), *Cognitive Electrophysiology* (pp. 183-210). Boston: Birkhaeuser.
- Kutas, M. & Van Petten, C. (1988). Event-related brain potential studies of language. *Advances in Psychophysiology*, 3, 139-187.

- Lambrech, K. (1994). *Information structure and sentence form. Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
- Ledoux, K., Gordon, P. C., Camblin, C., & Swaab, T. Y. (2007). Coreference and lexical repetition: Mechanisms of discourse integration. *Memory and Cognition*, 35, 801-815.
- Leech, G. (1981). *Semantics: The study of meaning* (2nd ed.). London: Pelican.
- Levine, W. H., Guzmán, A. E., & Klin, C. M. (2000). When anaphor resolution fails. *Journal of Memory and Language*, 43, 594-617.
- Levinson, S., C. (1987a). Minimization and conversational inference. In J. Verschueren & M. Bertucelli-Papi (Eds.), *The pragmatic perspective*. Amsterdam: John Benjamins.
- Levinson, S., C. (1987b). Pragmatics and the grammar of anaphora: A partial pragmatic reduction of Binding and Control phenomena. *Journal of Linguistics*, 23, 379-434.
- Look, R. (2003). Les fonctions des propositions subordonnées relatives "appositives" en discours. *Anglophonia*, 12, 113-131.
- Look, R. (2007). Appositive relative clauses and their functions in discourse. *Journal of Pragmatics*, 39, 336-362.
- MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language*, 32, 692-715.
- MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, 9, 157-201.
- MacDonald, M. C. (1999). Distributional information in language comprehension, production, and acquisition: three puzzles and a moral. In B. MacWhinney (Ed.), *The emergence of language* (pp. 177-196). Mahwah, NJ: Lawrence Erlbaum.
- MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, 24, 56-98.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676-703.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken-word recognition. *Cognition*, 25, 71-102.
- Marslen-Wilson, W. D. & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, 8, 1-71.
- Matsui, T. (1993). Bridging reference and the notions of 'topic' and 'focus'. *Lingua*, 90, 49-68.
- Maynell, L. A. (1999). *Effect of pitch accent placement on resolving relative clause ambiguity in English*. Paper presented at the 12th Annual CUNY, New York.
- Maynell, L. A. (2000). *Prosodic effects on relative clause attachment*. Paper presented at the 13th Annual CUNY, La Jolla.
- McDonald, J. L. & MacWhinney, B. (1995). The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language*, 34, 543-566.
- McKoon, G., Gerrig, R. J., & Greene, S. B. (1996). Pronoun resolution without pronouns: Some consequences of memory-based text processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 919-932.
- McKoon, G. & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, 99, 440-466.
- McKoon, G. & Ratcliff, R. (2003). Meaning through syntax: Language comprehension and the reduced relative clause construction. *Psychological Review*, 110, 490-525.

- McKoon, G., Ratcliff, R., Ward, G., & Sproat, R. (1993). Syntactic prominence effects on discourse processes. *Journal of Memory and Language*, 32, 593-607.
- McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes*, 12, 137-176.
- McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, 38, 283-312.
- Meseguer, E., Carreiras, M., & Clifton, C., Jr. (2002). Overt reanalysis strategies and eye movements during the reading of mild garden path sentences. *Memory and Cognition*, 30, 551-561.
- Miltsakaki, E. (2004a). *A centering analysis of relative clauses in English and Greek*. Paper presented at the 28th Penn Linguistics Colloquium, University of Pennsylvania, PA.
- Miltsakaki, E. (2004b). Toward an aposynthesis of topic continuity and intrasentential anaphora. *Computational Linguistics*, 28, 319-355.
- Mitchell, D. C. (1989). Verb guidance and other lexical effects in parsing. *Language and Cognitive Processes*, 4, SI 123-154.
- Mitchell, D. C. (1994). Sentence parsing. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics*. San Diego, CA: Academic Press.
- Mitchell, D. C. & Brysbaert, M. (1998). Challenges to recent theories of cross-linguistic variation in parsing: Evidence from Dutch. In D. Hillert (Ed.), *Sentence processing: A cross-linguistic perspective*. San Diego, CA: Academic Press.
- Mitchell, D. C., Cuetos, F., Corley, M. M. B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical record. *Journal of Psycholinguistic Research*, 24, 469-488.
- Mitchell, D. C., Cuetos, F., & Zagar, D. (1990). Reading in different languages: Is there a universal mechanism for parsing sentences? In D. A. Balota, G. B. Flores d'Arcais & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 285-302). Hillsdale: Lawrence Erlbaum.
- Mitchell, D. C. & Green, D. W. (1978). The effects of context and content on immediate processing in reading. *The Quarterly Journal of Experimental Psychology*, 30, 609-636.
- Mulkern, A. E. (1996). The game of the name. In T. Fretheim & J. K. Gundel (Eds.), *Reference and referent accessibility* (pp. 235-250). Amsterdam: John Benjamins.
- Neville, H. J., Mills, D. L., & Lawson, D. S. (1992). Fractioning language: Different neural subsystems with different sensitive periods. *Cerebral Cortex*, 2, 244-258.
- New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet : LEXIQUETM. *L'Année Psychologique*, 101, 447-462.
- Nicol, J. & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, 18, 5-19.
- Nieuwland, M. S., Otten, M., & Van Berkum, J. J. A. (2007). Who are you talking about? Tracking discourse-level referential processing with event-related brain potentials. *Journal of Cognitive Neuroscience*, 19, 228-236.
- Nieuwland, M. S., Petersson, K. M., & Van Berkum, J. J. A. (2007). On sense and reference: Examining the functional neuroanatomy of referential processing. *NeuroImage*, 37, 993-1004.
- Nieuwland, M. S. & Van Berkum, J. J. A. (2006). Individual differences and contextual bias in pronoun resolution: Evidence from ERPs. *Brain Research*, 1118, 155-167.

- O'Brien, E. J. & Albrecht, J. E. (1991). The role of context in accessing antecedents in texts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1991, 94-102.
- O'Brien, E. J., Raney, G. E., Albrecht, J. E., & Rayner, K. (1997). Processes involved in the resolution of explicit anaphors. *Discourse Processes*, 23, 1-24.
- Osterhout, L. & Holcomb, P. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31, 785-806.
- Osterhout, L., McLaughlin, J., & Bersick, M. (1997). Event-related brain potentials and human language. *TRENDS in Cognitive Sciences*, 1, 203-209.
- Osterhout, L. & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34, 739-773.
- Osterhout, L. & Nicol, J. (1999b). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes*, 14, 283-317.
- Palolahti, M., Leino, S., Jokela, M., Kopra, K., & Paavilainen, P. (2005). Event-related potentials suggest early interaction between syntax and semantics during on-line sentence comprehension. *Neuroscience Letters*, 384, 222-227.
- Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, 41, 427-456.
- Petrone, C., Colonna, S., Hemforth, B., d'Imperio, M., & Pynte, J. (2005). *Length effects in PP-attachment*. Paper presented at the 11th Annual Conference on Architectures and Mechanisms for Language Processing.
- Pickering, M. J. & Traxler, M. J. (1998). Plausibility and recovery from garden paths: An eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 940-961.
- Pickering, M. J., Traxler, M. J., & Crocker, M. W. (2000). Ambiguity resolution in sentence processing: Evidence against frequency-based accounts. *Journal of Memory and Language*, 43, 447-475.
- Poesio, M., Stevenson, R., Di Eugenio, B., & Hitzeman, J. (2004). Centering: a parametric theory and its instantiations. *Computational Linguistics*, 30, 309-363.
- Poesio, M. & Vieira, R. (1998). A corpus-based investigation of definite description use. *Computational Linguistics*, 24, 183-216.
- Pollatsek, A. & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental psychology: Learning, memory, and Cognition*, 21, 785-794.
- Prat-Sala, M. & Branigan, H. P. (2000). Discourse constraints on syntactic processing in language production: A cross-linguistic study of English and Spanish. *Journal of Memory and Language*, 42, 168-182.
- Prince, E. F. (1990). *Syntax and discourse: A look at resumptive pronouns*. Paper presented at the 16th annual meeting of the Berkeley Linguistic Society.
- Pynte, J. & Colonna, S. (2000). Decoupling syntactic parsing from visual inspection: The case of relative clause attachment in French. In A. Kennedy, R. Radach, D. Heller & J. Pynte (Eds.), *Reading as a perceptual process*. Oxford: Elsevier.
- Pynte, J. & Colonna, S. (2001). Competition between primary and non-primary relations during sentence comprehension. *Journal of Psycholinguistic Research*, 30, 569-599.
- Pynte, J. & Kennedy, A. (2007). The influence of punctuation and word class on distributed processing in normal reading. *Vision Research*, 47, 1215-1227.

- Pynte, J. & Prieur, B. (1996). Prosodic breaks and attachment decisions in sentence parsing. *Language and Cognitive Processes*, 11, 165-191.
- Race, D. S. & MacDonald, M. C. (2003). *The use of "that" in the production and comprehension of object relative clauses*. Paper presented at the 25th Annual Meeting of the Cognitive Science Society.
- Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin*, 85, 618-660.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.
- Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, 22, 358-374.
- Rayner, K. & Clifton, C., Jr. (2002). Language processing. In D. Medin (Ed.), *Stevens handbook of experimental psychology, Third edition: Vol 2, Memory and cognitive processes*. New York: John Wiley and Sons.
- Rayner, K. & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory and Cognition*, 14, 191-201.
- Rayner, K. & Frazier, L. (1987). Parsing temporarily ambiguous complements. *The Quarterly Journal of Experimental Psychology*, 39A, 657-673.
- Rayner, K., Garrod, S., & Perfetti, C. A. (1992). Discourse influences during parsing are delayed. *Cognition*, 45, 109-139.
- Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology*, 53, 1061-1080.
- Rayner, K. & Sereno, J. A. (1994). Eye movements in reading. Psycholinguistic studies. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 57-81). Sand Diego, CA: Academic Press.
- Rayner, K., Sereno, J. A., Morris, R. G., Schmauder, A. R., & Clifton, C., Jr. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, 4(SI 21-49).
- Reboul, A. (1989). Résolution de l'anaphore pronominal : Sémantique et/ou pragmatique. *Cahiers de linguistique française*, 10, 77-100.
- Reboul, A. (1995). La pragmatique à la conquête de nouveaux domaines : la référence. *L'information grammaticale*, 66, 32-37.
- Reboul, A. (1997). What (if anything) is accessibility? A relevance-oriented criticism of Ariel's Accessibility Theory of referring expressions. In J. F. Connolly (Ed.), *Discourse and pragmatics in functional grammar*. Berlin/New York: de Gruyter.
- Reboul, A. (2001). Foundations of reference and predication. In M. Haspelmath, E. König, W. Oesterreicher & W. Raible (Eds.), *Language typology and language universals. An international handbook* (pp. 509-522). Berlin: de Gruyter.
- Reichman, R. (1978). Conversational coherency. *Cognitive Science*, 2, 283-237.
- Riegel, M., Pellat, J.-C., & Rioul, R. (2004). *Grammaire méthodique du français*. Paris: Presses Universitaires de France.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192-233.

- Rosenbach, A. (2005). Animacy versus weight as determinants of grammatical variation in English. *Language*, 81, 613-644.
- Rosenbach, A. (2008). Animacy and grammatical variation - Findings from English genitive variation. *Lingua*, 118, 151-171.
- Sanford, A. & Garrod, S. (1989). What, when, an how? Questions of immediacy in anaphoric reference resolution. *Language and Cognitive Processes*, 4, SI 235-262.
- Sanford, A., Garrod, S., Lucas, A., & Henderson, R. (1983). Pronouns without explicit antecedents? *Journal of Semantics*, 2, 303-318.
- Schafer, A., Carter, J., Clifton, C., Jr., & Frazier, L. (1996). Focus in relative clause construal. *Language and Cognitive Processes*, 11, 135-163.
- Schafer, A., Speer, S. R., Warren, P., & White, S. D. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research*, 29.
- Schelstraete, M.-A. (1993). La conception du traitement syntaxique en compréhension de phrases. *L'Année Psychologique*, 93, 543-582.
- Schelstraete, M.-A. & Degand, L. (1998). Assignment of grammatical functions in French relative clauses. *Language Science*, 20, 163-188.
- Schmalhofer, F., McDaniel, M. A., & Keefe, D. (2002). A unified model for predictive and bridging inferences. *Discourse Processes*, 33, 105-132.
- Schmid, H. (1994). *Probabilistic part-of-speech tagging using decision trees*. Paper presented at the International Conference on New Methods in Language Processing.
- Schütze, C. T. & Gibson, E. (1999). Argumenthood and English prepositional phrase attachment. *Journal of Memory and Language*, 40, 409-431.
- Seppänen, A. (1999). Dialectal variation in English relativization. *Lingua*, 109, 15-34.
- Sheldon, A. (1974). On the role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior*, 13, 272-281.
- Sidner, C. L. (1979). *Towards a computational theory of definite anaphora comprehension in English discourse*. Massachusetts Institute of Technology.
- Sidner, C. L. (1981). Focusing for interpretation of pronouns. *American Journal of Computational Linguistics*, 7, 217-231.
- Singer, M. (1994). Discourse inference processes. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 479-515). San Diego, CA: Academic Press.
- Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, 23, 197-229.
- Solomon, E. S. & Pearlmutter, N. J. (2004). Semantic integration and syntactic planning in language production. *Cognitive Psychology*, 49, 1-46.
- Speer, S. R., Kjelgaard, M. M., & Dobroth, K. M. (1996). The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research*, 25, 249-271.
- Sperber, D. & Wilson, D. (1989). *La pertinence : communication et cognition*. Paris: Les Editions de Minuit.
- Spivey-Knowlton, M. J. & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, 55, 227-267.
- Spivey-Knowlton, M. J. & Tanenhaus, M. K. (1994). Referential context and syntactic ambiguity resolution. In C. Clifton, Jr., L. Frazier & K. Rayner (Eds.), *Perspectives in sentence processing* (pp. 415-439): Lawrence Erlbaum.

- Spivey-Knowlton, M. J., Trueswell, J. C., & Tanenhaus, M. K. (1993). Context effects in syntactic ambiguity resolution: Discourse and semantic influences in parsing reduced relative clauses. *Canadian Journal of Experimental Psychology*, *47*, 276-309.
- Spivey, M. J. & Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1521-1543.
- Stahlke, H. F. W. (1976). Which that. *Language*, *52*, 584-610.
- Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language*, *86*, 142-164.
- Steinhauer, K. & Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research*, *30*, 267-295.
- Stevenson, R. (2002). The role of salience in the production of referring expressions: A psycholinguistic perspective. *Journal of Psycholinguistic Research*, *23*, 295-322.
- Stevenson, R., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, *9*, 519-548.
- Stevenson, R., Knott, A., Oberlander, J., & McDonald, S. (2000). Interpreting pronouns and connectives: Interactions among focusing, thematic roles and coherence relations. *Language and Cognitive Processes*, *15*, 225-262.
- Streb, J., Rösler, F., & Hennighausen, E. (1999). Event-relative responses to pronoun anaphors in parallel and nonparallel discourse structures. *Brain and Language*, *70*, 273-286.
- Strohner, H., Sichelschmidt, L., Duwe, I., & Kessler, K. (2000). Discourse focus and conceptual relations in resolving referential ambiguity. *Journal of Psycholinguistic Research*, *29*, 497-516.
- Strube, G. & Hahn, U. (1999). Functional centering - Grounding referential coherence in information structure. *Computational Linguistics*, *25*, 309-344.
- Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, *48*, 542-562.
- Sturt, P., Scheepers, C., & Pickering, M. J. (2002). Syntactic ambiguity resolution after initial misanalysis: The role of recency. *Journal of Memory and Language*, *46*, 371-390.
- Tabor, W., Juliano, C., & Tanenhaus, M. K. (1997). Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes*, *12*, 211-271.
- Tanenhaus, M. K., Boland, J., Garnsey, S. M., & Carlson, G. N. (1989). Lexical structure in parsing long-distance dependencies. *Journal of Psycholinguistic Research*, *18*, 37-50.
- Tanenhaus, M. K., Carlson, G. N., & Trueswell, J. C. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes*, *4*, SI 211-234.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632-1634.
- Tanenhaus, M. K. & Trueswell, J. C. (1995). Sentence comprehension. In J. L. Miller & P. D. Eimas (Eds.), *Handbook in Perception and Cognition, Volume 11: Speech Language and Communication*: Academic Press.
- Tao, H. & McCarthy, M. J. (2001). Understanding non-restrictive *which*-clauses in spoken English, which is not an easy task. *Language Science*, *23*, 651-677.

Taraban, R. & McClelland, J. L. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *Journal of Memory and Language*, 27, 597-632.

Tellier, C. (1996). *Éléments de syntaxe du français : méthodes d'analyse en grammaire générative*. Montréal, Canada: Les Presses de l'Université de Montréal.

Temperley, D. (2003). Ambiguity avoidance in English relative clauses. *Language*, 79, 464-484.

Temperley, D. (2007). Minimization of dependency length in written English. *Cognition*, 105, 300-333.

Tetreault, J. R. (2001). A corpus-based evaluation of centering and pronoun resolution. *Computational Linguistics*, 27, 507-520.

Thornton, R., Gil, M., & MacDonald, M. C. (1998). Accounting for crosslinguistic variation: A constraint-based perspective. In D. Hillert (Ed.), *Syntax and semantics* (Vol. 31, pp. 211-225). San Diego: Academic Press.

Thornton, R. & MacDonald, M. C. (2003). Plausibility and grammatical agreement. *Journal of Memory and Language*, 48, 740-759.

Thornton, R., MacDonald, M. C., & Arnold, J. E. (2000). The concomitant effects of phrase length and informational content in sentence comprehension. *Journal of Psycholinguistic Research*, 29, 195-203.

Thornton, R., MacDonald, M. C., & Gil, M. (1999). Pragmatic constraints on the interpretations of complex noun phrases in Spanish and English. *Journal of Experimental Psychology: Learning, memory, and Cognition*, 25, 1347-1365.

Toole, J. (1996). The effect of genre on referential choice. In T. Fretheim & J. K. Gundel (Eds.), *Reference and referent accessibility* (pp. 263-290). Amsterdam: John Benjamins.

Traxler, M. J. (2005). Plausibility and verb subcategorization in temporarily ambiguous sentences: Evidence from self-paced reading. *Journal of Psycholinguistic Research*, 34, 1-30.

Traxler, M. J., Pickering, M. J., & Clifton, C., Jr. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, 39, 558-592.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language*, 35, 566-585.

Trueswell, J. C., Sekerina, I. A., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, 73, 89-134.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, 33, 285-318.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preferences from Garden-Paths. *Journal of Experimental psychology: Learning, memory, and Cognition*, 19, 528-553.

Tyler, L. K. & Marslen-Wilson, W. D. (1982). Processing utterances in discourse contexts: On-line resolution of anaphors. *Journal of Semantics*, 1, 297-314.

Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999a). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, 41, 147-182.

Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999b). When does gender constrain parsing? Evidence from ERPs. *Journal of Psycholinguistic Research*, 28, 565-571.

- Van Berkum, J. J. A., Brown, C. M., Hagoort, P., & Szwtserlood, P. (2003). Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology*, *40*, 235-248.
- Van Berkum, J. J. A., Koornneef, A. W., Otten, M., & Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research*, *1146*, 158-171.
- van den Broek, P. (1994). Comprehension and memory of narrative texts. Inferences and coherence. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 539-588). Sand Diego, CA: Academic Press.
- Van der Henst, J.-B. (2002). La perspective pragmatique dans l'étude du raisonnement et de la rationalité. *L'Année Psychologique*, *102*, 65-108.
- van Gompel, R. P. G. & Majid, A. (2004). Antecedent frequency effects during the processing of pronouns. *Cognition*, *90*, 255-264.
- van Gompel, R. P. G. & Pickering, M. J. (2001). Lexical guidance in sentence processing: A note on Adams, Clifton, and Mitchell (1998). *Psychological Bulletin & Review*, *8*, 851-857.
- van Gompel, R. P. G., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, *52*, 284-307.
- van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2000). Unrestricted race: A new model of syntactic ambiguity resolution. In A. Kennedy, R. Radach, D. Heller & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 621-648). Oxford: Elsevier.
- van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language*, *45*, 225-258.
- van Hoek, K. (1995). Conceptual reference points: A cognitive grammar account of pronominal anaphora constraints. *Language*, *71*, 310-340.
- Vigliocco, G. (2000). The anatomy of meaning and syntax. *Current Biology*, *10*(2), R78-80.
- Virtue, S., Haberman, J., Clancy, Z., Parrish, T., & Beeman, M. J. (2006). Neural activity of inferences during story comprehension. *Brain Research*, *1084*, 104-114.
- Vonk, W., Hustinx, L. G. M. M., & Simons, W. H. G. (1992). The use of referential expressions in structuring discourse. *Language and Cognitive Processes*, *7*, 301-333.
- Vosse, T. & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition*, *75*, 105-143.
- Walker, C. H. & Yekovitch, F. R. (1987). Activation and use of script-based antecedents in anaphoric reference. *Journal of Memory and Language*, *26*, 673-691.
- Walker, M. A. (1998). Centering, anaphora resolution, and discourse structure. In M. A. Walker, A. K. Joshi & E. F. Prince (Eds.), *Centering in discourse*. Oxford: Oxford University Press.
- Walker, M. A., Iida, M., & Cote, S. (1994). Japanese discourse and the process of centering. *Computational Linguistics*, *20*, 193-232.
- Walker, M. A., Joshi, A. K., & Prince, E. F. (1998). Centering in naturally-occurring discourse: An overview. In M. A. Walker, A. K. Joshi & E. F. Prince (Eds.), *Centering in discourse*. Oxford: Oxford University Press.

Walker, M. A. & Prince, E. F. (1996). A bilateral approach to Givenness: A hearer-status algorithm and a Centering algorithm. In J. K. Gundel & T. Fretheim (Eds.), *Referent and reference accessibility* (pp. 291-306). Amsterdam, Ph.: John Benjamins.

Warren, T. & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition*, 85, 79-112.

Wilson, D. (1998). Discourse, coherence and relevance: A reply to Rachel Giora. *Journal of Pragmatics*, 29, 57-74.

Wilson, D. & Matsui, T. (1998). Recent approaches to bridging: Truth, coherence, relevance. *UCL Working Papers in Linguistics*, 10, 173-200.

Wolf, F., Gibson, E., & Desmet, T. (2004). Discourse coherence and pronoun resolution. *Language and Cognitive Processes*, 19, 665-675.

Yang, C. L., Gordon, P. C., Hendrick, R., & Wu, J. T. (1999). Comprehension of referring expressions in Chinese. *Language and Cognitive Processes*, 14, 715-743.

Yekovitch, F. R. & Walker, C. H. (1978). Identifying and using referents in sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, 17, 265-277.

Yule, G. (1979). Pragmatically controlled anaphora. *Lingua*, 49, 127-135.

Yule, G. (1981). New, current and displaced entity reference. *Lingua*, 55, 41-52.

Zagar, D., Pynte, J., & Rativeau, S. (1997). Evidence for early-closure attachment on first-pass reading times in French. *The Quarterly Journal of Experimental Psychology*, 50A, 421-438.