



HAL
open science

Quand la nominalisation envahit la subordination : étude sur ‘ce que’ dans les propositions subordonnées en français, selon une approche comparative, diachronique et expérimentale

Ingrid Konrad

► To cite this version:

Ingrid Konrad. Quand la nominalisation envahit la subordination : étude sur ‘ce que’ dans les propositions subordonnées en français, selon une approche comparative, diachronique et expérimentale. Linguistique. Université Sorbonne Paris Cité, 2019. Français. NNT : 2019USPCC072 . tel-02971026

HAL Id: tel-02971026

<https://theses.hal.science/tel-02971026>

Submitted on 19 Oct 2020

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

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

Thèse de doctorat
de l'Université Sorbonne Paris Cité
Préparée à l'Université Paris Diderot
École doctorale Sciences du langage (ED 132)
Laboratoire de Linguistique Formelle / Labex EFL

Quand la nominalisation envahit la subordination

*Étude sur 'ce que' dans les propositions subordonnées en français,
selon une approche comparative, diachronique et expérimentale*

par **Ingrid KONRAD**

Thèse de doctorat de Linguistique Théorique, Formelle et Automatique
Dirigée par Caterina DONATI

Présentée et soutenue publiquement à Paris le 22 juillet 2019

Anne ABEILLÉ, (PRÉSIDENTE DU JURY)	Professeure,	LLF, CNRS - Université Paris Diderot
Adriana BELLETTI,	Professeure,	Université de Genève et <i>Università di Siena</i>
Carlo CECCHETTO,	Directeur de recherches,	SFL, CNRS - Université Paris 8
Caterina DONATI,	Professeure,	LLF, CNRS - Université Paris Diderot
Nino GRILLO,	<i>Lecturer,</i>	<i>University of York</i>
Barbara HEMFORTH,	Directrice de recherches,	LLF, CNRS - Université Paris Diderot
Maria LOBO,	<i>Professora Associada,</i>	<i>Universidade NOVA, Lisbonne</i>

Quand la nominalisation envahit la subordination

Étude sur ‘ce que’ dans les propositions subordonnées en français, selon une approche comparative, diachronique et expérimentale

Résumé : Cette thèse porte sur la syntaxe des propositions subordonnées impliquant ‘ce’ en français, à savoir certaines complétives, certaines relatives et certaines interrogatives indirectes. Il semblerait que la nominalisation ait envahi la subordination en français et je m’interroge sur le statut et le rôle de l’élément nominal (D) ‘ce’ dans ces différentes constructions. Une grande partie de cette thèse se focalise sur l’ambiguïté atypique entre relatives à tête allégée (CITKO, 2004) et interrogatives indirectes introduites par ‘ce qui/ce que’. A partir d’une étude diachronique du 11^{eme} au 17^{eme} siècles, de tests de répétition de phrases avec des enfants français âgés entre 3 et 6 ans, d’une comparaison avec le portugais à la fois sur le plan syntaxique et sur le plan de l’acquisition, et d’une expérience de lecture par présentation auto-segmentée avec des adultes francophones, j’exclus l’hypothèse d’une incorporation de ‘ce’ à ‘que’ donnant lieu à un nouveau mot wh- /səkə/. Le comportement syntaxique atypique, en synchronie, des interrogatives indirectes introduites par ‘ce qui/ce que’ conduit à les analyser comme des DPs. Je propose qu’il s’agit en réalité de questions déguisées en relatives. Cette proposition m’amène à réfléchir à deux problèmes : le problème de la sélection et le problème de l’interprétation. En m’appuyant sur l’analyse des complétives en P+ce+que par ZARING, 1992 ainsi que plusieurs données dans d’autres langues, je mets en évidence l’existence de CPs nominalisés et j’affirme que la distinction syntaxique entre DP et CP n’est pas catégorique. Quant à la question de comment une relative avec ‘ce’ comme tête peut être interprétée comme une question, je postule que ‘ce’ est explétif, donc dénué de référentialité, tout comme dans les complétives. Les deux structures relatives (vraies relatives et questions déguisées en relatives) sont dérivées par appariement impliquant un opérateur nul. Mais la relation de prédication entre l’opérateur et la tête ‘ce’ ne se concrétise que dans les vraies relatives, où ‘ce’ est référentiel. Ainsi, les propositions subordonnées impliquant ‘ce que’ en français ont toutes une structure DP+CP. Toutefois, l’élément D ‘ce’ n’est référentiel et interprétable sur le plan sémantique que dans les relatives à tête allégée.

Mots-clés : Relatives à tête allégée, interrogatives indirectes, complétives, ambiguïté syntaxique, diachronie, acquisition, traitement, incorporation, nominalisation, explétif.

When nominalization pervades subordination

A study on 'ce que' in subordinate clauses in French, through a comparative, diachronic and experimental approach

Abstract : This dissertation deals with the syntax of subordinate clauses involving 'ce' in French, namely a particular kind of complement clauses, a particular kind of relative clauses and a particular kind of indirect questions. It seems that nominalization has pervaded subordination in French, and I am questioning the status and the role of the nominal element (D) 'ce' in these constructions. A large part of this dissertation focuses on the syntactic ambiguity between light-headed relatives (CITKO, 2004) and indirect questions introduced by 'ce qui/ce que'. On the basis of a diachronic study ranging from the 11th century to the 17th century, of a number of sentence repetition tasks performed on French children, aged from 3 to 6 years of age and of a self-paced reading experiment performed with French adults, I exclude that 'ce' and 'que' have incorporated and that /səkə/ is a wh-word. A closer scrutiny of the atypical syntactic behaviour, in synchrony, of indirect questions introduced by 'ce qu-' leads me to analyse them as DPs. My proposal is that these structures are indeed questions disguised as relative clauses. This proposal brings me to reconsider two issues : the selection issue and the interpretation one. Building on the analysis of complement clauses in P+ce+que given by ZARING, 1992, as well as several data in other languages, I emphasise the existence of nominalized CPs, and I claim that the syntactic distinction between DP and CP is not categorical. As for the issue of how a relative clause with 'ce' as a head can get to be interpreted as a question, I assume that 'ce' is expletive, and thus devoid of referentiality, just like it is in complement clauses. Both relatives with 'ce que' (true relatives and questions disguised as relatives) are generated through a matching derivation involving a null operator. But the predicative relation holding between the operator and the head 'ce' only takes a value in true relatives, where 'ce' is referential. Concluding, subordinate clauses involving 'ce que' in French all have a DP+CP structure. However, the D element 'ce' is semantically referential and interpretable only in light-headed relatives.

Keywords : Light-headed relatives, indirect questions, complement clauses, syntactic ambiguity, diachrony, acquisition, parsing, incorporation, nominalization, expletive.

Entre
Ce que je pense
Ce que je veux dire
Ce que je crois dire
Ce que je dis
Ce que vous avez envie d'entendre
Ce que vous croyez entendre
Ce que vous entendez
Ce que vous avez envie de comprendre
Ce que vous croyez comprendre
Ce que vous comprenez
Il y a dix possibilités qu'on ait des difficultés à communiquer.
Mais essayons quand même...

Bernard WERBER
(L'Encyclopédie du Savoir Relatif et Absolu)

Remerciements

Durant ce long périple qu'est le doctorat, mon chemin a croisé celui de nombreuses personnes, qui ont contribué, d'une façon ou d'une autre, à l'élaboration de la présente thèse.

En premier lieu, je tiens bien sûr à remercier Caterina DONATI, ma directrice, pour avoir accepté de diriger mon travail de thèse et pour m'avoir guidée comme il se doit pendant presque quatre années de recherches. Je lui suis extrêmement reconnaissante pour sa disponibilité, pour son soutien constant, pour ses précieux conseils et pour ses commentaires toujours très pertinents. J'ai beaucoup appris avec elle, grâce à sa générosité intellectuelle.

Je tiens également à exprimer ma profonde gratitude envers Carla SOARES JESEL, qui m'a codirigée pour le mémoire de Master 2 et qui m'a beaucoup soutenue en toutes choses depuis de nombreuses années. J'ai eu la chance d'enrichir mes connaissances sur l'acquisition du langage grâce à elle, domaine qui me passionne depuis longtemps. Je la remercie aussi pour ses commentaires et suggestions m'ayant permis de retravailler la partie acquisition de la présente thèse.

Je remercie Gilles BOYÉ et Catherine MATHON, tous deux maîtres de conférences à l'Université Bordeaux Montaigne, qui ont cru en moi et qui m'ont conseillée du mieux que possible sur le parcours à suivre après la licence. C'est en partie grâce à eux que je suis arrivée jusqu'ici.

Je remercie mes rapporteurs, Maria LOBO et Carlo CECCHETTO, ainsi que Barbara HEMFORTH, Nino GRILLO, Anne ABEILLÉ et Adriana BELLETTI, qui ont accepté de faire partie de mon jury de thèse.

La préparation de cette thèse a été financée par un contrat doctoral (2015-2018), et j'ai également bénéficié d'une bourse de mobilité de trois mois (de mai à juillet 2017), grâce au Labex EFL, qui m'a permis de faire une partie de mes recherches à l'*Universidade NOVA* à Lisbonne. Je remercie Maria LOBO pour m'avoir accueillie dans ce nouvel environnement de recherches et pour m'avoir guidée durant cette période.

De plus, grâce au projet portant sur « Le développement de la syntaxe dans le processus d'acquisition monolingue et bilingue du français et du portugais européen » (38033ZL) du Programme PESSOA 2017/2018 (Ministère de l'Europe et des Affaires Étrangères, Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche (France), Fondation pour la Science et la Technologie (Portugal)), j'ai eu la chance d'être financée pour pouvoir poursuivre mes investigations sur l'acquisition des dépendances A-barre en portugais européen, durant quelques jours à Lisbonne.

Je remercie Margarida TOMAZ, doctorante à l'*Universidade NOVA* à Lisbonne, pour avoir fait passer aux enfants portugais les tests que j'ai mis en place.

Je remercie Yair HAENDLER, ingénieur de recherche à l'Université Paris Diderot, pour son aide considérable pour l'analyse des résultats des expériences en acquisition.

Je remercie Adrian STAUB, Professeur de Sciences psychologiques et neurologiques à UMass, Amherst, et Francesca FOPPOLO, Professeure au Département de Psychologie à Milan, *Università degli studi di Milano Bicocca*, avec qui j'ai eu la chance et le plaisir de collaborer.

Je remercie Barbara HEMFORTH, directrice de recherches à l'Université Paris Diderot pour son aide pour les analyses statistiques, pour ses commentaires et suggestions, m'ayant permis de retravailler le chapitre sur le traitement des ambiguïtés syntaxiques impliquant 'ce que' en français.

Je remercie Fahima MAHOUCHE pour m'avoir assisté pour la passation de la première expérience de lecture par présentation auto-segmentée, ainsi que Antoine HÉDIER pour avoir créé les items et fait la passation de la deuxième expérience de lecture par présentation auto-segmentée.

Je remercie toutes les personnes qui ont librement accepté de participer à mes expériences.

Je remercie Lieven DANCKAERT, chargé de recherches à l'Université de Lille, et Heather BURNETT, chargée de recherches à l'Université Paris Diderot, de m'avoir conseillée et guidée pour l'étude diachronique que j'ai entreprise dans la présente thèse. Je tiens également à remercier Olivier BONAMI, directeur du Laboratoire de Linguistique Formelle et professeur à l'Université Paris Diderot, pour avoir réalisé les graphiques impliquant le calcul d'une moyenne mobile dans le chapitre portant sur l'étude diachronique.

Je remercie Anne ABEILLÉ, Professeure de syntaxe à l'Université Paris Diderot, pour ses nombreux commentaires m'ayant notamment permis de compléter la partie sur la description des relatives et des interrogatives indirectes en français.

Je remercie aussi les ingénieurs d'études Doriane GRAS et Rachel SHEN, pour leur aide pour les analyses statistiques, Loïc LIÉGEOIS, pour ses conseils pour les études de corpus, Johan FERGUTH et Achille FALAISE, pour avoir créé des scripts perl permettant d'extraire les résultats des expériences implémentées sur Ibex Farm.

Je remercie la directrice de l'école maternelle de Nogaro, Nathalie PIQUEMAL, ainsi que l'équipe enseignante, pour leur accueil, leur gentillesse, leur disponibilité, ainsi que les parents ayant accepté que je fasse passer des tests à leur(s) enfant(s), et tous les enfants qui ont participé. De même, je remercie les directeurs des écoles à Santa Iria de Azoia, au Portugal, ainsi que les équipes enseignantes, les parents pour leur accord, et les enfants qui ont participé aux tests.

Je remercie Joan BUSQUETS pour ses jugements sur les phrases en catalan, Israel DE LA FUENTE pour ses jugements sur les phrases en espagnol, Eunice FERNANDES, pour ses jugements sur les phrases en portugais, Carmen DOBROVIE-SORIN pour ses jugements sur les phrases en roumain.

Je remercie Alain ROUVERET, Professeur de syntaxe émérite de l'Université Paris Diderot, pour les discussions constructives que j'ai pu avoir avec lui.

Je remercie David PESETSKY et Michelle SHEEHAN, que j'ai eu l'opportunité de rencontrer lors d'écoles d'été en linguistique à Lisbonne, qui m'ont accordé du temps et qui m'ont donné des conseils pour mes recherches.

Je remercie Klaus ABELS, Dominique SPORTICHE, Nicolas GUILLIOT, entre autres, que j'ai eu la chance de rencontrer lors de conférences et avec qui j'ai eu l'opportunité d'échanger sur mon sujet de recherche.

Je remercie bien sûr mes collègues doctorant.e.s Céline POZNIAK, Patty GARET, Jiaying HUANG, Sacha BENIAMINE, Milica DENIC, Aixiu AN, Charlotte HAUSER pour leur disponibilité, leur patience, leur solidarité et leur amitié, sans oublier Wei XIANG, Shrita HASSAMAL, Yiqin QIU, Xiaoqian ZHANG, Rémi GODEMENT-BERLINE, Gabriel THIBERGE, Xinyue YU, Saïda LOUCIF, Anisia POPESCU, Suzanne LESAGE, Sandro CAPO CHICHI, Etienne RIOU, et tou.t.e.s les autres.

Enfin, je remercie mes proches qui m'ont soutenue et encouragée pendant ces quatre années, et qui m'ont donné l'affection et la force nécessaires pour surmonter les difficultés rencontrées. Ceci s'adresse avant tout à mes parents, bien sûr, et à mon compagnon, Cédric, mais aussi à Manon, Delphine, Sarah, et Yacine, qui ont été là pour moi.

Table des matières

Table des figures	xxi
Liste des tableaux	xxvii
Liste des abréviations	xxxii
Introduction	1
0.1 Préliminaires	1
0.2 Une approche comparative, diachronique et expérimentale	3
0.3 Organisation de la thèse	5
I Syntaxe des relatives à tête allégée et interrogatives indirectes introduites par ‘ce qu-’	7
1 Les constructions relatives : ‘ce’, une tête allégée	9
1.1 Les relatives : une définition générale	9
1.2 Les relatives : une bipartition classique	11
1.2.1 Relatives pleines	11
1.2.1.1 Définition	11
1.2.1.2 Compétition entre deux analyses syntaxiques	11
1.2.1.2.1 Les problèmes de connectivité et de modification	11
1.2.1.2.2 L’analyse d’appariement (<i>matching analysis</i>)	12
1.2.1.2.3 L’analyse par montée de la tête (<i>head-raising analysis</i>)	12
1.2.1.2.3.1 Avantages et problèmes	14
1.2.2 Relatives libres	17
1.2.2.1 Définition	17
1.2.2.1.1 Les principales propriétés des relatives libres	17
1.2.2.1.1.1 L’absence d’antécédent	17
1.2.2.1.1.2 La position du mot <i>wh-</i> : un débat	18
1.2.2.1.1.3 Les effets de <i>matching</i>	20
1.2.2.1.1.4 L’interprétation des relatives libres	22

1.3	Relatives à tête allégée, un troisième type?	23
1.3.1	Relatives semi-libres ou relatives à tête allégée : une définition . . .	23
1.3.2	Les relatives à tête allégée en français	24
1.3.2.1	Description générale	24
1.3.2.2	Les différentes constructions	24
1.3.2.2.1	Classification de PIERRARD, 1988	24
1.3.2.2.2	Les trois têtes allégées du français	25
1.3.2.2.2.1	Celui + REL	25
1.3.2.2.2.2	Ce + REL	27
1.3.2.2.2.3	Là + REL	28
1.3.2.3	Comparaison avec les relatives libres	28
1.3.2.4	Comparaison avec les relatives pleines	30
1.3.2.5	Propriétés sémantiques	31
1.3.2.6	Structure syntaxique	32
1.3.2.6.1	‘Celui’ et ‘ce’, un traitement distinct	32
1.3.2.6.2	La nature de ‘ce que’	33
1.3.2.6.2.1	La nature de ‘ce’	33
1.3.2.6.2.2	La nature de ‘que’	35
1.3.2.6.2.3	Analyse syntaxique préliminaire	36
1.3.3	Les relatives à tête allégée dans différentes langues	37
1.3.3.1	Dans les langues romanes	37
1.3.3.2	Dans des langues d’autres familles	50
1.4	Quelques conclusions sur la nature des relatives à tête allégée	51
2	Une ambiguïté surprenante avec les interrogatives indirectes	57
2.1	Ressemblances entre relatives à tête allégée et interrogatives indirectes en français	57
2.2	Une ambiguïté typique : le cas de l’anglais	58
2.2.1	<i>What</i>	58
2.2.2	<i>The labeling account</i>	58
2.2.3	<i>The embedding account</i>	59
2.3	Une ambiguïté atypique : le cas du français	60
2.3.1	Une confirmation de l’ <i>embedding analysis</i> ?	60
2.3.2	Le problème : ni ‘que’ ni ‘quoi’	61
2.3.3	Analyses antérieures	65
2.3.3.1	GOLDSMITH, 1981	65
2.3.3.2	KOOPMAN, 1982	67
2.3.3.3	ERIKSSON, 1982, JONES, 1996	69
2.3.3.4	MULLER, 1989, MULLER, 1996	69
2.3.3.5	SALVESEN, 2008	70
2.3.4	‘Ce’, un renforçateur de ‘que’?	71
2.3.5	‘Ce que’, une partie de ‘qu’est-ce que’?	72
2.3.6	L’hypothèse de l’incorporation	73

2.3.6.1	La nature de 'ce'	74
2.3.6.2	La nature de 'que'	74
2.3.6.3	Vers une grammaticalisation?	74
2.4	Du français aux autres langues romanes	81
2.5	Conclusion intermédiaire	85

II Évolution diachronique des relatives à tête allégée et interrogatives indirectes introduites par 'ce qu-' 87

3	Étude diachronique : une apparition tardive et soudaine	89
3.1	'Ce qu-' dans une perspective diachronique	89
3.1.1	NIÉGER, 1975	89
3.1.2	LEARD, 1987	90
3.1.3	PIERRARD, 1995	90
3.2	Étude diachronique sur 'ce qu-'	93
3.2.1	Objectifs, questionnements	93
3.2.2	Prédictions	93
3.3	Une étude en deux temps	93
3.3.1	Une étude globale à partir d'un corpus étiqueté	93
3.3.1.1	Le corpus	93
3.3.1.2	Ce que j'ai cherché	93
3.3.1.3	Les résultats	94
3.3.1.4	Discussion	96
3.3.2	Une étude plus spécifique à partir d'un corpus non étiqueté	96
3.3.2.1	Le corpus	96
3.3.2.2	Ce que j'ai cherché	97
3.3.2.3	Les résultats	98
3.3.2.3.1	Ancien français	98
3.3.2.3.2	Moyen français	103
3.3.2.3.3	Français du 16ème siècle	107
3.3.2.3.4	Français classique : 1ère moitié du 17ème siècle	109
3.3.2.3.5	Français classique : 2ème moitié du 17ème siècle	112
3.3.2.3.6	Vue d'ensemble : du 11ème au 17ème siècle	112
3.3.2.3.7	Analyse statistique	119
3.3.2.4	Discussion	119
3.3.2.4.1	Un changement soudain	119
3.3.2.4.2	Des données en faveur de l'hypothèse de l'incorporation?	121
3.3.2.4.3	Vers une nouvelle différenciation des deux systèmes?	122
3.4	Conclusion	122

III Acquisition des relatives à tête allégée et interrogatives indirectes : du français au portugais 125

4	Expériences en acquisition : la sensibilité des enfants au double statut de <i>ce qu-</i> , ainsi qu'à la nature de la tête / mot-wh	127
4.1	L'acquisition des relatives à tête allégée et des interrogatives indirectes en français	127
4.1.1	Une seule étude traitant en partie de 'ce que' : BENTEA, DURRLEMAN et RIZZI, 2016	128
4.1.1.1	Background : les effets d'intervention	128
4.1.1.2	L'expérience de Bentea et al.	130
4.1.1.2.1	Objectif général	130
4.1.1.2.2	Design	130
4.1.1.2.3	Participants	130
4.1.1.2.4	Les résultats	130
4.2	Les questionnements du point de vue de l'acquisition	132
4.3	Données de production spontanée	133
4.3.1	Le corpus	133
4.3.2	Les résultats à partir de la production des enfants	133
4.3.3	Les résultats à partir de l'input	136
4.4	Expérience 1 : test de répétition de phrases contenant des subordonnées introduites par <i>ce qu-</i>	139
4.4.1	Prédictions	139
4.4.2	Design expérimental	139
4.4.2.1	Participants	139
4.4.2.2	Matériel	140
4.4.2.3	Procédure	141
4.4.3	Résultats	142
4.4.3.1	Analyse quantitative : la performance des enfants	142
4.4.3.2	Analyse qualitative : les types d'erreur	145
4.4.3.2.1	L'omission de 'ce'	145
4.4.3.2.2	L'utilisation d'une forme wh- complexe	147
4.4.3.2.3	Les autres types d'erreur	149
4.4.4	Discussion	152
4.4.4.1	Des résultats contre l'hypothèse de l'incorporation	152
4.4.4.2	Absence d'asymétrie sujet-objet	154
4.5	Retour sur les données de production spontanée	155
4.5.1	Le corpus	155
4.5.2	Les résultats à partir de la production des enfants	155
4.5.2.1	L'omission de 'ce'	157
4.5.2.2	L'utilisation de <i>kesk</i>	157
4.5.3	Les résultats à partir de l'input	158
4.6	Conclusion intermédiaire	158

4.7	Expérience 2 : test de répétition de phrases contenant des subordonnées introduites par <i>qui</i>	160
4.7.1	Prédictions	160
4.7.2	Design expérimental	160
4.7.2.1	Participants	160
4.7.2.2	Matériel	160
4.7.2.3	Procédure	162
4.7.3	Résultats	162
4.7.3.1	Analyse quantitative : la performance des enfants	163
4.7.3.2	Analyse qualitative : les erreurs que font les enfants	165
4.7.3.2.1	Élimination de l'enchâssement	165
4.7.3.2.2	Interrogatives clivées	167
4.7.3.2.3	Relativisation explicite	170
4.7.3.2.4	Autres subordonnées	171
4.7.3.2.5	Omission de 'qui'	173
4.7.3.2.6	Les autres types d'erreurs	175
4.7.3.2.7	Bilan	176
4.7.4	Discussion	176
4.8	Conclusion du Chapitre	178
5	Du français au portugais : l'acquisition de 'o que'	181
5.1	L'acquisition des relatives à tête allégée et des interrogatives indirectes en portugais	182
5.1.1	Etude portant sur la compréhension des relatives libres introduites par 'quem' : COSTA, GRILLO et LOBO, 2012	182
5.1.1.1	Objectif général	182
5.1.1.2	Design et prédictions	182
5.1.1.3	Participants	184
5.1.1.4	Les résultats	184
5.1.1.5	Le traitement par les adultes	185
5.2	Données de production spontanée	186
5.2.1	Le corpus	187
5.2.2	Les résultats	187
5.3	Expérience 1 : test de répétition de phrases contenant des subordonnées introduites par <i>o que</i>	190
5.3.1	Prédictions	190
5.3.2	Design expérimental	190
5.3.2.1	Participants	190
5.3.2.2	Matériel	190
5.3.2.3	Procédure	192
5.3.3	Résultats	192
5.3.3.1	Analyse quantitative : la performance des enfants	193
5.3.3.2	Analyse qualitative : les types d'erreur	195

	5.3.3.2.1	L'omission de 'o'	195
	5.3.3.2.2	L'utilisation d'une forme wh-complexe	197
	5.3.3.2.3	Les autres types d'erreur	199
5.3.4		Discussion	203
	5.3.4.1	'O que', un cas de grammaticalisation?	203
	5.3.4.2	L'asymétrie sujet-objet retrouvée	205
5.4		Retour sur les données de production spontanée	206
	5.4.1	Le corpus	206
	5.4.2	Les résultats	206
	5.4.2.1	L'omission de 'o'	208
	5.4.2.2	L'utilisation de 'o que é que'	208
5.5		Conclusion intermédiaire	209
5.6		Expérience 2 : test de répétition de phrases contenant des subordonnées introduites par <i>quem</i>	210
	5.6.1	Prédictions	210
	5.6.2	Design expérimental	210
	5.6.2.1	Participants	210
	5.6.2.2	Matériel	210
	5.6.2.3	Procédure	212
	5.6.3	Résultats	212
	5.6.3.1	Analyse quantitative : la performance des enfants	212
	5.6.3.2	Analyse qualitative : les types d'erreur	214
	5.6.3.2.1	Un problème de cafouillage concernant les verbes matrices	216
	5.6.4	Discussion	219
5.7		Conclusion	219
5.8		L'Hypothèse de la Hiérarchie du Poids de la Tête	221

IV Traitement des ambiguïtés syntaxiques impliquant 'ce que' 225

6	Expériences de lecture par présentation auto-segmentée : évitement des lacunes et similarité de traitement entre relatives à tête allégée et interrogatives indi- rectes		227
6.1	Traitement de l'ambiguïté entre complétives et relatives		228
	6.1.1	Le Principe de Chaîne Minimale	228
	6.1.2	Une ambiguïté temporaire en français avec 'ce que'	229
	6.1.2.1	L'interprétation de complétive	229
	6.1.2.2	L'interprétation de relative	232
	6.1.3	Traitement de l'ambiguïté complétive/relative	232
	6.1.3.1	Expérience 1 : Test de jugement d'acceptabilité	233
	6.1.3.1.1	Prédictions	234
	6.1.3.1.2	Design expérimental	234

6.1.3.1.2.1	Participants	234
6.1.3.1.2.2	Matériel	234
6.1.3.1.2.3	Procédure	235
6.1.3.1.3	Résultats	236
6.1.3.1.3.1	Influence potentielle de la fréquence sur les notes attribuées pour chaque conti- nuation	239
6.1.3.1.3.2	Potentiel effet de l'âge sur l'acceptabilité des complétives	241
6.1.3.1.4	Discussion	242
6.1.3.2	Expérience 2 : Tâche de lecture par présentation auto- segmentée	243
6.1.3.2.1	Design expérimental	243
6.1.3.2.1.1	Participants	243
6.1.3.2.1.2	Matériel et procédure	243
6.1.3.2.2	Résultats	244
6.1.3.2.3	Discussion	247
6.2	Et quand les deux structures sont des dépendances <i>filler-gap</i> ? : traitement des relatives à tête allégée et interrogatives indirectes	248
6.2.1	Expérience de lecture par présentation auto-segmentée	249
6.2.1.1	Prédictions	249
6.2.1.2	Design expérimental	250
6.2.1.2.1	Participants	250
6.2.1.2.2	Matériel	250
6.2.1.2.3	Procédure	251
6.2.1.3	Résultats	252
6.2.1.4	Discussion	258
6.2.2	Emploi de 'qu'est-ce qu-'	260
6.2.2.1	Dans FrWaC	260
6.3	Conclusion	262

V Retour à la syntaxe 263

7	Quand des interrogatives se déguisent en relatives 265
7.1	Un comportement syntaxique inhabituel 266
7.1.1	L'extraction 266
7.1.2	Les effets de <i>matching</i> 269
7.2	Le problème de l'étiquette 272
7.2.1	La sélection 272
7.2.1.1	Vision traditionnelle 272
7.2.1.2	Une approche alternative : Pesetsky (1982) 273

7.2.2	Origine nominale des compléments et existence de CPs nominalisés	274
7.2.2.1	Origine nominale des compléments	274
7.2.2.2	Existence de CPs nominalisés	274
7.2.2.2.1	Propriétés nominales du CP en français	275
7.2.2.2.2	CPs nominalisés dans des langues autres que le français	277
7.2.2.2.2.1	CPs nominalisés en espagnol	277
7.2.2.2.2.2	CPs nominalisés en grec moderne	278
7.2.2.2.2.3	CPs nominalisés en persan	279
7.3	Le problème de l'interprétation	280
7.3.1	Le cas particulier des 'questions déguisées' (<i>concealed questions</i>)	281
7.3.2	Un 'ce' explétif	282
	Conclusion Générale et Perspectives de Recherche	287
	Bibliographie Générale	297
	Bibliographie des sources citées dans le Chapitre 3	311
	Annexes	313
A	Avis Favorable par le Conseil d'évaluation éthique pour les recherches en santé (CERES) de l'Université Paris Descartes	315
B	Liste des items du test de répétition de phrases contenant des subordonnées introduites par <i>ce qu-</i> en français (Chapitre 4, Section 4.4)	317
B.1	Items d'entraînement	317
B.2	Items expérimentaux	317
B.2.1	Relatives	317
B.2.1.1	Sujet	317
B.2.1.2	Objet	317
B.2.2	Interrogatives indirectes	318
B.2.2.1	Sujet	318
B.2.2.2	Objet	318
B.3	Distracteurs	318
C	Liste des items du test de répétition de phrases contenant des subordonnées introduites par <i>qui</i> en français (Chapitre 4, Section 4.7)	319
C.1	Items d'entraînement	319
C.2	Items expérimentaux	319
C.2.1	Relatives libres sujet	319

C.2.2	Interrogatives indirectes sujet	320
C.3	Distracteurs	320
D	Liste des items du test de répétition de phrases contenant des subordonnées introduites par <i>o que</i> en portugais européen (Chapitre 5, Section 5.3)	321
D.1	Items d'entraînement	321
D.2	Items expérimentaux	322
D.2.1	Relatives	322
D.2.1.1	Sujet	322
D.2.1.2	Objet	322
D.2.2	Interrogatives indirectes	322
D.2.2.1	Sujet	322
D.2.2.2	Objet	323
D.3	Distracteurs	323
E	Liste des items du test de répétition de phrases contenant des subordonnées introduites par <i>quem</i> en portugais européen (Chapitre 5, Section 5.6)	325
E.1	Items d'entraînement	325
E.2	Items expérimentaux	325
E.2.1	Relatives libres sujet	325
E.2.2	Interrogatives indirectes sujet	326
E.3	Distracteurs	326
F	Liste des items du test de jugement d'acceptabilité ainsi que de l'expérience de lecture par présentation auto-segmentée portant sur le traitement de l'ambiguïté complétives/relatives (Chapitre 6, Section 6.1.3)	329
F.1	Items d'entraînement	329
F.2	Items expérimentaux	329
F.3	Distracteurs	332
F.3.1	Concordance temporelle	332
F.3.1.1	Distracteurs grammaticaux	332
F.3.1.2	Distracteurs agrammaticaux	332
F.3.2	Îlots	332
F.3.2.1	Distracteurs grammaticaux	332
F.3.2.2	Distracteurs agrammaticaux	332
F.3.3	Items à polarité négative	333
F.3.3.1	Distracteurs grammaticaux	333
F.3.3.2	Distracteurs agrammaticaux	333

G	Liste des items de l'expérience de lecture par présentation auto-segmentée portant sur le traitement de l'ambiguïté interrogatives indirectes/relatives (Chapitre 6, Section 6.2.2)	335
G.1	Items d'entraînement	335
G.2	Items expérimentaux	335
G.3	Distracteurs	343
H	Analyses statistiques pour l'étude diachronique (Chapitre 3, Section 3.3.2.3.7)	345
I	Analyses statistiques pour le premier test de répétition en français, avec 'ce qui/ce que' (Chapitre 4, Section 4.4.3)	347
I.1	Performance générale	348
I.2	Types d'erreur	349
I.2.1	Omission de 'ce'	349
I.2.2	Emploi de <i>kesk</i>	350
I.2.3	Stratégie SVO (phrase simple)	351
I.2.4	Emploi de 'ce que' ou 'ce qu'il' à la place de 'ce qui' dans les dépendances sujet	352
J	Analyses statistiques pour le deuxième test de répétition en français, avec 'qui' (Chapitre 4, Section 4.7.3)	353
J.1	Performance générale	354
J.2	Types d'erreur	355
J.2.1	Élimination de l'enchâssement	355
J.2.2	Interrogatives clivées	356
J.2.3	Relativisation explicite	357
J.2.4	Autres subordinées	358
J.2.5	Omission du mot <i>wh-</i> 'qui'	359
K	Analyses statistiques pour le premier test de répétition en portugais, avec 'o que' (Chapitre 5, Section 5.3.3)	361
K.1	Performance générale	362
K.2	Types d'erreur	363
K.2.1	Omission de 'o'	363
K.2.2	Emploi de 'o que è que'	364
L	Analyses statistiques pour le deuxième test de répétition en portugais, avec 'quem' (Chapitre 5, Section 5.6.3)	365
L.1	Performance générale	366

M	Analyses statistiques pour le test de jugement d'acceptabilité sur l'ambiguïté complétive/relative	
	(Chapitre 6, Section 6.1.3.1.3)	367
M.1	Modèle général	368
M.1.1	Avec la condition a (COMPL) comme niveau de comparaison et avec les sujets comme variables aléatoires	368
M.1.2	Avec la condition b (REL.AMB) comme niveau de comparaison et avec les sujets comme variables aléatoires	369
M.1.3	Avec la condition b (REL.AMB) comme niveau de comparaison et avec les items comme variables aléatoires	370
M.2	Conditions COMPL et REL.AMB seulement	371
M.2.1	Effet de fréquence	371
M.2.2	Effet du groupe d'âge	372
M.2.2.1	Condition COMPL <i>versus</i> condition REL.AMB chez les moins de 35 ans	373
M.2.2.2	Condition COMPL <i>versus</i> condition REL.AMB chez les plus de 35 ans	374
N	Analyses statistiques pour la première expérience de lecture par présentation auto-segmentée, sur l'ambiguïté COMPL/REL	
	(Chapitre 6, Section 6.1.3.2.2)	375
N.1	Au niveau du point de désambiguïstation (codé 0)	377
N.1.1	Avec la Condition a (COMPL) comme niveau de comparaison	377
N.1.2	Avec la Condition b (REL.AMB) comme niveau de comparaison	378
N.2	Au niveau de la région 'post.disamb' (codée +1)	379
N.2.1	Avec la Condition a (COMPL) comme niveau de comparaison	379
N.2.2	Avec la Condition b (REL.AMB) comme niveau de comparaison	380
N.3	Au niveau de la région 'post.post.disamb' (codée +2)	381
N.3.1	Avec la Condition a (COMPL) comme niveau de comparaison	381
N.3.2	Avec la Condition b (REL.AMB) comme niveau de comparaison	382
O	Analyses statistiques pour la deuxième expérience de lecture par présentation auto-segmentée, sur l'ambiguïté REL/INT	
	(Chapitre 6, Section 6.2.1.3)	383
O.1	Première comparaison : REL-ce qui/ce que <i>versus</i> INT-ce qui/ce que	385
O.1.1	Au niveau de la région critique, SUB word (codée 0)	385
O.1.2	Au niveau de la région critique+1, DET/AUX	386
O.1.3	Au niveau de la région critique+2, N3/AUX	387
O.1.4	Au niveau de la région critique+3, SUB V	388
O.1.5	Au niveau de la région critique+4, PREP	389
O.2	Deuxième comparaison : INT-ce qui/ce que <i>versus</i> INT-quand	390
O.2.1	Au niveau de la région critique, SUB word (codée 0)	390
O.2.2	Au niveau de la région critique+1, DET/AUX	391

O.2.3	Au niveau de la région critique+2, N3/AUX	392
O.2.4	Au niveau de la région critique+3, SUB V	393
O.2.5	Au niveau de la région critique+4, PREP	394

Table des figures

2.1	Les différents introducteurs des questions directes et indirectes, à temps fini et temps non fini, en position de sujet, d'objet et d'oblique (GOLDSMITH, 1981 : 555).	65
2.2	Structure syntaxique des interrogatives indirectes introduites par 'ce que' en français et règle d'insertion de 'ce' d'après GOLDSMITH, 1981 : 562.	66
3.1	Pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) <i>versus</i> 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé), pour la période Ancien Français.	94
3.2	Pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) <i>versus</i> 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé), pour la période Moyen Français.	95
3.3	Pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) <i>versus</i> 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la période Ancien Français (<1330).	99
3.4	Pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) <i>versus</i> 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la période Moyen Français (1330-1500).	104
3.5	Pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) <i>versus</i> 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la période de la Renaissance (1501-1600).	108
3.6	Pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) <i>versus</i> 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la première moitié du 17ème siècle (1601-1650).	110
3.7	Pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) <i>versus</i> 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la deuxième moitié du 17ème siècle (1651-1700).	113

3.8	Proportion de ‘ce qu-’ en fonction de chaque verbe matrice, du 11ème jusqu’à la fin du 17ème siècle (moyenne mobile de 101 ans)	115
3.9	Nombre d’occurrence de ‘ce qu-’ dans les textes en fonction de chaque verbe matrice, du 11ème jusqu’à la fin du 17ème siècle (somme mobile de 101 ans)	116
3.10	Proportion de ‘ce qu-’ en fonction de chaque verbe matrice : zoom sur la période qui va de 1600 à 1700 (moyenne mobile de 21 ans)	118
4.1	La grammaire de l’adulte <i>versus</i> la grammaire de l’enfant : schéma récapitulatif présenté par FRIEDMANN, BELLETTI et RIZZI, 2009 : 84.	129
4.2	Images utilisées par BENTEÀ, DURRLEMAN et RIZZI, 2016 pour leur tâche de choix de personnage (p.27).	131
4.3	Proportion de réponses correctes pour les relatives objet, en fonction des deux variables manipulées : NP et animacité, et en fonction de l’âge (BENTEÀ, DURRLEMAN et RIZZI, 2016 : 29).	131
4.4	Proportion de répétitions correctes (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Le panneau de gauche montre la performance pour les interrogatives indirectes et celui de droite celle pour les relatives à tête allégée. Les barres bleues indiquent la performance pour les dépendances objet, et les oranges celle pour les dépendances sujet.	144
4.5	Proportion d’omission de ‘ce’ (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Le panneau de gauche montre la proportion d’omissions du démonstratif ‘ce’ dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion d’omissions du démonstratif ‘ce’ dans les dépendances objet, et les oranges celle dans les dépendances sujet.	146
4.6	Proportion d’utilisation de <i>kεsk</i> (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Le panneau de gauche montre la proportion d’utilisation de <i>kεsk</i> dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion d’utilisation de <i>kεsk</i> dans les dépendances objet, et les oranges celle dans les dépendances sujet.	148
4.7	Proportion de production de phrases simples de type SVO (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Le panneau de gauche montre la proportion de production de phrases simples de type SVO dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion de production de phrases simples de type SVO dans les dépendances objet, et les oranges celle dans les dépendances sujet.	150

4.8	Proportion de production de ‘ce que’ ou ‘ce qu’il’ (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Le panneau de gauche montre la proportion de production de ‘ce que’ ou ‘ce qu’il’ dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion de production de ‘ce que’ ou ‘ce qu’il’ dans les dépendances objet, et les oranges celle dans les dépendances sujet. . . .	151
4.9	Répartition des différents types d’erreurs avec ‘ce qu-’	153
4.10	Proportion de répétitions correctes (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Les barres jaunes indiquent la performance pour les interrogatives indirectes, et les roses celle pour les relatives libres. . . .	164
4.11	Proportion d’éliminations de l’enchâssement (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Les barres bleues indiquent la proportion d’éliminations de l’enchâssement dans les interrogatives indirectes, et les roses celle dans les relatives libres.	167
4.12	Proportion d’interrogatives clivées (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Les barres vertes indiquent la proportion d’interrogatives clivées dans les interrogatives indirectes, et les rouges celle dans les relatives libres.	169
4.13	Proportion de relativisation explicite (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Les barres oranges indiquent la proportion de relativisation explicite dans les interrogatives indirectes, et les jaunes celle dans les relatives libres.	172
4.14	Proportion de production d’autres subordonnées (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Les barres vertes indiquent la proportion de production d’autres subordonnées dans les interrogatives indirectes, et les roses celle dans les relatives libres.	174
4.15	Proportion d’omission du mot wh- ‘qui’ (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Les barres bleues indiquent la proportion d’omission du mot-wh ‘qui’ dans les interrogatives indirectes, et les oranges celle dans les relatives libres.	175
4.16	Répartition des différents types d’erreurs avec ‘qui’	177
5.1	Un exemple de paire d’images utilisée dans le test de compréhension des relatives par COSTA, GRILLO et LOBO, 2012 : 149.	183
5.2	Caractéristiques des enfants pour l’étude de production spontanée (CAMPOS, 2015a : 18. Les informations indiquées de gauche à droite sont respectivement la tranche d’âge, le nombre de participants, la moyenne d’âge, le genre et le niveau de scolarité.	187
5.3	Proportion de répétitions correctes (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Le panneau de gauche montre la performance pour les interrogatives indirectes et celui de droite celle pour les relatives à tête allégée. Les barres bleues indiquent la performance pour les dépendances objet, et les oranges celle pour les dépendances sujet.	194

5.4	Proportion d'omission du déterminant 'o' (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Le panneau de gauche montre la proportion d'omissions du déterminant 'o' dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion d'omissions du déterminant 'o' dans les dépendances objet, et les oranges celle dans les dépendances sujet.	196
5.5	Proportion d'utilisation de 'o que e que' (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Le panneau de gauche montre la proportion d'utilisation de 'o que e que' dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion d'utilisation de 'o que e que' dans les dépendances objet, et les oranges celle dans les dépendances sujet.	198
5.6	Proportion d'omission de 'o+que' (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Le panneau de gauche montre la proportion d'omission de 'o+que' dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion d'omission de 'o+que' dans les dépendances objet, et les oranges celle dans les dépendances sujet.	200
5.7	Proportion d'emploi de 'do que' (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Le panneau de gauche montre la proportion d'emploi de 'do que' dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion d'emploi de 'do que' dans les dépendances objet, et les oranges celle dans les dépendances sujet.	202
5.8	Répartition des différents types d'erreurs avec 'o que'.	203
5.9	Proportion de répétitions correctes (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Les barres jaunes indiquent la performance pour les interrogatives indirectes, et les roses celle pour les relatives libres. . . .	213
5.10	Nombres d'occurrences de cafouillage concernant les 4 verbes matrices introducteurs divisés en deux types, VERB.PLUS et VERB.ONLY pour chacun des 44 enfants.	217
5.11	Proportion de répétitions correctes pour chaque verbe matrice, selon le type de structures.	218
5.12	Proportion de répétitions correctes (avec erreur-type ± 1) pour les dépendances sujet. Sur l'axe x, l'âge est représenté en années. Les panneaux de gauche montrent la performance des enfants français (bleu) et ceux de droite la performance des enfants portugais (rouge). Les panneaux du haut montrent la performance pour les relatives et ceux du bas celle pour les interrogatives indirectes (lignes continues pour 'qui'/'quem', c'est-à-dire pour les relatives libres, et lignes discontinues pour 'ce qui'/'o que', c'est-à-dire pour les relatives à tête allégée).	221

6.1	Résultats du test de jugement d'acceptabilité : notes moyennes attribuées pour chaque condition sur une échelle de 0 à 10.	236
6.2	Notes moyennes pour chaque verbe ambigu, pour les conditions (a) et (b). La ligne de régression montre la corrélation négative à travers les mêmes verbes dans les deux conditions.	238
6.3	Temps de lecture moyens pour les mots correspondant au point de désambiguïisation dans chaque condition, appelée 'disamb', et pour les régions précédant ce point ('pre.pre.disamb' et 'pre.disamb') et celles suivant ce point ('post.disamb' et 'post.post.disamb').	246
6.4	Expérience de lecture par présentation auto-segmentée - comparaison entre relatives et interrogatives introduites par 'ce qui/ce que' : temps de lecture moyens pour chaque condition (ce-que-INT, ce-que-REL, ce-qui-INT, ce-qui-REL) selon les régions de la phrase.	255
6.5	Expérience de lecture par présentation auto-segmentée - comparaison entre interrogatives indirectes introduites par 'ce qui/ce que' et interrogatives indirectes introduites par 'quand' : temps de lecture moyens pour chaque condition (ce-que-INT, quand-INT, ce-qui-INT) selon les régions de la phrase.	257
7.1	Structure d'une pseudo-relative d'après l'analyse de GRILLO et MOULTON, 2016 : 4.	295

Liste des tableaux

1.1	Propriétés syntaxiques des relatives avec tête allégée dans les langues romanes	40
1.2	Les différents introducteurs des différents types de relatives en français. . .	55
2.1	Possibilité d’avoir ‘élément nominal + C’ dans les questions extensionnelles et dans les questions intensionnelles dans les langues romanes	83
3.1	Nombre d’occurrences et pourcentages de ‘ce qu-’ (dans les relatives à tête allégée et les interrogatives indirectes) <i>versus</i> ‘qu-’ (dans les relatives libres et les interrogatives indirectes avec référent inanimé), pour la période Ancien Français.	94
3.2	Nombre d’occurrences et pourcentages de ‘ce qu-’ (dans les relatives à tête allégée et les interrogatives indirectes) <i>versus</i> ‘qu-’ (dans les relatives libres et les interrogatives indirectes avec référent inanimé), pour la période Moyen Français.	95
3.3	Présentation du corpus : dates et nombre de textes pour chaque période de la langue française.	97
3.4	Nombre d’occurrences et pourcentages de ‘ce qu-’ (dans les relatives à tête allégée et les interrogatives indirectes) <i>versus</i> ‘qu-’ (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la période Ancien Français (<1330).	98
3.5	Nombre d’occurrences et pourcentages de ‘ce qu-’ (dans les relatives à tête allégée et les interrogatives indirectes) <i>versus</i> ‘qu-’ (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la période Moyen Français (1330-1500).	103
3.6	Nombre d’occurrences et pourcentages de ‘ce qu-’ (dans les relatives à tête allégée et les interrogatives indirectes) <i>versus</i> ‘qu-’ (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la période de la Renaissance (1501-1600).	107
3.7	Nombre d’occurrences et pourcentages de ‘ce qu-’ (dans les relatives à tête allégée et les interrogatives indirectes) <i>versus</i> ‘qu-’ (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la première moitié du 17ème siècle (1601-1650).	109

3.8	Nombre d'occurrences et pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) <i>versus</i> 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la deuxième moitié du 17ème siècle (1651-1700).	112
4.1	Présentation des données concernant les quatre enfants pour l'étude de production spontanée dans le corpus CoLaJE.	133
4.2	Nombre d'occurrences de <i>ce qui</i> et <i>ce que</i> dans la production spontanée d'Anaé, en fonction de son âge. Type de structure indiqué entre crochets : AMB = ambigu, INT = interrogative indirecte, -SUB = non enchâssé.	134
4.3	Nombre d'occurrences de <i>ce qui</i> et <i>ce que</i> dans la production spontanée de Théophile, en fonction de son âge. Type de structure indiqué entre crochets : AMB = ambigu, INT = interrogative indirecte, -SUB = non enchâssé, REL = relative.	134
4.4	Nombre d'occurrences de <i>ce qui</i> et <i>ce que</i> dans la production spontanée d'Antoine, en fonction de son âge. Type de structure indiqué entre crochets : AMB = ambigu, -SUB = non enchâssé, REL = relative, CLIV = clivée.	134
4.5	Nombre d'occurrences de <i>ce qui</i> et <i>ce que</i> dans la production spontanée de Madeleine, en fonction de son âge. Type de structure indiqué entre crochets : AMB = ambigu, INT = interrogative indirecte, -SUB = non enchâssé, REL = relative, PRES = présentative.	135
4.6	Nombre d'occurrences de <i>ce qui</i> et <i>ce que</i> dans l'input d'Anaé, selon le type de structure.	137
4.7	Nombre d'occurrences de <i>ce qui</i> et <i>ce que</i> dans l'input de Théophile, selon le type de structure.	137
4.8	Nombre d'occurrences de <i>ce qui</i> et <i>ce que</i> dans l'input de Madeleine, selon le type de structure.	138
4.9	Nombre d'occurrences de <i>ce qui</i> et <i>ce que</i> dans l'input d'Antoine, selon le type de structure.	138
4.10	Scores moyens en fonction du type de phrase (Expérience 1).	143
4.11	Nombre d'occurrences d'omission de 'ce' et de l'utilisation de 'qu'est-ce que' ou 'qu'est-ce qui' dans la production d'Anaé en fonction de son âge.	156
4.12	Nombre d'occurrences d'omission de 'ce' et de l'utilisation de 'qu'est-ce que' ou 'qu'est-ce qui' dans la production de Théophile en fonction de son âge.	156
4.13	Nombre d'occurrences d'omission de 'ce' et de l'utilisation de 'qu'est-ce que' ou 'qu'est-ce qui' dans la production de Madeleine en fonction de son âge.	156
4.14	Nombre d'occurrences d'omission de 'ce' et de l'utilisation de 'qu'est-ce que' ou 'qu'est-ce qui' dans la production d'Antoine en fonction de son âge.	156
4.15	Nombre d'occurrences de 'qu'est-ce que' et 'qu'est-ce qui' dans les propositions subordonnées, dans l'input d'Anaé, de Théophile, de Madeleine et d'Antoine.	158
4.16	Scores moyens en fonction du type de phrase (Expérience 2).	163

5.1	Résultats globaux de l'expérience de COSTA, GRILLO et LOBO, 2012 : 150. . .	184
5.2	Résultats de l'expérience de COSTA, GRILLO et LOBO, 2012 : 151 : performance des enfants de 4 ans.	184
5.3	Résultats de l'expérience de COSTA, GRILLO et LOBO, 2012 : 151 : performance des enfants de 5 ans.	185
5.4	Performance des adultes (pourcentage de réponses correctes) pour l'expérience de lecture par présentation auto-segmentée de COSTA, GRILLO et LOBO, 2012 : 155.	186
5.5	Nombre d'occurrences de <i>o que</i> dans les subordinées (relatives ou interrogatives) et dans les questions racines dans la production des 40 enfants. .	188
5.6	Scores moyens en fonction du type de phrase (Expérience 1).	193
5.7	Nombre d'occurrences d'omission du déterminant 'o' et de l'utilisation de 'o que é que' dans la production des 40 enfants.	207
5.8	Scores moyens en fonction du type de phrase (Expérience 2).	213
6.1	Note moyenne attribuée lors du test de jugement d'acceptabilité et nombre d'occurrences dans FrWaC pour chaque continuation possible (complétive ou relative), pour chacun des 8 verbes ambigus.	240
6.2	Notes moyennes attribuées par les deux groupes d'âge (moins de 35 ans ; plus de 35 ans) pour chaque condition.	241
6.3	Encodage des régions pour les analyses des items expérimentaux de la tâche de lecture par présentation auto-segmentée (exemple 1). Le point de désambiguïsation est indiqué en gras.	245
6.4	Encodage des régions pour les analyses des items expérimentaux de la tâche de lecture par présentation auto-segmentée (exemple 2). Le point de désambiguïsation est indiqué en gras.	245
6.5	Encodage des régions pour les analyses des items expérimentaux de l'expérience de lecture par présentation auto-segmentée (exemple 1). Le point critique (mot subordonnant) est indiqué en gras.	253
6.6	Encodage des régions pour les analyses des items expérimentaux de l'expérience de lecture par présentation auto-segmentée (exemple 2). Le point critique (mot subordonnant) est indiqué en gras.	254
6.7	Nombre d'occurrences de 'qu'est-ce que' et 'qu'est-ce qui' dans le corpus FrWaC en fonction du verbe matrice.	261
7.1	Notes moyennes attribuées à chaque type de phrase selon que l'on extrait un adjectif ou bien un argument.	268

Liste des abréviations

ACC	Accusatif	M	Masculin
AMB	Ambigu	N	Neutre / Nom
ART	Article	N'	Nom'
AUX	Auxiliaire	NEG	Négation
C	Complémenteur	NOM	Nominatif
C°	Complémenteur°	NP	Syntagme Nominal
C'	Complémenteur'	OBJ	Objet
CL	Clitique	Op	Opérateur
CLIV	Clivée	PF	Parfait
COD	Complément d'Objet Direct	PL	Pluriel
COI	Complément d'Objet Indirect	PP	Syntagme Prépositionnel
COMP	Complémenteur	PREP	Préposition
CP	Syntagme de Complémenteur	PRES	Présentative
D	Déterminant	PRO	Pronom / Grand PRO
D'	Déterminant'	PRON	Pronom
DAT	Datif	PRS	Présent
DEF	Défini	PRT	Particule
DEM	Démonstratif	PST	Passé
DET	Déterminant	PTCP	Participe
DP	Syntagme de Déterminant	Q	Question
EXPL	Explétif	REL	Terme Relatif / Relative
FinP	Syntagme de Finitude	S	Phrase
FUT	Futur	S'	Phrase'
GEN	Génitif	SBJV	Subjonctif
GRDV	Forme Verbale au Gérondif	SG	Singulier
IMP	Impératif	Spec	Spécificateur
INF	Infinitif / Interrogative	SUB	Mot Subordonnant / Subordonnée
INT	Interrogatif/Interrogative	T	Temps
IntP	Syntagme d'Interrogation	TP	Syntagme de Temps
IP	Syntagme de Flexion	V	Verbe
IPF	Imparfait	WH	Mot Interrogatif/Relatif

Introduction

0.1 Préliminaires

A l'origine de ce travail se trouve la volonté de comprendre le statut et le rôle de 'ce' dans les propositions subordonnées en français. On rencontre ce démonstratif dans quatre constructions enchâssées en français : les relatives, les interrogatives indirectes, les complétives et les exclamatives, comme illustré en (1).

- (1) a. **Relative**
Léa déguste **ce** que son mari a cuisiné.
- b. **Interrogative indirecte**
Léa se demande **ce** que son mari a cuisiné.
- c. **Complétive**
Léa s'attend à **ce** que son mari prépare le repas pour ce soir.
- d. **Exclamative**
C'est fou **ce** que son mari cuisine bien!

Il semble évident que 'ce' ne revêt pas le même sens et ne joue pas le même rôle dans ces quatre constructions, comme l'illustrent les périphrases qui suivent (2).

- (2) a. **Relative**
Léa déguste **la chose** que son mari a cuisiné.
- b. **Interrogative indirecte**
Léa se demande **quelle chose** son mari a cuisiné.
- c. **Complétive**
Léa s'attend au **fait** que son mari prépare le repas pour ce soir.
- d. **Exclamative**
C'est fou **comme** son mari cuisine bien!

Parmi ces quatre constructions, nous nous focaliserons dans la présente thèse sur les trois premières seulement, laissant ainsi de côté les exclamatives, dans lesquelles 'ce que' fonctionne comme un adverbe de quantité.

- (3) a. C'est fou ce qu'il a de chance d'être ici!
- b. C'est fou combien il a de chance d'être ici!

Dans les trois autres constructions, ‘ce que’ s’impose comme un substitut à ‘que’.

- (4) a. **Relative**
*Léa déguste que son mari a cuisiné.
- b. **Interrogative indirecte**
*Léa se demande que son mari a cuisiné.
- c. **Complétive**
*Léa s’attend à que son mari prépare le repas pour ce soir.

Comment ne pas être intrigué par cet élément D envahissant la subordination en français? Pourquoi le mot *wh-* ou le complémenteur ‘que’ ne peut-il pas à lui seul introduire certaines relatives libres, interrogatives indirectes, et complétives en français? En quoi ‘ce’ est-il nécessaire?

Plus spécifiquement, c’est l’ambiguïté syntaxique entre relatives et interrogatives indirectes introduites par ‘ce qu-’ qui a éveillé ma curiosité de linguiste. Cette ambiguïté, qui a souvent été mise de côté dans la littérature, requiert une explication syntaxique. Une proposition telle que (5), avec le démonstratif ‘ce’ suivi du complémenteur *qui/que*, ressemble clairement à une relative à tête allégée (CITKO, 2004). Pourtant, cette proposition peut être interprétée de deux manières : soit comme une relative à tête allégée, soit comme une interrogative indirecte.

- (5) ce que Cédric a écrit
- a. **Relative à tête allégée**
Je lis [ce que Cédric a écrit].
- b. **Interrogative indirecte**
Je me demande [ce que Cédric a écrit].

Un verbe tel que ‘lire’, qui sélectionne un DP, est satisfait avec une relative à tête allégée qui est un DP complexe. Au contraire, un verbe comme ‘se demander’, qui est interrogatif, devrait sélectionner un CP. Comment la sélection d’un verbe de ce type peut-elle être satisfaite avec une proposition introduite par un élément D? A la question du label vient s’ajouter la question du trait *wh-*. Certes, ‘ce’ est un élément D, mais en plus, il est non *wh-*. Quant à ‘que’ qui le suit directement, il est aussi non *wh-* car il s’agit ici du complémenteur. Si ni ‘ce’ ni ‘que’ n’est *wh-*, comment analyser alors ces mystérieuses interrogatives indirectes?

La même ambiguïté se retrouve avec l’anglais ‘what’ mais elle est bien moins surprenante.

- (6) what Cédric wrote
- a. **Relative libre**
I am reading [what Cédric wrote].
- b. **Interrogative indirecte**
I am wondering [what Cédric wrote].

‘What’ est un élément D, certes, mais il est wh-. Cette double nature lui permet naturellement d’introduire une relative libre ou une interrogative indirecte.

0.2 Une approche comparative, diachronique et expérimentale

Ma thèse s’inscrit dans le modèle de la grammaire générative, qui a été inaugurée par Noam Chomsky au milieu des années 50 (voir CHOMSKY, 1961, CHOMSKY, 1968, CHOMSKY, 1973, CHOMSKY, 1986a, CHOMSKY, 1988, CHOMSKY, 1993, CHOMSKY, 2002, CHOMSKY, 2011, CHOMSKY, 2013, entre autres). L’idée sous-jacente est que les phrases sont générées par un ensemble de procédures subconscientes qui font partie de nos capacités cognitives. Le but de la grammaire générative est de modéliser ces procédures. Cette modélisation se fait au moyen de règles grammaticales. Ces règles nous indiquent, pas à pas, comment assembler les mots pour former une phrase, et la comprendre.

Comme l’entreprise que je mène dans la présente thèse est essentiellement descriptive, je n’utilise pas les formalismes du minimalisme, version plus récente de la grammaire générative cherchant notamment à développer l’économie des dérivations et des représentations syntaxiques (voir CHOMSKY, 2000, CHOMSKY, 2008, CHOMSKY, 2014, entre autres), ni ceux de l’approche cartographique (voir RIZZI, 1997, RIZZI, 2001, RIZZI, 2004, CINQUE, 2006, CINQUE et RIZZI, 2008, entre autres), approche qui propose une structure très riche de projections fonctionnelles supposée être universelle. J’ai ainsi décidé de représenter un CP simple dans tous mes arbres syntaxiques.

Pour tenter de faire la lumière sur le problème théorique évoqué dans les Préliminaires, j’ai choisi d’explorer plusieurs terrains de recherche, ce qui rend ma démarche originale :

1. la diachronie

Explorer le développement diachronique de ‘ce que’ pourra permettre de mieux comprendre le comportement syntaxique de ‘ce que’ en synchronie. Pour citer LIGHTFOOT, 2018 : « *nothing in syntax makes sense except in the light of change* ». ¹

2. l’acquisition par les enfants

Explorer l’acquisition des relatives à tête allégée et interrogatives indirectes introduites par ‘ce qu-’ en français est également une source d’information cruciale. La façon dont les enfants acquièrent ces structures ainsi que la façon dont ils analysent la séquence ‘ce que’, comparée à un mot wh- sans élément D tel que ‘qui’, permettra de mieux comprendre le statut syntaxique de ‘ce’ dans les interrogatives indirectes.

3. le traitement par les adultes

Explorer le traitement des différentes structures impliquant ‘ce que’ pourra permettre

1. Il s’agit d’une paraphrase du célèbre adage de DOBZHANSKY, 1973 : « *nothing in biology makes sense except in the light of evolution* ».

de déceler s'il existe différents 'ce que' sur le plan syntaxique. Je me focaliserai en particulier sur deux ambiguïtés syntaxiques, celle entre complétives et relatives d'une part et celle entre relatives à tête allégée et interrogatives indirectes d'autre part. La première ambiguïté confrontera la structure sans lacune (complétive) avec une structure avec lacune (relative). Son traitement permettra de déterminer s'il existe une préférence pour l'une ou l'autre des structures. Quant au traitement de la deuxième ambiguïté, qui confrontera deux structures avec lacune, il permettra de vérifier si l'analyseur traite différemment les relatives à tête allégée et les interrogatives indirectes introduites par 'ce que'. Le traitement de ces dernières sera lui-même comparé au traitement des interrogatives indirectes introduites par un mot *wh-* sans élément D (comme 'quand'). Comme dans le cas de l'acquisition, l'objectif de cette dernière comparaison sera de voir si les interrogatives indirectes introduites par 'ce qu-', qui semblent atypiques, sont traitées de la même façon que des interrogatives indirectes typiques, introduites par un mot *wh*.

Par ailleurs, ma thèse, bien que se centrant sur le français, comporte une dimension comparative avec les autres langues romanes, et en particulier le portugais européen. L'exploration des propriétés syntaxiques de structures similaires en italien, espagnol, portugais, catalan et roumain, va me permettre de dégager des similarités et des différences particulièrement intéressantes entre le français 'ce que' et le portugais 'o que'. En effet, nous retrouvons en portugais la même ambiguïté qu'en français.

- (7) o que compraste
 'ce que tu as acheté'
- a. Adoro o que compraste.
 'J'aime ce que tu as acheté.'
- b. Pergunto-me o que compraste.
 'Je me demande ce que tu as acheté.'

Au-delà de la possibilité d'introduire une interrogative indirecte, 'o que' en portugais peut également introduire une interrogative directe, ce qui est impossible en français avec 'ce que'. Ceci suggère que 'o que' en portugais se rapproche davantage d'un élément *wh-* que 'ce que' en français.

- (8) O que compraste?
- (9) *Ce que tu as acheté?

Pour cette raison, j'ai décidé de comparer l'acquisition de 'ce qu-' en français avec celle de 'o que' en portugais, afin de déterminer si l'on relève un schème développemental similaire ou non entre les deux langues, en ce qui concerne les relatives à tête allégée et les interrogatives indirectes introduites par 'ce qu-'/'o que'.

0.3 Organisation de la thèse

Ma thèse est organisée en cinq parties : la Partie I, qui correspond à une description syntaxique des relatives à tête allégée et des interrogatives indirectes introduites par ‘ce qu-’ en français ; la Partie II, dans laquelle j’étudie l’évolution diachronique des mêmes structures en français ; la Partie III, qui présente des données d’expériences sur l’acquisition de ces structures par les enfants français âgés entre 3 et 6 ans, en comparaison avec l’acquisition de structures similaires par les enfants portugais de la même tranche d’âge ; la Partie IV, qui porte sur le traitement des ambiguïtés syntaxiques impliquant ‘ce que’ en français ; et la Partie V, dans laquelle je retourne à l’analyse syntaxique des interrogatives indirectes introduites par ‘ce que’.

Dans le Chapitre 1, je donne une définition générale des relatives et je présente la bipartition classique entre relatives pleines et relatives libres. Je passe ensuite à la description des relatives à tête allégée en français. En comparant leurs propriétés avec celles des relatives pleines d’une part et celles des relatives libres d’autre part, j’en déduis qu’elles constituent une classe à part au sein des relatives. Après avoir fait une comparaison avec des structures similaires dans d’autres langues romanes, je termine ce chapitre en avançant quelques conclusions sur la nature des relatives à tête allégée.

Dans le Chapitre 2, je présente l’ambiguïté typique entre relatives et interrogatives indirectes que l’on observe avec ‘what’ en anglais, avant de passer à l’ambiguïté atypique du français (avec ‘ce que’). Je discute ensuite plusieurs hypothèses sur l’analyse des interrogatives indirectes introduites par ‘ce qu-’ en français. Je termine par une comparaison avec des structures similaires dans plusieurs langues romanes.

Dans le Chapitre 3, je fais d’abord le point sur ce que nous savons sur ‘ce qu-’ dans une perspective diachronique, puis je souligne l’intérêt de mener une étude diachronique d’envergure sur ‘ce qu-’. Je décris ensuite les deux temps de mon étude diachronique : une étude globale à partir d’un corpus étiqueté dans un premier temps puis une étude plus spécifique à partir d’un corpus non étiqueté dans un deuxième temps.

Dans le Chapitre 4, je présente d’abord la littérature sur l’acquisition des relatives à tête allégée et des interrogatives indirectes en français. Je présente ensuite des données de production spontanée puis je décris le test de répétition de phrases contenant des subordinées introduites par ‘ce qu-’ que j’ai mis en place. Puis je compare les résultats de ce test avec un deuxième test de répétition, portant cette fois-ci sur des phrases contenant des subordinées introduites par ‘qui’, qui est clairement wh- et sans élément D.

Le Chapitre 5 est parallèle au Chapitre 4. Comme dans le cas du français, je présente d’abord la littérature sur l’acquisition des relatives à tête allégée et des interrogatives indirectes en portugais. Je présente ensuite des données de production spontanée puis je décris

le test de répétition de phrases contenant des subordonnées introduites par ‘o que’ que j’ai mis en place. Puis je compare les résultats de ce test avec un deuxième test de répétition, portant cette fois-ci sur des phrases contenant des subordonnées introduites par ‘quem’, qui est clairement wh- et sans élément D. A la fin de ce chapitre, avant de clôturer la partie acquisition, je propose l’Hypothèse de la Hiérarchie du Poids de la Tête, permettant de rendre compte du schème d’acquisition des différents types de relatives. Je propose également une adaptation de cette hypothèse aux interrogatives indirectes.

Le Chapitre 6 se centre sur le traitement de deux ambiguïtés impliquant ‘ce que’ en français. Je décris d’abord l’ambiguïté entre relatives et complétives et je présente les résultats d’un test de jugement d’acceptabilité ainsi que d’une tâche de lecture par présentation auto-segmentée que j’ai mis en place. Je présente ensuite une autre tâche de lecture par présentation auto-segmentée portant sur le traitement de l’ambiguïté entre relatives à tête allégée et interrogatives indirectes introduites par ‘ce qu-’.

Enfin, dans le Chapitre 7, je montre que les interrogatives indirectes introduites par ‘ce que’ se comportent en fait comme des relatives sur le plan syntaxique. Ceci m’amène à faire face à deux problèmes, du moins en apparence : le problème de la sélection et le problème de l’interprétation. Je montrerai d’une part que la distinction entre DP et CP n’est pas aussi catégorique qu’on pourrait le penser, et, d’autre part, je proposerai que ‘ce’ en français peut avoir un usage explétif, indépendamment de son usage référentiel.

Première partie

Syntaxe des relatives à tête allégée et
interrogatives indirectes introduites par
'ce qu-'

Chapitre 1

Les constructions relatives : ‘ce’, une tête allégée

1.1 Les relatives : une définition générale

Une relative est une proposition non racine à base verbale, qui modifie un syntagme nominal, et qui est généralement introduite par un complémenteur ou un terme relatif. Elle contient soit un élément vide (*gap*), comme en (1), soit un pronom résomptif, comme en (2). Le syntagme nominal (NP) qu’elle modifie est appelé ‘antécédent’ ou ‘tête’ de la relative. Le terme relatif, mais également l’élément vide (noté *e*) ou pronom résomptif à l’intérieur de la relative, entrent généralement dans une relation de coréférence avec la tête de la relative.

(1) La femme_{*i*} [qui *e_i* est assise à côté de John] est ma voisine.

(2) Cărțile_{*i*} [pe care_{*i*} le_{*i*}-am cumpărat] sunt pasionante.
livres-ART.DEF.NOM.PL PREP REL PRON-ai acheté sont passionnants
‘Les livres que j’ai achetés sont passionnants.’

L’exemple en (1) contient une relative indiquée entre crochets. Le NP modifié est ici ‘femme’, qui est interprété comme étant le sujet du verbe de la relative ‘est assise’. On postule donc l’existence d’un élément vide en position de sujet à l’intérieur de la relative. Dans ce cas, on parle de ‘relative sujet’. De plus, les relatives sujet étant introduites en français par le complémenteur, qui se trouve dans C, on admet généralement l’existence d’un opérateur nul dans la position de spécifieur de CP, coréférent à la tête de la relative. Dans l’exemple en (2), qui est en roumain, le NP modifié ‘cărțile’ est interprété comme l’objet direct du verbe de la relative, ‘am cumpărat’. Dans ce cas on parle de ‘relative objet’. Contrairement au français, la stratégie de relativisation employée en (2) en roumain implique un pronom résomptif, ‘le’, et non pas un élément vide. Ce pronom résomptif fait référence à la tête de la relative, qui s’accorde en genre et en nombre avec ce dernier. Ici, comme ‘cărțile’ est féminin pluriel, et que nous sommes dans un cas de relative objet, le pronom utilisé est un clitique accusatif féminin pluriel. Une traduction littérale en français serait : ‘Les livres que je les ai achetés sont passionnants.’ Notons que nous pouvons avoir un pronom résomptif

en français aussi, notamment dans les relatives introduites par 'dont' comme en (3), mais la stratégie la plus généralisée en français dans les relatives est celle avec élément vide.

- (3) une femme_i [dont_i il est clair qu'on a pas besoin d'elle_i]

En français, comme en roumain, les relatives sont postnominales, c'est-à-dire qu'elles suivent le NP qu'elles modifient. La stratégie de relativisation standard en français est de laisser un *gap*. Les relatives sujet et objet direct sont introduites par le complémenteur 'que' qui prend la forme 'qui' lorsque le *gap* est en position sujet (KAYNE, 1976).

- (4) La fille [que Léo a embrassée] est très belle.

- (5) La fille [qui est assise là-bas] est très belle.

Les fonctions de l'élément vide ne se restreignent pas aux fonctions de sujet et d'objet direct en français, mais comme il s'agit d'une présentation très générale des relatives, je ne ferai pas mention des autres fonctions ici.¹

Notons par ailleurs que les relatives peuvent être introduites soit par un complémenteur, soit par un pronom relatif. Plusieurs propriétés permettent de distinguer les deux types d'introducteurs. Dans le cas du français, 'lequel' par exemple, qui est un pronom relatif, ne se comporte pas de la même façon que 'dont'. C'est d'ailleurs pour cela que GODARD, 1988 analyse 'dont' comme un complémenteur. Contrairement au complémenteur, un pronom relatif est sensible à certaines propriétés de la tête de la relative. Par exemple, 'lequel' varie en genre et en nombre, alors que 'dont' est insensible aux propriétés de la tête : comparer les exemples en (6) et ceux en (7) (à partir de GODARD, 1988 : 74).

- (6) a. le voisin **auquel** je me suis adressé
b. les voisins **auxquels** je me suis adressé
c. la voisine à **laquelle** je me suis adressé

- (7) a. le voisin **dont** je me souviens
b. les voisins **dont** je me souviens

De plus, alors que les pronoms relatifs sont à la fois relatifs et interrogatifs, les complémenteurs sont seulement relatifs (GODARD, 1988 : 75).

- (8) Lequel de ces livres as-tu lu ?

- (9) *Dont te souviens-tu ?

Une autre propriété distinguant le pronom relatif du complémenteur est que seul le premier est compatible avec l'infinitif. Avec le complémenteur, le verbe de la relative doit obligatoirement être à temps fini.

1. Pour une description détaillée de la syntaxe des relatives en français, voir notamment GODARD, 1988.

- (10) Voilà le garçon auquel s'adresser.
 (11) *Voilà la chose dont se souvenir.

Enfin, seul le pronom relatif peut être utilisé dans des constructions avec *pied-piping*.

- (12) le pinceau avec l'aide duquel j'ai repeint le mur
 (13) *le pinceau avec l'aide dont j'ai repeint le mur.

'Qui' en français n'est pas toujours le complémenteur dans les relatives. Lorsqu'il est précédé d'une préposition justement, il s'agit du pronom relatif, qui est spécialisé pour les antécédents [+animé] (voir GODARD, 1988 : 74-75).

- (14) a. le voisin à qui j'ai raconté cette histoire
 b. *l'arbre sur qui je m'étais appuyé

1.2 Les relatives : une bipartition classique

Traditionnellement, on fait la distinction entre relatives pleines et relatives libres ou relatives sans antécédent. Compte tenu de la définition donnée en Section 1.1, il semble difficile de concevoir l'existence de relatives sans antécédent. Bien que cela puisse apparaître comme une contradiction théorique, nous savons que de telles structures existent.

1.2.1 Relatives pleines

1.2.1.1 Définition

On appelle relatives 'pleines' les relatives dont l'antécédent est un NP 'plein' de type D+NP, avec un déterminant et un nom réalisés, comme en (15).

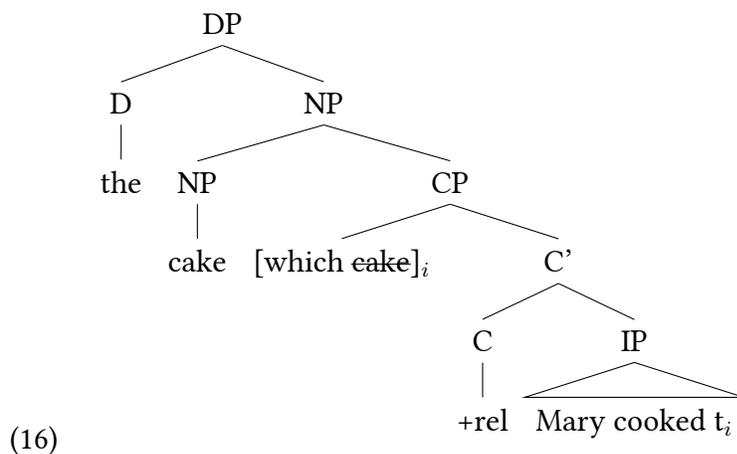
- (15) La fille [que Paul a vue hier] a disparu.

1.2.1.2 Compétition entre deux analyses syntaxiques

1.2.1.2.1 Les problèmes de connectivité et de modification Deux problèmes majeurs ont été soulevés dans le cadre de la grammaire générative en ce qui concerne l'analyse des relatives pleines. Le premier est celui de la connectivité. En effet, la tête de la relative, qui est comme un pivot, a un double rôle : en plus d'être la tête de la relative, elle revêt une fonction au sein de la relative, et cette fonction dépend de la position dans laquelle la tête de la relative est interprétée. Par exemple, en (15), la tête 'la fille' est interprétée comme étant l'objet du verbe 'voir' à l'intérieur de la relative. La question centrale est donc : quelle est la relation syntaxique entre la tête de la relative et le site de relativisation situé à l'intérieur de la relative ?

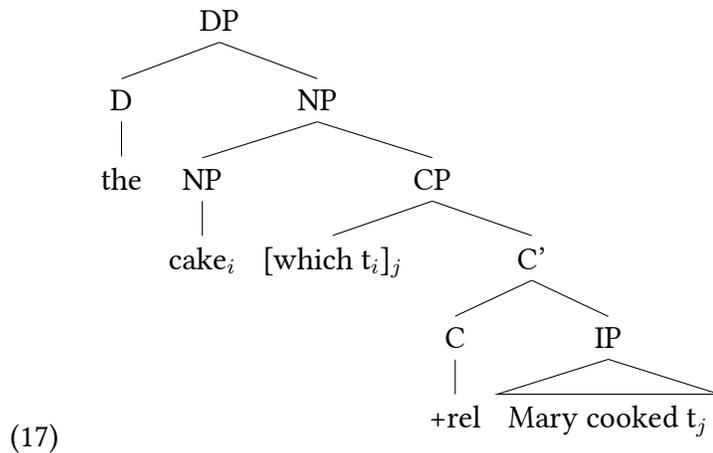
Le deuxième problème est celui de la modification. Nous avons vu qu'une proposition relative modifie la tête de la relative. La question ici est : comment la proposition relative est-elle reliée à la tête de la relative sur le plan syntaxique ? Est-ce qu'il s'agit d'une adjonction, d'une simple modification, ou bien d'une complémentation ? La relative, qui est un CP, est-elle adjointe/complément d'un NP ou d'un DP ?

1.2.1.2.2 L'analyse d'appariement (*matching analysis*) L'analyse traditionnelle des relatives est l'analyse d'appariement, connue sous le nom de *matching analysis* en anglais, et défendue entre autres par LEES, 1961, CHOMSKY, 1965, puis SAUERLAND, 2003. L'idée est qu'il y a deux têtes, une externe à la relative et une interne à la relative, les deux entretenant une relation de coréférence. La copie interne de la tête est supprimée sur la branche phonologique sous identité avec celle externe. Il est donc important de comprendre que, selon cette analyse, aucun mouvement de la tête n'est impliqué.

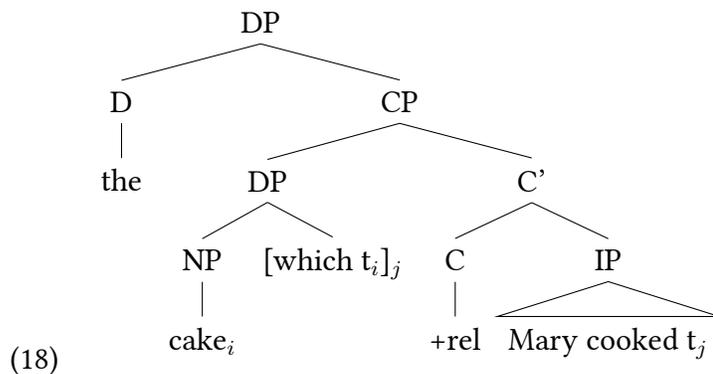


Notons qu'une autre version possible de la *matching analysis* implique le mouvement d'un opérateur en [Spec, CP] (*weak matching*).

1.2.1.2.3 L'analyse par montée de la tête (*head-raising analysis*) Une analyse alternative, proposée à l'origine par SCHACHTER, 1973, VERGNAUD, 1974, puis revisitée plus tard par KAYNE, 1994 et BIANCHI, 1999, BHATT, 2002 entre autres, est l'analyse par montée de la tête, connue sous le nom de *head-raising analysis* en anglais. Selon cette analyse, une seule tête est présente, et non pas deux comme dans l'analyse d'appariement. Cette tête unique est générée directement à l'intérieur de la relative, et se déplace à l'extérieur de la relative, par un mouvement de montée de la tête vers la périphérie gauche.



D'après la perspective antisymétrique de la syntaxe de KAYNE, 1994, la tête du DP, c'est-à-dire D, sélectionne la relative comme son seul complément. Le NP est généré directement à l'intérieur de la relative, laissant une lacune dans la place originelle à l'intérieur de la relative. La structure n'est donc pas DP+NP+CP (voir structure en (17)), mais DP+CP (voir structure en (18)).



Notons la présence de deux mouvements distincts : le premier est le mouvement du syntagme wh-, ici 'which cake', dans la position [Spec, CP]; le second, qui suit le premier, est le mouvement du nom, ici 'cake', jusque dans la position de tête de la relative.

Bien que cela puisse paraître surprenant qu'un D prenne pour complément un CP, des données dans plusieurs langues semblent justifier cela. Par exemple, le déterminant *derjenige* en allemand requiert la présence d'une relative (CP), mais pas d'un N/NP.

- (19) derjenige (Mann) *(der dort sitzt)
 the+that man who there sits
 'The very man (/person/one) who is sitting there'

La présence de relatives à tête interne dans des langues telles que le japonais, le lakhota ou le quechua, sont également un argument en faveur de D+CP. Dans ce type de relatives,

la tête de la relative est contenue à l’intérieur de la relative elle-même, au lieu d’être externe à la relative. La proposition est assemblée avec le déterminant pour former un constituant plus large qui fonctionne comme un DP d’un point de vue externe, indiquant que la relative est le complément du D (cf. exemple (20)) en Lakhota, tiré de WILLIAMSON, 1987, dont la structure est en partie représentée en (21)).

- (20) Mary owiza wa kage ki he ophewathu
 Mary quilt a make the I buy
 ‘I bought the quilt that Mary made.’

- (21) [_{DP}[_{CP} Mary[_{DP} owiza wa]kage]ki]he ophewathu

1.2.1.2.3.1 Avantages et problèmes Je vais présenter dans cette section les avantages de l’analyse par montée (à partir de ALEXIADOU, LAW, MEINUNGER et WILDER, 2000), qui sont donc des problèmes pour l’analyse d’appariement. Inversement, les problèmes de l’analyse par montée sont des avantages pour l’analyse d’appariement.

L’argument majeur en faveur de l’analyse par montée concerne les effets de reconstruction, qui apparaissent seulement lorsqu’il y a un déplacement via mouvement (CHOMSKY, 1993). Observons l’exemple en (22).

- (22) the picture of himself that John painted

‘Himself’ est un pronom réfléchi, et nous savons d’après le Principe A de la Théorie du Liage, que ce type de pronom doit avoir un antécédent local. Le fait que (22) soit grammatical prouve que ‘picture of himself’ est interprété dans la position à l’intérieur de la relative, comme objet du verbe ‘painted’. Cet effet de reconstruction est la preuve d’un mouvement de la tête de la relative.

- (23) the *picture of himself*_j that John painted *t*_j

De la même façon, l’analyse par montée peut rendre compte de l’agrammaticalité de (24).

- (24) *the *portrait of him*_j that John painted *t*_j

‘Him’ est un pronom non anaphorique qui obéit au Principe B de la Théorie du Liage : un pronom peut avoir un antécédent mais pas localement. En particulier, un pronom ne peut pas être c-commandé par un antécédent. L’agrammaticalité de (24) ne peut s’expliquer que si l’on admet une montée de la tête ‘portrait of him’. En effet, on observe ici un effet de reconstruction. ‘Him’ n’a pas l’air *a priori* d’être c-commandé par ‘John’, mais si ‘portrait of him’ est interprété dans la position à l’intérieur de la relative, en tant qu’objet du verbe ‘painted’, le pronom se retrouve c-commandé par ‘John’. Qui dit effet de reconstruction dit opération de mouvement. Nous avons donc ici un fort argument en faveur de l’analyse par montée de la tête. Naturellement, cela est problématique pour l’analyse d’appariement qui ne postule pas de mouvement de la tête de la relative.

Les évidences d'effets de reconstruction de la tête sont nombreux et ils ne concernent pas que les pronoms. Voilà en (25) un autre exemple sur la portée des quantifieurs, qui montre qu'un quantifieur dans la position de tête de la relative peut avoir une portée étroite sur un quantifieur situé à l'intérieur de la relative.

(25) I met the **two patients** that *every doctor* will examine today. [$\forall > 2$]

En effet, en (25), chaque docteur va examiner les deux patients en question. Encore une fois, cela montre qu'il y a eu un déplacement de la tête de la relative, à savoir 'two patients'. A comparer avec (26).

(26) I met **two patients** that *every doctor* will examine today. [$*\forall > 2$]

La différence entre (25) et (26) est que dans le premier cas, 'two' fait partie de la tête qui monte, car il y a le déterminant externe 'the', alors que dans le deuxième cas, 'two' est le déterminant externe et il ne fait pas partie de la tête qui subit le mouvement.²

Un autre problème pour l'analyse d'appariement concerne les phénomènes d'attraction inverse (*attractio inversa* ou *inverse attraction* : voir BIANCHI, 1999, BIANCHI, 2000, entre autres). En latin et en ancien haut allemand, entre autres, la tête de la relative pouvait prendre le cas de la position dans laquelle il est interprétée au sein de la relative. Cela est illustré par les exemples en latin en (27) et en ancien haut allemand (28), tirés respectivement de BIANCHI, 2000 : 59 et PITTNER, 1995 : 200.

(27) Urbem quam statuo vestra est.
ville-ACC REL-ACC (j')ai.fondé votre est
'La ville que j'ai fondée est à vous'

(28) Den liebsten bulen den ich hab der leit
le-ACC plus.cher-ACC ami-ACC REL-ACC 1SG ai PRON-NOM repose
beim wirt im keller.
PREP+ART auberge dans-la cave
'Le plus cher ami que j'ai est dans la cave de l'auberge.'

En (27), la tête de la relative 'urbem' n'est pas au cas nominatif, mais au cas accusatif, qui est le cas de la lacune à l'intérieur de la relative ('j'ai fondé la ville'). En (28), la tête de la relative 'bulen' n'est pas au cas nominatif, mais au cas accusatif, qui est là aussi le cas de la lacune à l'intérieur de la relative ('le plus cher ami que j'ai').

Ces phénomènes ont souvent été illustrés par des données de langues mortes, mais récemment, de nouvelles données concernant l'islandais et le finnois d'Ingrie ont été décrites respectivement par WOOD, SIGURÐSSON et NOWENSTEIN, 2017 et KHOLODILOVA, 2013. Si l'on admet que la tête de la relative est générée à l'extérieur de la relative, il semble très difficile de rendre compte de ces phénomènes d'attraction du cas.

2. Les données présentées ici ne sont problématiques que pour certaines versions de l'analyse d'appariement des relatives, en particulier celles impliquant un opérateur (*weak matching*).

Un autre argument en faveur de la montée de la tête concerne les expressions idiomatiques. Il est généralement admis que les parties nominales des expressions idiomatiques sont générées comme compléments du verbe de l'expression, mais elles ne peuvent pas être générées de façon indépendante (voir (29b)).

- (29) **Make headway** = 'faire des progrès'
- a. How much headway did they make?
 - b. *The headway was insufficient.
 - c. The headway that we made was insufficient.
 - d. *We made the headway that was insufficient.

La grammaticalité de (29c) et l'agrammaticalité de (29d) montrent que la tête de la relative, 'headway', est interprétée à l'intérieur de la relative, et a été déplacée dans la position de tête de la relative.

La possibilité d'extrapolation des relatives est problématique pour l'analyse par montée de la tête, en particulier dans l'approche de KAYNE, 1994. En effet, dans les cas d'extrapolation, la séquence D+N est séparée du reste de la construction relative. Or, D+N n'est pas un constituant selon cette approche.

- (30) We will discuss **the claim** tomorrow **that John made yesterday**.

Cela est au contraire prédit par l'analyse d'appariement, étant donné que D+NP est un constituant et que la relative CP est un adjectif, qui, en tant que tel, peut être extrapolé.

Par ailleurs, on peut se demander ce qui déclencherait la montée de la tête, mais aussi se questionner sur la nature du constituant qui peut monter en [Spec, CP]. Un déterminant n'est pas censé pouvoir se combiner directement avec un CP. BORSLEY, 1997 a sévèrement critiqué la version kaynienne de l'analyse par montée, en s'appuyant sur les relatives de l'anglais, en particulier les relatives introduites par 'that' en anglais : l'analyse de KAYNE, 1994 prédit l'existence de deux déterminants dans ce cas, alors qu'un seul déterminant est visible en surface.

- (31) a. *The man the that I saw.
b. *The the man that I saw.

De plus, un DP ne peut avoir qu'un seul cas grammatical d'assigné. Mais l'analyse kaynienne implique qu'un même DP ait deux cas, un assigné par le verbe de la proposition matrice et un assigné par le verbe de la relative.

1.2.2 Relatives libres

1.2.2.1 Définition

Les relatives libres³, contrairement aux relatives pleines, sont introduites par un mot *wh-* nu, sans antécédent explicite, comme on peut le voir en (32).

(32) Qui accusera le gouvernement sera puni.

Même si l'antécédent n'est pas visible, la proposition a bien l'interprétation d'une relative, qui correspondrait à la relative pleine en (33).

(33) La personne [qui accusera le gouvernement] sera punie.

Les relatives libres peuvent avoir une interprétation définie mais également une interprétation de choix libre (LARSON, 1987). Cette dernière interprétation est la seule disponible dans les relatives universellement quantifiées, introduites en anglais par un mot *wh-* auquel on attache le suffixe *-ever*.

(34) I will go **wherever** you will go.
je AUX-FUT aller n'importe-où tu AUX-FUT aller
'J'irai (partout) où tu iras.'

Le français ne dispose pas de suffixe aussi productif qu'en anglais. En effet, seul 'quiconque' existe (*quoiconque, *oùconque...).

(35) Quiconque accusera le gouvernement sera puni.

1.2.2.1.1 Les principales propriétés des relatives libres

1.2.2.1.1.1 L'absence d'antécédent

Les relatives libres sont qualifiées ainsi car aucun antécédent n'est visible en surface. La nature de ces propositions est ambiguë car, formellement, elles sont introduites par un mot *wh-* et ressemblent donc à une question, mais d'un point de vue distributionnel, elles apparaissent dans les mêmes contextes qu'un syntagme nominal (voir (36)).

(36) a. **Qui dit cela** ment.
b. Marie ment.

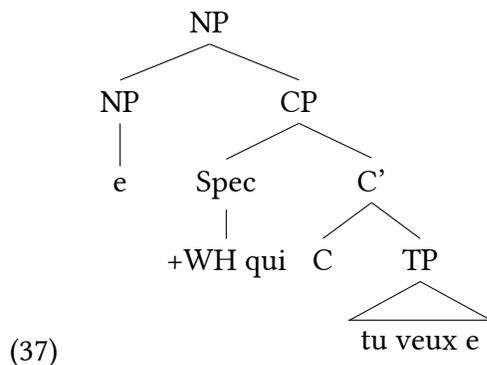
En (36a), la relative, indiquée en gras, peut être substituée par un DP ('Marie'), comme illustré en (36b).

3. Pour une description détaillée de la syntaxe des relatives libres et de leurs propriétés particulières telles que les effets de *matching*, voir notamment BRESNAN et GRIMSHAW, 1978 et RIEMSDIJK, 2000. En ce qui concerne les propriétés des relatives libres en français, voir HIRSCHBUHLER, 2016, livre publié à partir de sa dissertation datant de 1978.

1.2.2.1.1.2 La position du mot wh- : un débat

Du fait de l'absence d'antécédent, la question centrale qui se pose d'un point de vue syntaxique est la suivante : quelle est la position du mot wh- ? est-il dans sa position standard, c'est-à-dire dans la position de [Spec, CP] ou bien dans la position de tête de la relative ? Cela fut l'objet d'un débat vers la fin des années 80 et début des années 90, et la question reste ouverte. Les deux analyses concurrentes qui ont été proposées sont la *Comp proposal* et la *Head proposal*.

GROOS et VAN RIEMSDIJK, 1981 défendent l'hypothèse selon laquelle le mot wh- reste dans la position de spécifieur de CP. La position d'antécédent n'est donc remplie par aucun matériel lexical. Selon cette proposition, il n'y a pas de grande différence, du point de vue structural, entre les relatives libres et les relatives pleines (introduites par un constituant relatif). La structure serait donc la suivante :



Le statut de la tête vide est problématique. Une question qui découle directement de cette analyse est : quelle est la nature de l'élément nul ? comment le caractériser ? GROOS et VAN RIEMSDIJK, 1981 postulent que la position de l'antécédent, à savoir NP⁴, est occupée par un élément vide de nature pronominale, soit *pro*. Cette analyse ne permet pas de rendre compte aisément des phénomènes de *matching* observés dans les langues⁵. Si 'qui' est dans le spécifieur de CP, comment expliquer qu'il doive satisfaire la sélection du verbe de la proposition matrice ?

- (38) a. *J'inviterai à qui tu lanceras un défi.
 b. Je parlerai à qui tu lanceras un défi.

Par ailleurs, puisque cette analyse est similaire à celle des relatives pleines, comment expliquer alors l'absence de *matching* dans les relatives pleines ?

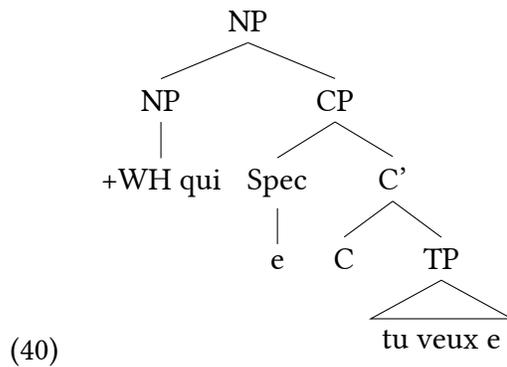
- (39) a. J'inviterai la personne à qui tu lanceras un défi.

4. Par souci de cohérence, je maintiens les étiquettes données par GROOS et VAN RIEMSDIJK, 1981, mais notons, que, depuis ABNEY, 1987, on parlerait plutôt de DP, et non plus de NP.

5. Nous allons présenter les effets de *matching* de façon plus détaillée dans la sous-section dédiée (1.2.2.1.1.3).

b. Je parlerai à la personne à qui tu lanceras un défi.

Selon l'hypothèse alternative, appelée *Head-proposal*, notamment défendue par BRESNAN et GRIMSHAW, 1978, l'élément introducteur occupe la position d'antécédent. Cette fois-ci, c'est donc la position Spec, CP qui est vide. On aurait alors la structure suivante :



Le mot *wh-* s'intègre donc dans la proposition principale, et non pas dans la proposition relative. Cette analyse permet de mieux rendre compte du phénomène de *matching*. En effet, si le syntagme *wh-* est dans la position de la tête de la relative, alors il doit satisfaire la sélection du verbe de la proposition supérieure comme le font les têtes de toutes les relatives, que ce soit en termes de catégorie ou en termes de cas.

Cette hypothèse semble compatible avec la *Raising Analysis* ou 'Analyse à Montée' des relatives. En effet, BRESNAN et GRIMSHAW, 1978 proposent que le syntagme *wh-* est généré dans la position d'antécédent à la base et que le mot *wh-*, qui est effacé *in situ*, monte ensuite dans la position de spécifieur de CP pour arriver finalement dans la position de tête de la relative.

Toutefois, cette hypothèse prédit l'existence d'une relative comme (41) (HIRSCHBÜHLER et RIVERO, 1983) : rien n'empêche de remplir la position C avec le complémenteur 'que' puisque tout le COMP est vide.

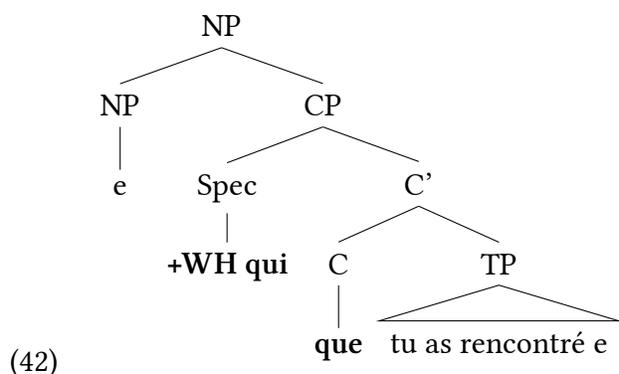
(41) *Qui que tu as rencontré est architecte.

Or la phrase en (41) est agrammaticale en français standard.

La *Comp-proposal*, elle, peut expliquer l'agrammaticalité de cette structure en appliquant le filtre du COMP doublement rempli (*doubly filled COMP filter*). Selon ce filtre, le COMP ne peut pas être rempli par deux éléments à la fois. Autrement dit les positions [Spec, CP] et C ne peuvent pas être toutes les deux remplies.⁶ Comme selon la *Comp-proposal*, le COMP

6. Notons que la nature précise de ce filtre, qui est une généralisation se basant sur de nombreuses données empiriques, reste encore à déterminer. Il est vrai que, dans les dépendances A-barre, comme les relatives et les interrogatives indirectes, les deux positions en question ne peuvent pas être toutes deux remplies.

est déjà rempli par le mot *wh-*, qui se situe dans la position de [Spec, CP], alors il devient évident qu'une phrase telle que (41) ne puisse pas être possible (voir structure en (42)). Cette donnée du français semble donc aller dans le sens de la *Comp proposal*.



Un autre point problématique de cette analyse est qu'elle enfreint le critère thêta (CHOMSKY, 1981)⁷, car elle assigne à un argument (le syntagme dans COMP) plus d'un rôle thématique (celui assigné par le verbe de la proposition matrice, et celui assigné par le verbe de la proposition relative).

1.2.2.1.1.3 Les effets de *matching*

Une des propriétés des relatives libres les plus abondamment décrites dans la littérature concerne les effets de *matching*. En effet, une des caractéristiques des relatives libres est la contrainte d'une compatibilité fonctionnelle entre le verbe de la proposition principale et le verbe de la proposition relative. Cette contrainte n'est pas présente dans les relatives pleines. Ce *matching* est de nature catégorielle, mais aussi de nature casuelle, dans le cas des langues casuelles.

- (43) a. Parle à qui tu t'adresses.
 b. *Parle à qui je regarde.
- (44) a. Elle a rencontré qui a fait un spectacle hier.

-
- (1) a. *J'apprécie qui que tu as invité.
 b. *Je me demande qui que tu as invité.

Toutefois, en français, dans les concessives par exemple, les deux positions peuvent être remplies sans que cela conduise à un résultat agrammatical.

- (2) a. Où que tu ailles, je viendrai avec toi.
 b. Quoi que tu fasses, je te défendrai toujours.

7. Theta-Criterion

Each argument bears one and only one theta-role, and each theta-role is assigned to one and only one argument. [CHOMSKY, 1981 : 36]

- b. *Elle a rencontré à qui tu parlais hier.
- (45) a. Parle à la femme que je regarde.
b. Elle a rencontré l'homme à qui tu parlais hier.

Les exemples donnés en (43) et (44) illustrent cette nécessité de compatibilité fonctionnelle. Le verbe 'parle' de la principale en (43a) sélectionne un COI, donc un PP formellement et le verbe 's'adresser' de la relative a les mêmes exigences, donc 'à qui', qui est un PP, satisfait la sélection des deux verbes. Quant au verbe 'a rencontré' de la principale en (44a), il attend un COD, donc un DP formellement, et le verbe 'a fait' attend un sujet de nature nominale, donc le mot wh- 'qui', qui est un DP, satisfait la sélection des deux verbes.

Cependant, en (43b) et (44b), nous avons deux cas d'incompatibilité fonctionnelle. En effet, en (43b), le verbe 'regarde' attend un COD, donc un DP. La présence de la préposition 'à' devant la relative pose alors problème. La préposition 'à' suivie du mot wh- 'qui' satisfait la sélection du verbe de la proposition supérieure, 'parler', mais pas celle du verbe de la relative, 'regarde'. La phrase est donc agrammaticale. En (44b), la situation est inversée, en ce sens que 'à qui' satisfait la sélection du verbe de la relative, 'parlais', mais pas celle du verbe de la principale, 'a rencontré', qui se construit avec un DP et non pas un PP. Cette incompatibilité conduit, là encore, à un résultat agrammatical. Rien de tel n'est observable dans les relatives pleines (45).

Le *matching* casuel fonctionne de façon similaire au *matching* catégoriel. Je vais illustrer ce phénomène en reprenant des exemples de VAN RIEMSDIJK, 2006, dans le cas de l'allemand.

- (46) Ich vertraue wem du vertraust.
1SG faire-confiance.1SG REL.datif 2SG faire-confiance.2SG
Je fais confiance à qui tu fais confiance.
- (47) *Ich nehme wem du vertraust.
1SG prendre.1SG REL.datif 2SG faire-confiance.2SG
*Je prends à qui tu fais confiance.

Le principe est le même. Le verbe 'vertrauen' (faire confiance) se construit avec un datif. En (46), le pronom relatif 'wem' est au cas datif donc il permet de satisfaire à la fois les exigences casuelles du verbe de la principale et celles du verbe de la relative. Par contre, en (47), le verbe 'nehmen' (prendre) se construit avec un accusatif, d'où l'agrammaticalité de la phrase.

Il est important de préciser que ces contraintes de *matching* ne sont pas universelles. Par exemple, en serbo-croate et en grec moderne, qui sont des langues *pro-drop*, le *matching* n'est pas obligatoire dans les relatives libres sujet (voir (48) et (49), à partir de IZVORSKI, 1996 : 3).

- (48) [Kome god budeš pomogao] osvojiće šampionat.
 who-DAT ever be-FUT helped will-win-3SG championship
 Whoever you help will win the championship.
- (49) [Oplanu milisi] tha kerdisi to vradio.
 who-GEN talk-2SG will win-3SG the prize
 Whoever you talk to will win the prize.

Le *mismatch* ici est de nature casuelle. En effet, le sujet d'une phrase est censé recevoir le cas nominatif. Mais ici, le mot *wh-*, qui est en position de sujet au sein de la proposition matrice, est marqué du cas datif en serbo-croate dans l'exemple (48) et du cas génitif en grec moderne dans l'exemple (49). Ainsi, dans ces deux exemples, le mot *wh-* introduisant la relative libre ne satisfait que les exigences casuelles du verbe situé à l'intérieur de la relative ('aider+DAT' en serbo-croate, 'parler+GEN' en grec moderne), mais pas celui du verbe de la proposition matrice.

1.2.2.1.1.4 L'interprétation des relatives libres

Du point de vue sémantique, les relatives libres ont une interprétation propre, qui se distingue de celle des relatives pleines. D'après PIERRARD, 1988, un trait saillant est celui de l'indéfinitude. L'introducteur de ce type de relative a la valeur d'un pronom indéfini. On peut rattacher cela à l'absence d'antécédent. Comme le fait remarquer PIERRARD, 1988, à un minimum de réalisation de l'antécédent correspond un maximum d'ouverture du sémantisme de l'introducteur. Mais on pourrait aussi penser que cette indéfinitude est liée au fait que l'introducteur soit le 'qui' interrogatif.

Il est vrai que la relative libre peut avoir une interprétation indéfinie (choix libre), et donc, dans ce cas, il n'est pas question d'individualiser une entité : il y a absence de spécification.

- (50) a. **Qui** veut réussir doit travailler dur.
 b. **Quiconque** veut réussir doit travailler dur.

Comme le montre l'exemple en (50a), l'introducteur désigne toute personne qui souhaite réussir, et non pas quelqu'un en particulier. Le mot *wh-* 'qui' a ici un trait [- défini]. On peut remplacer 'qui' par 'quiconque', sans que cela affecte le sens de la phrase (voir (50b)).

Mais nous savons que cette interprétation de choix libre n'est pas la seule interprétation possible.⁸ Les relatives libres peuvent aussi désigner une entité (un individu, un lieu, un moment...) et donc avoir une interprétation définie. Considérons l'exemple suivant :

- (51) **Qui** a dit cela est un menteur.

8. Voir notamment JACOBSON, 1995 et GODARD, A paraître pour la double interprétation des relatives libres (définie ou de choix libre).

En (51), le mot *wh-* ‘qui’ a un trait [+ défini] car il renvoie à une personne en particulier. La phrase ne signifie pas : toute personne qui a dit cela est un menteur, mais plutôt : la personne qui a dit cela est un menteur. Il y aurait donc un trait [+ spécifique] d’associé à l’introducteur ‘qui’. Ce dernier renvoie à une personne en particulier, mais dont on ne connaît pas ou ne précise pas l’identité. Il en est de même dans l’exemple (52).

(52) Je n’aime pas qui elle fréquente à la bibliothèque.

1.3 Relatives à tête allégée, un troisième type ?

1.3.1 Relatives semi-libres ou relatives à tête allégée : une définition

Comme le fait remarquer REBUSCHI, 2001, cette bipartition entre relatives pleines et relatives libres n’est pas suffisamment fine. A mi-chemin entre relatives pleines et relatives libres, il existe dans certaines langues des relatives qu’il appelle « semi-libres » (*semi-free relatives*). S’appuyant sur les données du basque, il les définit de la façon suivante : il s’agit de relatives dans lesquelles un élément phonétiquement réalisé en dehors de la relative est présent, mais ne porte aucune information lexicale.

(53) Gizonak e irrakuri du.en — .a
 homme-SG-k e lire AUX-en — SG
 ‘Ce que l’homme a lu.’

Notons que le basque est une langue SOV et que les relatives sont prénominales. Dans l’exemple (53)⁹, c’est la marque du singulier ‘.a’ qui se trouve dans la position d’antécédent de la relative, mais aucune information lexicale ne lui est rattachée.

Même si son analyse porte sur le basque, il donne également l’exemple espagnol suivant :

(54) El que ayuda a los ciegos.
 le COMP aide ACC les aveugles
 ‘Celui qui aide les aveugles.’

Cet exemple montre aussi que la position D est remplie mais qu’il n’y a pas de N explicite.

CITKO, 1999, CITKO, 2004 utilise une autre terminologie, qui me semble plus appropriée : elle parle de *light-headed relatives*, que l’on peut traduire par ‘relatives à tête allégée’. Elle affirme qu’en plus de la distinction entre relatives pleines et relatives libres, qui nous est familière, beaucoup de langues autorisent des relatives ayant pour antécédent un pronom démonstratif. Elle donne un exemple en polonais (CITKO, 1999, p.1) :

(55) Jan śpiewa to co Maria śpiewa.
 Jean chante DEM WH Marie chante
 ‘Jean chante ce que Marie chante.’

9. En basque, ‘k’ est la marque du sujet d’un verbe transitif ou d’un verbe inergatif.

Comme définition préliminaire, je dirai qu'une relative à tête allégée désigne toute relative restrictive présentant en surface un antécédent non plein, allégé, de type D, et donc plus réduit qu'un syntagme nominal composé d'un déterminant et d'un nom. Notons que cette définition exclut les relatives dites « pronominales », ayant pour antécédent un pronom personnel fort (souvent de première ou deuxième personne).

- (56) **Moi** [qui n'aime pas la musique électronique], je ne viendrai pas à ce festival à Amsterdam.
- (57) **Toi** [qui ne mens jamais], tu oses me raconter des salades à cet instant précis.

1.3.2 Les relatives à tête allégée en français

1.3.2.1 Description générale

En français, les relatives à tête allégée ont pour tête un démonstratif, comme en (58). En quelque sorte, ces relatives se situent à mi-chemin entre les relatives pleines et les relatives libres. Elles sont souvent classées au sein des relatives libres, dans une sous-classe, mais leurs propriétés et leurs particularités syntaxiques (que nous détaillerons dans les sous-sections suivantes) semblent indiquer qu'elles ne peuvent pas être assimilées aux relatives libres, et qu'une classe à part est nécessaire.

- (58) J'ai téléchargé **ce** [que tu m'as fait écouter].

1.3.2.2 Les différentes constructions

1.3.2.2.1 Classification de PIERRARD, 1988

PIERRARD, 1988 classe ces structures parmi les « relatives sans antécédent », mais il fait la distinction entre deux paradigmes : le paradigme synthétique (qui, que, quoi, où), qui correspond aux structures que j'appelle « relatives libres » et le paradigme analytique (celui qui, ce que, là où), qui correspond aux structures que j'appelle « relatives à tête allégée ».

Au sein même du paradigme que PIERRARD, 1988 qualifie d'analytique se trouvent deux paradigmes de démonstratifs : le paradigme de celui/celle/ceux/celles et le paradigme de celui/ce/là.

1. Premier paradigme : celui, celle, ceux, celles

- (59) J'aime celui [que tu as acheté e].
- (60) Je déteste celle [qui e est venue hier].
- (61) Achète ceux [qui e te plaisent].

(62) Revends celles [que tu trouves e inutiles].

2. Deuxième paradigme : celui/ce/là

(63) Celui [qui e veut maigrir] doit faire beaucoup de sport.

(64) Fais ce que tu veux e.

(65) Emmène-moi là où tu veux e.

Le démonstratif du premier paradigme (voir exemples (59) à (62)) indique le genre et le nombre, et peut être [+ animé] ou [- animé]. Comme on peut le remarquer, ‘celle’ en (60) a un trait [+ animé], alors que ‘celui’, ‘ceux’ et ‘celles’ en (59), (61) et (62) ont un trait [- animé]. De plus, tous ont un trait [+ spécifique] et [+ anaphorique]. Par contre, le démonstratif ‘celui’ du deuxième paradigme (voir exemple (63)) est toujours [+ animé] et [+ générique]. ‘Ce’ est un démonstratif ayant un trait [- animé] et un trait [+ générique] (voir (64)). Quant au locatif ‘là’, il se caractérise par les traits [+ locatif] et [+ générique] (voir (65)).

Bien sûr, cette classification que fait PIERRARD, 1988 est de nature sémantique. Mais je ne trouve pas ce classement pertinent. Peut-on affirmer qu’il y a deux pronoms ‘celui’ distincts ? Il semble plutôt que ce soit l’interaction de ‘celui’ avec la modalité de la phrase ainsi que le temps utilisé qui fait que l’on a soit un emploi générique, soit un emploi spécifique. Quand le verbe est au passé, comme en (59), ce n’est pas la lecture générique qui prend le dessus. Par contre, quand le verbe est au présent, comme en (63), et qu’il y a en plus un verbe modal comme ‘vouloir’, la lecture spécifique est écartée. La fonction grammaticale semble elle aussi jouer un rôle. En effet, le fait que la relative modifiant ‘celui’ soit en position de sujet au sein de la proposition matrice dans l’exemple (63) favorise l’emploi générique.

1.3.2.2.2 Les trois têtes allégées du français

Les relatives avec tête allégée en français regroupent différentes constructions. Comme nous l’avons évoqué dans la section précédente, la tête peut être un démonstratif, comme ‘celui’ ou ‘ce’, ou bien un locatif comme ‘là’. Quant au terme introducteur de la relative, il peut être le complémenteur ou un terme relatif. Cela va dépendre de la fonction de la lacune à l’intérieur de la relative.

1.3.2.2.2.1 Celui + REL

‘Celui’ dans la construction celui + REL est analysé, selon la tradition grammaticale, comme un pronom déterminé par une relative. Mais, d’après GOUGENHEIM, 1965 et PIERRARD, 1988 entre autres, ‘celui’ serait plutôt un déterminant ou un article. HIRSCHBÜHLER

et LABELLE, 1990, eux, suggèrent que 'celui' représente la tête d'un syntagme nominal modifié par une relative. Une autre possibilité serait de dire qu'il constitue un déterminant d'un syntagme nominal dont la tête n'aurait pas de réalisation phonétique, mais serait présente dans la structure sous-jacente. Est-ce que cette construction a un statut structural fondamentalement distinct des relatives pleines ?

'Celui' n'apparaît jamais sans modification restrictive, autrement dit on ne peut le trouver que comme tête de relative.

(66) *Celui aime bien les lasagnes.

Il semblerait que dans tous les contextes où il peut apparaître, il puisse commuter avec un SN.

(67) a. **Celui** que tu as invité hier est très gentil.

b. **Le garçon** que tu as invité hier est très gentil.

(68) a. J'aime bien **celle** que tu as essayée en premier.

b. J'aime bien **la robe** que tu as essayée en premier.

Comme le montrent HIRSCHBÜHLER et LABELLE, 1990, il ne semble pas avoir la distribution d'un article. Par exemple, il ne peut jamais être modifié par un adjectif.

(69) La bleue *versus* *Celle bleue

CORBLIN, 1995 (p.123) et COMBETTES, 1995 (p.94) indiquent également que (70) est agrammatical, comparé à (71).¹⁰

(70) *celui rouge

(71) le rouge

'Celui' requiert donc une détermination, comme c'est le cas des groupes nominaux du type 'le seul N' ou 'le simple N'.

(72) a. *Je parlerai à la seule personne.

b. Je parlerai à la seule personne qui me respecte et m'a toujours respecté.

Nous rejoignons HIRSCHBÜHLER et LABELLE, 1990 et postulons que 'celui' a une valeur référentielle. Il s'agit donc bien d'un pronom. Il est un véritable antécédent dans le cas des relatives avec tête allégée. D'un point de vue sémantique, « il désigne un individu spécifique appartenant à un ensemble particulier, mais sans fournir d'informations permettant d'identifier cet individu » (HIRSCHBÜHLER et LABELLE, 1990, p.119).

10. Notons que l'on trouve plusieurs attestations de celui+ADJ sur le Web.

(1) Et **celle bleue** juste derrière **celle verte** qu'est ce que c'est ?

(Source : commentaire sur www.tinkercad.com)

1.3.2.2.2 Ce + REL

Le statut syntaxique de ‘ce’ n’est pas aussi évident, comparé à ‘celui’. D’ailleurs, HIRSCHBÜHLER et LABELLE, 1990 ne le cachent pas : « Quant à ‘ce’, l’analyse nous paraît plus ouverte. » (p. 110).

Contrairement à ‘celui’, qui n’apparaît jamais sans modification restrictive, ‘ce’ peut être employé devant un nom.

(73) Ce livre est à Lucas.

Il est faible et non autonome, car il ne peut pas apparaître seul ni être modifié par autre chose qu’une relative.

(74) *Ce de Julie.

‘Ce’ désigne un référent inanimé, il renvoie à une chose indéfinie et peut être remplacé par la tête pleine ‘la chose’.¹¹ Il peut être la tête d’une relative sujet (voir (75)), objet direct (voir (76)) ou génitive (voir (77)), sachant que chaque relative peut elle-même être l’objet (voir exemples en a) ou le sujet (voir exemples en b) au sein de la proposition matrice.

(75) a. J’écris ce [qui e te fait plaisir].
b. Ce [qui e te rend heureux] me rend heureuse.

(76) a. J’achèterai ce [que tu veux e].
b. Ce [que Tom t’a dit e] n’est pas vrai.

(77) a. J’ai raconté à Pauline ce [dont tu m’as parlé e hier].
b. Ce [dont tu prétends avoir besoin e] me paraît futile.

Lorsque la lacune est en position de sujet à l’intérieur de la relative, c’est la forme ‘qui’ du complémenteur qui est employée en français. Dans l’exemple en (75a), ce+REL est l’objet direct du verbe principal ‘J’écris’, alors que dans l’exemple en (75b), ce+REL est le sujet du verbe principal ‘rend’. Lorsque la lacune est en position d’objet direct à l’intérieur de la relative, c’est le complémenteur ‘que’ qui est utilisé en français. Dans l’exemple en (76a), ce+REL est l’objet direct du verbe principal ‘J’achèterai’, alors que dans l’exemple en (76b), ce+REL est le sujet du verbe principal ‘est’. Lorsque les verbes construisent leur complément avec la préposition ‘de’, le pronom relatif ‘dont’ est requis. Nous avons dans ce cas une relative génitive, mais sans relation de possession exprimée. Dans l’exemple en (77a), ce+REL est l’objet direct du verbe principal ‘J’ai raconté’, alors que dans l’exemple en (77b), ce+REL est le sujet du verbe principal ‘paraît’.

11. En ce qui concerne la forme forte de ‘ce’, ‘ça’, CORBLIN, 1995 parle de référence « indistincte ». Il utilise le même qualificatif pour les formes fortes ‘ceci’ et ‘cela’ dans CORBLIN, 1987.

1.3.2.2.3 Là + REL

Une troisième tête légère utilisée dans les relatives en français est 'là'. Son statut de tête peut être discutable étant donné que la tête d'une relative, par définition, est de nature nominale, et que 'là' est plutôt de nature adverbiale. La relative qui suit cette tête est une relative locative introduite par le pronom relatif 'où'.

(78) J'irai là [où tu iras e].

D'un point de vue sémantique, la phrase en (78) est similaire à la relative libre en (79).

(79) J'irai [où tu iras e].

Bien que cela soit étonnant, la présence du locatif 'là' ne semble pas apporter de nuance sémantique. Étant un déictique, on s'attendrait à ce qu'il ajoute un trait de spécificité, mais cela n'est pas le cas. Encore une fois, en fonction de la modalité, l'interprétation sera soit définie (comme en (80) et (81)) soit de choix libre (comme en (78) et (79)) et ce, que ce soit avec les relatives à tête allégée ou les relatives libres.

(80) Retrouve-moi là [où on s'est vus e la première fois].

(81) Retrouve-moi [où on s'est vus e la première fois].

Peut-être que la présence de 'là' introduit une différence de registre, ce dernier étant plus soutenu avec 'là' que sans.

1.3.2.3 Comparaison avec les relatives libres

Le but de cette sous-section ainsi que de la sous-section suivante est de montrer la nécessité d'envisager un troisième type pour les relatives avec tête allégée. En effet, malgré les points communs qu'elles partagent avec les relatives libres d'une part et les relatives pleines d'autre part, certaines différences font que les relatives avec tête allégée ne peuvent être strictement assimilées ni aux premières, ni aux deuxièmes. Commençons par la comparaison avec les relatives libres.

Comme le fait remarquer CИТКО, 2004, il y a des raisons de croire que les relatives avec tête allégée sont tout simplement un type de relative libre. La raison principale est le fait que l'antécédent ne soit pas 'plein'. En effet, il n'est pas constitué d'un élément D et d'un élément N, mais seulement d'un élément D. Dans certaines langues, ce sont les mêmes mots wh- qui sont utilisés dans les deux types de relatives. Les exemples polonais qui suivent le montrent :

(82) Jan śpiewa co Maria śpiewa.
 Jan sings what Maria sings
 'John sings what Mary sings.'

- (83) Jan śpiewa **to co** Maria śpiewa.
 Jan sings this what Maria sings
 ‘John sings what Mary sings.’

En (82), nous avons une relative libre, introduite par le mot *wh-* ‘*co*’. En (83), il s’agit d’une relative avec tête allégée. La tête est le démonstratif ‘*to*’ et ensuite c’est le même mot *wh-* (‘*co*’) qui introduit la relative. En français, la ressemblance entre les deux types de relatives n’est pas aussi flagrante. En effet, l’introducteur n’est pas le même. Observons les exemples suivants :

- (84) a. **Qui** veut réussir doit travailler dur.
 b. Invite **qui** tu veux.
- (85) a. Celui **qui** veut réussir doit travailler dur.
 b. Invite celle **que** tu veux.

Comme on peut le constater, l’introducteur des relatives libres a les mêmes propriétés que l’interrogatif ‘*qui*’ (voir (84)) : il est [+humain] et il est utilisé pour les dépendances sujet et les dépendance objet. Par contre, dans le cas des relatives avec tête allégée (voir (85)), l’introducteur est le complémentateur. On retrouve en effet l’alternance ‘*que/qui*’ selon la fonction au sein de la relative.¹²

Par ailleurs, sur le plan sémantique, les relatives à tête allégée semblent être proches des relatives libres. Nous avons vu plus haut que les relatives libres peuvent avoir une interprétation de choix libre ou bien une interprétation définie. On retrouve ces deux interprétations dans le cas des relatives à tête allégée (cf. Section 1.3.2.5 sur les propriétés sémantiques des relatives à tête allégée).

Un autre point divergent qui me semble important est l’absence d’effets de *matching* dans le cas des relatives avec tête allégée. Comme nous l’avons montré ci-dessus, dans les relatives libres, le mot *wh-* doit satisfaire la sélection (catégorielle et casuelle) à la fois du verbe de la principale et du verbe de la proposition relative.¹³ Ceci n’est pas vrai de l’introducteur des relatives avec tête allégée qui ne subit pas de telles contraintes.

- (86) a. Parle à qui tu t’adresses.

12. Voir KAYNE, 1976 pour une analyse détaillée de ces formes du complémentateur en français.

13. Dans certains cas en français, on peut avoir l’impression que cette nécessité de *matching* n’est pas requise dans les relatives libres, comme en (1).

(1) Parle avec qui je te demande.

En effet, le verbe ‘demander’ ne peut pas sélectionner la préposition ‘avec’. Mais cette absence d’effet de *matching* n’est qu’une illusion, étant donné qu’à l’intérieur de la relative, il y a ellipse du verbe ‘parler’, qui lui est satisfait par la sélection de la préposition ‘avec’.

(2) Parle avec qui je te demande ~~de parler~~.

b. *Parle à qui je regarde.

(87) Parle à celui que je regarde.

Une dernière différence que l'on peut citer, en s'appuyant sur CITKO, 2004, est que les relatives avec tête allégée sont incompatibles avec *-ever*.

(88) *Celui quiconque accuse le gouvernement sera puni.

1.3.2.4 Comparaison avec les relatives pleines

Nous venons d'argumenter dans la direction suivante : les relatives avec tête allégée ne peuvent pas être assimilées aux relatives libres. L'autre possibilité qui reste encore envisageable est de les assimiler aux relatives pleines. Cette deuxième option semble plus plausible car, après tout, même si la tête est allégée, il y en a quand même une. En ce qui concerne les effets de *matching* et l'incompatibilité avec *-ever*, les relatives pleines et les relatives avec tête allégée fonctionnent de la même façon.

(89) Parle à la fille que je regarde.

(90) *La personne quiconque accuse le gouvernement sera punie.

CITKO, 2004 signale qu'en polonais les relatives pleines et les relatives avec tête allégée diffèrent en ce qui concerne la sélection du pronom relatif. En français, nous avons vu que l'introducteur des relatives avec tête allégée (pour les fonctions sujet et objet) est le complément, comme dans les relatives pleines. Les exemples ci-dessous semblent indiquer qu'il y a une identité systématique entre l'introducteur d'une relative pleine et l'introducteur d'une relative avec tête allégée.

(91) a. Emmène-moi à l'endroit **où** tu veux.

b. Emmène-moi là **où** tu veux.

(92) a. La fille **dont** je t'ai parlé est médecin.

b. Celle **dont** je t'ai parlé est médecin.

c. Ce **dont** je t'ai parlé est important.

(93) a. Le camarade avec **lequel** il a joué est sympathique.

b. Celui avec **lequel** il a joué est sympathique.

Toutefois, en position oblique, lorsque la tête est 'ce', 'lequel' ne peut pas être utilisé : le résultat est agrammatical. Il faut alors utiliser le pronom 'quoi'. C'est exactement la situation inverse avec une tête pleine : 'lequel' est la seule possibilité ; l'emploi de 'quoi' est agrammatical.

(94) a. Le ballon avec **lequel** tu as joué [...]

b. *Le ballon avec **quoi** tu as joué [...]

c. *Ce avec **lequel** tu as joué [...]

d. Ce avec **quoi** tu as joué [...]

Notre comparaison entre les relatives avec tête allégée et les relatives libres d'une part, et entre les relatives avec tête allégée et les relatives pleines d'autre part, nous a conduit aux conclusions suivantes : les relatives avec tête allégée ne peuvent pas être assimilées aux relatives libres ; les relatives avec tête allégée sont plus proches des relatives pleines que des relatives libres.

1.3.2.5 Propriétés sémantiques

D'après PIERRARD, 1988, alors que les relatives libres tendent à exprimer une référence générique, les relatives avec tête allégée tendent à exprimer une référence spécifique. Et cela est dû à la présence du pronom démonstratif. En effet, comme le rappelle PIERRARD, 1988, tout pronom démonstratif implique dans son sémantisme une connexion référentielle avec un élément présent dans la situation d'énonciation. 'Celui' est donc prédisposé à produire des emplois référentiels spécifiques. Il est caractérisé par son sens de désignation.

Mais, de la même façon que les relatives libres introduites par 'qui' peuvent avoir une interprétation définie (équivalent à 'la personne qui/que') ou bien une interprétation de choix libre (équivalent à 'quiconque'), les relatives avec tête allégée introduites par 'celui' peuvent également avoir ces deux interprétations. Soit ces dernières expriment une référence indicative, due au sens de désignation du démonstratif, soit elles indiquent une référence descriptive (voir (96)). En effet, l'être ou l'objet auquel on réfère n'est pas nécessairement présent dans la situation de communication. L'emploi générique est alors autorisé. Même si syntaxiquement, les relatives avec tête allégée ne sont pas des relatives libres, elles le sont au niveau interprétatif (PIERRARD, 1988).

Reprenons les exemples (59) et (63), répétés ici en (95) et (96) par souci de commodité.

(95) J'aime celui que tu as acheté.

(96) Celui qui veut maigrir doit faire beaucoup de sport.

En (95), le pronom démonstratif 'celui' renvoie à un objet spécifique. Nous pouvons imaginer qu'il s'agit d'un jean. La phrase en (95) est remplaçable par la relative pleine en (97), sans que cela change le sens :

(97) J'aime le jean que tu as acheté.

Cependant, en (96), le pronom démonstratif 'celui' est générique. Comme nous l'avons fait remarquer plus haut, cela est en particulier dû à la présence de modaux tels que 'vouloir' et 'devoir' et à l'emploi du présent de l'indicatif. La phrase en (96) est alors équivalente à une relative libre introduite par 'qui' (98) ou 'quiconque' (99), voire par une relative pleine avec pour tête 'toute personne' (100).

(98) Qui veut maigrir doit faire beaucoup de sport.

(99) Quiconque veut maigrir doit faire beaucoup de sport.

(100) Toute personne qui veut maigrir doit faire beaucoup de sport.
Il s'agit d'emplois génériques.

Par contre, les relatives avec tête allégée précédées par 'ce' tendent à indiquer une référence « descriptive » (dans les termes de PIERRARD, 1988; on dirait plutôt « définie »). Le locatif 'là' réalise une référence spécifique définie qui indique un lieu particulier dont l'identification plus poussée n'est pas requise ou est assurée par la situation d'énonciation. Mais, comme nous l'avons vu, associé au pronom relatif 'où', il ne désigne pas forcément un endroit spécifique. En (78), répété en (101), le sens est : j'irai à n'importe quel endroit où toi tu iras. Cela tend vers une interprétation de choix libre : localisation vague ou abstraite. Mais en (80), répété en (102), le sens est : retrouve-moi à cet endroit précis, particulier où on s'est vus pour la première fois. Cela tend vers une interprétation définie : localisation précise et concrète.

(101) J'irai là où tu iras.

(102) Retrouve-moi là où on s'est vus la première fois.

1.3.2.6 Structure syntaxique

Comme nous venons de le voir, il y a deux éléments clés qui précèdent ou font partie d'une relative avec tête allégée en français : un élément démonstratif et le complémenteur. Lorsque l'on s'intéresse à la structure syntaxique de ce type de relatives, des questions surgissent, notamment :

- Faut-il analyser 'celui qui/que' ou 'ce qui/que' comme un tout indissociable, correspondant à un introducteur de relatives, ou bien faut-il dissocier ces deux éléments et les associer à deux positions différentes dans l'arbre ?
- Faut-il analyser 'celui' ou 'ce' comme antécédent de la relative ?
- Y a-t-il un NP vide ?
- Les relatives avec 'celui' et celles avec 'ce' ont-elles la même analyse ?
- Devrait-on considérer que ces relatives sont des DPs ou bien des CPs ?

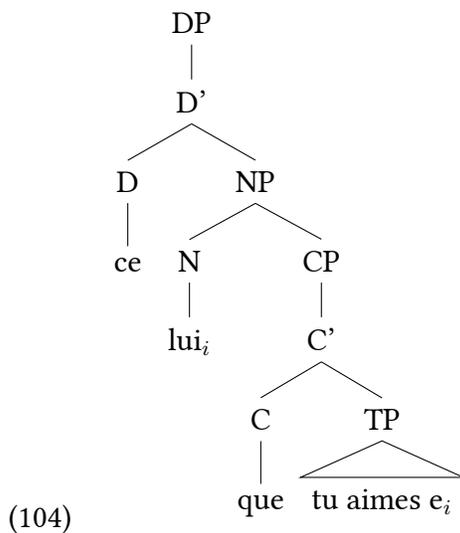
1.3.2.6.1 'Celui' et 'ce', un traitement distinct Il semble évident que 'celui' et 'ce' ne sont pas à traiter sur le même plan. Tout d'abord, 'celui' n'est pas faible comparé à 'ce'. De plus, 'celui' et 'ce' possèdent une structure interne bien différente : 'ce' est clairement plus léger que 'celui', il s'agit d'un D nu. Au contraire, 'celui' possède une structure plus complexe, étant donné que l'on peut le décomposer en ce+lui (de même pour 'celle' en c(e)+elle, 'ceux' en c(e)+eux, 'celles' en c(e)+elles).¹⁴ Il est donc constitué d'un élément D mais également d'un pronom fort, que l'on peut analyser comme un NP. Ainsi, analyser 'celui qu-'

14. GROSS, 1968 : 51 est le premier à avoir défendu cette analyse de 'celui' en ce+lui (D+N). VELAND, 1990 et ZRIBI-HERTZ, 1996 adoptent également cette analyse.

comme un tout indissociable n'est pas pertinent, et nous verrons dans le Chapitre 2 qu'il est au contraire envisageable d'analyser 'ce qu-' comme un tout (cf. Section 2.3.4.3.). Si 'celui' est dissociable de 'qu-', cela signifie que 'celui' dans celui+REL est à analyser comme la tête de la relative. Et, si 'celui' est décomposable en un D+NP, cela signifie qu'il s'agit en réalité d'une tête pleine, et non pas d'une tête légère. De plus, 'celui' est compatible avec 'lequel' dans les relatives obliques avec *pied-piping*, comme les têtes pleines.

- (103) a. Le stylo avec lequel j'écris est un cadeau de mon père.
b. Celui avec lequel j'écris est un cadeau de mon père.

La structure syntaxique de celui+REL serait alors la suivante :



Pour cette raison, je vais désormais me focaliser dans cette thèse seulement sur ce+REL, qui sont les 'vraies' relatives à tête allégée du français. Arrêtons-nous un instant sur la nature de 'ce' et la nature de 'que', avant de proposer une analyse syntaxique préliminaire de ce type de relatives en français.

1.3.2.6.2 La nature de 'ce que'

1.3.2.6.2.1 La nature de 'ce'

'Ce' est un démonstratif faible, qui a une fonction anaphorique ou déictique.

- (105) Le discours qu'il vient de faire est dénué de sens. Ce n'est pas intéressant.

Dans l'exemple en (105), 'ce' fait référence au 'discours'.

'Ce' comme pronom est faible et, de ce fait, il ne peut pas être employé seul (voir (39)), il ne peut pas être disloqué (voir (107)), il ne peut pas être modifié (voir (108)), il ne peut pas être complément de préposition (voir (109)), contrairement à sa variante forte 'ça', ou bien 'cela'.

(106) Qu'est-ce que tu veux? *Ce / ça.

(107) *Ce / ça, j'aime bien.

(108) Qu'est-ce que tu veux que je t'apporte? - *Ce / ça de rouge sur le meuble, tu vois?

(109) J'écris avec *ce / ça.

De plus, 'ce' ne peut pas non plus être en position d'objet direct (voir (110)), il est restreint à la position de sujet et il ne peut être que sujet du verbe 'être' (voir (111)).

(110) Je prends ça / *ce avec moi.

(111) a. C'est une bonne idée d'ajouter de la cannelle à ton dessert.

b. *Ce vaut la peine d'ajouter de la cannelle à ton dessert.

Dans les relatives, 'ce' a un trait [+ inanimé] et il peut alors être remplacé par un antécédent plein de type 'la chose' :

(112) J'aime ce que tu as acheté. → J'aime la chose que tu as achetée.

... mais pas systématiquement (PIERRARD, 1993).

(113) Je fais ce qui me plaît. → ?la chose, ?une chose, ?toutes les choses, ? quelque chose

(114) Je dépense ce que tu dépenses. → *la chose, la quantité (d'argent)

Il existe clairement un 'ce que' quantitatif non exclamatif, où 'ce' n'est pas référentiel, ce qui a amené certains auteurs à traiter 'ce' comme un quantificateur.¹⁵

(115) Je dépense en nourriture ce que tu dépenses en boisson.

Parfois, 'ce' semble pleinement garder sa valeur de démonstratif.

(116) Au nom de la France, je déclare formellement ce qui suit : [...].

Notons également que 'ce' peut aussi être employé comme antécédent phrastique. Mais dans ce cas, il s'agit de relatives appositives, et rappelons que dans ce travail nous nous centrons sur les relatives de type restrictif.

(117) Il est parti tôt, ce qui m'a déplu.

15. Sur les *amount relatives*, voir notamment GROSU et LANDMAN, 1998.

1.3.2.6.2.2 La nature de ‘que’

‘Que’ dans ‘ce que’ est bel et bien le complémenteur car il n’a pas la distribution du mot interrogatif, et ce, par rapport à deux aspects. Le premier aspect concerne la façon dont ‘que’ alterne avec ‘qui’ dans ‘ce que/ce qui’.

- (118) a. L’homme que j’ai rencontré est intelligent.
b. La fille qui mange est ma sœur.
- (119) a. Que veux-tu ?
b. Qui aimes-tu ?

KAYNE, 1976¹⁶ a été le premier à affirmer que le ‘que’ des propositions relatives n’était pas le mot interrogatif, mais le complémenteur¹⁷. Comme on peut le voir en (119b), le mot wh- ‘qui’ est [+ humain] et le mot wh- ‘que’ est [+ inanimé]. Par contre, le complémenteur ‘que’ est utilisé dans des propositions contenant une lacune en position d’objet et la forme ‘qui’ du complémenteur est employée dans des propositions contenant une lacune en position de sujet. Cette propriété du complémenteur est illustrée en (118).¹⁸

Si l’on observe maintenant la façon dont ‘ce que’ alterne avec ‘ce qui’ dans les relatives, on se rend compte que cette alternance dépend de la fonction de l’élément relativisé, et non pas du caractère animé du référent.

- (120) a. J’aime ce que tu as cuisiné.
b. Ce qui ne nous tue pas nous rend plus fort.

Comme le montrent les exemples en (120), ‘ce que’ est utilisé quand la lacune est en position d’objet au sein de la proposition subordonnée, et ‘ce qui’ est employé lorsque la lacune est en position de sujet au sein de la proposition subordonnée.

Le deuxième aspect concerne l’incompatibilité avec l’infinitif. Un fait bien connu est que le mot wh- ‘que’ est compatible avec l’infinitif (illustré en (121)) alors que le ‘que’ C° ne l’est pas (illustré en (122)).

- (121) a. Que faire ?
b. Je me demande que faire.

16. Voir aussi les travaux plus récents de KOOPMAN et SPORTICHE, 2008 et SPORTICHE, 2011.

17. Notons, de façon anecdotique, que Kayne analyserait désormais ‘que’ comme un pronom relatif et non plus comme un complémenteur, de la même façon que ‘that’ en anglais serait un pronom relatif (KAYNE, 2014).

18. Rappelons que le ‘qui’ des relatives pleines n’est pas systématiquement le complémenteur. En effet, après préposition, il s’agit du mot wh-, qui est spécialisé pour les DPs [+animé].

(1) le jeune homme avec qui elle est allée au cinéma

- (122) a. *Le garçon que aimer est intelligent.
 b. *Je crois que venir.

Concernant cet aspect, 'ce que/ce qui' dans les relatives se comporte encore une fois comme le complémenteur, et non pas comme le mot *wh-*.

- (123) *J'aime ce que cuisiner.

L'exemple en (123) montre que 'ce que/ce qui' n'est pas compatible avec l'infinitif.

Comme conclusion intermédiaire, on peut dire que 'ce' est un élément faible, nominal, qui semble garder sa valeur de démonstratif dans les relatives. Par ailleurs, nous avons deux arguments solides prouvant que 'que', toujours dans ces mêmes constructions, est le complémenteur, et non pas le mot interrogatif.

1.3.2.6.2.3 Analyse syntaxique préliminaire

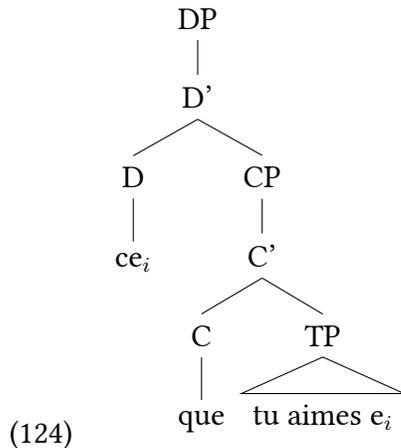
Lors de notre comparaison des relatives à tête allégée avec les relatives pleines et les relatives libres, nous avons conclu que les structures qui nous intéressent ici partageaient davantage de points communs avec les relatives pleines qu'avec les relatives libres. La question légitime qui se pose alors d'un point de vue syntaxique est la suivante : si les relatives introduites par 'ce que' sont des relatives pleines, quel est l'élément qui se déplace ?

Au moins quatre analyses différentes sont envisageables :

1. DP+CP, 'ce' étant l'élément nominal qui monte dans la position de tête de la relative
2. DP+NP+CP, avec 'ce' externe et un NP nul qui monte dans la position de tête de la relative
3. DP+CP, où CP est décomposable en Op+C
4. DP+NP+CP, où CP est décomposable en Op+C

Les analyses 1. et 2. sont des analyses par montée, alors que celles en 3. et 4. sont des analyses par appariement. Dans l'analyse par montée de la tête, c'est le nom qui monte dans la position de tête de la relative ou en Spec CP, selon que la structure est DP+NP+CP ou DP+CP. Dans les relatives à tête allégée, il n'y a qu'un élément D, sans N. Soit on postule qu'un élément D peut subir la montée (1.), soit on reste cohérent avec l'analyse pour les relatives pleines et on postule que c'est un N nul qui subit la montée (2.). Dans la version via appariement, c'est un opérateur *wh* nul qui monte en Spec CP. La différence entre 3. et 4. est la présence ou non, dans la structure sous-jacente, d'un NP avec un N nul dans la position de la tête.

Par souci d'homogénéité (entre les différents types de relatives) et par souci de parsimonie (éviter de postuler un élément vide), j'analyse les relatives à tête allégée comme en 1. : des DPs, dans lesquelles seul le D est rempli. Ce DP sélectionne un CP, dans lequel la tête C est remplie par le complémenteur et le complément IP par le reste de la relative. L'élément vide qui se trouve à l'intérieur de la relative entre dans une relation de coréférence avec le démonstratif. C'est l'élément D qui monte, qui se déplace, non pas en Spec CP, mais dans la position D.



Notons bien qu'il s'agit ici d'une analyse préliminaire, et que la dérivation de 'ce qu-' sera davantage discutée dans le Chapitre 2 à la lumière de l'ambiguïté qui existe avec les interrogatives indirectes. En particulier, nous réfléchirons à quel élément se déplace dans ce type de relatives et nous évaluerons la pertinence d'une analyse de 'ce qu-' comme un tout indissociable. En effet, le fait que 'ce que' puisse également introduire des questions indirectes semble remettre en cause cette analyse préliminaire et aller davantage en faveur du mouvement d'un opérateur nul en Spec CP...

1.3.3 Les relatives à tête allégée dans différentes langues

Dans cette sous-section, je vais étudier de plus près les propriétés syntaxiques des relatives à tête allégée en français, en les comparant avec celles des relatives à tête allégée dans d'autres langues romanes, afin de mieux comprendre la nature de ce type particulier de relatives peu décrit dans la littérature.

1.3.3.1 Dans les langues romanes

On trouve des relatives à tête allégée dans les autres langues romanes. Cette tête n'est pas systématiquement un démonstratif, elle peut aussi être un déterminant neutre, ou bien un déterminant défini masculin. Voici un exemple de relative de ce type dans quelques langues romanes, dont la traduction en français est 'J'aime ce que tu as acheté.'¹⁹

19. Les données présentées dans toute cette sous-section dans les différentes langues romanes sont des communications personnelles de locuteurs natifs venant de :

Portugais

- (125) Adoro o que compraste.
adorer-1SG ART.DEF que acheter.PST.2SG

Italien

- (126) Mi piace quello che hai comprato.
1SG.DAT aimer-3SG DEM que AUX.PRS.2SG acheter.PST.PTCP
- (127) Mi piace ciò che hai comprato.
1SG.DAT aimer-3SG DEM que AUX.PRS.2SG acheter.PST.PTCP

Espagnol

- (128) A mi me gusta lo que compraste.
PREP 1SG.DAT 1SG.CL.DAT aimer-3SG ART.DEF.N que acheter.PST.2SG

Catalan

- (129) M’agrada el que has comprat.
1SG.DAT-aimer-3SG ART.DEF.M que AUX.PRS.2SG acheter.PST.PTCP

Comme le montrent les gloses, en portugais et en catalan, la tête de ces relatives, lorsque le référent est inanimé et indéfini, est l’article défini masculin ; en espagnol, il s’agit de l’article neutre. En italien, c’est un démonstratif qui est utilisé, comme en français. Mais il y a deux possibilités, comme on peut le voir en (126) et (127). La différence entre ‘quello’ et ‘ciò’ est que le premier est fort alors que le deuxième est faible.

Le Tableau 1.1 fait le point sur les propriétés syntaxiques de ces relatives à tête allégée dans ces quatre langues romanes, et en français. Il s’agit de critères binaires : pour chaque critère dans chaque langue, un - est indiqué si la propriété n’est pas présente dans la langue ; un + si cette propriété est présente dans la langue. Les différents critères qui me semblent pertinents ici pour comparer ces structures dans les différentes langues romanes sont les suivants.

- Article défini
- Démonstratif
- Tête exclusive aux REL
- Element faible/clitique

-
1. Israel de La Fuente pour l’espagnol
 2. Caterina Donati pour l’italien
 3. Eunice Fernandes pour le portugais
 4. Joan Busquets pour le catalan
- que je tiens à remercier.

- COMP
- Infinitif
- Modification locale
- Quantification
- Coordination avec tête pleine
- Coordination avec REL libre
- Préposition + REL
- Clivées
- Phrases présentatives

Les deux premiers critères concernent la nature de la tête de la relative : s'agit-il d'un article défini ? d'un démonstratif ? Bien sûr, ces deux critères sont complémentaires : si dans une langue *x*, la tête est un article défini, alors c'est qu'elle n'est pas un démonstratif. Par 'tête exclusive aux REL', j'entends : l'élément nominal (article ou démonstratif) peut-il être employé dans d'autres contextes ou bien est-il restreint aux relatives ? Le quatrième critère est lié mais distinct : l'élément nominal est-il un élément faible ou bien un élément fort ? La cinquième propriété, intitulée 'COMP', concerne la nature du terme introducteur de la relative. Le symbole + indiquera qu'il s'agit du complémenteur ; le symbole - indiquera qu'il ne s'agit pas du complémenteur, donc d'un terme interrogatif.

Le critère suivant a pour but de déterminer si le verbe de la relative à tête allégée peut être à l'infinitif dans une langue donnée. Les deux critères suivants concernent la modification locale et la quantification. Est-ce que la séquence 'élément nominal + introducteur REL' peut être interrompue par un élément de nature syntaxique différente d'une relative ? Cette même séquence peut-elle être précédée d'un quantifieur ?

Viennent ensuite deux critères de coordination. D'une part, est-ce que, dans la langue *x*, les relatives à tête allégée peuvent être coordonnées avec des relatives à tête pleine, de type D+N ? D'autre part, est-ce que les relatives de ce type peuvent être coordonnées avec une relative libre ?

Un autre critère de comparaison pertinent est : est-ce qu'une préposition peut précéder une relative à tête allégée ?

Enfin, les deux derniers critères concernent des structures voisines des relatives. Peut-on retrouver la même séquence 'élément nominal + terme introducteur' que l'on a dans ces relatives dans les clivées et dans les phrases présentatives ?

Critères syntaxiques	Français	Espagnol	Italien	Portugais	Catalan
Article défini	-	+	-	+	+
Démonstratif	+	-	+	-	-
Tête exclusive aux REL	-	-	-	-	-
Élément faible/clitique	+	+	-	+	?
COMP	+	+	+	+	+
Infinitif	-	-	-	+	-
Modification locale	-	-	-	-	-
Quantification	+	+	+	+	+
Coordination avec tête pleine	+	+	+	+	+
Coordination avec REL libre	-	-	?	+	-
Préposition + REL	+	+	+	+	-
Clivées	+	+	+	+	+
Phrases présentatives	+	+	+	+	+

TABLE 1.1 – Propriétés syntaxiques des relatives avec tête allégée dans les langues romanes

Étudions ces critères de plus près, en les illustrant avec des exemples dans chaque langue romane. En ce qui concerne la nature de la tête, comme nous l'avons fait remarqué précédemment, il s'agit d'un article neutre en espagnol, de l'article défini masculin en portugais et en catalan et d'un démonstratif en français et en italien.

Propriétés : nature de la tête ; tête exclusive aux relatives

Français

(130) ce pantalon

Espagnol

(131) lo importante
ART.N important
ce qui est important

Italien

(132) quello sbaglio
DEM erreur
cette erreur

(133) Ciò è vero.
DEM être.PRS.3SG vrai
Cela est vrai.

Portugais

(134) o menino
ART.DEF.M garçon
le garçon

Catalan

(135) el nen
ART.DEF.M garçon
le garçon

Les exemples ci-dessus permettent également d'illustrer le critère 'Tête exclusive aux REL', qui est négatif dans chacune des langues romanes étudiées. En effet, à partir du moment où l'élément nominal en question est utilisé dans d'autres contextes que les relatives, cette tête n'est pas exclusivement employée dans les relatives.

Ces éléments D sont des éléments faibles en espagnol, français et portugais. En effet, ils ne peuvent pas être employés seuls. En italien, 'quello' est fort alors que 'ciò' est faible (comme je me base sur 'quello', un - est indiqué dans le Tableau 1.1 pour l'italien). Quant au catalan, il semblerait aussi qu'il s'agisse d'un élément faible, mais l'élément fort est très proche ('ell' : pronom personnel et pronom fort de troisième personne du singulier).

Propriété : élément faible/clitique

Français

(136) Quel livre as-tu lu cette semaine ? *Ce.

Espagnol

(137) Qué libro has leído ? *Lo.
quel livre AUX.PRS.2SG lire.PST.PTCP ? ART.N

Italien

(138) Quale libro hai letto questa settimana ? Quello.
quel livre AUX.PRS.2SG lire.PST.PTCP cette semaine ? DEM
Quel livre as-tu lu cette semaine ? Celui-là.

Portugais

(139) Qual livro já leste ? *O.
quel livre ADV lire.PST.2SG ? ART.DEF.M
Quel livre as-tu déjà lu ?

Catalan

(140) Quin llibre has llegit ? *El.
quel livre AUX.PRS.2SG lire.PST.PTCP ? ART.DEF.M

Il semblerait que ce soit le complémenteur qui introduit ces relatives à tête allégée dans les cinq langues romanes, bien que la frontière entre mot *wh-* et complémenteur ne soit pas si nette dans les langues romanes, étant donné l'homophonie entre les deux. Toutefois, en français, la distribution entre 'que' et 'qui' (lacune en position d'objet direct *versus* lacune en position de sujet) nous permet de trancher en faveur du complémenteur.

Il est impossible d'avoir un verbe à l'infinitif à l'intérieur des relatives à tête allégée, sauf en portugais.

Propriété : compatibilité avec l'infinitif

Français

(141) *J'ai ce que lire.

Espagnol

(142) *Tengo lo que leer.
avoir-PRS.1SG ART.N C lire-INF

Italien

(143) *Ho quello che leggere.
avoir.PRS.1SG PREP DEM que lire-INF

Portugais

(144) Tenho o que fazer.
avoir-PRS.1SG ART.DEF.M que faire-INF
J'ai de quoi faire.

Catalan

(145) *Tinc el que llegir.
avoir.PRS.1SG ART.DEF.M que lire-INF

Quant à la modification locale, cela ne paraît pas possible dans aucune des langues étudiées.

Propriété : modification locale*Français*

(146) * Je recopie ce que d'autre tu as écrit.

Espagnol

(147) *Copio lo que otro has escrito.
recopier-PRS.1SG ART.N C autre AUX.PRS.2SG écrire-PST.PTCP

Italien

(148) *Ricopo quello che altro hai scritto.
recopier.PRS.1SG DEM que autre AUX.PRS.2SG écrire-PST.PTCP

Portugais

(149) *Copio o que de mais escreveste.
recopier-PRS.1SG ART.DEF.M que de plus écrire-PST.2SG

Catalan

- (150) *Copio el que d’altre has escrit.
 recopier.PRS.1SG ART.DEF.M que d’autre AUX.PRS.2SG écrire-PST.PTCP

Nous ne rentrerons pas ici dans les détails des différences syntaxiques entre ‘quello’ et ‘ciò’ en italien, mais il est intéressant de remarquer que ‘ciò’, contrairement à ‘quello’, peut être modifié localement.

- (151) Ricopo ciò che altro hai scritto.
 recopier.PRS.1SG DEM que autre AUX.PRS.2SG écrire-PST.PTCP

Dans les cinq langues romanes en question, la tête légère peut être précédée du quantifieur universel ‘tout’.

Propriété : quantification*Français*

- (152) Je déteste **tout** ce que tu as acheté.

Espagnol

- (153) Odio **todo** lo que compraste.
 détester.PRS.1SG tout ART.N C acheter-PST.2SG

Italien

- (154) Odio **tutto** quello che hai comprato.
 détester.PRS.1SG tout DEM que AUX.PRS.2SG acheter-PST.PTCP

Portugais

- (155) Odeio **tudo** o que compraste.
 détester.PRS.1SG tout ART.DEF.M que acheter-PST.2SG

Catalan

- (156) Odio **tot** el que has comprat.
 détester.PRS.1SG tout ART.DEF.M que AUX.PRS.2SG acheter-PST.PTCP

Il est possible de coordonner une tête pleine avec une relative à tête allégée dans les langues romanes étudiées, ce qui montre bien que ces relatives sont à analyser comme des DPs.²⁰

20. Ce test à lui seul ne permet probablement pas de déterminer la nature syntaxique des relatives à tête allégée. En effet, depuis SAG, GAZDAR, WASOW et WEISLER, 1985, nous savons qu’il est possible de coordonner

Propriété : coordination avec tête pleine*Français*

(157) J'ai écrit **la lettre** et ce que tu voulais que j'écrive.

Espagnol

(158) Escribí **la carta** y lo que querías que
 écrire-PST.1SG la lettre et ART.N C vouloir-IPF.2SG COMP
 escribiera.
 écrire-IPF.SBJV.1SG

Italien

(159) Ho scritto **la lettera** e quello che volevi che
 AUX.PRS.1SG écrire.PST.PTCP la lettre et DEM que vouloir-IPF.2SG COMP
 scrivessi.
 écrire-IPF.SBJV.1SG

Portugais

(160) Escrevi **a carta** e o que querias que
 écrire-PST.1SG la lettre et ART.DEF.M COMP vouloir-IPF.2SG COMP
 escrevesse.
 écrire-IPF.SBJV.1SG

Catalan

(161) He escrit **la carta** i el que volies que
 AUX.PRS.1SG écrire.PST.PTCP la lettre et DEM que vouloir-IPF.2SG COMP
 escrigués.
 écrire-IPF.SBJV.1SG

des catégories différentes si elles ont la même fonction. Par exemple, en français, on peut coordonner un adjectif et un syntagme prépositionnel, comme en (1).

(1) Lucie est contente et de bonne humeur.

Mais observons le contraste entre (2) et (3), qui nous semble intéressant dans notre cas.

(2) J'ai écrit [la lettre] et [ce que tu voulais que j'écrive].

(3) ??J'ai écrit [la lettre] et [que Marie devrait travailler davantage].

Le verbe 'écrire' peut sélectionner un DP ou un CP (complétive). Le fait que le DP 'la lettre' puisse être coordonné avec une relative à tête allégée (comme en (2)) mais pas avec une complétive (comme en (3)) nous indique que la première, contrairement à la seconde, est un DP.

Par contre, si l'on coordonne la relative à tête allégée avec un mot *wh-* (introduceur de relatives libres), cela conduit à un résultat agrammatical dans toutes les langues, sauf en portugais.²¹

Propriété : coordination avec mot *wh-*

Français

(162) ??J'aime **qui** et ce que tu aimes.

Espagnol

(163) *Me gusta **quién** y lo que te gusta.
1SG.DAT plaire-PRS.3SG qui et ART.N COMP 2SG.DAT plaire-PRS.3SG

Italien

(164) ?Amerò **chi** e quello che mi dirai d'amare
aimer-FUT.1SG qui et DEM que 1SG.DAT dire-FUT.2SG PREP-aimer-INF

Portugais

(165) Adoro **quem** e o que tu adoras.
aimer-PRS.1SG qui et ART.DEF.M que 2SG aimer-PRS.2SG

Catalan

(166) *M'agrada **qui** i el que t'agrada.
1SG.DAT-plaire-PRS.3SG qui et ART.DEF.M COMP 2SG.DAT-plaire-PRS.3SG

21. Là encore, ce critère de coordination à lui seul ne permet pas de tirer des conclusions sur la nature de 'ce que', mais il semble mettre en évidence une différence intéressante entre le français et le portugais. Notons d'ailleurs que, pour le français, on ne trouve dans FrWaC (corpus de Français du Web, BARONI, BERNARDINI, FERRARESI et ZANCHETTA, 2009) que 3 résultats pour la coordination de 'qui' et 'ce qui' et 9 résultats pour la coordination de 'qui' et 'ce que'.

(1) Nous espérons que vous trouverez toutes les informations qui vous indiqueront qui et ce que nous sommes.

(Source : http://corentincelton.php.fr/corentin/01_hospital.html)

(2) Vous saviez parfaitement qui et quoi elle visait.

(Source : <http://lemessenger.online.fr/Italiano/Scuola3.html>)

Il est vrai que ce n'est pas forcément évident de coordonner quelque chose d'animé avec quelque chose d'inanimé, mais il y a davantage de résultats pour la coordination de 'qui' et 'quoi' (68 résultats). Pour des coordinations plus naturelles comme 'où' et 'quand' ou 'quand' et 'comment', on trouve encore plus de résultats évidemment, respectivement 624 et 619 résultats.

La séquence D+C est compatible avec une préposition, autrement dit ces relatives à tête allégée peuvent être complément de préposition, sauf en catalan.

Propriété : préposition + REL

Français

(167) Je pense à ce que tu m'as dit.

Espagnol

(168) Estoy pensando en lo que me dijiste.
être-PRS.1SG penser-GRDV PREP ART.N COMP 1SG.DAT dire-PST.2SG

Italien

(169) Penso a quello che voglio fare.
penser-PRS.1SG PREP DEM que vouloir-PRS.1SG faire.INF
Je pense à ce que je veux faire.

Portugais

(170) Penso no que me disseste.
penser-PRS.1SG PREP+ART.DEF.M que 2SG.DAT dire-PST.2SG

Catalan

(171) *Penso en el que em digueres.
penser-PRS.1SG PREP ART.DEF.M COMP 1SG.DAT dire-PST.2SG

Enfin, dans les cinq langues, la séquence D+C peut être utilisée dans les clivées et phrases présentatives. Les deux structures étant proches, je choisis de mettre seulement un exemple dans chaque langue pour illustrer ces deux dernières propriétés.

Propriétés : phrases présentatives et clivées

Français

(172) Voilà ce que j'ai dessiné.

Espagnol

(173) Aquí está lo que dibujé.
ici être-PRS.3SG ART.N COMP dessiner-PST.1SG

Italien

- (174) Ecco quello che ho disegnato.
voilà DEM que AUX.PRS.1SG dessiner-PST.PTCP

Portugais

- (175) Isto é o que desenhei.
cela être.PRS.3SG ART.DEF.M que dessiner-PST.1SG

Catalan

- (176) Això és el que he dibuixat.
cela être.PRS.3SG ART.DEF.M que AUX.PRS.1SG dessiner-PST.PTCP

Globalement, on observe de nombreux points communs entre les relatives à tête allégée dans les cinq langues romanes en question. Ces propriétés semblent confirmer qu’il s’agit bien de DPs. Mais on peut constater que le portugais se détache un peu des autres langues romanes, en particulier en ce qui concerne les propriétés suivantes : en portugais, on peut utiliser un verbe à l’infinitif dans ce type de relatives ; on peut coordonner un mot *wh*- avec ‘o que’. Cela suggère qu’en portugais, ‘o que’ serait peut-être de nature *wh*-, peut-être suite à un processus de grammaticalisation de cette séquence D+C. Le statut de ‘o que’ sera approfondi dans le Chapitre 2.

Par ailleurs, le cas du roumain est intéressant car il dispose des deux stratégies, et ce de façon quasi systématique.

- (177) Imi place ceea ce ai cumparat.
1SG.DAT plaire.PRS.3SG DEM WH AUX.PRS.2SG acheter-PST.PTCP
J’aime ce que tu as acheté.
- (178) Fă ce vrei !
faire.IMP.2SG WH vouloir.PRS.2SG
Fais ce que tu veux !

Contrairement aux autres langues romanes, qui n’autorisent que la stratégie ‘relative à tête allégée’ lorsque le référent est inanimé, le roumain dispose également de relatives libres. Mais l’alternance ne semble pas aussi libre qu’on pourrait le penser. Dans certains cas, il semblerait préférable d’ajouter ‘ceea’ alors que dans d’autres, cela semble plus naturel d’utiliser ‘ce’ seul. D’après Grosu, dans le Chapitre 10 de la grammaire de référence de la langue roumaine, éditée par Dobrovie-Sorin et Giurgea (DOBROVIE-SORIN et GIURGEA, 2013) :

« L’emploi de ‘ceea ce’ est préféré par rapport à l’emploi de ‘ce’ (voire requis) dans au moins deux situations : (i) lorsque la relative libre est prédicative et (ii) lorsqu’elle est dite ‘transparente’

[...] » (p. 638)²²

Chaque cas est illustré avec un exemple roumain ci-après.

(i) lorsque la relative libre est prédicative (p. 627)

- (179) Ion est astăzi **ceea ce** a sperat dintotdeauna mama lui că va fi
 Ion is today that what has hoped always mother his that would be
 cândva : respectat, admirat, și acoperit de onoruri.
 some-day : respected, admired, and full of honours
 ‘Ion is today what his mother always hoped he would be some day : respected,
 admired and full of honors.’

(ii) lorsque qu’il s’agit d’une relative dite ‘transparente’ (p.639)

- (180) Ion mănâncă **ceea ce** pare a fi un cotlet de porc (dar ar putea în
 Ion eats that what seems to be a steak of pork (but would can in
 principiu să fie și o falcă de măgar)
 principle CONJ be-SUBJ also a jaw of donkey
 ‘Ion eats what looks like a pork steak (but in principle it could also be a donkey
 jaw).’

La nature de ‘ceea’ et la nature de ‘ce’ sont également différentes comparées aux autres langues romanes étudiées. Comme les gloses l’indiquent dans les exemples (177) et (178), ce n’est pas le complémenteur qui introduit ce type de relatives en roumain mais le mot wh-, équivalent à l’anglais ‘what’. Quant à ‘ceea’, ce mot n’existe pas tel quel en roumain. On ne le trouve que dans les relatives à tête allégée. Mais sa ressemblance avec le démonstratif ‘cea’ (démonstratif féminin singulier, nominatif et accusatif, équivalent de ‘celle’ en français) est assez troublante. Le statut de tête allégée est discutable dans le cas du roumain, comme le note DINDELEGAN, 2016 (p.481)

« La combinaison de ‘ce’ avec le démonstratif singulier féminin ‘cea’ est pratiquement grammaticalisée en roumain moderne. Il est employé comme un connecteur relatif neutre spécialisé, surtout dans les appositives. Même dans les relatives restrictives, le composant démonstratif a perdu sa référentialité et n’est plus antécédent de la relative. »²³

Notons que la question de la grammaticalisation se pose également dans le cas de ‘ce qu-’ en français. Cette hypothèse sera justement détaillée et discutée dans le Chapitre 2, Section 2.3.6.

22. Il s’agit ici d’une traduction en français de la citation originale qui est en anglais.

23. Il s’agit ici d’une traduction en français de la citation originale qui est en anglais.

1.3.3.2 Dans des langues d’autres familles

Les relatives à tête allégée n’existent pas que dans les langues romanes. On trouve ce type de relatives dans des langues d’autres familles : dans des langues germaniques, telles que l’allemand et le néerlandais ; dans des langues slaves, telles que le polonais ; dans des langues sémitiques, comme l’arabe (dialecte de Mardin) ; dans des langues iraniennes, telles que le tati (dialecte de Chali).²⁴

Allemand

- (181) Mary isst **das** was (auch) John isst.
 Mary manger.PRS.3SG DEM WH- (aussi) John manger.3PRS.3SG
 Marie mange ce que Jean mange.

Néerlandais

- (182) Marie eet **dat** wat Jan eet.
 Marie manger.PRS.3SG DEM WH- John manger.3PRS.3SG
 Marie mange ce que Jean mange.

Polonais²⁵

- (183) Zrobię **to**, co Maria mi kazała.
 faire-1SG DEM, WH- Marie 1SG.DAT dire
 Je ferai ce que Marie me dira de faire.

Arabe, dialecte de Mardin (exemple de GRIGORE, 2007 : 239, tiré de la thèse de GANDON, 2016 : 300)²⁶

- (184) Ana ma-fahəmtu ayš la kəl-qał.
 PRO.1SG NEG-compris.1SG PRO.INT L PF-dire.PST.3SGM
 Je n’ai pas compris ce qu’il avait dit.

Tati, dialecte de Chali (exemple de YAR-SHATER, 1969 : 179, tiré de la thèse de GANDON, 2016 : 163)

- (185) **Em** ke šoma bat.
 DEM.SGM SUB 2PL dire.PST.2SG
 Ce que tu as dit.

Ces données dans des langues autres que les langues romanes nous permet de réaliser qu’une relative à tête allégée sujet ou objet n’est pas systématiquement introduite par un complémenteur, mais parfois par un mot wh-.

24. Je cite ici ces langues de façon anecdotique. J’ai pu trouver un exemple de relative à tête allégée pour chacune de ces langues dans la littérature.

25. Les trois exemples (181), (182) et (183) proviennent de CITKO, 2004.

26. La glose L dans l’exemple (184) correspond à une particule des dialectes arabes qui a, entre autres, une fonction de relativiseur.

1.4 Quelques conclusions sur la nature des relatives à tête allégée

Les relatives à tête allégée partagent certaines propriétés avec les relatives pleines et d'autres avec les relatives libres, mais il semblerait qu'une classe à part soit nécessaire. Dans le cas du français, nous considérons que les constructions *ce+REL* sont des relatives à tête allégée, mais pas les constructions *celui+REL*, que nous assimilons aux relatives pleines.

Une relative à tête allégée peut être introduite soit par un complémenteur, soit par un mot *wh-*. Dans les langues romanes, c'est le complémenteur qui est généralement utilisé²⁷, tandis que dans d'autres langues (allemand, néerlandais, polonais, par exemple), c'est un mot *wh-* qui est employé. Quant à la tête elle-même, si l'on ne regarde que les exemples dans les langues autres que les langues romanes, on pourrait penser que la tête est toujours un démonstratif. Ce qui est faux, puisqu'en espagnol, en catalan et en portugais, il s'agit d'articles/pronoms. L'exemple en (184) est à part, comparé à tous les autres exemples, toutes familles de langues confondues. En effet, les deux éléments clés de la relative à tête allégée sont un pronom interrogatif, suivi d'une particule de relativisation. Il y aurait donc trois configurations différentes possibles dans le cas des relatives à tête allégée.

- D+C (surtout dans les langues romanes)
- D+wh (dans des langues d'autres familles)
- wh+C (cas de l'arabe, dialecte de Mardin)

Par ailleurs, il est important de noter que les relatives à tête allégée et les relatives libres sont parfois complémentaires, mais pas toujours. Il est vrai que certaines langues ont la possibilité de faire des relatives libres lorsque le référent est inanimé, mais pas des relatives à tête allégée. C'est le cas de l'anglais. D'autres langues, comme le français, n'ont pas d'autre possibilité que d'avoir recours à une relative à tête allégée, lorsque le référent est inanimé. Mais dans certaines langues, comme l'allemand, les deux types de relatives sont possibles.

Anglais

- (186) I like what you bought.
 PRO.1SG aimer.PRS WH- PRO.2SG acheter.PST
 J'aime ce que tu as acheté.

Français

- (187) a. J'aime ce que tu as acheté.

27. Du moins lorsque la lacune est en position de sujet ou d'objet direct. En effet, en français, après une préposition, c'est le mot *wh-* 'quoi' qui est employé dans les relatives à tête allégée.

(1) Je note ce à quoi je pense.

- b. *J'aime que tu as acheté.
 c. *J'aime quoi tu as acheté.

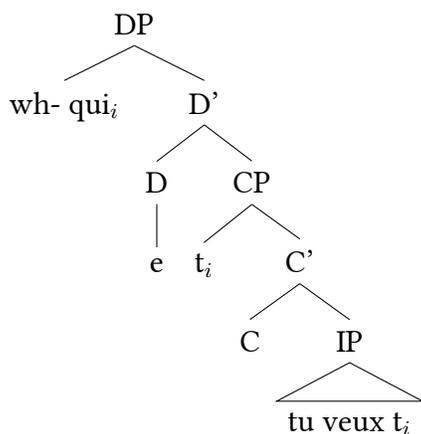
Allemand

- (188) Mary isst **das** was (auch) John isst.
 Mary manger.PRS.3SG DEM WH- (aussi) John manger.3PRS.3SG
 Marie mange ce que Jean mange.
- (189) Mary isst **was** (auch) John isst.
 Mary manger.PRS.3SG WH- (aussi) John manger.3PRS.3SG
 Marie mange ce que Jean mange.

Ce qui est intéressant, c'est que l'anglais, qui ne dispose que des relatives libres aujourd'hui, a eu, à une certaine époque, des relatives à tête allégée. Comme le font remarquer (CECCHETTO et DONATI, 2015 : 53-54), en anglais archaïque, on pouvait trouver des relatives introduites par l'élément wh- 'what', précédé de l'élément D 'that'. L'exemple Shakespearien (extrait de *Romeo and Juliet* II, ii, 1-2) qui suit le montre.

- (190) **That which** we call a rose / By any other name will smell as sweet.

Ces données sont une évidence en faveur d'une analyse syntaxique unifiée des relatives. L'idée est que la structure est toujours DP+CP, et donc un élément D (au minimum) est toujours là, du moins dans la structure sous-jacente. Dans le cas des relatives pleines, la tête est pleine, donc de type D+N. Dans le cas des relatives à tête allégée, la tête est légère, de type D. Dans le cas des relatives libres, en apparence, il n'y a ni D ni N, mais la structure sous-jacente est bien DP + CP car la position D est toujours là. CAPONIGRO, 2002 avait déjà proposé qu'un D est présent même dans les relatives libres. L'idée est la suivante : le mot wh- se déplace dans le spécifieur de DP, et un D silencieux prend comme complément un CP :



Comme on le voit sur l'arbre syntaxique en (191), l'élément wh- qui introduit la relative libre se déplace dans la position spécifieur de CP puis dans la position spécifieur de DP. Ce

deuxième déplacement, de CP à DP, a lieu pour licencier la tête silencieuse D. Notons que, dans le cas du français, l'élément qui introduit une relative libre est le même élément que celui qui introduit une question enchâssée (ils ont tous les deux un trait [+ animé]), alors qu'une relative pleine est introduite par un complémenteur, qui est donc sous C.

Il est intéressant d'observer, que, même si Caponigro ne parle pas des relatives telles que (190), la structure qu'il propose est tout à fait compatible avec (190) : dans ce cas, le mot *wh-* 'which' resterait dans la position [Spec, CP], et la position D serait remplie par le démonstratif 'that'. Par ailleurs, cette analyse des relatives libres mettrait fin au célèbre débat entre la *Comp Analysis* et la *Head Analysis* dans le cas des relatives libres (cf. Sous-Section 1.3.2.3).

Le squelette serait donc toujours le même, que ce soit dans le cas des relatives libres ou dans le cas des relatives à tête allégée : un DP suivi d'un CP. Par contre, les positions remplies par des éléments lexicaux ne sont pas les mêmes. Dans le cas des relatives libres, c'est la position de [Spec, DP] qui est remplie par le mot *wh*, et la position D est vide, de même que la position C. Dans le cas des relatives à tête allégée, c'est l'inverse : la position D est remplie par la tête légère, et la position C est remplie par le complémenteur. La position [Spec, DP] est vide. En ce qui concerne les relatives pleines, un NP intermédiaire est nécessaire entre le DP et le CP pour accueillir le nom.

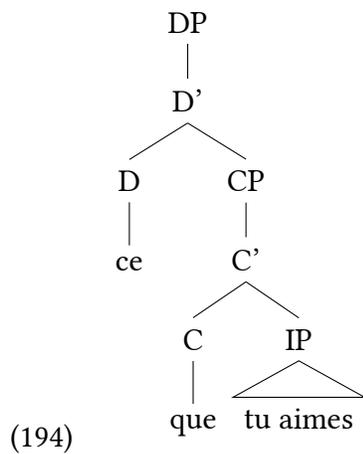
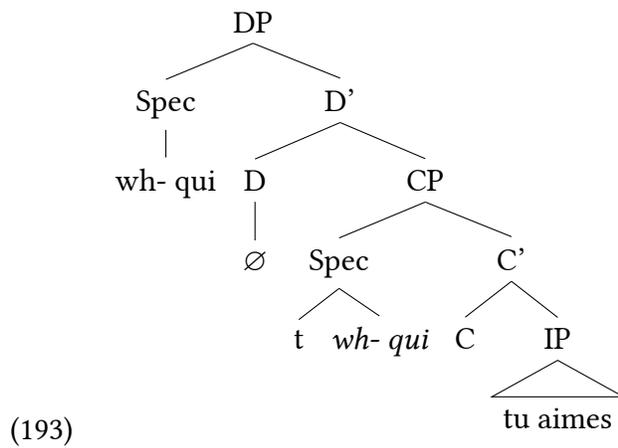
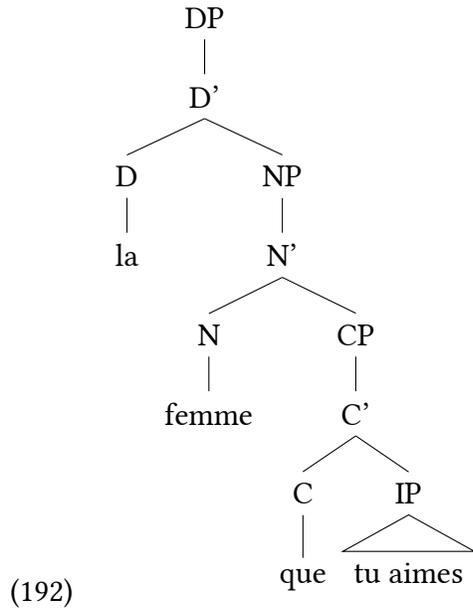
On pourrait proposer une structure maximale équivalente pour les trois types de relatives, c'est-à-dire une structure DP+NP+CP. Or, dans le cas des relatives libres, il n'y a aucune raison de postuler un N vide. Quant aux relatives à tête allégée, on pourrait être facilement tentés de postuler un N vide, étant donné que l'élément D est un démonstratif ou un article défini : la position N serait là dans la structure sous-jacente, le nom n'étant par contre pas épelé²⁸. Toutefois, lorsque nous avons fait la comparaison avec les relatives libres d'une part, et les relatives pleines d'autre part, il s'est avéré que les relatives à tête allégée partageaient plusieurs propriétés avec les relatives libres. Par ailleurs, comme nous le verrons dans le Chapitre 2 qui suit, ces relatives à tête allégée introduites par 'ce qu-' sont ambiguës en français avec les interrogatives indirectes introduites par 'ce qu-', exactement comme les relatives libres²⁹. Il me semble que cela est une raison suffisante pour abandonner l'idée d'un N sous-jacent dans le cas des relatives à tête allégée, du moins pour le français.

28. Sur les DPs sans nom, réduits à un déterminant (DWNs, *Determiners Without a Noun*), et leur interprétation, voir notamment SLEEMAN, 2003 et CORBLIN, MARANDIN et SLEEMAN, 2004.

29.

- (1) a. qui tu veux inviter
 b. Invite [_{DP} qui tu veux inviter].
 c. Je me demande [_{CP} qui tu veux inviter].

Ainsi, les représentations syntaxiques d'une relative pleine, d'une relative libre et d'une relative à tête allégée seraient respectivement les structures en (192), (193) et (194).



Le Tableau 1.2 permet de faire le bilan sur les différents introducteurs des différents types de relatives en français.

Introduceur	REL pleines	REL libres	REL à tête allégée
qui WH-		✓	
qui COMP	✓		✓ (ce qui)
que COMP	✓		✓ (ce que)
dont	✓		✓ (ce dont)
où	✓	✓	✓ (là où)
quand	✓	✓	
PRÉP+qui	✓	✓	
PRÉP+quoi	✓	✓	✓ (ce+PRÉP+quoi)
PRÉP+lequel	✓		

TABLE 1.2 – Les différents introducteurs des différents types de relatives en français.

L'un des premiers rôles de 'ce' en français est donc celui de tête allégée dans certaines relatives. Nous allons voir dans le Chapitre suivant que 'ce' est également présent dans certaines interrogatives indirectes.

Chapitre 2

Une ambiguïté surprenante avec les interrogatives indirectes ¹

2.1 Ressemblances entre relatives à tête allégée et interrogatives indirectes en français

Nous avons vu qu'il existait des ressemblances formelles entre les relatives libres et les interrogatives indirectes. Certaines relatives à tête allégée sont elles aussi homophones à certaines questions *wh-* enchâssées, appelées 'interrogatives indirectes partielles' dans les grammaires traditionnelles. Les structures en question sont celles introduites par 'ce qu-'.¹

- (1) **ce que tu as écrit e.**
- (2) Je lis **ce que tu as écrit e.**
- (3) Je me demande **ce que tu as écrit e.**

Une proposition subordonnée comme en (1) peut être interprétée de deux façons différentes, soit comme une relative (2), soit comme une interrogative indirecte (3). La seule différence entre (2) et (3) est le verbe de la proposition supérieure, qui peut sélectionner une relative, comme le verbe 'aimer', ou une interrogative indirecte comme le verbe 'se demander'. Mais les propositions subordonnées sont strictement identiques, du moins en surface.

On retrouve la même ambiguïté avec 'ce qui'...

- (4) **ce qui e te dérange.**
- (5) Je note **ce qui e te dérange.**
- (6) Je me demande **ce qui e te dérange.**

1. J'ai présenté des versions antérieures de ce qui est détaillé dans ce chapitre à deux conférences internationales, *The 30th Going Romance* (KONRAD, 2016) et *The 14th Workshop on Syntax, Semantics and Phonology* (KONRAD, 2017b). J'ai également écrit un article sur cette ambiguïté entre relatives et interrogatives indirectes introduites par 'ce qu-', qui sera publié en août 2019 (KONRAD, A paraître).

... et avec ‘ce dont’.

- (7) **ce dont tu as besoin e.**
- (8) Je note **ce dont tu as besoin e.**
- (9) Je me demande **ce dont tu as besoin e.**

Sur le plan sémantique, le démonstratif ‘ce’ ne revêt pas le même sens dans les interrogatives indirectes partielles que dans les relatives. Dans ces dernières (2, 5, 8), ‘ce’ dénote une entité. Mais, dans les interrogatives indirectes (3, 6, 9), ‘ce que/qui/dont’ dénote, en forme logique, un ensemble de propositions qui sont vraies dans le monde réel et qui sont équivalentes à la proposition qu’il y a une entité inanimée que tu as écrite (3) / qui te dérange (6), dont tu as besoin (9). Ainsi, dans un cas (les relatives) on réfère à un objet alors que dans l’autre (les interrogatives indirectes) on réfère à l’identité de l’objet.

2.2 Une ambiguïté typique : le cas de l’anglais

2.2.1 *What*

Avant d’expliquer en quoi cette ambiguïté est atypique, et de tenter d’en donner une analyse syntaxique, arrêtons-nous un instant sur une ambiguïté typique, celle que l’on trouve en anglais avec le mot wh- ‘what’.

- (10) **what you bought e.**
- (11) I like **what you bought e.**
- (12) I wonder **what you bought e.**

La structure en (10) peut être interprétée comme une relative libre, dans le cas où le verbe de la proposition matrice sélectionne un DP (11), ou bien comme une interrogative indirecte, dans le cas où il sélectionne un CP (12). Cette ambiguïté est typique étant donné que l’on retrouve systématiquement cette ambiguïté avec les éléments wh-, et même en français, avec d’autres mots wh- (comme ‘qui’, entre autres).

Il existe au moins deux explications différentes permettant de rendre compte de cette ambiguïté : une explication en termes d’étiquetage, *the labeling account*, et une explication plus traditionnelle, en termes d’enchâssement, que nous appellerons *the embedding account*.

2.2.2 *The labeling account*

Selon cette première approche, il n’y a qu’une seule structure correspondant à (10), avec deux possibilités d’étiquetage. La première possibilité est que l’élément wh- donne le label à la structure. Comme cet élément est en fait un élément de type D, on obtient un DP.

Cette structure pourra alors être sélectionnée par des verbes tels que *to read* ('lire') qui sélectionnent un DP. Dans ce cas-là, on aura l'interprétation de relative. La deuxième possibilité est que l'élément C, la sonde du mouvement wh-, donne le label à la structure. Dans ce cas on obtient un CP et donc cette proposition pourra être sélectionnée par des verbes comme *to wonder* ('se demander') qui ont besoin d'un CP. On aura alors l'interprétation d'interrogative indirecte.²

(13) I like [_{DP} what you bought e].

(14) I wonder [_{CP} what you bought e].

2.2.3 *The embedding account*

Selon cette deuxième approche, il existe bel et bien deux structures différentes pour (10), et ces structures ne sont identiques que de façon superficielle. Quand la structure est interprétée comme un DP (dans le cas des relatives), un D nul est postulé. Ce D nul enchâsse le CP, qui est identique à celui que l'on a dans le cas de l'interrogative indirecte.

(15) I like [_{DP} D [_{CP} what you bought e]].

(16) I wonder [_{CP} what you bought e].

Dans le cas d'une interrogative indirecte, on a donc une structure plus simple, qui n'est pas enchâssée sous un DP.

Il est vrai que le *labeling account* est davantage parsimonieux puisqu'il évite le fait de devoir postuler un élément vide. Toutefois, cette approche exige une théorie et un algorithme de l'étiquetage plus complexes puisqu'il autorise des ambiguïtés d'étiquetage.

Selon le *labeling account*, le mot wh- doit forcément être une tête : un syntagme n'a pas de pouvoir d'étiquetage. Cela implique que les relatives libres introduites par un syntagme wh- de type *-ever + N* (comme en (17)) doivent être analysées autrement.

(17) I will buy whichever book you will buy.

DONATI et CECCHETTO, 2011, en se basant sur les propriétés syntaxiques de ces relatives dites « libres » (BATTYE, 1989), postulent qu'il s'agit en réalité de relatives pleines³ :

1. Les items en *-ever* peuvent avoir un usage absolu, ce qui n'est pas le cas des relatives libres ordinaires (confronter (18) et (19)).

(18) John would go out with whichever woman.

2. Pour une description détaillée de cette approche de l'étiquetage, voir les articles CECCHETTO et DONATI, 2010 et DONATI et CECCHETTO, 2011, ainsi que le livre des mêmes auteurs, CECCHETTO et DONATI, 2015.

3. Voir CAPONIGRO, 2019 pour une critique de l'analyse de ces structures proposée par DONATI et CECCHETTO, 2011.

- (19) *John would go out with which/what woman.
2. Les items en *-ever* peuvent être suivis du complémenteur introduisant les relatives pleines, alors que cela n'est pas possible dans les relatives libres ordinaires (confronter (20) et (21)).⁴
- (20) John would read whichever book that he happened to put his hands on.
- (21) *John would read which/what (book) to put his hands on.
3. Le troisième argument est la possible co-occurrence d'un item en *-ever* avec un pronom relatif. La co-occurrence avec un pronom relatif conduit à un résultat agrammatical dans le cas des relatives libres ordinaires (confronter (22) et (23)).
- (22) Whichever boy to whom I speak keeps telling me the same thing.
- (23) *Which/what (boy) to whom I speak keeps telling me the same thing.

Selon l'*embedding account*, où le mot *wh-* des relatives libres est en [Spec, CP] comme dans les relatives pleines, la syntaxe devrait être la même. Comment expliquer alors que seules les relatives libres (pas les relatives pleines) soient contraintes aux phénomènes de *matching* (rappelons que cette contrainte est que la forme du mot *wh-* doit pouvoir satisfaire la sélection catégorielle et casuelle du verbe de la proposition matrice et de celui de la relative)?⁵

2.3 Une ambiguïté atypique : le cas du français

2.3.1 Une confirmation de l'*embedding analysis*?

Les relatives à tête allégée du français semblent être une confirmation, *prima facie*, de l'*embedding account*. En effet, le D est explicite et non pas silencieux. Ceci suggère que la position D vide postulée pour les relatives libres de l'anglais peut être remplie dans certaines langues par un déterminant ou démonstratif. Mais, en français, les interrogatives indirectes avec référent inanimé sont également introduites par 'ce qu-'. Cela pose problème car l'*embedding account* ne postule pas d'élément D dans le cas des interrogatives indirectes. Ce qui différencie les deux structures est justement le fait que dans les relatives il y ait cet élément D, mais pas dans les interrogatives.

4. Une forme équivalente à 'whichever' existe en français (quelconque) mais elle est dissociée de l'emploi relatif.

(1) Thomas a rencontré une femme quelconque.

(2) *Antoine serait capable de lire quelconque livre qu'il aurait trouvé à la bibliothèque.

5. Voir HUDDLESTON, PULLUM et PETERSON, 2002, CAPONIGRO, 2019, qui montrent que les relatives peuvent être introduites par des syntagmes *wh-*, contrairement aux prédictions de DONATI et CECCHETTO, 2011.

L'ambiguïté de la structure en 'ce que' est au contraire plausible du point de vue du *labeling account*, étant donné que soit un élément D soit la sonde C peuvent donner le label à une même structure. Mais ce qui pose problème ici est l'absence de trait wh- : 'ce' est un élément nominal [-wh] et 'que' n'est pas le mot wh- ici mais le complémenteur.

Par ailleurs, notons que, indépendamment du point de vue que l'on adopte, cette ambiguïté du français reste surprenante. En effet, on sait qu'une interrogative indirecte est un CP. On sait aussi que des verbes comme 'se demander' sélectionnent un CP, et non pas un DP.⁶ Ni une tête pleine ni la tête allégée 'celui' ne peuvent introduire une interrogative indirecte. Seul 'ce qu-' est possible.

(24) *Je me demande le livre que tu as acheté.

(25) *Je me demande celui que tu as acheté.

(26) Je me demande ce que tu as acheté.

Si 'se demander' requiert un CP, comment expliquer la présence de l'élément nominal, donc non-interrogatif, 'ce' ? Quelle est sa place dans la structure ? Est-il dans le spécifieur de CP, alors qu'il est [-wh] ? Cela est peu probable.

'Que' ou sa version forte 'quoi' devrait introduire les questions indirectes en français, mais comme nous le verrons dans la sous-section suivante, ni 'que' ni 'quoi' ne sont possibles dans ce contexte syntaxique. Au-delà de cette lacune du français, 'que' ou 'quoi' devraient également pouvoir introduire une relative libre, ce qui n'est pas le cas. Pourquoi le français n'a-t-il pas de structures analogues à celles introduites par 'what' en anglais ?

2.3.2 Le problème : ni 'que' ni 'quoi'

En français, 'que' seul, contrairement à 'qui', n'est pas capable d'introduire une interrogative indirecte.

(27) a. Qui aimes-tu ?

b. Je me demande qui tu aimes.

(28) a. Que veux-tu ?

b. *Je me demande que tu veux.

c. Je me demande ce que tu veux.

6. Comme nous le verrons dans le Chapitre 7, certains DPs interprétés comme des questions peuvent être sélectionnés par des verbes tels que 'se demander'. Il s'agit des *concealed questions* ou 'questions déguisées' (voir notamment BAKER, 1968, GRIMSHAW, 1979, HEIM, 1979, ROMERO, 2006 et FRANA, 2017).

(1) Je me demande l'heure qu'il est.

Comment expliquer cela ? Plusieurs linguistes affirment que ‘que’ est en fait un élément faible (OBENAUER, 1976 ; GOLDSMITH, 1978 ; BOUCHARD et HIRSCHBÜHLER, 1986 ; FRIEDEMANN, 1990), qui, en tant que clitique, a besoin de s’accrocher à T, c’est-à-dire au verbe conjugué. C’est pour cela qu’il requiert l’inversion clitique-verbe, qui est le résultat du mouvement de T à C. Et une telle inversion est impossible dans des contextes enchâssés.

- (29) a. *Que tu veux ?
 b. Que veux-tu ?
 c. *Je me demande que veux-tu.

Nous venons de donner une explication probable de l’impossibilité d’utiliser ‘que’ seul dans les interrogatives indirectes en français⁷. Une question qui vient à l’esprit maintenant est alors la suivante : qu’en est-il de ‘quoi’, la version forte de ‘que’ ? Pourquoi ‘quoi’ non plus ne peut-il pas être employé dans des interrogatives indirectes ? Notons que ‘quoi’, comme ‘que’, n’est pas autorisé non plus dans les relatives libres.

- (30) *Je me demande quoi tu veux.
 (31) *J’aime quoi tu as cuisiné.

L’agrammaticalité de (30) et de (31) peut s’expliquer par le fait que ‘quoi’ nu ne peut pas se déplacer : il doit rester *in situ* (OBENAUER, 1976).

- (32) a. *Quoi vas-tu dessiner ?
 b. Tu vas dessiner quoi ?

Par ‘quoi nu’, j’entends ‘quoi’ seul, c’est-à-dire ‘quoi’ lorsqu’il n’est pas précédé par une préposition, et lorsqu’il n’est pas modifié. La séquence ‘préposition + quoi’ peut en effet se déplacer ou bien rester *in situ*.

- (33) a. Avec quoi vas-tu dessiner le chat ?
 b. Tu vas dessiner le chat avec quoi ?

‘Quoi’, lorsqu’il est précédé par une préposition, peut d’ailleurs être employé dans les interrogatives indirectes, mais plus difficilement dans les relatives libres, où l’insertion d’un PP impliquant ‘ce’ est obligatoire, du moins en français standard.

- (34) Je me demande (ce) sur quoi tu travailles.
 (35) a. ??Je m’attendais à quoi tu t’attendais.
 b. Je m’attendais à ce à quoi tu t’attendais.

Comme l’a montré ABEILLÉ, 2008 : 92, ‘quoi’ peut se déplacer lorsqu’il est modifié par ‘d’autre’ dans les questions racines...

7. Nous verrons dans le Chapitre 3, portant sur la diachronie de ‘ce que’, que le mot *wh-* ‘que’ était l’introducteur d’interrogatives indirectes en Ancien Français et ce, pendant plusieurs siècles, avant que ‘ce que’ ne s’impose définitivement.

(36) Quoi d'autre te satisferait ?

...et cela est également possible dans les questions indirectes.

(37) Je me demande quoi d'autre te satisferait.

Si 'quoi' doit rester *in situ*, nous savons maintenant pourquoi il ne peut pas apparaître dans des relatives libres objet, comme en (31). Il est bien connu que les relatives libres ne peuvent jamais impliquer de wh- *in situ* (CECCHETTO et DONATI, 2015). L'agrammaticalité de l'exemple (38) n'est pas spécifique au français.

(38) *J'aime tu as cuisiné quoi.

Même des langues comme le chinois mandarin qui utilisent le wh- *in situ* comme stratégie de formation des questions wh-, n'autorisent pas de wh- *in situ* dans les relatives libres.

(39) Nei jinwan xiang chi shen-me ?
2SG ce-soir vouloir manger quoi
Tu veux manger quoi (ce soir) ?

(40) *Wo xihuan ni zuo de shen-me.
1SG aimer 2SG faire DE quoi

Quant aux interrogatives indirectes, 'quoi' peut y être accepté *in situ*, du moins dans certaines variétés du français non standard.

(41) !Je me demande tu veux quoi.

Mais le statut des cas comme (41) est plutôt controversé dans la littérature (voir notamment ADLI, 2006, SHLONSKY, 2012; COVENEY, 1995, COVENEY, 1997, MATHIEU, 2004, DELAVEAU, CAPPEAU et DAGNAC, A paraître).

Ainsi, on observe un réel contraste en français au sein des interrogatives indirectes, selon que le référent est animé ou inanimé. Lorsqu'il est animé, le mot wh- est le même dans les questions wh- directes, dans les questions wh- indirectes et dans les relatives libres.

- (42) a. Qui a porté des chocolats ?
b. Je me demande qui a porté des chocolats.
c. J'inviterai qui portera des chocolats.

Lorsque le mot wh- est inanimé, la question wh- directe peut être introduite par la forme wh- complexe 'qu'est-ce que / qu'est-ce qui' ou bien par 'que', nécessitant l'inversion cli-tique. Une troisième possibilité est d'employer 'quoi', qui reste *in situ*.

- (43) a. Qu'est-ce que tu fais ce soir ?
b. Que fais-tu ce soir ?

c. Tu fais quoi ce soir ?

Mais les questions wh- indirectes non obliques et les relatives libres ne peuvent être introduites ni par ‘que’ ni par ‘quoi’.

- (44) a. *Je me demande que fais-tu ce soir.
 b. *Je me demande que tu fais ce soir.
 c. !Je me demande tu fais quoi ce soir.
 d. *Je me demande quoi tu fais ce soir.
- (45) a. *Je photographie que fais-tu.
 b. *Je photographie que tu fais.
 c. *Je photographie tu fais quoi.
 d. *Je photographie quoi tu fais.

Elles sont introduites par ‘ce que / ce qui’.

- (46) a. Je me demande ce que tu fais ce soir.
 b. Je photographie ce que tu fais.

Pourquoi cette différence ? Quelle est la raison de la présence de ce ‘ce’ ?

Bien que la norme soit d’employer ‘ce que’ dans les interrogatives indirectes, il est important de noter qu’en français informel, il est possible d’étendre la stratégie des interrogatives directes, à savoir l’emploi de ‘qu’est-ce qu-’ aux interrogatives indirectes (en plus de la stratégie avec ‘quoi’ *in situ* comme en (47d)).

- (47) a. Qu’est-ce que tu veux ?
 b. !Je me demande qu’est-ce que tu veux.
 c. Qu’est-ce qui te ferait plaisir ?
 d. !Je me demande qu’est-ce qui te ferait plaisir.

Dans certaines variétés, comme en québécois, l’emploi de ‘qu’est-ce qu-’ (/kɛsk/), qui est d’ailleurs très fréquent, va même jusqu’à s’étendre aux relatives (NIÉGER, 1975).^{8 9}

- (48) Je vais faire qu’est-ce que tu veux.

Dans cette sous-section, nous avons montré qu’il existe une réelle lacune dans le paradigme des relatives libres et interrogatives indirectes en français : ni ‘que’ ni ‘quoi’ ne peuvent faire surface dans ces structures. Pour combler ce vide, nous employons ‘ce que’ (du moins en français standard), une forme dont la curieuse nature nécessite des investigations davantage approfondies.

8. *Kesk* serait alors un mot wh- qui aurait les mêmes fonctions et la même distribution que ‘what’ en anglais.

9. Cet emploi de ‘qu’est-ce que’ dans les relatives est généralisé en québécois mais on le trouve également en français informel, en France métropolitaine (voir étude dans FrWac dans le Chapitre 6, Section 6.2.2), mais c’est assez marginal.

Dans les prochaines sous-sections, nous allons justement proposer une analyse de cette construction en ‘ce que’, en essayant de comprendre sa nature, mais aussi comment une telle forme peut être légitimée dans des questions indirectes. Mais avant cela, je vais résumer les analyses majeures qui ont été proposées par plusieurs syntacticiens dans les années 80-90. Même si la syntaxe des constructions en ‘ce que’ a bénéficié de très peu d’attention, on peut trouver quelques analyses dans la littérature, qui sont des tentatives de réponses aux questions que nous avons soulevées.

2.3.3 Analyses antérieures

2.3.3.1 GOLDSMITH, 1981

GOLDSMITH, 1981 est le premier à tenter de résoudre ce problème, à savoir pourquoi la question indirecte est si différente de la question directe dans le cas où le pronom *wh-* est inanimé. Il fait le bilan, dans un tableau, des différents introducteurs de questions directes et indirectes, à temps fini et à temps non fini, que le *gap* soit en position de sujet, d’objet ou d’oblique (voir Figure 2.1).

Table 1

				<i>Finite Questions</i>		
<i>Direct Questions</i>	<i>Subject</i>	<i>Object</i>	<i>Oblique</i>			
Animate	qui	qui	qui			
Inanimate	(qu’est-ce qui)	(qu’est-ce que) or que*	quoi			
				<i>Nonfinite Questions</i>		
<i>Direct Questions</i>	<i>Subject</i>	<i>Object</i>	<i>Oblique</i>			
Animate	qui	qui	qui			
Inanimate	que* (?quoi)	quoi	quoi			
				<i>Nonfinite Questions</i>		
<i>Indirect Questions</i>	<i>Subject</i>	<i>Object</i>	<i>Oblique</i>			
Animate	qui	qui	qui			
Inanimate	ce qui	ce que	quoi			
				<i>Nonfinite Questions</i>		
<i>Indirect Questions</i>	<i>Subject</i>	<i>Object</i>	<i>Oblique</i>			
Animate	qui	qui	qui			
Inanimate	ce qui	ce que	quoi			
		(??que)				

The forms marked with an asterisk are clitic forms; see text. The forms marked with question marks represent variation in the data; see Hirschbühler (1978) and text.

FIGURE 2.1 – Les différents introducteurs des questions directes et indirectes, à temps fini et temps non fini, en position de sujet, d’objet et d’oblique (GOLDSMITH, 1981 : 555).

D'après GOLDSMITH, 1981, dans les questions directes, 'que' est la forme faible de 'quoi', de la même façon que 'te' et 'me' sont les formes faibles de 'toi' et 'moi'.

Son analyse des questions indirectes introduites par 'ce que' repose sur une règle d'insertion de 'ce', qui permet d'insérer 'ce' dans un NP qui domine un COMP commençant par 'que', comme l'illustre la Figure 2.2. Cette règle, dit-il, est requise par la syntaxe du français.

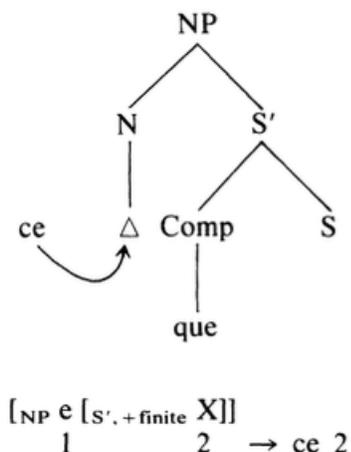


FIGURE 2.2 – Structure syntaxique des interrogatives indirectes introduites par 'ce que' en français et règle d'insertion de 'ce' d'après GOLDSMITH, 1981 : 562.

Cette règle du français est d'après lui nécessaire afin de pouvoir expliquer la présence de 'ce' dans les complétives (nous reviendrons sur ces structures dans le Chapitre 6.).

(49) Je m'attends à ce que Jean le fasse.

Revenons aux interrogatives indirectes. Il analyse donc 'que' comme étant le complémenteur dans les questions indirectes introduites par 'ce que'. En plus de cette règle d'insertion de 'ce', il propose une autre règle, qui s'applique en second lieu, et qui change la catégorie de l'interrogatif 'quoi' en un pronom [-WH]. Cet interrogatif transformé en pronom relatif est ensuite supprimé suite à l'application d'une règle indépendante (*Quoi deletion*).

La dérivation, étape par étape, d'une question indirecte selon l'analyse de GOLDSMITH, 1981 serait la suivante :

– Mouvement wh-

(50) Je me demande S' [quoi_i que [tu veux NP [e] $_i$]]
+wh

– Insertion d'un noeud NP

(51) Je me demande $_{NP}$ [$N_{S'}$ [$_{+wh}$ $_{wh}$ **quoi**_i que [tu veux $_{NP}$ [e]_i]]]]

– Insertion de ‘ce’

(52) Je me demande $_{NP}$ [ce $_{S'}$ [$_{+wh}$ $_{wh}$ **quoi**_i que [tu veux $_{NP}$ [e]_i]]]]

– Changement du trait [+WH] en [-WH]

(53) Je me demande $_{NP}$ [ce $_{S'}$ [$_{-wh}$ $_{wh}$ **quoi**_i que [tu veux $_{NP}$ [e]_i]]]]

– Effacement du pronom [-WH]

(54) Je me demande $_{NP}$ [ce $_{S'}$ [que [tu veux $_{NP}$ [e]_i]]]]

L’analyse de GOLDSMITH, 1981 est un peu faible de par la simple stipulation d’une règle d’insertion de ‘ce’. Mais l’intuition intéressante, que nous finirons par considérer dans le Chapitre 7, est que les interrogatives indirectes introduites par ‘ce que’ sont de fait des relatives.

2.3.3.2 KOOPMAN, 1982

Peu de temps après GOLDSMITH, 1981, KOOPMAN, 1982 propose à son tour une analyse de ces constructions. Je vais reprendre son raisonnement dans les grandes lignes. Notons que la théorie du Gouvernement et du Liage de Chomsky est à son apogée à ce moment-là. L’analyse de Koopman découle directement de la notion de gouvernement. KOOPMAN, 1982 fait l’hypothèse que l’élément déplacé dans une question directe comme en (55) est un PRO, sachant que PRO peut apparaître dans un COMP non gouverné.

(55) Que fais-tu ?

La représentation sous-jacente d’un exemple comme (55) sera alors la suivante :

(56) $_{S'}$ [$_{COMP}$ [$_{wh}$ $_{wh}$ $_{PRO}$ _i que][tu fais $_{e_i}$]]]

Comme nous sommes dans un cas de mouvement wh-, l’élément déplacé, ici PRO, possède un trait [+wh].

S’applique ensuite une règle de transfert du trait wh- sur le complémenteur, cette règle étant une variante de la règle proposée par PESETSKY, 1982.

(57) $_{COMP}$ [$_{+wh}$ $_{PRO}$ que] \rightarrow $_{COMP}$ [$_{+wh}$ PRO que]

Il est important de préciser que PRO ne transfère que son trait wh- à ‘que’, et pas ses autres traits pronominaux.

Afin d'expliquer pourquoi un exemple tel que (58) ne peut pas recevoir de sens inanimé, Koopman adopte la variante de la règle que-qui (KAYNE, 1976) proposée par AOUN, HORNSTEIN et SPORTICHE, 1981.

(58) *Qui a été fait ?

L'idée est que le complémenteur 'qui', contrairement à 'que', est un gouverneur propre car il est engendré avec un index. La structure sous-jacente de l'exemple (58) est donc la suivante.

(59) $PRO_{-anim} \text{ qui}_i^+ e_i \text{ a été fait}$

'Qui' gouverne PRO et PRO ne peut pas être gouverné donc l'exemple (58) est agrammatical. A la place, le français a recours à une forme wh- complexe, 'qu'est-ce qui'.

(60) Qu'est-ce qui a été fait ?

Koopman propose d'étendre cette analyse de mouvement de PRO aux questions indirectes et aux relatives libres. Une fois que le mouvement de PRO ainsi que le transfert du trait wh- ont eu lieu, on obtient la structure suivante.

(61) Je me demande S' [$COMP$ [PRO_i que [tu fais e_i]]
+wh

Le verbe 'se demander' a besoin d'un complément de type interrogatif. Pour vérifier cela, le verbe doit avoir accès au COMP. Cela signifie, en termes de gouvernement, qu'il doit y avoir gouvernement du verbe dans COMP. PRO va alors se retrouver gouverné par le verbe, ce qui est problématique puisque PRO ne peut pas être gouverné. Pour réparer cela, le français a la possibilité d'épeler les traits de PRO. C'est alors que le PRO inanimé, qui a perdu son trait [+wh] au profit de 'que', est épelé comme la forme non marquée des pronoms inanimés, à savoir 'ce'. On obtient alors (62).

(62) Je me demande ce que tu fais.

Koopman montre qu'on peut étendre le même raisonnement aux relatives libres.

(63) J'aime NP [NP [e] S' [PRO_i que [tu as fait NP [e] $_i$]]]
+wh

Le pronom wh- doit être accessible aux règles de vérification de cas, afin de pouvoir rendre compte des effets de *matching*. Encore une fois, le gouvernement entre en jeu, car qui dit vérification de cas dit relation de gouvernement. PRO est épelé comme 'ce'.

(64) J'aime ce que tu fais.

L'analyse de KOOPMAN, 1982 est intéressante : on ne sait pas pourquoi 'ce' est là dans les interrogatives indirectes, et postuler que 'ce' correspond à l'épellation d'un PRO est attirant (plus que la simple stipulation d'une règle d'insertion d'un noeud NP). Toutefois, le problème majeur de cette analyse est qu'elle ne permet pas d'expliquer pourquoi 'ce que' peut également introduire une relative en position de sujet dans la phrase.

(65) Ce que Marie a fait est inacceptable.

En effet, PRO ici n'est pas gouverné. La position sujet est d'ailleurs la position typique de PRO. Alors pourquoi 'ce' est-il obligatoire dans ce cas ?

(66) *Que Marie a fait est inacceptable.

D'autres grammairiens et linguistes n'ont pas forcément essayé d'expliquer pourquoi on a 'ce' dans les interrogatives indirectes, mais se sont interrogés sur la nature même de ces interrogatives indirectes introduites par 'ce que'. Est-ce que le 'ce que' des interrogatives indirectes est identique au 'ce que' des relatives, d'un point de vue syntaxique ?

2.3.3.3 ERIKSSON, 1982, JONES, 1996

Olof Eriksson, dans son article intitulé *Il m'a dit ce qu'il pense : interrogative ou relative ?*, affirme qu'une phrase de ce type est à considérer comme une relative, et ce, indépendamment de sa valeur sémantique. Il analyse donc 'ce' comme un antécédent, au même titre qu'un antécédent plein de type D+N.

JONES, 1996, dans son livre *Foundations of French syntax*, postule aussi une identité syntaxique entre les relatives et interrogatives indirectes introduites par 'ce que'. Il affirme en effet que '*French resorts to a construction involving a 'free' relative which masquerades as an indirect question*' (JONES, 1996 : 496).

2.3.3.4 MULLER, 1989, MULLER, 1996

Claude Müller aussi mentionne les « interrogatives indirectes partielles ayant pour antécédent le mystérieux *ce* neutre ». Comme ERIKSSON, 1982, Müller analyse 'ce' comme un antécédent dans des phrases comme (67).

(67) Je sais ce qui est arrivé.

MULLER, 1989 rappelle que 'ce' alterne avec 'quoi' lorsqu'il y a ellipse du verbe, comme en (68) ou lorsque le verbe est à l'infinitif, comme en (69).

(68) Il est arrivé quelque chose, mais je ne sais pas **quoi**.

(69) Je me demande **quoi** faire.

Par ailleurs, 'ce' est facultatif lorsqu'il y a une préposition...

(70) Je sais (ce) sur quoi tu travailles.

...sauf avec la préposition 'de'.

(71) J'ignore (*ce) de quoi il est question.

Dans ce cas, on peut aussi employer 'ce dont'.

(72) J'ignore ce dont il est question.

Encore une fois, il est impossible d'utiliser 'ce' si le verbe est à l'infinitif.

(73) Je ne sais pas (*ce) sur quoi travailler.

D'après MULLER, 1996, 'ce que' dans les interrogatives indirectes a le sens de 'l'identité des choses que'.

(74) Je me demande ce que Paul a acheté. ⇒ Je me demande l'identité des choses que Paul a achetées.

L'élément 'ce' des interrogatives indirectes aurait une double interprétation :

1. démonstratif d'identification
2. antécédent de la relative sous-jacente

Sa position est que la subordination se fait bel et bien sur le modèle de la relativisation. Toutefois, la structure se distingue de celle des relatives par le double rôle de l'antécédent particulier 'ce'. J'arriverai à des conclusions similaires dans le Chapitre 7, à savoir que ces interrogatives indirectes sont en fait des relatives sur le plan syntaxique.

2.3.3.5 SALVESEN, 2008

Dans son article intitulé *Ce que and ce que are two different things*, SALVESEN, 2008 affirme au contraire que les similarités entre *ce que* relatif et *ce que* interrogatif ne sont que superficielles, et que leur structure syntaxique sous-jacente est distincte. Elle va jusqu'à affirmer qu'il y a deux compléments différents.

D'après elle, il est impossible de dire qu'il s'agit de deux constructions identiques, car une relative libre n'autorise pas de verbe à l'infinitif, contrairement à une interrogative indirecte.

(75) Je ne sais que faire.

Elle précise qu'une proposition à temps non finie ne peut jamais être introduite par 'ce que'.

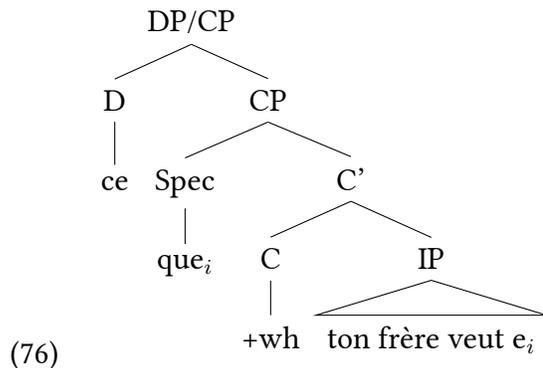
Elle se place dans une approche cartographique, et propose que les interrogatives indirectes introduites par 'ce que' sont comme toutes les autres interrogatives avec un élément wh- qui s'est déplacé dans IntP ou WhP, et avec la position FinP remplie par le complément 'que'. Elle suggère que 'ce' fait partie de la phrase matrice, mais ajoute : « *If this is the case, it is difficult to understand what mechanism could possibly attach an interrogative clause to a cataphoric pronoun. I leave this question for future work.* » (p.20)

Dans la section qui suit, je vais présenter plusieurs hypothèses que l'on peut formuler pour tenter de résoudre ce problème.

2.3.4 'Ce', un renforçateur de 'que' ?

Une première hypothèse assez intuitive que l'on pourrait faire est la suivante : étant donné le caractère faible de 'que', 'ce' pourrait être une sorte d'élément de soutien, de renforçateur du clitique 'que'.

Si 'que' dans 'ce que' était justement le mot *wh-* 'que' et 'ce' un élément de soutien, les lectures de relative libre et d'interrogative indirecte pourraient être le résultat d'une ambiguïté d'étiquetage, comme illustré dans l'arbre syntaxique en (76). Si 'ce' donne l'étiquette, la structure résultante est un DP, donc une relative ; si 'que' fournit l'étiquette, la structure résultante est un CP, donc une interrogative indirecte.¹⁰



Mais, comme nous l'avons montré dans le Chapitre 1 (Section 1.3.2.6.2.2), 'que' dans 'ce que' n'a pas la distribution du mot interrogatif, mais bien celle du complémenteur. Cela est vrai pour les relatives à tête allégée, mais également pour les interrogatives indirectes. En effet, l'alternance entre 'ce que' et 'ce qui' dépend là encore de la fonction de l'élément relativisé, et non pas du caractère animé du référent.

- (77) a. Je me demande ce que tu as fait aujourd'hui.
 b. Je me demande ce qui te tracasse.

10. Cette distinction entre deux 'que' a récemment été remise en question par de nombreux travaux (MANZINI et SAVOIA, 2003 pour l'italien 'che', KATO et NUNES, 2009 pour le portugais brésilien 'que', KAYNE, 2014 pour l'anglais 'that', RINKE et ABMANN, 2017 pour le portugais européen 'que', entre autres). 'Que' serait toujours *wh-*, et varierait selon le type de variable qu'il lie (argumentale et de constituant dans les interrogatives indirectes et les relatives libres ; propositionnelle dans les complétives). Si cela est vrai, alors cette analyse en termes d'étiquetage serait valable. Même si cette analyse est intéressante, je reste sceptique vis-à-vis de l'extension de cette analyse à 'que' en français. En particulier, du fait de l'alternance que-qui, je continue de penser qu'il existe bien un complémenteur 'que' différent du mot *wh-* 'que'.

Comme le montrent les exemples en (77), ‘ce que’ est utilisé quand la lacune est en position d’objet au sein de la proposition subordonnée, et ‘ce qui’ est employé lorsque la lacune est en position de sujet au sein de la proposition subordonnée.

Un autre aspect concerne l’incompatibilité avec l’infinitif. Comme dans le cas des relatives à tête allégée, ‘ce que’ n’est pas compatible avec un verbe à l’infinitif dans les interrogatives indirectes, ce qui prouve bien que ‘que’ dans ‘ce que’ est le complémenteur et pas le mot *wh-*.

(78) *Je me demande ce que cuisiner.

Rappelons d’ailleurs que ‘ce dont’ peut également introduire des questions indirectes et que, selon l’analyse de GODARD, 1988, présentée dans le Chapitre 1 (Section 1.1), ‘dont’ est un complémenteur. Cela nous permet de régulariser *ce+qui/que/dont* et donc cela vient renforcer le fait que c’est le complémenteur qui suit la tête légère ‘ce’.

2.3.5 ‘Ce que’, une partie de ‘qu’est-ce que’?

Une explication possible est que ‘ce que’ dans les interrogatives indirectes soit une partie de ‘qu’est-ce que’. L’idée serait que la partie ‘qu’est’ est élidée (ni prononcée ni présente dans la structure sous-jacente), ou bien seulement non prononcée mais présente dans la structure sous-jacente, ce qui permettrait de conserver le trait *wh-*. En effet, si on analyse chaque élément constitutif de cette formule, on observe que le premier ‘que’ est le mot *wh-*, suivi du verbe ‘être’, suivi du démonstratif faible ‘ce’, puis du complémenteur. Nous pouvons le confirmer avec la nature de l’alternance entre ‘que’ et ‘qui’. Le premier ‘que’ alterne avec ‘qui’ selon que le référent est [+humain] ou [-humain], alors que le deuxième ‘que’ alterne avec ‘qui’ selon que le *gap* est objet ou sujet.

(79) a. **Qu**’est-ce que tu as vu? - Un chat errant.

b. **Qui** est-ce que tu as vu? - Marie.

(80) a. Qu’est-ce **que** tu as fait e?

b. Qu’est-ce **qui** e te dérange?

Le trait *wh-* de cette formule est donc bien porté par le premier ‘que’.

Ainsi, cette hypothèse permettrait de comprendre pourquoi ‘ce que’ serait possible dans des interrogatives indirectes : il y a bien un trait *wh-* dans la structure sous-jacente, dans un [Spec, CP] supérieur à ‘ce que’.

Toutefois, l’un des problèmes de cette hypothèse est que la partie élidée n’est pas « naturelle » en quelque sorte. En effet, ‘ce’ est un clitique accroché au verbe ‘être’ comme dans la forme ‘c’est’ du présentatif, mais ici il y a inversion. Le découpage entre ‘est’ et ‘ce’ est peu probable car ce n’est pas un constituant.

De plus, ‘qu’est-ce que’ semble avoir subi un processus de grammaticalisation, et semble désormais analysable comme un tout, /kɛsk/. En ancien français, le ‘que’ wh- était sûrement fort, et, comme le signale ROUQUIER, 2002, l’élément ‘ce’ pouvait aisément « se promener » dans la séquence en question ¹¹.

- (81) a. que ce est que vos dites ?
b. que est ce que vos dites ?

D’ailleurs, en français parlé, c’est plutôt le deuxième ‘que’ qui est élide dans la séquence, c’est-à-dire le complémenteur.

- (82) Qu’est-ce tu fais ?

La troisième analyse que j’ai proposée pour rendre compte de l’ambiguïté entre relatives et interrogatives indirectes introduites par ‘ce que’ est ce que j’appelle « l’hypothèse de l’incorporation » (KONRAD, A paraître).

2.3.6 L’hypothèse de l’incorporation

Rappelons que la construction ‘ce que’ est une construction ambiguë (83). Cette structure peut en effet être une proposition relative ou bien une interrogative indirecte, comme la construction ‘what’ en anglais (84).

- (83) a. J’aime ce que tu as cuisiné.
b. Je me demande ce que tu as cuisiné.
- (84) a. I like what you cooked.
b. I wonder what you cooked.

Mais l’ambiguïté en (83) est plus surprenante que l’ambiguïté en (84), en particulier à cause du statut nominal de la construction ‘ce que’. Comme nous l’avons déjà évoqué dans la Section 2.6, la disponibilité d’une telle construction dans un contexte sélectionnant une question n’est pas attendue, étant donné que les *verba interrogandi* exigent un CP de nature interrogative. La question essentielle à laquelle nous allons tenter de répondre dans cette sous-section est donc : comment ‘ce que’, qui a une composante nominale, non interrogative, peut-il être légitimé dans une interrogative indirecte ?

Essayer de comprendre la nature de ‘ce que’ implique d’essayer de comprendre la nature des deux éléments qui le constituent, à savoir ‘ce’ d’une part, et ‘que’ d’autre part.

11. Pourtant, d’autres éléments à la même époque semblent indiquer le contraire : on peut trouver, en moyen français, l’orthographe ‘qu’esse que’ (exemple donné par ROUQUIER, 2002 : 114).

(1) Sotouart, qu’esse que tu masches entre tes dens [...] ?
(*Les Vigiles de Triboulet*, p.226, ca. 1480)

2.3.6.1 La nature de ‘ce’

Nous avons déjà vu dans le Chapitre 1 (Section 1.3.2.6.2.1) que ‘ce’ est un démonstratif faible. Il s’agit donc d’un élément nominal, sans trait wh.

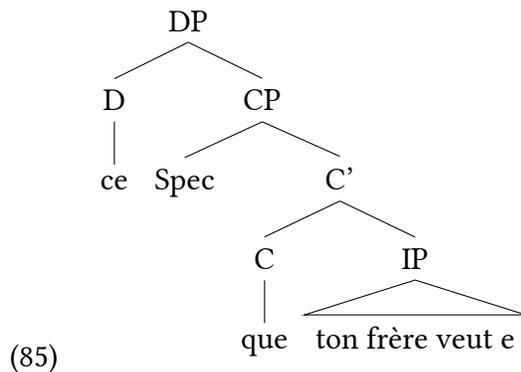
2.3.6.2 La nature de ‘que’

Nous venons de postuler plus haut que le ‘que’ présent dans ‘ce que’ était bel et bien le complémenteur et non pas le mot wh-.

La question centrale qui se pose désormais est alors la suivante : si ni ‘ce’ ni ‘que’ n’a de trait [+wh], comment expliquer que ‘ce que’ soit autorisé dans une question wh- ? Il y a bel et bien un *gap* dans ces interrogatives indirectes, alors quel est l’élément qui se déplace ?

2.3.6.3 Vers une grammaticalisation ?

Comme nous le verrons plus en détails dans le Chapitre 3., consacré à l’étude diachronique que nous avons faite, et selon NIÉGER, 1975, ‘ce que’ est d’abord apparu dans des propositions relatives, avant d’avoir été utilisé dans des interrogatives indirectes, où ‘que’ seul était employé. Partant de cette observation, il semble raisonnable de prendre comme point de départ pour notre analyse une structure relative, avec une tête D (‘ce’) et le complémenteur (‘que’).



La question qui se pose maintenant est la suivante : quel élément se déplace ? En effet, aucun index n’est apparent dans la structure en (85). Est-ce que l’élément vide [e] est une trace de ‘que’ ? Non, puisqu’ici c’est le complémenteur et non pas le mot wh-. On serait tentés de postuler qu’il s’agit d’une trace de ‘ce’¹². Mais ‘ce’ n’est pas wh- et qui dit mouvement wh- dit mouvement d’un élément wh-...

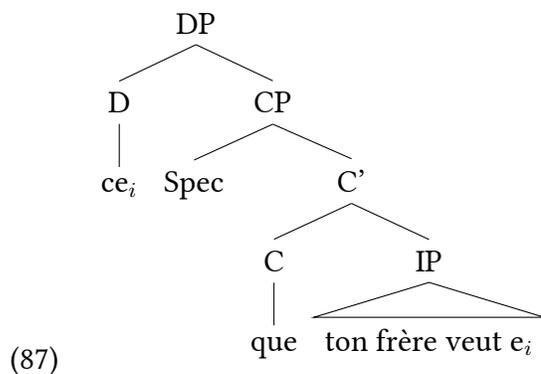
12. Dans une relative, nous avons bien un mouvement de ‘ce’, du moins selon l’analyse par montée décrite dans le Chapitre 1, Section 1.2.1.2.3.

Si ‘ce’ se déplaçait, cela conduirait à un effet d’intervention dans le cadre de la Minimalité Relativisée, qui est un principe de localité énoncé par RIZZI, 1990, puis RIZZI, 2004. Il stipule qu’une relation syntaxique ne peut pas être établie entre deux éléments X et Y si Z est un élément structurellement similaire à X et qui intervient entre X et Y¹³. Ainsi, un élément intervenant, qui porte les mêmes traits que l’élément qui se déplace, va donner lieu à un effet de Minimalité Relativisée. Cela permet par exemple de rendre compte de l’impossibilité d’extraire un élément wh- par-dessus un autre élément wh.

(86) # **Comment**_j tu te demandes **quel problème**_i résoudre e_i e_j ?

Dans l’exemple (86), la phrase est agrammaticale si ‘comment’ est interprété comme un ad-joint du verbe ‘résoudre’. En effet, dans ce cas, ‘comment’, qui a un trait [+wh], rencontre, en se déplaçant, le syntagme wh- ‘quel problème’ sur son chemin. Le principe de Minimalité Relativisée permet directement d’expliquer l’agrammaticalité de cette phrase : ‘quel problème’ agit comme un intervenant sur la trajectoire de mouvement de ‘comment’.¹⁴

Gardons cela en tête et observons maintenant ce qu’il se passerait si ‘ce’ se déplaçait. La structure serait celle en (87).



Le problème avec cette dérivation est que, si ‘ce’ lui-même se déplace, alors on s’attendrait à un effet d’intervention dans les dépendances objet introduites par ‘ce que’. En effet, le mouvement serait déclenché par un trait nominal, trait partagé avec le sujet, possible intervenant, qui devrait interrompre la dérivation, ce qui conduirait à un résultat agrammatical¹⁵.

13. **Relativized Minimality** : $X \alpha$ -governs Y only if there is no Z such that

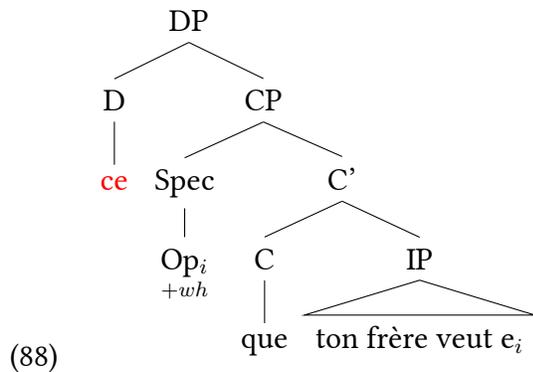
- (i) Z is a typical potential α -governor for Y ,
- (ii) Z c-commands Y and does not c-command X .

[RIZZI, 1990 : 7]

14. Notons que les arguments et les éléments *D-linked* (reliés au discours), échappent à cette généralisation selon laquelle on ne peut pas extraire un élément wh- par-dessus un autre élément wh-. Mais il existe des tentatives d’incorporer ces cas dans une approche dite *featural Relativized Minimality* (voir RIZZI, 2018, entre autres).

15. Précisons que les relatives pleines objet sont elles aussi grammaticales, ce qui signifie qu’il n’y a pas d’intervention du sujet non plus. L’absence d’intervention s’explique par le fait que, selon l’analyse par mon-

Mais les propositions objet introduites par ‘ce que’ sont parfaitement grammaticales en français, tout comme les propositions sujet correspondantes. Nous devons alors assumer que l’élément qui se déplace dans de telles structures n’est pas ‘ce’. On serait donc tentés de rapprocher ‘ce que’ de l’anglais ‘what’, du fait de l’absence d’effet d’intervention. Nous proposons que ‘ce’ est assemblé à la structure en tant que tête externe D. Ce qui se déplace serait alors un opérateur wh-, similaire à ‘what’, comme on peut le voir dans l’arbre en (88).

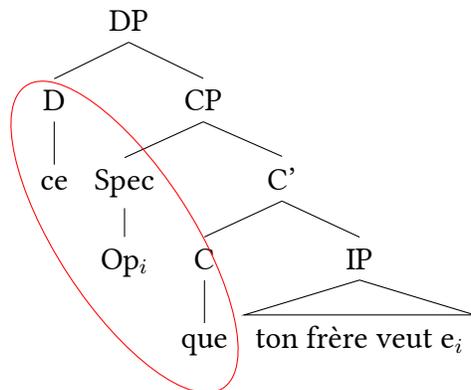


Dans cette configuration, du fait que le DP soit assemblé de façon externe et qu’il y ait un mouvement d’opérateur, aucun effet d’intervention n’est déclenché par le sujet. Mais qu’en est-il des interrogatives indirectes? La structure en (88) est toutefois problématique pour une question car il s’agit d’un DP.

Nous pourrions postuler que ‘ce’ et ‘que’ s’incorporent l’un avec l’autre, et avec l’opérateur A-barre, comme en (89). Toutes les conditions sont réunies pour un processus d’incorporation (BAKER, 1988). Nous avons :

- a) une séquence de têtes (D et C)
- b) une séquence d’éléments faibles
- c) des éléments adjacents

tée, il y a d’abord un mouvement wh. Le sujet n’ayant pas de trait wh, il n’agit pas comme intervenant. Ensuite, il y a la montée de la tête : c’est le NP seul qui se déplace, et le D est assemblé de façon externe à la structure. C’est donc un NP qui croise un DP sur son trajet, et non pas un DP qui croise un autre DP. Quant aux relatives libres objet, elles ont un statut un peu particulier. Les résultats d’un test de jugement d’acceptabilité que j’ai mis en place ont montré qu’elles ne sont pas très naturelles en français (voir KONRAD, A paraître).

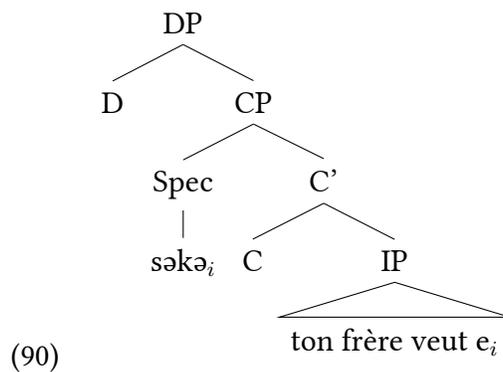


(89)

Le résultat de ce processus d'incorporation serait une forme hybride /səkə/, qui garde les propriétés des trois éléments, i.e. les propriétés de l'élément nominal 'ce', les propriétés de l'opérateur (wh-) et les propriétés du complémenteur 'que'¹⁶.

Mais d'un point de vue structural, cette incorporation peut donner lieu à deux configurations :

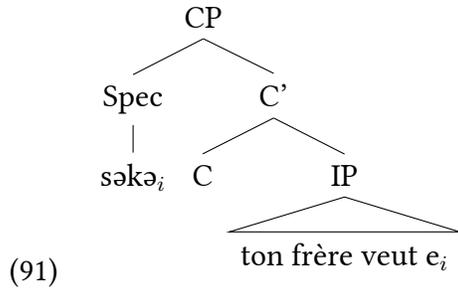
- a) Si la projection D est retenue dans la structure, alors nous avons une relative libre. Cette configuration est présentée en (90).



(90)

- b) Si la projection nominale n'est pas retenue dans la structure, alors nous avons une interrogative indirecte. Ce deuxième cas de figure est représenté en (91).

16. Notons que l'analyse proposée ici est en faveur de l'analyse par appariement des relatives. Du fait de la présence de l'opérateur nul, l'incorporation implique trois éléments. Mais, seuls deux éléments, 'ce' et 'que', seraient impliqués dans une autre version de l'incorporation, s'appuyant sur l'analyse par montée. Si 'ce' subissait la montée, alors il n'y aurait pas d'opérateur nul. Le problème de cette analyse, est, comme nous l'avons vu, que l'on s'attendrait à des effets d'intervention du sujet, ce qui n'est pas le cas.



D'après KONRAD, A paraître, les relatives libres possèdent dans leur structure sous-jacente une tête D externe comme (90). Ainsi, les propositions relatives introduites par 'səkə' en français sont similaires aux relatives libres introduites par 'what' en cela qu'elles sont toutes les deux sélectionnées par une tête nulle D selon l'*embedding account*. Mais les éléments 'səkə' et 'what' sont de nature différente. Alors que 'what' est pleinement interrogatif, i.e. il a un trait wh- fort, 'səkə' est partiellement interrogatif, i.e. il a un trait wh- faible. Ce trait wh- faible ne peut faire surface que dans un contexte où il est sélectionné, lorsqu'un verbe interrogatif peut le valuer. C'est pour cela que 'səkə' ne peut pas introduire de question directe¹⁷, comme en (92).

(92) *Ce que tu aimes ?

Une incorporation de ce type est certainement attestée dans d'autres langues romanes, comme le portugais et le florentin, mais à un degré différent. En effet, dans ces langues, contrairement au français, l'équivalent de 'ce que' peut introduire une question directe. C'est le cas de 'o que' en portugais et de 'icché' en florentin, comme illustré dans les exemples (93) et (94), respectivement.

(93) O que quer o senhor ?
 le que vouloir.3SG le monsieur
 Qu'est-ce que veut le monsieur ?

(94) Icché vói che facciano ?
 le-que vouloir.2SG que faire.3PL
 Qu'est-ce que tu veux qu'ils fassent ?

De plus, 'o que' et 'icché' ne sont pas incompatibles avec un verbe à l'infinitif, ce qui prouve leur nature wh-.

(95) O que fazer ?
 le que faire.INF

17. Sauf question de clarification.

(1) ce que je préfère ? Je préfère le chocolat noir.

Que faire ?

- (96) Icché fare ?
le-que faire.INF
Que faire ?

Cela peut être expliqué comme le résultat d'un processus de grammaticalisation (ROBERTS et ROUSSOU, 2003), processus qui se traduit ici par une incorporation de l'élément D : 'o que' et 'icché' sont ré-analysés comme des éléments wh-. C'est pourquoi ces éléments peuvent entrer à la fois dans une interrogative indirecte et dans une proposition relative, comme on peut le voir dans les exemples (97a), (97b) et (98a), (98b) en portugais et en florentin, respectivement.

- (97) a. O senhor pergunta o que caiu na sala.
le monsieur demande.PST.3SG le que tomber.PAST.3SG dans-la salle
Le monsieur demande ce qui est tombé dans la salle.
- b. O senhor apanha o que caiu na sala.
le monsieur attrape.PST.3SG le que tomber.PAST.3SG dans-la salle
Le monsieur attrape ce qui est tombé dans la salle.
- (98) a. Vorrei sapere icché la viene a fare sempre
vouloir.COND.1SG savoir.INF le-que elle venir.PST.3SG à faire.INF toujours
qui.
ici
Je voudrais savoir ce qu'elle fait toujours par ici.
- b. Fà icché ti pare.
faire.IMP.2SG le-que te plaire.PST.3SG
Fais ce qu'il te plaît.

Mais il reste à déterminer si le 'que'/'che' est le complémenteur ou le mot wh-, ce qui n'est pas facile en particulier dans les langues romanes étant donné l'homophonie entre les deux. Si nous reprenons les arguments que nous avons discutés pour le français dans la Section 2.3.4, nous pourrions conclure qu'il s'agit de l'élément wh-. Mais il existe l'alternative suivante : si 'que'/'che' est le complémenteur, alors on doit postuler la présence d'un opérateur wh- nul. Autrement dit, à l'origine, avant incorporation, 'o que'/'icché' est analysable soit comme D+wh-, soit comme D+Op_{wh-}+C, comme les structures (86) et (87) postulées plus haut. Nous laissons cette question ouverte.

Après ce processus de ré-analyse, la séquence D+wh- ou D+Op_{wh-}+C est interprétée comme une unité, *o-que* ou *icché*, et non plus comme des unités distinctes. Selon cette analyse, ‘o que’ et ‘icché’ sont de nouveaux items wh- (cf. MEDEIROS JUNIOR, 2016 pour le portugais brésilien, MUNARO, 2001 pour les dialectes du Nord-Ouest de l’Italie, MANZINI et SAVOIA, 2005 et MANZINI et SAVOIA, 2011 pour diverses variétés romanes). Cela peut alors être perçu comme un autre cas d’ambiguïté d’étiquetage : l’élément wh- peut donner ou non son étiquette à la structure, ce qui donne lieu à une interrogative directe ou bien à une proposition relative. Rappelons qu’en français, ‘ce que’ ne peut pas introduire de question directe et donc une telle analyse ne peut pas s’étendre complètement au cas du français. La grammaticalisation aurait lieu dans toutes les langues romanes que nous avons évoquées, mais à des stades plus (portugais, florentin) ou moins (français) avancés. Le français serait alors dans une sorte de stade intermédiaire, vers la naissance d’un nouvel élément complètement wh-. La seule différence entre le français et les autres langues romanes évoquées serait la force de l’élément wh-.

D’autres éléments pourraient-ils aller dans le sens de cette hypothèse d’incorporation ? Il y a le fait que l’insertion de matériel lexical entre ‘ce’ et ‘que’ n’est pas possible sans que le sens en soit modifié. ‘Ce’ est indissociable de ‘que’. Il ne peut pas apparaître sans modification restrictive, sauf dans le cas où il est cliticisé au verbe ‘être’.

(99) *Ce me tracasse.

(100) C’est magnifique.

Encore plus intéressant, l’insertion de matériel lexical (comme un adverbe) était possible en français médiéval, ce qui montre qu’il y a bien eu une évolution (nous verrons ceci plus en détails dans le Chapitre 3, consacré à l’étude diachronique de ‘ce qu-’ dans les relatives et les interrogatives indirectes).

(101) Et pensai a *ce longuement qu’il m’iert* avenu. [Froiss. Esp. 625-626, cité par Jokinen 1986 : 172]

Par ailleurs, notons que ‘ce que’ est également utilisé dans des phrases exclamatives, phrases dans lesquelles ‘ce’ n’a en rien la fonction d’un démonstratif. Il semble préférable d’analyser ‘ce que’ comme un tout dans des exemples du type (102).

(102) Ce que c’est beau !

Là encore, il est important de préciser que cet emploi n’apparaît qu’au XIX^{ème} siècle.

Toutefois, des propriétés syntaxiques importantes de ‘ce que’ semblent indiquer que cette hypothèse n’est pas la bonne analyse. Outre le fait que ‘ce que’ ne peut pas introduire de questions racines, il y a le fait que ‘ce que’ ne puisse pas être coordonné à un mot wh- tel que ‘qui’. Nous savons qu’il n’est possible de coordonner deux éléments que s’ils sont de même nature syntaxique.

(103) ??J’aime **qui** et **ce que** tu aimes.

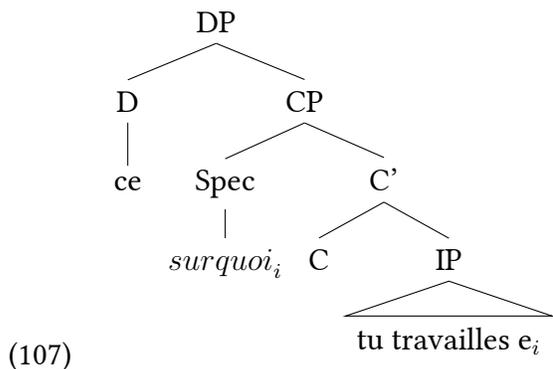
(104) ??J'aime **ce que** et **qui** tu aimes.

Par ailleurs, on trouve des structures dans lesquelles 'ce' est compatible avec un élément wh-, lorsqu'il y a *pied-piping*. Notons que 'ce' est optionnel ici.

(105) Je me demande **(ce) sur quoi** tu travailles.

(106) Je me demande **(ce) à quoi** tu penses.

Ces exemples montrent clairement que la projection D peut être présente et réalisée en dehors du CP dans les interrogatives indirectes, ce qui est problématique dans la perspective d'une incorporation où ce+que seraient dans [Spec, CP].



2.4 Du français aux autres langues romanes

Dans le Chapitre 1, nous avons vu que les relatives à tête allégée existent dans les autres langues romanes et nous avons décrit leurs différences et similitudes syntaxiques. Une question que l'on peut se poser est la suivante : est-ce que cette ambiguïté surprenante entre relative et interrogative indirecte est propre à 'ce que' en français, ou bien retrouve-t-on cette ambiguïté dans d'autres langues romanes, qui ont des relatives à tête allégée ?

Afin de répondre à cette question, il est important de prendre en compte une distinction sémantique que nous n'avons pas évoquée jusqu'à présent. On fait en effet la distinction entre deux types d'interrogatives indirectes, selon que l'on a une interprétation intensionnelle ou bien une interprétation extensionnelle. Dans le premier cas, d'après SUÑER, 1993, lorsque le verbe de la proposition matrice a une force intensionnelle, cela signifie qu'il s'intéresse à l'intension du complément, c'est-à-dire à la question elle-même. Un exemple d'un verbe de ce type est 'se demander'. Dans le second cas, lorsque le verbe de la proposition matrice a une force extensionnelle, cela signifie qu'il s'intéresse à l'extension du complément, c'est-à-dire à la réponse à la question. Un exemple d'un verbe de ce type est 'savoir'.

(108) Marie se demande qui a téléphoné.

⇒ cela est vrai si et seulement si Marie veut savoir, et ne sait pas encore, laquelle des possibles réponses à *qui a téléphoné* est vraie

(109) Marie sait qui a téléphoné.

⇒ cela est vrai si et seulement si Marie connaît la réponse vraie à *qui a téléphoné*

Cette distinction sémantique, qui vient à l'origine de GROENENDIJK et STOKHOF, 1982, a été reprise en d'autres termes par GINZBURG et SAG, 2000. Ils font la différence entre les propositions ouvertes (« *questions* »), sélectionnées par des verbes tels que *wonder* ('se demander'), et les questions résolues (« *facts* »), sélectionnées par les autres verbes, tels que *say* ('dire') et *find out* ('trouver'). Ils proposent une sélection de nature plutôt sémantique (interrogatives de type question *versus* interrogatives de type proposition). GODARD et ABEILLÉ, 2011 ont montré qu'en français, seules celles après les verbes 'chercher', 'se demander' et 'ne pas savoir' sont interprétées comme des propositions ouvertes. Les autres, après des verbes tels que 'dire', 'trouver' et 'savoir', sont interprétées comme des questions résolues, donc comme des faits.

Comme on peut le voir dans le Tableau 2.1, certaines langues romanes, comme le portugais, le catalan et l'occitan, présentent la même ambiguïté qu'en français, et la séquence D+C peut introduire des questions extensionnelles et des questions intensionnelles. Par contre, dans d'autres langues romanes comme l'italien et l'espagnol, l'ambiguïté n'existe pas avec de *pures* interrogatives, c'est-à-dire celles intensionnelles, introduites par des verbes comme 'se demander'. Par contre, la séquence D+C peut être utilisée dans les questions extensionnelles, qui ressemblent en fait davantage à des relatives, et donc à des DPs.¹⁸

Le roumain semble être un cas un peu à part au sein des langues romanes. Nous avons déjà vu, que, pour les relatives avec référent inanimé, seul le roumain, parmi les langues romanes, avait la possibilité d'employer des relatives libres, à la *what*, introduites par le mot *wh- ce*, ou la possibilité d'utiliser des relatives à tête allégée, introduites par la séquence D+wh, *ceea ce*. En ce qui concerne les interrogatives indirectes, il est impossible en roumain d'employer *ceea ce*, que ce soit dans les questions extensionnelles ou dans les questions intensionnelles.

18. Les données présentées ci-après sont des communications personnelles qui m'ont été fournies par des locuteurs natifs, venant de :

1. Eunice Fernandes pour le portugais
2. Joan Busquets pour le catalan
3. Caterina Donati pour l'italien
4. Israel de La Fuente pour l'espagnol
5. Carmen Dobrovie-Sorin pour le roumain

que je tiens à remercier.

Langues	Q extensionnelles	Q intensionnelles
Français	✓	✓
Portugais	✓	✓
Catalan	✓	✓
Occitan	✓	✓
Italien	✓	X
Espagnol	✓	X
Roumain	X	X

TABLE 2.1 – Possibilité d’avoir ‘élément nominal + C’ dans les questions extensionnelles et dans les questions intensionnelles dans les langues romanes

Français

- (110) a. *Question extensionnelle* : Je sais **ce que** tu as acheté.
 b. *Question intensionnelle* : Je me demande **ce que** tu as acheté.

Portugais

- (111) a. *Question extensionnelle* : Sei **o que** compraste.
 b. *Question intensionnelle* : Pergunto-me **o que** compraste.

Catalan

- (112) a. *Question extensionnelle* : Sé **el que** has comprat.
 b. *Question intensionnelle* : Em pregunto **el que** has comprat.

Occitan

- (113) a. *Question extensionnelle* : Que sabi **ço qu’**as crompat.
 b. *Question intensionnelle* : Que’s demandan **ço qu’**as crompat.

Italien

- (114) a. *Question extensionnelle* : So **quello que** hai comprato.
 b. *Question intensionnelle* : *Mi chiedo **quello que** hai comprato.

Espagnol

- (115) a. *Question extensionnelle* : Sé **lo que** compraste.
 b. *Question intensionnelle* : *Me pregunto **lo que** compraste.

Roumain

- (116) a. *Question extensionnelle* : *Stiu ceea ce ai cumparat.
 b. *Question intensionnelle* : *Ma întreaba ceea ce ai cumparat.

Il est intéressant de noter que le français a connu une époque où il y avait une distinction nette entre questions extensionnelles et questions intensionnelles : comme nous le verrons plus en détails dans le Chapitre 3, sur la diachronie de ‘ce que’, ‘ce que’ s’est imposé dans les questions extensionnelles avant de s’imposer dans les questions intensionnelles. Le français était donc comme l’espagnol et l’italien de ce point de vue.

Au vu des différences déjà observées entre le français et le portugais dans le cadre de l’hypothèse de l’incorporation, il me paraît intéressant d’étudier les propriétés syntaxiques des questions intensionnelles dans les deux langues. En effet, ‘o que’ semble être réanaly-sable comme un mot *wh-* en portugais, de façon plus convaincante que ‘ce que’ en français. Les propriétés pertinentes sont les suivantes :

1. **Question INF** : est-ce que le verbe de l’interrogative indirecte introduite par cette séquence D+C peut être à l’infinitif?
2. **Matching** : est-ce que ces interrogatives indirectes sont contraintes par la nécessité de *matching*?
3. **Coordination avec mot wh-** : est-ce que la séquence D+C peut être coordonnée avec un mot *wh-* dans les interrogatives indirectes ?

Illustrons chaque propriété avec des exemples dans les deux langues. Notons + si la propriété est présente dans la langue, et - si elle est absente.

Question INF

Français : -

- (117) *Je me demande ce que manger.

Portugais : +

- (118) Pergunto-me o que comer.

Matching

Français : -

- (119) *Je me demande de ce que tu vas parler.

Portugais : +

(120) Pergunto-me do (*de+o*) que estás a falar.

Coordination avec mot wh-

Français : -

(121) ??Je me demande qui et ce que tu détestes.

(122) ??Je me demande ce que et qui tu détestes.

Portugais : +

(123) Pergunto-me quem e o que odeias.

(124) Pergunto-me o que e quem odeias.

On remarque l'absence de ces propriétés en français et leur présence en portugais, ce qui semble aller dans le sens d'une réanalyse de 'o que' en un mot wh- dans le cas du portugais, mais pas dans le cas du français.

2.5 Conclusion intermédiaire

Pour conclure, nous avons vu dans ce Chapitre que les subordonnées introduites par 'ce qu-' sont tout aussi ambiguës en français que les subordonnées introduites par 'what' en anglais, ce qui est problématique et surprenant d'un point de vue syntaxique. Bien que ce problème ait souvent été laissé de côté dans la littérature, des analyses intéressantes ont été proposées dans le passé, sans pour autant pouvoir expliquer entièrement les données. Après avoir discuté la possibilité de 'ce' comme renforçateur de 'que' et la possibilité de l'élision de 'ce que' dans la formule 'qu'est-ce que', nous avons décrit l'hypothèse de l'incorporation qui semble attractive.

Afin d'éclaircir le statut syntaxique de ces interrogatives indirectes introduites par 'ce que', j'ai mené, dans un premier temps, une étude diachronique, consistant à observer de plus près l'évolution et l'apparition de 'ce' dans les relatives et les interrogatives indirectes. Cette étude est détaillée dans le Chapitre 3 qui suit, qui constitue à lui seul la deuxième partie de ma thèse. Dans un deuxième temps, j'ai étudié l'acquisition de ces structures par les enfants français (Chapitre 4), afin de déterminer quelle analyse ils en font, en faisant une comparaison avec le portugais européen (Chapitre 5), où 'o que' semble avoir subi une grammaticalisation. Dans un troisième temps (Chapitre 6), je me suis centrée sur le traitement de ces structures par les adultes, qui peut également être révélateur de l'analyse syntaxique qu'ils en font. Dans un quatrième (et dernier) temps, je me suis focalisée sur certaines propriétés syntaxiques et sémantiques en synchronie.

Ainsi, l'étude diachronique, les tests en acquisition, les expériences sur le traitement, et les propriétés en synchronie des interrogatives indirectes introduites par 'ce que' devraient aider à déterminer si ces structures sont à analyser comme étant de réelles interrogatives indirectes introduites par un élément *wh- səka*, situé en Spec CP, ou bien si ces interrogatives sont en réalité ce qu'elles semblent être : des DPs, donc des relatives...

Deuxième partie

Évolution diachronique des relatives à
tête allégée et interrogatives indirectes
introduites par ‘ce qu-’

Chapitre 3

Étude diachronique : une apparition tardive et soudaine ¹

3.1 ‘Ce qu-’ dans une perspective diachronique

La littérature sur ‘ce qu-’ en diachronie est plus riche que la littérature sur ‘ce qu-’ en synchronie. La plupart des travaux sont assez anciens et descriptifs, sans qu’il y ait une étude systématique de l’ancien français au 17^{ème} siècle, sans analyse statistique. Parmi ces études, je vais présenter celles de NIÉGER, 1975, LEARD, 1987 et PIERRARD, 1995.

3.1.1 NIÉGER, 1975

Monique Niéger, dans son article intitulé ‘L’interrogation indirecte : étude diachronique’ (1975), mentionne déjà une évolution qui retient notre attention : « l’extension des groupes *ce qui, ce que* aux interrogations indirectes ». La section 3. de son article s’intitule ‘Influence des relatives sur les interrogatives indirectes’. ‘Ce qu-’ a donc d’abord été employé dans les relatives, puis ce système s’est étendu aux interrogatives indirectes, introduites auparavant par ‘qui’ et ‘que’ seuls, sans le démonstratif ‘ce’. Elle précise que ce changement dans les interrogatives indirectes se produit vers le milieu du 17^{ème} siècle.

« Brunot rapporte de nombreux autres exemples qui nous permettent de situer le changement vers le milieu du 17^{ème} siècle. Il est évident que le remplacement de ‘que’ par ‘ce que’ n’a pas été automatique et qu’il ne faut pas s’étonner de rencontrer, à cette époque, deux systèmes qui se font concurrence. » (p.9)

Elle signale également, qu’aux 16^{ème} et 17^{ème} siècles, les formes en ‘qu’est-ce que / qu’est-ce qui / qui est-ce qui / qui est-ce que’, employées pour introduire des interrogatives

1. L’étude diachronique décrite dans le présent chapitre a été présentée lors d’une conférence dédiée à la linguistique historique, les *Rencontres lyonnaises des jeunes chercheurs en linguistique historique* (voir KONRAD, 2019).

directes, sont fréquemment utilisées pour introduire des interrogatives indirectes, mais sans inversion : que c'est que / que c'est qui / qui c'est que / qui c'est qui. Il s'agit en réalité d'interrogatives clivées en quelque sorte, mais Niéger n'emploie pas cette terminologie.

« Cependant, à cause de la répugnance des grammairiens et du bon usage en général pour les formes en 'qu'est-ce que', cette évolution a été contrecarrée et la langue a cherché à exprimer les mêmes nuances à l'aide de formes simples. On y réussit en généralisant le système des relatives aux interrogatives : les formes 'ce qui', 'ce que' auront l'avantage d'exprimer clairement la fonction en précisant la nature [- animé] du sujet ou de l'objet (ce/Ø). » (p.11)

Ainsi, NIÉGER, 1975 met en évidence l'influence considérable des relatives sur les interrogatives indirectes, le système des relatives s'étant étendu à celui des interrogatives indirectes, avec le remplacement du 'que' interrogatif par 'ce que'.

3.1.2 LEARD, 1987

Jean-Marcel Léard s'est lui aussi intéressé à 'ce que' en diachronie, en particulier dans son article intitulé 'Le développement de *ce que* en français et l'évolution du relatif - interrogatif - exclamatif' (1987). Il qualifie à juste titre 'ce que' de forme « bizarre et polyvalente ». Il affirme que seuls les emplois relatifs de 'ce que' ont été étudiés de façon abondante. 'Ce que' s'est étendu dans les emplois relatifs, mais aussi dans les emplois interrogatifs et les emplois quantitatifs.

D'après lui, le système du relatif-interrogatif est constitué de formes simples (c'est-à-dire sans 'ce') jusqu'à la fin du 12^{ème} siècle. Lui aussi signale que 'que' a fortement reculé dans l'interrogation directe et indirecte et qu'il est remplacé par 'ce que', 'qu'est-ce que'.

Il postule que 'ce' apparaît pour donner un statut neutre à une forme prédicative. 'Ce' enlève ainsi à 'qui' sa valeur prédicative, par exemple, et donc sa valeur animée. Cela n'est pas clair : qu'entend-il par « valeur prédicative » ? Il invite à ne pas traiter 'ce que', 'ce qui' de la même manière que 'ce sur quoi', 'ce à quoi' (nous avons déjà évoqué ces constructions dans le Chapitre 2., Section 2.3.6.3, et nous y reviendrons dans le Chapitre 7).

Par ailleurs, il insiste sur le fait que 'ce que' n'a pas qu'une simple valeur de relatif, mais qu'il a aussi une valeur quantitative. 'Ce que' peut alors être paraphrasé par 'la quantité que', et non pas seulement et systématiquement par 'la chose que'.

3.1.3 PIERRARD, 1995

Michel Pierrard a également écrit plusieurs articles sur 'ce que', que ce soit en synchronie ou en diachronie. L'article pertinent pour le présent chapitre est celui intitulé 'Système

et évolution du système : l'évolution de *ce que* introducteur de subordonnées'. 'Ce que' entretient des rapports complexes avec 'que' :

« *Le fonctionnement du morphème complexe 'ce que' doit nous interpeller puisqu'il s'impose successivement comme structure parallèle, comme substitut ou comme alternative à certains emplois de 'que'.* » (p.1)

De façon légitime, Pierrard se pose la question suivante : peut-on attribuer aux divers emplois de 'ce que' une valeur morpho-sémantique unique ou bien sont-ils le produit de l'homonymie, autrement dit du hasard ? Il est vrai que l'on ne trouve que très peu d'études qui ont tenté de proposer une analyse qui prenne en compte les différents emplois de 'ce que' dans les subordonnées, rappelés ci-dessous.

1. Conjonctif²

(1) Je m'attends à ce qu'il parte vivre à l'étranger.

2. Relatif

(2) J'ai lu ce que tu as écrit.

3. Interrogatif

(3) Il ne sait pas ce qu'il veut.

4. Exclamatif

(4) C'est fou ce qu'elle est belle !

On se concentre ici sur les emplois relatif et interrogatif, qui sont les deux emplois qui nous intéressent dans le cadre de l'ambiguïté étudiée.

En ce qui concerne l'emploi relatif, PIERRARD, 1995 nous apprend que 'ce que' est apparu avant 'ce qui', qui n'apparaît qu'à partir du 13^{ème} siècle. Cela veut dire que, en Ancien Français, 'ce que' est employé tant dans les dépendances objet que dans les dépendances sujet. Par ailleurs, il affirme que le choix entre 'ce que' et 'que' est plutôt libre en Ancien Français, et que, en Moyen Français, 'ce que' est définitivement établi. A l'époque où les formes en 'qu-' et celles en 'ce qu-' étaient possibles, il existait tout de même une nuance interprétative, d'après PIERRARD, 1995 : 'ce qu-' avait un sens davantage assertif ('cela, à savoir qui/que').

De plus, 'ce' serait facultatif en Ancien Français et garderait une certaine autonomie. Dans de nombreux cas, 'ce' semble être un morphème autonome à l'époque, d'où la séparation possible de 'ce' et 'que' par un verbe, un adverbe ou un prédicat.

2. La terminologie est de Pierrard. On dirait plutôt 'complétif'.

- (5) Tu as tout **ce** bien entendu **que** Sevestre m'a répondu. [exemple donné par PIERRARD, 1995 : 115]
Equivalent en Français Moderne : Tu as bien entendu tout ce que Sevestre m'a répondu.

Non seulement, 'ce' semble être autonome à l'époque, mais il semble aussi être fort. Il pouvait en effet être employé dans des contextes où 'cela/ceci' sont requis. L'apparition au 16ème siècle de ces formes fortes va d'ailleurs de pair avec l'affaiblissement de 'ce'. Au 17ème siècle, on ne trouve quasiment plus d'emplois de 'ce' séparé de 'que'.

Quant à l'emploi interrogatif, PIERRARD, 1995 parle d'une extension du 'qui' interrogatif à d'autres emplois de l'inanimé. Ces emplois subsistent jusqu'au 17ème siècle, mais ils portent à la confusion entre deux informations distinctes : d'une part le caractère animé et d'autre part la fonction (sujet/objet). D'où l'alignement, finalement, sur le système des relatives, avec les formes 'ce qui' et 'ce que', pour les fonctions de sujet et d'objet, respectivement. Il affirme que le remplacement définitif et obligatoire de 'que/qui' par 'ce que/ce qui' a lieu à la fin du premier tiers du 17ème siècle, et que cela est en grande partie dû au lien traditionnel entre les interrogatives indirectes et les relatives et au poids de la norme imposant 'ce que'.

A noter que, toujours d'après PIERRARD, 1995, les formes 'ce qu-' apparaissent dès l'Ancien Français, mais comme des formes parmi tout un tas d'autres alternatives, comme les formes 'qu-', des formes disloquées, des formes combinées.

- (6) si li demande **ki ce est qui** gist en la litière [Lancelot 13.2, exemple tiré de PIERRARD, 1995, p. 117]

Ces formes alternatives se trouvent dans les interrogatives indirectes mais également dans les interrogatives directes. Aux 16ème et 17ème siècles apparaissent également des constructions identiques à celles utilisées dans l'interrogation directe, c'est-à-dire des interrogatives indirectes introduites par les formes 'qui/que est-ce qui/que'.

Les formes 'ce qu-' finiront par s'imposer en langue standard aux 16ème et 17ème siècles. Cette évolution a été favorisée par la transformation de 'ce' en élément enclitique (comme dans 'c'est'). Entre 'ce que' et 'qu'est-ce que', la norme imposera également 'ce que' dans les interrogatives indirectes, ce qui mène à une différenciation entre les formes de l'interrogation directe et celles de l'interrogation indirecte.

Il est impossible de faire le tour de toutes les études portant sur la diachronie de 'ce qu-', mais je tiens à mentionner également la thèse de Magali Rouquier intitulée 'Les Ce qu-phrases en ancien français' (ROUQUIER, 1988), l'article de Pierre Kunstmann décrivant 'Le système du relatif-interrogatif en ancien français' (KUNSTMANN, 1990) et l'article de Laurie Zaring et Paul Hirschbühler 'Qu'est-ce que *ce que*? The diachronic evolution of a French complementizer' (ZARING et HIRSCHBÜHLER, 1997), entre autres.

3.2 Étude diachronique sur ‘ce qu-’

3.2.1 Objectifs, questionnements

Le but de mener une étude diachronique sur ‘ce qu-’ en français est de déterminer quand et selon quelle chronologie une structure ouvertement nominale est passée à pouvoir exprimer une sémantique interrogative assumant une distribution de phrase. Les données diachroniques devraient également permettre de vérifier la validité (ou du moins la plausibilité) de l’hypothèse de l’incorporation présentée dans le Chapitre 2 (Section 2.3.5).

3.2.2 Prédictions

Si ‘ce que’ dans les interrogatives indirectes est à analyser comme un mot *wh- səkə* (hypothèse de l’incorporation), on s’attend à ce que ‘ce que’ soit utilisé en parallèle dans les deux structures (relatives et interrogatives indirectes). Si, au contraire, ‘ce que’ est bien une séquence D+C, on s’attend à ce que cette séquence apparaisse d’abord dans les relatives. On s’attend par ailleurs à une certaine résistance à étendre cette séquence de nature clairement nominale dans les interrogatives indirectes.

3.3 Une étude en deux temps

3.3.1 Une étude globale à partir d’un corpus étiqueté

3.3.1.1 Le corpus

Dans un premier temps, j’ai conduit une étude de corpus globale à partir de MCVF (*Modéliser le Changement : les Voies du Français*), corpus du français ancien, structuré et annoté morphosyntaxiquement, et dirigé par France Martineau, de l’Université d’Ottawa. Ce corpus contient 19 textes pour la période Moyen-Âge, que j’ai choisi de subdiviser en deux périodes : Ancien Français (<1330, 13 textes) et Moyen Français (>1330, 6 textes), afin de pouvoir repérer une évolution éventuelle dans le passage de l’Ancien Français au Moyen Français.

3.3.1.2 Ce que j’ai cherché

J’ai utilisé le logiciel Tregex (abréviation de *tree regular expressions*), qui permet d’interroger des *treebanks* assez facilement, en utilisant des expressions régulières. Il propose une interface graphique pour éditer les requêtes et visualiser les résultats en texte et sous forme d’arbres syntaxiques. Comme le corpus MCVF est annoté morphosyntaxiquement, j’ai pu chercher toutes les relatives introduites par ‘ce qui / ce que’ ou ‘qui / que’, ainsi que toutes les interrogatives indirectes introduites par les mêmes formes.

3.3.1.3 Les résultats

Les résultats de cette première étude sont présentés dans le Tableau 3.1 et la Figure 3.1 pour la période Ancien Français et dans le Tableau 3.2 et la Figure 3.2 pour la période Moyen Français.

	CE QU-	QU-
<i>RELATIVES</i>	124 (89,2%)	15 (10,8%)
<i>INTERROGATIVES</i>	1 (2,2%)	45 (97,8%)

TABLE 3.1 – Nombre d’occurrences et pourcentages de ‘ce qu-’ (dans les relatives à tête allégée et les interrogatives indirectes) *versus* ‘qu-’ (dans les relatives libres et les interrogatives indirectes avec référent inanimé), pour la période Ancien Français.

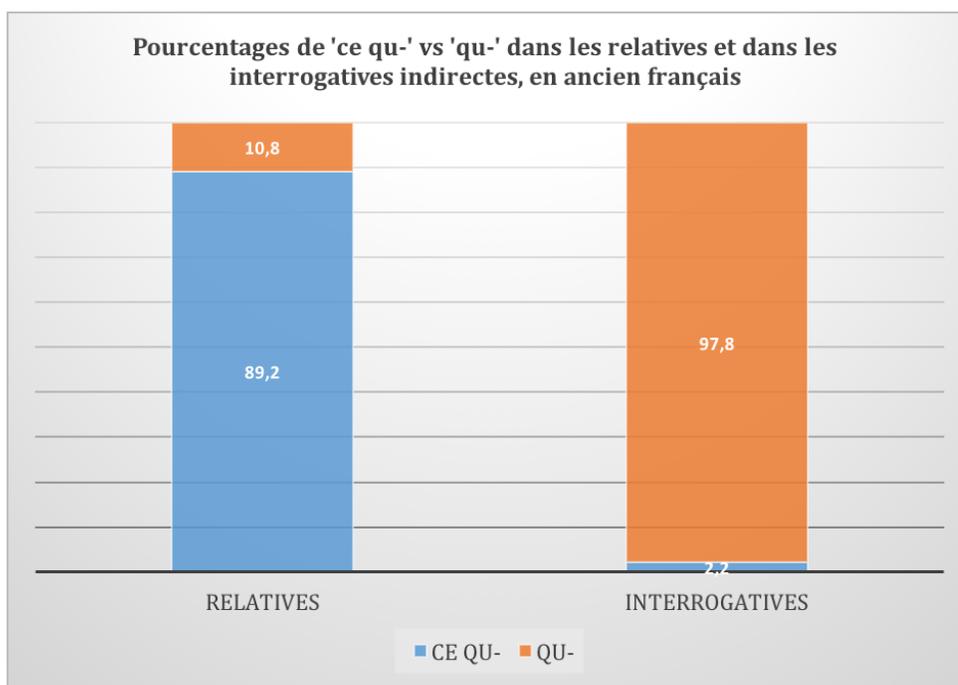


FIGURE 3.1 – Pourcentages de ‘ce qu-’ (dans les relatives à tête allégée et les interrogatives indirectes) *versus* ‘qu-’ (dans les relatives libres et les interrogatives indirectes avec référent inanimé), pour la période Ancien Français.

	CE QU-	QU-
<i>RELATIVES</i>	295 (97,4%)	8 (2,6%)
<i>INTERROGATIVES</i>	8 (6,5%)	116 (93,5%)

TABLE 3.2 – Nombre d’occurrences et pourcentages de ‘ce qu-’ (dans les relatives à tête allégée et les interrogatives indirectes) *versus* ‘qu-’ (dans les relatives libres et les interrogatives indirectes avec référent inanimé), pour la période Moyen Français.

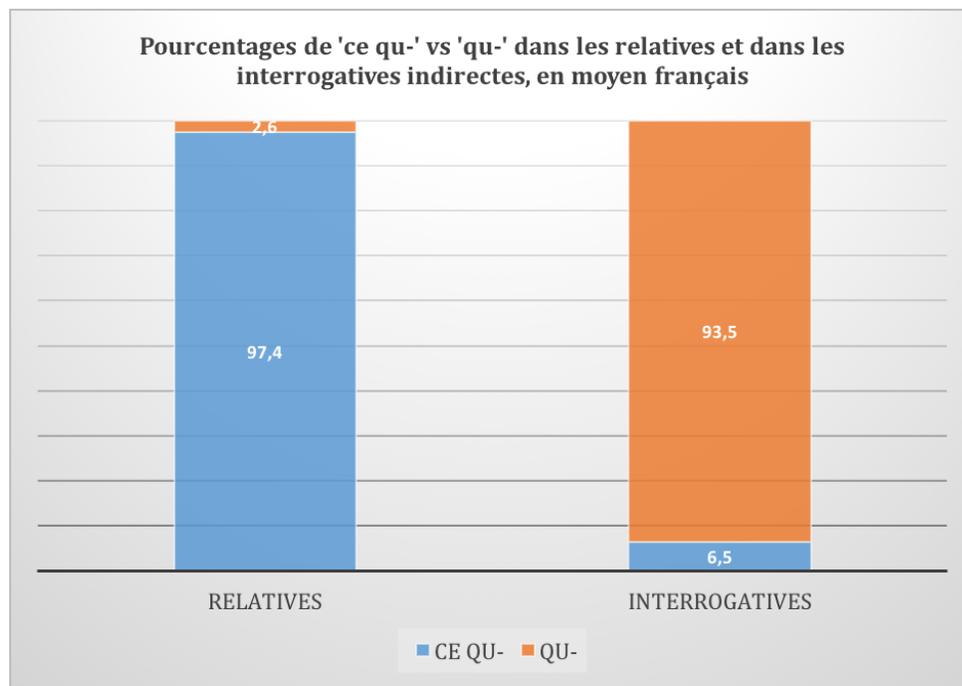


FIGURE 3.2 – Pourcentages de ‘ce qu-’ (dans les relatives à tête allégée et les interrogatives indirectes) *versus* ‘qu-’ (dans les relatives libres et les interrogatives indirectes avec référent inanimé), pour la période Moyen Français.

Comme on peut le voir de façon très nette sur les Figures 3.1 et 3.2, ‘ce qu-’ est largement préféré dans les relatives et ‘qu-’ dans les interrogatives indirectes. Le changement est minime entre l’Ancien Français et le Moyen Français, mais notons quand même que, en Ancien Français, l’emploi de ‘qu-’ dans les relatives était plus fréquent. Dès le Moyen Français, ‘ce qu-’ s’impose dans les relatives. Par contre, dans les interrogatives indirectes, le pourcentage de ‘ce qu-’ est encore très très bas. Donc cela confirme que ‘ce qu-’ s’impose tôt dans les relatives, et ce, avant de s’imposer dans les interrogatives indirectes. Quelques exemples sont donnés en (7) et (8) pour les relatives et en (9) et (10) pour les interrogatives indirectes.

Relatives

- (7) et leur **dit ce que** je li avoie dit [Jean de JOINVILLE (1307). *Mémoires ou Vie de Saint Louis*, p.323]
- (8) car il **fera por lui ce que** il porra [ANONYME (1223). *Queste del Saint Graal*. Anonyme, p.43]

Interrogatives indirectes

- (9) Il me **demanda que** il m'en sembloit [Chrétien de TROYES (1179). *Yvain ou Le Chevalier au Lion*, p.79]
- (10) et je ne **sai que** m'avenra [ANONYME (1223). *Queste del Saint Graal*. Anonyme, p.43]

3.3.1.4 Discussion

Cette première étude de corpus nous a permis de confirmer que 'ce qu-' apparaît dès l'Ancien Français dans les relatives et qu'il reste très rare dans les interrogatives indirectes. Mais l'inconvénient de ce corpus pour notre étude est qu'il contient très peu de données après la période 'Moyen-Âge' : 1 seul texte annoté morphosyntaxiquement pour la période 'Renaissance', et 3 textes annotés pour la période Français Classique. Nous ne pouvons donc pas vérifier la situation après le Moyen-Âge, surtout dans le cas des interrogatives indirectes où 'ce qu-' est très peu utilisé. C'est pour cela que j'ai mené une deuxième étude à partir d'un autre corpus, Frantext, sur une période plus longue. Cette étude est plus spécifique étant donné que Frantext n'est pas étiqueté : je n'ai pas pu relever toutes les relatives et toutes les interrogatives indirectes.

3.3.2 Une étude plus spécifique à partir d'un corpus non étiqueté

3.3.2.1 Le corpus

Le corpus que j'ai constitué pour la présente étude provient de Frantext, qui est une base de données de textes français, rassemblant des textes littéraires et philosophiques, mais aussi scientifiques et techniques. Frantext a été développé par les chercheurs du laboratoire ATILF-CNRS (Analyse et Traitement Automatique de la Langue Française). Dans le cadre de cette étude, j'ai fait mes recherches période par période, en commençant par l'Ancien Français et en allant jusqu'au 17^{ème} siècle (inclus). J'ai suivi le découpage en périodes de Frantext³. Pour chaque période, j'ai effectué mes recherches dans tous les textes de la période, sans sélectionner des textes d'un genre particulier, car cela ne semblait pas pertinent pour la présente étude. La description du corpus est donnée dans le Tableau 3.3.

3. Découpage de l'ancienne version de Frantext. En effet, une nouvelle version a vu le jour au cours de l'été 2018 ce qui a entraîné une réorganisation de la base de données et donc des changements en ce qui concerne le classement des textes.

Période	Dates	Nombre de textes
<i>Ancien Français</i>	<1330	80
<i>Moyen Français</i>	1330-1500	276
<i>Français de la Renaissance</i>	1501-1600	181
<i>Français classique (1)</i>	1601-1650	232
<i>Français classique (2)</i>	1650-1700	375

TABLE 3.3 – Présentation du corpus : dates et nombre de textes pour chaque période de la langue française.

3.3.2.2 Ce que j’ai cherché

L’ancienne version de Frantext (avant 2018) n’étant pas entièrement étiquetée et ne me permettant pas de chercher toutes les relatives introduites par ‘ce qu-’ ou ‘qu-’ d’une part et toutes les interrogatives indirectes introduites par ‘ce qu-’ ou ‘qu-’ d’autre part, j’ai choisi de relever les subordonnées introduites par ‘ce qu-’ ou ‘qu-’ sélectionnées par trois verbes en particulier : ‘demander’, ‘savoir’ et ‘faire’. Les deux premiers sélectionnent des interrogatives indirectes, mais de types différents (questions extensionnelles *versus* questions intensionnelles, distinction discutée dans la Section 2.4 du Chapitre 2); le troisième sélectionne des relatives.

Rappelons que les relatives peuvent être en position de sujet ou bien en position d’objet au sein de la proposition matrice.

(11) **Ce que tu m’as dit l’autre jour** m’a surprise.

(12) Je n’ai pas compris **ce que tu m’as dit l’autre jour**.

Comme illustré dans les exemples ci-dessus, la relative, indiquée en gras, est sujet du verbe matrice ‘surprendre’ en (11) alors qu’en (12), elle est l’objet direct du verbe matrice ‘comprendre’. Je n’ai relevé que celles en position d’objet pour la présente étude, étant donné que je me centre sur l’ambiguïté entre relatives et interrogatives indirectes, et que ces dernières sont nécessairement enchâssées sous un verbe.

Notons également que j’ai utilisé des expressions régulières afin d’obtenir toutes les formes de chaque verbe. Le deuxième élément de la séquence de recherche correspondait au démonstratif ‘ce’ ou au ‘que/qui’ interrogatif. Le troisième correspondait au complémenteur ‘que/qui’. J’ai également demandé à ce que chaque séquence soit adjacente l’une à l’autre, pour exclure les exemples de type (13).

(13) J’ai essayé de **faire ce même gâteau** au chocolat **que** nous avons dégusté chez ta soeur.

Nous avons bien ici nos trois éléments, indiqués en gras, mais il s'agit d'une relative pleine, pas d'une relative à tête allégée. Exiger l'adjacence des éléments permet donc d'éviter ce genre de 'bruit' indésirable.

Les données obtenues ont tout de même nécessité un tri manuel. En effet, en cherchant 'faire que', ou 'savoir que', on obtient de façon inévitable des complétives (sans *gap*) de type (14) et (15).

(14) Paul sait que Marie va venir demain.

(15) Le roi fait que la reine soit protégée. [sens : fait en sorte que]

J'ai également dû exclure les structures de type : clitique NEG + faire + que + verbe INF, comme en (16).

(16) Cela ne fait que l'importuner davantage.

Parfois, 'fait' est non pas une forme verbale du verbe 'faire', mais le nom 'fait'.

(17) C'est un fait que je ne nie pas.

Ces structures, qui sont en fait des relatives pleines, ne nous intéressent pas ici.

J'ai donc trié les données à la main et relevé les occurrences pertinentes pour la présente étude.

3.3.2.3 Les résultats

3.3.2.3.1 Ancien français Les résultats pour la période Ancien Français, c'est-à-dire avant 1330, sont présentés dans le Tableau 3.4, eux-mêmes représentés dans la Figure 3.3.

	CE QU-	QU-
<i>DEMANDER</i>	6 (11,5%)	46 (88,5%)
<i>SAVOIR</i>	23 (11%)	187 (89%)
<i>FAIRE</i>	82 (61,7%)	51 (38,3%)

TABLE 3.4 – Nombre d'occurrences et pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) *versus* 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la période Ancien Français (<1330).

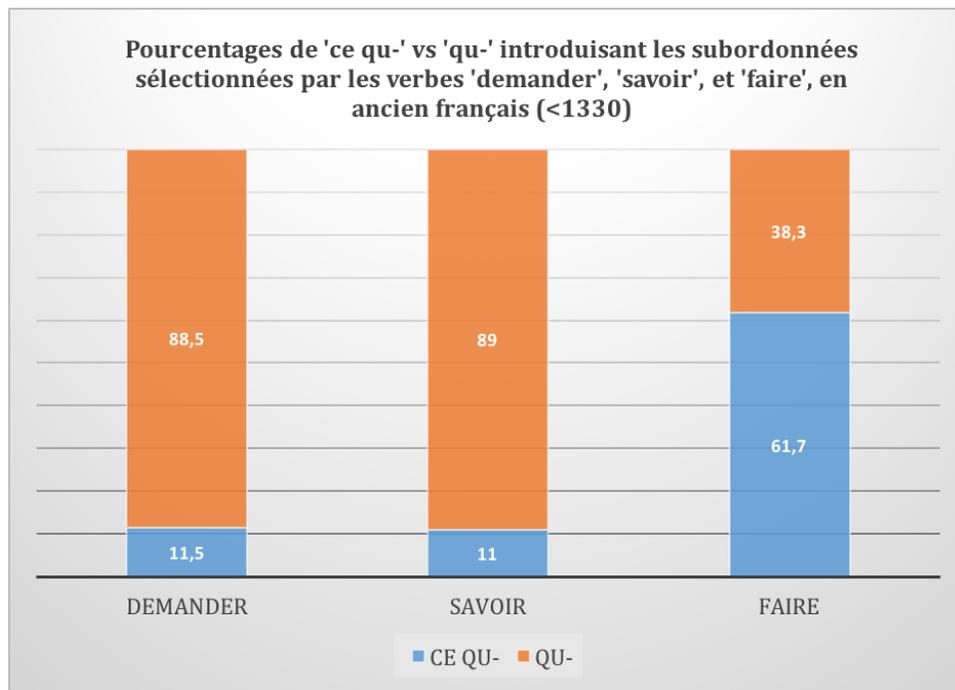


FIGURE 3.3 – Pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) *versus* 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la période Ancien Français (<1330).

On peut observer qu'en Ancien Français, la grande majorité des interrogatives indirectes (environ 90%) sont introduites par 'que' seul ou 'qui' seul (selon que le *gap* est en position d'objet ou de sujet au sein de la subordonnée). Cela est vrai pour les questions extensionnelles et les questions intensionnelles, puisqu'aucune différence n'est visible entre les interrogatives indirectes sélectionnées par le verbe 'demander' et celles sélectionnées par le verbe 'savoir'.⁴

- (18) *En Ancien Français* : et li **demandèrent que** il a veu [ANONYME (1223). *Queste del Saint Graal*. Anonyme, p.258]
Equivalent en Français Moderne supposé : et ils lui demandèrent ce qu'il avait vu
- (19) *En Ancien Français* : Li rois **demande que** il ont. [ANONYME (1155). *Eneas*. T. 1, p.116]
Equivalent en Français Moderne supposé : Le roi demande ce qu'ils ont.
- (20) *En Ancien Français* : Ne **sai que** je die [Gautier de COINCI (1223b). *Miracles de Notre-Dame*. T. 4, p.287-288]
Equivalent en Français Moderne supposé : Je ne sais pas ce que je dis.
- (21) *En Ancien Français* : Et **savez que** je vos promet? [Wauchier de DENAIN (1210). *Deuxième continuation de Perceval*, p.59]
Equivalent en Français Moderne supposé : Et savez-vous ce que je vous promets?

Même si l'emploi de 'ce qu-' est marginal à l'époque, on trouve quand même quelques occurrences dès le 11ème siècle.

- (22) *En Ancien Français* : Et lors **demanda** messire Rogier **ce que** il estoit avenu de la gent et de la compaignie de monseignor Jehan de Tournay. [ANONYME (1322). *Chroniques de Morée*, p.305-306]
Equivalent en Français Moderne supposé : Et à ce moment-là, messire Rogier demanda ce qu'il était advenu de ceux qui accompagnent monseigneur Jehan de Tournay.
- (23) *En Ancien Français* : Car il **sevent ce qu'** il donna et quanque a feire commanda; [Robert de BORON (1199). *Le Roman de l'Estoire dou Graal*, p.76-78]
Equivalent en Français Moderne supposé : Car ils savent ce qu'il donna et tout ce qu'il ordonna de faire;

En ce qui concerne les relatives, on note une tendance inverse : 'ce qu-' est préféré à 'qu-', mais l'écart n'est pas aussi important que dans les interrogatives indirectes. En effet, environ 60% des relatives sélectionnées par le verbe 'faire' sont introduites par 'ce qu-'. Dans 40% des cas, on a donc des relatives libres introduites par 'qu-' seul.

4. Notons quand même que les pourcentages concernant le verbe 'demander' pourraient être surévalués. En effet, 'demander' peut être aussi ambigu, surtout à la deuxième personne du singulier, et avec des modaux dans la subordonnée.

(1) Demande-moi ce que tu veux.

- (24) *En Ancien Français* : por ce, ne poent faire ce que il volent [ANONYME (1250). *Del Confortement de Philosophie*, p.63]
Equivalent en Français Moderne supposé : pour cela, ils ne peuvent pas faire ce qu'ils veulent
- (25) *En Ancien Français* : Sire, vostre prisoniers sui por fere ce que vos voldroiz [Chrétien de TROYES (1181b). *Perceval*, p.376a]
Equivalent en Français Moderne supposé : Sire, je suis votre prisonnier pour faire ce que vous voudrez
- (26) *En Ancien Français* : Faisons que devons faire [Alexandre de PARIS (1180). *Roman d'Alexandre*, branche 3, p.183]
Equivalent en Français Moderne supposé : Faisons ce que nous devons faire
- (27) *En Ancien Français* : Ne pués el faire que t'as fait ? [Gautier d'ARRAS (1180). *Eracle*, p.184]
Equivalent en Français Moderne supposé : Ne peut-il pas faire ce que tu as fait ?

Ainsi, on peut dire que le système des relatives est similaire à celui des interrogatives indirectes en Ancien Français, même si on tend vers une différenciation des deux systèmes. Cette ambiguïté possible entre les relatives et les interrogatives indirectes rappelle l'ambiguïté typique que nous avons avec 'what' en anglais (discutée dans le Chapitre 2, Section 2.2).

Un point important est qu'à la fois 'ce' et 'que' semblent être des éléments forts à l'époque, alors qu'ils sont faibles aujourd'hui. Observons les données ci-dessous.

- (28) *En Ancien Français* : **Que** vous en sanble ? [Adam de LA HALLE (1285). *Le jeu de Robin et Marion*, p.32]
Equivalent en Français Moderne supposé : Qu'en pensez-vous ?
- (29) *En Ancien Français* : Tetraptota **que** est ? [ANONYME (1325). *Donat (M2)*, p.116]
Equivalent en Français Moderne supposé : Qu'est-ce qu'un Tetraptota ?

On peut en effet remarquer que les emplois du 'que' interrogatif étaient plus étendus en Ancien Français qu'en Français moderne. 'Que' semble être davantage autonome à l'époque.

Quant à 'ce', il peut lui aussi être employé seul, sans se cliticiser sur le verbe 'être'. Il peut revêtir le sens de 'cela', démonstratif fort du français moderne.

- (30) *En Ancien Français* : Qui ce refuse il n'est pas sages. [Chrétien de TROYES (1176). *Cligès*, p.54e]
Equivalent en Français Moderne supposé : Qui refuse cela n'est pas sage.
- (31) *En Ancien Français* : Ce dit chascuns : [...] [Chrétien de TROYES (1176). *Cligès*, p.59a]
Equivalent en Français Moderne supposé : Chacun dit cela : [...]

- (32) *En Ancien Français* : ce vos voel conter [...] [Chrétien de TROYES (1177). *Lancelot ou Le Chevalier à la Charrette*, p.38d]
Equivalent en Français Moderne supposé : Je veux vous raconter cela [...]

Un autre fait que je souhaite souligner est la possibilité d'employer 'ce que' en position oblique suivi d'un verbe à l'infinitif, ce qui conduirait à un résultat agrammatical en français moderne.

- (33) *En Ancien Français* : mais ele n'avoit de ce que faire [Guillaume de LORRIS (1230). *Le Roman de la Rose*, p.50-54]

En français moderne, en position oblique, seul 'quoi' peut être utilisé dans ce cas.

- (34) a. J'ai de quoi faire.
 b. *J'ai de ce que faire.

De plus, certaines données montrent que l'incorporation (présentée dans le Chapitre 2, Section 2.3.5) n'a certainement pas eu lieu en Ancien Français - si elle a eu lieu. Nous avons déjà évoqué la possibilité d'insérer du matériel lexical (comme des adverbes) entre 'ce' et 'que' (cf. exemple (99) dans le Chapitre 2, répété ici en (35) par souci de commodité).

- (35) Et pensai a ce **longuement** qu'il m'iert avenu. [Froiss. Esp. 625-626, cité par Jokinen 1986 : 172]

Vient s'ajouter à cela un fait qui est très mystérieux : pendant une courte période (fin du 12ème siècle)⁵, 'ce que' pouvait introduire une interrogative directe.

- (36) *En Ancien Français* : Dame, fait il, ce que puet estre ? [ANONYME (1155). *Eneas*. T. 1, p.82-84]
Equivalent en Français Moderne supposé : Dame, fait-il, qu'est-ce que cela peut être ?
- (37) *En Ancien Français* : Ce que est et de coi vos sert ? [Chrétien de TROYES (1181a). *Conte du Graal*, p. 381v]
Equivalent en Français Moderne supposé : Qu'est-ce que c'est et à quoi cela vous sert ?

Cela est exactement à l'opposé de ce à quoi on s'attendrait s'il y avait incorporation de l'élément 'ce'. En effet, si 'ce que' est à l'origine une séquence D+C qui a subi un processus de grammaticalisation se manifestant par l'incorporation de 'ce' à 'que' (/səkə/), alors l'emploi de səkə dans une interrogative directe devrait être le stade ultime de cette grammaticalisation, et non pas le tout premier stade. Nous y reviendrons.

5. J'ai cherché, et effectivement, on ne trouve plus de questions directes introduites par 'ce que' en Moyen Français.

Notons que ‘qu’est-ce qu-’ pouvait également introduire des questions directes en Ancien Français, mais il n’était pas analysable comme un tout, contrairement à aujourd’hui, où /kɛsk/ est clairement grammaticalisé et perçu comme un bloc indécomposable. Cela est sûrement lié au fait que ‘ce’ et ‘que’ étaient des éléments forts qui se sont affaiblis.

- (38) *En Ancien Français* : que est ce que vos dites et que est ce que vos me demandez? [ANONYME (1230). *La mort le roi Artu*, p.61-62]
Equivalent en Français Moderne supposé : Qu’est-ce que vous me dites et qu’est-ce que vous me demandez?

On pouvait d’ailleurs insérer du matériel, comme l’apostrophe au maître en (39), entre ‘qu’est-ce’ et ‘que’.

- (39) *En Ancien Français* : qu’est ce, biaux maistres, que vos dites? [Gautier de COINCI (1223a). *Miracles de Notre-Dame*. T. 2, p.138-140]
Equivalent en Français Moderne supposé : Qu’est-ce que vous me dites, beau maître?

3.3.2.3.2 Moyen français Les résultats pour la période Moyen Français, c’est-à-dire entre 1330 et 1500, sont présentés dans le Tableau 3.5, eux-mêmes représentés dans la Figure 3.4.

	CE QU-	QU-
DEMANDER	14 (17,1%)	68 (82,9%)
SAVOIR	38 (8,6%)	404 (91,4%)
FAIRE	388 (95,6%)	18 (4,4%)

TABLE 3.5 – Nombre d’occurrences et pourcentages de ‘ce qu-’ (dans les relatives à tête allégée et les interrogatives indirectes) *versus* ‘qu-’ (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la période Moyen Français (1330-1500).

La situation pour la période du Moyen Français n’évolue pas beaucoup en ce qui concerne les interrogatives indirectes : ‘ce’ reste sporadique. La règle est d’employer ‘que’ seul. Cela est vrai pour les interrogatives intensionnelles...

- (40) *En Moyen Français* : Il luy fut **demandé que** il y venoit faire. [Philippe de COMMYNES (1490). *Mémoires*. T. 2, p.43]
Equivalent en Français Moderne supposé : Il lui fut demandé ce qu’il y venait faire.
- (41) *En Moyen Français* : [...] **demendans que** sa femme fait. [Eustache DESCHAMPS (1394). *Le Miroir de mariage*, p.121]
Equivalent en Français Moderne supposé : [...] demandant ce que fait sa femme.

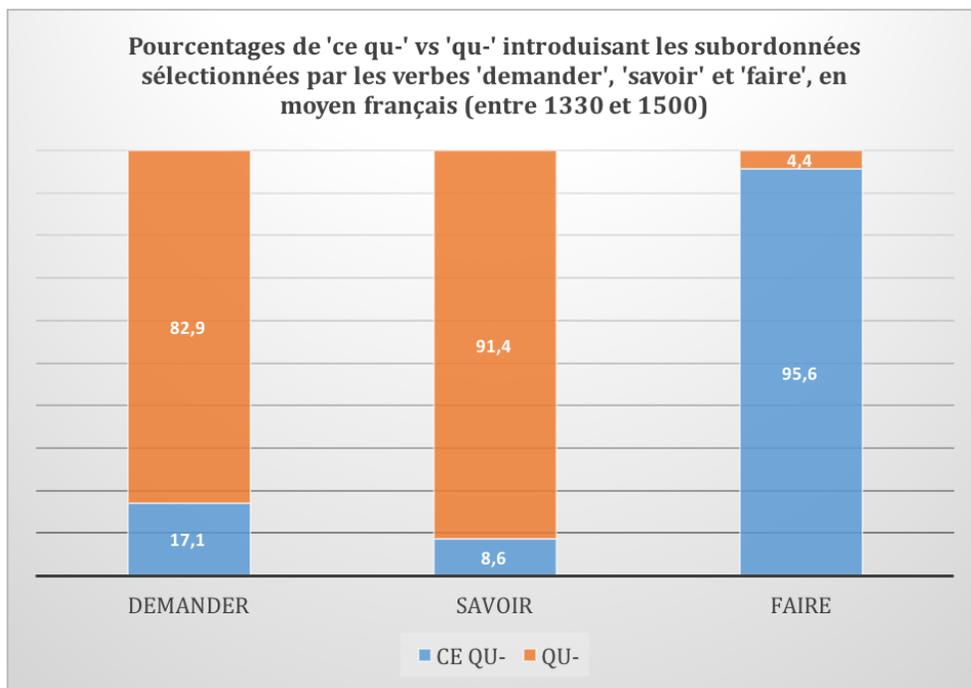


FIGURE 3.4 – Pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) *versus* 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la période Moyen Français (1330-1500).

- (42) *En Moyen Français* : Le roy dist : ‘**Demande ce que** tu voudras, car j’en suis bien content.’⁶ [Jean BAGNYON (1468). *L’Histoire de Charlemagne*, p.6]
Equivalent en Français Moderne supposé : Le roi dit : ‘Demande ce que tu voudras, car j’en suis bien content.’

... comme pour les interrogatives extensionnelles. Voilà quelques exemples avec ‘savoir’.

- (43) *En Moyen Français* : on ne **savoit que** il estoient devenu [Jean FROISSART (1400). *Chroniques*, p.337]
Equivalent en Français Moderne supposé : on ne savait ce qu’ils étaient devenus
- (44) *En Moyen Français* : je ne **scay que** mes amys diront [Jean RÉGNIER (1432). *Les Fortunes et adversitez*, p.123]
Equivalent en Français Moderne supposé : je ne sais pas ce que diront mes amis
- (45) *En Moyen Français* : Et se j’**eusse sceu ce que** je sçay, vous n’y fussiez pas venu a temps! [Raoul LEFEVRE (1460). *L’Histoire de Jason*, p.211]
Equivalent en Français Moderne supposé : Et si j’avais su ce que je sais, vous ne serez pas venu à temps!

Le changement majeur concerne les relatives : alors qu’en Ancien Français, les relatives libres coexistaient avec les relatives à tête allégée (même si ces dernières avaient déjà un avantage sur les premières), ces dernières s’imposent jusqu’à faire quasiment disparaître les relatives libres introduites par ‘qu-’ (18 occurrences, soit 4% seulement).

- (46) *En Moyen Français* : **Fais ce que** tu dois et adviengne ce qu’il pourra. [Jean de BUEIL (1461). *Le Jouvenciel*. T. 2, p.205]
Equivalent en Français Moderne supposé : Fais ce que tu dois faire, et advienne que pourra.
- (47) *En Moyen Français* : je estoie jeune fille et **fasioe ce que** mes pere et mere me disoient [ANONYME (1390). *Les Quinze joies de mariage*, p.38]
Equivalent en Français Moderne supposé : j’étais jeune fille et je faisais ce que mes père et mère me disaient de faire

L’exemple en (46) est intéressant car il comporte un usage de ‘ce qu-’ que l’on n’emploie plus aujourd’hui : ‘advienne ce qu’il pourra’. L’expression qui s’est maintenue est justement une relative libre ‘advienne que pourra’, une des seules relatives libres un peu figées encore introduites par ‘que’ en français moderne, avec ‘coûte que coûte’ et ‘vaille que vaille’. Ce sont les trois seules expressions, qui, à ma connaissance, sont un résidu du ‘que’ interrogatif pouvant introduire des relatives libres, à la manière de ‘what’ en anglais.

6. Notons que cet exemple est ambigu et pourrait très bien être une relative. C’est la cas pour d’autres exemples relevés, où le modal ‘vouloir’ est employé à l’intérieur de la subordonnée. Cela renforce encore plus la rareté de l’usage de ‘ce que’ dans les interrogatives indirectes. En effet, il y a peu d’exemples avec ‘demander ce que’, et parmi ces quelques exemples, plusieurs sont ambigus avec une structure relative.

Par ailleurs, ‘ce qui’ semble apparaître après ‘ce que’. Alors qu’il était quasiment absent dans les textes du Moyen Français, on trouve plusieurs occurrences de subordinées introduites par ‘ce qui’ dans les textes du Moyen Français, que ce soit des interrogatives indirectes ou bien des relatives. Quelques exemples ci-dessous.

- (48) *En Moyen Français* : Et **demander ce qui** est juste [Guillaume de MACHAUT (1341). *Remède de Fortune*, p.139]
- (49) *En Moyen Français* : car je ne **sçavois ce qui** pourroit survenir à mon maistre [Philippe de COMMYNES (1497). *Mémoires*. T. 3, p.142]
Equivalent en Français Moderne supposé : car je ne savais ce qui pourrait advenir à mon maître
- (50) *En Moyen Français* : nous **ferons ce qui** vous plaira [ANONYME (1365). *Miracle de Nostre Dame d’Amis et d’Amille*, p.62]

Notons que les dépendances sujet sont tout de même moins fréquentes que les dépendances objet. 6 des 14 interrogatives indirectes introduites par ‘ce qu-’ et sélectionnées par le verbe ‘demander’ en Moyen Français sont des dépendances sujet introduites par ‘ce qui’ ; 14 sur 38 pour celles sélectionnées par le verbe ‘savoir’. 53 des 388 relatives sélectionnées par ‘faire’ sont introduites par ‘ce qui’. On trouve d’ailleurs plusieurs exemples de relatives introduites par ‘ce que’ (employé comme par défaut, à la place de ‘ce qui’) dans le cas où le *gap* est en position de sujet.

- (51) *En Moyen Français* : et se il fait **ce que** mieulx est [...] [Pierre CRAPILLET (1450). *Cur deus homo*, p.248]
Equivalent en Français Moderne supposé : et s’il fait ce qui est le mieux
- (52) *En Moyen Français* : [...] pour continuer et faire **ce que** dit est [Clément de FAUQUEMBERGUE (1433). *Journal*. T. 3, p.2]
Equivalent en Français Moderne supposé : pour continuer et faire ce qui est dit

On peut supposer qu’en Ancien Français, seul ‘ce que’ était utilisé, que le *gap* soit en position d’objet ou de sujet, et qu’en Moyen Français, ‘ce qui’ apparaît, d’où l’observation d’un certain nombre d’occurrences de subordinées introduites par ‘ce qui’. Il reste encore en Moyen Français quelques dépendances sujet introduites par ‘ce que’ (tout changement prend un certain temps avant de se mettre en place).

Dans les interrogatives indirectes, où ‘ce’ est encore très marginal, on aura ‘qui’ seul lorsque le *gap* est en position de sujet.

- (53) *En Moyen Français* : Le roy d’Angleterre fut fort esbahy de ceste venue si soubdaine et luy **demanda qui** l’amenoit ; [Philippe de COMMYNES (1490). *Mémoires*. T. 2, p.53]
Equivalent en Français Moderne supposé : Le roi d’Angleterre fut très surpris de cette venue si soudaine et lui demanda ce qui l’amenait ;

- (54) *En Moyen Français* : Je ne sçay **qui** m'en avendra. [ANONYME (1346). *Miracle de un pape qui vendi le basme*, p.379]
Equivalent en Français Moderne supposé : Je ne sais pas ce qui m'arrivera.

Une ambiguïté pouvait alors survenir étant donné que les relatives libres sont également introduites par 'qui' lorsque le référent est [+ humain].

- (55) je ne sais qui te rend malheureux

Un exemple tel que (55) pouvait alors avoir deux significations à l'époque : soit 'je ne sais pas quelle personne te rend malheureux', soit 'je ne sais pas quelle chose te rend malheureux'.

3.3.2.3.3 Français du 16ème siècle Les résultats pour la période de la Renaissance, c'est-à-dire le français du 16ème siècle, sont présentés dans le Tableau 3.6, eux-mêmes représentés dans la Figure 3.5.

	CE QU-	QU-
<i>DEMANDER</i>	12 (24%)	38 (76%)
<i>SAVOIR</i>	74 (31,5%)	161 (68,5%)
<i>FAIRE</i>	156 (98,1%)	3 (1,9%)

TABLE 3.6 – Nombre d'occurrences et pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) *versus* 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la période de la Renaissance (1501-1600).

Au 16ème siècle, 'ce qu-' continue de progresser dans les interrogatives indirectes, mais reste quand même assez marginal. On peut remarquer que l'écart entre 'demander' et 'savoir' commence à se creuser. Quelques exemples ci-dessous.

- (56) et luy **demande ce que** nous voulons obtenir de luy [Guy de BRUÈS (1557). *Les Dialogues de Guy de Bruès contre les nouveaux académiciens*, p.221]
- (57) pour luy **demande ce qui** nous est nécessaire [Guy de BRUÈS (1557). *Les Dialogues de Guy de Bruès contre les nouveaux académiciens*, p.127]
- (58) Tu **demandes que** c'est que l'enfer? [Pierre-Victor PALMA-CAYET (1598). *L'Histoire prodigieuse du Docteur Fauste*, p.86]
- (59) ... l
edit Eduart, qui luy **demanda qui** le mouvoit de venir en Angleterre. [Jean LE CLERC (1502). *Interpolations et variantes de la Chronique scandaleuse*, p.277]

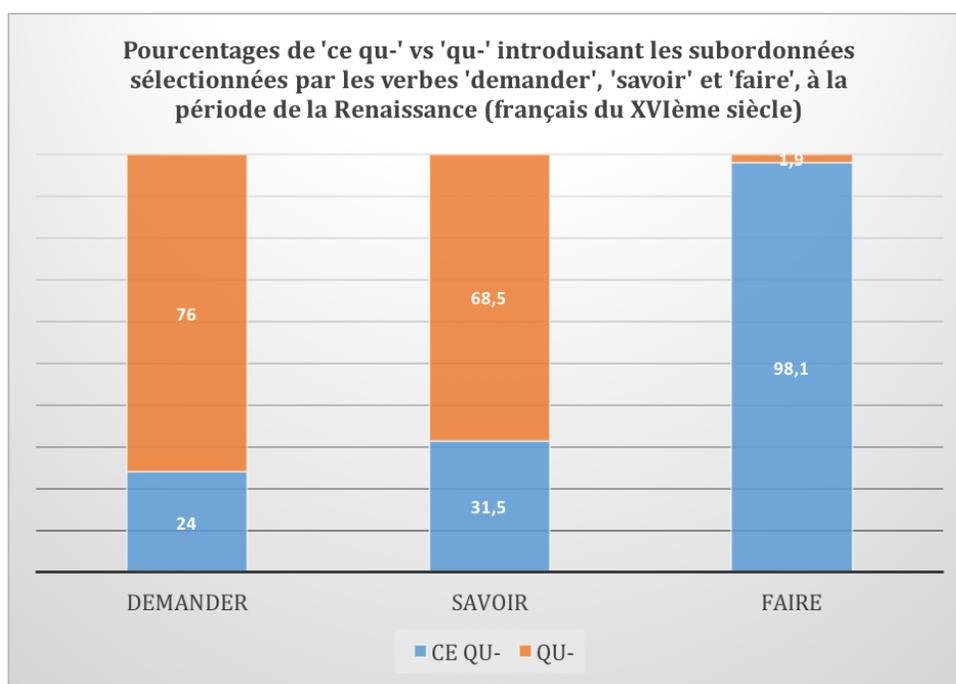


FIGURE 3.5 – Pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) *versus* 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la période de la Renaissance (1501-1600).

- (60) je ne sçay **ce que** j'ay voulu dire [Michel de MONTAIGNE (1592). *Essais*. T. 1, livres 1 et 2, p.566]
- (61) et est tousjours desireux de sçavoir **ce qui** luy doit advenir [Pierre BELON (1555). *L'Histoire de la nature des oyseaux*, p.68]
- (62) *En Français du 16ème* : Si c'est un enfant qui juge, il ne sçait **que** c'est ; [Michel de MONTAIGNE (1592). *Essais*. T. 1, livres 1 et 2, p.504]
En Français Moderne : Si c'est un enfant qui juge, il ne sait ce que c'est ;
- (63) *En Français du 16ème* : [...] dont elle fut en si grande admiration qu'elle ne sçavoit **qui** luy estoit advenu [Jean LE CLERC (1502). *Interpolations et variantes de la Chronique scandaleuse*, p.277]
En Français Moderne : [...] dont elle fut en si grande admiration qu'elle ne savait ce qui lui était advenu

Quant aux relatives, la stratégie de relative libre (avec référent inanimé), comme en (66), a quasiment disparu (plus que 2% environ).

- (64) je feray **ce que** tu dis [ANONYME (1521). *Le Violier des histoires romaines moralisées*, p.94]
- (65) Car ordinairement il fait **ce qui** luy plaist [André de RIVAUDEAU (1566). *Aman : tragédie sainte*, p.102-103]
- (66) *En Français du 16ème* : et sagement fait **que** la plus part des gens fait [Philibert de VIENNE (1548). *Le Philosophe de court*, p.78]
En Français Moderne : et sagement fait ce que la plupart des gens font

3.3.2.3.4 Français classique : 1ère moitié du 17ème siècle Les résultats pour la première moitié du 17ème siècle, c'est-à-dire le français classique, sont présentés dans le Tableau 3.7, eux-mêmes représentés dans la Figure 3.6.

	CE QU-	QU-
DEMANDER	35 (44,3%)	44 (55,7%)
SAVOIR	305 (75,5%)	99 (24,5%)
FAIRE	234 (99,6%)	1 (0,4%)

TABLE 3.7 – Nombre d'occurrences et pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) versus 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la première moitié du 17ème siècle (1601-1650).

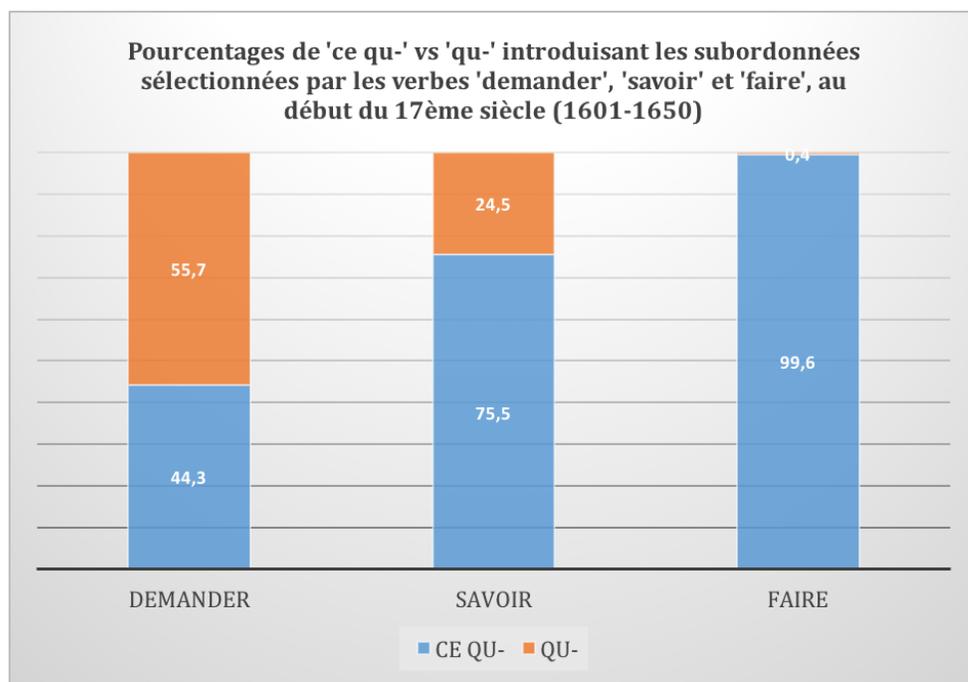


FIGURE 3.6 – Pourcentages de ‘ce qu-’ (dans les relatives à tête allégée et les interrogatives indirectes) *versus* ‘qu-’ (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la première moitié du 17^{ème} siècle (1601-1650).

Au début du 17^{ème} siècle, on observe une augmentation considérable de l'emploi de 'ce qu-' dans les interrogatives indirectes : +20% dans le cas de celles sélectionnées par le verbe 'demander', et +44% dans le cas de celles sélectionnées par le verbe 'savoir'. Un fossé se creuse donc entre les questions intensionnelles et les questions extensionnelles. Cela n'est pas surprenant. En effet, étant donné qu'un verbe comme 'savoir' se focalise sur l'extension du complément, c'est-à-dire sur la réponse à la question, et pas sur la question elle-même, on s'attendait à ce que l'utilisation du démonstratif 'ce' s'impose dans les interrogatives indirectes extensionnelles avant de s'imposer dans les questions intensionnelles. C'est bien ce que l'on constate au début du 17^{ème} siècle. Voilà quelques exemples d'interrogatives indirectes de l'époque en question.

Questions intensionnelles

- (67) je luy **demanday ce que** le Roy luy avoit dict en partant [Jean HEROARD (1608). *Histoire particulière de Louis XIII*, p.553]
- (68) Il me faut **demander ce qui** m' est necessaire, quoyque mon naturel repugne à recevoir. [Claude MALLEVILLE (1649). *Oeuvres poétiques*, p.64]
- (69) *En Français Classique* : Me de Montglat qui l'avoit oui crier a table me **demanda que** c'est qu'il avoit. [Jean HEROARD (1608). *Histoire particulière de Louis XIII*, p.336]
En Français Moderne : Mme de Monglat qui l'avait entendu crier à table me demanda ce que c'est qu'il avait.
- (70) *En Français Classique* : je ne peux m' empescher de **demander que** c' estoit [Honoré D'URFÉ (1610). *L'Astrée*. T. 2, livres 1 à 12, p.223]
En Français Moderne : je ne pus m'empêcher de demander ce que c'était

Questions extensionnelles

- (71) si tu **sçavois ce que** c'est que d'aimer [Tristan L'HERMITE (1639). *Panthée*, p.158]
- (72) je **sais ce qui** l'amène [Pierre CORNEILLE (1639). *L'Illusion comique*, p.620]
- (73) *En Français Classique* : si cela ne vous contente, je ne **sçay que** vous pouvez desirer de moy [Honoré D'URFÉ (1631). *L'Astrée*. T. 3, livres 1 à 12, p.258]
En Français Moderne : si cela ne vous satisfait pas, je ne sais ce que vous pouvez espérer de moi
- (74) *En Français Classique* : il ne **sçait qui** le fasche, ou qui luy fait plaisir, et luy-mesme en son coeur n'entend point son desir [Théophile de VIAU (1621). *Oeuvres poétiques*, 1ère partie, p.73-74]
En Français Moderne : il ne sait pas ce qui le met en colère, ou ce qui lui fait plaisir, et lui-même ne comprend pas ce que son cœur désire

‘Ce qu-’ s’est clairement imposé dans les relatives.

- (75) il faut **faire ce que** le pere spirituel dira ; [Saint-François de SALES (1619). *Introduction à la vie dévote*, p.118]
- (76) Il est besoin de **faire ce qui** fait besoin. [Antoine de MONTCHRESTIEN (1615). *Traicté de l’oeconomie politique*, p.184]

3.3.2.3.5 **Français classique : 2ème moitié du 17ème siècle** Les résultats pour la deuxième moitié du 17ème siècle (français classique) sont présentés dans le Tableau 3.8, eux-mêmes représentés dans la Figure 3.7.

	CE QU-	QU-
<i>DEMANDER</i>	81 (100%)	0 (0%)
<i>SAVOIR</i>	559 (97,1%)	17 (2,9%)
<i>FAIRE</i>	300 (99%)	3 (1%)

TABLE 3.8 – Nombre d’occurrences et pourcentages de ‘ce qu-’ (dans les relatives à tête allégée et les interrogatives indirectes) *versus* ‘qu-’ (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la deuxième moitié du 17ème siècle (1651-1700).

On constate qu’à la fin du 17ème siècle, la situation est quasiment celle du français moderne. ‘Ce qu-’ s’est finalement imposé dans les interrogatives indirectes, qu’il s’agisse de questions extensionnelles ou intensionnelles. Ainsi, l’absence du démonstratif ‘ce’ est désormais considérée comme agrammaticale, que ce soit dans les relatives ou dans les interrogatives indirectes. Voilà quelques exemples de la fin du 17ème siècle.

- (77) Il n’y a qu’à **demander ce que** veut le roi. [Madame de SÉVIGNÉ (1675). *Correspondances*. T. 1, p.314]
- (78) Il est vrai que je **sçay ce qui** se passe [Roger de BUSSY-RABUTIN (1681). *Les Lettres de messire Roger de Rabutin, comte de Bussy*. T. 1, p.61]
- (79) J’ai **fait ce que** j’ai pu pour brouiller vos esprits ; [Pierre CORNEILLE (1682). *La Galerie du Palais*, p.109]

3.3.2.3.6 **Vue d’ensemble : du 11ème au 17ème siècle** La Figure 3.8 indique la proportion de ‘ce qu-’ en fonction de chaque verbe matrice (‘demander’, ‘savoir’, ‘faire’), sur toute la période étudiée, c’est-à-dire du 11ème siècle jusqu’au 17ème siècle. Cela permet d’avoir une vue d’ensemble sur l’évolution de ‘ce qu-’ au fil du temps. Ce graphique se base sur le calcul d’une moyenne mobile. Une moyenne mobile permet de ‘lisser’ en quelque sorte une

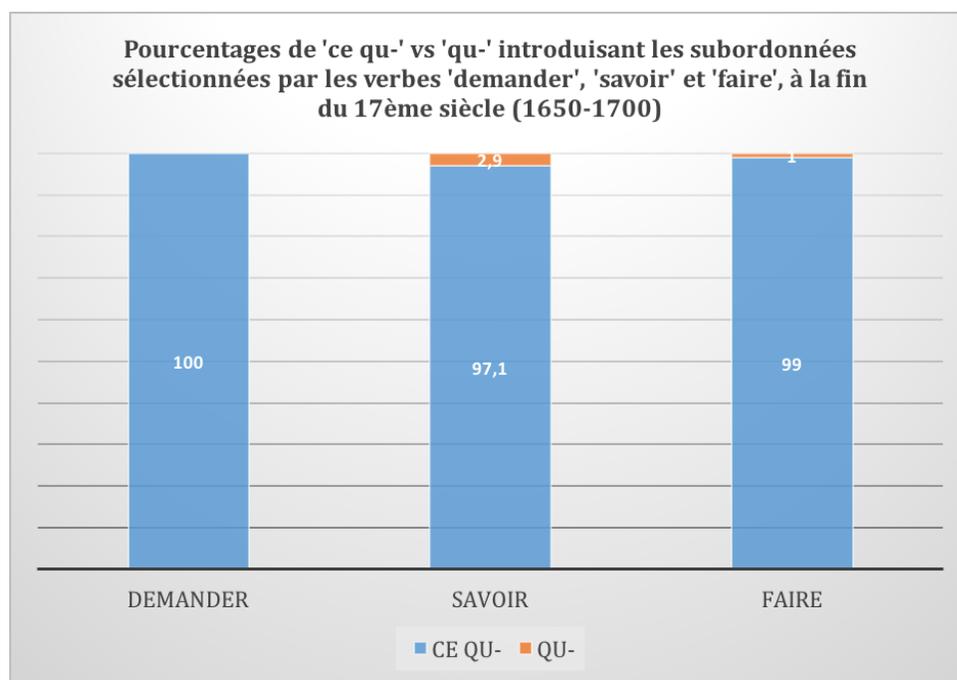


FIGURE 3.7 – Pourcentages de 'ce qu-' (dans les relatives à tête allégée et les interrogatives indirectes) *versus* 'qu-' (dans les relatives libres et les interrogatives indirectes avec référent inanimé) en fonction de chaque verbe, pour la deuxième moitié du 17ème siècle (1651-1700).

série de valeurs exprimées en fonction du temps. Elle permet d'éliminer les fluctuations les moins significatives, et de souligner les tendances sur le long terme. Ici il s'agit de la proportion moyenne d'emploi de 'ce qu-' (comparé à l'emploi de 'qu-' seul), et cette proportion moyenne est recalculée de façon continue. Comme l'indique la Figure 3.8, la moyenne mobile est de 101 ans, ce qui signifie que cette proportion moyenne est recalculée de façon continue, en prenant en compte à chaque fois les valeurs 50 années avant et 50 années après. Ce calcul est très utile pour la présente étude. En effet, comme il y a peu voire pas d'occurrences pour certaines années, c'est difficile de pouvoir visualiser l'évolution sous la forme d'une courbe linéaire sans passer par le calcul d'une moyenne mobile.

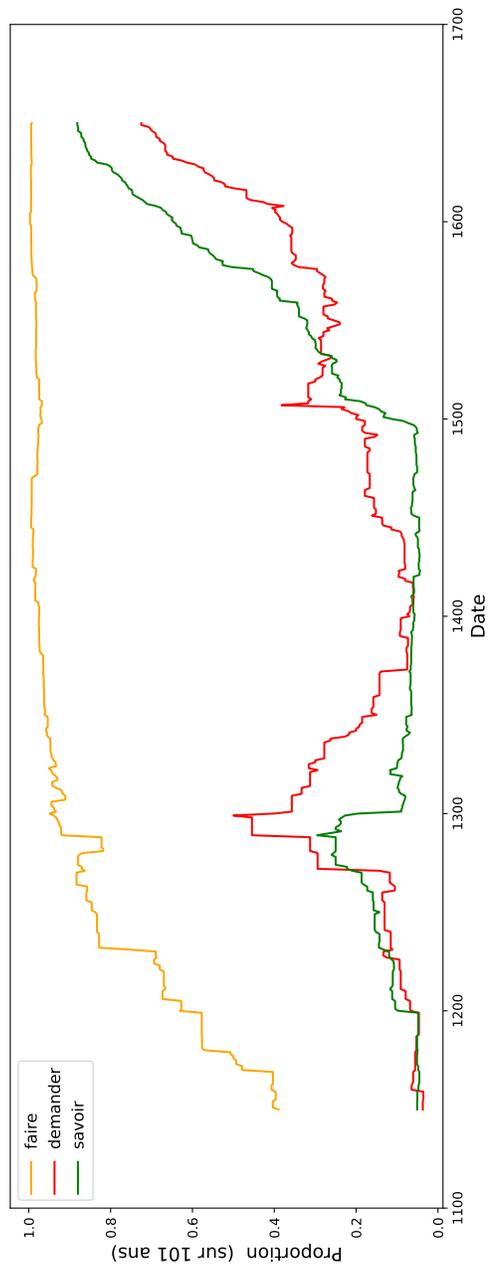


FIGURE 3.8 – Proportion de ‘ce qu-’ en fonction de chaque verbe matrice, du 11ème jusqu’à la fin du 17ème siècle (moyenne mobile de 101 ans)

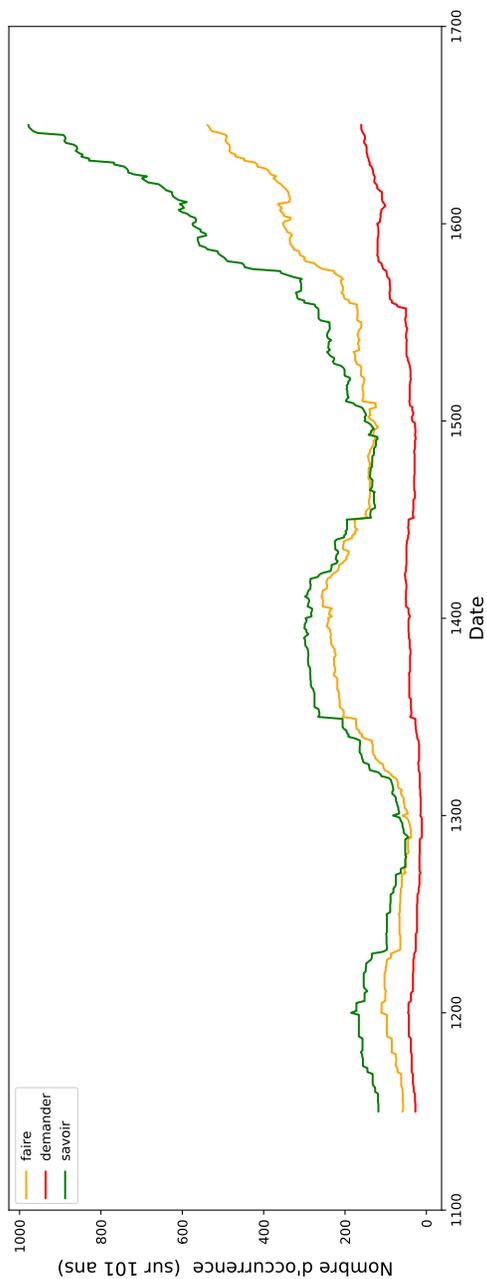


FIGURE 3.9 – Nombre d'occurrence de 'ce qu-' dans les textes en fonction de chaque verbe matrice, du 11ème jusqu'à la fin du 17ème siècle (somme mobile de 101 ans)

Comme la quantité de données est variable d'un verbe à l'autre, il est important, pour bien interpréter la Figure 3.8, d'observer en parallèle la Figure 3.9, qui indique le nombre d'occurrence de 'ce qu-' en fonction de chaque verbe matrice, au fil des siècles. Ici, ce n'est pas une moyenne mobile, mais une somme mobile, puisqu'on représente la quantité de données et non pas la proportion d'utilisation de 'ce qu-'. La fenêtre est la même (101 ans).

Comme le reflète très nettement la Figure 3.8, 'ce qu-' est davantage utilisé dans les relatives que dans les interrogatives indirectes. En effet, on peut constater un écart considérable entre 'faire' d'une part, et 'demander' et 'savoir' d'autre part. Pour le verbe 'faire', 'ce qu-' alterne avec 'qu-' en ancien français, mais il s'impose rapidement : on peut voir que dès 1300 'qu-' n'est quasiment plus utilisé sans être précédé du démonstratif 'ce'. Si l'on compare 'savoir' et 'demander', on constate que 'ce qu-' s'impose dans les interrogatives extensionnelles avant de s'imposer dans les interrogatives intensionnelles. On relève en effet une croissance constante de la proportion d'utilisation de 'ce qu-' à partir de la première moitié du 16ème siècle avec 'savoir', et seulement à partir de la première moitié du 17ème siècle avec 'demander'.

En ce qui concerne la quantité de données (voir Figure 3.9), globalement, on peut observer que nous avons moins de données en ancien français, et que nous avons davantage de données à partir de la deuxième moitié du 16ème siècle. Cela est valable pour les 3 verbes. Notons que vers 1300 nous avons relevé que très peu d'occurrences dans les textes, que ce soit avec le verbe 'faire', 'savoir' ou 'demander', et cela est reflété dans la Figure 3.8 par des pics de courbes un peu étranges vers 1300. De plus, il est important de noter que la quantité de données est moindre pour le verbe 'demander' que pour les verbes 'faire' et 'savoir', et ce sur toute la période (du 11ème au 17ème siècles). Il faut donc davantage relativiser la proportion de 'ce qu-' avec le verbe 'demander' étant donné le peu d'occurrences que nous avons.

Nous avons vu que 'ce qu-' a mis beaucoup de temps à s'imposer dans les interrogatives intensionnelles. Plus particulièrement, nous avons observé un changement brutal et soudain vers le milieu du 17ème siècle. Rappelons que nous passons de 44% de 'ce qu-' à 100% de 'ce qu-' avec le verbe 'demander' de la première moitié à la deuxième moitié du 17ème siècle. Afin de mieux percevoir ce changement, faisons un zoom sur la période 1600-1700, toujours via le calcul d'une moyenne mobile, avec une fenêtre moins large cette fois-ci (fenêtre de 21 ans, prenant en compte les données 10 années avant et 10 années après à chaque fois que la moyenne est recalculée). Les données pour cette période sont représentées sur la Figure 3.10.

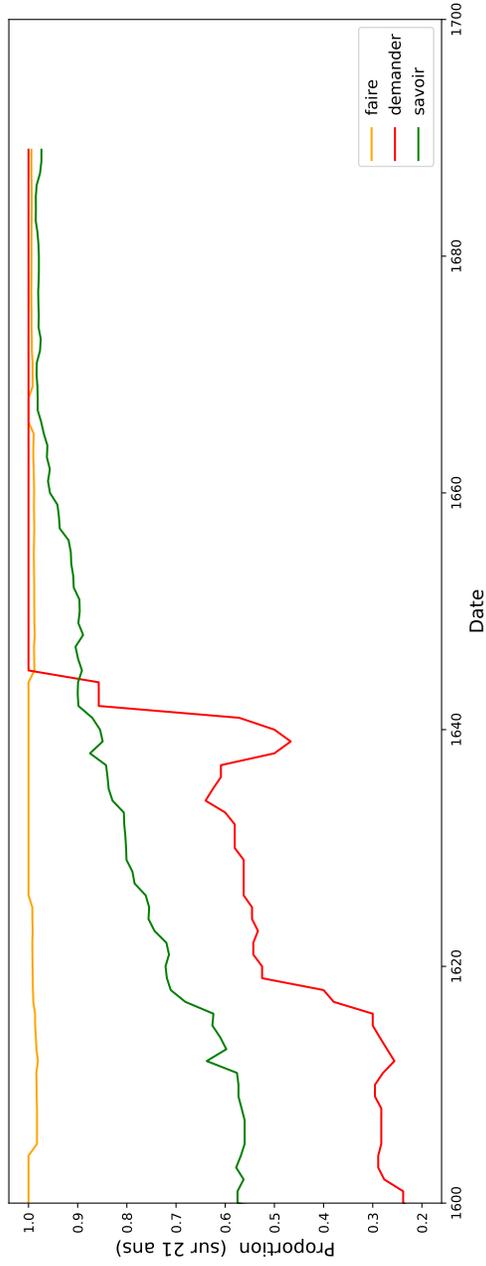


FIGURE 3.10 – Proportion de ‘ce qu-’ en fonction de chaque verbe matrice : zoom sur la période qui va de 1600 à 1700 (moyenne mobile de 21 ans)

Le changement est désormais perceptible de façon très nette avec le pic descendant vers 1640, puis une montée en flèche. On voit bien que vers 1647, ‘ce qu-’ est désormais la norme dans les interrogatives indirectes intensionnelles. Encore une fois, comme nous avons beaucoup plus de données pour le verbe ‘savoir’ que pour le verbe ‘demander’, on trouve encore quelques occurrences de ‘ce qu-’ avec ‘savoir’ : le changement ne se fait probablement pas de façon aussi radicale et brutale.

3.3.2.3.7 Analyse statistique J’ai fait un modèle linéaire mixte simple dans R, en utilisant la fonction *lm*, du paquet *stats*. La variable dépendante est PRESENCECE, les deux variables indépendantes sont DATE et VERBE. Il s’agit de vérifier si la présence du démonstratif ‘ce’ varie en fonction de la date d’une part, en fonction du verbe matrice d’autre part (‘faire’, ‘demander’, ‘savoir’), mais aussi s’il y a une interaction entre la date et le verbe.

Il s’est avéré que tous les effets et interactions sont significatifs. Ainsi, la présence de ‘ce’ varie en fonction de la date ($p < .001$). Le fait qu’il y ait le démonstratif ou non varie également en fonction du verbe matrice, autrement dit en fonction de la structure (relative ou interrogative indirecte). Plus précisément, la différence entre les verbes ‘demander’ et ‘faire’ (donc entre interrogatives indirectes et relatives) est significative ($p < .001$), de même que la différence entre ‘demander’ et ‘savoir’ (donc entre les deux types d’interrogatives indirectes : $p = .0003$). Les interactions sont significatives ($p < .001$) : cela signifie que la différence entre les verbes ‘demander’ et ‘faire’ en ce qui concerne la présence ou non de ‘ce’ dépend de la date. Il en est de même pour la différence entre ‘demander’ et ‘savoir’.

3.3.2.4 Discussion

3.3.2.4.1 Un changement soudain Un changement brutal et soudain se produit vers le milieu du 17^{ème} siècle : alors que pendant la première moitié du siècle, ‘ce qu-’ n’a toujours pas atteint les 50% dans les interrogatives indirectes, le ‘que’ interrogatif disparaît à l’improviste pendant la deuxième moitié du siècle, avec 100% de ‘ce qu-’ dans les interrogatives indirectes. ‘Ce’ s’impose donc massivement en une période de temps très courte dans les interrogatives indirectes.

Un changement si brutal et soudain est forcément dû à une intervention externe à la langue, normative de quelque type. Il est vrai que 1635 marque un tournant dans l’histoire de la langue française avec la fondation de l’Académie Française par Richelieu, dont le rôle est clairement établi dès le départ :

« La principale fonction de l’Académie sera de travailler, avec tout le soin et toute la diligence possibles, à donner des règles certaines à notre langue et à la rendre pure, éloquente et capable de traiter les arts et les sciences. » [article XXIV]

Au 17^{ème} siècle, on accorde une grande importance aux langues classiques et la langue normative imposée par l'Académie doit avoir la clarté et l'éloquence que l'on accorde au latin.

Plus particulièrement, l'intervention de certains grammairiens ou remarqueurs du grand siècle peut être à l'origine de ce changement. En effet, en faisant quelques recherches, on se rend compte que Vaugelas a très probablement eu une grande influence en ce qui concerne le remplacement de 'qu-' par 'ce qu-' dans les interrogatives indirectes. FOURNIER, 1998, le signale dans sa *Grammaire du français classique*, p.210 :

« Au XVII^e siècle, 'que' peut encore être connecteur percontatif comme en ancien français [...] et au XVI^e siècle [...]. Mais cet emploi, condamné par Vaugelas, est en régression très sensible dès le tournant du siècle et ne se maintient guère qu'en fonction objet dans les percontatives à l'infinitif⁷ (c'est la seule distribution possible en français moderne) [...] et c'est le supplétisme par 'ce que' qui est usuel. »

Mais qui était Vaugelas? Claude Favre de Vaugelas (1585-1650) était un grammairien du 17^{ème} siècle, célèbre pour avoir écrit ses *Remarques sur la langue françoise utiles à ceux qui veulent bien parler et bien écrire* (VAUGELAS, 1647), qui datent de 1647 (au tournant du siècle justement). Il a été un des membres de l'Académie Française dès sa création. Il était passionné par l'étude, la réflexion, les discussions grammaticales et il a rapidement acquis la réputation de l'un des hommes de France parlant le plus correctement la langue française à l'époque et l'un des hommes de France connaissant le mieux les règles de la langue française à l'époque. Son autorité en matière de langue française a subsisté durant tout le siècle.⁸ Il portait un jugement très négatif sur l'usage populaire de la langue, condamné par l'Académie Française. Il affirme dans la préface de ses *Remarques* que le bon usage correspond à « la façon de parler de la plus saine partie de la Cour, conformément à la façon d'écrire de la plus saine partie des Auteurs du temps. » Voilà l'extrait intitulé 'Que c'est', dans lequel il impose l'usage de 'ce que', venant détrôner le 'que' interrogatif.

« On ne dit plus gueres maintenant *que c'est*, comme l'on disoit autrefois. On dit *ce que c'est*. Par exemple, M. de Malherbe dit, *Il n'y a point de loy qui nous apprenne que c'est, que l'ingratitude..* Aujourd'huy l'on dit, *qui nous apprenne ce que c'est que, [...]* » (p.173)

La question que l'on se pose naturellement est : pourquoi donc imposer le remplacement de 'que' par 'ce que'? En quoi l'usage de 'ce que' dans les interrogatives indirectes serait

7.

(1) *Je ne sais que manger.*

8. Source : Wikipédia

plus éloquent, plus pur, plus clair que l'emploi de 'que' ? Car comme nous l'avons déjà fait remarquer à plusieurs reprises dans la présente thèse, l'emploi de 'ce que' est au contraire obscur et mystérieux dans les interrogatives indirectes. Une des hypothèses que l'on peut faire est la suivante : on impose 'ce que' par analogie avec la construction latine, étant donné l'importance accordée au latin à l'époque, langue que l'on trouve claire, pure et éloquente. En latin, c'est le pronom relatif neutre 'quod' qui introduisait les relatives aujourd'hui introduites par 'ce que'.

- (80) age **quod** agis
'Fais (bien) ce que tu fais.'

Ce pronom relatif neutre pouvait être, à titre facultatif, accompagné du pronom démonstratif neutre 'id', correspondant à 'ce'.

- (81) age **id quod** agis
'Fais (bien) ce que tu fais.'

Les interrogatives indirectes, elles, sont introduites par le pronom interrogatif neutre 'quid'.

- (82) dic mihi **quid** cogites
'Dis-moi ce que tu penses.'

BEAUZÉE, 1819, dans sa *Grammaire Générale*, affirme que 'quid' est par essence démonstratif, et il analyse d'ailleurs 'quid' en le décomposant en 'quod + id', ce qui est très intéressant car cela correspond à notre 'ce que' des relatives. Si 'quid' est perçu comme comportant un élément de nature nominale et si le latin est perçu comme une langue prestigieuse sur laquelle il faut se baser pour parler un français pur, alors on comprend pourquoi les grammairiens normatifs du 17^{ème} siècle tels que Vaugelas ont imposé l'usage de 'ce que' dans les interrogatives indirectes.

3.3.2.4.2 Des données en faveur de l'hypothèse de l'incorporation ? Les données diachroniques nous montrent que 'ce qu-' a clairement un statut nominal et qu'il met du temps à s'imposer dans les interrogatives indirectes. Le fait que l'insertion de matériel lexical entre 'ce' et 'qu-' ait été possible en ancien français et ne le soit plus en français moderne, et le fait que 'ce' ait perdu son autonomie semblent tous deux aller dans la direction de l'hypothèse de l'incorporation, du moins dans les interrogatives indirectes. De plus, le fait que 'ce que' ait pu introduire des interrogatives directes pendant une certaine période montre qu'il y a eu une tentative d'incorporation de 'ce' et 'que' au 12^{ème} siècle, mais cela ne s'est pas maintenu. On observe clairement une résistance à étendre cette séquence de nature nominale aux interrogatives indirectes. 'Ce que' finit par s'imposer dans les interrogatives indirectes, mais pas dans les questions directes : il s'agit clairement d'une séquence D+C, comme dans les relatives. Nous verrons d'ailleurs que les enfants semblent analyser 'ce que' comme une séquence D+C (Chapitre 4.), et que même en synchronie, ces interrogatives partagent de curieuses propriétés avec les relatives, contrairement aux autres interrogatives indirectes (Chapitre 7.).

3.3.2.4.3 Vers une nouvelle différenciation des deux systèmes ? Il existe une réelle interdépendance entre le système des relatives et le système des interrogatives indirectes. De l'ancien français au français moderne, les relatives ont clairement eu une influence sur les interrogatives avec le remplacement de 'que' interrogatif par 'ce que' dans les interrogatives indirectes. Comme le fait déjà remarquer NIÉGER, 1975, il semblerait qu'en français contemporain soit amorcé un mouvement inverse, c'est-à-dire que l'on se dirigerait à nouveau vers une différenciation des deux systèmes. En effet, 'qu'est-ce que', à l'origine utilisé dans les interrogatives directes, est de plus en plus employé dans les interrogatives indirectes, du moins dans la langue parlée, non standard (déjà signalé dans le Chapitre 2, voir également étude dans FrWaC dans le Chapitre 6).

On pourrait même aller plus loin, après cette nouvelle différenciation, et retourner vers deux systèmes interdépendants, où les interrogatives indirectes cette fois-ci auraient une influence sur les relatives. En effet, 'qu'est-ce que' est étendu aux relatives en français québécois et on peut supposer que cette extension puisse se produire en français parlé également en France. On observe même une tendance à remplacer 'ce que' par 'qu'est-ce que' dans les exclamatives, toujours en français parlé, non-standard.

(83) qu'est-ce qu'elle est belle !

Comme l'affirme NIÉGER, 1975 :

« L'extension est sans doute liée, au départ, aux liens étroits entre les relatives et les interrogatives, mais la tendance semble maintenant s'inscrire dans un phénomène plus général. » (p.13).

3.4 Conclusion

Cette étude diachronique confirme le statut nominal de 'ce qu-', qui apparaît en premier lieu dans les relatives. Il faut attendre le milieu du 17^{ème} siècle pour qu'il s'étende aux interrogatives indirectes. Ce changement, qui est soudain, est probablement dû à l'influence de remarqueurs et grammairiens accordant une grande importance à la norme, tels que Vaugelas. Cette séquence semble bel et bien analysable comme une séquence D+C et les données diachroniques prises dans leur ensemble ne viennent pas confirmer l'hypothèse de l'incorporation.

Toujours dans le but d'éclaircir le statut syntaxique de cette ambiguïté entre relatives et interrogatives indirectes et de vérifier la plausibilité de l'hypothèse de l'incorporation, j'ai étudié, dans un deuxième temps, l'acquisition de 'ce qu-' par les enfants. En effet, la façon dont est acquis ce morphème complexe peut être un révélateur de son statut syntaxique. Les expériences que j'ai mises en place dans le domaine de l'acquisition du langage et leurs résultats sont détaillés dans la Partie III de la présente thèse, contenant deux chapitres, un

portant sur le français et un portant sur le portugais. Cette comparaison est pertinente étant donné les différences syntaxiques observées entre les deux langues romanes.

Troisième partie

Acquisition des relatives à tête allégée et interrogatives indirectes : du français au portugais

Chapitre 4

Expériences en acquisition : la sensibilité des enfants au double statut de *ce qu-*, ainsi qu'à la nature de la tête / mot-wh^{1 2}

Comme précisé précédemment, je suis allée dans deux directions différentes afin de faire la lumière sur le statut et l'analyse syntaxiques des mystérieuses interrogatives indirectes introduites par 'ce qu-'. La première consistait à faire une étude diachronique, étude que je viens de détailler dans le Chapitre 3 qui précède. La seconde consistait à étudier les données de l'acquisition. En effet, le schéma d'acquisition peut être un indicateur, voire un révélateur de la structure syntaxique d'une construction donnée. C'est cette deuxième direction que je vais explorer dans le présent Chapitre.

4.1 L'acquisition des relatives à tête allégée et des interrogatives indirectes en français

De nombreuses études se sont centrées sur l'acquisition des relatives en français (LABELLE, 1990, GUASTI et SHLONSKY, 1995, LABELLE, 1996, FRAGMAN, 2000, entre autres) d'une part et sur l'acquisition des questions en français d'autre part (HULK, 1995, PLUNKETT, 1999, PLUNKETT, 2000, HULK et ZUCKERMAN, 2000, HAMANN, 2006, STRIK, 2007, HAMANN, 2008, JAKUBOWICZ et STRIK, 2008, entre autres), voire sur l'acquisition des deux types de struc-

1. Le matériel et la procédure des expériences qui seront décrites dans le présent chapitre ont été approuvés par le **Conseil d'évaluation éthique pour les recherches en santé** (CERES), de l'Université Paris Descartes, sous la présidence de Jacqueline Fagard (voir Avis Favorable par CERES dans les appendices, Annexe A).

2. J'ai présenté les résultats d'expériences détaillées dans cette partie Acquisition à deux conférences internationales, *The 42nd Boston University Conference on Language Development (BUCLD)* (KONRAD, HAENDLER et DONATI, 2017b) et *The Romance Turn IX* (KONRAD, 2018a), ainsi qu'à la 10ème Journée des Doctorants de l'Université Paris Diderot (KONRAD, 2017a) et au Colloque organisé en l'honneur de Janet Fodor à l'Université Paris Diderot (KONRAD, HAENDLER et DONATI, 2017a). J'ai également publié un article dans les *proceedings* de BUCLD, en collaboration avec Yair Haendler et Caterina Donati (KONRAD, HAENDLER et DONATI, 2018).

tures (BENŢEA et DURRLEMAN, 2013, entre autres). Toutefois, à ma connaissance, aucune d'entre elles n'a étudié l'acquisition des constructions présentant 'ce que' en français, en particulier l'ambiguïté entre relatives et interrogatives indirectes (décrite dans le Chapitre 2). Les études sur les relatives se centrent la plupart du temps sur l'asymétrie entre relatives sujet et relatives objet. Et souvent sur les relatives pleines, rarement sur les relatives à tête allégée ou sur les relatives libres. Quant aux études sur les interrogatives, elles se focalisent sur l'acquisition des questions racines ou bien des questions à longue distance, et pas sur les interrogatives indirectes.

4.1.1 Une seule étude traitant en partie de 'ce que' : BENTEÀ, DURRLEMAN et RIZZI, 2016

4.1.1.1 Background : les effets d'intervention

Un fait bien connu est que les dépendances A-barre (notamment les relatives et les questions wh-) sont plus faciles à traiter/produire/comprendre lorsque le *gap* est en position sujet que lorsqu'il est en position objet, en français mais aussi dans beaucoup d'autres langues (parmi les nombreuses études à ce sujet, nous pouvons citer FRAGMAN, 2000 pour le français; GOODLUCK et TAVAKOLIAN, 1982 pour l'anglais; COSTA, LOBO et SILVA, 2011 pour le portugais européen; FRIEDMANN et NOVOGRODSKY, 2004 pour l'hébreu; ADANI, 2011 pour l'italien; ADANI, SEHM et ZUKOWSKI, 2012 pour l'allemand).

Une explication possible, de nature grammaticale, est que les enfants ont des difficultés avec certaines configurations impliquant un mouvement, dues au fait que l'objet est déplacé en croisant un sujet avec lequel il partage des traits morphosyntaxiques. Autrement dit, ce ne sont pas toutes les dépendances objet qui sont difficiles finalement, mais celles dans lesquelles le sujet et l'objet ont une matrice de traits similaire. Inversement, si le sujet et l'objet ne partagent pas de traits morphosyntaxiques, les dépendances en question devraient être plus faciles à traiter. Cette explication en termes d'effets d'intervention est basée sur le principe de localité appelé 'Minimalité Relativisée', qui est opératoire dans la grammaire adulte (RIZZI, 1990). Ce que dit ce principe est qu'une relation locale ne peut être établie entre X et Y s'il y a un intervenant Z du même type structural que X et qui est un candidat potentiel pour la relation en question (par exemple, le mouvement A-barre)³.

FRIEDMANN, BELLETTI et RIZZI, 2009, ont montré, à partir de tâches de compréhension portant sur des relatives et sur des questions wh- à la fois sujet et objet en hébreu, que les enfants peinaient à comprendre seules les dépendances A-barre objet dont l'élément déplacé était un DP plein (de type D+N, par exemple : 'la fille' pour les relatives; 'quelle

3. *Relativized Minimality* : $X \alpha$ -governs Y only if there is no Z such that

- (i) Z is a typical potential α -governor for Y ,
- (ii) Z c-commands Y and does not c-command X .

[RIZZI, 1990, p.7]

filles' pour les questions wh-). Ainsi, dans les cas où le sujet et l'objet sont des DPs pleins, la performance est significativement différente de la performance dans les cas où l'objet est un DP plein mais le sujet est de nature pronominale. Le trait $[\pm NP]$ exprime la restriction lexicale et est un des traits pertinents pour la computation. En français, la relative en (1) devrait être plus difficile que celle en (2). De même dans le cas des questions wh- : (3) devrait être plus complexe sur le plan du traitement pour les enfants que (4).

- (1) *la fille* [_{+NP}] [*que lamaman coiffe e*] [_{+NP}]
 (2) *la fille* [_{+NP}] [*qu' elle coiffe e*] [_{-NP}]
 (3) *Quelle fille lamaman coiffe e ?* [_{+NP}] [_{+NP}]
 (4) *Quelle fille elle coiffe e ?* [_{+NP}] [_{-NP}]

Les exemples (2) et (4) ne sont pas problématiques pour l'enfant : comme le sujet est de nature pronominale, il ne partage pas de trait de restriction lexicale, noté [+NP], avec l'objet qui se déplace. Par contre, en (1) et en (3), le sujet partage le trait de restriction lexicale [+NP] avec l'objet qui se déplace. Le sujet agit alors comme un élément intervenant. De ce fait, le traitement de ces dépendances objet est plus complexe pour les enfants.

D'après les auteurs, la spécificité des enfants par rapport aux adultes n'est pas de nature grammaticale, mais de nature computationnelle. Le système de l'enfant est pourvu de capacités computationnelles plus faibles que le système de l'adulte. Il s'agit d'un problème de calcul : calculer une relation de sous-ensemble ou sur-ensemble est plus difficile pour l'enfant, car cela exige plus de ressources cognitives (plus de mémoire, par exemple). Le système de l'enfant contiendrait ainsi une version plus stricte de la Minimalité Relativisée que le système de l'adulte.

			Adult grammar	Child grammar		
(a)	+A	+A	<+A>	*	*	(identity)
(b)	+A,+B	+A	<+A,+B>	ok	*	(inclusion)
(c)	+A	+B	<+A>	ok	ok	(disjunction)

FIGURE 4.1 – La grammaire de l'adulte *versus* la grammaire de l'enfant : schéma récapitulatif présenté par FRIEDMANN, BELLETTI et RIZZI, 2009 : 84.

Comme l'illustre la Figure 4.1, les adultes n'ont des difficultés à traiter les dépendances objet que dans les cas où la matrice de traits du sujet est identique à celle de l'objet (voir (a)). Par contre, la relation d'inclusion (lorsque la matrice de traits du sujet contient un ou des traits présents dans la matrice de traits de l'objet) est également problématique pour les enfants (voir (b)). Enfin, la configuration de disjonction (aucun trait de la matrice de traits

du sujet ne se retrouve dans la matrice de traits de l'objet) n'est problématique ni pour les adultes ni pour les enfants (voir (c)). Bien évidemment, cette approche prédit que les dépendances sujet sont plus faciles que les dépendances objet du fait de l'absence d'intervenant.

Au sein des relatives objet, la prédiction est que les relatives libres objet devraient être plus faciles à traiter que les relatives pleines objet avec un sujet plein, si le trait pertinent est le trait [\pm NP]. En effet, dans le cas des relatives libres objet avec un sujet plein, le mot *wh-* 'qui' a un trait [-NP] (pas de restriction lexicale) alors que le sujet a un trait [+NP] (restriction lexicale). Cette différence en termes de restriction lexicale entre le sujet et l'objet ne rend pas difficiles les relatives libres objet. Par contre, dans le cas des relatives pleines objet avec un sujet plein, à la fois l'objet et le sujet ont un trait [+NP] (restriction lexicale), il s'agit d'une identité en termes de restriction lexicale, relation difficile à calculer.

4.1.1.2 L'expérience de Bentea et al.

4.1.1.2.1 Objectif général BENTEA, DURRLEMAN et RIZZI, 2016 ont pour objectif de mieux comprendre la nature et l'impact du trait [\pm NP], ainsi que son interaction avec l'animacité. Ils ont mis en place des tests de compréhension de relatives et de questions *wh-* objet en français.

4.1.1.2.2 Design Comme FRIEDMANN, BELLETTI et RIZZI, 2009, ils ont mis en place une tâche de choix de personnage, à partir d'images. Je vais seulement présenter les items et les résultats concernant les relatives (pas ceux concernant les questions) car ce sont les structures qui nous intéressent ici. Ils ont manipulé deux traits de la tête : le trait [\pm NP] et le trait [\pm animé]. Le sujet est toujours [+NP] et [+animé]. Voilà des exemples d'images (voir Figure 4.2), ainsi que les items associés (BENTEA, DURRLEMAN et RIZZI, 2016 : 24-25).

- (5) Montre-moi la dame que la fille embrasse. [+NP] et [+animé]
- (6) Montre-moi la balle que la fille tape. [+NP] et [-animé]
- (7) Montre-moi celle que la fille embrasse. [-NP] et [+animé]
- (8) Montre-moi ce que la fille tape. [-NP] et [-animé]

4.1.1.2.3 Participants Ont participé à cette étude 61 enfants, présentant un développement typique, âgés entre 5 et 11 ans, testés dans deux écoles primaires à Genève, en Suisse.

4.1.1.2.4 Les résultats Les résultats concernant les relatives sont représentés dans la Figure 1.2. Globalement, on peut voir que les relatives objet avec tête [+NP] sont plus difficiles que les relatives objet avec tête [-NP]. De plus, dans le cas des relatives objet avec tête [+NP], le fait qu'il y ait un *mismatch* de traits en ce qui concerne l'animacité semble faciliter la compréhension. En effet, lorsque l'objet est [-animé], contrairement au sujet qui lui est [+animé], la performance est plus élevée.

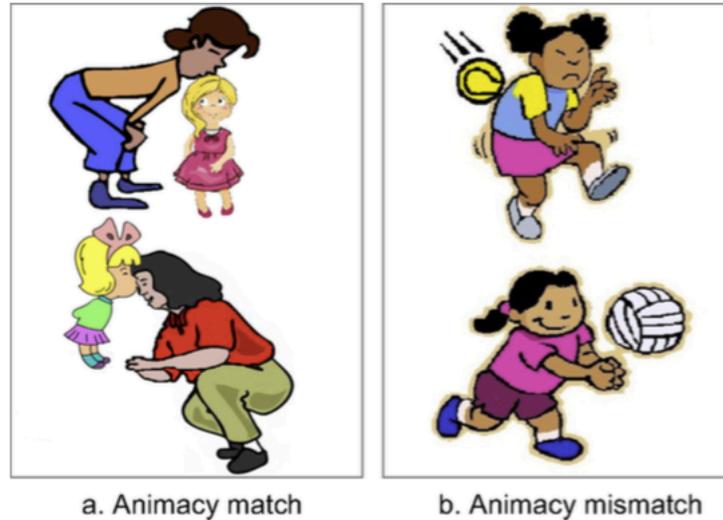


Fig. 1. Pictures used in the character-selection task.

FIGURE 4.2 – Images utilisées par BENTEA, DURRLEMAN et RIZZI, 2016 pour leur tâche de choix de personnage (p.27).

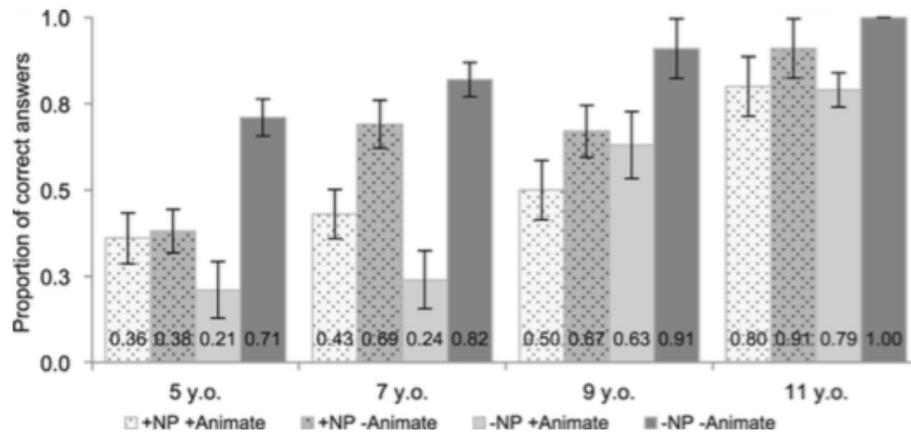


FIGURE 4.3 – Proportion de réponses correctes pour les relatives objet, en fonction des deux variables manipulées : NP et animacité, et en fonction de l'âge (BENTEA, DURRLEMAN et RIZZI, 2016 : 29).

Les résultats qui nous intéressent le plus ici sont ceux pour la condition [-NP, -animé] car cela correspond aux relatives objet introduites par 'ce que'. On peut observer que cette condition semble être celle qui pose le moins de problème aux enfants. Ce qui est étonnant est que les relatives objet introduites par 'celle que' sont les plus difficiles pour les enfants,

bien que la tête soit [-NP]. Il y a un contraste très net entre ‘celle’ et ‘ce’, les relatives avec ‘celle’ étant moins bien comprises par les enfants que celles avec ‘ce’. D’après les auteurs, cela est dû au fait que ‘celui’ ou ‘celle’ est en fait décomposable en ce+lui ou ce+elle, et donc en D+NP en quelque sorte, alors que ‘ce’ est un simple D (nous l’avons d’ailleurs suggéré dans le Chapitre 1, Section 1.3.2.6.1). Cela voudrait dire qu’en réalité ‘celle/celui’ est [+NP], et non pas [-NP]. Notons que, si leur analyse est la bonne, on devrait trouver des effets d’intervention lorsque le sujet est ‘celle-ci/celui-ci’ et l’objet [+NP].

(9) Montre-moi la **ball**e que **celui-ci** tape.

De plus, il est plus difficile d’avoir accès au référent du discours de ‘celle’ comparée à ‘la fille’, ce qui expliquerait également le contraste entre les deux.

Pour résumer en ce qui concerne les relatives à tête allégée, cette étude de BENTEA, DURR-LEMAN et RIZZI, 2016 semble indiquer que les enfants n’ont pas de difficultés particulières à comprendre les relatives objet introduites par ‘ce que’, ce qui est prédit par l’analyse en termes d’intervention de FRIEDMANN, BELLETTI et RIZZI, 2009, étant donné qu’il y a un *mis-match* pour le trait [±NP] et pour le trait [±animé] (l’objet ‘ce’ est [-NP] et [-animé] alors que le sujet est [+NP] et [+animé]). Un des points faibles de cette étude est le nombre limité d’items : seuls 2 items sont des relatives introduites par ‘ce que’. De plus, les enfants les plus jeunes ont 5 ans, donc on ne sait pas si, avant 5 ans, les enfants francophones sont en mesure de comprendre de telles structures. Par ailleurs, dans tous leurs items, le sujet est [+NP], mais il aurait été intéressant d’étudier l’interaction avec les traits du sujet. En particulier, ‘ce’ étant [-NP], on s’attend à ce qu’une phrase comme (10), avec un sujet pronominal (donc [-NP]), soit difficile pour les enfants.

(10) Montre-moi **ce** qu’il attrape.

4.2 Les questionnements du point de vue de l’acquisition

Nous venons de remarquer que la littérature est très pauvre au sujet de l’acquisition des subordonnées introduites par ‘ce que’ en français. Aucune étude ne s’est centrée sur le double usage de ‘ce que’, à savoir dans les relatives et les interrogatives indirectes.

Les questions que l’on peut se poser, du point de vue de l’acquisition, sur la comparaison entre relatives à tête allégée et interrogatives indirectes introduites par ‘ce qu-’ sont les suivantes :

1. Comment et quand les enfants francophones acquièrent-ils ce double usage (relatif/interrogatif) de la séquence ce+que ?
2. Est-ce que ‘ce qu-’ est acquis de façon parallèle dans les deux structures (relatives et interrogatives indirectes), ou pas ?
3. Est-ce que les enfants francophones analysent ‘ce qu-’ comme une séquence D+C ou bien comme un mot wh-, [səkə] ?

4. Si ‘ce qu-’ pose des difficultés aux enfants francophones, quelle(s) stratégie(s) syntaxique(s) utilisent-ils à la place de ‘ce qu-’ ?

Afin de répondre à ces questions, j’ai regardé dans un premier temps des données de production spontanée, qui pourraient nous donner un premier aperçu sur l’acquisition de ces structures. Mais ‘ce qu-’ étant très peu fréquent dans le discours des enfants, ces données se sont révélées insuffisantes. Dans un second temps, j’ai mis en place un test de répétition de phrases, dont les résultats permettent de répondre en grande partie aux questionnements ci-dessus.

4.3 Données de production spontanée

4.3.1 Le corpus

Le corpus étudié est celui du projet ANR CoLaJE (Communication Langagière chez le Jeune Enfant), dont la responsable scientifique est Aliyah Morgenstern (MORGENSTERN et PARISSE, 2007). Le corpus comporte des suivis longitudinaux de productions spontanées d’enfants, depuis la naissance jusqu’à 5-7 ans, selon les cas. Dans la présente étude, j’ai compté les occurrences de ‘ce que’ et ‘ce qui’ dans la production de quatre enfants : Anaé, Théophile, Madeleine et Antoine. Les transcriptions étant au format CHAT, cela permet de pouvoir faire directement des recherches dans les fichiers à l’aide des programmes CLAN du projet CHILDES (MACWHINNEY, 2000).

Enfant	Tranche d’âge	Nombre de fichiers	Durée des enregistrements
Anaé	de 2;06.27 à 5;10.30	20	20h
Théophile	de 2;07.04 à 4;11.11	18	18h
Madeleine	de 2;06.10 à 6;11.27	18	18h
Antoine	de 3;00.21 à 6;03.08	14	14h

TABLE 4.1 – Présentation des données concernant les quatre enfants pour l’étude de production spontanée dans le corpus CoLaJE.

4.3.2 Les résultats à partir de la production des enfants

Regardons les résultats. Les tableaux ci-dessous montrent les occurrences de ‘ce que’ et ‘ce qui’ dans la production spontanée de chaque enfant, tout en indiquant l’âge de l’enfant pour chaque occurrence.

Âge	ce qui	ce que
2;11.22	1 [AMB]	0
4;00.13	0	2 [AMB, INT]
4;08.09	1 [INT]	0
5;01.21	0	2 [AMB, AMB]
5;07.06	3 [-SUB, -SUB, -SUB]	1 [INT]
TOTAL	5	5

TABLE 4.2 – Nombre d’occurrences de *ce qui* et *ce que* dans la production spontanée d’Anaé, en fonction de son âge. Type de structure indiqué entre crochets : AMB = ambigu, INT = interrogative indirecte, -SUB = non enchâssé.

Âge	ce qui	ce que
3;08.15	1 [AMB]	0
3;11.15	0	1 [AMB]
4;01.24	0	2 [AMB, -SUB]
4;03.29	0	2 [-SUB, AMB]
4;06.20	0	3 [AMB, -SUB, AMB]
4;09.07	1 [AMB]	4 [INT, AMB, REL, -SUB]
TOTAL	2	12

TABLE 4.3 – Nombre d’occurrences de *ce qui* et *ce que* dans la production spontanée de Théophile, en fonction de son âge. Type de structure indiqué entre crochets : AMB = ambigu, INT = interrogative indirecte, -SUB = non enchâssé, REL = relative.

Âge	ce qui	ce que
4;08.02	1 [REL]	4 [AMB, CLIV, CLIV, REL]
5;02.13	1 [AMB]	1 [AMB]
5;07.11	0	2 [AMB, -SUB]
5;11.18	0	1 [-SUB]
TOTAL	2	8

TABLE 4.4 – Nombre d’occurrences de *ce qui* et *ce que* dans la production spontanée d’Antoine, en fonction de son âge. Type de structure indiqué entre crochets : AMB = ambigu, -SUB = non enchâssé, REL = relative, CLIV = clivée.

Âge	ce qui	ce que
2;08.05	0	1 [-SUB]
3;00.28	1 [-SUB]	2 [AMB, PRES]
3;03.02	2 [-SUB, -SUB]	1 [REL]
3;09.07	1 [INT]	0
3;09.28	0	2 [-SUB, REL]
4;01.27	0	1 [INT]
4;10.03	1 [-SUB]	1 [-SUB]
5;01.05	0	2 [AMB, PRES]
5;05.08	0	2 [-SUB, -SUB]
6;05.14	4 [REL, AMB, REL, -SUB]	2 [INT, AMB]
6;11.27	2 [-SUB, AMB]	3 [AMB, -SUB, REL]
TOTAL	11	17

TABLE 4.5 – Nombre d’occurrences de *ce qui* et *ce que* dans la production spontanée de Madeleine, en fonction de son âge. Type de structure indiqué entre crochets : AMB = ambigu, INT = interrogative indirecte, -SUB = non enchâssé, REL = relative, PRES = présentative.

Comme on peut le voir, les nombres d’occurrences sont très petits. Globalement, on peut noter que les enfants français produisent rarement des subordinées introduites par ‘ce qu-’, entre 2 et 6 ans. D’après les données d’Anaé et de Madeleine, certains enfants commenceraient à produire de telles structures un peu avant 3 ans. Dans l’ensemble, on peut constater que ‘ce qui’ est utilisé un peu moins fréquemment que ‘ce que’ par les enfants.

Quant à la distribution entre relatives et interrogatives indirectes, qui nous intéresse ici, on ne peut pas vraiment déterminer d’après ces données si ‘ce qu-’ est davantage employé dans les relatives ou bien dans les interrogatives, d’une part parce les données sont trop peu nombreuses et d’autre part, parce que, dans beaucoup de cas, le verbe de la proposition matrice est un verbe ambigu. Le verbe ‘regarder’ est le verbe qui sélectionne très fréquemment une subordinée introduite par ‘ce qu-’ dans ces données.

- (11) Regarde ce que je fais avec la paille. [Anaé, 4;00.13]
- (12) Regarde ce que j’ai fait. [Théophile, 4;01.24]
- (13) Regarde ce que j’ai fait pour maman. [Madeleine, 6;11.27]
- (14) Papa, regarde ce que j’ai fait. [Antoine, 5;07.11]

Ces phrases sont ambiguës puisqu’elles peuvent avoir l’interprétation d’un DP, c’est-à-dire d’une relative (regarde **la chose** que je fais) ou bien celle d’un CP, c’est-à-dire d’une interrogative indirecte (regarde **quelle chose** je fais). On retrouve la même ambiguïté avec les verbes ‘dire’ et ‘voir’.

(15) Tu vas voir ce que je vais découper ! [Antoine, 4;08.02]

Parmi ces occurrences de ‘ce qu-’, on trouve quand même plusieurs interrogatives indirectes, sans ambiguïté (9 au total, dans la production des quatre enfants), la plupart du temps introduites par le verbe ‘savoir’.

(16) Je sais pas ce que ça serait mais ça serait pas supermag. [Madeleine, 6;05.14]

Dans très peu de cas, la proposition introduite par ‘ce qu-’ est en position de sujet au sein de la proposition matrice. Dans ces cas, il s’agit forcément de relatives.

(17) mais ce qui est en haut c’est pour toi [Madeleine, 3;03.02]

Par ailleurs, au sein des structures relatives, on trouve des cas particuliers : quelques structures présentatives et clivées.

(18) voilà ce que j’ai dessiné [Madeleine, 5;01.05]

(19) c’est ce que je te demande de faire [Antoine, 4;08.02]

La tête ‘ce’ est modifiée par le quantifieur ‘tout’ dans quelques relatives.

(20) où tout ce qui était grand est devenu petit [Anaé, 5;07.06]

(21) tout ce que je mangeais [Théophile, 4;06.20]

Ces données de production spontanée peu importantes quantitativement ne permettent pas de répondre à un ou plusieurs des questionnements évoqués dans la section précédente. La seule conclusion que l’on peut faire est que les propositions introduites par ‘ce qu-’ sont peu fréquentes dans la parole enfantine, entre 2 et 6 ans. Cette faible fréquence peut à son tour laisser sous-entendre que ces structures posent des difficultés aux enfants de cette tranche d’âge. Mais peut-être aussi qu’ils n’ont tout simplement pas les contextes pour avoir ces structures.

Il serait intéressant de vérifier si les enfants sont exposés à ces structures. Les adultes produisent-ils ces structures dans le discours adressé à l’enfant ? Si oui, trouve-t-on une différence quantitative entre les relatives introduites par ‘ce que’ et les interrogatives indirectes introduites par ‘ce que’ ?

4.3.3 Les résultats à partir de l’input

Regardons les résultats relatifs à la parole adressée à l’enfant dans le même corpus. Les tableaux ci-dessous montrent le nombre d’occurrences de ‘ce que’ et ‘ce qui’ dans l’input de chaque enfant, en indiquant s’il s’agit d’interrogatives indirectes, de relatives ou de structures ambiguës, pouvant être interprétées comme les premières ou les dernières.

(22) je sais pas ce que tu as vu toi. [INT ; mère d’Anaé ; 4;01.18]

	CE QUE	CE QUI
INT	22	2
REL	10	2 (+1 clivée)
AMB	30	6
TOTAL	62	10 (+1 clivée)

TABLE 4.6 – Nombre d’occurrences de *ce qui* et *ce que* dans l’input d’Anaé, selon le type de structure.

- (23) tu **pré pares** les [//] **ce que** tu veux emmener en vacances. [REL; mère d’Anaé; 4;00.13]
- (24) **regarde ce que** tu fais [AMB; mère d’Anaé; 2;10.24]
- (25) **ramasse ce qui** est tombé par terre [REL; mère d’Anaé; 3;01.07]

	CE QUE	CE QUI
INT	12	2
REL	11 (+1 clivée)	2
AMB	20	4
TOTAL	43 (+1 clivée)	8

TABLE 4.7 – Nombre d’occurrences de *ce qui* et *ce que* dans l’input de Théophile, selon le type de structure.

- (26) ah ben je **sais pas ce que** tu veux [INT; mère de Théophile; 4;09.07]
- (27) ... m
 ais je peux **faire** tout **ce que** je veux c’est génial [REL; mère de Théophile; 2;11.28]
- (28) ouais mais toi tu choisis tu peux **choisir ce que** tu veux lui donner [AMB; observateur de Théophile; 4;11.11]
- (29) je **sais pas** du tout **ce qui** se passe dans ta vie, **ce qui** se passe à l’école, je sais rien du tout de ça [INT; observateur de Théophile; 4;06.20]
- (30) je **me demande ce que** ça va être ! [INT; observateur de Madeleine; 4;01.27]
- (31) il faut pas **défaire ce que** ta sœur fait parce que là ça va pas bien se passer [REL; mère de Madeleine; 3;03.02]
- (32) alors tu fais une fois pis tu **regardes ce que** la flèche montre [AMB; mère de Madeleine; 2;06.10]

	CE QUE	CE QUI
INT	22	1
REL	6 (+2 clivées)	2
AMB	21	3
TOTAL	49 (+2 clivées)	6

TABLE 4.8 – Nombre d’occurrences de *ce qui* et *ce que* dans l’input de Madeleine, selon le type de structure.

	CE QUE	CE QUI
INT	21	1
REL	11 (+5 clivées)	4
AMB	30	0
TOTAL	62 (+5 clivées)	5

TABLE 4.9 – Nombre d’occurrences de *ce qui* et *ce que* dans l’input d’Antoine, selon le type de structure.

- (33) on **enlève ce qui** y a autour [REL; mère de Madeleine; 2;08.05]
- (34) moi je **sais pas du tout ce que** c’est [INT; grand-mère d’Antoine; 4;00.09]
- (35) tu **manges ce que** tu as là et on verra ensuite [REL; tante d’Antoine; 3;02.24]
- (36) **montre-moi ce que** tu faisais avec le petit garçon [AMB; mère d’Antoine; 5;11.18]
- (37) en fait i(l) **fait tout ce qui** est en bois [REL; père d’Antoine; 6;03.08]

Globalement, on peut remarquer que les enfants sont bel et bien exposés à des subordonnées introduites par ‘ce qu-’ : 73 occurrences en tout dans l’input d’Anaé ; 51 dans l’input de Théophile ; 55 dans l’input de Madeleine ; 67 dans l’input d’Antoine. Notons également que ‘ce qui’ est beaucoup moins fréquent dans l’input que ‘ce que’ : 10 occurrences de ‘ce qui’ contre 62 occurrences de ‘ce que’ dans l’input d’Anaé ; 8 *versus* 43 dans l’input de Théophile ; 6 *versus* 49 dans l’input de Madeleine ; 5 *versus* 62 dans l’input d’Antoine. En ce qui concerne la distribution de ‘ce que’, on trouve davantage d’occurrences de ‘ce que’ dans les interrogatives indirectes que dans les relatives dans l’input des quatre enfants. Mais on relève une proportion importante de ‘ce que’ dans des contextes ambigus, c’est-à-dire des subordonnées sélectionnées par des verbes pouvant sélectionner un DP (donc une relative) ou un CP (donc une interrogative indirecte), tels que ‘regarder’, ‘montrer’, ‘dire’.

Nous savons maintenant que les enfants français ne produisent pas beaucoup de subordonnées introduites par ‘ce qu-’, bien qu’ils soient exposés à ces structures. Supposons qu’ils

ont des difficultés à produire de telles structures, la question qui se pose est la suivante : quelle(s) stratégie(s) syntaxique(s) utilisent-ils à la place de ‘ce qu-’ ? quel(s) type(s) d’erreur font-ils s’ils ne réalisent pas correctement les structures cibles ? Afin de tenter de répondre à ces questions, j’ai mis en place un test de répétition de phrases contenant des subordonnées introduites par ‘ce qu-’.

4.4 Expérience 1 : test de répétition de phrases contenant des subordonnées introduites par *ce qu-*

4.4.1 Prédiction

De manière générale, du fait de la faible fréquence dans les données de production spontanée, on s’attend à ce que les enfants français âgés entre 3 et 6 ans aient des difficultés avec ces subordonnées introduites par ‘ce qu-’. Rappelons que ces structures sont ambiguës dans la grammaire de l’adulte et que nous cherchons à déterminer si les enfants analysent ce+que comme un élément wh- ou bien comme une séquence D+C. L’idée est alors de comparer l’acquisition de ‘ce qu-’ avec celle d’un mot wh-, sans élément D, comme ‘qui’.

Si les enfants analysent ce+que comme un élément wh- *səka* (ce qui irait dans le sens de l’hypothèse de l’incorporation présentée dans le Chapitre 2), on s’attend à trouver des tendances similaires dans l’acquisition de ‘ce qu-’ et dans l’acquisition de ‘qui’. Si au contraire ils analysent ce+que comme une séquence impliquant un déterminant suivi d’un CP, ils devraient mieux répéter les relatives (où une structure nominale est sélectionnée) que les interrogatives indirectes (une structure nominale devrait être incompatible dans ce cas). On s’attend alors à ce que des erreurs de substitution (avec des mots wh-) surviennent plus fréquemment dans les interrogatives indirectes (où un élément wh- est requis) que dans les relatives (où un élément wh- n’est pas forcément requis). Dans ce cas, on s’attend également à des divergences dans l’acquisition de ‘ce qu-’ comparée à l’acquisition de ‘qui’, qui est un mot wh-.

4.4.2 Design expérimental

4.4.2.1 Participants

J’ai testé 52 enfants monolingues français, tous présentant un développement typique du langage, et âgés entre 3 et 6 ans. Tous les participants ont été testés à l’Ecole Maternelle de Nogaro, dans le département du Gers. Un accord parental a été obtenu avant que je fasse passer l’expérience aux enfants. Je précise que trois des enfants ont été exclus de l’analyse. Pour deux d’entre eux, la raison est qu’ils avaient des difficultés même pour les *fillers* ou distracteurs, qui étaient des phrases très simples. Le troisième enfant a été exclu à cause d’un réel manque de concentration pendant la tâche.

4.4.2.2 Matériel

L'expérience consistait en un design regroupant 4 conditions expérimentales, obtenues en croisant deux facteurs : le facteur TYPE DE PHRASE (relatives ou interrogatives indirectes) et le facteur TYPE DE GAP (*gap* sujet ou objet). Pour chaque condition, j'ai construit 4 items, ce qui conduit à 16 items expérimentaux au total. A cela s'ajoutent 10 *fillers* ou distracteurs. Il y avait également 3 items d'entraînement au début de l'expérience, afin de s'assurer que les enfants, en particulier les plus jeunes, comprenaient bien la tâche à réaliser.

Les items ont été construits en utilisant un vocabulaire qui fait référence à l'environnement de l'enfant (en particulier, à l'école ou à la maison), afin de s'assurer que les enfants soient familiers avec les mots et les événements décrits dans les phrases à répéter. Un point important est que j'ai contrôlé la longueur des items, en terme de nombre de syllabes (chaque item contenait 12 syllabes), dans le but de contrôler des effets potentiels liés à la mémoire.

La structure de chaque item était la suivante, pour les dépendances objet :

(38) DP sujet - V principal - ce que - DP sujet - V subordonnée

et pour les dépendances sujet :

(39) DP sujet - V principal - ce qui - V subordonnée - adjectif

Les items ont été randomisés avec Excel, de façon à ce qu'il y ait autant de listes que de participants. Ainsi, à chaque session de test, l'enfant participant à l'expérience était exposé exactement aux mêmes items que les autres, mais dans un ordre différent.

Voilà deux exemples d'items pour chacune des quatre conditions expérimentales.

1. Relatives à tête allégée avec *gap* sujet

(40) Le papa mange ce qui est tombé du placard.

(41) Le monsieur attrape ce qui a roulé par terre.

2. Interrogatives indirectes avec *gap* sujet

(42) Papa sait pas ce qui est tombé du placard.

(43) Le monsieur demande ce qui a roulé par terre.

3. Relatives à tête allégée avec *gap* objet

(44) L'élève attrape ce que la maîtresse a lancé.

(45) Le maître efface ce que la fille a dessiné.

4. Interrogatives indirectes avec *gap* objet

(46) L'élève demande ce que la maîtresse a lancé.

(47) Le maître demande ce que la fille a dessiné.

Comme l'illustrent ces exemples d'items, les deux types de structures, à savoir les relatives et les interrogatives indirectes, sont construits de façon parallèle : la subordonnée est identique, seule la proposition matrice change, en particulier le verbe matrice, qui sélectionne soit un DP (une relative), soit un CP (une interrogative indirecte).

La liste des verbes purement interrogatifs (non ambigus, sélectionnant exclusivement un CP) est très réduite en français (et dans les langues en général), comparée à la liste de verbes sélectionnant un DP, qui est bien moins limitée. Pour les interrogatives indirectes, j'ai utilisé les verbes 'demander', 'deviner' et '(ne) pas savoir'. Il était difficile d'en trouver un quatrième qui ne soit pas ambigu et qui soit connu des enfants entre 3 et 6 ans. 'Ignorer' par exemple est un verbe interrogatif sélectionnant un CP, mais il est trop complexe pour les enfants, qui ne le connaissent peut-être pas. Par parallélisme, je n'ai utilisé que trois verbes différents pour les relatives, même s'il existe beaucoup de possibilités : 'attraper', 'manger', 'effacer'.

Les *fillers* étaient des phrases courtes, simples, sans proposition subordonnée.

(48) La fille adore la peinture.

(49) Marie est amoureuse de Lucas.

Quant aux items d'entraînement, il est important de préciser qu'ils étaient aussi longs (12 syllabes) et complexes que les items expérimentaux. Cependant, ils ne contenaient pas de proposition subordonnée introduite par 'ce qu-' : il s'agissait de complétives introduites par 'que'. Ces items, bien qu'ils soient juste des items d'entraînement, sont un peu comme une condition de contrôle dans ce test. En effet, à travers la performance des enfants sur ces items-là, on peut déterminer s'ils ont des difficultés avec 'ce qu-' en particulier, ou bien, plus généralement, avec des phrases complexes et la subordination.

(50) Mamie pense que papi a trop mangé de soupe.

4.4.2.3 Procédure

Les enfants ont été testés par moi-même et ce de façon individuelle. La tâche était très simple : elle consistait à écouter une phrase que je prononçais et à la répéter. Bien sûr, il fallait une dimension ludique pour amuser les enfants et les motiver, sans quoi la répétition

aurait pu leur paraître très ennuyante. Je présentais donc aux enfants deux marionnettes, qui adorent entendre des histoires, mais plutôt celles racontées par les enfants, pas par les adultes. Les enfants étaient alors encouragés à interagir avec l'une des marionnettes (au choix), et à répéter chaque phrase à la marionnette.

Par ailleurs, à chaque essai, après avoir répété la phrase (correctement ou non), ils étaient autorisés à piocher, dans un sac, un cube en bois coloré afin de construire quelque chose, au fur et à mesure (un château, une tour...). Les réponses des enfants ont été enregistrées afin de pouvoir les réécouter et noter/vérifier leurs réponses. De manière générale, les enfants étaient impliqués et trouvaient la tâche amusante. L'expérience durait environ 15 minutes par enfant, sans pause.

Il a été prouvé qu'un test de répétition implique, outre la capacité de mémoire, la connaissance grammaticale et la compréhension de l'enfant (LUST, FLYNN et FOLEY, 1996). En effet, l'enfant ne devrait pas être capable de reconstruire la phrase qu'il a entendue, dans le but de la répéter, si la structure de cette phrase ne faisait pas partie de la grammaire de l'enfant à ce stade spécifique d'acquisition du langage. Mes résultats, comme nous allons le voir, confirment la validité et la pertinence de ce type de test.

4.4.3 Résultats

Je vais présenter les résultats de ce test de répétition, en deux temps : dans un premier temps, je détaillerai l'analyse quantitative, afin de voir si les enfants ont des difficultés ou non avec 'ce que', si ces difficultés se réduisent avec l'âge, et si ces difficultés concernent plutôt une structure que l'autre (relative ou interrogative indirecte). Dans un second temps, je ferai une analyse qualitative, en me focalisant sur les types d'erreurs que les enfants font lorsqu'ils répètent mal la phrase cible. Là encore, il sera intéressant de déterminer si tel ou tel type d'erreur concerne davantage les relatives ou bien les interrogatives indirectes.

4.4.3.1 Analyse quantitative : la performance des enfants

La phrase de réponse que les enfants produisaient à chaque essai a été codée comme une variable dépendante binaire : 1 si la répétition était correcte et 0 si elle ne l'était pas. Pour déterminer cela, je me suis principalement focalisée sur la répétition correcte de 'ce qu-', qui est la séquence clé ici. Cela signifie que si un enfant remplaçait un mot par un autre de même catégorie mais que tout le reste était correct, je comptais quand même 1 : par exemple, le DP 'la mamie' au lieu du DP 'la maman'. De même s'ils remplaçaient le verbe matrice par un autre verbe sans que cela implique un changement de structure (relative/interrogative indirecte). Par contre, si 'ce que' était bien réalisé mais que la phrase était agrammaticale parce que l'enfant avait omis le verbe matrice ou le verbe enchâssé par exemple), alors je comptais 0. De même si le changement de verbe matrice impliquait un changement de structure ('mange' au lieu de 'demande', par exemple).

Le tableau 4.10 indique le score moyen, entre 0 et 1, pour chaque type de phrase, à savoir les items d'entraînement, les items expérimentaux, et les distracteurs.

Type de phrase	Score moyen
items d'entraînement	0.79
items expérimentaux	0.45
distracteurs	0.96

TABLE 4.10 – Scores moyens en fonction du type de phrase (Expérience 1).

On peut voir que les enfants atteignent quasiment une performance maximale pour les distracteurs, qui sont des phrases simples et courtes. Quant aux items expérimentaux, la performance n'est que de 0.45. Ils ont bien mieux répété les items d'entraînement, qui étaient également des phrases complexes et aussi longues que les items expérimentaux. Les enfants présentent donc des difficultés particulières avec les subordinées introduites par 'ce que' en français.

Les données concernant les items expérimentaux ont été analysées avec des modèles linéaires mixtes généralisés, avec la fonction de lien *logit* (BAAYEN, DAVIDSON et BATES, 2008, JAEGER, 2008), en utilisant le paquet *lme4* (BATES, MÄCHLER, BOLKER et WALKER, 2015) dans R (RCORETEAM, 2017).

Le facteur TYPE DE PHRASE a été codé 0.5 pour relative et -0.5 pour interrogative indirecte. De même pour le facteur TYPE DE GAP : 0.5 pour un *gap* objet et -0.5 pour un *gap* sujet. De plus, la covariable ÂGE a été insérée dans le modèle en tant que variable continue centrée. Nous avons en effet inclus tous les participants dans un seul groupe, et vérifié l'effet de l'âge sur leur performance.

Dans la partie 'effets fixes', nous avons estimé les effets principaux, ainsi que toutes les interactions possibles de TYPE DE PHRASE, TYPE DE GAP et ÂGE. Dans la partie 'effets aléatoires', nous avons inclus seulement un intercept pour les sujets et un pour les items, dans le but d'éviter un surparamétrage et, par conséquent, des résultats de modèle qui seraient ininterprétables (BATES, KLIEGL, VASISHTH et BAAYEN, 2015, MATUSCHEK et al., 2017).

La proportion de répétitions correctes à travers les conditions et en fonction de l'âge est illustrée dans la Figure 4.4.

Comme on peut le voir sur la Figure 4.4, il y a une augmentation générale de répétitions correctes au fur et à mesure que les enfants avancent en âge. Cela est reflété par l'effet principal de l'ÂGE, qui est significatif (coef=2.46, SE=0.59, z=4.12, p<.001). Ce qui est intéressant est que, même si cela n'est pas directement visible sur le graphique, la différence entre les deux types de structures est significative. On observe en effet un effet principal significatif

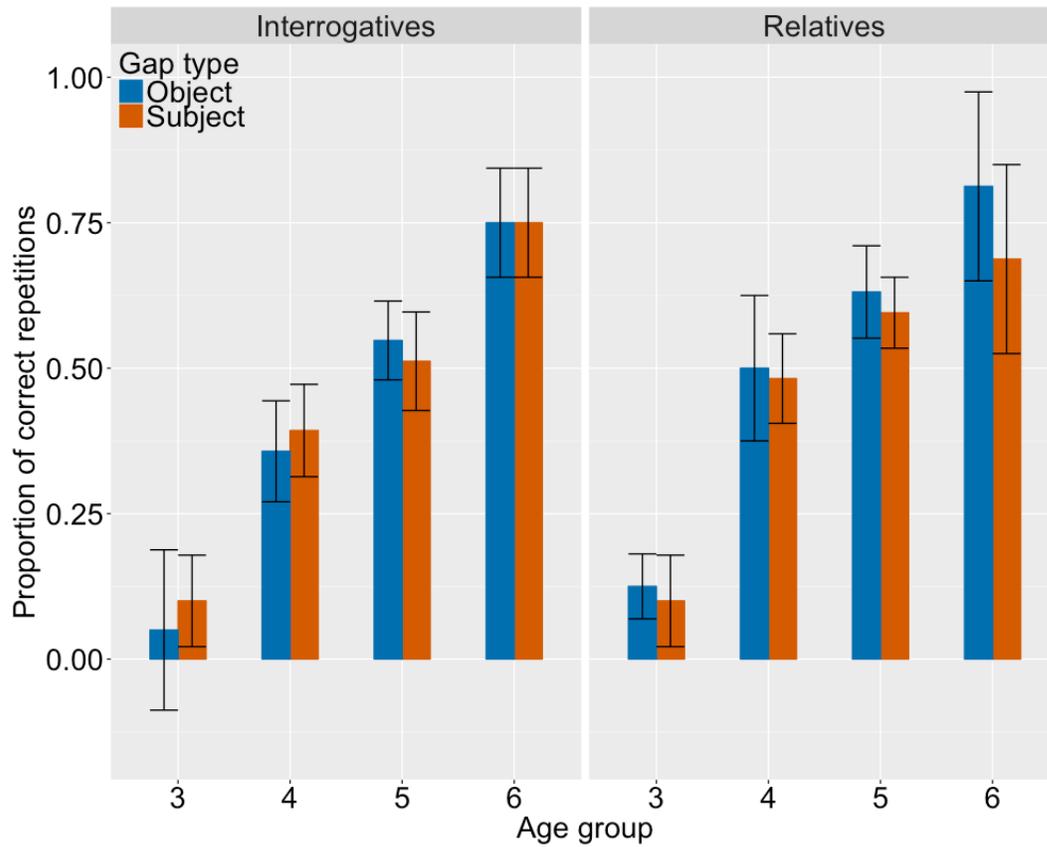


FIGURE 4.4 – Proportion de répétitions correctes (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Le panneau de gauche montre la performance pour les interrogatives indirectes et celui de droite celle pour les relatives à tête allégée. Les barres bleues indiquent la performance pour les dépendances objet, et les oranges celle pour les dépendances sujet.

du TYPE DE PHRASE : les enfants ont mieux répété les relatives que les interrogatives indirectes (coef=0.86, SE=0.32, $z=2.74$, $p=0.006$). Par ailleurs, aucune différence n'est relevée entre la performance des enfants pour les dépendances objet et leur performance pour les dépendances sujet, comme en témoigne l'effet non significatif de TYPE DE GAP (coef=0.15, SE=0.31, $z=0.47$, $p=0.63$). De plus, aucune des interactions dans le modèle n'était significative (toutes les valeurs $p > 0.05$).

4.4.3.2 Analyse qualitative : les types d'erreur

Regardons maintenant les types d'erreurs que les enfants font lorsqu'ils ne répètent pas bien la phrase cible. Nous avons vu que la performance générale est assez faible. Le taux d'erreur est de 55%, ce qui est beaucoup (432 erreurs sur un total de 784 observations). Les erreurs produites par les enfants ont été classées dans plusieurs catégories. Notons que beaucoup de types différents ont été relevés. Ici, je ne vais évoquer que les principaux types d'erreurs (les plus récurrents). Une grande partie des erreurs concerne la production de la séquence 'ce que'.

Un point important est que les enfants ont très bien répété les items d'entraînement (proportion moyenne de répétitions correctes : 0.79), comparés aux items expérimentaux. Cela montre bien que leurs erreurs ne sont pas dues à des compétences cognitives pas encore assez développées (telles que la mémoire ou l'attention), mais que ces erreurs proviennent de réelles difficultés avec la séquence 'ce que'. Car, rappelons-le, les items d'entraînement étaient également des phrases complexes et ils étaient aussi longs que les items expérimentaux.

4.4.3.2.1 L'omission de 'ce' Le type d'erreur le plus proéminent est l'omission de 'ce'. En effet, 43,5% des erreurs (188 sur 432) impliquent l'omission de l'élément D (voir exemple (51)), dont 11,6% (50 sur 432) correspondent à l'omission de D combinée à une ou plusieurs autres erreurs (comme dans l'exemple (52)).

(51) *Phrase cible* : Le maître demande **ce que** la fille a dessiné.

Phrase réponse : Le maître demande **que** la fille a dessiné. [Carla, 5;3]

(52) *Phrase cible* : Le maître demande **ce qui** a roulé par terre.

Phrase réponse : Le monsieur demande **que** il a roulé par terre. [Timéo, 5;1]

En (52), l'enfant a omis 'ce', mais en plus, il emploie 'que' au lieu de 'qui'.

La proportion d'omission de 'ce' en fonction du TYPE DE PHRASE, du TYPE DE GAP et de l'ÂGE est représentée dans la Figure 4.5.⁴ L'omission de 'ce' conduit à un résultat agrammatical dans la grammaire adulte, mais nous avons vu dans le Chapitre 3 que jusqu'au 17^{ème}

4. Dans la Figure 4.5, la proportion d'omission de 'ce' est calculée par rapport à toutes les phrases réponses données, correctes ou fausses, et non pas seulement parmi les erreurs.

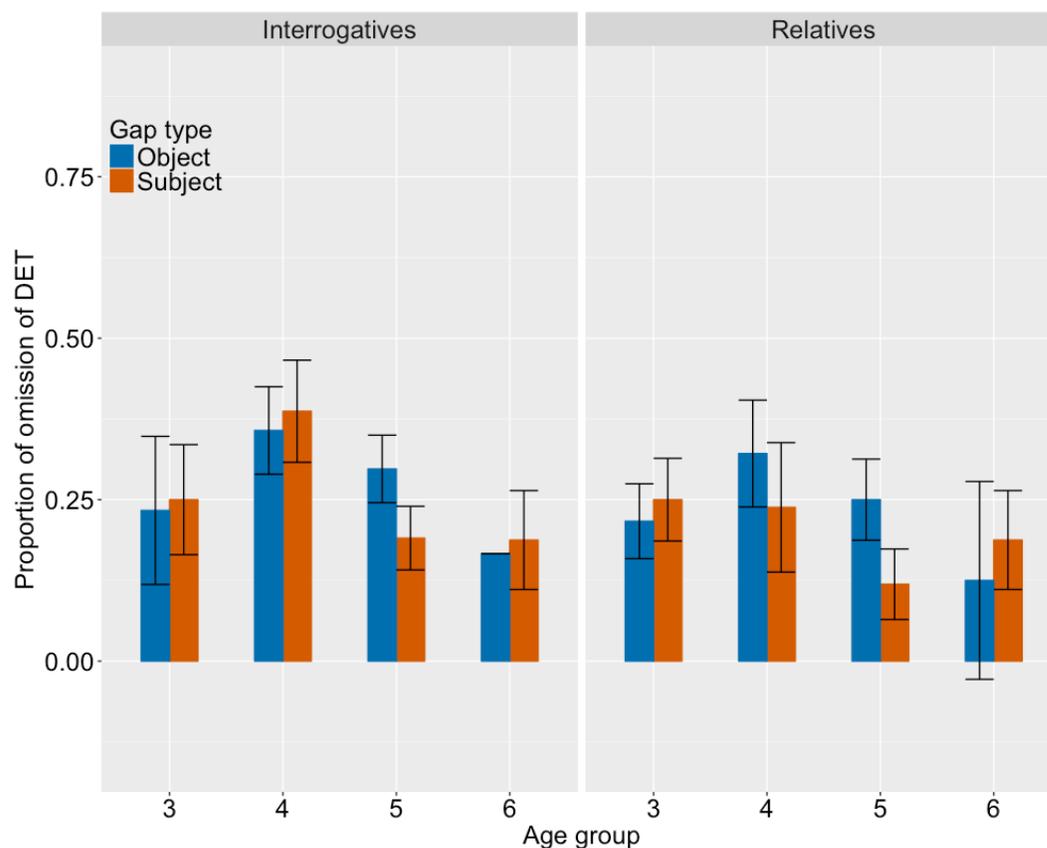


FIGURE 4.5 – Proportion d’omission de ‘ce’ (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Le panneau de gauche montre la proportion d’omissions du démonstratif ‘ce’ dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion d’omissions du démonstratif ‘ce’ dans les dépendances objet, et les oranges celle dans les dépendances sujet.

siècle, 'ce' était absent des interrogatives indirectes portant sur un référent [+ inanimé]. Ce qui est intéressant est que les enfants ont tendance à omettre 'ce' davantage dans les interrogatives indirectes que dans les relatives, bien que cet effet soit marginal (coef=-0.65, SE=0.39, z=-1.67, p=0.096). Aucune différence n'a été relevée entre les dépendances sujet et les dépendances objet, ou bien en relation avec l'âge (toutes les valeurs p > 0.05). Cela signifie que les enfants n'omettent pas plus 'ce' dans les dépendances sujet ou objet d'une part, et, que cette omission n'est pas forcément moins fréquente au fur et à mesure du développement d'autre part.

4.4.3.2.2 L'utilisation d'une forme wh- complexe Le deuxième type d'erreur le plus récurrent, après l'omission de 'ce' est l'utilisation de la forme wh- complexe 'qu'est-ce que' / 'qu'est-ce qui', que je vais noter *kɛsk* en me basant sur la prononciation (voir exemples (53) et (54)). Il correspond à 17,1% des erreurs (74 sur 432). L'omission de 'ce' et l'emploi de *kɛsk* représentent à eux seuls 60,6% des erreurs (262 erreurs sur 432).

- (53) *Phrase cible* : Papa sait pas **ce qui** est tombé du placard.
Phrase réponse : Papa sait pas **qu'est-ce qui** est tombé du placard. [Sarah, 5;6]
- (54) *Phrase cible* : La dame mange **ce que** la maman a cuisiné.
Phrase réponse : La dame mange **qu'est-ce que** maman a cuisiné. [Juliette, 4;10]

La proportion d'utilisation de *kɛsk* en fonction du TYPE DE PHRASE, du TYPE DE GAP et de l'ÂGE est représentée dans la Figure 4.6.⁵ Cette forme wh- est employée en français pour introduire des questions directes, pour interroger sur un référent [+ inanimé]. On utilise 'qu'est-ce que' si le *gap* est en position d'objet et 'qu'est-ce qui' s'il est en position de sujet.

- (55) Qu'est-ce que tu as mangé ?
- (56) Qu'est-ce qui te tracasse en ce moment ?

En français standard, on ne peut pas introduire une interrogative indirecte par *kɛsk*, 'ce qu-' étant la « norme ». Toutefois, dans le français non standard, il est fréquent d'employer cette forme wh- complexe dans les subordinées.⁶

- (57) !Marie m'a demandé qu'est-ce que t'as fait hier.

Notons que dans la très grande majorité des cas, les enfants remplaçaient 'ce que' ou 'ce qui' par le bon équivalent, c'est-à-dire qu'ils ne remplaçaient pas 'ce que' par 'qu'est-ce qui' ou 'ce qui' par 'qu'est-ce que', sauf cas rares.

- (58) *Phrase cible* : Le monsieur attrape **ce qui** a roulé par terre.
Phrase réponse : Le monsieur attrape **qu'est-ce que** a roulé par terre. [Laurena, 5;6]

5. Comme c'était le cas dans la Figure 4.5, la proportion d'utilisation de *kɛsk* est calculée par rapport à toutes les phrases réponses données, correctes ou fausses, et non pas seulement parmi les erreurs, dans la Figure 4.6.

6. Sur les interrogatives en français dans une approche variationniste, voir notamment COVENEY, 1997.

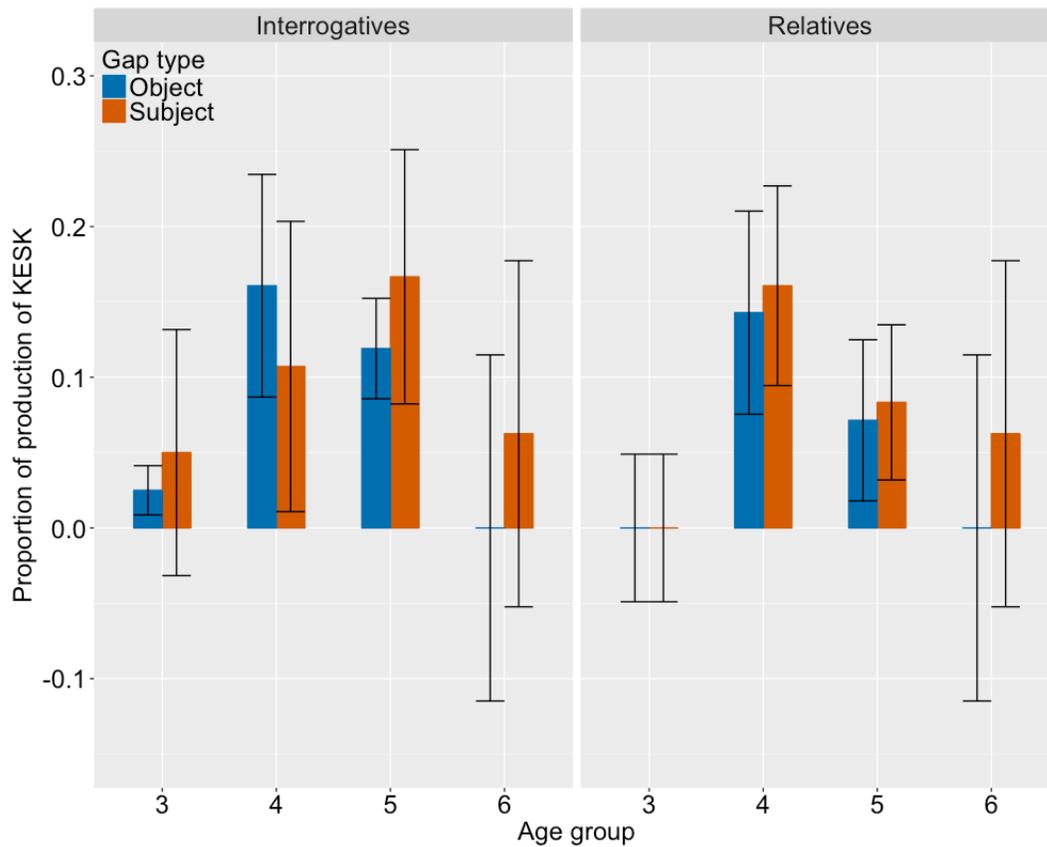


FIGURE 4.6 – Proportion d’utilisation de *kesk* (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Le panneau de gauche montre la proportion d’utilisation de *kesk* dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion d’utilisation de *kesk* dans les dépendances objet, et les oranges celle dans les dépendances sujet.

Comme dans le cas de l'omission de 'ce', les enfants ont tendance à choisir la stratégie *kɛsk* plus souvent dans les interrogatives indirectes que dans les relatives, mais encore une fois, il s'agit d'un effet marginal (coef=-0.64, SE=0.35, z=-1.83, p=0.067). Aucune différence n'a été relevée entre les dépendances sujet et les dépendances objet ou en fonction de l'âge en ce qui concerne la production de *kɛsk* (toutes les valeurs p > 0.05). Les enfants n'utilisent donc pas *kɛsk* davantage dans les dépendances sujet ou bien les dépendances objet, et son emploi ne diminue pas forcément avec l'âge.

4.4.3.2.3 Les autres types d'erreur Les 39,4% d'erreurs restantes (170 sur 432) correspondent à plusieurs types d'erreur différents, chacun assez peu fréquent comparés aux deux premiers types d'erreurs. Parmi eux, citons la production de phrases simples, sans subordonnées, de type SVO. Ce type d'erreur correspond à environ 10% des erreurs (34 sur 432). Ce sont principalement les enfants les plus jeunes (les 3 ans) qui produisent ce genre de phrases sans subordination, comme le montre la Figure 4.7.

- (59) *Phrase cible* : La nounou attrape **ce que** bébé a caché.
Phrase réponse : Nounou a caché le bébé. [Lauryne, 3;2]

Parfois, l'objet est même absent, ce qui donne une simple structure SV, qui est agrammaticale.

- (60) *Phrase cible* : L'élève demande **ce que** la maîtresse a lancé.
Phrase réponse : La maîtresse elle a lancé. [Lorenzo, 3;11]

Il y a donc un effet significatif de l'ÂGE (coeff=-5.77, SE=2.02, z=-2.86, p=0.004) puisque seuls les enfants les plus âgés n'emploient pas cette stratégie de simplification. On note également un effet marginal du TYPE DE GAP, avec une tendance à produire des phrases de type SVO davantage dans les dépendances objet (coeff=-6.53, SE=3.42, z=-1.91, p=0.056). Enfin, on trouve une interaction entre l'ÂGE et le TYPE DE GAP, qui est significative, étant donné que cette stratégie est davantage employée dans les dépendances objet et par les enfants de 3 ans (coeff=-5.75, SE=2.49, z=-2.31, p=0.02).

Un autre type d'erreur marginal concerne les dépendances sujet seulement et correspond à l'emploi de 'ce que' (cf. exemple (61)) ou bien 'ce qu'il' (cf. exemple (62)), que j'analyse comme 'ce que' + un pronom résomptif dans la position de sujet, à la place du *gap*) à la place de 'ce qui'.⁷

- (61) *Phrase cible* : L'enfant devine **ce qui** est sorti de la boîte.
Phrase réponse : L'enfant devine **ce que** est sorti de la boîte. [Johan, 5;2]
- (62) *Phrase cible* : L'enfant devine **ce qui** est sorti de la boîte.
Phrase réponse : L'enfant devine **ce qu'il** est sorti de la boîte. [Louise, 5;1]

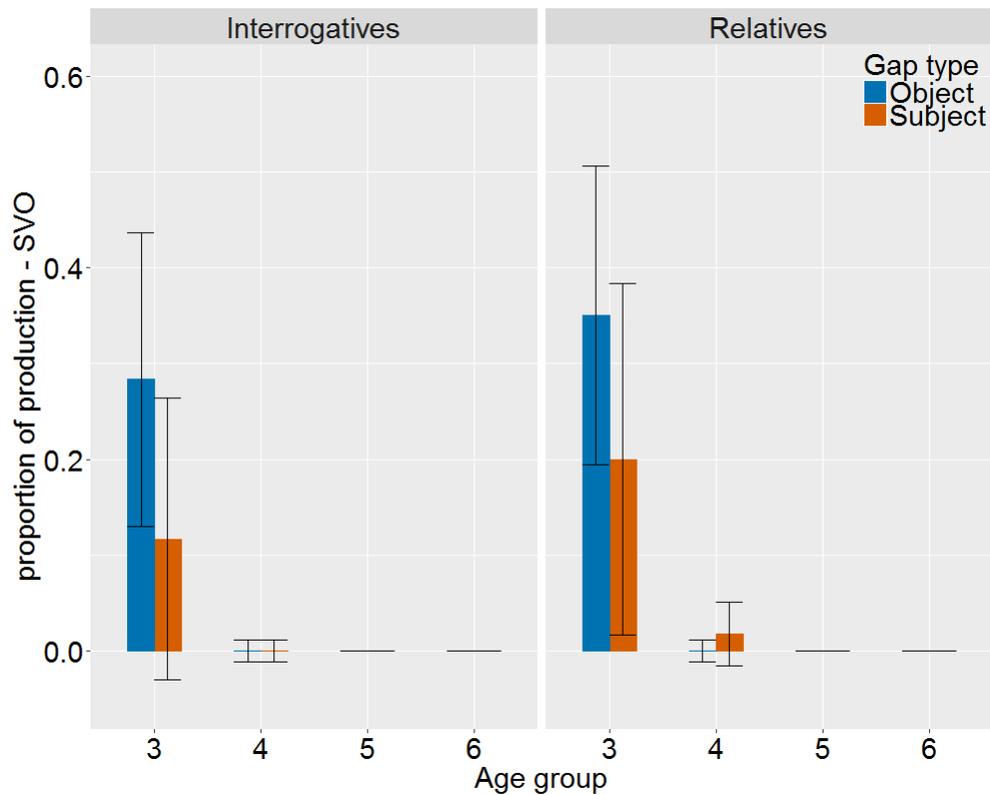


FIGURE 4.7 – Proportion de production de phrases simples de type SVO (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Le panneau de gauche montre la proportion de production de phrases simples de type SVO dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion de production de phrases simples de type SVO dans les dépendances objet, et les oranges celle dans les dépendances sujet.

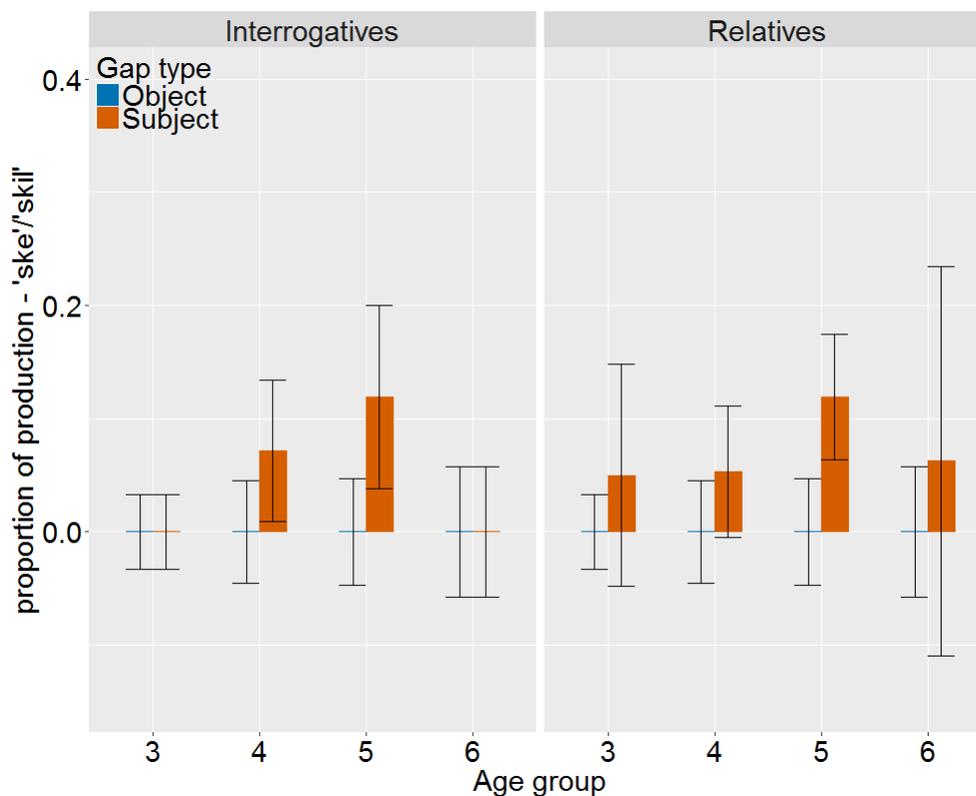


FIGURE 4.8 – Proportion de production de ‘ce que’ ou ‘ce qu’il’ (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Le panneau de gauche montre la proportion de production de ‘ce que’ ou ‘ce qu’il’ dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion de production de ‘ce que’ ou ‘ce qu’il’ dans les dépendances objet, et les oranges celle dans les dépendances sujet.

Ce type d'erreur représente seulement 6,9% (30 sur 432) des erreurs (dont 5,1% pour l'emploi de 'ce que' (22 sur 30) et 1,8% pour l'emploi de 'ce qu'il' (8 sur 30)). Aucun effet n'est significatif (pas d'effet du TYPE DE PHRASE ni du TYPE DE GAP ni de l'ÂGE). La proportion de production de 'ce que' ou 'ce qu'il' parmi toutes les réponses (correctes ou non) est représentée dans le Figure 4.8.

Enfin, j'ai relevé de nombreuses autres types d'erreurs marginaux, parmi lesquels je peux citer l'omission de la séquence D+C (23 sur 432, soit 5,3% des erreurs), le remplacement du verbe interrogatif de la proposition matrice par un verbe sélectionnant un DP, donc une relative (2 sur 432), comme illustré dans l'exemple (63), mais aussi la substitution de l'élément D 'ce' (tête allégée) par un DP plein indéfini, de type D+N, c'est-à-dire une tête pleine (7 sur 432), comme dans l'exemple (64), ou encore le changement du mot subordonnant, où 'ce que' est remplacé par un autre terme introducteur (3 sur 432, voir exemple (65)).

- (63) *Phrase cible* : La nounou **devine** ce que bébé a caché.
Phrase réponse : Nounou **répète** ce que bébé a caché. [Juliette, 4;10]
- (64) *Phrase cible* : Le monsieur attrape ce qui a roulé par terre.
Phrase réponse : Le monsieur attrape **le truc** qui a roulé par terre. [Candice, 3;11]
- (65) *Phrase cible* : Le papa mange **ce qui** est tombé du placard.
Phrase réponse : Le papa mange **quand** c'est tombé de le placard. [Inaya, 3;9]

Le graphique circulaire ci-après (Figure 4.9) permet d'avoir un aperçu général de la répartition des différents types d'erreur en ce qui concerne ce premier test de répétition.

4.4.4 Discussion

Les résultats de cette première expérience montrent que les enfants français ont des difficultés à répéter des phrases contenant des subordonnées introduites par 'ce qu-' : les enfants de 3 ans, bien sûr, mais même les enfants plus âgés. En effet, la performance des enfants de 5 ans est tout juste au-dessus de 0,5. Cette difficulté est cependant modulée par la structure syntaxique : les enfants ont été davantage performants pour la répétition de relatives que pour la répétition d'interrogatives indirectes. La nature de la dépendance, c'est-à-dire si la dépendance implique un *gap* sujet ou bien un *gap* objet, ne semble pas jouer de rôle significatif. Ce sont ces deux résultats que je vais commenter dans les deux sous-sections qui suivent.

4.4.4.1 Des résultats contre l'hypothèse de l'incorporation

La première observation que nous pouvons faire est que les enfants ont davantage de difficultés à répéter les interrogatives indirectes introduites par 'ce qu-' que les relatives

7. Ces deux erreurs (emploi de 'ce que' ou 'ce qu'il' à la place de 'ce qui') laissent penser que certains enfants n'ont pas encore acquis la règle que → qui du français adulte (RIZZI, 1990, KOOPMAN et SPORTICHE, 2008, entre autres). Ils utilisent donc soit un 'que' invariable suivi d'un *gap*, soit un 'que' suivi d'un explétif.

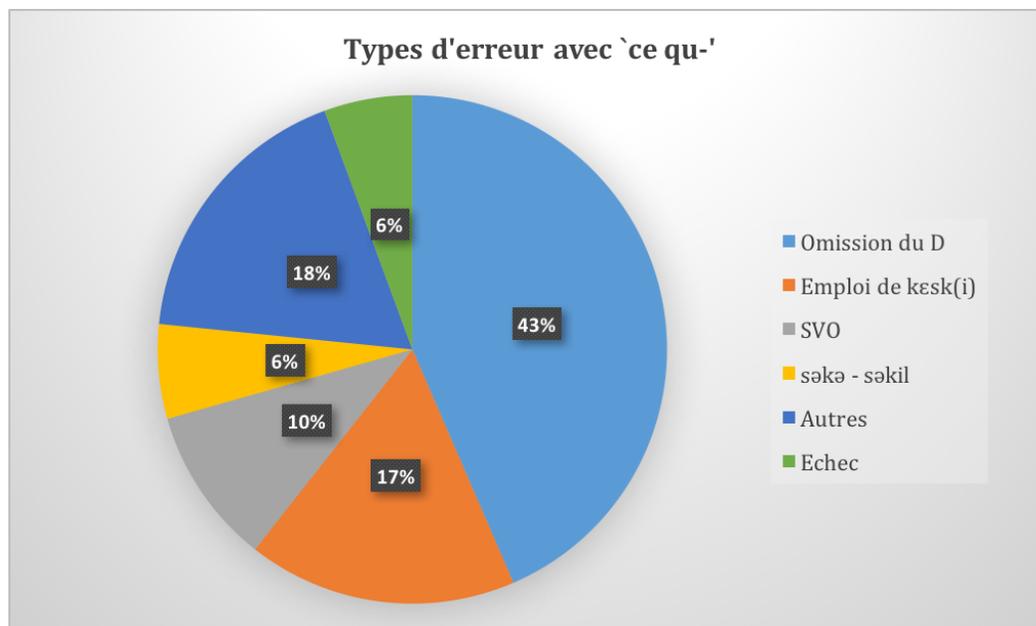


FIGURE 4.9 – Répartition des différents types d'erreurs avec 'ce qu-'

introduites par la même séquence. Cette préférence pour les relatives peut être un indice en faveur de l'analyse de 'ce que' comme D+C. Cela serait d'autant plus évident si les enfants de 3-6 ans n'avaient pas de difficultés particulières avec les interrogatives indirectes introduites par le mot wh- 'qui' (Expérience 2). Cela signifierait que ce ne sont pas les interrogatives indirectes en général qui sont complexes pour eux, mais celles introduites par 'ce que'.

Par ailleurs, l'analyse qualitative (types d'erreur) semble aller dans la même direction. L'omission apparente de l'élément D 'ce', qui est observée davantage dans les interrogatives indirectes que dans les relatives, peut en fait être interprétée comme une erreur de substitution. L'idée est que les enfants remplacent la séquence dérangement 'ce que' par le mot wh- 'que' (nous savons que le complémenteur et le mot wh- sont homophones en français). Rappelons que l'emploi du mot wh- 'que' n'est pas acceptable dans ce contexte dans la grammaire adulte, parce que sa nature de clitique exige une adjacence avec l'élément fléchi qui se trouve dans C (cf. (66)), et l'inversion n'est pas possible dans un contexte enchâssé (cf. (67)).

- (66) a. Que veux-tu ?
 b. *Que tu veux ?
- (67) a. *Je me demande que veux-tu.
 b. *Je me demande que tu veux.

Les enfants sont prêts à enfreindre cette règle (ou à l'ignorer) dans le but d'éliminer cet élément D 'ce', qui leur pose problème dans les interrogatives indirectes. Notons l'existence

d'une connexion frappante entre nos données diachroniques (présentées dans le Chapitre 3) et nos données d'acquisition : de l'ancien français jusqu'au milieu du 17^{ème} siècle, les interrogatives indirectes étaient introduites par un 'que' interrogatif. La grammaire des enfants semble donc correspondre à ce qui a en fait été la règle pendant plusieurs siècles. Nos données viennent renforcer l'idée selon laquelle le processus d'acquisition (ou d'apprentissage) d'une langue est similaire à celui de l'évolution d'une langue. Autrement dit, le développement de certaines formes linguistiques dans la grammaire de l'enfant peut refléter le développement diachronique de ces mêmes formes dans la langue que l'enfant est en train d'acquérir (voir les travaux de BARON, 1977, ROMAINE, 1989, SLOBIN, 2002, SLOBIN, 2004, entre autres).

L'autre stratégie principale que l'on a pu observer chez les enfants est le remplacement de la séquence 'ce qu-' par l'élément wh- complexe 'qu'est-ce qu-'. Encore une fois, le fait que les enfants optent pour cette option au lieu d'utiliser 'ce que', et davantage dans les interrogatives indirectes que dans les relatives, est cohérent avec l'hypothèse qu'ils perçoivent 'ce que' comme étant quelque chose de plus qu'un simple élément wh-, quelque chose qui ne semble pas être interrogatif : une structure DP.

Ainsi, ces résultats pris globalement (meilleur taux de réussite pour les relatives, omission de 'ce' plus fréquente dans les interrogatives indirectes, emploi de la forme wh- 'qu'est-ce qu-' plus fréquente dans les interrogatives indirectes) m'amènent à discréditer l'hypothèse que ce+que a subi un processus de réanalyse en tant qu'élément wh- en français (du moins chez l'enfant), comme cela s'est produit dans plusieurs variétés romanes (en présupposant l'existence d'une continuité entre le système de l'enfant et le système de l'adulte). Une fois cette hypothèse abandonnée reste l'hypothèse que ces interrogatives indirectes introduites par 'ce que' sont ce qu'elles semblent être, c'est-à-dire des relatives. Des arguments sur les plans syntaxique et sémantique en faveur de cette analyse seront discutés dans le Chapitre 7.

4.4.4.2 Absence d'asymétrie sujet-objet

Le deuxième résultat qui attire notre attention est que les dépendances sujet introduites par 'ce qui' ne semblent pas être plus faciles à répéter pour les enfants que les dépendances objet introduites par 'ce que'. Cela peut paraître surprenant au premier abord. En effet, nous savons que les dépendances sujet sont plus faciles à traiter que les dépendances objet.

Cependant, nous avons vu que la difficulté des dépendances objet est modulée par la matrice de traits du sujet et de l'élément déplacé (FRIEDMANN, BELLETTI et RIZZI, 2009 et de nombreuses études ultérieures). Comme nous l'avons vu au début de chapitre, un trait pertinent d'après FRIEDMANN, BELLETTI et RIZZI, 2009 est le trait de restriction lexicale.

Avec cette approche en tête, nous pouvons revenir aux résultats du test de répétition, en particulier l'absence d'asymétrie sujet-objet. Dans tous les cas de dépendances objet, l'objet et le sujet présentent une double différence en traits : l'objet, 'ce', est inanimé alors que le sujet est toujours animé (par exemple : 'la maîtresse'); l'objet est un déterminant nu, donc [-NP], alors que le sujet est spécifié lexicalement ([+NP]). Cette configuration est illustrée en (68).

(68) $\begin{array}{c} ce \\ [-anim, -NP] \end{array} [_{CP} \text{ que } \begin{array}{c} la\text{ fille} \\ [+anim, +NP] \end{array} \text{ a dessiné t}]$

(68) est exactement le genre de configuration dans laquelle les dépendances objet ne sont pas plus difficiles à traiter que les dépendances sujet, selon l'approche de l'intervention du sujet exposée précédemment. L'absence de cette asymétrie dans mon test, qui est finalement loin d'être surprenante, confirme l'importance de la composition en traits du sujet intervenant dans les dépendances objet. Cela va d'ailleurs dans la direction des résultats obtenus par BENTEA, DURRLEMAN et RIZZI, 2016, discutés dans la Section 4.1.1.2.

4.5 Retour sur les données de production spontanée

Revenons un instant sur les données de production spontanée. Nous avons relevé quelques occurrences de 'ce qu-' dans la production de quatre enfants de 2-3 ans jusqu'à 6 ans. Toutefois, nous ne savions pas quelle stratégie ils utilisaient à la place de 'ce qu-'. Nous savons désormais, grâce au test de répétition, que les enfants entre 3 et 6 ans ont des difficultés avec 'ce qu-' et que les deux principaux types d'erreurs qu'ils font sont l'omission (en apparence) de l'élément D 'ce' et la substitution de 'ce qu-' par l'élément wh- complexe 'qu'est-ce qu-'. Il serait donc intéressant de vérifier si l'on retrouve ces deux types d'erreur dans les données de production spontanée.

4.5.1 Le corpus

Le corpus est le même que celui présenté dans la sous-section 4.3.1, à savoir le corpus CoLaJE (MORGENSTERN et PARISSÉ, 2007). Pour cette étude, j'ai cherché les occurrences de 'que' ou 'qui' à la place de 'ce que' ou 'ce qui', ainsi que les occurrences de 'qu'est-ce que' ou 'qu'est-ce qui' à la place de 'ce que' ou 'ce qui' dans les données de production spontanée des quatre mêmes enfants, à savoir Anaé, Théophile, Madeleine et Antoine.

4.5.2 Les résultats à partir de la production des enfants

Il s'avère que nous retrouvons ces deux types d'erreurs dans la production spontanée, mais les occurrences sont très peu nombreuses, comme on peut le voir dans les Tableaux 4.11, 4.12, 4.13 et 4.14.

Tranche d'âge	Omission de 'ce'	Emploi de <i>kεsk</i>
2;06.27	4	0
3;01.07	1	0
3;06.27	2	0
4;00.13	1	0
TOTAL	8	0

TABLE 4.11 – Nombre d'occurrences d'omission de 'ce' et de l'utilisation de 'qu'est-ce que' ou 'qu'est-ce qui' dans la production d'Anaé en fonction de son âge.

Tranche d'âge	Omission de 'ce'	Emploi de <i>kεsk</i>
2;10.01	1	0
3;02.00	1	0
3;10.00	0	3
3;11.15	1	0
TOTAL	3	3

TABLE 4.12 – Nombre d'occurrences d'omission de 'ce' et de l'utilisation de 'qu'est-ce que' ou 'qu'est-ce qui' dans la production de Théophile en fonction de son âge.

Tranche d'âge	Omission de 'ce'	Emploi de <i>kεsk</i>
6;11.27	0	1
TOTAL	0	1

TABLE 4.13 – Nombre d'occurrences d'omission de 'ce' et de l'utilisation de 'qu'est-ce que' ou 'qu'est-ce qui' dans la production de Madeleine en fonction de son âge.

Tranche d'âge	Omission de 'ce'	Emploi de <i>kεsk</i>
3;07.01	1	0
4;00.09	1	0
4;02.19	1	1
4;08.02	1	0
5;11.18	1	0
TOTAL	5	1

TABLE 4.14 – Nombre d'occurrences d'omission de 'ce' et de l'utilisation de 'qu'est-ce que' ou 'qu'est-ce qui' dans la production d'Antoine en fonction de son âge.

4.5.2.1 L'omission de 'ce'

On trouve quelques cas où l'enfant semble clairement omettre 'ce', ou bien, comme nous l'avons suggéré dans la discussion des résultats de l'Expérience 1, substituer la séquence D+C par le mot wh- 'que'.

- (69) c'est que j'a mis [Anaé, 2;06.27]
Phrase cible supposée : c'est ce que j'ai mis
- (70) je sais pas que c'est [Antoine, 4;00.09]
Phrase cible supposée : je sais pas ce que c'est
- (71) ... je fais que j'ai envie de faire [Antoine, 4;08.02]
Phrase cible supposée : je fais ce que j'ai envie de faire
- (72) ... alors regarde que j(e) fais [Antoine, 5;11.18]
Phrase cible supposée : alors regarde ce que je fais

Les occurrences de ce type d'erreur sont trop peu nombreuses pour que nous puissions déterminer si l'omission de l'élément D est plus fréquente dans les relatives ou bien dans les interrogatives indirectes. En tout cas, ces quelques exemples d'omission de 'ce' ne se trouvent pas exclusivement dans des relatives ou bien exclusivement dans des interrogatives indirectes.⁸ On trouve même une occurrence de l'omission de 'ce' à l'intérieur de 'qu'est-ce que'.

- (73) qu'est que c'est que ça? [Théophile, 2;10.28]

4.5.2.2 L'utilisation de *kɛsk*

On trouve également quelques cas où l'enfant utilise la forme wh- complexe 'qu'est-ce que' ou bien 'qu'est-ce qui', à la place de 'ce que' / 'ce qui'.

- (74) regarde qu'est-ce que j'ai dit [Théophile, 3;10.00]
Phrase cible supposée : regarde ce que j'ai dit
- (75) veux voir qu'est-ce que c'est [Théophile, 3;10.00]
Phrase cible supposée : je veux voir ce que c'est

8. Notons, de manière anecdotique, qu'Anaé, à l'âge de 2;10.24, a employé 'que' seul pour introduire une question directe, sans inversion clitique, ce qui est agrammatical dans la grammaire adulte (voir (1)). J'ai également relevé deux questions directes de ce type dans les données de Théophile également (voir (2) et (3)).

- (1) que t(u) as acheté?
- (2) que c'est ça? [2;10.01]
- (3) j'ai dit moi? [4;03.29]

Comme pour l’omission de ‘ce’, l’utilisation de *kɛsk* n’est pas assez fréquente dans les données de production des quatre enfants pour voir s’il y a une différence entre relatives et interrogatives indirectes de ce point de vue là. Mais il est important de noter que ces erreurs sont aussi peu nombreuses que les formes cibles. En effet, on ne peut pas affirmer que les enfants font peu d’erreurs ici, étant donné que nous avons relevé peu d’occurrences des formes cibles (cf. Section 4.3).

4.5.3 Les résultats à partir de l’input

L’omission de ‘ce’ est agrammaticale dans la grammaire adulte. Par contre, on trouve quelques exemples de ‘qu’est-ce que’ (dépendances objet) ou ‘qu’est-ce qui’ (dépendances sujet) dans la production des proches d’Anaé, de Théophile et d’Antoine (voir Tableau 4.15 et exemples (76)-(81)).

	qu’est-ce que	qu’est-ce qui
Input Anaé	7	2
Input Théophile	2	3
Input Madeleine	0	0
Input Antoine	8	0

TABLE 4.15 – Nombre d’occurrences de ‘qu’est-ce que’ et ‘qu’est-ce qui’ dans les propositions subordonnées, dans l’input d’Anaé, de Théophile, de Madeleine et d’Antoine.

- (76) tu dis à Aliyah **qu’est-ce que** tu vas faire pendant les vacances [mère d’Anaé, 4;04.10]
- (77) i(l) veut jardiner (.) alors voyons voir **qu’est-ce qui** s(e) passe [mère d’Anaé, 3;10.00]
- (78) je comprends pas **qu’est-ce que** tu veux? [mère de Théophile, 4;11.11]
- (79) mais pas trop regarde **qu’est-ce qui** t’attend là dans la poêle [mère de Théophile, 3;04.10]
- (80) mystère et boule de gomme c’est quand on n’a pas encore compris **qu’est-ce que** c’est[observateur d’Antoine, 4;05.16]
- (81) je te filme pour savoir **qu’est-ce que** tu vas faire ce soir [observateur d’Antoine, 5;02.13]

4.6 Conclusion intermédiaire

Pour conclure, les données d’acquisition discréditent l’hypothèse de l’incorporation dans le cas de ‘ce que’ en français : les enfants de 3-6 ans n’analysent pas *ce+que* comme un élément *wh-*. Ils sont moins performants en ce qui concerne la répétition d’interrogatives indirectes qu’en ce qui concerne la répétition de relatives introduites par ‘ce qu-’. Ils semblent

donc être conscients qu'un élément *wh-* est requis dans les interrogatives indirectes et ils sont en attente d'un élément de ce type. On peut en déduire qu'ils analysent *ce+que* comme une séquence complexe, de nature nominale, qui est de ce fait davantage acceptée dans un environnement de relative.

Par ailleurs, l'analyse qualitative de nos résultats va également dans cette direction : les deux types d'erreur majeurs, à savoir l'omission de 'ce', que j'analyse comme la substitution de la séquence *ce+que* par le mot *wh-* 'que', ainsi que l'emploi de la forme *wh-* complexe *kɛsk*, tendent à apparaître davantage dans les interrogatives indirectes que dans les relatives. Il est important de préciser que ces deux principales erreurs ne semblent pas être un artefact de la tâche, c'est-à-dire du test de répétition, étant donné qu'on les retrouve dans les données de production spontanée (même si les occurrences sont peu nombreuses).

Ces conclusions que nous tirons concernent le système de l'enfant et nous en déduisons que, dans le système de l'adulte, l'hypothèse de l'incorporation n'est pas valide non plus, dans le cadre d'une continuité entre les deux systèmes. Mais il se pourrait que, en français, durant le processus d'acquisition, on passe d'une phase dans laquelle on analyse *ce+que* comme une séquence complexe, de type D+C, à une phase où on réanalyse cette séquence comme un simple élément *wh-*, *səkə*. Toutefois, cela n'est pas vraisemblable car cela impliquerait que l'on passe d'une construction plus complexe à une construction plus simple. Il serait plus naturel de tendre vers une complexification d'une structure plutôt que vers une simplification dans le passage à la grammaire adulte.

La connexion entre les données diachroniques (voir Chapitre 3) et les données d'acquisition est frappante : 'ce que' est d'abord apparu dans les relatives et les enfants ont moins de difficultés avec 'ce que' dans les relatives ; 'que' seul a été employé pendant des siècles pour introduire des interrogatives indirectes, et les enfants, dans notre test de répétition, ont justement tendance à omettre 'ce' dans les interrogatives indirectes.

Le raisonnement que j'ai fait jusqu'ici m'amène à soutenir que cet avantage significatif des relatives est dû à la présence de l'élément nominal 'ce'. Bien sûr, une explication alternative pourrait être que les relatives sont de toute façon acquises avant les interrogatives indirectes. A ma connaissance, aucune étude n'a montré cela, c'est-à-dire qu'on ne sait pas si, de manière générale, les relatives libres sont acquises avant les interrogatives indirectes ou si, au contraire, les interrogatives indirectes sont acquises avant les relatives libres.

Le deuxième test de répétition que j'ai mis en place, avec pour items expérimentaux des relatives et interrogatives indirectes introduites par 'qui', qui est un mot *wh-* et qui n'est pas précédé par un élément de nature nominale, devrait nous permettre de déterminer cela. Le but est de vérifier si cet avantage des relatives se maintient avec de simples éléments *wh-*. Les résultats de cette deuxième expérience permettront par ailleurs de déterminer

si l'on retrouve une tendance similaire d'acquisition entre 'ce qu-' et 'qui', c'est-à-dire de déterminer si les enfants analysent 'ce qu-' comme un élément wh- ou non.

4.7 Expérience 2 : test de répétition de phrases contenant des subordonnées introduites par *qui*

4.7.1 Prédications

L'Expérience 1 a montré que les enfants français entre 3 et 6 ans ont des difficultés avec 'ce qu-', qui est un morphème complexe. Il est important de faire une comparaison avec des subordonnées introduites par un mot wh- sans élément nominal, comme 'qui', afin de déterminer si 'ce que' se comporte comme un élément wh-.

Dans le cas de la présente expérience, on s'attend à ce que les enfants de la même tranche d'âge soient davantage performants car les subordonnées contenues dans les items expérimentaux sont introduites par un simple élément wh-, à savoir 'qui'. Par ailleurs, en ce qui concerne la différence de performance entre les relatives et les interrogatives indirectes, il ne devrait pas y avoir de différence significative entre les deux structures dans cette deuxième expérience, ou du moins pas d'avantage des relatives. En effet, notre hypothèse est que l'avantage en faveur des structures relatives que nous avons observé dans l'Expérience 1 est dû à la présence de l'élément nominal 'ce'.

4.7.2 Design expérimental

4.7.2.1 Participants

J'ai testé 55 enfants monolingues français, tous présentant un développement typique du langage, et âgés entre 3 et 6 ans. Tous les participants ont été testés à l'Ecole Maternelle de Nogaro, dans le département du Gers. Bien sûr, un accord parental a été obtenu avant que je fasse passer l'expérience aux enfants. Je précise que quelques-uns de ces enfants avaient déjà participé à l'Expérience 1, un an plus tôt, mais la plupart ne s'en souvenaient pas.

4.7.2.2 Matériel

Dans cette deuxième expérience, seul le facteur TYPE DE PHRASE a été retenu, avec les deux mêmes conditions : relatives / interrogatives indirectes. Les items expérimentaux ne sont que des extractions sujet, car nous savons que les relatives libres objet ne sont pas très acceptables en français (voir KONRAD, A paraître). En effet, dans le cadre de mon mémoire de Master, j'avais mis en place un test de jugement d'acceptabilité concernant plusieurs types de relatives, test passé par 93 adultes, locuteurs natifs du français. La note moyenne pour les relatives libres objet n'était que de 3,5 sur une échelle de 1 à 10. Afin de ne pas biaiser les résultats, j'ai créé 16 items, dont 8 contenant des relatives et 8 contenant des interrogatives

indirectes, avec *gap* sujet seulement. J'ai également ajouté 10 *fillers* ou distracteurs. Il y avait également 5 items d'entraînement au début de l'expérience, afin de s'assurer que les enfants, en particulier les plus jeunes, comprenaient bien la tâche à réaliser.

Comme dans l'Expérience 1, les items ont été construits en utilisant un vocabulaire qui fait référence à l'environnement de l'enfant (en particulier, à l'école ou à la maison), afin de s'assurer que les enfants soient familiers avec les mots et les événements décrits dans les phrases à répéter. J'ai contrôlé la longueur des items, en terme de nombre de syllabes (chaque item contenait 12 syllabes), dans le but de neutraliser des effets potentiels liés à la mémoire. Ces items étaient donc aussi longs que les items de l'Expérience 1. Tout a donc été contrôlé, dans la mesure du possible, afin d'éviter tout facteur de différenciation autre que la complexité de l'élément *wh-* introducteur, entre les items de l'Expérience 1 et ceux de l'Expérience 2.

Notons par ailleurs que les relatives libres peuvent avoir une interprétation définie (*la personne qui*) ou bien une interprétation de choix libre (*toute personne qui*). Cette dernière interprétation étant généralement plus répandue, j'ai décidé de mettre le verbe matrice au futur proche et le verbe de la relative au futur simple, afin de favoriser cette lecture, le but étant de rendre les relatives libres, qui restent quand même des structures un peu marginales (du moins en français), les plus naturelles possible. Comparez la relative libre en (82), sans l'usage du futur, et celle en (83), avec l'emploi du futur.

- (82) Le monsieur punit qui a craché par terre.
 ⇒ Le monsieur punit celui/la personne qui a craché par terre.
- (83) Le monsieur va punir qui crachera par terre.
 ⇒ Le monsieur va punir quiconque crachera par terre.

Alors qu'en (82), l'interprétation définie est la seule interprétation possible, en (83), c'est l'interprétation de choix libre qui est mise en avant.

Dans le cas des interrogatives indirectes, le futur n'est pas forcément approprié. J'ai préféré utiliser le passé composé pour le verbe de la subordonnée dans ce cas, et le présent de l'indicatif pour le verbe matrice, comme illustré en (86).

- (84) ??Le maître va demander qui dira un gros mot.
- (85) ??Le maître demande qui dira un gros mot.
- (86) Le maître demande qui a dit un gros mot.

La structure de chaque item était la suivante :

- (87) DP sujet - V matrice - qui - V subordonnée - DP objet / adjoint

Comme on peut le voir en (87), le verbe de la subordonnée n'est pas nécessairement un verbe transitif. C'est pour cela que dans certains cas, le verbe de la subordonnée est suivi d'un adjectif, et dans d'autres cas, il est suivi d'un DP objet. Les items ont été randomisés avec Excel, de façon à ce qu'il y ait autant de listes que de participants. Ainsi, à chaque session de test, l'enfant participant à l'expérience était exposé exactement aux mêmes items que les autres, mais dans un ordre différent.

Voilà deux exemples d'items pour chacune des deux conditions expérimentales :

1. Relatives libres

(88) L'enfant va inviter qui gagnera le jeu.

(89) Tatie va gronder qui écrira sur le livre.

2. Interrogatives indirectes

(90) La nounou ne sait pas qui a gagné le jeu.

(91) Papi se demande qui a écrit sur le livre.

Les *fillers* étaient des phrases courtes, simples, sans proposition subordonnée.

(92) Le jeu est facile et rigolo.

(93) Julie aime bien manger des pâtes.

Quant aux items d'entraînement, comme dans l'Expérience 1, il s'agissait de phrases complexes aussi longues (12 syllabes) que les items expérimentaux. Cependant, ils ne contenaient pas de proposition subordonnée introduite par le mot *wh-* 'qui' : il s'agissait de complétives introduites par le complémenteur 'que'.

(94) Papa pense que maman a fait de bons gâteaux.

4.7.2.3 Procédure

La procédure est identique à celle de l'Expérience 1 (cf. 4.4.2.3).

4.7.3 Résultats

Dans cette sous-section, je vais présenter les résultats de ce deuxième test de répétition, et ce, en deux temps : dans un premier temps, je me centrerai sur l'analyse quantitative (la performance des enfants) ; dans un second temps, je me focaliserai sur l'analyse qualitative (les types d'erreurs que les enfants font lorsqu'ils répètent mal la phrase cible).

4.7.3.1 Analyse quantitative : la performance des enfants

Comme pour l'Expérience 1, la phrase de réponse que les enfants produisaient à chaque essai a été codée comme une variable dépendante binaire : 1 si la répétition était correcte et 0 s'il y avait une ou plusieurs erreurs. Pour déterminer cela, je me suis principalement centrée sur la répétition correcte de 'qui'. Mais si 'qui' a été bien répété, avec un ordre de mots correct, mais que la phrase était agrammaticale pour une raison indépendante, j'ai compté 0.

Le tableau 4.16 indique le score moyen, entre 0 et 1, pour chaque type de phrase, à savoir les items d'entraînement, les items expérimentaux, et les distracteurs.

Type de phrase	Score moyen
items d'entraînement	0.67
items expérimentaux	0.67
distracteurs	0.97

TABLE 4.16 – Scores moyens en fonction du type de phrase (Expérience 2).

Comme c'était le cas pour l'Expérience 1, on peut noter que les enfants atteignent quasiment une performance maximale pour les distracteurs. Par contre, pour cette deuxième expérience, on relève une performance équivalente entre les items d'entraînement et les items expérimentaux (et au-delà de 0.5), ce qui montre que les enfants n'ont pas de difficultés particulières avec les subordonnées introduites par 'qui'.

Comme pour l'Expérience 1, les données ont été analysées avec des modèles linéaires mixtes généralisés, avec la fonction de lien *logit* (BAAYEN, DAVIDSON et BATES, 2008 et JAEGER, 2008), en utilisant le *package lme4* (BATES, MÄCHLER, BOLKER et WALKER, 2015) dans R (RCORETEAM, 2017).

Le facteur TYPE DE PHRASE a été codé 0.5 pour interrogative indirecte et -0.5 pour relative. De plus, la covariable ÂGE a été insérée dans le modèle en tant que variable continue centrée. Nous avons en effet inclus tous les participants dans un seul groupe et vérifié s'il y avait un effet de l'âge sur leur performance. Dans la partie 'effets fixes', nous avons estimé les effets principaux, ainsi que l'interaction éventuelle de TYPE DE PHRASE et ÂGE. Dans la partie 'effets aléatoires', nous avons inclus seulement un intercept pour les sujets et un pour les items, afin d'éviter un suparamétrage, et par conséquent, des résultats de modèle qui seraient ininterprétables (BATES, KLIEGL, VASISHTH et BAAYEN, 2015, MATUSCHEK et al., 2017).

La proportion de répétitions correctes à travers les deux conditions, et en fonction de l'âge, est illustrée dans la Figure 4.10.

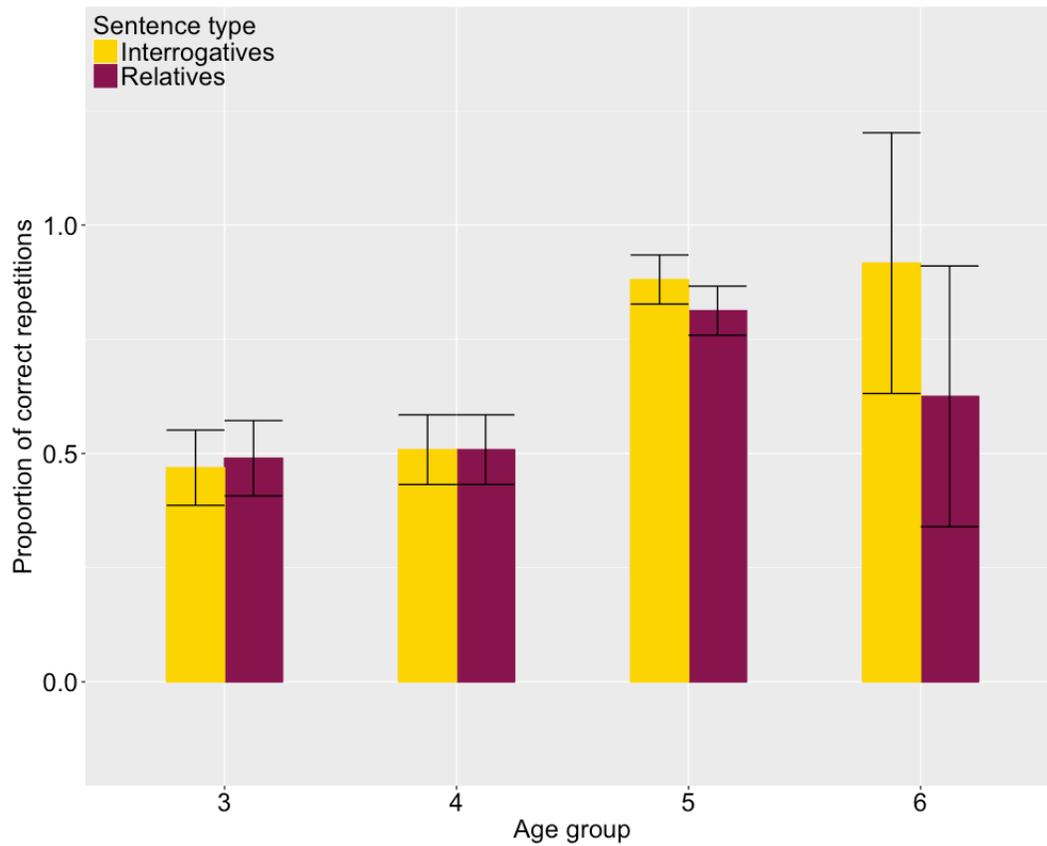


FIGURE 4.10 – Proportion de répétitions correctes (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Les barres jaunes indiquent la performance pour les interrogatives indirectes, et les roses celle pour les relatives libres.

Comme on peut le voir sur la Figure 4.10, il y a une augmentation de répétitions correctes au fur et à mesure que les enfants prennent de l'âge, et en particulier à partir de 5 ans. En effet, la performance des enfants de 4 ans ne semble pas très différente de celle des enfants de 3 ans, et un changement se produit vers 5 ans. Cela est reflété par l'effet principal de l'ÂGE, qui est significatif (coef=1.40, SE=0.43, $z=3.27$, $p=0.001$). Ce qui est frappant est que la performance générale est bien meilleure pour les items de l'Expérience 2 contenant des subordinées introduites par 'qui' que pour les items de l'Expérience 1, contenant des subordinées introduites par 'ce qu-'. Le score moyen global pour l'Expérience 1 était de 0.45 ; ici, il est de 0.67. Ces résultats rejoignent d'autres résultats présentés dans des travaux antérieurs montrant que les enfants ont une meilleure performance avec les relatives libres qu'avec les relatives pleines en français (FLYNN et LUST, 1980, KONRAD, 2015 (non publié)). L'idée que les relatives libres sont plus précoces apparaît également dans des travaux sur le mandarin (PACKARD et al., 1988) et le coréen (LEE, 1991).

Par ailleurs, dans notre expérience, la différence entre les deux types de structures (relatives libres et interrogatives indirectes) est significative puisqu'on relève un effet principal du TYPE DE PHRASE, les enfants répétant mieux les interrogatives indirectes que les relatives libres introduites par 'qui' (coef=0.67, SE=0.28, $z=2.44$, $p=0.015$). Il y a également une interaction significative entre le TYPE DE PHRASE et l'ÂGE, qui est visible sur le graphique (coef=0.87, SE=0.23, $z=3.86$, $p=0.0001$). En effet, cette différence entre les deux types de structures n'apparaît que chez les enfants de 5 ans, et de façon encore plus nette chez les enfants de 6 ans.

4.7.3.2 Analyse qualitative : les erreurs que font les enfants

Regardons maintenant les types d'erreurs que les enfants font dans les cas où ils ne répètent pas correctement la phrase cible. La performance générale étant plus élevée que dans l'Expérience 1, le taux d'erreur est plus bas cette fois-ci. Mais comme pour l'Expérience 1, de nombreux types d'erreurs différents ont été relevés. Je ne retiens toutefois que les principaux types d'erreurs, c'est-à-dire les plus récurrents. Là encore, la majorité des erreurs concerne la production de la séquence-clé, à savoir le mot wh- 'qui'. J'ai relevé en tout 292 erreurs (sur un total de 880 observations).

4.7.3.2.1 Élimination de l'enchâssement Le type d'erreur le plus proéminent est ce que l'on peut appeler une élimination de l'enchâssement. 26% des erreurs (76 sur 292) impliquent des structures sans enchâssement. Au sein de ce type d'erreur, je distingue plusieurs sous-catégories, les quatre principales étant les suivantes :

1. Omission du verbe principal + omission de 'qui'

⇒ cette sous-catégorie représente 57,9% des erreurs de type 'élimination de l'enchâssement' (44 sur 76).

(95) *Phrase cible* : Maman se demande qui a trouvé le trésor.

Phrase réponse : Maman il a trouvé le trésor. [Mathys, 4;6]

- (96) *Phrase cible* : Le monsieur va punir qui crachera par terre.
Phrase réponse : Le monsieur il va cracher par terre. [Eryn, 3;8]

2. Omission du verbe principal seulement

⇒ cette sous-catégorie représente 19,7% des erreurs de type ‘élimination de l’enchâssement’ (15 sur 76).

- (97) *Phrase cible* : La dame demande qui a craché par terre.
Phrase réponse : La maman qui a craché par terre. [Cynthia, 4;10]

- (98) *Phrase cible* : Max va applaudir qui trouvera le trésor.
Phrase réponse : Max qui va trouver le trésor. [Vincent, 3;1]

3. Coordination des deux propositions avec ‘et’

⇒ cette sous-catégorie représente 9,2% des erreurs de type ‘élimination de l’enchâssement’ (7 sur 76).

- (99) *Phrase cible* : Papi se demande qui a écrit sur le livre.
Phrase réponse : Papi a écrit et demande sur le livre. [Eryn, 3;8]

- (100) *Phrase cible* : Papa va applaudir qui fera une roulade.
Phrase réponse : Papa il va applaudir et fait une belle roulade. [Lauryne, 4;1]

4. Omission du verbe de la proposition subordonnée + omission de ‘qui’

⇒ cette sous-catégorie représente 5,3% des erreurs de type ‘élimination de l’enchâssement’ (4 sur 76).

- (101) *Phrase cible* : La maîtresse va punir qui dira un gros mot.
Phrase réponse : La maîtresse va punir le gros mot. [Jeanne, 3;5]

- (102) *Phrase cible* : Max va applaudir qui trouvera le trésor.
Phrase réponse : Max il va applaudir le trésor. [Cynthia, 4;10]

On s’attend à ce que ce soit les enfants les plus jeunes (les 3 et 4 ans) qui fassent ce type d’erreur. Cela est confirmé par la Figure 4.11, et par l’analyse statistique : on trouve un effet principal de l’âge, qui est significatif (coeff=-1.97, SE=0.58, z=-3.39, p=0.0007). La différence entre les interrogatives et les relatives n’est pas significative en ce qui concerne ce type d’erreur (coeff=-0.02, SE=0.58, z=-0.039, p=0.97). De plus, aucune interaction n’est relevée entre l’ÂGE et le TYPE DE PHRASE (coeff=0.59, SE=0.41, z=1.43, p=0.15).

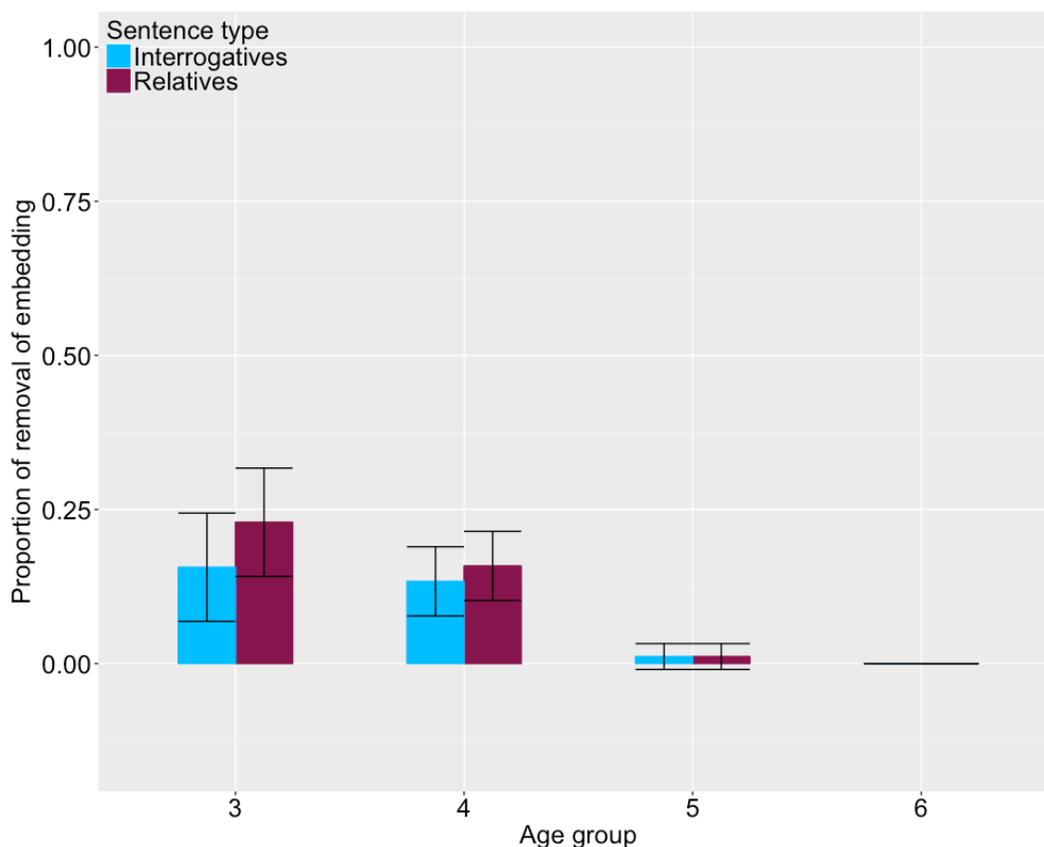


FIGURE 4.11 – Proportion d’éliminations de l’enchâssement (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Les barres bleues indiquent la proportion d’éliminations de l’enchâssement dans les interrogatives indirectes, et les roses celle dans les relatives libres.

4.7.3.2.2 Interrogatives clivées Le deuxième type d’erreur le plus fréquent, après l’élimination de l’enchâssement, est la stratégie qui consiste à faire des interrogatives clivées au lieu d’interrogatives indirectes classiques. Ce type d’erreur représente 17,8% des erreurs (52 sur 292). J’ai relevé plusieurs sous-catégories d’interrogatives clivées, les principales étant les suivantes.

1. **Omission du verbe matrice et mot *wh- in situ*, avec ou sans complémenteur** \Rightarrow cette sous-catégorie représente 38,5% des erreurs de type ‘interrogatives clivées’ (20 sur 52).

(103) *Phrase cible* : Le maître demande qui a dit un gros mot.
Phrase réponse : Le maître c’est qui qui a dit un beau mot. [Vincent, 3;1]

(104) *Phrase cible* : Maman se demande qui a trouvé le trésor.
Phrase réponse : Maman c’est qui a trouvé le trésor. [Lauryne, 4;1]

2. Avec *wh- in situ* et absence du complémenteur : ‘c’est qui’

⇒ cette sous-catégorie représente 19,2% des erreurs de type ‘interrogatives clivées’ (10 sur 52).

(105) *Phrase cible* : Amélie ne sait pas qui a fait une roulade.

Phrase réponse : Amélie elle ne sait pas c’est qui a fait une roulade. [Léona, 4;0]

(106) *Phrase cible* : Maman se demande qui a trouvé le trésor.

Phrase réponse : Maman demande c’est qui trouvé le trésor. [Sylvain, 3;11]

3. Avec *wh- in situ* et présence du complémenteur : ‘c’est qui qui’

⇒ cette sous-catégorie représente 19,2% des erreurs de type ‘interrogatives clivées’ (10 sur 52).

(107) *Phrase cible* : Papi se demande qui a écrit sur le livre.

Phrase réponse : Papi se demande c’est qui qui a écrit sur le livre. [Luna, 5;7]

(108) *Phrase cible* : La dame demande qui a craché par terre.

Phrase réponse : La dame demande c’est qui qui a craché par terre. [Vincent, 3;1]

4. Avec extraction du mot *wh-* et absence du complémenteur : ‘qui c’est’ ⇒ cette sous-catégorie représente 13,5% des erreurs de type ‘interrogatives clivées’ (7 sur 52).

(109) *Phrase cible* : Amélie ne sait pas qui a fait une roulade.

Phrase réponse : Amélie ne sait pas qui c’est a fait une roulade. [Théo, 4;9]

5. Avec extraction du mot *wh-* et présence du complémenteur : ‘qui c’est qui’ ⇒ cette sous-catégorie représente 5,8% des erreurs de type ‘interrogatives clivées’ (3 sur 52).

(110) *Phrase cible* : La dame demande qui a craché par terre.

Phrase réponse : La dame demande qui c’est qui a craché par terre. [Grégoire, 4;10]

Bien sûr, on s’attend à ce que ce type d’erreur (interrogatives clivées) n’apparaisse que dans les interrogatives indirectes, mais cela n’est pas exclusivement le cas, comme on peut le voir sur la Figure 4.12. Voilà deux exemples de stratégie d’interrogatives clivées après un verbe sélectionnant un DP, donc une relative.

(111) *Phrase cible* : Papa va applaudir qui fera une roulade.

Phrase réponse : Papa va applaudir qui c’est va faire une roulade. [Théo, 4;9]

(112) *Phrase cible* : Mamie va gronder qui se cachera dehors.

Phrase réponse : Mamie il va gronder c’est qui se cache dehors. [Sylvain, 3;11]

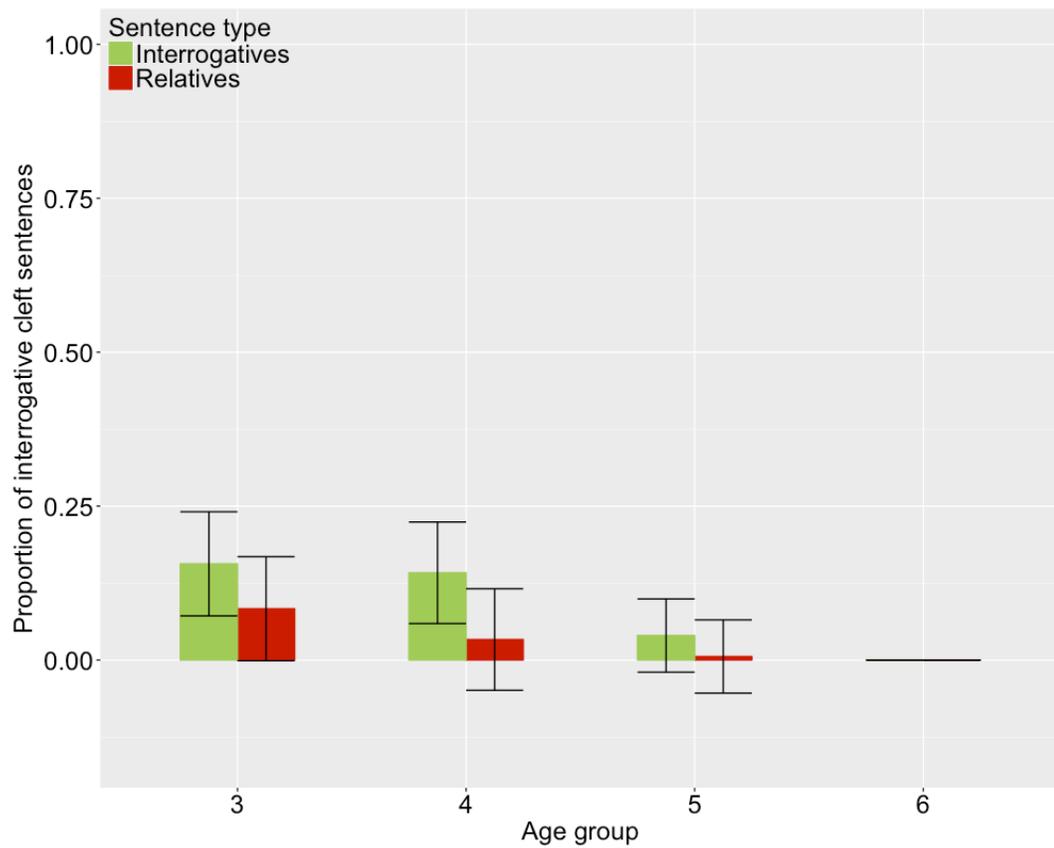


FIGURE 4.12 – Proportion d’interrogatives clivées (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Les barres vertes indiquent la proportion d’interrogatives clivées dans les interrogatives indirectes, et les rouges celle dans les relatives libres.

On observe un effet de l'ÂGE, qui est significatif (coeff=-1.398, SE=0.67, z=-2.087, p=.0369) : cette stratégie syntaxique (interrogatives clivées) est surtout utilisée par les enfants de 3 et 4 ans, cela diminue fortement chez les enfants de 5 ans, jusqu'à totalement disparaître chez les 6 ans. En ce qui concerne le TYPE DE PHRASE, on trouve un effet principal significatif (coeff=2.04, SE=0.55, z=3.71, p=0.0002) : les enfants utilisent ces structures clivées davantage dans les interrogatives indirectes que dans les relatives. Dans les relatives, cet emploi semble se réduire avec l'âge, mais il n'y a pas d'interaction significative entre le TYPE DE PHRASE et l'ÂGE (coeff=0.07, SE=0.47, z=0.15, p=0.88).

4.7.3.2.3 Relativisation explicite Le troisième type d'erreur le plus récurrent, après l'élimination de l'enchâssement et les interrogatives clivées, est ce que j'appelle relativisation explicite. Ce type d'erreur représente 16,8% des erreurs (49 sur 292). En effet, dans les relatives libres, il n'y a pas d'antécédent apparent. Alors une stratégie est d'en ajouter un. Là encore, il existe plusieurs sous-catégories différentes, en fonction de la nature de l'antécédent qui est ajouté (tête légère, tête pleine).

1. Ajout d'une tête pronominale au singulier, de type 'celui'

⇒ cette sous-catégorie représente 40,8% des erreurs de type 'relativisation explicite' (20 sur 49).

(113) *Phrase cible* : Tatie va gronder qui va écrire sur le livre.

Phrase réponse : Tatie va gronder celui qui va écrire sur le livre. [Nina, 5;10]

(114) *Phrase cible* : Le monsieur va punir qui crachera par terre.

Phrase réponse : Le monsieur va punir celui qui va cracher par terre. [Candice, 4;10]

2. Ajout d'une tête pronominale au pluriel, de type 'ceux'⁹

⇒ cette sous-catégorie représente 34,7% des erreurs de type 'relativisation explicite' (17 sur 49).

(115) *Phrase cible* : L'enfant va inviter qui gagnera le jeu.

Phrase réponse : L'enfant va inviter ceux qui gagnent le jeu. [Jeanne, 3;5]

(116) *Phrase cible* : Papa va applaudir qui fera une roulade.

Phrase réponse : Papa va applaudir ceux qui fait une roulade. [Oihana, 6;0]

9. La répétition étant orale par nature, il semble difficile de déterminer si les enfants ajoutent une tête pronominale inanimée, au singulier, de type 'ce', ou bien une tête pronominale au pluriel de type 'ceux', les deux étant homophones. Mais le fait même que les enfants ajoutent une tête pronominale indique qu'ils comprennent le sens de cette relative libre. De plus, ils sont conscients de la nature inanimée de 'ce' puisqu'ils le remplaçaient parfois par 'la chose' ou 'quelque chose' dans l'Expérience 1. J'en déduis qu'ici ils utilisent la tête pronominale au pluriel 'ceux' pour éviter la relative libre. Ils savent que l'agent est humain pour des prédicats tels que 'cracher par terre', 'faire une roulade', 'gagner le jeu'. L'emploi de 'ce' me paraît donc peu probable dans les exemples (115)-(118).

Comme le montre l'exemple (116), l'emploi de la tête légère au pluriel 'ceux' n'engendre pas forcément un accord adéquat en nombre avec le verbe de la subordonnée.

3. **Ajout d'une tête pronominale au pluriel de type 'ceux' + omission de 'qui'**

⇒ cette sous-catégorie représente 10,2% des erreurs de type 'relativisation explicite', et n'apparaît que chez un seul enfant (5 sur 49).

(117) *Phrase cible* : Le monsieur va punir qui crachera par terre.

Phrase réponse : Le monsieur (v)a punir ceux a craché par terre. [Léona, 4;0]

(118) *Phrase cible* : Tom va inviter qui portera des bonbons.

Phrase réponse : Tom a invité ceux a porté des bonbons. [Léona, 4;0]

4. **Ajout d'une tête pleine**

⇒ cette sous-catégorie représente 6,1% des erreurs de type 'relativisation explicite' (3 sur 49).

(119) *Phrase cible* : L'enfant va inviter qui gagnera le jeu.

Phrase réponse : L'enfant va inviter l'enfant qui gagnera le jeu. [Yaël, 4;7]

Bien sûr, on s'attend à ce que ce type d'erreur (relativisation explicite) se trouve exclusivement dans les relatives. Globalement cela est vrai mais il y a deux ou trois cas isolés, comme on peut le voir sur la Figure 4.13.

(120) *Phrase cible* : La dame demande qui a craché par terre.

Phrase réponse : La dame demande celui qui a craché par terre. [Julia, 6;0]

L'*output* du modèle linéaire mixte généralisé indique un effet principal du TYPE DE PHRASE (coeff=-3.90, SE=0.79, z=-4.94, p<.001), mais pas d'effet principal de l'ÂGE (coeff=0.73, SE=0.73, z=1.001, p=0.317), même si, d'après le graphique, les enfants de 6 ans semblent utiliser cette stratégie un peu plus que les enfants des tranches d'âge inférieures. Il est évident que l'on ne relève aucune interaction entre le TYPE DE PHRASE et l'ÂGE (coeff=-0.56, SE=0.68, z=-0.82, p=0.41).

4.7.3.2.4 Autres subordonnées Un autre type d'erreur relevé est la production de subordonnées d'autres types. En effet, les enfants ont parfois transformé les subordonnées introduites par 'qui', qui sont des compléments du verbe, en d'autres subordonnées, telles que des subordonnées conditionnelles ou temporelles. Ce type d'erreur représente 10,6% des erreurs (31 sur 292). Je distingue principalement deux grandes sous-catégories, en fonction du terme introducteur :

1. **Si (quelqu'un)** ⇒ cette sous-catégorie représente 51,6% des erreurs de type 'autres subordonnées' (16 sur 31).

(121) *Phrase cible* : Tatie va gronder qui écrira sur le livre.

Phrase réponse : Tatie a grondé si a écrit sur le livre. [Léona, 4;0]

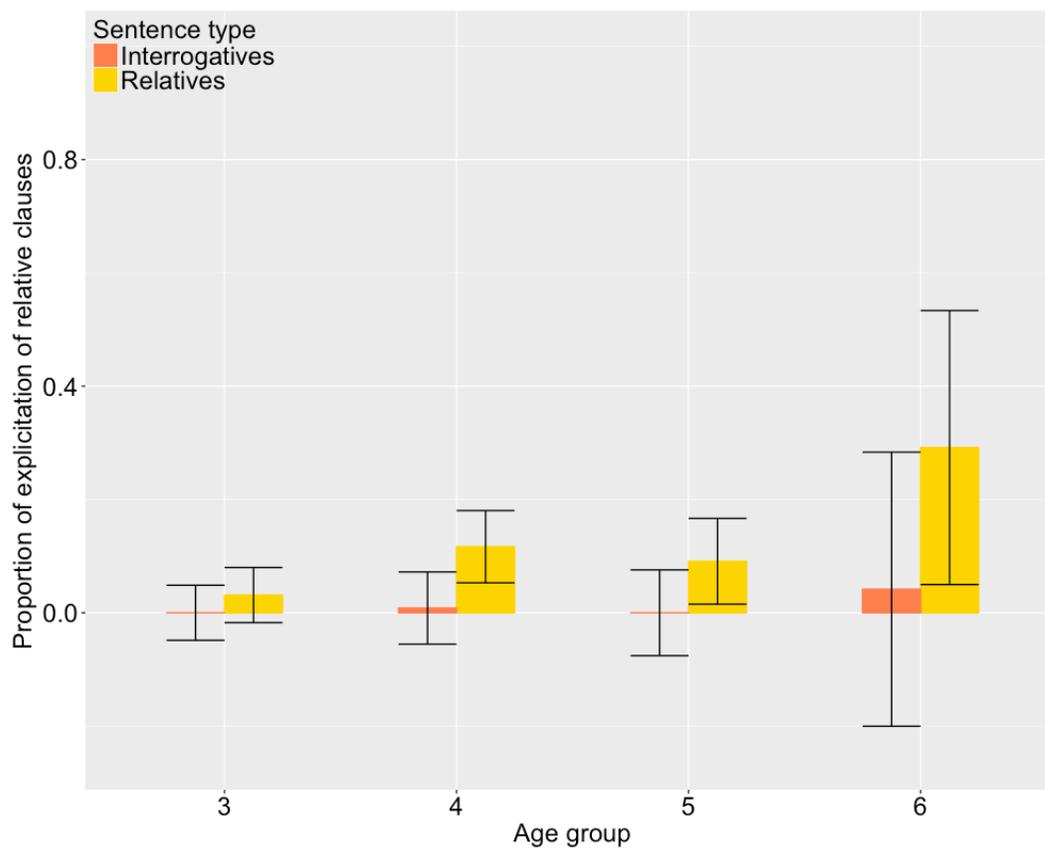


FIGURE 4.13 – Proportion de relativisation explicite (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Les barres oranges indiquent la proportion de relativisation explicite dans les interrogatives indirectes, et les jaunes celle dans les relatives libres.

(122) *Phrase cible* : La nounou ne sait pas qui a gagné le jeu.
Phrase réponse : La nounou il sait pas si quelqu'un a gagné le jeu. [Timéo, 6;0]

(123) *Phrase cible* : Le monsieur va punir qui crachera par terre.
Phrase réponse : Le monsieur il va punir si il crache par terre. [Lauryne, 4;1]

2. **Quand (quelqu'un)** ⇒ cette sous-catégorie représente 38,7% des erreurs de type 'autres subordonnées' (12 sur 31).

(124) *Phrase cible* : Mélanie devine qui a porté des bonbons.
Phrase réponse : La mamie devine quand elle a porté des bonbons. [Jeanne, 3;5]

(125) *Phrase cible* : La maîtresse va punir qui dira un gros mot.
Phrase réponse : La maîtresse va punir quand quelqu'un dira un gros mot. [Alexis, 5;5]

(126) *Phrase cible* : Max va applaudir qui trouvera le trésor.
Phrase réponse : Max va applaudir quand il va trouver le trésor. [Julianne, 3;8]

L'effet de l'ÂGE n'est que marginal (coeff=-1.02, SE=0.56, z=-1.82, p=0.0695) : on peut voir que cette stratégie est employée par les enfants de tout âge, mais les enfants de 5 ans un peu moins. Par ailleurs, les enfants ont tendance à transformer la subordonnée complément introduite par 'qui' en d'autres subordonnées davantage dans les relatives libres que dans les interrogatives indirectes (comme on le voit sur la Figure 4.14). L'effet du TYPE DE PHRASE est significatif (coeff=-2.02, SE=0.69, z=-2.94, p=0.003). Aucune interaction n'est relevée entre l'ÂGE et le TYPE DE PHRASE (coeff=-0.79, SE=0.52, z=-1.54, p=0.12).

4.7.3.2.5 **Omission de 'qui'** Enfin, le cinquième et dernier type d'erreur majeur est l'omission du mot wh- 'qui'.

(127) *Phrase cible* : La nounou ne sait pas qui va gagner le jeu.
Phrase réponse : La nounou elle sait pas elle va gagner le jeu. [Mathys, 4;6]

(128) *Phrase cible* : Tom va inviter qui portera des bonbons.
Phrase réponse : Tom va inviter portera des bonbons. [Eléa, 5;4]

(129) *Phrase cible* : La dame demande qui a craché par terre.
Phrase réponse : La dame il dit il a craché par terre. [Mathys, 4;6]

Ce type d'erreur représente 9,9% des erreurs (29 sur 292). Dans la très grande majorité des cas, cette omission n'est pas combinée à d'autres erreurs (93,1% des cas). Mais on trouve deux cas isolés d'omission de 'qui' combinées avec l'omission du verbe de la proposition matrice (6,9% des cas).

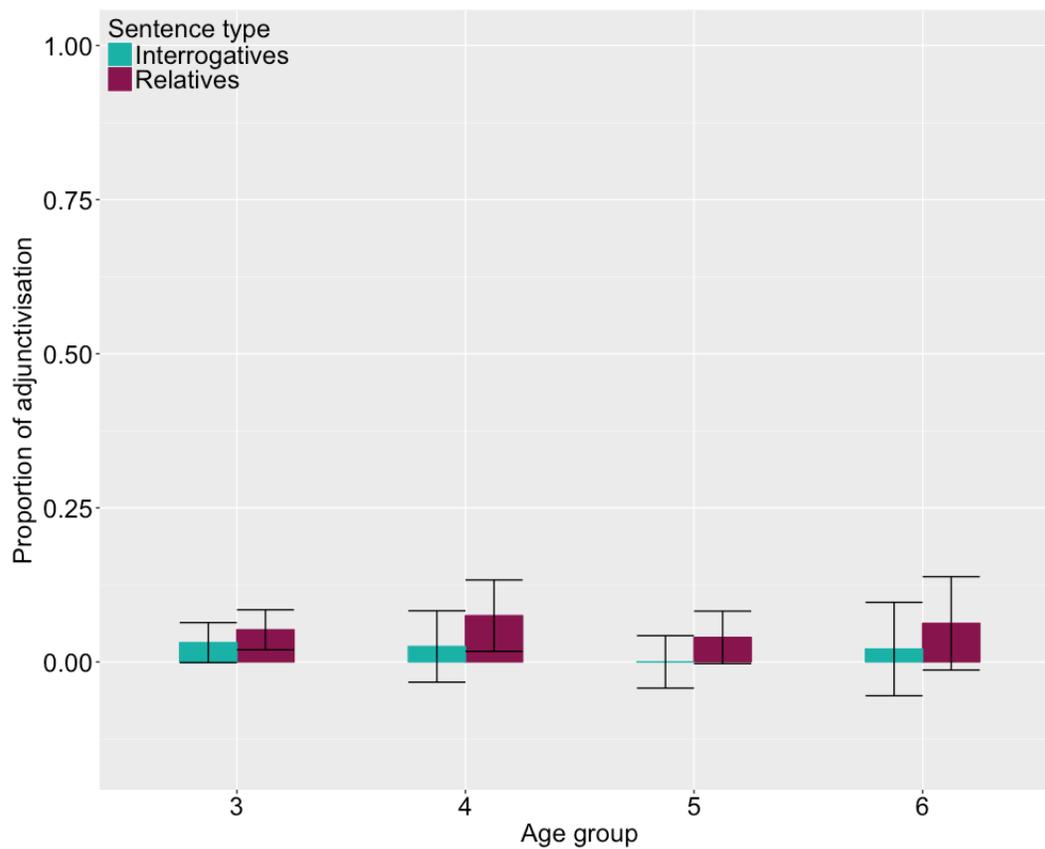


FIGURE 4.14 – Proportion de production d'autres subordonnées (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Les barres vertes indiquent la proportion de production d'autres subordonnées dans les interrogatives indirectes, et les roses celle dans les relatives libres.

Comme pour le type d'erreur 'élimination d'enchâssement', on s'attend à ce que ce soit les enfants les plus jeunes (3-4 ans) qui omettent le mot wh. Regardons la répartition de ces omissions en fonction des l'âge et du type de phrase (Figure 4.15).

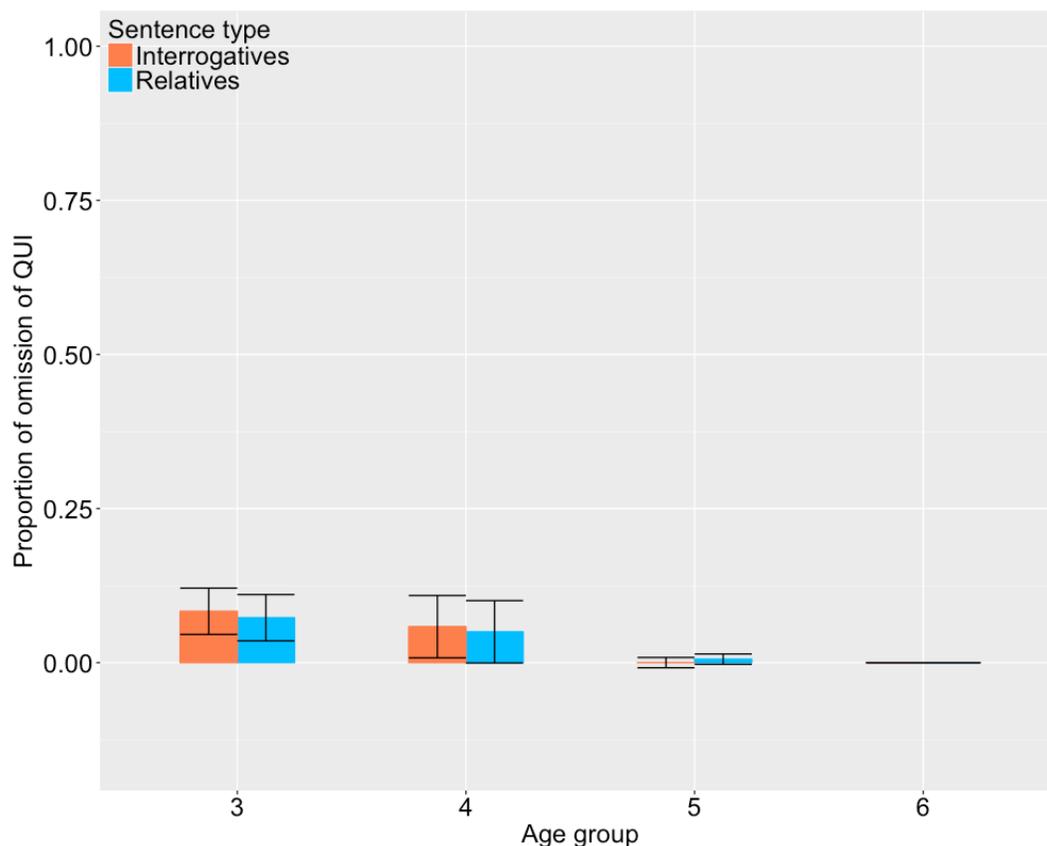


FIGURE 4.15 – Proportion d'omission du mot wh- 'qui' (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Les barres bleues indiquent la proportion d'omission du mot-wh 'qui' dans les interrogatives indirectes, et les oranges celle dans les relatives libres.

Comme on peut le voir sur la Figure 4.15, et confirmé par l'analyse statistique, on note un effet principal de l'ÂGE, qui est significatif (coeff=-1.69, SE=0.75, z=-2.25, p=0.025). On observe également que cette omission n'est pas plus fréquente dans un type de phrase que dans l'autre, d'où l'absence d'effet significatif du TYPE DE PHRASE (coeff=-0.54, SE=0.84, z=-0.65, p=0.52). De plus, aucune interaction n'a été relevée entre l'ÂGE et le TYPE DE PHRASE (coeff=-0.65, SE=0.70, z=-0.93, p=0.35).

4.7.3.2.6 Les autres types d'erreurs Les 5 types d'erreurs principaux que je viens de décrire représentent à eux seuls 81,1% des erreurs. Parmi les 18,9% d'erreurs restantes (soit

55 erreurs sur les 292), on trouve plusieurs types d'erreurs marginaux (pas plus de 10 occurrences de chacun de ces types d'erreurs marginaux), mais aussi des cas isolés. Je peux citer par exemple :

1. **Changement du verbe de la proposition matrice (sans aucune autre erreur combinée)**

- (130) *Phrase cible* : La dame demande qui a craché par terre.
Phrase réponse : La dame dit qui a craché par terre. [Jules, 5;4]
- (131) *Phrase cible* : Maman se demande qui a trouvé le trésor.
Phrase réponse : Maman se trouve qui a trouvé le trésor. [Inaya, 4;8]

2. **Réduplication du 'qui'**

- (132) *Phrase cible* : Amélie ne sait pas qui a fait une roulade.
Phrase réponse : Amélie **qui** sait pas **qui** a fait une roulade. [Emma, 5;7]
- (133) *Phrase cible* : Le monsieur va punir qui crachera par terre.
Phrase réponse : Le monsieur **qui** va punir **qui** crachera par terre. [Emerance, 5;3]

4.7.3.2.7 **Bilan** Le graphique circulaire ci-après (Figure 4.16) permet d'avoir un aperçu général de la répartition des différents types d'erreur en ce qui concerne ce deuxième test de répétition.

4.7.4 Discussion

Les résultats de cette deuxième expérience montrent que les enfants français, entre 3 et 6 ans, ont beaucoup moins de difficultés à répéter des phrases contenant des subordinées introduites par 'qui' que par 'ce qu-'. En ce qui concerne la différence entre les deux types de phrase, on ne trouve aucun avantage pour les relatives cette fois-ci. Les deux structures semblent acquises de façon parallèle, comme prédit, du moins chez les enfants de 3 et 4 ans.

Ce qui est intéressant est que, chez les enfants de 6 ans, le *pattern* est inversé : ce sont les relatives libres qui sont plus difficiles que les interrogatives indirectes. On peut expliquer cela par le fait qu'ils se rapprochent de plus en plus de la grammaire adulte, en quelque sorte, car, comme nous l'avons mentionné précédemment, les relatives libres en général ne sont pas très naturelles en français. Cela est d'ailleurs mis en avant par le fait que l'un des types d'erreur majeurs soit l'explicitation de relatives, grâce à l'ajout d'une tête légère de type 'celui' dans les relatives, stratégie qui est principalement utilisée par les enfants de 6 ans. LEE, 1991, se basant sur l'acquisition des relatives en coréen, avait aussi trouvé que les enfants ajoutent parfois des têtes aux relatives libres.

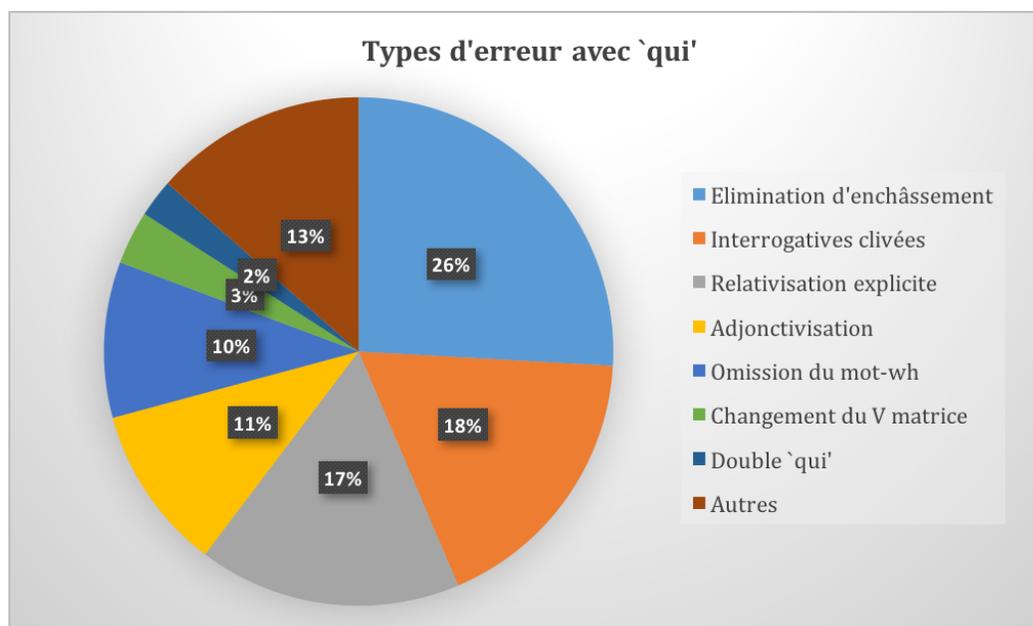


FIGURE 4.16 – Répartition des différents types d'erreurs avec 'qui'

Dans les deux tests de répétition, nous avons remarqué que les enfants les plus jeunes (3 ans) avaient tendance à éliminer l'enchâssement (10% des erreurs dans l'Expérience 1 et 26% des erreurs dans l'Expérience 2). Par ailleurs, l'omission du mot (ou de la séquence) subordonnant(e) se retrouve aussi dans les deux tests : l'omission de 'ce qu-' correspondait à 5% des erreurs dans l'Expérience 1, et l'omission de 'qui' correspondait à 10% des erreurs dans l'Expérience 2. Toutefois, la proportion d'omission de 'ce qu-' est deux fois moins importante que la proportion d'omission de 'qui'. En effet, nous avons vu dans la première expérience, que les enfants omettaient très souvent 'ce' plutôt que la séquence toute entière 'ce qu-'. Si l'enfant analysait 'ce qu-' comme un tout, comme un mot wh- similaire à 'qui', alors le pourcentage d'omission de 'ce qu-' aurait été bien plus important, et celui d'omission de 'ce' bien moindre.

Quant aux formes complexes wh-, impliquant la séquence interrogative 'est-ce qu-' (c'est-à-dire 'qu'est-ce que/qu'est-ce qui' dans l'Expérience 1 et 'qui est-ce qui' dans l'Expérience 2 à la place respectivement de 'ce que/ce qui' et de 'qui'), on les retrouve dans les deux expériences (17,1% des erreurs pour l'Expérience 1, 17,8% des erreurs pour l'Expérience 2). Il est intéressant de noter que les enfants semblent analyser différemment 'qu'est-ce qu-' et 'qui est-ce qui'. Dans la deuxième expérience, les enfants n'ont pas produit une seule fois la séquence canonique 'qui est-ce qui', normalement employée dans les questions racines : soit ils n'inversaient pas le clitique 'ce' ('qui c'est qui'), soit ils le plaçaient au tout début de la séquence, créant ainsi des interrogatives clivées ('c'est qui qui'). Par contre, dans la première expérience, à aucun moment les enfants n'ont inversé le clitique ('que c'est que')

ou placé le clitique et la copule en début de séquence ('c'est que que'). Cela semble indiquer que les enfants analysent 'qu'est-ce qu-' comme un tout grammaticalisé, contrairement à 'qui est-ce qui'.

Les enfants de 3 et 4 ans semblent avoir des difficultés avec l'enchâssement de propositions, c'est pourquoi ils tentent d'éliminer le mot *wh-* et/ou le verbe de la proposition ou le verbe de la proposition subordonnée. Cela était déjà le cas avec 'ce qu-' (stratégie appelée SVO) mais la différence de performance est très grande. En effet, dans le cas de ce deuxième test de répétition, la proportion de répétitions correctes pour les enfants de 3 ans est proche de 0.5 alors qu'elle était bien en-dessous de 0.25 pour le premier test de répétition.

Globalement, ces résultats semblent confirmer que, dans le cas de la première expérience, l'avantage pour les relatives était dû à la présence de l'élément nominal 'ce', étant donné que cet avantage disparaît si le terme introducteur est un mot *wh-* simple, comme 'qui'.

4.8 Conclusion du Chapitre

Nous avons vu que les enfants français entre 3 et 6 ans avaient beaucoup de difficultés avec 'ce qu-', et moins avec des subordonnées introduites par 'qui'. Les résultats de l'Expérience 1 semblent invalider l'hypothèse de l'incorporation de *ce+que*, et présentent de fortes similitudes avec l'évolution de 'ce que' (voir étude diachronique présentée dans le Chapitre 3). Par ailleurs, l'avantage des relatives n'est présent que dans le cas de 'ce qu-'. On trouve même un *pattern* inverse avec 'qui', les interrogatives indirectes étant globalement mieux répétées que les relatives libres.

Un point important à souligner est que cette différence entre les deux types de structures est indépendant de la position du *gap* au sein de la proposition subordonnée. Friedmann et al. (2009), entre autres, ont montré que la nature de la tête (en termes de traits) pouvait moduler la difficulté des relatives objet. Les deux expériences que j'ai mises en place semblent indiquer que la nature de la tête joue un rôle de manière générale, c'est-à-dire même dans les relatives sujet. En effet, que l'introducteur soit de nature nominale (*ce qu-*) ou de nature *wh-* (*qui*) a une influence sur la performance des enfants. Ces derniers sont donc sensibles à la matrice de traits de l'élément qui introduit une relative ou une interrogative indirecte.

Il serait intéressant d'une part de faire une expérience similaire à l'Expérience 1 avec 'o que', étant donné qu'en portugais ce morphème complexe a sûrement subi un processus de grammaticalisation. Cela va être développé dans le chapitre qui suit, à savoir le Chapitre 5.

D'autre part, on peut se demander ce qu'il se passe sur le plan du traitement chez les adultes. Autrement dit, est-ce que les adultes préfèrent également 'ce qu-' dans les relatives plutôt que dans les interrogatives indirectes ? Et est-ce qu'ils emploient 'qu'est-ce que' à la

place de 'ce que' dans les subordonnées en question? Je répondrai à ces questions, à partir d'une expérience de lecture par présentation auto-segmentée (*self-paced reading task*) et d'une étude de corpus dans le Chapitre 6.

Chapitre 5

Du français au portugais : l'acquisition de 'o que'

Comme nous l'avons vu dans le Chapitre 2 (Section 2.4), 'o que' en portugais semble avoir subi une grammaticalisation en élément interrogatif, du moins dans le cas des interrogatives indirectes, contrairement au français 'ce qu-'. Le but du présent chapitre est de vérifier si ces différences syntaxiques observées entre 'ce qu-' et 'o que' se reflètent dans l'acquisition de ces morphèmes dans les deux structures subordonnées qui nous intéressent, à savoir les relatives et les interrogatives indirectes. Nous venons de voir dans le Chapitre 4 que les données de l'acquisition ne vont pas dans le sens d'une incorporation de l'élément 'ce' à 'que' dans le cas du français. Cela était attendu au vu du comportement syntaxique de 'ce qu-'. Qu'en est-il de l'acquisition de 'o que' en portugais ? En particulier, on peut se poser les questions suivantes :

1. Est-ce que 'o que' est acquis de façon parallèle dans les deux structures (relatives et interrogatives indirectes) ou pas ?
2. Est-ce que le *pattern* d'acquisition de 'o que' est différent du *pattern* d'acquisition de 'ce qu-' ?
3. Est-ce que les enfants portugais analysent 'o que' comme un tout, un mot wh- comme 'what' en anglais ?
4. Les enfants portugais ont-ils des difficultés avec 'o que' ? Si oui, quelle(s) stratégie(s) syntaxique(s) utilisent-ils à la place de 'o que' ?

Comme dans le cas du français, j'ai d'abord étudié des données de production spontanée afin d'avoir un premier aperçu général sur l'acquisition de ces structures. Dans un second temps, j'ai mis en place un test de répétition de phrases similaire à celui pour le français, afin de pouvoir répondre aux questions que je viens de soulever.

5.1 L'acquisition des relatives à tête allégée et des interrogatives indirectes en portugais

Comme dans le cas du français, plusieurs études se sont centrées sur l'acquisition des questions *wh-* et/ou des relatives en portugais européen (SOARES, 2003, SOARES, 2004, SOARES, 2006, BAIÃO et LOBO, 2014, COSTA, FRIEDMANN, SILVA et YACHINI, 2014, entre autres), mais aucune d'entre elles n'a étudié l'acquisition du double usage de 'o que' en portugais. Les études sur les dépendances A-barre se focalisent souvent sur l'asymétrie entre dépendances sujet et dépendances objet (par exemple CEREJEIRA, 2010 dans le cas des questions *wh-* et COSTA, LOBO et SILVA, 2011 dans le cas des relatives).

Par ailleurs, la très grande majorité de ces travaux portent sur les relatives pleines, rarement sur les relatives à tête légère ou les relatives libres. A ma connaissance, une étude porte sur la compréhension des relatives libres en portugais européen, mais celles à référent [+humain] : il s'agit des relatives introduites par le mot *wh-* 'quem' (*qui* en français). Rien n'est dit sur les relatives à référent [-inanimé], à savoir celles introduites par 'o que'. Je vais présenter les résultats majeurs de ce papier de COSTA, GRILLO et LOBO, 2012 dans la sous-section qui suit.

5.1.1 Etude portant sur la compréhension des relatives libres introduites par 'quem' : COSTA, GRILLO et LOBO, 2012

5.1.1.1 Objectif général

Nous avons vu précédemment, que, comme le prévoit l'hypothèse de la Minimalité Relativisée et des travaux de FRIEDMANN, BELLETTI et RIZZI, 2009, les relatives libres sont plus faciles à comprendre et à produire que les relatives pleines. Le but de COSTA, GRILLO et LOBO, 2012 est de comparer la compréhension des relatives pleines et libres par les enfants et les adultes, afin de déterminer s'il s'agit d'un problème développemental ou bien si cette préférence pour les relatives libres se retrouve dans le traitement par les adultes également. Plus particulièrement, ils veulent vérifier l'importance du trait [\pm NP] de restriction lexicale dans l'approche de FRIEDMANN, BELLETTI et RIZZI, 2009, dont nous avons déjà parlé dans le Chapitre 4 (Section 4.1.1.1). Est-ce que l'on trouve des effets de minimalité même en l'absence de restriction lexicale sur l'un des deux DPs ?

5.1.1.2 Design et prédictions

COSTA, GRILLO et LOBO, 2012 ont mis en place une tâche de choix de personnage à partir d'images. 4 conditions expérimentales ont été manipulées, dans un design 2×2 : relatives pleines / relatives libres ; *gap* sujet / *gap* objet. Il y avait 20 items par condition, ce qui fait un total de 80 phrases expérimentales. Un exemple pour chaque condition est donné en (1) [COSTA, GRILLO et LOBO, 2012 : 149].

- (1) a. **Relative pleine sujet**
Mostra-me o menino que seca o hipopótamo.
'Montre-moi le garçon qui sèche l'hippopotame'.
- b. **Relative pleine objet**
Mostra-me o menino que o hipopótamo seca.
'Montre-moi le garçon que l'hippopotame sèche'.
- c. **Relative libre sujet**
Mostra-me quem seca o hipopótamo.
'Montre-moi qui sèche l'hippopotame'.
- d. **Relative libre objet**
Mostra-me quem o hipopótamo seca.
'Montre-moi qui l'hippopotame sèche'.

Les deux images correspondantes aux items donnés en exemple sont présentées ci-dessous.

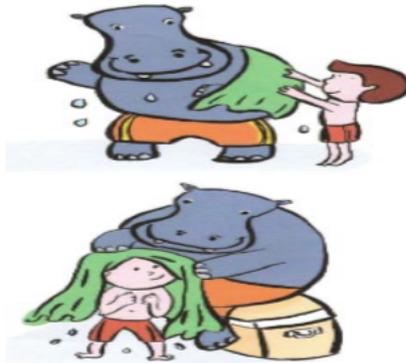


Fig. 1. An example for a picture pair used in the comprehension task.

FIGURE 5.1 – Un exemple de paire d'images utilisée dans le test de compréhension des relatives par COSTA, GRILLO et LOBO, 2012 : 149.

Dans le cas des relatives libres, notons que 'quem' (*qui*), même s'il ne contient pas de restriction lexicale ([-NP] en les termes de FRIEDMANN, BELLETTI et RIZZI, 2009), possède une riche structure interne : [+D], [+animé], [+humain], singulier, nominatif/accusatif. Cela le distingue du complémenteur 'que'. La Minimalité Généralisée (GRILLO, 2005, GRILLO, 2008)¹

1. L'hypothèse générale est la suivante : les difficultés qu'ont les enfants et les aphasiques agrammatiques, soit aphasiques de Broca, avec les dépendances impliquant un mouvement peuvent être expliquées en tant qu'effet de minimalité qui serait généré par une sous-spécification de la composition en traits morphosyntaxiques normalement associée aux catégories syntaxiques. Concrètement, dans les relatives objet (par exemple), une correspondance partielle en traits entre l'objet déplacé et le sujet DP intervenant pourrait générer un effet de minimalité et donc engendrer des difficultés de compréhension.

prédit que le fait de déplacer 'quem' par-dessus le sujet DP produirait quand même un effet d'intervention (malgré le *mismatch* en ce qui concerne la restriction lexicale), dû à la correspondance en traits partielle entre les deux DPs (sujet et objet) : [+D], [+animé], [+humain], singulier, nominatif/accusatif.

5.1.1.3 Participants

40 enfants locuteurs natifs du portugais européen, âgés entre 4;0 et 5;11, ont participé à cette étude. Le groupe contrôle d'adultes était constitué de 20 locuteurs natifs, âgés entre 24 et 47 ans.

5.1.1.4 Les résultats

En ce qui concerne les résultats, notons tout d'abord une performance maximale pour les adultes et ce, dans les quatre conditions.

Les Tableaux 5.1, 5.2 et 5.3 indiquent respectivement la performance globale pour tous les enfants, la performance des enfants de 4 ans seulement, et la performance des enfants de 5 ans seulement.

	REL pleines	REL libres
Sujet	98,1%	94%
Objet	81,8%	82,25%

TABLE 5.1 – Résultats globaux de l'expérience de COSTA, GRILLO et LOBO, 2012 : 150.

Les résultats globaux révèlent une asymétrie sujet-objet dans les deux conditions (relatives pleines et relatives libres). De plus, si l'on se centre sur les relatives objet seulement, il n'est pas vérifié que les relatives libres objet sont plus faciles à comprendre que les relatives pleines objet pour les enfants.

	REL pleines	REL libres
Sujet	96,5%	91,5%
Objet	76,75%	83,5%

TABLE 5.2 – Résultats de l'expérience de COSTA, GRILLO et LOBO, 2012 : 151 : performance des enfants de 4 ans.

	REL pleines	REL libres
Sujet	99,75%	96,5%
Objet	87%	87%

TABLE 5.3 – Résultats de l'expérience de COSTA, GRILLO et LOBO, 2012 : 151 : performance des enfants de 5 ans.

Si l'on sépare les résultats des enfants de 4 ans de ceux des enfants de 5 ans, on remarque que l'avantage pour les relatives libres objet sur les relatives pleines objet est présent chez les 4 ans mais disparaît chez les 5 ans, avec une performance équivalente pour les deux types de dépendances objet.

5.1.1.5 Le traitement par les adultes

Les auteurs ont décidé d'étudier le traitement de ces mêmes structures par les adultes via une tâche de mesure *online*, en particulier une tâche de lecture par présentation auto-segmentée (*self-paced reading task*), avec des questions de compréhension.

Ils ont créé 20 ensembles d'items expérimentaux, dans les 4 conditions. Un exemple pour chaque condition est donné en (2) [COSTA, GRILLO et LOBO, 2012 : 153].

(2) a. **Dépendance libre sujet**

A advogada contou ao juiz quem estava a corromper o político com dinheiro público.

'L'avocate a raconté au juge qui s'apprêtait à corrompre le politicien avec de l'argent public.'

b. **Dépendance libre objet**

A advogada contou ao juiz quem o político estava a corromper com dinheiro público.

'L'avocate a raconté au juge qui le politicien s'apprêtait à corrompre avec de l'argent public.'

c. **Dépendance pleine sujet**

A advogada contou ao juiz que político estava a corromper o presidente com dinheiro público.

'L'avocate a raconté au juge quel politicien s'apprêtait à corrompre le président avec de l'argent public.'

d. **Dépendance pleine objet**

A advogada contou ao juiz que político o presidente estava a corromper com dinheiro público.

'L'avocate a raconté au juge quel politicien le président s'apprêtait à corrompre avec de l'argent public.'

26 adultes locuteurs natifs du portugais européen, étudiants et membres du personnel de l'*Universidade NOVA* à Lisbonne ont participé à cette expérience.

Les résultats sont présentés dans le Tableau 5.4. Ils révèlent une asymétrie entre dépendances wh- sujet et dépendances wh- objet : les dépendances wh- sujet ont été lues plus rapidement que les dépendances wh- objet et la performance en ce qui concerne la compréhension est meilleure également pour les dépendances wh- sujet. Quant à la différence entre dépendances wh- pleines et dépendances wh- libres, il s'avère que le temps de lecture est plus court pour les dépendances wh- libres et que globalement il y a plus de bonnes réponses aux questions de compréhension pour les dépendances wh- libres que pour les dépendances wh- pleines.

	Dépendances pleines	Dépendances libres
Sujet	77,6%	87,2%
Objet	64,0%	90,4%

TABLE 5.4 – Performance des adultes (pourcentage de réponses correctes) pour l'expérience de lecture par présentation auto-segmentée de COSTA, GRILLO et LOBO, 2012 : 155.

Ainsi, les auteurs concluent que, bien que considérablement réduits, les effets d'intervention ne disparaissent pas dans les dépendances introduites par un pronom relatif sans restriction lexicale (relatives libres). Leurs résultats confirment que plus les deux DPs (sujet et objet) sont similaires en termes de traits, plus les effets d'intervention sont forts dans le cas des dépendances A-barre objet.

Retenons pour la suite, que, globalement, les relatives libres introduites par 'quem' semblent être plutôt bien comprises par les enfants portugais de 4-5 ans, et que l'asymétrie entre relatives libres sujet et relatives libres objet est bel est bien présente. La performance des enfants de 5 ans pour les relatives libres objet est équivalente à celle pour les relatives pleines objet, ce qui suggère que 'quem' requiert tout de même un certain coût de traitement, peut-être à cause de sa riche composition en traits.

5.2 Données de production spontanée

Nous venons de voir que la littérature est assez pauvre au sujet de l'acquisition des subordinées introduites par 'o que' en portugais européen. Aucune étude ne s'est focalisée sur le double usage de 'o que', à savoir dans les relatives et les interrogatives indirectes. Avant de mettre en place un test de répétition similaire à celui en français, j'ai regardé des données de production spontanée.

5.2.1 Le corpus

Le corpus étudié dans le cas du portugais est un corpus qui a été constitué en 2015 par deux étudiants de Master de l'Université NOVA à Lisbonne (CAMPOS, 2015a et CAMPOS, 2015b). Elles ont enregistré 50 enfants portugais âgés entre 3;5 ans et 7;10 ans dans le cadre de leur mémoire de Master.² Les données ont été recueillies à partir de dialogues entre l'enfant et l'investigatrice, d'interactions entre l'enfant et l'investigatrice au cours d'un jeu symbolique, mais aussi de narration d'histoires à partir de séquences d'images.

Tabela 2 - Caracterização da amostra da população.

Faixa etária	Número de participantes	Média da idade	Género		Nível de Escolaridade	
			M	F	PE	E
[3;00 – 3;11]	10	3A 8M	1	9	10	0
[4;00 – 4;11]	10	4A 5M	6	4	10	0
[5;00 – 5;11]	10	5A 6M	5	5	10	0
[6;00 – 6;11]	10	6A 5M	6	4	5	5
[7;00 – 7;11]	10	7A 3M	6	4	0	10

FIGURE 5.2 – Caractéristiques des enfants pour l'étude de production spontanée (CAMPOS, 2015a : 18). Les informations indiquées de gauche à droite sont respectivement la tranche d'âge, le nombre de participants, la moyenne d'âge, le genre et le niveau de scolarité.

5.2.2 Les résultats

Regardons les résultats de cette étude. Le Tableau 5.5 indique le nombre d'occurrences de 'o que' par enfant. Plus précisément, une première colonne indique le nombre d'occurrences de 'o que' introduisant une proposition subordonnée, qu'il s'agisse d'une relative ou d'une interrogative indirecte, et une deuxième colonne indique le nombre d'occurrences de 'o que' introduisant une question racine.

2. Comme on peut le voir dans la Figure 5.2 (CAMPOS, 2015a), il y a en tout 50 enfants, mais, en ce qui concerne mon étude, je n'ai pas analysé les données des 10 enfants de 7 ans. En effet, je me centre sur les enfants âgés entre 3 et 6 ans, comme pour mon étude sur le français. J'ai donc compté les occurrences de 'o que' dans la production de ces 40 enfants. Les transcriptions ayant été faites au format CHAT, j'ai pu directement faire des recherches dans les fichiers à l'aide des programmes CLAN de CHILDES (MACWHINNEY, 2000).

Enfant	Âge	'o que' dans une SUB	'o que' dans une Q racine
1	3;5.01	0	5
2	3;7.07	0	1
3	3;7.19	0	0
4	3;8.06	1	4
5	3;9.03	0	8
6	3;9.10	2	3
7	3;9.24	2	6
8	3;10.19	1	1
9	3;10.20	1	0
10	3;11.13	0	3
11	4;0.11	1	0
12	4;2.17	2	1
13	4;4.12	2	1
14	4;4.18	0	2
15	4;6.01	2	0
16	4;6.02	0	3
17	4;7.14	0	4
18	4;8.15	0	0
19	4;9.06	2	3
20	4;11.16	0	0
21	5;2.15	1	0
22	5;2.26	1	0
23	5;3.10	1	0
24	5;6.18	0	3
25	5;7.04	3	0
26	5;8.22	1	2
27	5;9.13	2	0
28	5;9.14	0	1
29	5;9.27	2	0
30	5;11.09	0	0
31	6;1.15	0	1
32	6;2.02	1	0
33	6;2.08	2	0
34	6;3.10	2	0
35	6;4.16	3	0
36	6;7.17	1	0
37	6;7.23	1	0
38	6;10.18	1	0
39	6;11.14	2	1
40	6;11.19	0	0
TOTAL		40	53

TABLE 5.5 – Nombre d'occurrences de *o que* dans les subordonnées (relatives ou interrogatives) et dans les questions racines dans la production des 40 enfants.

Comme le montre le Tableau 5.5, il y a juste quelques occurrences par enfant, voire pas d'occurrences chez certains enfants, mais au total, j'ai relevé 40 occurrences de 'o que' dans des subordonnées et 53 occurrences de 'o que' dans des questions racines. On peut remarquer que les enfants plus jeunes produisent des questions racines introduites par 'o que' mais pas forcément encore des subordonnées introduites par 'o que'. Cela va dans le sens de ce que SOARES, 2006 avait trouvé dans sa thèse : les interrogatives racines introduites par 'o que' sont beaucoup plus précoces que les mêmes interrogatives enchâssées.

En ce qui concerne les subordonnées introduites par 'o que' produites par les enfants, il s'avère que parmi les 40 subordonnées, 17 sont des relatives et 6 des clivées³, 9 sont des interrogatives indirectes (introduites par '(não) saber', signifiant (*ne pas*) *savoir*), et 8 sont ambiguës (introduites par des verbes tels que 'ver' *voir*, 'olhar' *regarder* ou 'dizer' *dire*). Si l'on considère que les clivées sont un cas particulier de relatives, on se retrouve alors avec 23 relatives et seulement 9 interrogatives indirectes. Voilà quelques exemples :

- (3) o menino (es)tava a **faze(r)** o (1)que a mãe (es)tava a dize(r) [relative ; 4;9.06]
'Le garçon s'apprêtait à faire ce que la mère s'apprêtait à dire.'
- (4) <e posso> [/] e posso **construir** o (1)que eu quiser [relative ; 5;7.04]
'Et je peux construire ce que je voudrai.'
- (5) é o (1)que eu mais adoro ver [clivée ; 5;2.26]
'C'est ce que j'aime voir le plus.'
- (6) **não sabia** o que a mãe disse [interrogative indirecte ; 3;9.24]
'Je ne savais pas ce que la maman avait dit.'
- (7) **não sei** o (1)que é isso [interrogative indirecte ; 6;7.23]
'Je ne sais pas ce que c'est.'
- (8) **olha** o que fizeram à mesa [structure ambiguë ; 3;9.10]
'Regarde ce qu'ils ont fait à table.'
- (9) ab(r)i a boca pa(ra) **ve(r)** o (1)que se passava [structure ambiguë ; 4;6.01]
'J'ai ouvert la bouche pour voir ce qui se passait.'

On trouve également de nombreux exemples de 'o que' introduisant des questions racines.

- (10) o que é isto? [question racine fréquente, produite par plusieurs enfants, voire plusieurs fois par le même enfant]
'Qu'est-ce que c'est?'
- (11) o que foi? [question racine ; 5;9.14]
'Qu'est-ce que c'était?'

3. Je ne rentre pas dans les détails des sous-classes de clivées ici, étant donné que mon travail ne porte pas sur ces constructions. Mais notons qu'il existe plusieurs types de clivées en portugais européen. Pour une typologie, voir LOBO, SANTOS et SOARES-JESEL, 2016.

D'après nos données, 'o que' est plus fréquent dans les questions racines que dans les subordonnées. En ce qui concerne la distribution dans les subordonnées, 'o que' est plus fréquent dans les relatives que dans les interrogatives indirectes. Cela est sûrement dû au fait que l'emploi de 'o que é que' soit possible dans les interrogatives indirectes mais pas dans les relatives. D'après SOARES, 2006, les interrogatives enchâssées sont majoritairement en 'é que' à partir de 3;5.

5.3 Expérience 1 : test de répétition de phrases contenant des subordonnées introduites par *o que*

5.3.1 Prédications

Il est intéressant de remarquer que nous avons trouvé bien plus d'occurrences de 'o que' en portugais que d'occurrences de 'ce qu-' en français dans les données de production spontanée. Bien sûr, il s'agit de corpora différents, mais cela pourrait être un indice que les enfants portugais âgés entre 3 et 6 ans auraient moins de difficultés avec 'o que' que les enfants français de la même tranche d'âge n'en ont avec 'ce qu-'. Dès le début, les enfants portugais produisent 'o que' dans les questions racines. Si cette possibilité est l'indice d'une grammaticalisation de *o+que* en un élément *wh-*, alors on s'attend à trouver des tendances similaires dans l'acquisition de 'o que' et dans l'acquisition de 'quem', mot *wh-* sans élément D. Étudions dans un premier temps les résultats du test de répétition de subordonnées introduites par 'o que', et, dans un deuxième temps, comparons ces résultats avec ceux du test de répétition de subordonnées introduites par 'quem'.

5.3.2 Design expérimental

5.3.2.1 Participants

Ont participé à cette expérience 55 enfants monolingues portugais, tous présentant un développement typique du langage, et âgés entre 3 et 6 ans. Ils ont tous été testés à Santa Iria de Azoia, près de Lisbonne, au sein de deux écoles. Bien entendu, un accord parental a été obtenu au préalable pour chacun des enfants ayant participé. Je précise que trois des enfants ont été exclus de l'analyse car leur prononciation n'était pas assez nette, et ce, tout au long de l'expérience, pour que l'on puisse déterminer avec précision ce qu'ils disaient et/ou ils avaient des difficultés même avec les items d'entraînement et/ou *fillers*.

5.3.2.2 Matériel

Les conditions expérimentales sont les mêmes que pour l'Expérience 1 en français, présentée dans le Chapitre 4 : il s'agit d'un design 2*2, croisant les facteurs TYPE DE PHRASE (relatives ou interrogatives indirectes) et TYPE DE GAP (*gap* sujet ou objet). Il y a 4 items par condition donc 16 items en tout, auxquels se mélangent 10 *fillers*. 5 items d'entraînement

ont été présentés au début de l'expérience afin de s'assurer que les enfants comprenaient bien la tâche.

Le matériel correspond à une traduction/adaptation du matériel français par des linguistes, locutrices natives du portugais européen (Carla SOARES, Margarida TOMAZ et Maria LOBO). L'idéal aurait été de conserver également la longueur des items français, à savoir 12 syllabes. Mais comme plusieurs des mots de vocabulaire de base requis sont plus longs en portugais qu'en français, les items portugais sont plus longs d'une syllabe que les items français. Ils sont donc tous constitués de 13 syllabes.

Voilà deux exemples d'items pour chacune des quatre conditions expérimentales, correspondant à des traductions/adaptations des items français donnés comme exemples dans le Chapitre 4, Section 4.4.2.2 (de (22) à (29)) [voir Annexe D pour la liste complète des items].

1. Relatives à tête allégée avec *gap* sujet

- (12) A tia come o que caiu do armário.
'La tante mange ce qui est tombé de l'armoire.'
- (13) O senhor apanha o que caiu na sala.
'Le monsieur attrape ce qui est tombé dans la salle.'

2. Interrogatives indirectes avec *gap* sujet

- (14) A tia não sabe o que caiu do armário.
'La tante ne sait pas ce qui est tombé de l'armoire.'
- (15) O senhor pergunta o que caiu na sala.
'Le monsieur demande ce qui est tombé dans la salle.'

3. Relatives à tête allégée avec *gap* objet

- (16) O senhor apanha o que o rapaz lhe atirou.
'Le monsieur attrape ce que le garçon lui a lancé.'
- (17) O pai apaga o que a menina desenhou.
'Le papa efface ce que la fille a dessiné.'

4. Interrogatives indirectes avec *gap* objet

- (18) A tia pergunta o que o rapaz atirou.
'La tante demande ce que le garçon a lancé.'
- (19) O pai pergunta o que a menina desenhou.
'Le papa demande ce que la fille a dessiné.'

Comme pour le français, les items ont été randomisés avec Excel, de façon à ce qu'il y ait autant de listes que de participants.

Les *fillers* ont également été traduits-adaptés à partir des phrases françaises : il s'agit de phrases plus courtes et simples (sans subordination). Voilà deux exemples correspondants aux exemples (30) et (31) du Chapitre 4, Section 4.4.2.2.

- (20) A menina gosta de pintar.
'La fille aime bien peindre.'
- (21) A Maria gosta do João.
'Marie est amoureuse de Jean.'

Comme dans le cas du français, les items d'entraînement étaient des phrases complexes aussi longues que les items expérimentaux (13 syllabes). Les 3 items d'entraînement ont été traduits-adaptés à partir des items français (complétives introduites par 'que'). J'ai ajouté deux items d'entraînement comparé à l'expérience française. Je me suis en effet rendue compte lors de la passation de l'expérience française qu'il fallait allonger la phase d'entraînement pour que les enfants aient le temps de bien comprendre la tâche avant de passer aux items expérimentaux. Encore une fois, ces items sont importants car ils permettront de déterminer si les enfants portugais ont des difficultés avec 'o que' en particulier, ou bien plus généralement avec des phrases complexes et la subordination.

- (22) A neta pensa que o avô comeu muito pão.
'La petite fille pense que le grand-père a mangé beaucoup de pain.'

5.3.2.3 Procédure

La procédure est la même que celle pour le test de répétition français correspondant (cf. Chapitre 4, Section 4.4.2.3). Par contre, le portugais n'étant pas ma langue maternelle, c'est une locutrice native qui a fait passer le test aux enfants portugais, mais Margarida Tomaz (de l'Université NOVA, à Lisbonne). Je l'ai accompagnée pour la passation de l'expérience sans intervenir.

5.3.3 Résultats

Dans cette sous-section, je vais présenter les résultats de ce test de répétition, en deux temps, comme dans le cas du français : je commencerai par l'analyse quantitative des résultats, afin de voir si les enfants portugais ont des difficultés avec 'o que', s'il y a un effet de l'âge, et si la performance diffère selon le type de structure (relative ou interrogative indirecte). Ensuite, je ferai l'analyse qualitative, en me centrant sur les types d'erreurs que les enfants portugais font lorsqu'ils ne répètent pas la phrase cible correctement. Là encore, il sera intéressant de déterminer si tel ou tel type d'erreur apparaît plutôt dans tel ou tel type de phrase.

5.3.3.1 Analyse quantitative : la performance des enfants

Comme pour le français, les données ont été analysées avec des modèles linéaires mixtes généralisés, avec la fonction de lien *logit* (BAAYEN, DAVIDSON et BATES, 2008, JAEGER, 2008), en utilisant le paquet *lme4* (BATES, MÄCHLER, BOLKER et WALKER, 2015) dans R (RCORE-TEAM, 2017).

Le codage des réponses données par les enfants est le même que pour le français : 1 pour une répétition correcte et 0 pour une répétition incorrecte, toujours en se focalisant principalement sur la séquence clé, à savoir ‘o que’.

Le facteur TYPE DE PHRASE a été codé 0.5 pour interrogative indirecte et -0.5 pour relative. De même pour le facteur TYPE DE GAP : 0.5 pour *gap* objet et -0.5 pour *gap* sujet. De plus, la covariable ÂGE a été insérée dans le modèle en tant que variable continue centrée. J’ai en effet inclus tous les participants dans un seul groupe, et vérifié l’effet de l’âge sur leur performance.

Dans la partie ‘effets fixes’, nous avons estimé les effets principaux, ainsi que toutes les interactions possibles de TYPE DE PHRASE, TYPE DE GAP et ÂGE. Dans la partie ‘effets aléatoires’, nous avons inclus seulement un intercept pour les sujets et un pour les items, dans le but d’éviter un surparamétrage et, par conséquent, des résultats de modèle qui seraient ininterprétables (BATES, KLIEGL, VASISHTH et BAAYEN, 2015, MATUSCHEK et al., 2017).

Le tableau 5.6 indique le score moyen, entre 0 et 1, pour chaque type de phrase, à savoir les items d’entraînement, les items expérimentaux, et les distracteurs.

Type de phrase	Score moyen
items d’entraînement	0.85
items expérimentaux	0.77
distracteurs	0.99

TABLE 5.6 – Scores moyens en fonction du type de phrase (Expérience 1).

On peut remarquer que la performance pour les items expérimentaux est assez élevée et très proche de celle pour les items d’entraînement. Les enfants portugais n’ont donc pas de grandes difficultés à répéter des phrases contenant des subordinées introduites par ‘o que’. Quant aux distracteurs, la performance est très élevée, atteignant presque le score de 1.

La proportion de répétitions correctes à travers les conditions et en fonction de l’âge est illustrée dans la Figure 5.3.

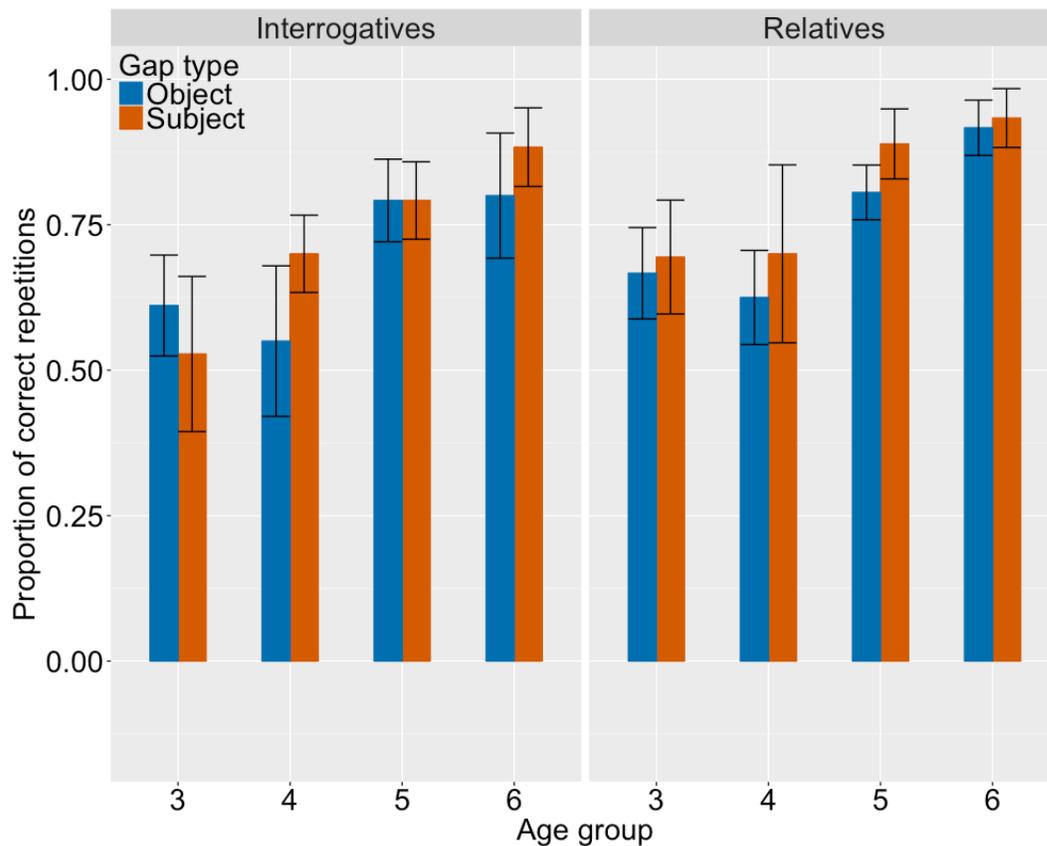


FIGURE 5.3 – Proportion de répétitions correctes (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Le panneau de gauche montre la performance pour les interrogatives indirectes et celui de droite celle pour les relatives à tête allégée. Les barres bleues indiquent la performance pour les dépendances objet, et les oranges celle pour les dépendances sujet.

Comme on peut le voir sur la Figure 5.3, il y a tout d'abord une augmentation générale de répétitions correctes au fur et à mesure que les enfants prennent de l'âge. Cela est reflété par l'effet principal de l'ÂGE, qui est significatif (coeff=1.40, SE=0.51, $z=2.73$, $p=0.006$). Ensuite, en ce qui concerne la différence entre relatives et interrogatives indirectes, elle s'avère être significative, et ce, en faveur des relatives (coeff=-1.02, SE=0.28, $z=-3.59$, $p=0.0003$) : la performance des enfants portugais est donc meilleure lorsqu'il s'agit de répéter des relatives que lorsqu'il faut répéter des interrogatives indirectes. Ces résultats sont cohérents avec ceux présentés par SOARES, 2006 sur 'o que' (données de production spontanée). Elle montre en effet dans sa thèse que les relatives introduites par 'o que' commencent à être réalisées vers 2;11 ans et que les interrogatives indirectes introduites par 'o que' commencent à être réalisées un peu plus tardivement, à partir de 3 ans.

Par ailleurs, dans notre expérience, on observe un effet du TYPE DE GAP : les enfants portugais ont moins de difficultés à répéter des phrases contenant des subordinées introduites par 'o que' lorsque le *gap* est en position sujet que lorsqu'il est en position d'objet (coeff=-0.70, SE=0.28, $z=-2.53$, $p=0.011$). Enfin, aucune des interactions dans le modèle n'était significative (toutes les valeurs $p > 0.05$). Notons que la performance générale des enfants portugais pour les items expérimentaux est bien meilleure que celle des enfants français. En effet, le taux d'erreur n'est que de 23%. (rappelons qu'il était de 55% pour le français). Les enfants portugais ne semblent pas avoir de difficultés particulières avec 'o que' : leur performance pour les items expérimentaux est élevée (0.77) et elle se rapproche de celle pour les items d'entraînement (0.85).

5.3.3.2 Analyse qualitative : les types d'erreur

Regardons maintenant les types d'erreurs qu'ont fait les enfants portugais lorsqu'ils ne répétaient pas correctement la phrase cible. La diversité des erreurs est moins grande dans le cas du portugais que dans le cas du français. Mais, comme en français, la plupart des erreurs concerne la production de la séquence clé, à savoir 'o que'. 200 erreurs ont été relevées en tout (sur un total de 832 observations).

5.3.3.2.1 L'omission de 'o' Le type d'erreur le plus fréquent, est, comme en français, l'omission de l'élément D, ici 'o'. En effet, 54,5% (109 sur 200) des erreurs impliquent l'omission de l'élément D (voir exemples (23) et (24)).

- (23) *Phrase cible* : A mãe adivinha o que o menino partiu.
 'La maman devine ce que le garçon a cassé.'
Phrase réponse : A mãe adivinha que o menino partiu. [Lourenço, 4;5]
 'La maman devine que le garçon a cassé.'
- (24) *Phrase cible* : O pai pergunta o que a menina desenhou.
 'Le papa demande ce que la fille a dessiné.'
Phrase réponse : O pai pergunta que a menina desenhou. [Simão, 3;11]
 'Le papa demande que la fille a dessiné.'

L'utilisation de 'que' tout seul dans les interrogatives indirectes est possible en portugais, mais il requiert l'inversion sujet-verbe.

(25) O pai pergunta que desenhou a menina.

Les productions des enfants avec omission de 'o' étant sans inversion, elles sont agrammaticales.

La proportion d'omission du déterminant en fonction du TYPE DE PHRASE, du TYPE DE GAP et de l'ÂGE est représenté dans la Figure 5.4⁴.

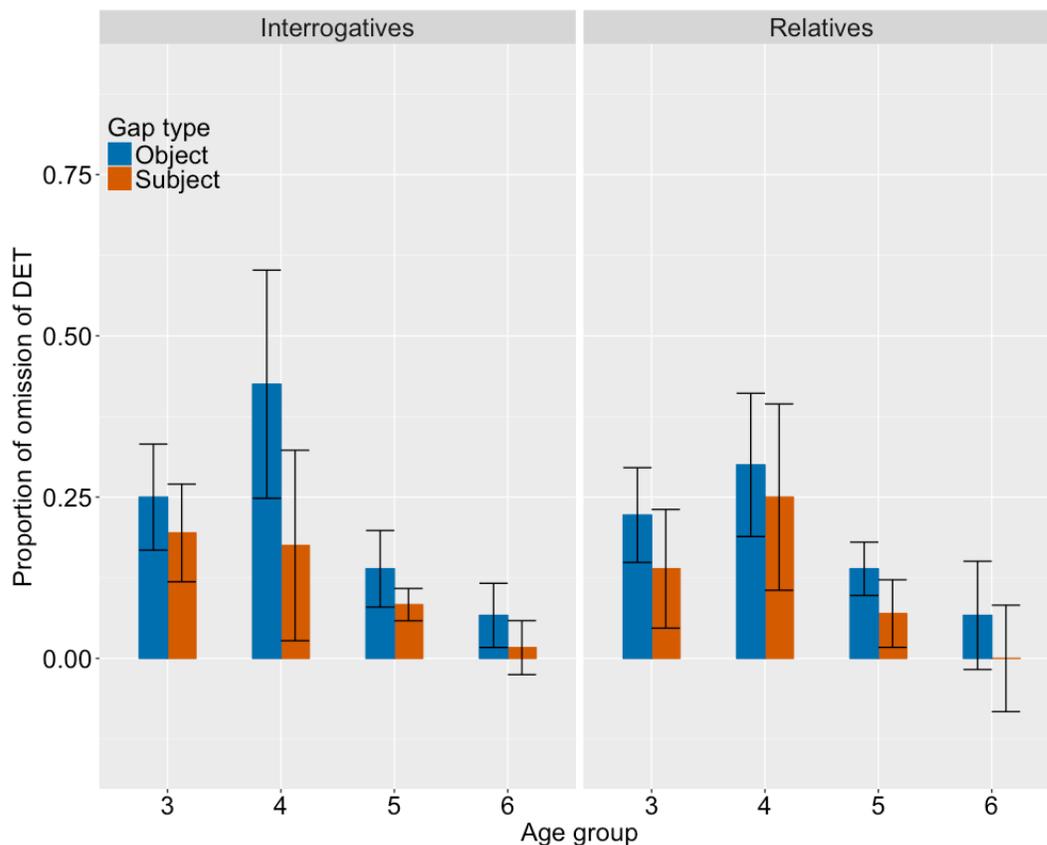


FIGURE 5.4 – Proportion d'omission du déterminant 'o' (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Le panneau de gauche montre la proportion d'omissions du déterminant 'o' dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion d'omissions du déterminant 'o' dans les dépendances objet, et les oranges celle dans les dépendances sujet.

4. Dans la Figure 5.4, la proportion d'omission du déterminant est calculée par rapport à toutes les phrases réponses données, correctes ou fausses, et non pas seulement parmi les erreurs.

L'analyse statistique indique que les enfants portugais ont tendance à omettre le déterminant plus dans les dépendances objet que dans les dépendances sujet, l'effet du TYPE DE GAP étant significatif (coeff=1.95, SE=0.458, $z=4.258$, $p<0.001$). Aucune différence n'a été relevée entre les relatives et les interrogatives indirectes, ce qui signifie que les enfants n'omettent pas plus le déterminant dans un type de phrase que dans l'autre. Il n'y a pas non plus d'effet de l'ÂGE, ce qui veut dire que les enfants portugais n'omettent pas forcément le déterminant moins fréquemment au fur et à mesure du développement. Bien que cet effet ne soit pas significatif, on peut voir d'après la Figure 5.4 que les enfants les plus jeunes (3-4 ans) omettent davantage le déterminant que les enfants les plus âgés (5-6 ans). Enfin, aucune des interactions n'est significative (toutes les valeurs $p > 0.05$).

5.3.3.2.2 L'utilisation d'une forme wh-complexe Le deuxième type d'erreur le plus récurrent, après l'omission de l'article défini 'o', est l'utilisation de la forme wh-complexe 'o que é que', correspondant au français 'qu'est-ce que' (voir exemples (26) et (27)). Il correspond à 13,5% des erreurs (27 sur 200). L'omission de 'o' et l'emploi de 'o que é que' représentent à eux seuls 68% des erreurs (136 sur 200).

- (26) *Phrase cible* : O senhor pergunta o que caiu na sala.
'Le monsieur demande ce qui est tombé dans la salle.'
Phrase réponse : O senhor pergunta o que é que caiu na sala. [Guillermo, 5;7]
'Le monsieur demande qu'est-ce qui est tombé dans la salle.'
- (27) *Phrase cible* : O avô não sabe o que a prima cozinhou.
'Le grand-père ne sait pas ce que la cousine a cuisiné.'
Phrase réponse : O avô não sabe o que é que a prima cozinhou. [Beatriz, 6;1]
'Le grand-père ne sait pas qu'est-ce que la cousine a cuisiné.'

La proportion d'utilisation de 'o que é que' en fonction du TYPE DE PHRASE, du TYPE DE GAP et de l'ÂGE est représentée dans la Figure 5.5.⁵

Cette forme wh- peut être employée pour introduire des questions directes (28) ou indirectes (29) en portugais.

- (28) a. O que queres?
'Ce que tu veux?'
- b. O que é que queres?
'Qu'est-ce que tu veux?'
- (29) a. Pergunto-me o que queres.
'Demande-moi ce que tu veux.'
- b. Pergunto-me o que é que queres.
'Demande-moi qu'est-ce que tu veux.'

5. Comme c'était le cas dans la Figure 5.4, la proportion d'utilisation de 'o que é que' est calculée par rapport à toutes les phrases réponses données, correctes ou fausses, et non pas seulement parmi les erreurs, dans la Figure 5.5.

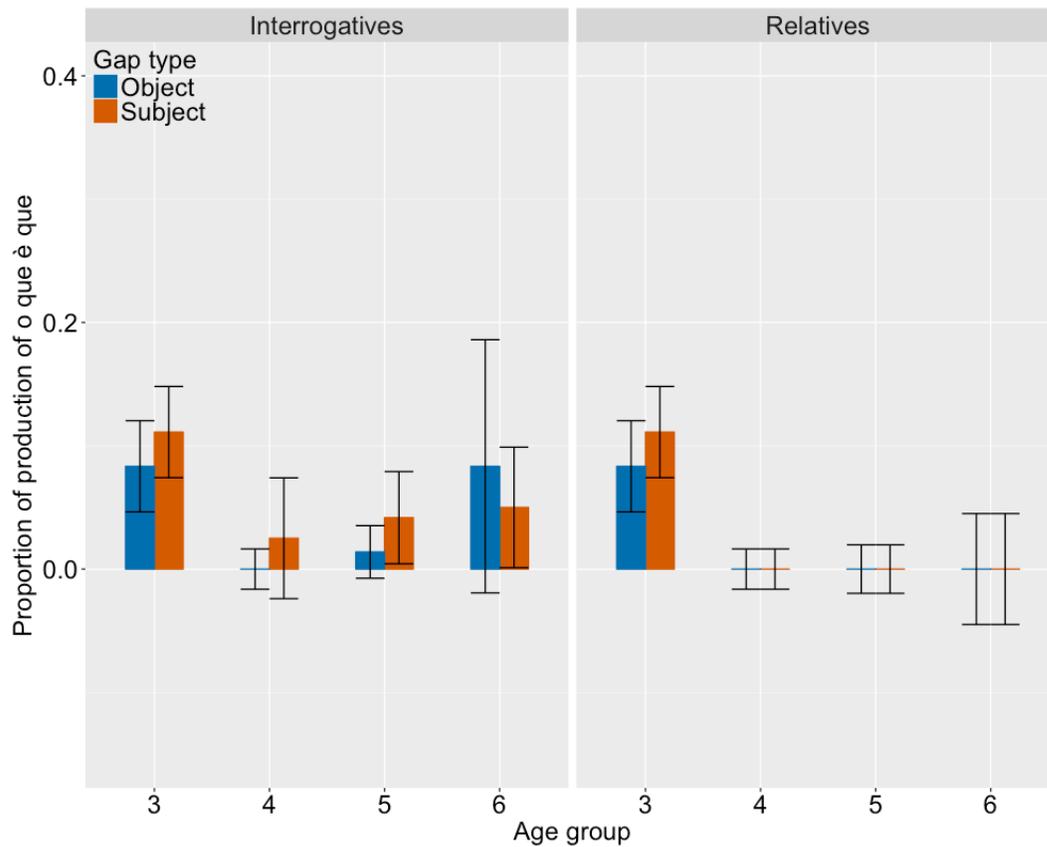


FIGURE 5.5 – Proportion d'utilisation de 'o que e que' (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Le panneau de gauche montre la proportion d'utilisation de 'o que e que' dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion d'utilisation de 'o que e que' dans les dépendances objet, et les oranges celle dans les dépendances sujet.

Que ce soit dans les questions racines ou bien les interrogatives indirectes, on peut employer soit ‘o que’ soit ‘o que é que’, pour interroger sur un référent [+ inanimé]. Ainsi, les interrogatives indirectes produites par les enfants avec ‘o que é que’ à la place de ‘o que’ ne sont pas agrammaticales, mais, étant donné que les enfants choisissent d’utiliser une autre stratégie syntaxique que celle de la phrase cible qu’ils entendent, j’ai compté ces occurrences comme des ”erreurs”.

Cela va dans le sens de ce que SOARES, 2006 avait trouvé : comme dans les questions racines, la stratégie ‘é que’ est adoptée dans la production d’interrogatives indirectes. A partir de 3;5, les interrogatives indirectes sont majoritairement en ‘é que’, comme les questions racines. ‘O que é que’ est donc plus fréquent que ‘o que’ dans les interrogatives indirectes d’après les données de SOARES, 2006. Il semble donc naturel que les enfants utilisent cette stratégie syntaxique dans le test de répétition.

En ce qui concerne l’analyse statistique, aucun des effets n’est significatif (toutes les valeurs $p > 0.05$). Cela veut dire que les enfants portugais n’emploient pas ‘o que é que’ davantage dans un type de structures (relatives ou interrogatives indirectes) que dans l’autre, ou bien davantage dans un type de dépendances que dans l’autre (dépendances sujet ou dépendances objet), et que l’emploi de cette forme wh- complexe n’est pas plus ou moins fréquent en fonction de l’âge. Pourtant, comme on peut le voir sur la Figure 5.5, il semblerait que les enfants de 3 ans auraient davantage tendance à avoir recours à cette stratégie syntaxique. De plus, globalement, tous âges confondus, il y a plus d’occurrences de ce type d’erreurs dans les interrogatives indirectes que dans les relatives. Sur les 27 occurrences de ‘o que é que’, 20 apparaissent dans les interrogatives indirectes et 7 dans les relatives. L’emploi de ‘o que é que’ est agrammatical dans les relatives. Si nous regardons de plus près ces 7 occurrences, on se rend compte qu’elles ont toutes été produites par un seul enfant, qui semble avoir généralisé la stratégie en ‘é que’ aux relatives (il a utilisé cette stratégie dans 14 items sur 16). Enfin, aucune des interactions entre les facteurs se révèle être significative (toutes les valeurs $p > 0.05$).

5.3.3.2.3 Les autres types d’erreur Les 32% d’erreurs restantes (64 sur 200) correspondent à plusieurs types d’erreurs différents, chacun assez peu fréquent. Parmi eux, je peux quand même citer les trois les plus fréquents après les deux premiers types d’erreurs que je viens de présenter : l’omission de D+C, l’emploi de ‘do que’ et le changement du verbe principal entraînant un changement de structure.

L’omission de la séquence entière ‘o que’ est tout de même assez fréquente, puisqu’elle représente 12% des erreurs (24 sur 200).

- (30) *Phrase cible* : O tio não sabe o que caiu do armario.
 ‘L’oncle ne sait pas ce qui est tombé de l’armoire.’
Phrase réponse : O tio não sabe caiu no armario. [Mario, 3 ;11]
 ‘L’oncle ne sait pas est tombé de l’armoire.’

- (31) *Phrase cible* : O senhor apanha o que o rapaz lhe atirou.
 'Le monsieur attrape ce que le garçon lui a lancé.'
Phrase réponse : O senhor apanha o rapaz lhe atirou. [David, 4;11]
 'Le monsieur attrape le garçon lui a lancé.'

La proportion d'omission de 'o+que' en fonction du TYPE DE PHRASE, du TYPE DE GAP et de l'ÂGE est représentée dans la Figure 5.6.

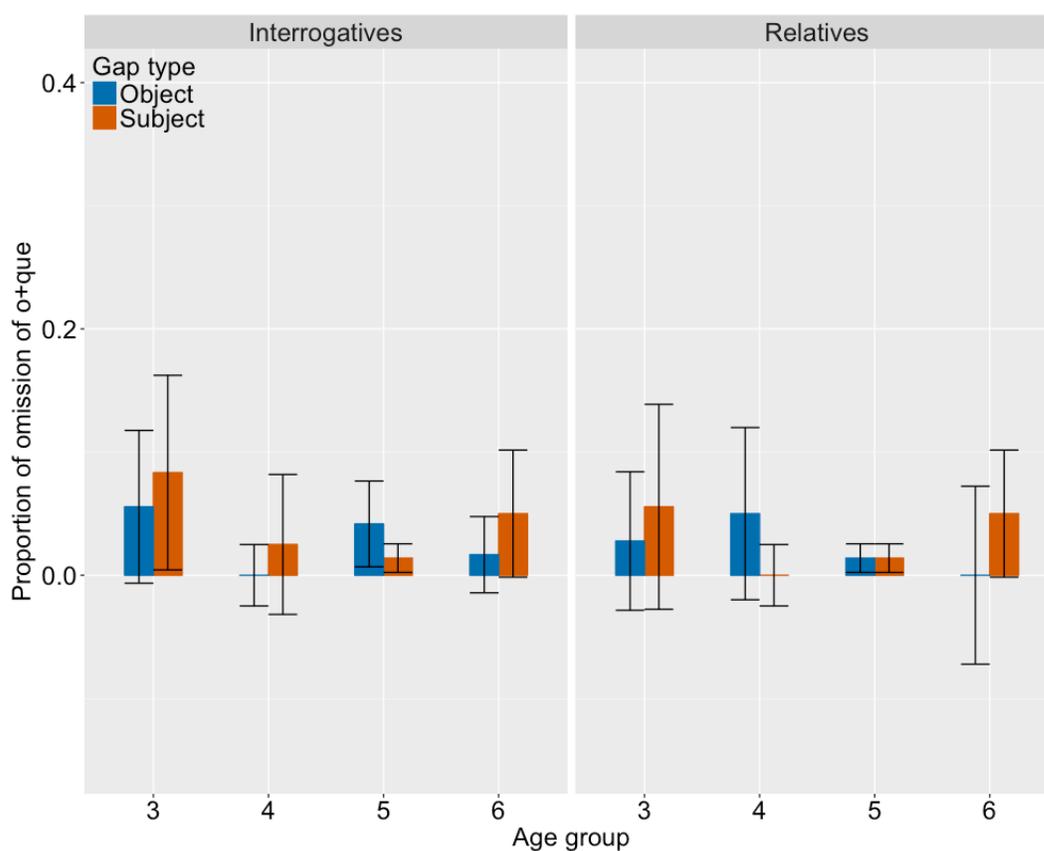


FIGURE 5.6 – Proportion d'omission de 'o+que' (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Le panneau de gauche montre la proportion d'omission de 'o+que' dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion d'omission de 'o+que' dans les dépendances objet, et les oranges celle dans les dépendances sujet.

Il y a un effet de l'ÂGE qui est marginal (coeff=-1.02, SE=0.57, $z=-1.80$, $p=0.07$). En effet, d'après la Figure 5.6, les enfants de 3 ans ont l'air d'omettre davantage o+que que les enfants des autres groupes d'âge, et ce, d'un point de vue global (interrogatives + relatives, dépendances sujet + dépendances objet). Aucun autre effet ne s'avère être significatif, ni aucune des interactions entre les facteurs (toutes les valeurs $p > 0.05$).

Un autre type d'erreur marginal est l'emploi de 'do que', qui correspond à la contraction de la préposition 'de' et de 'o que'. Mais son utilisation est agrammaticale dans nos items expérimentaux étant donné que les verbes de la proposition matrice ne sélectionnent pas de préposition 'de' (voir exemples (32) et (33)). Ce type d'erreur, qui représente 5,5% des erreurs (11 sur 200), est donc assez mystérieux, et apparaît aussi bien dans les relatives que dans les interrogatives indirectes.

- (32) *Phrase cible* : A criança apanha o que saiu da caixa.
 'L'enfant attrape ce qui est sorti de la boîte.'
Phrase réponse : A criança apanha do que saiu da caixa. [Matilde, 6;0]
 'L'enfant attrape de ce qui est sorti de la boîte.'
- (33) *Phrase cible* : A mãe adivinha o que o menino partiu.
 'La maman devine ce que le garçon a cassé.'
Phrase réponse : A mãe adivinha do que o menino partiu. [Miguel, 5;8]
 'La maman devine de ce que le garçon a cassé.'

La proportion de l'emploi de 'do que' à la place de 'o que', parmi toutes les réponses (correctes ou non), est représentée dans la Figure 5.7.

Comme cela est reflété par la Figure 5.7, il n'y a en réalité que deux enfants, un de 5 ans et un de 6 ans, qui ont fait ce type d'erreur-là. L'analyse statistique ne donne rien, et ce n'est pas très intéressant ici car il n'y a que trop peu d'occurrences.

Enfin, j'ai relevé d'autres erreurs très marginales, correspondant à des cas isolés, parmi lesquelles je peux citer le changement du terme introducteur de la subordonnée (1 seul cas, voir (34)), le changement du verbe principal entraînant un changement de structure (1 seul cas, voir (35)), ou l'omission du sujet de la subordonnée (1 seul cas, voir (36)).

- (34) *Phrase cible* : A menina adivinha o que (e)stá no quadro.
 'La fille devine ce qui est au tableau.'
Phrase réponse : A menina adivinha onde é que stá no quadro. [Carolina, 4;1]
 'La fille devine où est-ce que c'est au tableau.'
- (35) *Phrase cible* : A tia pergunta o que o rapaz atirou.
 'La tante demande ce que le garçon a lancé.'
Phrase réponse : A tia apanha o que a menina atirou. [Vicente, 4;5]
 'La tante attrape ce que la fille a lancé.'
- (36) *Phrase cible* : A senhora come o que a filha cozinhou.
 'La dame mange ce que la fille a cuisiné.'
Phrase réponse : A filha come o que cozinhou. [Micael, 4;11]
 'La fille mange ce que (elle) a cuisiné.'

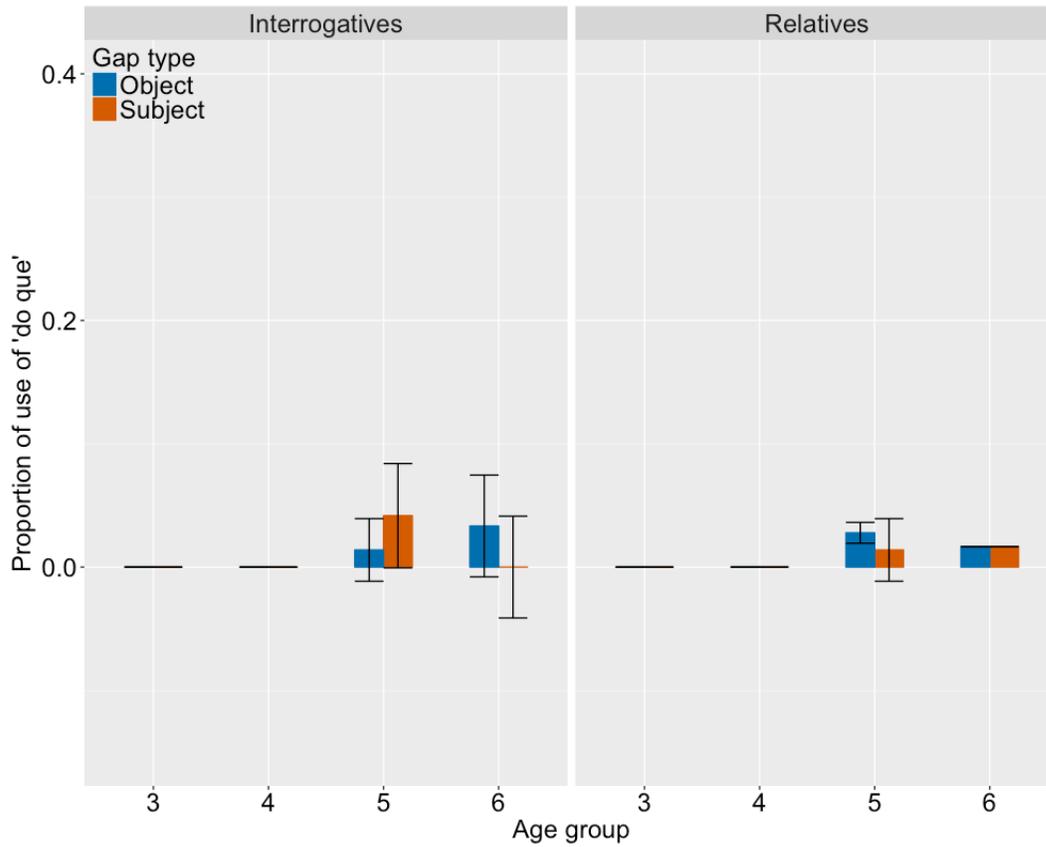


FIGURE 5.7 – Proportion d’emploi de ‘do que’ (avec erreur-type ± 1). Sur l’axe x, l’âge est représenté en années. Le panneau de gauche montre la proportion d’emploi de ‘do que’ dans les interrogatives indirectes et celui de droite celle dans les relatives à tête allégée. Les barres bleues indiquent la proportion d’emploi de ‘do que’ dans les dépendances objet, et les oranges celle dans les dépendances sujet.

Le graphique circulaire ci-après (Figure 5.8) permet d'avoir un aperçu général de la répartition des différents types d'erreurs en ce qui concerne ce premier test de répétition en portugais.

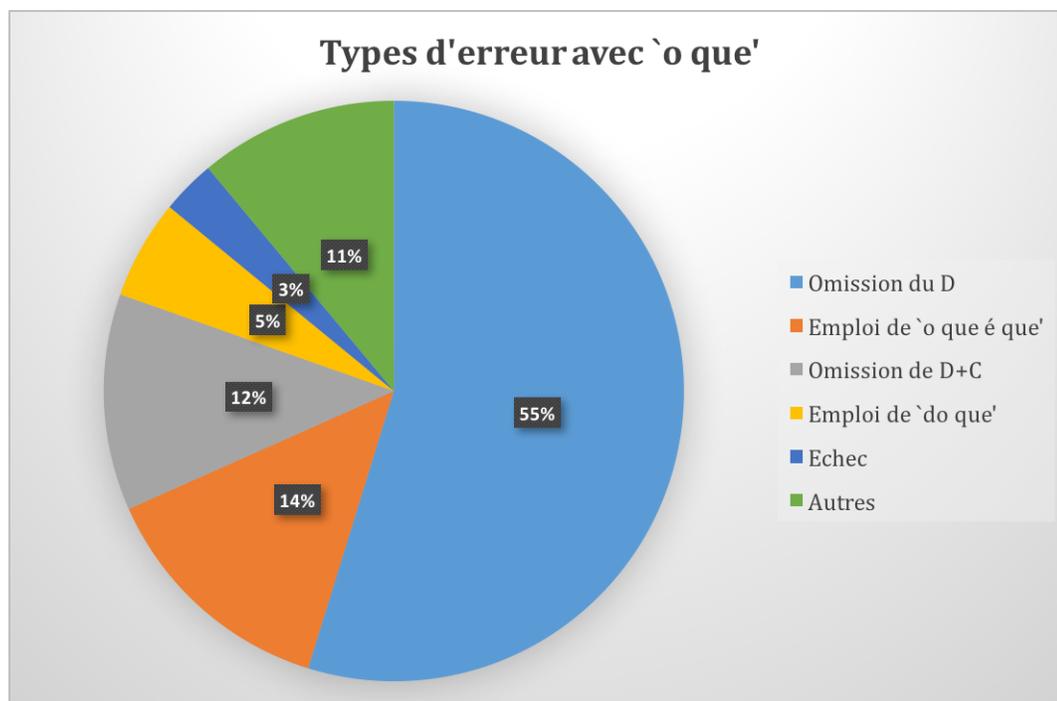


FIGURE 5.8 – Répartition des différents types d'erreurs avec 'o que'.

5.3.4 Discussion

Les résultats de cette première expérience en portugais, parallèle à l'Expérience 1 française, montrent que les enfants portugais ont beaucoup moins de difficultés avec 'o que' que les Français n'en ont avec 'ce qu-'. Toutefois, comme en français, la performance des enfants est modulée par la structure syntaxique, avec un avantage pour les relatives, et l'analyse qualitative est semblable dans les deux langues, du moins en ce qui concerne les deux principaux types d'erreurs. De plus, on retrouve en portugais l'asymétrie typique entre dépendances sujet et dépendances objet, les dépendances objet étant plus difficiles pour les enfants que les dépendances sujet. Je vais commenter ces résultats dans les sous-sections qui suivent.

5.3.4.1 'O que', un cas de grammaticalisation ?

Rappelons que 'o que' ne se comporte pas de la même façon que 'ce que' sur le plan syntaxique. En particulier, nous avons vu que l'hypothèse d'une incorporation de l'élément

D à 'que' était plus prometteuse dans le cas du portugais que dans le cas du français, 'o que' se comportant comme un élément wh-. Or, comme en français, on trouve dans le cas du portugais un avantage pour les relatives. Comment expliquer cela ? On peut envisager les possibilités qui suivent :

1. 'O que' est grammaticalisé et analysé comme un tout dans la grammaire adulte, mais les enfants âgés entre 3 et 6 ans l'analysent comme une séquence D+C. Cela implique alors de répondre aux questions suivantes : quand se produit le changement, le passage de la grammaire enfantine à la grammaire adulte ? Quel est le facteur (ou les facteurs) déclenchant ce changement ? Cela est peu convaincant, étant donné que je me place dans une perspective de continuité entre le système de l'enfant et celui de l'adulte. De plus, le fait que 'o que' soit utilisé dans les questions racines par les enfants ne va pas non plus dans cette direction.
2. 'O que' n'a en fait pas subi de processus de grammaticalisation, comme on le pensait, et donc, les enfants l'analysent comme une séquence D+C, tout comme les adultes.
3. 'O que' est un élément ambigu. Si l'on y réfléchit à nouveau, 'o que' ne se comporte comme un élément wh- que dans les interrogatives indirectes. En effet, dans le cas des relatives, il est difficile de maintenir que 'o que' est analysé comme un tout, à partir du moment où l'article 'o' peut varier ici en genre et en nombre ('o que' = celui que ; 'a que' = celle que ; 'os que' = ceux que ; 'as que' = celles que).

Je n'ai pas eu l'opportunité de tester des enfants anglais de la même tranche d'âge, mais, *a priori*, 'what' étant pour sûr un mot wh- (et sans élément D), il aurait été intéressant de vérifier si l'on retrouve ou non cet avantage pour les relatives en anglais. En revanche, je vais comparer, comme pour le français, l'acquisition de 'o que' avec 'quem' (mot wh- animé, sans D) en portugais.

La différence entre le français et le portugais n'est peut-être pas si radicale que cela, en ce sens qu'il ne s'agit pas d'une opposition nette de type absence d'incorporation *versus* présence d'incorporation, mais plutôt d'une différence plus nuancée de type absence d'incorporation générale en français *versus* présence d'incorporation dans les interrogatives indirectes seulement en portugais. Cela implique également que, finalement, 'o que' n'est pas totalement similaire à l'anglais 'what'.

L'analyse des erreurs semble aller dans cette direction également. En effet, en ce qui concerne l'omission de l'élément D ('ce' en français et 'o' en portugais), alors que les enfants français tendent à omettre 'ce' davantage dans les interrogatives indirectes que dans les relatives, aucune différence n'est relevée entre les deux types de structures dans le cas du portugais pour l'omission de 'o'. Cela signifie que cet élément nominal est moins problématique dans les interrogatives indirectes pour les enfants portugais que pour les enfants français, étant donné qu'ils savent que 'o que' peut introduire une question racine. Même si, en surface, ce morphème complexe a l'apparence d'un élément nominal suivi de 'que', les enfants portugais savent qu'ils peuvent l'utiliser comme un mot wh-, à la 'what'.

Quant à l'emploi de la forme *wh-* complexe ('qu'est-ce qu-' en français et 'o que é que' en portugais), nous avons vu que cette stratégie n'est employée dans les relatives que par un seul enfant portugais. Donc, si l'on exclut cet enfant, on retrouve en portugais la même asymétrie que pour le français : dans les deux langues, les enfants ont davantage tendance à employer une forme *wh-* complexe dans les interrogatives indirectes que dans les relatives. La situation est toutefois différente entre le français et le portugais. L'emploi de 'qu'est-ce qu-' en français est possible en français populaire, mais pas en français standard. Le fait que les enfants français tendent à employer cette forme dans les interrogatives indirectes montre qu'ils cherchent à introduire les subordonnées en question par une séquence clairement *wh-*. Or, dans le cas du portugais, il n'est pas question de registre ici. L'emploi de 'o que é que' est grammatical en portugais standard. Je ne crois pas que l'on puisse affirmer que 'o que é que' soit davantage *wh-* que 'o que', alors qu'en français, 'qu'est-ce que' est davantage *wh-* que 'ce que'.

5.3.4.2 L'asymétrie sujet-objet retrouvée

Rappelons que, dans le cas du français, les dépendances objet n'étaient pas plus difficiles pour les enfants que les dépendances sujet. J'avais expliqué cette absence d'asymétrie sujet-objet en termes d'intervention, selon les termes de FRIEDMANN, BELLETTI et RIZZI, 2009, l'idée étant que le sujet et l'objet présentaient une double inadéquation en traits.

Nous venons de découvrir qu'en portugais, cette asymétrie sujet-objet est bel et bien présente. Or, les items étant construits de façon parallèle dans les deux langues, si notre explication pour le français est la bonne, cette absence d'asymétrie aurait dû se retrouver en portugais : en portugais aussi, dans les dépendances objet, le sujet est [+ animé, +NP] alors que l'objet est [- animé, -NP].

Toutefois la différence entre les deux langues est qu'en français, le complémenteur prend deux formes distinctes selon que le *gap* est en position sujet ou en position d'objet : 'ce qui' *versus* 'ce que', alors que le portugais n'a qu'une seule et même forme dans les deux cas : 'o que'.

Il est possible que les enfants français analysent 'qui' dans 'ce qui' comme le mot *wh-* 'qui'. Nous savons que le mot *wh-* 'qui' est [+ humain]. Cela crée un conflit pour l'enfant car il sait que le démonstratif 'ce' est [+ inanimé]. Cette opposition en termes d'animacité peut engendrer des difficultés pour l'enfant dans le cas des dépendances sujet. Si l'on part du principe que les dépendances objet sont toujours un peu plus difficiles à traiter/comprendre/produire/répéter que des dépendances sujet, alors ce conflit en termes d'animacité ferait perdre aux dépendances sujet l'avantage qu'elles ont sur les dépendances objet. Cela rééquilibrerait la balance et expliquerait pourquoi, dans le cas du français, aucune asymétrie sujet-objet n'est relevée.

Par contre, dans le cas du portugais, comme la forme est la même pour les dépendances sujet et les dépendances objet, il n'y a pas un tel rééquilibrage. Donc les dépendances objet restent plus difficiles pour les enfants que les dépendances sujet.

5.4 Retour sur les données de production spontanée

Revenons maintenant sur les données de production spontanée, afin de déterminer si les types d'erreurs majeurs que les enfants portugais ont fait au cours du test de répétition se retrouvent dans la production spontanée. Plus précisément, est-ce que l'on retrouve l'omission du déterminant 'o' et l'emploi de la forme *wh-* complexe 'o que é que' dans des relatives et des interrogatives indirectes ?

5.4.1 Le corpus

Le corpus est le même que celui présenté dans la Sous-Section 5.2.1, à savoir le corpus constitué par CAMPOS, 2015a et CAMPOS, 2015b. Pour cette étude, j'ai cherché les occurrences de 'que' à la place de 'o que', ainsi que les occurrences de 'o que é que' à la place de 'o que' introduisant des propositions subordonnées, dans les données de production spontanée des 40 enfants.

5.4.2 Les résultats

Il s'avère que l'on retrouve quelques occurrences des deux types d'erreurs en question dans la production spontanée, comme on peut le voir dans le Tableau 5.7.

Enfant	Âge	Omission du D	Emploi de 'o que é que'
1	3;5.01	0	1
2	3;7.07	0	0
3	3;7.19	0	0
4	3;8.06	0	0
5	3;9.03	0	0
6	3;9.10	0	0
7	3;9.24	0	0
8	3;10.19	0	0
9	3;10.20	4	0
10	3;11.13	2	2
11	4;0.11	0	1
12	4;2.17	0	0
13	4;4.12	0	0
14	4;4.18	0	0
15	4;6.01	0	0
16	4;6.02	0	0
17	4;7.14	1	0
18	4;8.15	0	1
19	4;9.06	0	1
20	4;11.16	0	2
21	5;2.15	0	1
22	5;2.26	0	0
23	5;3.10	0	1
24	5;6.18	0	5
25	5;7.04	0	1
26	5;8.22	0	0
27	5;9.13	0	0
28	5;9.14	0	1
29	5;9.27	0	0
30	5;11.09	0	0
31	6;1.15	0	0
32	6;2.02	0	1
33	6;2.08	0	2
34	6;3.10	1	1
35	6;4.16	0	0
36	6;7.17	0	1
37	6;7.23	0	1
38	6;10.18	0	0
39	6;11.14	0	1
40	6;11.19	0	0
TOTAL		8	24

TABLE 5.7 – Nombre d'occurrences d'omission du déterminant 'o' et de l'utilisation de 'o que é que' dans la production des 40 enfants.

5.4.2.1 L'omission de 'o'

On ne trouve que 8 cas d'omission du déterminant (voir exemple (37)), ce qui est très peu, mais, tout de même, cela montre que cette erreur n'est pas juste un artefact du test de répétition.

- (37) depois é que eu conto [3;10.20]
 'Après, c'est que j'ai raconté.'
Phrase cible supposée : depois é o que eu conto
 'Après, c'est ce que j'ai raconté.'

5.4.2.2 L'utilisation de 'o que é que'

J'ai relevé 24 occurrences en tout de subordonnées introduites par 'o que é que'. Rappelons que cette stratégie syntaxique est grammaticale dans les interrogatives indirectes seulement.

- (38) eu já não me lembro **do que é que** eu fui lá fazer [4;11.16]
 'Je ne peux plus me souvenir de qu'est-ce que je suis allé faire là-bas.'
- (39) ver **o que é que** está aqui [4;8.15]
 'voir qu'est-ce qu'il y a ici.'
- (40) sabes **o que é que** é isso? [5;6.18]
 'tu sais qu'est-ce que c'est ça?'
- (41) e ele disse [///] <contou tudo à mãe> [/] contou tudo à mãe **o que é que** aconteceu .
 [5;2.15]
 'et il a dit [///] <a tout raconté à maman> [/] a tout raconté à maman qu'est-ce qui s'était passé.'

Parmi ces 24 occurrences de 'o que é que', 16 introduisent des interrogatives indirectes, 4 des relatives, 2 des clivées, et 7 des structures ambiguës sélectionnées par des verbes tels que 'ver', 'dizer', 'olhar'. Dans nos données, 'o que é que' est donc principalement utilisé par les enfants portugais dans les interrogatives indirectes. Mais 'o que' et 'o que é que' ne sont pas pour autant en distribution complémentaire. En effet, bien que 'o que' soit plus fréquent que 'o que é que' (40 occurrences *versus* 24) dans les enchâssées d'après nos données, on note une réelle alternance entre les deux.

Cette alternance se retrouve par ailleurs dans les questions racines. En effet, j'ai compté 66 occurrences de 'o que é que' introduisant une question directe, dans les données de production spontanée des 40 enfants, au total (rappel : 53 occurrences de 'o que' dans des questions racines).

- (42) o que é que aconteceu aqui? [3;9.03]
 'qu'est-ce qui s'est passé ici?'

- (43) o que é que ele (es)tá a fazer? [3;9.10]
 ‘qu’est-ce qu’il s’apprête à faire?’
- (44) e depois o que é que fazemos? [4;11.16]
 ‘et après, qu’est-ce que nous faisons?’

Il arrive qu’un même enfant utilise tantôt ‘o que’, tantôt ‘o que é que’ pour poser une question :

- (45) o que é que é isto? [5;6.18]
 ‘qu’est-ce que c’est ça?’
- (46) o que é isto? [5;6.18]
 ‘ce que c’est ça?’

D’après SOARES, 2006, les questions racines sans ‘é que’ émergent avant celles avec ‘é que’. Ces dernières ne sont pas fréquentes avant 3;5. A partir de 3;5, la situation s’inverse : elles deviennent beaucoup plus fréquentes que celles sans ‘é que’. La distribution entre ‘o que’ et ‘o que é que’ n’est pas aléatoire d’après SOARES, 2006. En effet, les questions racines en ‘é que’ seraient plus complexes et donc émergeraient plus tardivement selon l’Hypothèse de la Complexité Dérivationale, qui fait partie d’une hypothèse très générale selon laquelle l’enfant commence par maîtriser des structures moins complexes et évolue progressivement vers des structures plus complexes, dans le processus d’acquisition de sa langue maternelle (voir entre autres les travaux de JAKUBOWICZ et NASH, 2001, JAKUBOWICZ, 2003, JAKUBOWICZ, 2004 qui ont lancé les fondements de l’Hypothèse de la Complexité Dérivationale).

5.5 Conclusion intermédiaire

Pour conclure, les résultats de l’Expérience 1 en portugais sont similaires à ceux de l’Expérience 1 en français, en particulier en ce qui concerne l’avantage pour les relatives et les types d’erreurs les plus fréquents. Mais, d’un point de vue quantitatif, les enfants portugais ont fait une bien meilleure performance que les enfants français. Cela peut s’expliquer par le fait que les enfants portugais connaissent déjà ‘o que’, qui peut introduire des questions directes. Nous savons en effet que les interrogatives racines introduites par ‘o que’ sont beaucoup plus précoces que les mêmes interrogatives enchâssées (voir SOARES, 2006). Par ailleurs, les données d’acquisition prises globalement dans le cas du portugais ne vont pas clairement dans le sens de l’hypothèse de l’incorporation. Le processus de grammaticalisation ne concernerait que les interrogatives indirectes, et non pas les relatives, comme suggéré dans la discussion.

Comme pour le français, j’attribue cet avantage pour les relatives à la présence de l’élément nominal ‘o’, sachant qu’une explication alternative pourrait être que, de manière générale, les relatives sont de toute façon acquises avant les interrogatives indirectes. J’ai donc

décidé de mettre en place un deuxième test de répétition en portugais également, similaire à l'Expérience 2 française, avec pour items expérimentaux des relatives libres et interrogatives indirectes introduites par 'quem', qui est un mot wh- et qui n'est pas précédé par un élément de nature nominale. Le but, encore une fois, est de vérifier, si cet avantage des relatives se maintient avec des simples éléments wh-. L'Expérience 2 française a permis de confirmer que cet avantage ne se retrouve pas dans le cas d'éléments wh- simples. Qu'en est-il du portugais ?

5.6 Expérience 2 : test de répétition de phrases contenant des subordonnées introduites par *quem*

5.6.1 Prédiction

Comme dans le cas du français, il est important de faire une comparaison avec des subordonnées introduites par un mot wh- sans élément nominal, comme 'quem', afin de déterminer si 'o que' se comporte comme un élément wh-. Pour cette deuxième expérience, nos attentes sont les suivantes :

1. Par rapport à l'Expérience 1, les enfants portugais de la même tranche d'âge devraient être encore plus performants pour la répétition de phrases contenant des subordonnées introduites par 'quem', étant donné que le terme introducteur est un simple élément wh- ;
2. Il ne devrait pas y avoir de différence, en termes de performance, entre les deux types de subordonnées, à savoir relatives et interrogatives indirectes, ou du moins pas d'avantage pour les relatives. En effet, comme dans le cas du français, nous faisons l'hypothèse que l'avantage pour les relatives relevé dans l'Expérience 1 était dû à la présence de l'élément D 'o'.

5.6.2 Design expérimental

5.6.2.1 Participants

Ont participé à cette deuxième expérience 44 enfants monolingues portugais, tous présentant un développement typique du langage, et âgés entre 3 et 6 ans. Ils ont tous été testés à Santa Iria de Azoia, près de Lisbonne, au sein de deux écoles, comme pour l'Expérience 1. Un accord parental a été obtenu au préalable pour chacun des enfants ayant participé. Je précise que quelques-uns des enfants avaient déjà participé à l'Expérience 1, quelques mois plus tôt, mais la plupart ne se rappelaient pas forcément du 'jeu' ou des expérimentatrices.

5.6.2.2 Matériel

Les conditions expérimentales sont les mêmes que pour l'Expérience 2, en français, présentée dans le Chapitre 4 : seul le facteur TYPE DE PHRASE a été retenu (relatives *versus*

interrogatives indirectes). Les items contiennent tous des dépendances sujet. Il y a 8 items par type de phrase, donc 16 items en tout ; également 10 *fillers* et toujours 5 items d'entraînement.

Le matériel correspond à une adaptation du matériel français faite avec l'aide de linguistes, locutrices natives du portugais européen (Carla SOARES, Margarida TOMAZ, et Maria LOBO). Il ne s'agit pas vraiment d'une traduction étant donné que l'équivalent portugais de la plupart des mots de vocabulaire choisis pour les items français était plus long. Cela posait problème puisque la longueur a été contrôlée. Les items ont tous été construits de façon à faire 13 syllabes, comme ceux de l'Expérience 1. Outre cette contrainte de longueur, il fallait bien entendu faire attention à ce que les phrases paraissent naturelles et à ce que le vocabulaire utilisé soit maîtrisé/connu par des enfants âgés entre 3 et 6 ans. Comme pour l'expérience française, le futur a été utilisé pour les relatives libres, afin de favoriser l'interprétation de choix libre, plus naturelle, et le prétérit pour les interrogatives indirectes.

Voilà deux exemples d'items pour chacune des deux conditions expérimentales (voir Annexe E pour la liste complète des items de cette expérience).

1. Relatives libres

- (47) O avô castigará quem partir o vaso.
'Le grand-père punira qui cassera le vase.'
- (48) A Beatriz abraçará quem levar gomas.
'Béatrice prendra dans ses bras qui portera des bonbons.'

2. Interrogatives indirectes

- (49) O senhor adivinhou quem partiu o vaso.
'Le monsieur a deviné qui a cassé le vase.'
- (50) O diretor adivinhou quem levou gomas.
'Le directeur a deviné qui a porté des bonbons.'

Les *fillers* ont également été adaptés à partir des phrases françaises. Voilà deux exemples :

- (51) O jogo é divertido.
'Le jeu est amusant.'
- (52) A Matilde adora massa.
'Mathilde adore les pâtes.'

Idem pour les items d'entraînement, qui sont aussi longs et complexes que les items expérimentaux (13 syllabes).

- (53) O professor acha que a Clara desenha bem.
'Le professeur trouve que Clara dessine bien.'

5.6.2.3 Procédure

La procédure est la même que celle pour le test de répétition français correspondant (cf. Chapitre 4, Section 4.7.2.3). Par contre, comme dans le cas de L'Expérience 1 en portugais, c'est une locutrice native du portugais, Margarida TOMAZ (de l'Universidade NOVA, à Lisbonne) qui a fait passer le test aux enfants. Je l'ai accompagnée dans les écoles et j'ai supervisé la passation de l'expérience.

5.6.3 Résultats

Dans cette sous-section, je vais présenter les résultats de ce deuxième test de répétition, toujours en deux temps : l'analyse quantitative dans un premier temps, puis l'analyse qualitative dans un deuxième temps.

5.6.3.1 Analyse quantitative : la performance des enfants

Comme pour l'Expérience 1, les données ont été analysées avec des modèles linéaires mixtes généralisés, avec la fonction de lien *logit* (BAAYEN, DAVIDSON et BATES, 2008 et JAEGER, 2008), en utilisant le paquet *lme4* (BATES, MÄCHLER, BOLKER et WALKER, 2015) dans R (RCORETEAM, 2017).

La phrase de réponse que les enfants produisaient à chaque essai a été codée comme une variable dépendante binaire : 1 si la répétition était correcte et 0 s'il y avait une ou plusieurs erreurs. Pour déterminer cela, je me suis principalement centrée sur la répétition correcte de 'quem'. Mais si 'quem' est bien répété, avec un ordre des mots correct, mais que la phrase est agrammaticale pour une raison indépendante, j'ai compté 0.

Le facteur TYPE DE PHRASE a été codé 0.5 pour interrogative indirecte et -0.5 pour relative. De plus, la covariable ÂGE a été insérée dans le modèle en tant que variable continue centrée. Nous avons en effet inclus tous les participants dans un seul groupe et vérifié s'il y avait un effet de l'âge sur leur performance. Dans la partie 'effets fixes', nous avons estimé les effets principaux, ainsi que l'interaction éventuelle de TYPE DE PHRASE et ÂGE. Dans la partie 'effets aléatoires', nous avons inclus seulement un intercept pour les sujets et un pour les items, afin d'éviter un suparamétrage, et par conséquent, des résultats de modèle qui seraient ininterprétables (BATES, KLIEGL, VASISHTH et BAAYEN, 2015, MATUSCHEK et al., 2017).

Le Tableau 5.8 indique le score moyen, entre 0 et 1, pour chaque type de phrase, à savoir les items d'entraînement, les items expérimentaux, et les distracteurs.

On peut remarquer que la performance pour les items expérimentaux est très élevée et très proche de celle pour les items d'entraînement. Les enfants portugais n'ont pratiquement pas de difficultés à répéter des phrases contenant des subordonnées introduites par 'quem'. Quant à la performance pour les distracteurs, elle atteint pratiquement le score maximal.

Type de phrase	Score moyen
items d'entraînement	0.97
items expérimentaux	0.93
distracteurs	0.996

TABLE 5.8 – Scores moyens en fonction du type de phrase (Expérience 2).

La proportion de répétitions correctes à travers les deux conditions, et en fonction de l'âge, est illustrée dans la Figure 5.9.

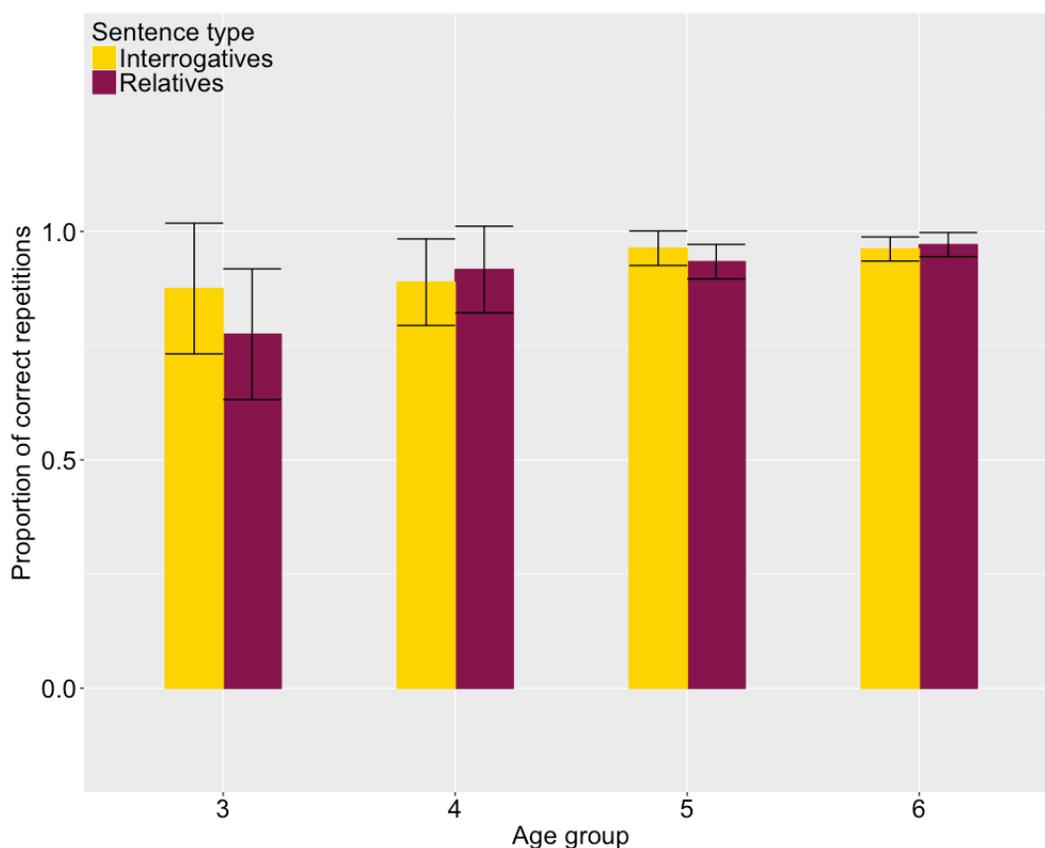


FIGURE 5.9 – Proportion de répétitions correctes (avec erreur-type ± 1). Sur l'axe x, l'âge est représenté en années. Les barres jaunes indiquent la performance pour les interrogatives indirectes, et les roses celle pour les relatives libres.

Ce qui est notable est que la performance des enfants portugais pour ce deuxième test de répétition est très bonne (score moyen de 0.93), même pour les enfants les plus jeunes. De ce fait, on ne voit qu'une légère évolution en fonction de l'âge car ces subordonnées introduites par 'quem' ont correctement été répétées dès le plus jeune âge. On trouve quand même un

effet principal de l'ÂGE qui est significatif (coeff=0.74, SE=0.30, $z=2.51$, $p=0.012$). Il semblerait que les interrogatives indirectes soient un peu plus faciles que les relatives pour les enfants de 3 ans, mais cette différence s'efface par la suite. Aucun effet du TYPE DE PHRASE n'a été relevé (coeff=0.10, SE=0.44, $z=0.22$, $p=0.827$). Aucune interaction n'est visible entre le TYPE DE PHRASE et l'ÂGE.

5.6.3.2 Analyse qualitative : les types d'erreur

En ce qui concerne l'analyse qualitative, je vais présenter les quelques types d'erreurs observés, mais, étant donné que le taux d'erreur n'est que de 6,82% (48 erreurs sur un total de 704 observations), il n'y a que très peu d'occurrences pour chaque type d'erreur. Je ne vais donc pas représenter la proportion pour chaque type dans un graphique, ni faire d'analyse statistique ici.

11 erreurs sur les 48 correspondent à l'omission du mot wh- 'quem' introduisant la proposition subordonnée.

- (54) *Phrase cible* : O senhor castigará quem rasgar o livro.
 'Le monsieur punira qui déchirera le livre.'
Phrase réponse : O senhor castigará rasgar o livro. [David, 5;5]
 'Le monsieur punira déchirera le livre.'
- (55) *Phrase cible* : O Duarte aplaudirá quem comprar um bolo.
 'Édouard applaudira qui achètera un gâteau.'
Phrase réponse : O Duarte aplaudirá comprou um bolo. [Matilde, 5;11]
 'Édouard applaudira achètera un gâteau.'

Plusieurs travaux ont montré que les enfants omettaient le mot wh- dans les questions racines, et ce, dans plusieurs langues (voir notamment YAMAKOSHI, 2002). Nous voyons ici que c'est également le cas dans les interrogatives indirectes, en portugais mais aussi en français (voir résultats de l'Expérience 2 dans le Chapitre 4).

Parfois, les enfants changeaient le verbe de la proposition matrice, ce qui pouvait impliquer un changement de structure (8 erreurs sur 48), ou bien conduire à une phrase inacceptable.

- (56) *Phrase cible* : A professora perguntou quem comeu tudo.
 'La maîtresse a demandé qui a tout mangé.'
Phrase réponse : A professora conta quem comeu tudo. [Martim, 3;7]
 'La maîtresse raconte qui a tout mangé.'
- (57) *Phrase cible* : O avô perguntou quem pintou o desenho.
 'Le grand-père a demandé qui a fait le dessin.'
Phrase réponse : O avô disse quem pintou o desenho. [Lourenço, 3;4]
 'Le grand-père a dit qui a fait le dessin.'

Une autre erreur consiste à remplacer ‘quem’ par la préposition ‘a’, ce qui conduit à un résultat agrammatical (5 cas parmi les 48 erreurs).

- (58) *Phrase cible* : O professor aplaudirá quem comer tudo.
 ‘Le maître applaudira qui mangera tout.’
Phrase réponse : *O professor aplaudirá a comer tudo. [Carolina, 4;7]
 ‘*Le maître applaudira à tout manger.’
- (59) *Phrase cible* : O Duarte aplaudirá quem comprar um bolo.
 ‘Édouard applaudira qui achètera un gâteau.’
Phrase réponse : *O Duarte aplaudirá a comprar o bolo. [Rafael, 6;2]
 ‘*Édouard applaudira à acheter un gâteau.’

Dans 4 cas sur 48, le mot wh- ‘quem’ a été remplacé par le complémenteur ‘que’, ce qui peut conduire à un résultat agrammatical dans certains contextes.

- (60) *Phrase cible* : A senhora beijará quem lavar a loiça.
 ‘La dame embrassera qui fera la vaisselle.’
Phrase réponse : *A senhora beijará que lavar a loiça. [Gonçalo, 5;8]
 ‘*La dame embrassera que fera la vaisselle.’
- (61) *Phrase cible* : O senhor adivinhou quem partiu o vaso.
 ‘Le monsieur a deviné qui a cassé le vase.’
Phrase réponse : *O senhor adivinhou que partiu o vaso. [Isabél, 4;9]
 ‘*Le monsieur a deviné qu’a cassé le vase.’

Un autre type d’erreur que je peux citer, et qui a 4 occurrences, est l’emploi d’une forme wh- complexe : ‘quem é que’, l’équivalent de ‘qui est-ce qui’ en français. Ces occurrences concernent seulement les interrogatives indirectes, pas les relatives. Cette stratégie syntaxique conduit à un résultat grammatical.

- (62) *Phrase cible* : O diretor adivinhou quem levou gomas.
 ‘Le directeur a deviné qui a porté des bonbons.’
Phrase réponse : O diretor adivinhou quem é que levou gomas. [Gabriel, 5;5]
 ‘Le directeur a deviné qui est-ce qui a porté des bonbons.’
- (63) *Phrase cible* : A senhora perguntou quem lavou a loiça.
 ‘La dame a demandé qui a fait la vaisselle.’
Phrase réponse : A senhora perguntou quem é que lavou a loiça. [Catarina, 6;5]
 ‘La dame a demandé qui est-ce qui a fait la vaisselle.’

Je termine avec deux autres types d’erreurs peu fréquents : l’omission du verbe de la proposition matrice (64) et l’emploi d’une subordonnée introduite par ‘se’ (si en français) au lieu de la subordonnée introduite par ‘quem’ (65).

- (64) *Phrase cible* : A Beatriz abraçará quem levar gomas.
 'Béatrice prendra dans ses bras qui portera des bonbons.'
Phrase réponse : *A Beatriz quem levar gomas. [Matilde, 5;11]
 '*Béatrice qui portera bonbons.'
- (65) *Phrase cible* : A menina perguntou quem rasgou o livro.
 'La fille a demandé qui a déchiré le livre.'
Phrase réponse : A menina perguntou se a menina rasgou o livro. [Lourenço, 3;4]
 'La fille a demandé si la fille a déchiré le livre.'

On trouve également quelques autres types d'erreurs isolés, pour lesquels je n'ai relevé qu'une seule occurrence, comme, par exemple, l'emploi de 'que ninguém' (*que personne* en français) à la place de 'quem'.

- (66) *Phrase cible* : O menino não soube quem trouxe o presente.
 'Le garçon ne sait pas qui a porté le cadeau.'
Phrase réponse : O menino não sabia que ninguém trouxe o presente.
 'Le garçon ne savait pas que quelqu'un avait porté le cadeau.' [Madalena, 4;9]

5.6.3.2.1 Un problème de cafouillage concernant les verbes matrices Il est important de noter, que, dans le cas des relatives libres, items dans lesquels le verbe matrice est au futur, les enfants portugais avaient du mal à prononcer ces verbes, et ce, assez fréquemment. Il s'agit des formes 'beijará' (*embrassera*), 'castigará' (*punira*), 'abraçará' (*prendera dans ses bras*) et 'aplaudirá' (*applaudira*). Plus précisément, ils prononçaient de manière très brouillonne ces verbes, mais cela n'a pas été compté comme une erreur, à partir du moment où tout le reste de la phrase était répété correctement. Plusieurs explications sont possibles :

1. Ces formes verbales sont tout simplement difficiles à prononcer pour un enfant, du fait de leur longueur (3-4 syllabes) [problème phonologique].
2. Les enfants connaissent les verbes 'beijar', 'castigar', 'abraçar' et 'aplaudir', mais c'est la forme du futur qu'ils ne connaissent pas et donc cela les perturbe [problème morphologique].
3. Les enfants ne connaissent pas le sens de ces verbes, et donc ils reproduisent les sons qu'ils ont entendu de façon approximative sans y associer un sens [problème lexical].

Contrairement à ce que l'on peut penser, ces quatre verbes sont très peu fréquents en portugais dans la production spontanée, d'après les données de CHILDES, tous corpus confondus (de 3 à 6 ans seulement, puisque cela correspond à la tranche d'âge des participants au test de répétition), ce qui est surprenant. J'ai trouvé seulement 6 occurrences de 'beijar', 5 de 'abraçar' et aucune pour 'aplaudir' et 'castigar'. Il est vrai que ces verbes peuvent avoir des équivalents périphrastiques, probablement plus fréquents, qui sont, respectivement, 'dar um beijo' (*faire un bisou*), 'dar um abraço' (*faire un câlin*), 'dar um castigo' (*donner une punition*) et 'bater palmas' (*claquer des mains*).

	BEIJAR		CASTIGAR		ABRACAR		APLAUDIR	
	VERB_ONLY	VERB_PLUS	VERB_ONLY	VERB_PLUS	VERB_ONLY	VERB_PLUS	VERB_ONLY	VERB_PLUS
1	0 0		0 0		0 0		0 0	
2	0 0		0 0		0 0		0 0	
3	0 0		0 0		0 1		1 0	
4	0 0		0 0		0 0		2 0	
5	0 0		1 1		0 1		0 2	
6	0 0		0 0		0 0		1 0	
7	0 0		1 0		0 0		2 0	
8	0 0		0 0		0 0		0 0	
9	0 0		0 0		1 0		0 0	
10	0 0		1 0		0 0		2 0	
11	0 0		1 0		1 0		1 0	
12	0 0		0 0		0 0		1 0	
13	0 0		0 0		0 0		0 0	
14	0 0		0 0		0 0		0 0	
15	0 0		0 0		0 0		2 0	
16	0 0		0 0		0 0		1 0	
17	0 0		0 0		0 0		0 0	
18	0 0		0 0		0 0		1 0	
19	0 0		0 0		0 0		0 0	
20	0 0		0 0		0 0		1 0	
21	0 0		1 1		0 0		1 0	
22	1 0		2 0		0 0		2 0	
23	0 0		0 0		0 0		1 0	
24	0 0		0 0		0 0		1 0	
25	0 0		0 0		0 0		1 0	
26	0 0		0 0		0 0		1 1	
27	0 0		0 0		0 0		0 0	
28	0 0		0 0		0 0		0 0	
29	0 0		0 0		0 0		0 0	
30	0 0		0 0		0 0		0 0	
31	0 0		0 0		0 0		1 0	
32	0 0		0 0		0 0		0 0	
33	0 0		0 0		0 0		0 0	
34	0 0		0 0		0 0		0 0	
35	1 0		1 1		1 1		2 0	
36	0 0		0 0		0 0		0 0	
37	0 0		0 0		0 0		0 0	
38	0 0		0 0		0 0		1 0	
39	0 0		1 1		0 0		0 1	
40	1 0		2 0		1 0		1 0	
41	0 0		0 0		0 0		1 0	
42	0 0		0 0		0 0		0 0	
43	0 0		1 0		0 0		1 0	
44	0 0		0 0		0 0		2 0	
TOTAL	3 0		12 4		4 3		31 4	
	3 errors in total (88 REL in total)		16 errors in total (88 REL in total)		7 errors in total (88 REL in total)		35 errors in total (88 REL in total)	

FIGURE 5.10 – Nombres d’occurrences de cafouillage concernant les 4 verbes matrices introducteurs divisés en deux types, VERB.PLUS et VERB.ONLY pour chacun des 44 enfants.

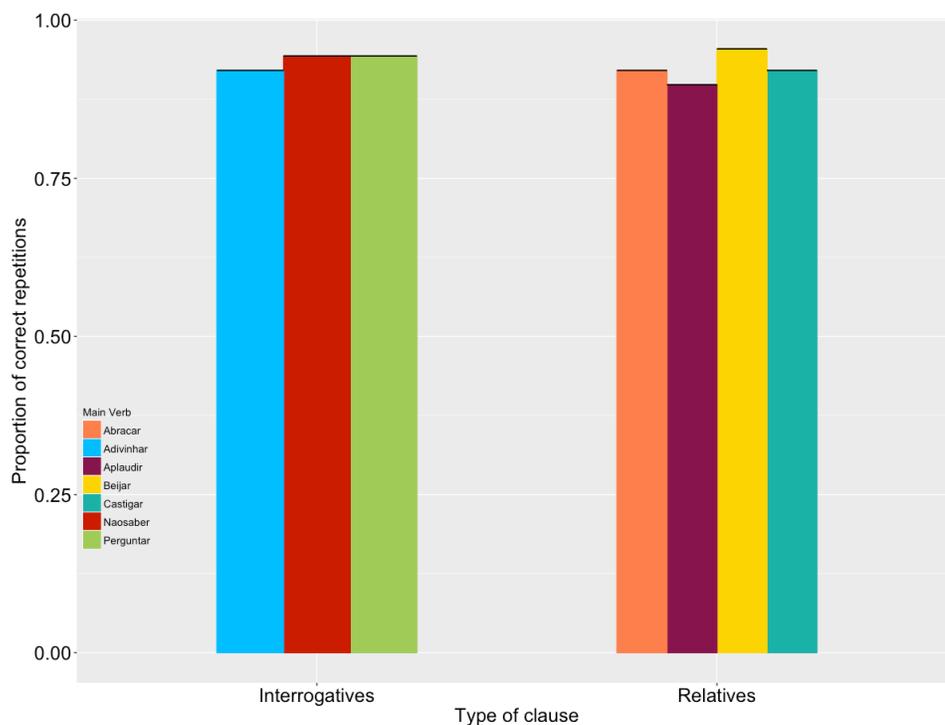


FIGURE 5.11 – Proportion de répétitions correctes pour chaque verbe matrice, selon le type de structures.

Il est important de vérifier si, cette prononciation très brouillonne du verbe est accompagnée d'une autre erreur dans la phrase de façon systématique. Autrement dit, est-ce que ces formes verbales, qui semblent être difficiles pour les enfants entre 3 et 6 ans, ont un effet boule de neige, et entraînent donc une ou plusieurs autres erreurs lors de la répétition des items expérimentaux ? Sur les 352 relatives au total, j'ai relevé 61 cas dans lesquels les enfants ont prononcé de façon très approximative le verbe de la proposition matrice. Seulement 11 cas sur ces 61 sont accompagnés d'une autre erreur, plus loin dans la phrase. Le détail en fonction de chaque verbe est représenté dans la Figure 5.10.

Comme on peut le voir sur la Figure 5.10, les verbes 'beijar' et 'abraçar' sont les deux verbes qui posent le moins de problème, et ce sont aussi les deux verbes pour lesquels on a trouvé quelques occurrences dans les données de production spontanée dans CHILDES. 'Aplaudir' est clairement le verbe qui pose le plus de problème avec 35 erreurs sur 88 relatives sélectionnées par ce verbe. Mais, ce qui est important est que pour 'aplaudir' mais aussi pour 'castigar', le nombre d'erreurs sur le verbe seulement (sur le plan de la prononciation) est nettement supérieur à celui impliquant une autre erreur. C'est pour cela que j'ai décidé de compter 1 (donc répétition correcte) dans les cas où seul le verbe matrice était prononcé de façon approximative mais que tout le reste a bien été répété, en particulier la séquence

‘quem’, qui nous intéresse ici. Bien sûr, je suis consciente qu’il s’agit d’une solution trouvée subséquemment, qui se base sur la plausibilité que ce cafouillage soit répandu indépendamment d’autres types d’erreurs. Il aurait sûrement mieux valu vérifier, avant l’expérience, que les enfants maîtrisaient le matériel lexical utilisé dans les items expérimentaux.

Une fois cette solution adoptée, regardons tout de même s’il y a de grandes différences de performance entre les verbes, ou non. Comme on peut le voir sur la Figure 5.11, les différences de performance entre les différents verbes matrices sont minimales.

5.6.4 Discussion

Les résultats de cette deuxième expérience révèlent que les enfants portugais, entre 3 et 6 ans, ont moins de difficultés à répéter des phrases contenant des subordonnées introduites par ‘quem’ que par ‘o que’. Comme nous l’avons déjà fait remarquer pour le français, dans les deux expériences en portugais, les enfants ont tendance à omettre le mot (ou la séquence) subordonnant(e) (12% des erreurs correspondaient à l’omission de ‘o que’ dans l’Expérience 1, 22,9% des erreurs correspondaient à l’omission de ‘quem’ dans l’Expérience 2). Ce pourcentage d’omission est presque deux fois moins important dans le cas de ‘o que’ car les enfants omettent fréquemment seulement l’élément ‘o’, et non pas toute la séquence (54,5% des erreurs correspondaient à l’omission de ‘o’ dans l’Expérience 1). Par ailleurs, une stratégie qui revient dans les deux tests est la stratégie ‘é que’ : 13,5% des erreurs correspondaient à l’emploi de ‘o que é que’ dans l’Expérience 1, 8,3% des erreurs correspondaient à ‘quem é que’ dans l’Expérience 2.

En ce qui concerne la différence cruciale entre les deux types de structures, aucun avantage pour les relatives n’est relevé dans l’Expérience 2 : les deux structures semblent acquises de façon parallèle. Cela ne semble donc pas être le cas que les relatives soient acquises avant les interrogatives indirectes lorsque les subordonnées en question sont introduites par de simples éléments *wh*-. L’avantage pour les relatives dans la première expérience était dû, comme nous l’avons postulé, à la présence de l’élément nominal ‘o’.

5.7 Conclusion

Pour conclure, l’Expérience 1 nous a permis d’éclaircir le statut syntaxique de ‘o que’ en portugais. En particulier, nous avons pu nous rendre compte que l’opposition entre le français et le portugais n’était pas si nette, et qu’il était peut-être trop hâtif de considérer qu’il y a absence d’incorporation dans un cas et processus de grammaticalisation/d’incorporation dans l’autre cas. L’analyse quantitative et qualitative des résultats nous a permis de nuancer cela, postulant que l’incorporation n’avait lieu que dans les interrogatives indirectes en portugais. ‘O que’ n’est donc pas équivalent à l’anglais ‘what’, mais il est entre le français ‘ce que’ et l’anglais ‘what’. Comme dans le cas du français, l’avantage pour les relatives n’a été relevé que dans l’Expérience 1, et pas dans l’Expérience 2. Ceci vient confirmer le

fait que, de manière générale, les relatives ne sont pas forcément préférées comparées aux interrogatives indirectes : elles sont plus faciles à répéter pour les enfants seulement lorsqu'il y a un élément D ('o'). Cela nous permet également de conclure que 'o que' n'est pas analysé par l'enfant de la même façon qu'un mot wh- simple ('quem').

5.8 L'Hypothèse de la Hiérarchie du Poids de la Tête

Les résultats des tests de répétition dans les deux langues, en français et en portugais, ont montré que les enfants étaient plus performants dans la répétition de phrases contenant des subordonnées introduites par un mot *wh-* sans élément D ('qui' en français et 'quem' en portugais) que dans la répétition de phrases contenant des subordonnées introduites par une séquence D+que ('ce qu-' en français et 'o que' en portugais). Un point important que je souhaite souligner est que cet avantage pour 'qui'/'quem' est valable même pour les dépendances sujet. On peut le voir sur la Figure 5.12, où seuls les résultats pour les dépendances sujet sont représentés.

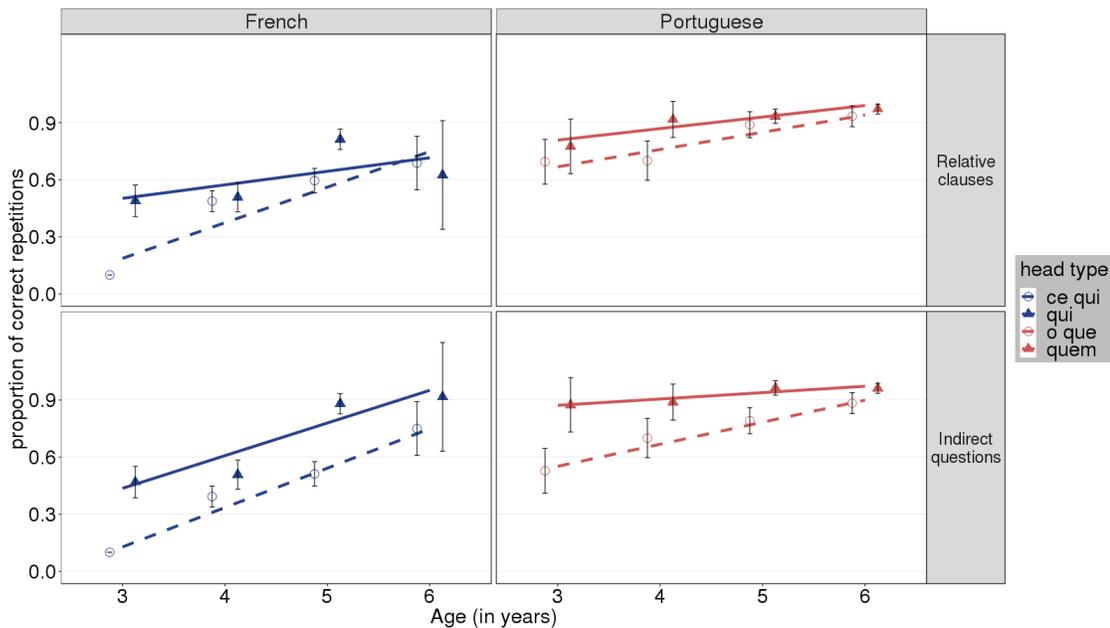


FIGURE 5.12 – Proportion de répétitions correctes (avec erreur-type ± 1) pour les dépendances sujet. Sur l'axe x, l'âge est représenté en années. Les panneaux de gauche montrent la performance des enfants français (bleu) et ceux de droite la performance des enfants portugais (rouge). Les panneaux du haut montrent la performance pour les relatives et ceux du bas celle pour les interrogatives indirectes (lignes continues pour 'qui'/'quem', c'est-à-dire pour les relatives libres, et lignes discontinues pour 'ce qui'/'o que', c'est-à-dire pour les relatives à tête allégée).

Ainsi, de manière générale, peu importe la position de la lacune, 'qui/quem' est plus facile que 'ce qu-'/'o que' pour les enfants français/portugais. La question que l'on peut alors se poser désormais est la suivante : pourquoi 'qui'/'quem' est-il plus facile pour les enfants que 'ce qu-'/'o que' ?

Une des différences saillantes entre 'qui'/'quem' d'une part et 'ce qu-'/'o que' d'autre part est que les premiers ont un trait [+animé] alors que les deuxièmes ont un trait [-animé]. On pourrait alors penser que le caractère animé de l'élément wh- facilite la répétition/production/compréhension de la dépendance par les enfants. Or, dans mon mémoire de Master (KONRAD, 2015 (non publié)), les résultats du test de compréhension (tâche d'identification d'images) portant sur les relatives pleines objet (introduites par une tête de type D+N), les relatives pronominales objet (introduite par une tête pronominale de type 'celui') et les relatives libres objet (introduites par le mot wh- 'qui') que j'ai mis en place semblent discréditer cette hypothèse. En effet, j'ai montré que les enfants âgés entre 3 et 6 ans comprenaient mieux les relatives libres introduites par 'qui' que les relatives pronominales introduites par 'celui', 'qui' et 'celui' ayant tous les deux un trait [+animé] dans les items que j'ai construits. Ainsi, 'qui' est plus facile que 'celui' et plus facile que 'ce qu-'. Le caractère [+animé] ne peut donc pas expliquer à lui seul la meilleure performance pour 'qui'.

Une autre propriété différenciant 'qui'/'quem' d'une part et 'ce qu-'/'o que' d'autre part est leur « poids syntaxique »⁶. Par « poids syntaxique » j'entends que plus un élément a de structure syntaxique, plus il est lourd. Et inversement, moins un élément a de structure syntaxique, plus il est léger.⁷ Du fait de la présence de l'élément D 'ce'/'o', 'ce qu-'/'o que' est plus lourd d'un point de vue syntaxique que 'qui'/'quem'. On pourrait avancer l'hypothèse que plus l'élément déplacé est lourd d'un point de vue syntaxique dans les dépendances A-barre, plus la dépendance est difficile à répéter, à produire, à comprendre pour l'enfant.

C'est ce que j'appelle « L'Hypothèse de la Hiérarchie du Poids de la Tête » (en anglais *The Head-Heaviness Hierarchy Hypothesis*). Cette hypothèse permet de rendre compte du *pattern* d'acquisition des différents types de relatives.

(67) **L'Hypothèse de la Hiérarchie du Poids de la Tête**

Plus la tête d'une relative est légère d'un point de vue syntaxique, plus la relative est facile à produire et à comprendre pour l'enfant.

DP+NP > DP > ∅

Relatives pleines > Relatives à tête allégée > Relatives libres

Ainsi, les relatives pleines sont plus difficiles à produire et à comprendre pour l'enfant que les relatives à tête allégée, qui sont elles-mêmes plus difficiles que les relatives libres. Cette hypothèse permet de rendre compte à la fois des résultats de mon mémoire de Master (KONRAD, 2015 (non publié)) et des résultats présentés dans les Chapitres 4 et 5 de la présente thèse.

Il est intéressant de remarquer que selon l'hypothèse de FRIEDMANN, BELLETTI et RIZZI, 2009 aussi, les relatives pleines devraient être plus difficiles que les relatives libres pour

6. Sur la notion de « poids grammatical », voir notamment WASOW, 1997.

7. Pour une vision alternative de la 'légèreté' en syntaxe, voir ABEILLÉ et GODARD, 2004, ABEILLÉ, 2008.

l'enfant. Mais leur hypothèse ne concerne que les dépendances objet avec un sujet avec restriction lexicale. De plus, ils ne s'attendraient pas à une différence entre relatives à tête allégée et relatives libres, étant donné que dans les deux cas l'objet déplacé est [-NP] (sans restriction lexicale). Pourtant, les résultats de mes expériences ont montré que les relatives libres étaient plus faciles pour les enfants que les relatives à tête allégée, et que cette différence est valable pour les dépendances objet mais aussi pour les dépendances sujet. Une autre différence entre les deux hypothèses concerne la nature même des éléments entrant en jeu. Selon l'hypothèse de FRIEDMANN, BELLETTI et RIZZI, 2009, la difficulté pour certaines relatives objet est perçue en termes de partage de traits entre l'objet déplacé et le sujet, le sujet agissant comme intervenant s'il partage un ou des traits avec le sujet. Mais, selon mon hypothèse, qui est décrite en termes de poids syntaxique, c'est le nombre de couches fonctionnelles pour la tête de la relative qui est pertinent.

Dans les tests de répétition que j'ai mis en place, je ne me centrais pas que sur les relatives. En effet, la moitié des items étaient des interrogatives indirectes. On peut alors se demander si une hiérarchie similaire pourrait être appliquée aux interrogatives indirectes. On suppose que oui. Si cela est vrai, on s'attend à ce que les interrogatives indirectes introduites par la séquence D+C ('ce que') soient plus difficiles que celles introduites par un syntagme wh- ('quel livre' / 'quelle fille'), qui seraient elles-mêmes plus difficiles que celles introduites par un mot wh- seul ('qui'). Il est évident qu'une séquence D+C et qu'un syntagme wh- sont plus lourds d'un point de vue syntaxique qu'un mot-wh- seul.

(68) DP+CP > CP[Dwh+N] > CP[wh]

Il est vrai que, depuis PESETSKY, 1987, nous faisons la distinction entre deux types de syntagmes wh- d'un point de vue sémantico-pragmatique, les syntagmes wh- reliés au discours (*D-linked*) d'une part et ceux qui ne sont pas reliés au discours (*non D-linked*) d'autre part. Les premiers, tels que 'quelle fille', impliquent l'existence d'un ensemble d'entités familières du type dénoté par le nom (un ensemble de filles déjà connues dans le cas de 'quelle fille'). Les seconds, tels que 'qui', n'ont pas d'implication particulière à propos d'entités déjà familières. Et il est vrai que les deux types de syntagmes wh- en question ne se comportent pas de la même manière sur le plan syntaxique. En particulier, les syntagmes wh- reliés au discours échappent aux Effets de Supériorité (*Superiority Effects*), contrairement aux syntagmes wh- non reliés au discours (PESETSKY, 1987), qui eux y obéissent. Ce contraste est illustré en (69).

- (69) a. John asked **which book**_{*i*} which man read *t*_{*i*}.
 b. *John asked **what**_{*i*} who read *t*_{*i*}.

Les Effets de Supériorité exigent que ce soit le syntagme wh- le plus haut dans la structure qui se déplace, lorsqu'il y a deux syntagmes wh-. Comme on le voit en (69b), le déplacement de 'what' est une violation de la Supériorité : c'est 'who', qui est plus haut, qui doit se déplacer. Par contre, le fait de déplacer 'which book', qui est plus bas, au lieu de 'which

man' en (69a) ne rend pas la phrase agrammaticale. On observe également des différences entre les deux types de syntagmes *wh-* sur le plan du traitement, avec une préférence pour les syntagmes *wh-* reliés au discours (voir par exemple FRAZIER et CLIFTON, 2002).

Connaissant ces faits, il semblerait alors étonnant que les questions indirectes introduites par un syntagme de type 'quelle fille' soient plus difficiles pour les enfants que celles introduites par le mot *wh-* 'qui', comme le prédirait mon hypothèse (voir hiérarchie en (68)). Mais des études en acquisition du langage, dans le cadre de la Minimalité Relativisée, ont montré que les questions *wh-* objet directes avec sujet [+NP] (avec restriction lexicale) étaient plus difficiles à comprendre pour les enfants lorsque l'objet était de type *Dwh+N* (BENȚEA et DURRLEMAN, 2013).

- (70) a. **Quel garçon** la fille pousse ?
 b. **Qui** la fille pousse ?

Encore une fois, le contraste entre (70a) et (70b) s'explique en termes d'intervention, le sujet et l'objet étant tous les deux avec restriction lexicale (trait [+NP]) en (70a) mais pas en (70b).

Il serait donc intéressant, dans de futurs travaux, de comparer les différents types d'interrogatives indirectes en français et dans d'autres langues, variant selon la lourdeur de l'introducteur, pour voir si ma prédiction est vérifiée et pour déterminer si mon hypothèse peut vraiment s'étendre ou non aux interrogatives indirectes, et ce, indépendamment de la position de la lacune. Il serait également intéressant de vérifier si cet effet de la hiérarchie du poids de la tête explique en fait les phénomènes d'intervention décrits par FRIEDMANN, BELLETTI et RIZZI, 2009, et se pose donc comme une approche alternative, ou bien s'il s'agit d'un effet qui s'ajoute à celui-ci, et est pour cela visible aussi quand la Minimalité Relativisée ne s'applique pas, c'est-à-dire dans les dépendances sujet.

Quatrième partie

Traitement des ambiguïtés syntaxiques impliquant ‘ce que’

Chapitre 6

Expériences de lecture par présentation auto-segmentée : évitement des lacunes et similarité de traitement entre relatives à tête allégée et interrogatives indirectes ¹

Nous venons de voir dans la partie précédente portant sur l’acquisition que les enfants français âgés entre 3 et 6 ans présentaient moins de difficultés à répéter des relatives que des interrogatives indirectes introduites par ‘ce qu-’, mais qu’ils ne présentaient pas de différence quand la structure à répéter était introduite par ‘qui’. Nous allons déterminer dans le présent Chapitre si l’on décèle également une préférence de traitement pour les relatives en ‘ce qu-’ chez les adultes ou non.

Nous savons que dans les cas d’ambiguïté entre une dépendance *filler-gap* et une structure sans *gap*, la deuxième structure est préférée. Cette préférence est gouvernée par des principes généraux de traitement tels que le ‘Principe de Chaîne Minimale’. Nous allons voir que ce principe est pertinent pour ce qui est du traitement d’une autre ambiguïté impliquant ‘ce que’, celle entre complétives et relatives à tête allégée. Or, dans le cas de l’ambiguïté syntaxique qui nous intéresse de plus près (relatives à tête allégée et interrogatives indirectes introduites par ‘ce qu-’), nous avons deux dépendances *filler-gap*. Le Principe de Chaîne Minimale ne peut alors pas prédire de préférence pour l’une ou l’autre de ces structures.

1. J’ai présenté les résultats d’expériences détaillées dans ce Chapitre au *Workshop Ambigo* (KONRAD, DONATI et al., 2018) ainsi qu’à trois conférences internationales, *Architectures and Mechanisms of Language Processing* (AMLaP) (FOPPOLO, KONRAD, BURATTIN et al., 2018), *Incontro di Grammatica Generativa* (IGG) (FOPPOLO et al., 2019a) et *Conference on Human Sentence Processing* (CUNY) (FOPPOLO et al., 2019b).

6.1 Traitement de l'ambiguïté entre complétives et relatives

6.1.1 Le Principe de Chaîne Minimale

Les dépendances syntaxiques impliquant un mouvement ont été un sujet central dans la recherche portant sur le traitement des phrases depuis les travaux pionniers de FODOR, 1978 et CLIFTON et FRAZIER, 1989. La plupart de ces travaux ont enquêté sur le processus de recherche d'une lacune à partir du moment où un élément déplacé est rencontré au cours du traitement progressif de la phrase et ont proposé qu'une recherche active d'une lacune est immédiatement enclenchée (*the Active Filler Strategy*, CLIFTON et FRAZIER, 1989).

The Active Filler Strategy est également capturée par la deuxième proposition du *Minimal Chain Principle* (Principe de Chaîne Minimale, DE VINCENZI, 1991), énoncé en (1).

- (1) *Minimal Chain Principle*
Avoid postulating unnecessary chain members at surface structure, but do not delay required chain members.

La première proposition du Principe de Chaîne Minimale établit que les chaînes (donc les dépendances *filler-gap* ne doivent pas être postulées par l'analyseur à moins que cela ne puisse être évité étant donné les règles de la grammaire, alors que dans la seconde proposition il est affirmé que la phrase doit être analysée en assumant qu'une chaîne, au cas où elle doit être postulée, est aussi courte que possible. La justification ultime de ces principes de traitement repose sur le besoin de l'analyseur d'optimiser ses ressources mémorielles : puisque maintenir du matériel non structuré dans la mémoire à court terme est coûteux d'un point de vue cognitif, le système a recours à des stratégies qui minimisent les coûts de traitement au cours de l'analyse de la phrase, et les lacunes non saturées sont évitées autant que possible et maintenues le moins longtemps possible.

Une illustration de la deuxième proposition du Principe de Chaîne Minimale est la difficulté de traitement qui serait rencontrée au niveau du verbe *eat* ('manger') en (2), ce qui suggère que l'analyseur a initialement computed l'interprétation étrange selon laquelle la famille mange un chien (STOWE, 1986). En effet, lorsqu'il tombe sur l'élément *wh-* (*which dog*), l'analyseur reconnaît la nécessité d'une chaîne et il se lance ainsi à la recherche d'une lacune, qui est d'abord postulée après le verbe parce qu'il s'agit de la première position disponible selon la grammaire de l'anglais. Cette analyse doit ensuite être révisée quand l'analyseur rencontre la préposition *with*, d'où un effet de phrase labyrinthe (*garden-path effect*).

- (2) **Which dog** did the family eat happily with e ?

Malgré beaucoup de preuves concordantes en faveur de la deuxième proposition du Principe de Chaîne Minimale, peu de tentatives ont été faites pour corroborer sa première proposition, selon laquelle une construction avec lacune doit être évitée si une construction sans lacune est disponible au cours du traitement incrémental.² Les expériences présentées dans ce chapitre contribuent au débat en testant tout d'abord deux structures minimalement différentes, qui diffèrent seulement par la présence ou l'absence d'une chaîne, la question expérimentale étant : est-ce que le postulat d'une chaîne est moins désirable comme l'affirme la première proposition du Principe de Chaîne Minimale ? Je vais, dans la section qui suit, présenter les deux structures que nous avons testées dans les deux expériences qui seront détaillées par la suite.

6.1.2 Une ambiguïté temporaire en français avec 'ce que'

Pour réfléchir sur ce principe structural, nous allons nous pencher avant tout sur une autre ambiguïté avec 'ce que', impliquant des complétives. Nous avons vu jusqu'à maintenant que 'ce que' pouvait introduire des dépendances A-barre : des relatives et des interrogatives indirectes. Mais l'emploi de 'ce que' n'implique pas forcément la présence d'une lacune. Ainsi, il n'existe pas que deux 'ce que', mais au moins un troisième 'ce que', différent des deux premiers de par l'absence d'une lacune. Les relatives introduites par 'ce que' peuvent être ambiguës avec les interrogatives indirectes introduites par 'ce que', mais également avec les complétives introduites par 'ce que'. C'est cette deuxième ambiguïté, impliquant une structure avec lacune (relative) et une sans lacune (complétive), que je vais décrire ici.

En français, 'ce que' peut introduire une complétive comme en (3a) ou une relative comme en (3b). Cela rend le segment (3) temporairement compatible avec une construction qui ne contient pas de lacune (la complétive, 3a) et une construction qui implique une lacune dans la position d'objet direct du verbe enchâssé (la relative, 3b).

- (3) Paul fait attention à ce que...
- a. ...sa femme regarde l'ordinateur.
 - b. ...sa femme regarde sur l'ordinateur.

6.1.2.1 L'interprétation de complétive

L'une des continuations possibles de (3) est une complétive.³ Dans ce cas, 'ce que' est en quelque sorte une forme spéciale du complémenteur 'que' et est obligatoire lorsqu'il introduit une complétive enchâssée sous les prépositions 'à', 'de', et parfois 'en'.⁴

2. Voir STAUB, FOPPOLO, DONATI et CECCHETTO, 2018 pour l'anglais.

3. Pour une analyse syntaxique détaillée des complétives en P+ce+que, voir notamment ZARING, 1992 et AUTHIER et REED, 2010.

4. Notons que le statut particulier de 'à' et 'de' en français a été discuté par HUOT, 1981, CINQUE, 1990, KAYNE, 1999, entre autres, notamment dans les constructions infinitives, où il s'agirait en fait de complémen-

- (4) Cédric veille à **ce que** Julie mange davantage de légumes.
- (5) Pauline se soucie **de ce que** son mari la trompe avec une autre femme.
- (6) Le problème réside **en ce que** aucun des deux ne veut faire un effort pour aller parler à l'autre.

'Ce que' peut alterner dans ces contextes avec la variante peut-être plus commune 'le fait que'.

- (7) Cédric veille **au fait que** Julie mange davantage de légumes.
- (8) Pauline se soucie **du fait que** son mari le trompe avec une autre femme.
- (9) Le problème réside **dans le fait qu'** aucun des deux ne veut faire un effort pour aller parler à l'autre.

L'alternance avec 'que' seul est possible dans les cas de 'de ce que' et 'à ce que', mais elle est non systématique dans le cas de 'à ce que'. Notons que l'effacement de la préposition et l'effacement de 'ce' vont de pair.⁵

- (10) a. Il se plaint **de ce que** tu ne sois pas venu hier.
b. Il se plaint **que** tu ne sois pas venu hier.
- (11) a. Elle se soucie **de ce que** son chat fasse des bêtises en son absence.
b. Elle se soucie **que** son chat fasse des bêtises en son absence.
- (12) a. Paul est content **de ce que** Marie vienne.
b. Paul est content **que** Marie vienne.
- (13) a. Paul a rêvé **de ce que** Marie viendrait.
b. Paul a rêvé **que** Marie viendrait.
- (14) a. Je fais attention à **ce que** le vase ne tombe pas.
b. Je fais attention **que** le vase ne tombe pas.
- (15) a. Elle s'oppose à **ce que** tu viennes à son mariage.
b. *Elle s'oppose **que** tu viennes à son mariage.
- (16) a. Étant maniaque, il aime à **ce que** tout soit parfait.
b. Étant maniaque, il aime **que** tout soit parfait.
- (17) a. Tu t'habitueras à **ce que** ses parents viennent vous rendre visite tous les dimanches.
b. *Tu t'habitueras **que** ses parents viennent vous rendre visite tous les dimanches.

L'alternance avec 'que' seul en (15) et (17) n'est pas possible car les verbes 's'opposer à' et 's'habituer à' sont des verbes à construction prépositionnelle. La préposition 'à' est requise, et donc l'insertion de 'ce' est obligatoire.

teurs.

5. Les exemples en (12) et (13) sont tirés de GROSS, 1975 : 69-70.

Mais on peut alors se demander : pourquoi ne peut-on pas avoir 's'opposer à que', autrement dit pourquoi l'insertion de 'ce' est-elle obligatoire après la préposition ? Nous savons qu'une préposition sélectionne un DP, marqué pour le cas. La position de complément de préposition est donc une position marquée pour le cas. Or, d'après le Principe de Résistance Casuelle (*Case-Resistance Principle*, STOWELL, 1981), un CP tensé ne peut pas recevoir de cas, donc il ne peut pas apparaître dans la position de complément de préposition. Ainsi, de la même façon qu'en anglais on ne peut pas avoir (18), en français on ne peut pas avoir (19), la différence étant qu'en anglais la seule possibilité pour rendre cette phrase grammaticale est d'utiliser 'the fact that', alors qu'en français on a également la possibilité (en plus de l'emploi de 'le fait que') d'utiliser 'ce que'.⁶ En français, 'ce' ajoute une couche DP à la structure permettant ainsi de ne pas violer le Principe de Résistance Casuelle. 'Ce' serait alors une sorte de nominalisateur de proposition.

(18) *I'm ashamed of that I neglected you.

(19) *J'ai honte de que je t'ai délaissé.

Toutefois, ce Principe de STOWELL, 1981 est à nuancer. En effet, qui dit « principe » dit « caractère universel ». Or, en espagnol et en portugais (entre autres), on peut très bien avoir un CP à temps fini après une préposition. D'ailleurs, l'emploi du déterminant neutre 'lo', présent dans les relatives à tête allégée, ne peut pas s'intercaler entre la préposition et le complémenteur. Il semblerait donc que cette propriété soit sujette à une paramétrisation selon les langues.

(20) a. Me opongo a que vengas.

b. *Me opongo a lo que vengas.

De plus, en français, 'ce' n'est pas nécessaire en position sujet (21) ni en position d'objet direct (22), qui sont pourtant des positions casuelles. L'emploi de 'ce' dans ces positions conduit même à un résultat agrammatical, ce qui serait difficile à expliquer pour Stowell.

(21) a. Qu'il soit venu m'ennuie.

6. Notons qu'en français, on peut avoir un CP tensé après plusieurs prépositions telles que 'pour' ou 'depuis'.

(1) Il insiste [pour [CP que tu viennes]].

(2) Il ne vient plus me voir [depuis [CP qu'il a déménagé]].

Toutefois, les prépositions autorisées dans ces cas ne sont pas des prépositions qui assignent un cas, contrairement à 'de' et 'à'. L'insertion de 'ce' n'est donc pas requise dans ces cas-là, et conduit d'ailleurs à un résultat agrammatical.

(3) *Il insiste pour ce que tu viennes.

(4) *Il ne vient plus me voir depuis ce qu'il a déménagé.

- b. *Ce qu'il soit venu m'ennuie.
 - c. Le fait qu'il soit venu m'ennuie.
- (22)
- a. Adrien attend que Nina s'en aille.
 - b. *Adrien attend ce que Nina s'en aille.

Notons que les cas nominatif et accusatif (correspondant respectivement aux positions sujet et objet direct) sont des cas structuraux alors que le cas oblique (après une préposition) est un cas inhérent. Il est intéressant de remarquer que, si l'on adopte cette distinction entre les cas (voir CHOMSKY, 1981, CHOMSKY, 1986b), on obtient la généralisation qui suit pour le français : l'emploi de 'ce' est nécessaire dans les positions de cas inhérent, mais pas dans celles de cas structural.

Cette première lecture appartient à un registre plutôt soutenu. En effet, les complétives en P+ce+que en français ont un côté assez formel.

6.1.2.2 L'interprétation de relative

Comme nous le savons, 'ce que' peut également introduire une relative à tête allégée (voir Chapitre 1), avec 'ce' comme tête légère et 'que/qui' comme le complémenteur usuel introduisant les relatives en français. Cette construction n'appartient pas à un registre de langue en particulier et elle peut être sélectionnée par n'importe quel verbe sélectionnant un DP. Comme il s'agit d'une relative, cette structure contient bien évidemment une lacune, contrairement à la complétive.

6.1.3 Traitement de l'ambiguïté complétive/relative

7

7. Les expériences décrites dans la présente sous-section sont également présentées dans l'article KONRAD, BURATTIN et al., Soumis, que nous avons soumis dans la revue *Syntax*. Dans cet article, nous avons également étudié le traitement d'une autre ambiguïté en italien, impliquant toujours une structure sans lacune et une structure avec lacune.

- (1) Gli studenti hanno capito che...
 les étudiants ont compris QUE
 'Les étudiants ont compris que...'
- a. ...fare gli esami è difficile.
 ...faire les examens est difficile
 '...faire les examens est difficile.'
 - b. ...fare agli esami difficili
 ...faire aux examens difficiles
 '...faire aux examens difficiles.'

La continuation en a. est une complétive (pas de lacune) alors que la continuation en b. est une interrogative indirecte (donc une dépendance *filler-gap*). Cette ambiguïté est possible en italien du fait de l'ambiguïté de 'che', qui peut être soit le complémenteur (comme en a.) soit le mot wh- (comme en b.). Le complémenteur

Il existe une certaine intersection entre les verbes qui sélectionnent une complétive introduite par 'ce que' et ceux qui sélectionnent une relative ayant pour tête légère 'ce', et c'est là que l'ambiguïté temporaire peut surgir. On peut alors se demander si l'analyseur fait un pari sur une analyse plutôt qu'une autre ou s'il a une préférence pour l'une de ces analyses. Si le Principe de Chaîne Minimale (DE VINCENZI, 1991) guide le traitement, alors on s'attend à ce que les relatives soient défavorisées dans les contextes ambigus puisqu'elles impliquent une dépendance *filler-gap*.

Certains verbes, comme 'regarder', 'approuver' ou encore 'noter', ne peuvent sélectionner que des DPs (donc des relatives) mais pas des CPs (donc pas de complétives). Il est alors intéressant de comparer le traitement de relatives dans des contextes ambigus avec celui de relatives dans des contextes non ambigus. Le traitement de relatives dans des contextes non ambigus devrait être plus facile étant donné qu'il n'y a pas de compétition avec une structure sans lacune.

6.1.3.1 Expérience 1 : Test de jugement d'acceptabilité

La première expérience est un test de jugement d'acceptabilité, dans lequel nous avons comparé des phrases contenant des propositions subordonnées introduites par 'ce que' qui impliquent une dépendance *filler-gap*, comme en (23b) et (23c) avec des phrases qui contiennent également des propositions subordonnées introduites par 'ce que' mais qui n'impliquent pas de dépendance *filler-gap*, comme (23a). Afin de moduler la disponibilité d'une interprétation de complétive, nous avons manipulé le type de verbe sélectionnant 'ce que', dans le but de comparer des fragments ambigus qui peuvent avoir comme continuations une complétive ou une relative avec des fragments non ambigus ne pouvant avoir qu'une relative comme continuation.

- (23) a. Luc veille à ce que Lou organise une surprise pour la fête.
 b. Luc veille à ce que Lou organise pour la fête.
 c. Luc filme ce que Lou organise pour la fête.

Il est vrai qu'une tâche de jugement d'acceptabilité n'est pas une mesure de traitement à strictement parler, étant donné que ce type de tâche ne permet pas d'étudier le traitement des phrases en temps réel. Mais supposons qu'il y a un effet de labyrinthe dans les cas où le verbe peut à la fois sélectionner une complétive et une relative, et supposons que cet effet est si fort qu'il peut se refléter dans les jugements d'acceptabilité des locuteurs.⁸ Nous pouvons alors considérer que les jugements d'acceptabilité nous fournissent une mesure indirecte

'que' est aussi homophone au mot *wh-* en français, mais en b. il serait plus naturel d'employer la forme forte du *wh-* 'que', à savoir 'quoi' : 'Les étudiants ont compris quoi faire aux examens difficiles.' La désambiguïsation en italien se fait juste après le verbe à l'infinitif ('fare'), selon qu'il y a une préposition devant le déterminant 'gli' ou non.

8. Sur la corrélation entre jugements d'acceptabilité et difficultés de traitement, voir notamment HOFMEISTER et SAG, 2010 et PHILLIPS, 2013.

d'une difficulté de traitement. Vérifions si c'est le cas avec une tâche de ce type dans un premier temps, tâche qui est très simple et rapide à mettre en place, et nous viendrons dans un second temps appuyer nos résultats avec une mesure de traitement en temps réel, en particulier une expérience de lecture par présentation auto-segmentée.

6.1.3.1.1 Prédications On s'attend à un effet de phrase-labyrinthe (*garden-path*) enclenchant un processus de réanalyse seulement dans les cas de relatives sélectionnées par un verbe ambigu. En effet, comme il existe dans ces cas d'ambiguïté une compétition avec une structure sans lacune, on s'attend à une préférence pour la complétive, qui ne contient pas de lacune. Le coût de cette réanalyse quand la continuation est une relative, ici mesuré de façon indirecte, devrait se refléter dans les jugements des participants : en particulier, on s'attend à des notes d'acceptabilité peu élevées pour ces relatives sélectionnées par un verbe ambigu, comparées aux complétives et comparées aux relatives sélectionnées par un verbe non ambigu.

6.1.3.1.2 Design expérimental

6.1.3.1.2.1 Participants 70 adultes français (dont 53 femmes), âgés entre 18 et 73 ans ($M=27$ ans) ont participé de façon volontaire à cette expérience. Une grande partie d'entre eux ont été recrutés via le RISC (Relais d'Information en Sciences de la Cognition) ; les autres ont été recrutés via les connaissances et les réseaux sociaux.

6.1.3.1.2.2 Matériel Le questionnaire comprenait 24 items expérimentaux⁹, comportant chacun trois conditions. Chaque participant a jugé 8 items par condition, une de chaque triplet, et 24 distracteurs (voir Annexe F pour la liste complète des items expérimentaux et des distracteurs).

En ce qui concerne les cas d'ambiguïté, 8 verbes matrices ont été choisis et réutilisés chacun trois fois :

1. veiller à ce que
2. tenir à ce que
3. faire attention à ce que
4. s'opposer à ce que
5. se soucier de ce que
6. être attentif à ce que
7. s'habituer à ce que
8. venir de ce que

9. L'item 1 a été exclu pour l'analyse des résultats, car, à cause d'une erreur dans le script sur Ibx Farm, il n'était pas parallèle aux autres items.

Dans les cas de non ambiguïté, les verbes sélectionnant seulement une relative (donc, par extension, un DP) étant plus nombreux, nous avons pu utiliser 24 verbes matrices différents (condition c.). Voilà un exemple de triplet expérimental :

- (24) a. Julien s'oppose à ce que son patron organise une réunion pour le mois prochain. [V ambigu, continuation de complétive]
 b. Julien s'oppose à ce que son patron organise pour le mois prochain. [V ambigu, continuation de relative]
 c. Julien est contre ce que son patron organise pour le mois prochain. [V non ambigu, relative]

Quant aux distracteurs, 12 étaient parfaitement grammaticaux et les 12 autres étaient agrammaticaux. Ces derniers comportaient des violations de concordance temporelle ou bien des violations d'îlots ou bien des items à polarité négative non autorisés. Les distracteurs grammaticaux contenaient également des îlots, des items à polarité négative, mais sans violations, ainsi que des réalisations correctes de concordance temporelle. Les différentes typologies de distracteurs sont illustrées en (25)-(30).

Concordance temporelle

- (25) L'université a décidé que tout le monde devait régulariser son inscription au mois d'octobre.
 (26) *Le témoin voudrait que la mariée arrêtaît de rire.

Îlots

- (27) Il y a des équations que tu ne peux résoudre que si tu connais les mathématiques.
 (28) *Ils m'ont donné une tablette que mon professeur se plaint si j'utilise en classe.

Items à polarité négative

- (29) Je n'ai jamais vu un arc-en-ciel comme celui de hier.
 (30) *Je n'ai pas jamais rencontré une personne avec un caractère aussi difficile.

Les distracteurs et items expérimentaux étaient intercalés dans un ordre randomisé pour chaque participant et étaient présentés après trois items d'entraînement.

6.1.3.1.2.3 Procédure Dans cette expérience en ligne, chaque phrase apparaissait au centre de l'écran, et les participants devaient la juger sur une échelle d'acceptabilité allant de 0 à 10, 0 signifiant « pas naturel du tout » et 10 signifiant « parfaitement naturel ». Il était demandé aux participants de ne pas juger le contenu des phrases, mais de prêter attention à leur forme et de déterminer si les phrases en question leur paraissaient naturelles ou non en français. L'expérience a été implémentée et administrée en ligne sur IbeX Farm (DRUMMOND, 2013) et durait environ 15 minutes.

6.1.3.1.3 Résultats¹⁰ Les résultats de cette première expérience sont représentés sur le Graphique 6.1. Lorsque les propositions sont sélectionnées par un verbe compatible avec les deux continuations (complétive et relative), les phrases incluant une complétive sans lacune (COMPL; moyenne : 6.53; écart-type = 1.68) ont reçu de meilleures notes comparées à celles incluant une relative avec une lacune (REL.AMB; moyenne = 5.95; écart-type = 1.54). Cependant, les phrases incluant une relative avec une lacune mais sélectionnée par un verbe ne pouvant sélectionner qu'une relative ont été les mieux notées parmi les trois conditions (REL.nonAMB; moyenne = 7.93; écart-type = 1.28). En particulier, on observe une différence claire et nette entre les notes attribuées aux phrases de la condition REL.AMB et celles attribuées aux phrases de la condition REL.nonAMB : les deux sont des relatives donc des dépendances *filler-gap*, mais elles diffèrent dans le fait qu'une complétive est temporairement autorisée en REL.AMB mais pas en REL.nonAMB.

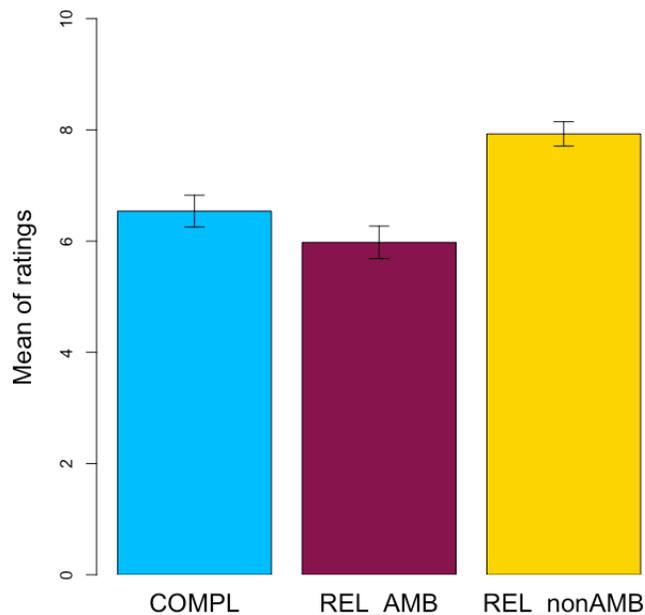


FIGURE 6.1 – Résultats du test de jugement d'acceptabilité : notes moyennes attribuées pour chaque condition sur une échelle de 0 à 10.

En ce qui concerne les distracteurs, notons que les phrases grammaticales ont reçu des notes très élevées (moyenne = 7.64, écart-type = 2.89) alors que celles agrammaticales ont reçu des notes très basses (moyenne = 2.25, écart-type = 3.13), ce qui était attendu.

10. Voir KONRAD, BURATTIN et al., Soumis pour les résultats du test de jugement d'acceptabilité portant sur l'ambiguïté similaire en italien.

En ce qui concerne l'analyse statistique, nous avons implémenté un modèle de régression ordinaire à effets mixtes avec une fonction lien *logit*, en utilisant la fonction *clmm()* dans le paquet *ordinal* (CHRISTENSEN, 2018). Il s'agit d'un modèle statistique spécialement conçu pour traiter des mesures dépendantes ordinales, qui ne peuvent pas être supposées représenter une échelle d'intervalle, comme c'est le cas avec l'échelle Likert utilisée dans notre étude. Nous avons utilisé des contrastes de traitement avec REL.AMB comme niveau de référence, permettant de tester directement l'hypothèse selon laquelle les notes dans cette condition sont moins élevées que celles dans chacune des deux autres conditions. Le modèle le plus maximal qui a convergé incluait la condition comme effet fixe, les sujets et items comme intercepts aléatoires et des variables aléatoires pour les sujets. L'analyse a révélé que les notes dans la condition COMPL étaient significativement plus élevées que celles dans la condition REL.AMB (Estimate = 0.3384, SE = 0.1236, $z = 2.738$, $p < .01$), et que les notes dans la condition REL.nonAMB étaient également significativement plus élevées que dans la condition REL.AMB (Estimate = 1.1976, SE = 0.1188, $z = 10.079$, $p < .0001$).

Malgré la différence significative entre les conditions COMPL et REL.AMB, notons tout de même que les notes pour la condition COMPL ne sont pas aussi élevées que celles dans la condition REL.nonAMB. La deuxième analyse que nous avons faite, avec cette fois-ci la condition COMPL comme niveau de référence, a confirmé que ces conditions diffèrent de façon significative (Estimate = 0.8592, SE = 0.1179, $z = 7.28$, $p < .0001$). Par ailleurs, si nous modifions notre modèle statistique initial en utilisant les items comme variables aléatoires, plutôt que les sujets (sachant qu'un modèle avec les deux ensembles de variables aléatoires ne converge pas), la différence entre COMPL et REL.AMB s'avère ne pas être significative (Estimate = 0.4100, SE = 0.5005, $z = 0.819$, $p = .413$).

En inspectant les données de plus près, nous avons trouvé une explication claire pour cette divergence : on observe une corrélation intéressante entre les notes attribuées pour COMPL et celles attribuées pour REL.AMB, en fonction du verbe qui sélectionne la proposition subordonnée (voir Figure 6.2). Plus précisément, nous avons relevé une corrélation négative (-0.71) significative ($p < .05$) pour chaque verbe ambigu entre les notes moyennes attribuées aux continuations de complétive et celles attribuées aux continuations de relatives, comme l'illustre la ligne de régression en rouge sur le Graphique 6.2.

Cela montre clairement que, lorsqu'un verbe autorise de façon très favorable la continuation de complétive (c'est-à-dire, reçoit des notes élevées pour la condition COMPL, comme le verbe 'tenir à ce que'), il y a un fort désavantage pour la continuation de relative, ce qui est reflété par les notes peu élevées pour la condition REL.AMB. Ce n'est que pour les verbes pour lesquels la continuation de complétive a reçu des notes basses (comme "se soucier de ce que") que la continuation de relative a reçu les notes les plus élevées. Dans ce dernier cas, il ne semble pas y avoir d'effet de labyrinthe tout simplement parce que la continuation de complétive n'a pas l'air d'être disponible pour la majorité des participants après les verbes en question.

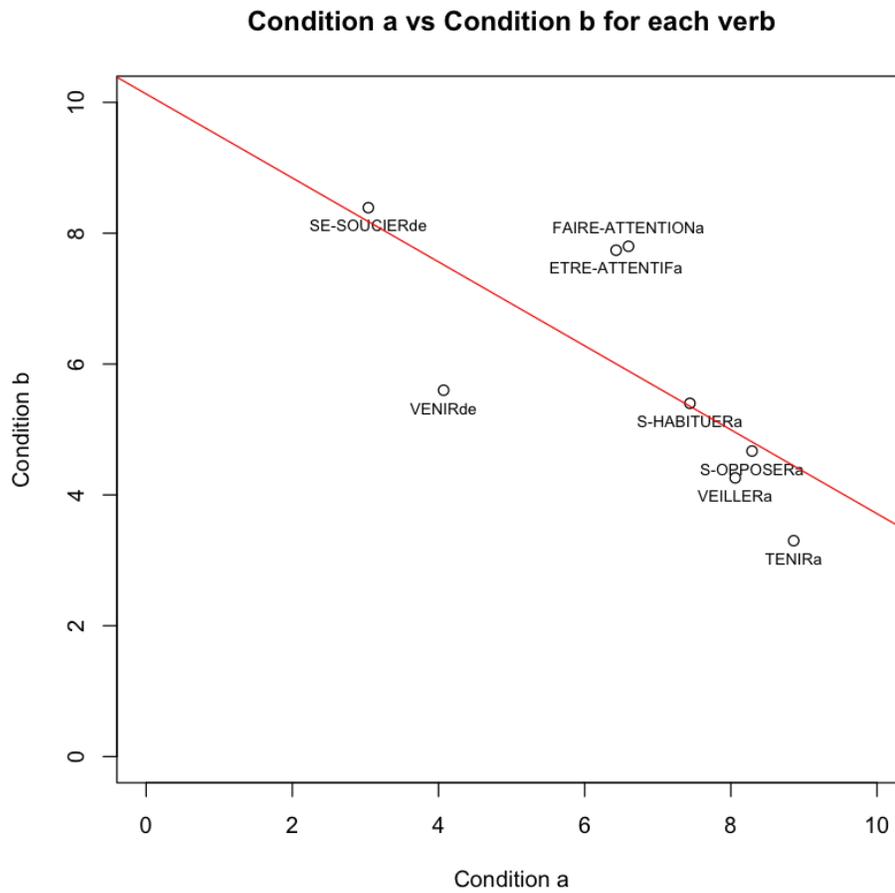


FIGURE 6.2 – Notes moyennes pour chaque verbe ambigu, pour les conditions (a) et (b). La ligne de régression montre la corrélation négative à travers les mêmes verbes dans les deux conditions.

6.1.3.1.3.1 Influence potentielle de la fréquence sur les notes attribuées pour chaque continuation Une question légitime à se poser est la suivante : cette corrélation peut-elle s'expliquer par la fréquence ? Autrement dit, est-ce que la continuation de relative est préférée pour le verbe 'se soucier de ce que' car c'est la continuation la plus fréquente ? Au contraire, est-ce que la continuation de relative est mal notée pour le verbe 'tenir à ce que' car c'est la continuation la moins fréquente ? Afin de vérifier cela, j'ai fait une étude de corpus à partir de FrWaC (BARONI, BERNARDINI, FERRARESI et ZANCHETTA, 2009)^{11 12}

1. **Veiller à ce que** : 9004 résultats, analyse des 200 premiers résultats
2. **Tenir à ce que** : 1098 résultats, analyse des 200 premiers résultats¹³
3. **Venir de ce que** : 614 résultats, analyse des 200 premiers résultats
4. **S'habituer à ce que** : 7 résultats
5. **Se soucier de ce que** : 46 résultats
6. **Être attentif à ce que** : 183 résultats
7. **S'opposer à ce que** : 649 résultats, analyse des 200 premiers résultats
8. **Faire attention à ce que** : 394 résultats, analyse des 200 premiers résultats

Regardons, pour chaque verbe, les résultats de cette étude de corpus tout en comparant les nombres d'occurrences avec les notes moyennes attribuées dans le test de jugement d'acceptabilité (voir Tableau 6.1).¹⁴

11. La taille du corpus en nombre de mots est de 1 328 628 428.

12. A la base, je comptais faire cette analyse dans Frantext, qui est un corpus davantage formel, mais il y avait beaucoup trop peu de données pour plusieurs des verbes ambigus en question. C'est pour cela que je me suis finalement tournée vers un corpus beaucoup moins formel, FrWaC, qui est un corpus de français du Web.

13. Sur ces 200 résultats, j'en ai exclu 9 car ils étaient non ambigus. En effet, le verbe matrice étant 's'en tenir à ce que', la continuation ne pouvait être qu'une relative.

(1) Je m'en tiens à ce que j'ai dit.

14. Pour chaque verbe, j'ai utilisé une requête de type CQL permettant de chercher toutes les formes de chaque verbe. Par exemple, pour 'veiller à ce que', la requête était la suivante : [lemma="veiller"][word="à"][word="ce"][word="que"].

		Acceptabilité	Nombre d'occurrences
Veiller à ce que	Complétive	8,06	200
	Relative	4,26	0
Tenir à ce que	Complétive	8,86	197
	Relative	3,30	3
Venir de ce que	Complétive	4,07	189
	Relative	5,60	11
S'habituer à ce que	Complétive	7,44	7
	Relative	5,40	0
Se soucier de ce que	Complétive	3,04	3
	Relative	8,39	43
Être attentif à ce que	Complétive	6,43	132
	Relative	7,74	51
S'opposer à ce que	Complétive	8,29	198
	Relative	4,62	2
Faire attention à ce que	Complétive	6,60	90
	Relative	7,80	110

TABLE 6.1 – Note moyenne attribuée lors du test de jugement d'acceptabilité et nombre d'occurrences dans FrWaC pour chaque continuation possible (complétive ou relative), pour chacun des 8 verbes ambigus.

On peut observer une corrélation entre fréquence et acceptabilité des deux continuations, complétive et relative, pour 5 verbes sur 8 : 'veiller à ce que', 'tenir à ce que', 'se soucier de ce que', 's'opposer à ce que' et 'faire attention à ce que'. Par contre, pour le verbe 'venir de ce que', la continuation de complétive a été assez mal notée (4,07 sur une échelle de 0 à 10), alors que c'est clairement la continuation la plus fréquente d'après les données de FrWaC (189 complétives contre 11 relatives). Notons quand même que la note moyenne attribuée à la continuation de relative n'est pas très élevée non plus donc nous n'excluons pas le fait qu'il y a une difficulté particulière associée à ce verbe-là, voire aux items incluant ce verbe, qui sont peut-être davantage complexes à traiter. En ce qui concerne le verbe 's'habituer à ce que', on ne peut pas vraiment tirer de conclusion étant donné la très faible quantité de données (seulement 7 résultats dans FrWaC). Notons quand même que ces 7 résultats correspondent tous à des complétives, qui est la continuation la mieux notée. Enfin, pour le verbe 'être attentif à ce que', la fréquence et l'acceptabilité vont dans le sens inverse : les deux continuations sont plutôt bien acceptées, avec une préférence pour la continuation de relative. Or, on trouve 132 complétives contre seulement 51 relatives dans FrWaC. Ainsi, il semblerait que la fréquence joue un rôle pour expliquer les préférences d'acceptabilité entre les continuations, mais ce rôle n'est pas absolu et reste à nuancer.

Afin de vérifier si l'on trouvait un effet significatif de la fréquence, nous avons implémenté, comme précédemment, un modèle de régression ordinaire à effets mixtes en utilisant la fonction *clmm()*. Il s'agissait d'un modèle maximal dans lequel les facteurs CONDITION et FREQUENCY ont été inclus comme prédicteurs, de même que l'interaction entre les deux.¹⁵ L'effet de fréquence s'est avéré être significatif (Estimate = 0.0201, SE = 0.0061, $z = 3.309$, $p = .0009$), de même que l'interaction entre la condition et la fréquence (Estimate = -0.0245, SE = 0.0095, $z = -2.579$, $p = .0099$). L'effet principal signifie qu'une continuation est plus acceptable lorsqu'elle est plus fréquente. Quant à l'interaction, elle signifie que la fréquence explique mieux les jugements pour les complétives que pour les relatives.

6.1.3.1.3.2 Potentiel effet de l'âge sur l'acceptabilité des complétives Un autre point intéressant est que l'acceptabilité des complétives introduites par 'ce que' (COMPL) varie en fonction de l'âge des participants, comme on peut le voir dans le Tableau 6.2.

	< 35 ans	> 35 ans
COMPL	6,33	7,99
REL.AMB	6,10	5,17
REL.nonAMB	7,88	8,23

TABLE 6.2 – Notes moyennes attribuées par les deux groupes d'âge (moins de 35 ans ; plus de 35 ans) pour chaque condition.

Les complétives introduites par 'ce que' (COMPL), qui appartiennent à un registre assez formel et soutenu, sont mieux acceptées par les participants au-delà de 35 ans. La différence entre COMPL et REL.AMB, c'est-à-dire la différence entre respectivement complétives et relatives sélectionnées par un verbe ambigu, est très nette chez les participants les plus âgés, alors qu'elle est minime chez les participants les plus jeunes. Comme la grande majorité des participants sont des étudiants, il y a un déséquilibre en nombre entre les deux groupes d'âge : seulement 9 participants âgés de plus de 35 ans contre 61 participants âgés de moins de 35 ans. C'est probablement en partie pour cette raison que, globalement, la différence entre COMPL et REL.AMB n'est pas très grande. Quant à la différence entre REL.AMB et REL.nonAMB, c'est-à-dire la différence entre les deux types de relatives, avec ou sans ambiguïté, elle est également davantage prononcée chez les participants les plus âgés. Autrement dit, l'effet de labyrinthe est encore plus fort chez ces derniers.

Afin de vérifier si l'on trouvait un effet significatif du groupe d'âge en ce qui concerne les complétives et les relatives sélectionnées par un verbe ambigu, nous avons implémenté un autre modèle de régression ordinaire à effets mixtes en utilisant la fonction *clmm()*. Il

15. Le facteur CONDITION a été codé 1 pour les complétives et 0 pour les relatives.

s'agissait d'un modèle quasiment maximal (sujets et items comme intercepts aléatoires, variables aléatoires pour les sujets seulement¹⁶) dans lequel les facteurs CONDITION et GROUP ont été inclus comme prédicteurs, de même que l'interaction entre les deux. Nous avons bien trouvé un effet principal du groupe d'âge (Estimate = 0.9499, SE = 0.3719, $z = 2.554$, $p = .0107$), et également une interaction entre la condition et le groupe d'âge (Estimate = -1.4574, SE = 0.3363, $z = -4.333$, $p < .001$). L'effet principal signifie que les jugements d'acceptabilité diffèrent entre les moins de 35 ans et les plus de 35 ans. Quant à l'interaction entre la condition et le groupe d'âge, elle s'explique par le fait qu'il n'y a pas d'effet de condition chez les moins de 35 ans. En effet, la différence entre les complétives et les relatives n'est significative que pour le groupe de participants ayant plus de 35 ans (Estimate = -1.6580, SE = 0.4702, $z = -3.526$, $p = .0004$), indiquant une préférence pour les complétives.

Il est important de souligner le fait que, malgré cette variabilité à travers les items d'une part et à travers les deux groupes d'âge d'autre part, une préférence globale pour la continuation de complétive, qui ne contient pas de lacune, a tout de même été trouvée sur la continuation de relative dans le modèle statistique de départ, incluant la condition comme effet fixe.

6.1.3.1.4 Discussion Les résultats de cette première expérience montrent que les participants évitent les phrases contenant une dépendance *filler-gap* seulement lorsque celles-ci sont en conflit avec une alternative sans lacune. Cela est mis en évidence par les notes basses pour la condition REL.AMB comparée à la condition COMPL d'une part, et, surtout, comparée à la condition REL.nonAMB d'autre part, malgré le fait que dans le cas de la dernière comparaison, les deux types de phrases contiennent une relative de même type. Ces résultats sont cohérents avec l'hypothèse de départ, rappelée ici : en lisant les phrases en REL.AMB, les participants adoptent initialement une analyse de complétive et ce n'est qu'à un stade ultérieur qu'ils sont forcés de ré-analyser la phrase comme une relative. Cette ré-analyse n'est pas requise en REL.nonAMB, étant donné la non disponibilité de la lecture de complétive après les verbes matrices utilisés dans cette condition.

Un résultat davantage inattendu est que les phrases de la condition COMPL ont été moins bien notées que les relatives sans ambiguïté (condition REL.nonAMB). On peut supposer que cela est dû au fait que la complétive en 'ce que' appartient à un registre relativement soutenu, ce qui n'est pas le cas des relatives en 'ce que' de la condition REL.nonAMB. Les phrases de la condition COMPL restent néanmoins mieux notées (du moins dans l'analyse par sujets) que les continuations conflictuelles contenant une dépendance *filler-gap* (condition REL.AMB).

En conclusion, ces résultats sont cohérents avec les prédictions de la deuxième proposition du Principe de Chaîne Minimale, selon lequel une lacune doit être évitée si possible.

16. Le modèle maximal incluant en plus les variables aléatoires pour les items ne convergait pas.

Plus précisément : si un verbe autorise une continuation sans lacune, alors ne pas postuler l'existence d'une lacune. Si un verbe n'autorise pas une continuation sans lacune, comme en REL.nonAMB dans notre expérience, alors poser l'existence d'une lacune (dans la position la plus proche possible). Pour ce qui est des constructions en 'ce que' qui nous intéressent, ces résultats nous confirment que 'ce que' est compatible avec des structures très différentes, face auxquelles les locuteurs réagissent de manière très différente.

6.1.3.2 Expérience 2 : Tâche de lecture par présentation auto-segmentée

La deuxième expérience est une tâche de lecture par présentation auto-segmentée conduite sur un groupe séparé de participants français en utilisant le même matériel que celui utilisé pour le test de jugement d'acceptabilité (Expérience 1). Même si l'effet de phrase labyrinthe est déjà visible avec une mesure indirecte telle que le jugement d'acceptabilité, nous voudrions vérifier si cet effet se retrouve également avec une mesure de traitement en temps réel, telle que la lecture par présentation auto-segmentée. De plus, ce type de tâche nous permettra de localiser avec précision le point où la difficulté de traitement apparaît.

6.1.3.2.1 Design expérimental

6.1.3.2.1.1 Participants 33 participants français (dont 29 femmes) âgés entre 18 et 39 ans ($M=22$ ans) ont participé à cette étude. Aucun de ces participants n'a fait partie de la première expérience. La plupart d'entre eux sont étudiants à Paris. Ils ont été recrutés via les connaissances et les réseaux sociaux, ou bien directement dans les locaux de l'Université Paris Diderot.

6.1.3.2.1.2 Matériel et procédure Nous avons repris le matériel (même items expérimentaux et mêmes distracteurs) de l'Expérience 1. Cette deuxième expérience a également été implémentée et administrée sur Ibex Farm.¹⁷ Les participants ont été testés dans une pièce au calme à l'Université Paris Diderot.^{18 19} Il leur était demandé de lire (silencieusement) des phrases mot par mot : ils devaient appuyer sur la barre d'espace de l'ordinateur pour faire apparaître le mot suivant sur l'écran. Leurs temps de lecture (en millisecondes) ont été automatiquement enregistrés. Dans le but de contrôler qu'ils prêtaient attention à ce qu'ils lisaient et qu'ils comprenaient ce qu'ils étaient en train de lire, nous avons créé des questions de compréhension (de type oui-non), mais de façon non systématique (pour un item sur trois et pour les distracteurs grammaticaux). Cela les forçait à rester concentrés durant la tâche et à ne pas appuyer sur la barre d'espace trop rapidement sans même lire

17. Comme précisé précédemment, cette expérience a été réalisée dans le cadre d'une collaboration avec Francesca Foppolo, Adrian Staub, Massimo Burattin, Carlo Cecchetto et Caterina Donati. Nous avons étudié une ambiguïté similaire en italien et comparé les résultats de l'expérience italienne avec ceux de l'expérience française. L'expérience italienne ayant été implémentée sur Ibex Farm avant l'expérience française, nous avons souhaité garder le même paradigme expérimental pour le français par souci de parallélisme.

18. Je tiens à remercier Fahima MAOUCHE de m'avoir aidée pour la passation de cette expérience.

19. Je remercie également le Labex EFL pour son financement m'ayant permis d'indemniser les participants.

ni comprendre les phrases en question. Ces questions de compréhension étaient les mêmes pour toutes les conditions d'un même item et elles portaient sur la proposition enchâssée. Notons qu'un participant a été exclu des analyses qui suivent étant donné que sa performance pour les questions de compréhension était en-dessous de 80%.

6.1.3.2.2 Résultats²⁰ Pour l'analyse des résultats, une étiquette a été assignée à chaque mot en fonction de sa position dans la phrase par rapport au point de désambiguïsation, qui a été codé comme étant le point 0 (voir Tableaux 6.3 et 6.4). Ce point de désambiguïsation est situé juste après le verbe de la subordonnée. Si un déterminant apparaît juste après le verbe enchâssé, cela signifie qu'il n'y a pas de lacune, donc la continuation est de façon non équivoque une complétive (condition (a)). Par contre, si une préposition apparaît juste après le verbe enchâssé, cela signifie qu'il y a une lacune dans la position d'objet, donc la continuation s'avère être une relative (conditions (b) et (c)). Le point de désambiguïsation est souligné pour chaque phrase de chaque item dans l'Annexe F.

20. Voir KONRAD, BURATTIN et al., Soumis pour les résultats de l'expérience de lecture par présentation auto-segmentée portant sur le traitement de l'ambiguïté similaire en italien.

	-8	-7	-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4
(a)	Léo	tient	à	ce	que	sa	soeur	offre	un	cadeau	à	leur	mère.
(b)	Léo	tient	à	ce	que	sa	soeur	offre	à	leur	mère.		
(c)	Léo	apprécie	ce	que	sa	soeur	offre	à	leur	mère.			

TABLE 6.3 – Encodage des régions pour les analyses des items expérimentaux de la tâche de lecture par présentation auto-segmentée (exemple 1). Le point de désambiguïsation est indiqué en gras.

	-8	-7	-6	-5	-4	-3	-2	-1	0	+1	+2	+3
(a)	Rémi	s'oppose	à	ce	que	son	enfant	demande	un	ordinateur	pour	Noël.
(b)	Rémi	s'oppose	à	ce	que	son	enfant	demande	pour	Noël.		
(c)	Rémi	mémorise	ce	que	son	enfant	demande	pour	Noël.			

TABLE 6.4 – Encodage des régions pour les analyses des items expérimentaux de la tâche de lecture par présentation auto-segmentée (exemple 2). Le point de désambiguïsation est indiqué en gras.

Notons que pour la condition (b), l'arrivée au point de désambiguïsation ne force pas complètement la lecture avec lacune. Le NP 'un cadeau' ou 'un ordinateur' pourrait être en effet être déplacé à la périphérie droite de la subordonnée. Mais cette possibilité de (*NP shift*) ne semble pas très naturelle en français et nous verrons qu'elle n'interfère pas avec nos résultats.

- (31) Léo tient à ce que sa soeur offre à leur mère **un cadeau**.
- (32) Rémi s'oppose à ce que son enfant demande pour Noël **un ordinateur**.

Les temps de lecture moyens pour les régions pertinentes qui nous intéressent ici sont représentés sur la Figure 6.3. Le point de désambiguïsation est appelé ‘disamb’. Ensuite, si l’on se déplace vers la gauche de l’axe des abscisses, on trouve les positions -1 (‘pre.disamb’) et -2 (‘pre.pre.disamb’); si l’on va vers la droite du même axe, on trouve les positions +1 (‘post.disamb’) et +2 (‘post.post.disamb’).

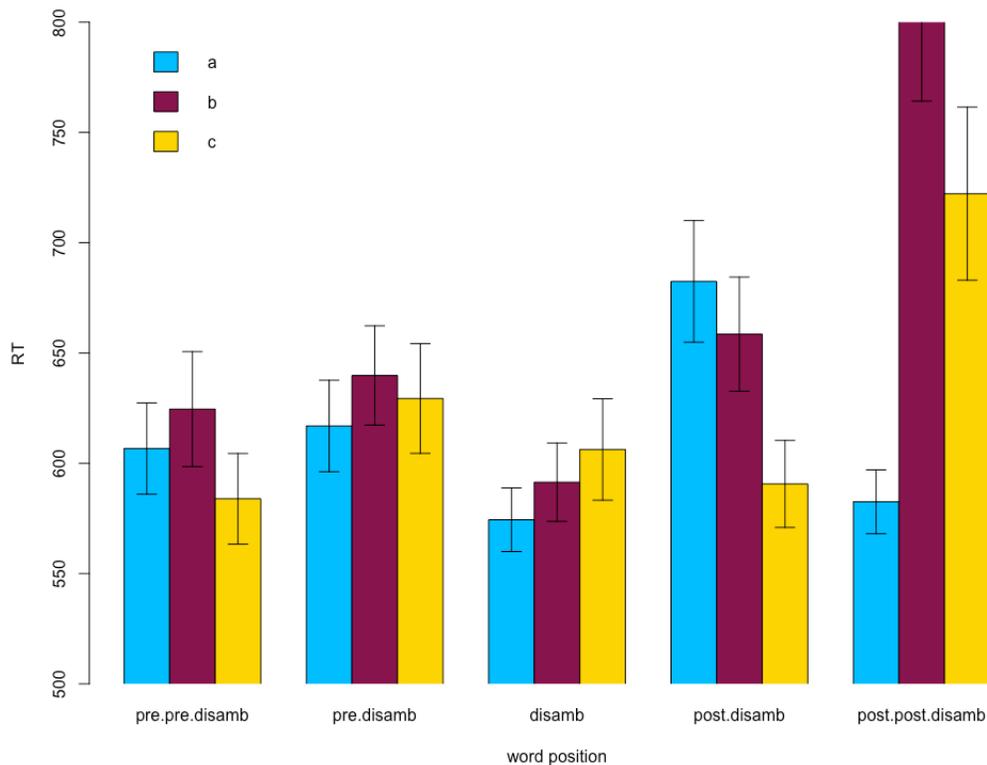


FIGURE 6.3 – Temps de lecture moyens pour les mots correspondant au point de désambiguïsation dans chaque condition, appelée ‘disamb’, et pour les régions précédant ce point (‘pre.pre.disamb’ et ‘pre.disamb’) et celles suivant ce point (‘post.disamb’ et ‘post.post.disamb’).

Les temps de lecture transformés en log ont été soumis à un modèle linéaire à effets mixtes en utilisant la fonction *lmer* du paquet *lme4* (BATES, MÄCHLER, BOLKER et WALKER, 2015) dans R. Le facteur CONDITION a été inclus comme prédicteur. La longueur des mots a également été ajoutée dans le modèle (voir Annexe N pour formule du modèle et *output*). Nous avons exécuté deux modèles maximaux de ce type sur chaque région de la phrase. Dans le premier modèle, la condition (a) [COMPL] était le niveau de référence. Il nous permettait de pouvoir comparer les conditions (a) [COMPL] et (b) [REL.AMB]. Nous avons

ensuite exécuté le même modèle mais en établissant la condition (b) [REL.AMB] comme niveau de référence ('cond2' au lieu de 'condition'), afin de pouvoir contraster les conditions (b) [REL.AMB] et (c) [REL.nonAMB]. Les régions qui nous intéressent pour l'analyse sont le point de désambiguïsation ainsi que les deux régions qui suivent. Les régions précédant le point de désambiguïsation ne sont pas pertinentes.

Comme le montrent les Tableaux 6.3 et 6.4, la longueur des mots diffère entre la condition (a) [COMPL] d'une part et les conditions (b) [REL.AMB] et (c) [REL.nonAMB] d'autre part, dans les deux régions suivant la désambiguïsation. Au niveau +1, il y a un nom dans la condition (a), qui est plus long que le déterminant dans les conditions (b) et (c). Cela pourrait expliquer pourquoi les temps de lecture sont plus longs pour (a) comparés à ceux pour (b) et (c), au point 'post.disamb'. On retrouve le schéma inverse au niveau +2 : il y a une préposition (plus courte) dans la condition (a) et un nom (plus long) dans les conditions (b) et (c). Cela pourrait encore une fois expliquer pourquoi les temps de lecture sont si courts pour (a) au point 'post.post.disamb', comparés à ceux pour (b) et (c). Il est donc crucial d'ajouter la longueur des mots dans le modèle.

On trouve des effets dans les deux régions suivant le point de désambiguïsation, mais pas au niveau du point de désambiguïsation même. Comme on peut voir sur le Graphique 6.3, la différence entre (b) et (c) semble surgir après le point de désambiguïsation. Cet effet de tâche d'huile (*spill-over effect*) est confirmé par l'analyse statistique : nous avons relevé une différence significative entre les deux conditions ($t = -2.238$, $p = .028$) un mot après le point de désambiguïsation ('post.disamb' sur le Graphique 6.3 et position +1 dans les Tableaux 6.3 et 6.4), les temps de lecture étant plus élevés pour la condition (b) que pour la condition (c). On trouve également un effet principal de longueur dans cette région de la phrase ($t = 3.231$, $p = .003$).

Passons maintenant aux effets dans la région codée +2 ('post.post.disamb'). Nous avons relevé une différence significative entre les conditions (a) et (b) ($t = 2.350$, $p = .025$) et une différence marginale entre les conditions (a) et (c) ($t = 1.713$, $p = .0965$), les temps de lecture étant plus élevés pour la condition (a) que pour les conditions (b) et (c). L'effet de longueur s'avère également être significatif ($t = 2.037$, $p = .048$). Les temps de lecture sont si courts pour la condition (a) sûrement parce que les mots à cet endroit précis de la phrase sont plus courts dans cette condition que dans les deux autres.

6.1.3.2.3 Discussion Les résultats de cette deuxième expérience vont dans la direction de ceux de la première expérience (test de jugement d'acceptabilité) discutée dans la section précédente. En particulier, nous avons décelé des difficultés de traitement pour la condition (b) [REL.AMB], qui ne se retrouvent pas dans la condition (c) [REL.nonAMB] : notes plus basses pour (b) [REL.AMB] que pour (c) [REL.nonAMB] dans le test de jugement d'acceptabilité; temps de lecture plus longs pour (b) [REL.AMB] que pour (c) [REL.nonAMB] dans la tâche de lecture par présentation auto-segmentée. Dans les deux conditions, la subordonnée

est une relative, c'est-à-dire une dépendance *filler-gap*, mais (b) [REL.AMB] est plus difficile à traiter que (c) [REL.nonAMB] à cause de la compétition avec une autre structure, qui ne contient pas de lacune, la complétive (condition (a) [COMPL]). C'est exactement ce à quoi on s'attendait. Les résultats de cette expérience ont également révélé un effet de tâche d'huile : cette difficulté de traitement pour (b) [REL.AMB] comparée à (c) [REL.nonAMB] n'apparaît pas au point même de désambiguïsation mais dans la région suivant immédiatement ce point, et elle est encore présente deux mots après le point de désambiguïsation.

Toutefois, rappelons que, dans le test de jugement d'acceptabilité, nous avons trouvé une différence subtile mais décelable entre (a) [COMPL] et (b) [REL.AMB], (a) étant mieux noté que (b), mais cette différence n'a pas été détectée dans la tâche de lecture par présentation auto-segmentée. Le fait que (a) [COMPL] diffère de (b) [REL.AMB] et (c) [REL.nonAMB] dans les deux régions suivant la désambiguïsation est probablement dû à une certaine variation entre les conditions en ce qui concerne la longueur des mots.

La nette différence entre (b) [REL.AMB] et (c) [REL.nonAMB] est une nouvelle évidence, en français, en faveur de la première proposition du Principe de Chaîne Minimale, montrant qu'une structure sans lacune est préférée sur une dépendance *filler-gap* lorsqu'il y a une ambiguïté possible entre les deux types de structures. Cette différence de traitement entre les complétives et les relatives introduites par 'ce que' nous indique bien que nous avons affaire ici à deux 'ce que' distincts, un avec lacune et un sans lacune.

6.2 Et quand les deux structures sont des dépendances *filler-gap*? : traitement des relatives à tête allégée et interrogatives indirectes

Nous venons de voir que, dans les cas d'ambiguïté entre la structure avec lacune (relative) et la structure sans lacune (complétive), cette dernière était préférée car l'analyseur évite autant que possible de postuler une lacune (Principe de Chaîne Minimale, De Vicenzi, 1991). Nous avons au moins deux 'ce que' distincts. Revenons maintenant à l'ambiguïté syntaxique sur laquelle nous nous focalisons dans cette thèse, celle entre relatives et interrogatives indirectes introduites par 'ce que'. Le traitement de ces structures est d'autant plus intéressant que nous sommes confrontés à la même séquence 'ce que' (avec lacune) dans les deux cas, du moins en apparence. Avons-nous affaire à deux 'ce que' distincts ici aussi? Autrement dit, avons-nous trois 'ce que' traités différemment? Les relatives et les interrogatives indirectes impliquant toutes deux une lacune, le Principe de Chaîne Minimale ne nous permet pas ici de pouvoir prédire une préférence pour l'une ou l'autre des structures puisqu'il n'y a pas de compétition entre une structure sans lacune et une structure avec lacune.

6.2.1 Expérience de lecture par présentation auto-segmentée

Rappelons que l'expérience que nous avons mise en place pour tester les enfants était un test de répétition. Ce design expérimental n'est pas adapté pour les adultes : la tâche est trop facile et non pertinente pour les adultes. Nous avons donc fait le choix de mettre en place une nouvelle expérience de lecture par présentation auto-segmentée afin de mesurer le traitement en temps réel de ces structures introduites par 'ce que' et de voir s'il y a une différence de traitement entre relatives et interrogatives indirectes. Nous avons également souhaité comparer le traitement d'interrogatives indirectes introduites par 'ce que' avec celui d'interrogatives indirectes introduites par un mot *wh-* comme 'quand'.

Cette expérience est similaire à celle avec les complétives, présentée dans la sous-section 6.1.3.2. L'intérêt de cette deuxième expérience est qu'elle nous permet de voir les préférences de l'analyseur dans des structures à l'apparence identique.

6.2.1.1 Prédiction

D'une part, étant donné que nous avons affaire ici à deux dépendances *filler-gap*, il ne devrait pas y avoir de différence due au traitement de la lacune et relativement au Principe de Chaîne Minimale.

D'autre part, si l'on suppose une continuité avec les résultats du test de répétition chez les enfants français de 3 à 6 ans, alors on s'attend à trouver chez les adultes français une préférence pour les relatives introduites par 'ce que' comparées aux interrogatives indirectes introduites par 'ce que'. En effet, pour ce qui est des enfants, nous avons des indices qu'ils traitent différemment les deux structures : performance plus élevée pour les relatives que pour les interrogatives indirectes ; tendance à omettre 'ce' et à employer une forme *wh-complexe* (*kɛsk*) davantage dans les interrogatives indirectes. On s'attend alors, pour ce qui est des adultes, à des temps de lecture plus élevés au niveau de 'ce que', voire dans la ou les régions qui suivent, lorsque le verbe matrice sélectionne un CP, donc une interrogative indirecte. De plus, si 'ce que', du fait de sa nature nominale, « dérange » les adultes comme il « dérange » les enfants dans les interrogatives indirectes, on s'attend à percevoir une différence de traitement entre les interrogatives indirectes introduites par 'ce que' et celles introduites par 'quand', qui est clairement *wh-* et sans élément D. Concrètement, un effet de surprise est attendu au moment où l'analyseur rencontre 'ce que' après un verbe interrogatif, sélectionnant un CP, mais pas après un verbe sélectionnant un DP, donc une relative.

Notons qu'il s'agit ici d'une ambiguïté syntaxique, en ce sens qu'une même subordonnée, introduite par 'ce qu-' peut être soit une relative, soit une interrogative en fonction des critères de sélection du verbe matrice (une relative si le verbe matrice sélectionne un DP ; une interrogative indirecte si le verbe matrice sélectionne un CP), mais pas d'une ambiguïté temporaire, comme c'était le cas avec l'ambiguïté entre complétives et relatives. C'est le

verbe matrice lui-même qui permet de déterminer très tôt si la subordonnée qui suit est une interrogative indirecte ou bien une relative. Ainsi, lors du traitement linéaire de ces phrases, il n’y a pas de réel point de désambiguïsation, mais plutôt une région critique correspondant au mot subordonnant. Il n’y a donc pas de réanalyse possible comme c’était le cas dans la première expérience de lecture par présentation auto-segmentée, mais plutôt un effet de surprise possible lorsque l’on rencontre ‘ce qui/ce que’ après un verbe interrogatif.

6.2.1.2 Design expérimental

6.2.1.2.1 Participants Ont participé à cette expérience 33 adultes français (dont 19 femmes) âgés entre 18 et 58 ans ($M=27$ ans). Ils ont été recrutés via les connaissances, les réseaux sociaux, ou bien directement à l’Université Paris Diderot.

6.2.1.2.2 Matériel Nous avons créé 50 items expérimentaux, comportant chacun cinq conditions :

- exp-a : relatives introduites par ‘ce que’
- exp-b : interrogatives indirectes introduites par ‘ce que’
- exp-c : interrogatives indirectes introduites par ‘quand’
- exp-d : relatives introduites par ‘ce qui’
- exp-e : interrogatives indirectes introduites par ‘ce qui’

Voilà un exemple de quintuple expérimental (voir Annexe G pour liste complète des items) :

- (33)
- a. Le directeur de thèse relit ce que son doctorant rédige pendant la semaine.
 - b. Le directeur de thèse ignore ce que son doctorant rédige pendant la semaine.
 - c. Le directeur de thèse ignore quand son doctorant rédige pendant la semaine.
 - d. Le directeur de thèse relit ce qui a été rédigé pendant la semaine.
 - e. Le directeur de thèse ignore ce qui a été rédigé pendant la semaine.

La condition exp-c permet de comparer les interrogatives indirectes introduites par ‘ce qu-’, c’est-à-dire par un élément D suivi du complémenteur, donc sans trait *wh-* en apparence, avec celles introduites par un mot *wh-*, sans élément D. Le mot *wh-* choisi pour cette comparaison est ‘quand’. Il est vrai que ‘quand’ est différent de ‘ce que’ ou ‘ce qui’ aussi du point de vue de la fonction : adjectif *versus* complément. Mais, en termes de longueur, ‘quand’ est aussi long que ‘ce que’ ou ‘ce qui’. ‘Qui’ aurait été préférable, car c’est aussi un argument, d’autant plus que nous avons comparé ‘ce qu-’ et ‘qui’ dans le Chapitre 4 sur l’acquisition de ces structures par les enfants. Mais le problème avec ‘qui’ est que l’on voit très clairement qu’il y a un élément de moins (l’élément D ‘ce’) comparé à ‘ce qui’. De plus, l’adjectif ‘quand’ permet de pouvoir garder toutes les conditions aussi parallèles que possibles les unes des autres. Au contraire, ‘qui’ étant [+animé] et ‘ce que’ [-animé], il aurait été difficile d’avoir une subordonnée identique entre les conditions.

Nous avons dû respecter certaines contraintes pour la création de ces items, en particulier :

- Choisir un DP sujet complexe, de type ‘le directeur de thèse’ ou ‘le fils du boulanger’, de façon à ne pas arriver trop rapidement à la région critique (le mot subordonnant).
- Pour le verbe de la subordonnée, nous avons dû faire attention à ce que la différence entre la forme du présent et la forme du participe passé ne soit pas trop grande en termes de longueur, dans le but d’avoir une différence minimale entre les conditions a), b), c) d’une part et les conditions d) et e) d’autre part (par exemple : rédige / rédigé). De plus, une autre contrainte concernant le verbe enchâssé était la nécessité d’avoir un verbe transitif qui puisse être employé de façon intransitive dans certains contextes. En effet, pour les conditions a), b), d) et e), le verbe doit être transitif (rédiger quelque chose). Par contre, pour la condition c) (quand-INT), l’emploi intransitif est requis (rédiger).
- Enfin, en ce qui concerne l’adjectif à la fin de la subordonnée, il fallait éviter les adjectifs temporels étant donné que la condition c) est une proposition de nature adjectivale introduite par ‘quand’, donc temporelle.

Par ailleurs, pour les verbes matrices, il était difficile de trouver des verbes purement interrogatifs, pouvant sélectionner un CP de façon non ambiguë. Pour cette raison, pour les conditions b), c) et e), c’est-à-dire pour les interrogatives indirectes, nous avons choisi 5 verbes, chacun ayant été utilisé 10 fois : ‘se demander’, ‘demander’, ‘ignorer’, ‘deviner’, ‘redemander’. Il était par contre plus facile de trouver des verbes pouvant sélectionner un DP (donc une relative) : pour les conditions a) et d), nous avons choisi 25 verbes et réutilisé chacun de ces verbes une fois.

Nous avons également créé 50 distracteurs, des complétives courtes et simples à traiter, de type :

- (34) Marie pense que Julie a mangé les chocolats.
- (35) Luc est certain que Léa est en vacances.

6.2.1.2.3 Procédure Cette expérience a été implémentée et administrée sur Ibex Farm. Les participants ont été testés dans une pièce au calme à l’Université Paris Diderot^{21 22}. Comme pour la précédente expérience du même type (cf. sous-section 6.1.3.2.1), il leur était demandé de lire, dans leur tête, des phrases mot par mot. Les items et distracteurs ont été randomisés pour chaque participant. Chaque participant n’a vu qu’une condition de chaque item (carré latin). Chaque item était suivi d’une question de compréhension, à laquelle il

21. Je tiens à remercier Antoine HEDIER de m’avoir aidée pour la création des items de cette expérience ainsi que pour sa passation.

22. Je remercie également le Labex EFL pour son financement m’ayant permis d’indemniser les participants.

fallait répondre par 'oui' ou par 'non'. Notons que trois participants ont été exclus des analyses qui suivent car leur performance aux questions de compréhension était inférieure à 80%.

6.2.1.3 Résultats

Pour l'analyse des résultats, une étiquette a été assignée à chaque mot en fonction de sa position dans la phrase, par rapport au point critique, à savoir l'introducteur de subordonnées ('ce qui', 'ce que' ou 'quand'), qui a été codé comme étant le point 0 (voir Tableaux 6.5 et 6.6). Les mots avant le mot subordonnant ont été codés -1, -2, -3, etc. ; ceux après le mot subordonnant ont été codés +1, +2, +3, etc.

-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6
D	N1	PREP	N2	matrix-V	SUB	D	N3	SUB-V	PREP	D	N4
(a)	Le directeur	de	thèse	relit	ce que	son	doctorant	rédige	pendant	la	semaine.
(b)	Le directeur	de	thèse	ignore	ce que	son	doctorant	rédige	pendant	la	semaine.
(c)	Le directeur	de	thèse	ignore	quand	son	doctorant	rédige	pendant	la	semaine.
D	N1	PREP	N2	matrix-V	SUB	AUX	AUX	SUB-V	PREP	D	N4
(d)	Le directeur	de	thèse	relit	ce qui	a	été	rédigé	pendant	la	semaine.
(e)	Le directeur	de	thèse	ignore	ce qui	a	été	rédigé	pendant	la	semaine.

TABLE 6.5 – Encodage des régions pour les analyses des items expérimentaux de l'expérience de lecture par présentation auto-segmentée (exemple 1). Le point critique (mot subordonnant) est indiqué en gras.

	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6
	D	N1	PREP	N2	matrix-V	SUB	D	N3	SUB-V	PREP	D	N4
(a)	Le	proviseur	du	collège	désapprouve	ce que	l'	équipe	filme	dans	la	cour.
(b)	Le	proviseur	du	collège	redemande	ce que	l'	équipe	filme	dans	la	cour.
(c)	Le	proviseur	du	collège	redemande	quand	l'	équipe	filme	dans	la	cour.
	D	N1	PREP	N2	matrix-V	SUB	AUX	AUX	SUB-V	PREP	D	N4
(d)	Le	proviseur	du	collège	désapprouve	ce qui	a	été	filmé	dans	la	cour.
(e)	Le	proviseur	du	collège	redemande	ce qui	a	été	filmé	dans	la	cour.

TABLE 6.6 – Encodage des régions pour les analyses des items expérimentaux de l'expérience de lecture par présentation auto-segmentée (exemple 2). Le point critique (mot subordonnant) est indiqué en gras.

Rappelons les deux comparaisons qui nous intéressent ici : la première est celle entre relatives introduites par ‘ce qui/ce que’ et interrogatives indirectes introduites par ‘ce qui/ce que’ ; la deuxième est celle entre interrogatives indirectes introduites par ‘ce qui/ce que’ et interrogatives indirectes introduites par ‘quand’. Les analyses et résultats pour chacune de ces comparaisons sont présentés ci-après.

Commençons avec la première comparaison. Les temps de lecture moyens pour les régions pertinentes qui nous intéressent ici sont représentés sur le Graphique 6.4. Le point critique est ‘SUB word’, c’est-à-dire le mot subordonnant. Les régions indiquées avant et après le mot subordonnant correspondent aux étiquettes données dans les Tableaux 6.5 et 6.6.

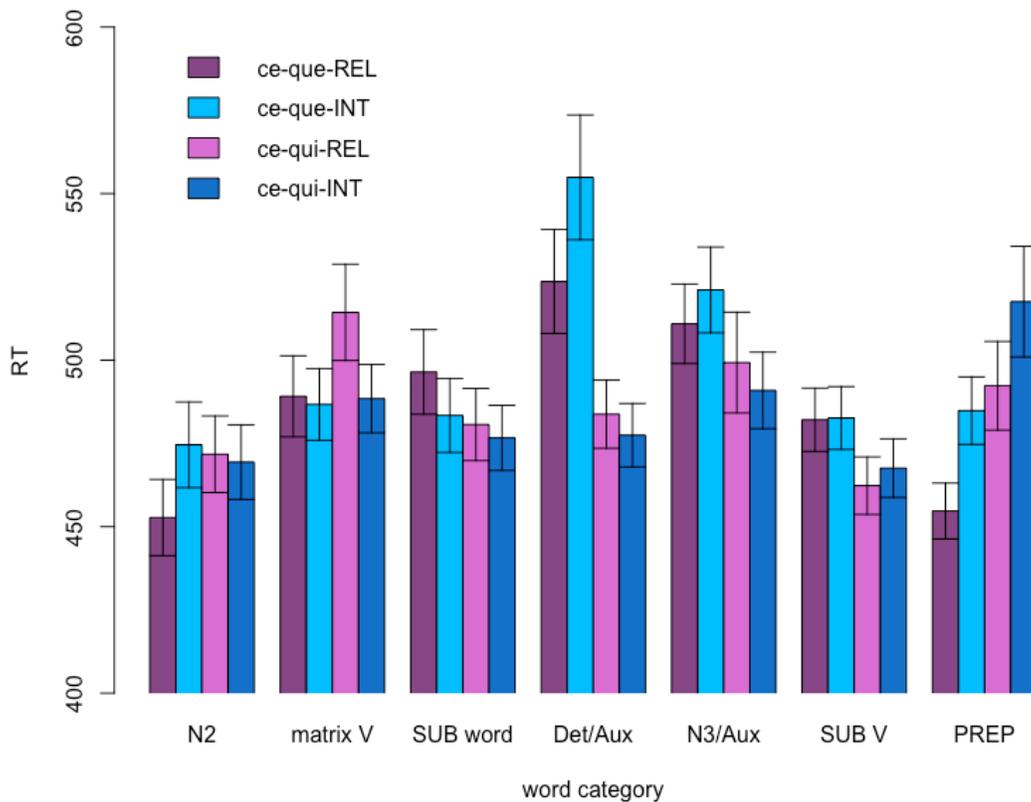


FIGURE 6.4 – Expérience de lecture par présentation auto-segmentée - comparaison entre relatives et interrogatives introduites par ‘ce qui/ce que’ : temps de lecture moyens pour chaque condition (ce-que-INT, ce-que-REL, ce-qui-INT, ce-qui-REL) selon les régions de la phrase.

Les temps de lecture transformés en log ont été soumis à un modèle linéaire à effets mixtes en utilisant la fonction *lmer* du paquet *lme4* dans R. Les facteurs PHRASETYPE [REL/INT] et GAPTYPE [SUJ/OBJ], qui ont été centrés²³, ont été inclus comme prédicteurs, de même que l'interaction entre les deux. La longueur des mots a également été ajoutée dans le modèle (voir Annexe O pour formule du modèle et *output*). Nous avons exécuté un modèle maximal de ce type sur chaque région de la phrase. Les régions qui nous intéressent pour l'analyse sont la région du mot subordonnant et les régions suivant le mot subordonnant. Les régions précédant le mot subordonnant, c'est-à-dire avant d'arriver au point critique, ne sont pas pertinentes.

On trouve des effets dans les quatre régions qui suivent la région critique, mais pas au niveau de la région critique même ('SUB word'). Dans les régions codées +1 ('Det/Aux') et +3 ('SUB V'), on observe seulement un effet marginal de GAPTYPE (région +1 : $t=-1.684$, $p=.0965$; région +3 : $t=-1.768$, $p=.0859$), signifiant que les temps de lecture sont plus élevés pour les dépendances objet que pour les dépendances sujet dans ces deux régions de la phrase. Quant aux régions +2 ('N3/Aux') et +4 ('PREP'), seul l'effet de longueur s'avère être significatif (région +2 : $t=3.958$, $p=.0001$; région +4 : $t=9.152$, $p<.001$), ce qui veut dire que les différences de temps de lecture observées entre les conditions sur le Graphique 6.4 au niveau de ces régions sont dues à des différences de longueur des mots. Notons que l'effet de longueur est également significatif au niveau du verbe matrice, juste avant la région critique ($t=2.869$, $p=.0051$).

Passons maintenant à la deuxième comparaison, en nous focalisant sur les différents types d'interrogatives indirectes, et en excluant donc les relatives. Les temps de lecture moyens pour les régions pertinentes qui nous intéressent ici sont représentés sur le Graphique 6.5.

Comme précédemment, les temps de lecture transformés en log ont été soumis à un modèle linéaire à effets mixtes en utilisant la fonction *lmer* du paquet *lme4* dans R. Nous avons pris comme niveau de comparaison la condition quand-INT. Les facteurs CEQUE [ce-que-INT/quand-INT] et CEQUI [ce-qui-INT/quand-INT], qui ont été centrés²⁴, ont été inclus comme prédicteurs. Comme précédemment, la longueur des mots a également été ajoutée dans le modèle (voir Annexe O pour formule du modèle et *output*). Nous avons exécuté un modèle maximal de ce type sur chaque région de la phrase. Encore une fois, les régions qui nous intéressent particulièrement pour l'analyse sont la région du mot subordonnant et les régions suivant le mot subordonnant.

23. Le facteur PHRASETYPE a été codé 1 pour relative, 0 pour interrogative; le facteur GAPTYPE a été codé 1 pour sujet et 0 pour objet.

24. Le facteur CEQUE a été codé 1 pour ce-que-INT, 0 pour quand-INT; le facteur CEQUI a été codé 1 pour ce-qui-INT et 0 pour quand-INT.

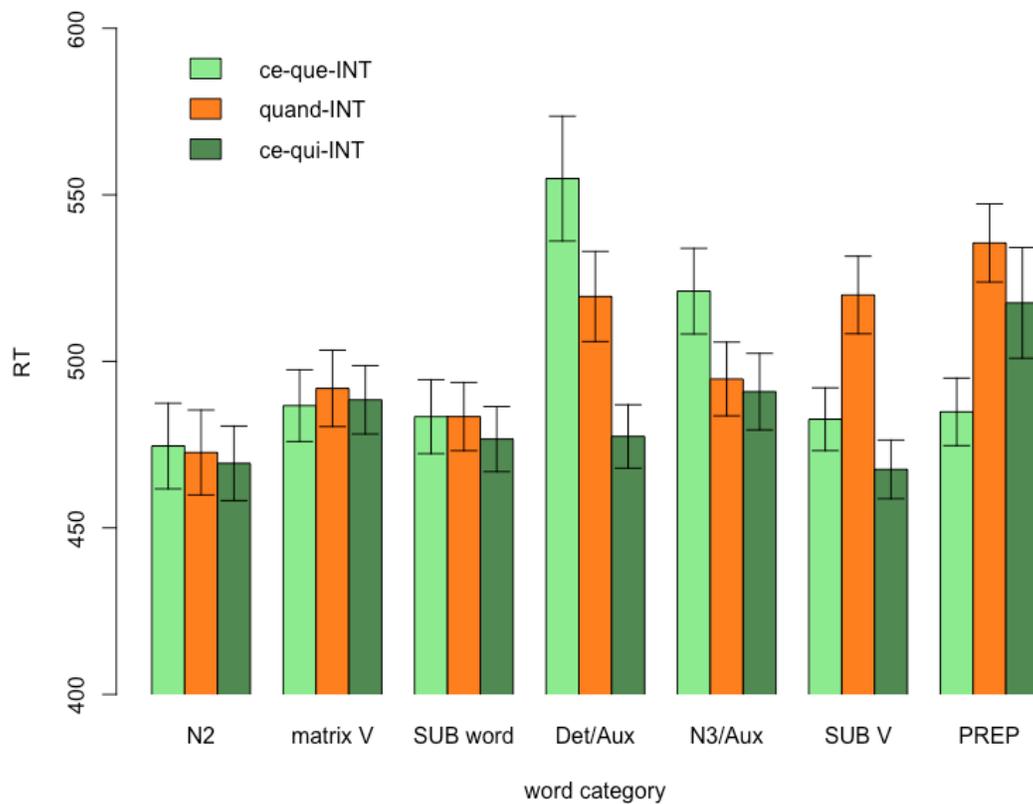


FIGURE 6.5 – Expérience de lecture par présentation auto-segmentée - comparaison entre interrogatives indirectes introduites par ‘ce qui/ce que’ et interrogatives indirectes introduites par ‘quand’ : temps de lecture moyens pour chaque condition (ce-que-INT, quand-INT, ce-qui-INT) selon les régions de la phrase.

Comme c'était le cas pour la première comparaison, on trouve des effets dans les quatre régions qui suivent la région critique, mais pas au niveau de la région critique elle-même ('SUB word'). Dans les régions codées +1 ('Det/Aux') et +2 ('N3/Aux'), seul l'effet de longueur s'avère être significatif (région +1 : $t=3.669$, $p=.0006$; région +2 : $t=5.081$, $p<.001$). Dans les régions codées +3 ('SUB V') et +4 ('PREP'), tous les effets sont significatifs. On trouve donc :

1. un effet de CEQUE
 région +3 ('SUB V') : $t=-2.249$, $p=.031$
 région +4 ('PREP') : $t=-2.786$, $p=.0079$
2. un effet de CEQUI
 région +3 ('SUB V') : $t=-2.947$, $p=.006$
 région +4 ('PREP') : $t=-2.047$, $p=.0462$
3. un effet de longueur
 région +3 ('SUB V') : $t=2.064$, $p=.0436$
 région +4 ('PREP') : $t=5.381$, $p<.001$

Cela signifie que, dans les régions +3 et +4, les temps de lecture sont significativement plus élevés pour quand-INT que pour ce-que-INT d'une part et ce-qui-INT d'autre part.

6.2.1.4 Discussion

Les résultats de cette expérience de lecture par présentation auto-segmentée viennent confirmer un fait bien connu concernant les dépendances A-barre : l'asymétrie sujet-objet. En effet, dans les régions +1 et +3, les temps de lecture pour les dépendances objet (ce-que-REL et ce-que-INT) sont plus élevés que ceux pour les dépendances sujet (ce-qui-REL et ce-qui-INT), ce qui reflète la complexité de traitement associée aux dépendances objet²⁵. Il est intéressant de remarquer que la région +3 correspond au verbe de la proposition subordonnée. L'analyseur est donc probablement déjà en train d'anticiper une lacune à ce moment-là, et cela est plus complexe pour les dépendances objet que pour les dépendances sujet.

Nous avons également trouvé un effet en ce qui concerne la comparaison entre les interrogatives indirectes introduites par 'ce qui/ce que', avec un élément nominal, et des interrogatives indirectes plus typiques, sans élément nominal, celles introduites par 'quand'. Les résultats de notre expérience indiquent que la présence d'un élément nominal ne rend pas le traitement des interrogatives indirectes plus complexes. C'est même l'effet inverse qui est observé puisque dans les régions +3 et +4, les temps de lecture sont plus élevés pour les interrogatives indirectes introduites par 'quand'. Outre le fait qu'il s'agisse d'un élément wh, sans élément nominal, 'quand' est également un adjectif alors que 'ce qu-' est un argument. Un adjectif serait-il, de façon générale, plus complexe à traiter qu'un argument? Comme le reporte AKAL, 2017 : 45, CLIFTON JR, SPEER et ABNEY, 1991 (entre autres) ont trouvé une

25. L'effet est marginal, certes, mais cela reste un effet observable.

préférence pour les arguments sur les adjoints sur le plan du traitement. Ils ont analysé les temps de lecture de phrases avec des syntagmes prépositionnels en anglais, ces syntagmes pouvant être analysés soit comme des arguments soit comme des adjoints syntaxiques. Des conditions de plausibilité permettent de faire la désambiguïsation. Il s'est avéré que, lorsque le syntagme prépositionnel était un argument, les temps de lecture étaient plus courts que lorsqu'il s'agissait d'un adjoint. Ainsi, dans notre expérience, le fait que les temps de lecture pour les interrogatives indirectes introduites par 'quand' soient plus longs que ceux pour les interrogatives indirectes introduites par 'ce qu-' est probablement dû au fait que 'quand', contrairement à 'ce qu-', est un adjoint.

Venons-en maintenant à l'effet zéro en ce qui concerne la comparaison entre les relatives et les interrogatives indirectes introduites par 'ce que'. Nous n'avons pas trouvé d'effet de surprise en ce qui concerne les interrogatives puisque les interrogatives indirectes introduites par 'ce qui/ce que' ne sont pas plus difficiles à traiter que les relatives introduites par 'ce qui/ce que' (pas de différences de temps de lecture) pour les adultes. Aucune différence entre les deux structures ne s'est révélée être significative dans aucune région de la subordonnée. Nous pourrions penser que cette absence d'effet est tout simplement due au fait que dans les deux cas nous avons des dépendances impliquant une lacune. Cela contraste avec les résultats de la première expérience de lecture par présentation auto-segmentée indiquant une pénalité pour les relatives comparée aux complétives, qui, elles, ne comportent pas de lacune. Cela vient renforcer le fait que cette pénalité observée pour les relatives était due à un effet de labyrinthe. De plus, dans la première expérience, il s'agissait d'une réelle ambiguïté en termes de traitement linéaire.

Mais on pourrait aller plus loin et affirmer que nous n'avons pas trouvé de différence entre interrogatives indirectes et relatives introduites par 'ce qu-' parce qu'il s'agit en fait d'une seule et même structure syntaxique, soit d'un DP suivi d'un CP, impliquant une lacune. Le coût de traitement serait le même. C'est vers cette hypothèse finale que nous allons tendre dans le Chapitre 7 en revenant à des propriétés syntaxiques (en synchronie) de ces interrogatives indirectes qui ont tout d'un DP, bien qu'elles soient interprétées comme des interrogatives indirectes.

Un autre point important concerne la comparaison entre la grammaire de l'enfant et la grammaire de l'adulte. Alors que les enfants présentaient davantage de difficultés pour la répétition des interrogatives indirectes introduites par 'ce qu-', les adultes ne présentent pas de difficulté de traitement particulière pour ces structures comparées aux relatives. Bien sûr, les résultats d'un test de répétition ne peuvent pas être directement comparables aux résultats d'une expérience de lecture par présentation auto-segmentée. La différence de traitement observée pourrait indiquer que les enfants n'ont pas encore intégré le fait qu'un élément D puisse, dans certaines circonstances, introduire un CP, le tout ayant la fonction de question. Le développement est d'ailleurs similaire à celui observé en diachronie : emploi de 'que' seul avant d'étendre 'ce que' aux interrogatives indirectes. Il se pourrait que les

enfants n'apprennent l'usage interrogatif de 'ce que' que plus tard, à l'école, grâce à un apprentissage explicite.

Comment expliquer cette différence de traitement entre les enfants et les adultes ? Une première hypothèse pourrait être la suivante : l'adulte est capable de donner deux étiquettes distinctes à la structure (DP *versus* CP), contrairement à l'enfant. Le problème est alors le passage de la grammaire de l'enfant à celle de l'adulte. Une deuxième hypothèse serait que cette structure n'est ambiguë syntaxiquement ni pour les enfants, ni pour les adultes : il s'agit d'un DP dans les deux cas. Mais seuls les adultes acceptent de faire suivre un verbe interrogatif par un DP. Il est impossible, sur la base des résultats de cette expérience, de trancher entre ces deux hypothèses. Un retour à la syntaxe est nécessaire (voir le Chapitre 7 qui suit).

6.2.2 Emploi de 'qu'est-ce qu-'

Rappelons que les deux erreurs majeures qu'ont fait les enfants lors du test de répétition étaient l'omission de 'ce' et l'emploi de la forme *wh-* complexe 'qu'est-ce que' (/kɛsk/). L'omission de 'ce' est agrammaticale dans la grammaire adulte et elle ne se retrouve pas dans l'input. Par contre, l'emploi de /kɛsk/ n'est pas agrammatical mais appartient à un français parlé, non-standard. Il serait intéressant de vérifier, de façon anecdotique, avant de clôturer ce chapitre, si /kɛsk/ en tant qu'introducteur d'interrogatives indirectes se retrouve dans un corpus de français du web (FrWaC, BARONI, BERNARDINI, FERRARESI et ZANCHETTA, 2009), qui est donc assez proche d'un français oralisé.

6.2.2.1 Dans FrWaC

Rappelons que nous avons relevé quelques occurrences de /kɛskə/ (dépendances objet) ou /kɛski/ (dépendances sujet) dans la production des proches d'Anaé, de Théophile et d'Antoine (voir Chapitre 4). On trouve également dans le corpus de français du Web FrWaC une petite centaine d'exemples de /kɛsk/ introduisant des subordonnées (109 en tout, voir Tableau 6.7).²⁶

- (36) Pour tous ceux qui **se demandent qu'est-ce que** je peux faire de mes journées.
(Source : clublhuillier.unblog.fr)
- (37) Je voudrais **savoir qu'est-ce que** je peux faire pour motiver mon demi-frère et ma soeur afin de prendre les choses en main avant que cela ne devienne trop grave [...]
(Source : alois.fr)
- (38) et maintenant le contour de mes articles s'est agrandi alors si tu pouvais jeter un coup d'oeil et me **dire qu'est-ce que** je dois modifier ? (Source : sakatruc.cowblog.fr)

26. Requêtes utilisées, de type CQL : [tag="V.+"] [word="qu"] [word="est"] [word="ce"] [word="que"] pour 'qu'est-ce que' dans les propositions subordonnées et [tag="V.+"] [word="qu"] [word="est"] [word="ce"] [word="qui"] pour 'qu'est-ce qui' dans les propositions subordonnées.

	qu'est-ce que	qu'est-ce qui
(se) demander	26	14
savoir	18	15
dire	7	6
comprendre	3	1
voir	1	3
montrer	3	0
expliquer	2	0
découvrir	1	0
trouver	0	1
déterminer	0	1
relativiser	1	0
faire	1	0
clivées	4	1
TOTAL	67	42

TABLE 6.7 – Nombre d’occurrences de ‘qu’est-ce que’ et ‘qu’est-ce qui’ dans le corpus FrWaC en fonction du verbe matrice.

- (39) **Je me demande qu’est-ce qui** peut le préoccuper autant...
(Source : conan-detective.forumpro.fr)
- (40) **J’aimerais savoir qu’est-ce qui** chez vous pourrait faire que vous coupiez subitement les ponts avec quelqu’un? [Source : parislgbt.free.fr]
- (41) Toujours dans mon effort pour que notre travail vous serve un peu à **voir qu’est-ce qui** intervient comme création en philosophie, [...] (Source : univ-paris8.fr)

Il est intéressant de remarquer que certains exemples (avec le verbe ‘faire’) sont des relatives (de façon non ambiguë) ou des structures proches (clivées). Comme nous l’avons évoqué à la fin du Chapitre 3, il semblerait qu’un mouvement inverse à celui de ‘ce que’ soit amorcé en français non standard avec /kɛsk/ : alors que ‘ce que’ est d’abord apparu dans les relatives puis s’est étendu aux interrogatives indirectes, /kɛsk/ qui, à la base introduit des questions directes, apparaît naturellement pour introduire des questions indirectes en français non standard, mais commence à s’étendre aux relatives. Le fait qu’une forme intrinsèquement interrogative commence à être utilisée dans les relatives est très intéressant, car cela montre bien qu’il n’existe pas d’étanchéité entre les interrogatives indirectes et les relatives, qu’un passage est possible entre les deux structures. Ce passage est d’ailleurs le miroir de l’extension de ‘ce qu-’, intrinsèquement de nature nominale, aux interrogatives indirectes.

Précisons que la proportion de ces interrogatives indirectes est minime comparée à celle des interrogatives indirectes standard introduites par ‘ce que’. En effet, si l’on prend les deux

verbes interrogatifs pour lesquels nous avons relevé le plus d'exemples de /kesk/, à savoir '(se) demander' et 'savoir', on trouve 4997 résultats dans FrWaC pour '(se) demander ce qu-' (3467 pour '(se) demander ce que' et 1530 pour '(se) demander ce qui') et 15128 résultats pour 'savoir ce qu-' (10420 pour 'savoir ce que' et 4708 pour 'savoir ce qui').^{27 28}

6.3 Conclusion

Pour conclure, le premier cas d'ambiguïté étudiée a montré que les relatives introduites par 'ce que', qui sont des dépendances *filler-gap*, sont moins bien jugées et plus lentes à lire que des complétives introduites par 'ce que', qui ne contiennent pas de lacune. Cela vient confirmer le Principe de Chaîne Minimale, principe selon lequel on évite de postuler une lacune, quand la grammaire le permet. Par contre, dans le deuxième cas d'ambiguïté (qui est une ambiguïté syntaxique mais pas au niveau du traitement linéaire), confrontant les relatives introduites par 'ce que' aux interrogatives indirectes introduites par 'ce que', le Principe de Chaîne Minimale n'était pas pertinent. Aucune préférence pour les relatives n'a été relevée contrairement aux interrogatives indirectes. Il s'agit certes de deux dépendances *filler-gap*, ce qui pourrait expliquer pourquoi leur traitement est similaire. Il n'y aurait alors pas trois 'ce que' distincts mais seulement deux 'ce que' (un sans lacune et un avec lacune).

Les enfants français entre 3 et 6 ans ont au contraire davantage de difficultés à répéter correctement des interrogatives indirectes que des relatives introduites par 'ce que'. L'enfant n'a peut-être pas encore intégré qu'une séquence de ce type, avec un élément nominal, pouvait être le complément d'un verbe interrogatif. Ce n'est peut-être que plus tard, lors d'un apprentissage explicite à l'école, qu'ils assimileront cela.

L'absence de différence de traitement entre les relatives et les interrogatives indirectes par les adultes est compatible avec une autre analyse que nous n'avons pas encore explorée jusqu'ici : ces deux structures introduites par 'ce qu-' ne sont peut-être qu'une seule et même structure. Ceci nous amène naturellement vers le Chapitre 7 et l'analyse finale de ces interrogatives indirectes : il s'agit en réalité de relatives à tête allégée, ayant pour tête le démonstratif 'ce'. Leurs propriétés syntaxiques vont être détaillées dans le Chapitre 7, montrant que ces interrogatives se comportent comme des relatives et non pas comme des interrogatives indirectes typiques.

27. Notons que la proportion de dépendances sujet est bien inférieure à la proportion de dépendances objet. Cette différence de fréquence est probablement due au fait que 'ce' est inanimé et qu'il est plus courant d'avoir un objet inanimé qu'un sujet inanimé.

28. Notons également que les données absolues sont probablement sous-estimées, à la fois pour l'emploi de 'qu'est-ce que/qu'est-ce qui' et celui de 'ce que/ce qui' dans les propositions subordonnées, car je n'ai pas cherché 'qu'est-ce qu'/'ce qu'', formes avec élision du schwa utilisées lorsque le mot d'après commence par un son vocalique.

Cinquième partie
Retour à la syntaxe

Chapitre 7

Quand des interrogatives se déguisent en relatives¹

Rappelons le problème central du présent travail. On trouve en français des interrogatives indirectes, qui ressemblent en tout et pour tout à des relatives à tête allégée : elles sont introduites par le démonstratif ‘ce’, suivi du complémenteur (qui/que). Cette stratégie vient combler l’impossibilité d’employer ‘que’ ou ‘quoi’ en français lorsque le référent est inanimé, dans les contextes enchâssés. Le problème est le suivant : ni ‘ce’ ni ‘qui/que’ n’étant wh-, d’où vient leur nature de question ? De plus, un verbe tel que ‘se demander’ sélectionne un CP, et non pas un DP, alors comment peut-il être satisfait par une proposition introduite par un élément D ?

Dans le Chapitre 2, nous avons formulé plusieurs hypothèses pour tenter de résoudre ce problème. La première était que ‘ce’ pouvait être un renforçateur du clitique ‘que’ (mot wh). La deuxième piste que nous avons explorée était celle d’analyser ‘ce que’ comme étant une partie de ‘qu’est-ce que’. Ensuite, nous avons exploré l’hypothèse de l’incorporation, hypothèse selon laquelle ‘ce que’ aurait subi un processus de grammaticalisation se manifestant concrètement par une incorporation de ‘ce’ avec un opérateur wh- nul et le complémenteur ‘que’, donnant naissance à *səkə*, nouveau mot wh-.

L’étude diachronique présentée dans le Chapitre 3 nous a révélé que ‘ce que’ était d’abord apparu dans les relatives, avant de s’imposer dans les interrogatives indirectes, tardivement, vers le milieu du 17ème siècle. ‘Ce que’ est donc intrinsèquement de nature nominale.

Les données d’acquisition présentées dans le Chapitre 4 nous ont montré que les enfants analysaient ‘ce que’ comme une séquence D+C dans les relatives, mais aussi dans les in-

1. J’ai présenté une partie des données décrites dans ce Chapitre ainsi qu’une version antérieure de l’analyse présentée dans ce Chapitre à deux conférences internationales, *Conference of the Student Organisation of Linguistics in Europe (ConSOLE)* (KONRAD, 2018b) et *Incontro di Grammatica Generativa (IGG)* (KONRAD, 2018c), mais aussi lors des *Rencontres lyonnaises des jeunes chercheurs en linguistique historique* (KONRAD, 2019).

terrogatives indirectes, où ils ont tendance à omettre le démonstratif ‘ce’ ou à employer la forme wh- complexe ‘qu’est-ce que’. Nos résultats ont également permis de déceler un avantage pour les relatives, mais seulement dans le cas de ‘ce que’, pas dans le cas de ‘qui’, mot wh- sans élément D.

Enfin, dans le Chapitre 6 portant sur le traitement des structures en question par les adultes, nous n’avons pas trouvé de différence significative entre les relatives à tête allégée et les interrogatives indirectes introduites par ‘ce qu-’, pour ce qui est des temps de lecture. ‘Ce que’ n’est donc pas plus difficile à traiter après un verbe interrogatif. Cette absence de différence de traitement pourrait peut-être être le reflet d’une absence de différence structurale.

Ces résultats nous amènent à formuler notre hypothèse finale qui est la suivante : ces mystérieuses interrogatives indirectes sont en réalité ce qu’elles semblent être, c’est-à-dire des relatives à tête allégée.

7.1 Un comportement syntaxique inhabituel

Pour déterminer la nature de ces « questions », nous devons donc aller au-delà de la simple observation de leur distribution, et leur appliquer des outils diagnostiques plus sophistiqués.

Je propose que nous examinions deux propriétés syntaxiques permettant habituellement de différencier relatives et interrogatives indirectes, et que nous observions le comportement des interrogatives indirectes introduites par ‘ce que’ vis-à-vis de ces deux propriétés : les effets d’îlot d’une part et les effets de *matching* d’autre part.

7.1.1 L’extraction

Une propriété syntaxique bien connue différenciant les phrases modifiant un élément nominal, comme les relatives, des phrases simples, y compris les interrogatives, concerne les effets d’îlot. En effet, alors qu’une relative est un îlot fort, une interrogative indirecte est un îlot faible^{2 3}. Cela signifie qu’on ne peut extraire aucun syntagme d’une relative sans que cela conduise à un résultat agrammatical ; par contre, l’extraction de certains syntagmes est

2. Voir Ross, 1967 pour une première description des différentes configurations d’îlots ; voir SZABOLCSI, 2006 pour une description détaillée des îlots forts et des îlots faibles.

3. Pour le français en particulier, GODARD, 1988, a montré que les interrogatives indirectes étaient des îlots faibles.

- (1) a. quel livre m’as-tu dit où ranger ?
 b. ? à quels enfants a-t-on décidé quel cadeau apporter ?
 c. ?? à quels enfants m’as-tu précisé quel cadeau tu as offert ?

possible dans le cas des interrogatives indirectes. Ce contraste est valable même pour les relatives libres et les interrogatives indirectes, qui sont pourtant superficiellement identiques (RIZZI, 1982). En particulier, l'extraction d'un argument est autorisée, mais l'extraction d'un adjoind rend la phrase agrammaticale. Cette propriété est illustrée dans les exemples qui suivent, avec des subordonnées introduites par 'qui', et des subordonnées introduites par 'où'.

Subordonnées introduites par 'qui'

Relatives libres

- (1) # C'est avec un cric_i que tu connais [qui sait réparer la voiture t_i]. [extraction d'un adjoind]
- (2) *C'est cette machine_i que tu connais [qui sait réparer t_i]. [extraction d'un argument]

Interrogatives indirectes

- (3) # C'est avec un cric_i que tu ignores [qui sait réparer la voiture t_i]. [extraction d'un adjoind]
- (4) C'est cette machine_i que tu ignores [qui sait réparer t_i]. [extraction d'un argument]

Subordonnées introduites par 'où'

Relatives

- (5) # C'est en avion_i que tu iras en train [où je te dirai d'aller t_i]. [extraction d'un adjoind]
- (6) *C'est ce tableau_i qu'il aime [où j'ai placé t_i]. [extraction d'un argument]

Interrogatives indirectes

- (7) # C'est en avion_i que tu ignores [où je te dirai d'aller t_i]. [extraction d'un adjoind]
- (8) C'est ce tableau_i qu'il ignore [où j'ai placé t_i]. [extraction d'un argument]

La question cruciale ici est la suivante : comment les interrogatives introduites par 'ce que' se comportent-elles vis-à-vis de l'extraction ? Comme des interrogatives, c'est-à-dire des îlots faibles, ou bien comme des relatives, c'est-à-dire des îlots forts ? Vérifions d'abord que les relatives introduites par 'ce que' sont bien des îlots forts.

Comme le montrent les exemples en (1), tirés de GODARD, 1988 : 87, l'extraction d'un argument est possible dans les interrogatives indirectes en français, bien qu'il y ait des facteurs qui facilitent l'extraction, comme, par exemple, le fait que le verbe à l'intérieur de l'interrogative indirecte soit à l'infinitif (d'où le contraste entre (1b) et (1c)).

Subordonnées introduites par ‘ce que’

Relatives

- (9) # C’est avec du fusain_i que tu aimes [ce que Paul a dessiné t_i]. [extraction d’un adjoind]
- (10) *C’est à Marie_i que tu aimes [ce que Paul a donné t_i]. [extraction d’un argument]

Les relatives introduites par ‘ce que’, en tant que relatives, sont donc bien des îlots forts : ni l’extraction d’un adjoind ni l’extraction d’un argument ne sont possibles sans que cela conduise à un résultat agrammatical.

Observons maintenant le comportement des interrogatives introduites par ‘ce que’.

Interrogatives

- (11) # C’est avec du fusain_i que tu ignores [ce que Paul a dessiné t_i]. [extraction d’un adjoind]
- (12) *C’est à Marie_i que tu ignores [ce que Paul a donné t_i]. [extraction d’un argument]

Aussi surprenant que cela puisse paraître, ce type d’interrogative indirecte se comporte bel et bien comme des relatives, étant donné qu’on ne peut extraire ni un adjoind ni un argument.

Notons que les données ci-dessus sont basées sur mes propres jugements et sur ceux de 10 collègues francophones, qui ne sont pas forcément syntacticiens et qui ne sont pas spécialistes des îlots. Je leur ai demandé de noter l’acceptabilité des phrases ci-dessus sur une échelle de 0 à 10, 0 correspondant à ‘pas naturel du tout’ et 10 correspondant à ‘parfaitement acceptable’. Le tableau 7.1 indique les notes moyennes pour chaque type de phrase et selon la nature du syntagme extrait (adjoind ou argument).

	Extraction ADJOIND	Extraction ARGUMENT
INT ce que	2	2,7
INT où	3	6
INT qui	0,9	3,6
REL ce que	2,2	1,2

TABLE 7.1 – Notes moyennes attribuées à chaque type de phrase selon que l’on extrait un adjoind ou bien un argument.

On voit bien que les INT *ce que* se comportent différemment des interrogatives indirectes typiques, introduites par les mots *wh-* ‘qui’ et ‘où’. En effet, alors que dans les premières on n’observe pas de grande différence de jugements entre l’extraction d’un adjectif et l’extraction d’un argument, l’extraction de l’argument est nettement plus acceptable dans les deuxièmes. Ces résultats donnent seulement une indication. Il faudrait bien évidemment vérifier cela dans des conditions expérimentales, avec plusieurs items pour chaque condition, avec des distracteurs et avec beaucoup plus de participants.

Mais il semblerait que les INT *ce que* sont des îlots forts, alors que les INT *où* et les INT *qui* sont des îlots faibles. Nous savons que, dans les relatives, ni l’extraction de l’argument ni l’extraction de l’adjectif ne sont possibles sans que cela conduise à un résultat agrammatical. Les jugements indiqués dans le tableau 7.1 montrent bien que les REL *ce que* se rapprochent des INT *ce que*, puisque ni dans un cas, ni dans l’autre, l’extraction d’un argument n’est pas largement plus acceptable que l’extraction d’un adjectif. Cela nous indiquerait que ces interrogatives sont des DP complexes, et non pas de simples CPs. Aurait-on une autre évidence qui permettrait de consolider cette analyse ?

7.1.2 Les effets de *matching*

Une deuxième propriété syntaxique différencie clairement certaines relatives des interrogatives indirectes. Il s’agit des effets de *matching* (discuté pour la première fois par GRIMSHAW, 1977). On sait que seules les relatives libres sont contraintes par la nécessité de *matching* (cf. Section 1.2.2.1.1.3 du Chapitre 1).

(13) Je travaille sur quoi tu travailles.

(14) *J’apprécie sur quoi tu travailles.

L’exemple (13) est grammatical en français, car le verbe de la proposition matrice, tout comme le verbe de la proposition subordonnée, sélectionne un PP (*travailler sur qqch*). Par contre, l’exemple (14) est agrammatical car le verbe de la proposition matrice ne sélectionne pas un PP, mais un DP (*apprécier qqch* et non pas *apprécier sur qqch*), tandis que le verbe de la proposition enchâssée sélectionne toujours un PP. Il y a donc une contradiction de Cas (*mismatching*) qui n’est pas tolérée.

Il n’y a aucune contrainte de ce type dans le cas des interrogatives indirectes. L’exemple ci-dessous est parfaitement grammatical.

(15) Je me demande sur quoi tu travailles.

Ceci suggère que, dans les interrogatives indirectes, la préposition est interne au CP : elle est dans la position de [Spec, CP] puisqu’il y a *pied-piping* en français. Par contre, les relatives libres étant des DP, la seule position possible pour la préposition est externe au DP, dans une position sélectionnée par le verbe matrice.

Dans les relatives pleines, si une préposition précède la tête de la relative, alors il faut que la préposition satisfasse la sélection du verbe matrice, comme l'illustre le contraste entre (16) et (17).

(16) Je travaille **sur** le sujet que tu m'as indiqué.

(17) *J'apprécie **sur** le sujet sur lequel tu travailles.

En effet, en (16), le verbe 'travailler' sélectionne un PP et donc la phrase est grammaticale. Par contre, en (17), le verbe 'apprécier' sélectionne un DP, et donc la préposition 'sur' n'est pas appropriée et la phrase est agrammaticale. Mais lorsque la préposition suit la tête, elle satisfait naturellement la sélection du verbe enchâssé mais elle n'a pas à satisfaire la sélection du verbe matrice.

(18) J'apprécie le sujet **sur** lequel tu travailles.

La grammaticalité de (18) s'explique par le fait qu'une couche DP externe (la tête de la relative) satisfait le verbe matrice 'apprécier'. D'ailleurs, deux prépositions peuvent être présentes indépendamment, l'une sélectionnée par le verbe matrice et l'autre par le verbe enchâssé (voir (19)).

(19) Son travail tend **vers** le sujet **sur** lequel tu travailles.

Les relatives à tête allégée se comportent comme les relatives pleines vis-à-vis du *matching*. Si une préposition précède la tête légère 'ce', alors elle doit être correctement sélectionnée par le verbe matrice, car la préposition est visible de l'extérieur du DP relatif.

(20) Je travaille **sur** ce que tu m'as indiqué.

(21) *J'apprécie **sur** ce que tu travailles.

Par contre, lorsque la préposition suit la tête légère 'ce', elle n'a pas à satisfaire la sélection du verbe matrice, du fait de la présence d'une tête. On peut alors postuler que 'sur quoi' est en [Spec, CP].

(22) J'apprécie ce **sur** quoi tu travailles.

Ceci montre bien que 'ce' est externe au CP et satisfait la sélection du verbe matrice, 'apprécie', qui sélectionne un DP.

La contrainte de *matching* observée dans les relatives libres est donc due à l'absence de tête : le mot *wh-* doit satisfaire à la fois la sélection du verbe matrice et celle du verbe enchâssé. Cette contrainte n'est pas valable dans les relatives à tête allégée car, par définition, la tête est bien là, bien que légère.

Passons maintenant aux interrogatives indirectes. Regardons si les interrogatives indirectes introduites par ‘ce que’ sont contraintes par ces effets de *matching* ou non, lorsqu’il y a *pied-piping*. Si ce sont des CPs, comme les interrogatives indirectes typiques, alors elles ne devraient obéir à aucune contrainte de ce type (voir (15)).

(23) Je m’interroge sur ce que tu fais.

(24) *Je me demande sur ce que tu travailles.

En (23), la préposition ‘sur’ est correctement sélectionnée par le verbe matrice (‘s’interroger sur’), et donc la phrase est grammaticale. Ceci contraste avec l’exemple (24), où l’interrogative indirecte introduite par ‘ce que’ est précédée par une préposition (‘sur’) qui n’est pas sélectionnée par le verbe matrice, et donc la phrase est agrammaticale. Ainsi, il y a bel et bien un effet de *matching* dans le cas des interrogatives indirectes introduites par ‘ce que’. En (24), l’interrogative indirecte en ‘ce que’ est précédée par une préposition correctement sélectionnée par le verbe enchâssé, mais pas par le verbe matrice : le résultat est agrammatical. La préposition est donc externe au CP et accessible à la sélection du verbe principal, comme dans une relative à tête allégée.

Ces interrogatives indirectes se comportent exactement comme les relatives à tête allégée : si la préposition précède ‘ce’, comme en (24), alors elle doit satisfaire la sélection du verbe matrice. Mais si ‘ce’ précède la préposition, la phrase devient grammaticale.

(25) Je me demande (ce) sur quoi tu travailles.

L’exemple (25) montre assez clairement que ‘ce’ est externe au CP : c’est ‘quoi’ qui fait le mouvement A-barre et ‘ce’ est un D externe qui ajoute une couche DP au CP. Précisons que ‘ce’ dans ‘ce sur quoi’ est facultatif dans les interrogatives indirectes. Notons que cette possibilité n’est pas restreinte à la préposition ‘sur’.

- (26) a. Je me demande (ce) à quoi tu penses.
 b. Je me demande (ce) **en** quoi tu crois.
 c. Je me demande (ce) **contre** quoi tu te bats.
 d. Je me demande (ce) **pour** quoi tu te bats.
 e. Je me demande (ce) **avec** quoi tu as construit cela.
 f. Je me demande (ce) **vers** quoi cela tend.

Ces données empiriques semblent indiquer, encore une fois, que ces interrogatives particulières sont à analyser comme des DPs et non pas comme des CPs. Nous allons d’ailleurs voir dans la section qui suit que la distinction entre DP et CP n’est pas aussi nette qu’on pourrait le penser.

Dans les sections qui suivent, je vais présenter deux problèmes qui se posent si l'on analyse les interrogatives indirectes introduites par 'ce que' comme des relatives à tête allégée, et montrer qu'il s'agit en réalité de faux problèmes. Le premier est le problème de l'étiquette : comment un verbe interrogatif, qui sélectionne un CP, peut-il être satisfait avec une relative à tête allégée, qui est un DP ? Le deuxième est celui de l'interprétation : comment une relative peut-elle être interprétée comme une interrogative indirecte ?

7.2 Le problème de l'étiquette

7.2.1 La sélection

7.2.1.1 Vision traditionnelle

Il est communément admis que les traits de sélection des verbes distinguent entre un argument nominal et un argument de type phrastique au niveau purement syntaxique (sélection catégorielle). Par exemple, un verbe tel que 'ouvrir' sélectionne un argument de type nominal, un DP, alors qu'un verbe comme 'penser' sélectionne un argument de type phrastique, un CP.

GRIMSHAW, 1979 propose que, en plus de cette sélection syntaxique, les prédicats sont également dotés d'un trait de sélection sémantique. On parle alors de s-sélection (sélection sémantique) : en plus de sélectionner la catégorie syntaxique de leurs arguments, les prédicats sélectionnent aussi leur contenu sémantique. L'hypothèse de GRIMSHAW, 1979 est basée sur le fait suivant : il existe des prédicats qui sélectionnent le même type de complément sémantique mais la réalisation syntaxique de leur complément n'est pas forcément la même. C'est le cas par exemple de *ask* et *wonder* en anglais, ou de 'demander' et 'se demander' en français. Les deux verbes prennent une question comme complément (sélection sémantique), mais la réalisation de cette question est différente sur le plan syntaxique : 'demander' autorise soit un CP soit un DP, comme illustré en (27), alors que 'se demander' requiert un CP, comme le montrent les exemples en (29).

- (27) a. Cédric m'a demandé [CP quelle heure il était].
 b. Cédric m'a demandé [DP l'heure].
- (28) a. Le client lui a demandé [CP quel prix cela coûtait].
 b. Le client lui a demandé [DP le prix].
- (29) a. Cédric se demande [CP quelle heure il est].
 b. *Cédric se demande [DP l'heure].
- (30) a. Le client se demande [CP quel prix ça coûte].
 b. *Le client se demande [DP le prix].

On admet généralement que le verbe ‘se demander’ ne peut pas sélectionner de DP. Or, comme me l’a fait remarquer Klaus Abels⁴, un exemple tel que (31) est parfait, et pourtant il s’agit bien d’un DP.

(31) Je me demande [DP la même chose].

Toutefois, le DP ‘la même chose’ en (31) n’est pas à mettre sur le même plan que le DP ‘l’heure’ (29b). Dans le premier cas, il s’agit en effet d’une proforme pour le CP tout entier ‘quelle heure il est’, comme dans l’exemple (32).

(32) Je me le demande.

Le fait qu’un pronom puisse pronominaliser un CP montre bien que la distinction catégorielle entre DP et CP n’est pas si nette.

7.2.1.2 Une approche alternative : Pesetsky (1982)

PESETSKY, 1982 postule que la sous-catégorisation syntaxique peut-être éliminée de la grammaire en faisant appel à d’autres aspects de la théorie de la grammaire déjà existants, et plus spécifiquement la Théorie du Cas. Il suggère que les verbes qui sélectionnent (d’un point de vue sémantique) un argument d’un certain type tel que question, proposition, exclamation, sont tous en principe compatibles avec à la fois des phrases et des NPs, c’est-à-dire à la fois des CPs et des DPs selon la terminologie actuelle. Cependant, les DPs ne peuvent apparaître que s’ils sont marqués pour le cas, c’est-à-dire que si l’item lexical en question (le verbe) assigne le cas accusatif. Les CPs n’ont au contraire pas besoin de cas (STOWELL, 1981). Ainsi, les propriétés de la sélection catégorielle peuvent être dérivées des propriétés de la sélection sémantique. Par exemple, *ask* peut assigner un cas à son complément alors que *wonder* ne peut pas. Comme les DPs ont besoin d’un cas⁵, alors il s’ensuit que *wonder* ne peut pas prendre de DP comme complément.

Cette perspective a été contestée par ROTHSTEIN, 1992, entre autres. Comme le fait remarquer ROTHSTEIN, 1992, la proposition de Pesetsky reste spéculative puisqu’il n’y a pas de moyen permettant de prédire, d’une façon générale, si un verbe est marqué [+Cas] ou non. L’idée de faire la connexion entre la sélection et le cas est intéressante, d’autant plus que nous avons vu que, en français, dans les complétives introduites par *P+ce+que*, le démonstratif ‘ce’ est probablement là pour des raisons casuelles. Toutefois, si Pesetsky avait raison, le verbe ‘se demander’ ne pourrait pas avoir de DP comme complément, et le même problème se poserait pour le pronom accusatif en (32). Or, nos interrogatives indirectes déguisées en relatives (introduites par ‘ce que’) sont clairement des DPs. ‘Se demander’ aurait-il alors un trait [+Cas]?

4. Lors de la conférence ConSOLE à UCL, à Londres, en février 2018.

5. Le Filtre du Cas stipule que tout DP explicitement réalisé doit avoir un cas d’assigné (*Case Filter*, CHOMSKY, 1981).

7.2.2 Origine nominale des complémenteurs et existence de CPs nominalisés

Dans cette sous-section, je vais prouver, en m'appuyant sur des données du français mais aussi de nombreuses autres données dans plusieurs langues, que cette distinction catégorielle aussi catégorique entre DP et CP n'est pas pertinente.

7.2.2.1 Origine nominale des complémenteurs

Un fait frappant est que les complémenteurs ressemblent très souvent à des éléments nominaux (ROBERTS et ROUSSOU, 2003, MANZINI et SAVOIA, 2003, MANZINI, 2014, BAUNAZ et LANDER, 2018). Le complémenteur anglais 'that' est homophone au démonstratif 'that'.

- (33) a. I think **that** you knew it.
 b. **That** skirt is black.

En allemand et en néerlandais, le complémenteur 'dass/dat' est également homophone au démonstratif 'das/dat' (BAUNAZ et LANDER, 2018 : 3).

Le complémenteur italien 'che' est homophone au mot wh- 'che' et au déterminant wh-'che' sélectionnant des NPs complexes interrogatifs (MANZINI et SAVOIA, 2003 : 87).

- (34) a. Mi hanno detto **che** vieni domani.
 1SG.DAT ont dit que viens demain
 'Ils m'ont dit que tu viens demain.'
 b. **Che** fai?
 que fais
 'Que fais-tu ?'
 c. **Che** camicia hanno portato?
 quelle chemise ont porté
 'Quelle chemise ont-ils porté ?'

En finnois, le complémenteur 'että' n'est pas homophone avec le démonstratif 'tä', mais on peut noter l'existence d'une proximité phonétique. D'ailleurs, historiquement, 'että' est dérivé du démonstratif *e- (forme reconstruite, que l'on retrouve en hongrois par exemple, avec le démonstratif 'ez'). Le complémenteur finnois semble donc bel et bien avoir une origine nominale (BAUNAZ et LANDER, 2018 :5).

7.2.2.2 Existence de CPs nominalisés

Dans plusieurs langues, les CPs en position de sujet / d'objet ou en position oblique sont précédés par un élément de type D, donc un élément de nature nominale. Cette nominalisation permet probablement aux CPs de pouvoir recevoir un cas. Dans les sections qui

suivent, je présente des données en français puis dans d'autres langues telles que l'espagnol, à partir de PICALLO, 2002, le grec moderne, à partir de ROUSSOU, 1991, et le persan, à partir de ÖHL et LOTFI, 2007, montrant que la nominalisation de CPs n'est pas restreinte au français et qu'elle n'est pas un cas rare dans les langues.

7.2.2.2.1 Propriétés nominales du CP en français Nous avons déjà vu qu'en français, certaines complétives en position oblique (après une préposition) sont obligatoirement introduites par l'élément nominal 'ce'.

- (35) Je veillerai à ce qu'il se sente comme chez lui ici.
 (36) Je tiens à ce que Luc me rembourse l'argent qu'il m'a emprunté.
 (37) Le problème vient de ce que le chanteur ne s'est pas présenté sur scène.

La présence de 'ce' conduit à un résultat agrammatical dans les complétives en position objet (voir (38), ZARING, 1992 : 57).

- (38) Je regrette *ce qu'il ne nous ait pas dit la vérité.

ZARING, 1992 décrit les propriétés de P+ce+CP dans les complétives en français. Elle postule que la préposition dans ces constructions n'est pas une réelle préposition, mais plutôt une manifestation morphologique d'assignation du Cas. Quant au statut de 'ce', qui nous intéresse particulièrement ici, ZARING, 1992 analyse 'ce' comme un type de déterminant phrastique dans ces constructions, comme un D° prenant pour complément un CP tensé.⁶

Elle montre que dans les complétives en question, 'ce' n'est pas un pronom référentiel et il ne peut pas être employé seul, sans la proposition subordonnée (voir (39), ZARING, 1992 : 56). Il a un statut de clitique.

- (39) T'habitués-tu à ce qu'elle dorme pendant la journée?
 *Oui, je m'habitue à ce.

De plus, 'ce' doit être obligatoirement adjacent à la proposition enchâssée : on ne peut pas insérer de matériel lexical, tel un adverbe, entre 'ce' et 'que' (voir (40), ZARING, 1992 : 56).

- (40) Je m'habitue à ce (*évidemment) qu'elle dorme pendant la journée.

Une autre caractéristique de ces constructions est que P+ce+CP agit comme un constituant. En effet, elles peuvent être disloquées à droite (voir (41), ZARING, 1992 : 67) et il est généralement admis que des éléments disloqués doivent être des constituants.

6. Voir notamment HUOT, 1981 et AUTHIER et REED, 2010 pour une analyse alternative, selon laquelle 'à ce que' et 'de ce que' dans les complétives seraient des complémenteurs agglomérés.

(41) J'y veillerai, à ce qu'il se couche de bonne heure.

En ce qui concerne l'extraction, P+ce+CP est un îlot : on ne peut pas extraire un adjectif sans que cela conduise à un résultat agrammatical (voir (42), ZARING, 1992 : 64).

(42) ***Comment**_i est-ce que tu t'habitues à ce qu'elle fasse la vaisselle t_i ?

L'extraction d'un argument semble davantage acceptable que l'extraction d'un adjectif, mais les jugements des locuteurs ne sont pas uniformes (voir (43) et (44), ZARING, 1992 : 65).

(43) **Qu'est-ce que**_i tu t'habitues à ce qu'elle fasse t_i toute seule ?

(44) ?**Dans quel pays**_i ont-ils renoncé à ce que l'ONU envoie des soldats t_i ?

Par ailleurs, il est généralement admis que les arguments phrastiques (CPs) n'ont ni cas structural (STOWELL, 1981) ni spécification en traits de personne/nombre/genre (CHOMSKY, 1995, IATRIDOU et EMBICK, 1997). Ce postulat est généralement corroboré par le fait qu'un prédicat ne semble pas pouvoir s'accorder avec ce type de sujet. Or, des données du français vont dans le sens opposé : les exemples (45) et (46), tirés de DAVIES et DUBINSKY, 2001 : 259-260, montrent que l'accord pluriel est possible.

(45) *Que le défilé continue* ou *qu'il soit annulé* a/ont été discuté(s) par les mêmes gens à différentes occasions.

(46) *Séjourner dans les montagnes* et *longer la côte* me paraissent une façon admirable pour connaître la vraie France.

Cela est également confirmé par des données trouvées par MOURET dans sa thèse de doctorat (MOURET, 2007), indiquant que les sujets infinitifs coordonnés permettent aussi bien l'accord singulier que l'accord pluriel et ce, même au sein d'un même article (voir exemples en (47)).

(47) a. Ne pas savoir se servir d'un ordinateur (66,5%), travailler à temps réduit (64,9%) et être âgé de plus de quarante-cinq ans (52,3%) **constitue** un handicap, lors d'une promotion.

b. A l'inverse, être un homme (48,6%) et avoir des diplômes (85,6%) **sont** manifestement des atouts [...].

Un autre fait troublant, toujours en français, concerne les causatives. Dans ces constructions, l'argument causataire peut être soit accusatif soit datif, selon qu'il y a un autre DP dans le domaine ou non.

(48) a. Elle l'a fait partir.

b. *Elle **lui** a fait partir.

(49) a. Elle **lui** a fait manger **des choux de Bruxelles**.

b. *Elle l'a fait manger **des choux de Bruxelles**.

En français standard, le causataire est au cas accusatif, comme en (48a), s'il n'y a pas de DP objet dans la proposition infinitive. Par contre, le causataire est au cas datif, comme en (49a), qui est le cas marqué, lorsqu'il y a un DP objet, déjà marqué accusatif, dans la proposition infinitive, ici 'les choux de Bruxelles'.⁷ Le cas du causataire est donc un cas dépendant dans le cadre de la Théorie du cas dépendant (*Dependent case theory*, BÁRÁNY et SHEEHAN, 2019)⁸.

Ce qui est particulièrement intéressant pour notre discussion ici est qu'un CP est en mesure de forcer le cas datif dans les constructions causatives, exactement comme un DP.

- (50) a. Elle **lui** a fait dire [qu'elle sortait].
 b. *Elle l'a fait dire [qu'elle sortait].

7.2.2.2.2 CPs nominalisés dans des langues autres que le français On trouve des CPs (sans lacune) nominalisés dans d'autres langues que le français, mais la distribution n'est pas la même. En effet, en français, un élément D est obligatoire devant un CP en position oblique. Mais en espagnol, en grec ou encore en persan, les CPs sont nominalisés en position de sujet et d'objet direct.

7.2.2.2.2.1 CPs nominalisés en espagnol On trouve en espagnol des phrases nominalisées sélectionnées par des prédicats factifs et ayant la structure *El que + IP*, littéralement 'le que + IP', comme l'illustre l'exemple (51) [PICALLO, 2002 : 119].

- (51) **El que** creas **que** hay fantasmas en la azotea carece de lógica.
 le COMP croire.2SG COMP il-y-a fantômes dans le grenier manque de logique
 'Le fait que tu croies qu'il y a des fantômes dans le grenier manque de logique.'

7. Notons que LAMIROY et CHAROLLES, 2010 ont toutefois fait remarquer que, lorsqu'on consulte un moteur de recherche comme Google, on observe une confusion entre 'le' et 'lui' dans les causatives avec 'faire', montrant que les deux clitiques sont utilisés dans des cas contraires à la norme.

- (1) Puis, discrètement elle me dit que ça **lui** fait rigoler car elle adore qu'on soit brutal avec elle.
 (2) Mon fils de 7 mois et demi refusait le biberon du matin et du soir (...) et pour le soir plus de biberons je **le** fais manger un petit pot à base de laits et 2 petit suisses et tout se passe bien.

En (1) [LAMIROY et CHAROLLES, 2010 : 2143], le clitique datif 'lui' est employé à la place du clitique accusatif 'la'. Inversement, en (2) [LAMIROY et CHAROLLES, 2010 : 2142], le clitique accusatif 'le' est employé à la place du clitique datif 'lui'.

8.

(1) **Dependent case theory**

Morphologically marked cases (ACC/ERG/DAT) results from a relationship between two DPs rather than from a relationship between a head and a DP.

Bien que ces constructions aient souvent été analysées comme DP+NP+CP, avec une tête nominale phonologiquement nulle correspondant à l'entrée lexicale *hecho* ('fait') [voir DEMONTE, 1977 et PLANN, 1981, entre autres], PICALLO, 2002 propose une structure DP+CP. D'après elle, cette structure permet de rendre compte de l'absence obligatoire de la préposition *de* ('de'), qui est normalement requise pour tous les CPs compléments de noms en espagnol. Cette préposition devrait apparaître même si le nom est phonologiquement nul, ce qui n'est pas le cas.

- (52) a. Lamento el hecho *(de) que no me saludara.
regretter.1SG le fait de COMP NEG me saluer-IPF.SBJV.3SG
- b. Lamento el (*de) que no me saludara.
regretter.1SG le de COMP NEG me saluer-IPF.SBJV.3SG

7.2.2.2.2 CPs nominalisés en grec moderne En grec moderne, les CPs nominalisés sont précédés de l'article défini neutre singulier *to* et sont introduits par le complémenteur *oti/pos* (voir (53) et (54)) ou la particule *na* (voir (55)) ou bien un syntagme *wh-* (voir (56)).

- (53) *to oti ehis filus simeni pola.*
the-NOM that have-2SG friends-ACC mean-3SG much
'That you have friends means a lot.'
- (54) *to oti perase to ksero.*
the-ACC that passed-3SG it-ACC know-1SG
'That he passed I know.'
- (55) *To na ise politikos apeti ...*
the-NOM prt be-2SG politician-NOM requires-3SG
'For you to be a politician requires...'
- (56) *To an tha fighi dhen To gnorizo.*
the-NOM whether AUX leave-3SG not it-ACC know-1SG
'Whether he will leave, I don't know.'

(53) et (55) sont des exemples de CPs nominalisés en position sujet au sein de la proposition matrice ; (54) et (56) sont des exemples de CPs nominalisés en position objet (Roussou, 1991 : 78). De la même façon qu'en français aucun élément ne peut intervenir entre 'ce' et 'que', rien ne peut intervenir entre le D et le CP en grec moderne, comme on peut le voir en (57) [Roussou, 1991 : 78].

- (57) **to oreo oti efighe.*
the-NOM nice that left-3SG

Roussou propose aussi que ces CPs ont une structure DP (D+CP) et que la nominalisation a lieu à des fins d'assignation de cas. Elle montre en effet que ces phrases nominalisées apparaissent dans des positions où le Cas est assigné et donc où la présence du déterminant

est obligatoire pour pouvoir recevoir le cas. Par contre, lorsque le Cas n'est pas requis, on n'observe pas d'insertion du déterminant, comme l'illustrent les exemples (58) et (59).

- (58) *ksero to oti efighe.
 know-1SG the-ACC that left-3SG
 'I know that he left.'
- (59) *protimo to na fevgho.
 prefer-1SG the-ACC PRT leave-1SG
 'I prefer to leave.'

Ici, le Cas n'est pas requis car la sélection d'un CP est spécifiée dans le cadre de la sous-catégorisation du verbe matrice (ROUSSOU, 1991 : 93-94). Ainsi, la présence du déterminant 'to' conduit à un résultat agrammatical.

7.2.2.2.3 CPs nominalisés en persan On rencontre des structures similaires en persan, langue SOV, dans lesquelles les CPs sont précédés du déterminant défini *in* et introduits par le complémenteur *ke*.

- (60) **In** ke gorbeha sir dust darand tabi'i ast
 DET that cats milk like have natural is
 'That cats like milk is natural.'
- (61) Man **in** ke gorbeha sir dust darand *(ra) midanam.
 I DET that cats milk like have OBJ know
 'I know that cats like milk.'

(60) est un exemple de CP nominalisé en position sujet au sein de la proposition matrice; (61) est un exemple de CP nominalisé en position objet (ÖHL et LOTFI, 2007 : 1). Il est obligatoire d'utiliser le déterminant 'in' dans le cas d'objets préverbaux de type CP. Un point à souligner est que le marqueur d'objet direct 'ra', qui est normalement employé avec les DPs objet, est également obligatoire avec les CPs en position d'objet (voir (61)).

ÖHL et LOTFI, 2007 postulent que les CPs préverbaux en persan sont en fait des DPs : avant la réalisation du cas abstrait, un CP doit être sélectionné par D°. D° sélectionne des syntagmes de catégorie [N]. Ils proposent alors que le complémenteur 'ke' est en réalité une tête nominale sélectionnant des phrases. 'Ke' porte donc un trait nominal [+N].

D'après eux, il n'y a pas de raison que les arguments phrastiques et les arguments nominaux soient sujets à des conditions distinctes vis-à-vis du Cas, juste parce que les premiers n'ont pas de marquage casuel explicite dans beaucoup de langues. Ils suggèrent d'ailleurs une modification du Principe du Cas⁹ :

9.

- (62) **Case Principle amended by ÖHL et LOTFI, 2007 : 13**
Every argument chain must be licensed by abstract case.

Le rôle des subordonnants est d'offrir la possibilité aux arguments phrastiques d'entrer dans une chaîne avec une position casuelle. Leur proposition, au-delà des données du persan, est que les CPs qui se trouvent dans une position casuelle sont universellement sélectionnés par D° pour permettre la réalisation du cas abstrait. D° sélectionnant des syntagmes de catégorie [N], l'implication est la suivante : seules les langues qui ont des compléments nominaux peuvent avoir des phrases dans des positions casuelles.

Les auteurs rappellent que Kayne, en 1984, avait déjà suggéré que C était une catégorie potentiellement nominale (KAYNE, 1984), et que Müller en 1993, avait proposé que les CPs étaient des NPs (MÜLLER, 1993)¹⁰.

Ainsi, en français et dans d'autres langues, le CP peut être précédé par un D, sans N. Il ne s'agit donc pas d'une nominalisation *stricto sensu*, mais plutôt d'une nominalisation purement fonctionnelle.

7.3 Le problème de l'interprétation

Nous venons de voir que, d'un point de vue syntaxique, la distinction entre DP et CP n'est pas si nette. Nous avons décrit plusieurs exemples de CPs nominalisés dans plusieurs langues. Pour qu'un CP puisse être employé comme sujet ou complément ou après une préposition, il faut qu'il y ait nominalisation. Ainsi, les phrases ont peut-être toujours une origine nominale. Peut-être qu'il existe des catégories nominales fortes (DP, NP) et des catégories nominales faibles (CP). Si on revient donc au français, on peut voir que le problème de l'étiquette est peut-être un faux problème : DP et CP ne sont pas deux catégories distinctes, et le fait qu'un DP puisse satisfaire la sélection d'un verbe interrogatif suivrait. Mais cette distinction entre DP et CP reste valable sur le plan sémantique. Typiquement, contrairement aux CPs, les DPs sont référentiels. Toutefois, dans le cas des complétives discuté pour le français, 'ce' n'a pas forcément ce caractère référentiel. La question qui se pose désormais si nous analysons les interrogatives indirectes introduites par 'ce que' comme des relatives à tête allégée est donc : comment une phrase relative avec 'ce' comme tête peut-elle être interprétée comme une question ?

(1) **Case Principle (CHOMSKY et LASNIK, 1993 : 561)**
Every realised DP/NP must be assigned abstract case. A chain is visible for Θ -marking if it contains a case-position.

10.

(1) **The NP-shell Hypothesis (MÜLLER, 1993)**
All embedded finite clauses are NPs with a phonetically empty head.

7.3.1 Le cas particulier des ‘questions déguisées’ (*concealed questions*)

Un cas intéressant de DPs interprétés comme des questions est celui des dites *concealed questions* (BAKER, 1968, GRIMSHAW, 1979, HEIM, 1979, ROMERO, 2006, NATHAN, 2006, FRANA, 2006, FRANA, 2017), que je traduis par « questions déguisées ».

Certains DPs sont appelés « questions déguisées » parce qu'ils peuvent être paraphrasés par des interrogatives indirectes de façon intuitive.

- (63) a. Marie a oublié **la capitale de la Norvège**.
 b. Marie a oublié **quelle était la capitale de la Norvège**.

Le DP indiqué en gras ‘la capitale de la Norvège’ en (63a) peut être interprété comme le CP ‘quelle était la capitale de la Norvège’ (63b). La question majeure qui se pose concernant ces structures est : est-ce que ces DPs dénotent des questions ?

Parmi les différentes analyses de ces questions déguisées, nous en retenons une en particulier, qui semble intéressante pour les structures sur lesquelles nous nous focalisons. Selon l'approche en question, ces DPs particuliers sont perçus comme des questions camouflées (BAKER, 1968, GRIMSHAW, 1979). L'idée générale est que les questions déguisées sont en réalité des questions, et que le matériel non prononcé est supprimé par un processus d'ellipse.

- (64) Marie a oublié [CP ~~quelle était~~ la capitale de la Norvège].

D'après BAKER, 1968, ces questions déguisées ne sont possibles qu'avec des prédicats qui sélectionnent des questions *wh-* (tels que ‘savoir’, ‘oublier’, mais pas ‘croire’, ‘penser’). GRIMSHAW, 1979 montre au contraire que les questions déguisées ne sont pas autorisées avec n'importe quel prédicat sélectionnant des questions *wh*. Et c'est là que nous revenons à sa proposition de faire la distinction entre la sous-catégorisation syntaxique et la sélection sémantique de traits du verbe enchâssant la question déguisée. Elle propose que ces structures sont des DPs sur le plan syntaxique mais, sur le plan sémantique (et seulement sur le plan sémantique), il s'agit de questions. C'est pour cela que les questions déguisées ne peuvent pas apparaître sous un verbe tel que ‘se demander’ : bien qu'il sélectionne une question sur le plan sémantique, il ne peut sélectionner qu'un CP sur le plan syntaxique, et pas un DP référentiel. Ainsi, les questions déguisées sont traitées comme des questions par BAKER, 1968 à la fois sur le plan syntaxique et sur le plan sémantique, mais seulement sur le plan sémantique par GRIMSHAW, 1979.

Les questions déguisées sont possibles avec des DPs simples mais aussi des DPs complexes, c'est-à-dire des relatives. Mais certaines restrictions sémantiques semblent entrer en jeu. En particulier, en français, on peut observer un contraste frappant entre des DPs contenant un nom davantage ‘concret’ et des DPs contenant un nom davantage ‘abstrait’, lorsqu'ils sont enchâssés sous le verbe ‘se demander’ (comparer exemples en (65) avec ceux en (66)).

- (65) a. ??Je me demande **la fille** que tu as vue.
 b. ??Je me demande **le livre** que tu as acheté.
 c. ??Je me demande **le plat** que tu as cuisiné.
- (66) a. Je me demande **le temps** qu'il fait à Paris.
 b. Je me demande **la sanction** qu'il va te donner.
 c. Je me demande **le mensonge** qu'il va inventer cette fois-ci.

On pourrait supposer que plus un nom est concret (en ce sens que l'on peut le voir, le sentir, le toucher, le goûter ou l'écouter), plus l'interprétation en tant qu'entité définie est proéminente, moins le DP complexe contenant ce nom peut être interprété comme une question et donc, moins il semble naturel de l'enchâsser sous un verbe tel que 'se demander'. Par exemple, on peut voir/sentir/toucher/goûter un plat, mais on ne peut ni voir, ni sentir, ni toucher, ni goûter un mensonge.

Il est également intéressant de noter que la relative permet l'interprétation de question, car, si l'on ne garde que la tête de la relative, aucun des DPs définis des exemples ci-dessus ne peut être sélectionné par 'se demander'.

- (67) a. *Je me demande la fille.
 b. *Je me demande le livre.
 c. *Je me demande le plat.
 d. *Je me demande le temps.
 e. *Je me demande la sanction.
 f. *Je me demande le mensonge.

J'analyse les questions déguisées en (66) à la BAKER, 1968, c'est-à-dire comme des CPs avec ellipse de la partie interrogative.

- (68) a. Je me demande [CP ~~quel est~~ le temps qu'il fait à Paris].
 b. Je me demande [CP ~~quelle est~~ la sanction qu'il va te donner].
 c. Je me demande [CP ~~quel est~~ le mensonge qu'il va inventer cette fois-ci].

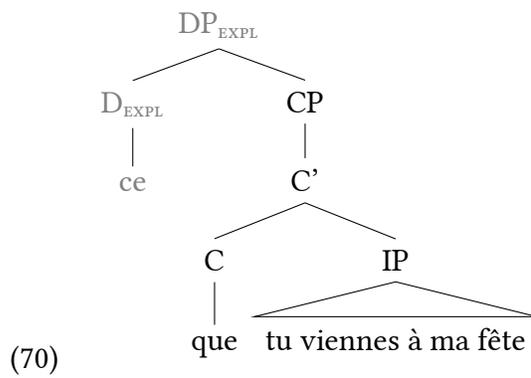
7.3.2 Un 'ce' explétif

Revenons maintenant aux interrogatives indirectes introduites par 'ce qu-' et en particulier à leur interprétation. Elles ressemblent fortement aux dites *concealed questions* : il s'agit de DPs sur le plan syntaxique mais leur interprétation est celle d'une question. Il ne s'agit par contre pas de DPs simples mais de DPs complexes, c'est-à-dire de relatives : ce sont des questions déguisées en relatives. Il est intéressant de noter que l'analyse de BAKER, 1968 serait applicable aux interrogatives indirectes introduites par 'ce que', si l'on suppose que 'ce que' est une partie de 'qu'est-ce que' et que la partie 'qu'est', qui n'est pas prononcée, est supprimée via un processus d'ellipse. Mais nous avons déjà remis en cause la plausibilité

de cette hypothèse dans le Chapitre 2 : le découpage entre 'est' et 'ce' est peu probable car ce n'est pas un constituant ; 'qu'est-ce que' en français contemporain est analysable comme un tout ; en français parlé, c'est plutôt le complémenteur (donc le deuxième 'que' dans la séquence) qui a tendance à être élide.

Il est important de souligner que 'ce' en français n'est pas forcément référentiel. Nous avons déjà vu que, dans les complétives, 'ce' est inséré dans le CP pour des raisons casuelles (après une préposition). Mais 'ce' n'est pas référentiel dans ces constructions et son ajout ne change pas le type d'objet, qui est toujours une proposition. Appelons ce 'ce' dénué de sens 'explétif'.

(69) Je tiens à [ce que tu viennes à ma fête].

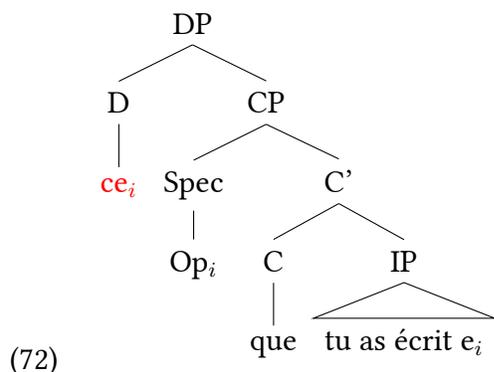


Passons maintenant au statut de 'ce' dans les structures avec lacune, à savoir les relatives et les interrogatives indirectes. Dans les relatives, 'ce' est la tête de la relative et il est référentiel. Il peut être remplacé par un DP plein. On peut ajouter un N et la phrase reste grammaticale.

- (71) a. J'ai lu **ce** que tu as écrit.
 b. J'ai lu **le livre** que tu as écrit.
 c. J'ai lu **ce livre** que tu as écrit.

Nous avons confronté dans le Chapitre 1 les deux principales analyses des relatives, à savoir l'analyse par montée et l'analyse d'appariement, et nous avons postulé une structure DP+CP pour les relatives à tête allégée en français. Dans le Chapitre 2, nous nous sommes questionnés sur l'élément qui se déplaçait dans les relatives à tête allégée : est-ce 'que' ? est-ce 'ce' ? ou bien ni l'un ni l'autre, mais plutôt un opérateur nul ? 'Que' ne peut pas se déplacer dans la position de [Spec, CP] ici car il s'agit du complémenteur et non pas du mot *wh*-. Nous savons cela du fait de l'alternance entre *qui/que* selon que la lacune est en position de sujet ou en position d'objet, mais également parce que 'ce *qui/ce que*' est incompatible avec l'infinitif. Nous avons ensuite envisagé que 'ce' se déplaçait, par montée dans la position de tête de la relative. Or le problème avec une telle dérivation serait que l'on s'attendrait à un effet d'intervention dans les dépendances objet, étant donné que 'ce',

qui se déplace, partage avec le sujet un trait nominal. Le sujet devrait alors agir comme un intervenant et interrompre la dérivation, ce qui conduirait à un résultat agrammatical. Or, les relatives à tête allégée introduites par ‘ce que’, c’est-à-dire avec lacune en position d’objet, sont parfaitement grammaticales en français. Nous étions donc arrivés à la conclusion que ‘ce’ est assemblé à la structure en tant que tête D externe, et que l’élément qui se déplace est alors un opérateur wh- nul. La structure des relatives à tête allégée est illustrée à nouveau en (72).



Il s’agit d’une analyse par appariement et non pas par montée de la tête. En effet, la tête ‘ce’ n’a pas pour position d’origine la position à l’intérieur de la relative, elle est assemblée à la relative de façon externe. L’opérateur nul se déplace dans la position de [Spec, CP], et établit une relation de prédication avec la tête ‘ce’ (comme l’indique l’indexation), de façon à ce que ‘ce’ soit interprété comme l’antécédent de l’opérateur nul.

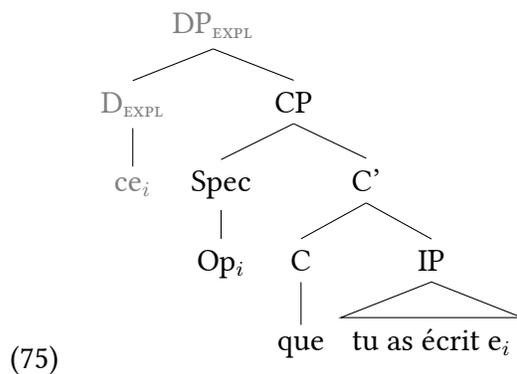
Dans les interrogatives indirectes introduites par ‘ce qui/ce que’, comme dans les relatives à tête allégée, il n’y a pas non plus d’élément wh- : ‘ce’ est un élément D, ‘qui/que’ est le complémenteur. Nous avons vu dans la première section du présent chapitre que ces interrogatives indirectes sont des questions déguisées en relatives et qu’elles ont des propriétés de DP. Or, un verbe interrogatif sélectionne un CP, pas un DP. Mais de nombreuses données dans diverses langues, impliquant des CPs nominalisés, montrent que finalement il n’existe pas de distinction catégorielle entre DP et CP. Seule la différence sémantique entre type référentiel d’une part et type propositionnel ou type question d’autre part reste pertinente. Les complétives en P+ce+que en français nous ont permis de découvrir que ‘ce’ en français n’est pas forcément référentiel : il peut être explétif. Nous postulons que ‘ce’ dans les interrogatives indirectes, tout comme dans les complétives, est également explétif et non pas référentiel. Le contraste net entre (73) et (74) met en évidence le caractère non référentiel de ‘ce’.

(73) Je me demande ce que tu as écrit.

- (74) a. ??Je me demande ce livre que tu as écrit.
 b. ??Je me demande le livre que tu as écrit.
 c. *Je me demande ce livre.

d. *Je me demande le livre.

Un verbe interrogatif tel que 'se demander' sélectionne un argument de type question et non pas référentiel. Le fait que 'se demander' soit satisfait avec la sélection de 'ce' mais pas avec la sélection de 'ce livre' ou 'le livre' montre bien que 'ce' n'est pas du même type sémantique que 'ce livre' ou 'le livre', qui eux, sont clairement référentiels. L'idée est donc qu'il existe en français un 'ce' dénué de référentialité lorsqu'il n'est pas combiné avec un N. La structure des interrogatives indirectes introduites par 'ce qui/ce que' serait donc identique à celle des relatives à tête allégée en français. La seule différence est la suivante : dans les relatives, 'ce' est référentiel, alors que dans les interrogatives indirectes, 'ce' est explétif. Ce 'ce' explétif est indiqué en gris dans la représentation en (75) pour montrer qu'il est dénué de référentialité. Je note le D des interrogatives indirectes D_{EXPL} pour le distinguer du D référentiel des relatives.



Dans les deux cas (relatives à tête allégée et interrogatives indirectes), il s'agit d'une analyse par appariement. Seulement, dans le cas des interrogatives indirectes, 'ce' étant explétif, il ne fournit pas son interprétation à l'opérateur : une relation de prédication avec un élément dénué de sens n'aboutit à rien sur le plan sémantique. Le résultat est que l'opérateur recevra sa valeur interrogative du verbe qui le sélectionne. 'Ce' est donc totalement invisible/ininterprétable sur le plan sémantique mais bel et bien visible sur le plan syntaxique. Mais la couche D est bel et bien présente dans la structure syntaxique, comme le montre le statut d'îlot fort des interrogatives indirectes introduites par 'ce qui/ce que' et leur comportement vis-à-vis du *matching*.

Il est intéressant de remarquer que 'ce' en français peut d'ailleurs être utilisé dans certaines structures impersonnelles, en alternance avec le pronom impersonnel 'il', comme

illustré dans les exemples (76) et (77).^{11 12}

- (76) a. Ce serait bien d'analyser leur distribution.
 b. Il serait bien d'analyser leur distribution.
- (77) a. C'est difficile pour lui d'admettre qu'il a tort.
 b. Il est difficile pour lui d'admettre qu'il a tort.

Pour conclure, dans les subordonnées impliquant 'ce que' en français, 'ce' n'est référentiel que dans les relatives. Dans les interrogatives indirectes, tout comme dans les complétives, 'ce' est explétif et il ne change donc pas le type d'objet sur le plan sémantique : le CP après 'ce' dans les complétives reste de type propositionnel et le CP après 'ce' dans les interrogatives indirectes reste une question.

Nous analysons les interrogatives indirectes introduites par 'ce qui/ce que' comme des relatives à tête allégée, c'est-à-dire comme des DPs sur le plan syntaxique. Ces relatives sont dérivées par appariement impliquant un opérateur nul, et non pas par montée de la tête. L'élément 'ce' étant explétif dans les interrogatives indirectes, la relation de prédication entre l'opérateur et 'ce' n'aboutit à rien. C'est le verbe interrogatif qui donne l'interprétation wh- à l'opérateur.

Notre analyse présente plusieurs avantages : d'une part, le fait que ces interrogatives indirectes soient analysées comme des relatives à tête allégée sur le plan syntaxique permet de rendre compte des propriétés atypiques de ces interrogatives indirectes (îlot fort, sensibilité aux effets de *matching*). D'autre part, elle permet d'expliquer l'interprétation de question, malgré la couche DP. Ces interrogatives déguisées en relatives ne sont interprétées comme des questions que de façon indirecte. De par la vacuité de 'ce', c'est le verbe interrogatif qui donne l'interprétation wh- à l'opérateur, d'où l'impossibilité d'avoir 'ce que' dans des questions racines.

11. Voir CORBLIN, 1994 sur 'ça' en français, qui peut également être employé dans certaines structures impersonnelles, dans un registre informel.

(1) ça pleut/ça neige

(2) Ça m'étonnerait qu'il dise la vérité.

12. Voir également CARLIER, 1996 sur l'interprétation générique du syntagme nominal disloqué au moyen de 'ce' ou 'ça' en français.

Conclusion Générale et Perspectives de Recherche

Récapitulatif

L'exploration des relatives à tête allégée en français mais aussi dans d'autres langues m'a menée aux conclusions suivantes :

- i. Les relatives à tête allégée se distinguent des relatives pleines et des relatives libres et elles constituent donc un type à part au sein des relatives.
- ii. Comme leur nom l'indique, la tête des relatives de ce type est légère, c'est-à-dire qu'elle correspond à un élément de type D (démonstratif ou article, selon les langues), et elle a un trait [-animé]. Quant à l'introducteur de la relative, il s'agit du complémenteur dans la plupart des langues romanes (excepté en roumain) mais d'un mot *wh-* dans des langues d'autres familles, telles que les langues germaniques.
- iii. Certaines langues, comme l'anglais, disposent des relatives libres mais pas des relatives à tête allégée (anglais). D'autres langues, comme le français, n'ont pas d'autre possibilité que d'avoir recours aux relatives à tête allégée, quand le référent est inanimé. Enfin, dans d'autres langues, comme le roumain et l'allemand, les deux types de relatives sont possibles.
- iv. Alors que les relatives pleines ont une structure DP+NP+CP, j'ai proposé que les relatives libres et les relatives à tête allégée avaient toutes deux une structure DP+CP. Seulement, dans les premières, la position D est vide alors qu'elle est remplie par la tête dans les deuxièmes. Et inversement, dans les premières la position [Spec,DP] est remplie par le mot *wh-* (CAPONIGRO, 2002) alors qu'elle est vide dans les deuxièmes.
- v. Nous avons postulé que les relatives à tête allégée en français sont dérivées par appariement, avec mouvement d'un opérateur nul en [Spec, CP].

J'ai ensuite décrit l'ambiguïté atypique qui existe en français entre les relatives à tête allégée et les interrogatives indirectes, 'ce qu-' pouvant introduire l'une ou l'autre des structures. Ces interrogatives indirectes, qui ont une interprétation de question *wh-*, sont atypiques sur le plan syntaxique du fait de la présence d'un élément de type D ('ce') et de l'absence de mot *wh-* ('qui/que' étant le complémenteur). J'ai exploré plusieurs hypothèses

pour tenter de faire la lumière sur leur structure syntaxique, parmi lesquelles l'hypothèse de l'incorporation (KONRAD, A paraître). L'idée est que 'ce que', à l'origine une séquence D+C, a subi un processus de grammaticalisation se manifestant par une incorporation de l'élément 'ce', de l'opérateur et de 'que', aboutissant à un nouveau mot wh- *səkə*, situé en [Spec, CP]. Cette incorporation donne lieu à deux configurations : une relative libre si la projection D (désormais vide) est retenue dans la structure sous-jacente ; une interrogative indirecte si cette projection n'est pas retenue dans la structure sous-jacente.

Dans le but d'éclaircir le statut syntaxique des interrogatives indirectes introduites par 'ce qu-' en français, j'ai mené une étude diachronique d'envergure (du 11ème au 17ème siècle) consistant à observer l'évolution et l'apparition de 'ce' dans les relatives et les interrogatives indirectes. Les conclusions suivantes se sont dégagées des résultats de cette étude :

- i. Il existait en ancien français des relatives libres introduites par 'qui/que' seul (avec référent inanimé), sans 'ce'.
- ii. 'Ce' apparaît en premier lieu dans les relatives et s'y impose très tôt (dès la fin du 13ème siècle). 'Ce qu-' a donc clairement un statut nominal.
- iii. Il faut attendre le milieu du 17ème siècle pour que l'emploi de 'ce qu-' s'étende aux interrogatives indirectes. Ce changement est soudain et est probablement la conséquence d'interventions normatives de la part de remarqueurs et grammairiens de l'époque tels que Vaugelas. Ce changement a probablement été facilité par l'affaiblissement de 'ce' d'une part et l'affaiblissement du mot wh- 'que' d'autre part.
- iv. 'Ce que' pouvait introduire des questions directes pendant une courte période (12ème siècle), ce qui pourrait être le signe d'une esquisse d'incorporation de 'ce' et 'que', mais cela ne s'est pas maintenu.

Ainsi, cette étude ne vient pas confirmer l'hypothèse d'une grammaticalisation de 'ce que' en un pronom interrogatif.

Toujours dans le but de faire la lumière sur cette ambiguïté syntaxique atypique, j'ai étudié l'acquisition des relatives et interrogatives indirectes introduites par 'ce qu-' par des enfants français âgés entre 3 et 6 ans, en la comparant à celle des mêmes structures introduites par un mot wh- sans élément D, à savoir 'qui'. J'ai pu tirer les conclusions suivantes à partir des résultats de deux tests de répétition :

- i. Les enfants français de 3 à 6 ans n'analysent pas 'ce qu-' comme un élément wh- mais bien comme comprenant un élément de nature nominale (performance plus élevée pour les relatives que pour les interrogatives indirectes, tendance à omettre 'ce' et à employer *kəsk* dans les interrogatives indirectes).
- ii. Les données d'acquisition de 'ce qu-' reflètent l'évolution diachronique de 'ce qu-'.
- iii. Les enfants français ont beaucoup moins de difficultés avec des subordonnées introduites par 'qui' qu'avec des subordonnées introduites par 'ce qu-'.

- iv. L'avantage pour les relatives dans le cas de 'ce qu-' est dû à la présence de 'ce' étant donné que nous n'avons pas retrouvé cet avantage dans le cas des subordonnées introduites par le mot wh- 'qui'.

Au vu des propriétés syntaxiques de 'o que' en portugais, qui sont distinctes du français 'ce qu-' et qui vont clairement dans la direction d'une grammaticalisation (dans les interrogatives indirectes, 'o que' est compatible avec l'infinitif et il peut entrer dans une relation de coordination avec un autre mot wh- ; il peut également introduire une question directe), j'ai comparé l'acquisition de 'ce qu-' avec celle de 'o que'. J'ai donc mis en place un test de répétition en portugais, similaire à celui du français. Comme pour le français, j'ai comparé 'o que' avec un mot wh- sans élément D, 'quem'. Les principales conclusions que j'ai tirées à partir des résultats des tests de répétition sont les suivantes :

- i. Les enfants portugais âgés entre 3 et 6 ans ont bien moins de difficultés avec 'o que' que les enfants français n'en ont avec 'ce qu-' et ce, dès le plus jeune âge.
- ii. Comme les enfants français, ils sont sensibles à la présence de l'élément D. Mais sa présence dans les interrogatives indirectes est moins problématique pour les enfants portugais que pour les enfants français.
- iii. L'avantage pour les relatives dans le cas de 'o que' est dû à la présence de 'o', car nous n'avons pas retrouvé cet avantage dans le cas des subordonnées introduites par le mot wh- 'quem'.

J'ai également mis en évidence le fait que cet avantage pour 'qui/quem' comparé à 'ce qu-/o que' est valable même pour les dépendances sujet, ce qui montre que les enfants sont sensibles à la nature de l'élément qui introduit une relative ou une interrogative indirecte. Afin de rendre compte du *pattern* d'acquisition des différents types de relatives, j'ai proposé l'Hypothèse de la Hiérarchie du Poids de la Tête. L'idée est que plus la tête d'une relative est légère d'un point de vue syntaxique, plus la relative est facile à produire et à comprendre pour l'enfant. Cela implique la hiérarchie suivante : les relatives pleines sont plus difficiles pour l'enfant que les relatives à tête allégée, qui sont elles-mêmes plus complexes à traiter que les relatives libres.

Afin de déterminer s'il existait des différences de traitement entre les différentes structures impliquant 'ce que', je me suis focalisée sur deux ambiguïtés syntaxiques : celle entre complétives et relatives à tête allégée d'une part et celle entre relatives à tête allégée et interrogatives indirectes d'autre part. Deux premières expériences (test de jugement d'acceptabilité et première expérience de lecture par présentation auto-segmentée) ont montré qu'il existait au moins deux 'ce que' distincts : un 'ce que' sans lacune (complétive) et un 'ce que' avec lacune (relative), et que le premier était préféré parce que l'analyseur évite autant que possible de postuler une lacune. Ainsi, dans les cas de compétition entre une structure avec lacune et une structure sans lacune, cette dernière est préférée, en vertu du Principe de Chaîne Minimale. Quant à la deuxième ambiguïté syntaxique, aucun principe de traitement ne pouvait prédire une préférence pour l'une ou l'autre de ces structures car les deux

comportent une lacune, et la structure est désambiguïlée dès son introduction par le verbe. Les résultats d'une deuxième expérience de lecture par présentation auto-segmentée ont montré que les adultes, contrairement aux enfants, ne sont pas dérangés par un élément D dans les interrogatives indirectes. Aucun avantage pour les relatives n'a été relevé. Nous avons postulé que cette absence de différence entre relatives à tête allégée et interrogatives indirectes indiquait qu'il s'agissait d'une seule et même structure syntaxique.

J'en suis arrivée à étudier à nouveau de plus près les propriétés syntaxiques des interrogatives indirectes introduites par 'ce qu-' en synchronie. Sur la base de leur comportement syntaxique inhabituel pour des questions les rapprochant de DPs complexes, donc de relatives (îlot fort, sensibilité aux effets de *matching*), j'ai proposé qu'il s'agissait en réalité de questions déguisées en relatives. Deux problèmes se sont naturellement posés pour notre analyse : le problème de la sélection d'une part et le problème de l'interprétation d'autre part. Le premier problème est que, en théorie, un verbe interrogatif sélectionne un CP et non pas un DP. Or, nous avons prouvé, en nous appuyant sur plusieurs données dans plusieurs langues impliquant la nominalisation de CPs, que la distinction syntaxique entre DP et CP n'est pas si catégorique. Le problème de la sélection est donc un faux problème. Le deuxième problème est qu'une relative avec 'ce' comme tête puisse être interprétée comme une question. Cela n'est pas un problème nouveau car les sémanticiens y ont été confrontés avec les *concealed questions*. Mais nous avons vu que l'analyse par ellipse, bien qu'attrayante, n'était pas plausible dans le cas des interrogatives indirectes introduites par 'ce qu-'.

J'ai proposé que 'ce' dans les interrogatives indirectes, tout comme dans les complétives, était explétif, dénué de référentialité. Seul le 'ce' des relatives est référentiel. J'analyse les relatives à tête allégée et les interrogatives indirectes en ce que' comme des relatives dérivées par appariement impliquant un opérateur nul. Dans les deux cas, cet opérateur entre dans une relation de prédication avec la tête 'ce'. Sauf que cette relation ne se concrétise que dans les relatives où 'ce' est référentiel. Dans le cas des interrogatives, 'ce' est comme invisible sur le plan sémantique et le verbe interrogatif donne l'interprétation wh- à l'opérateur. Cette analyse permet d'expliquer à la fois pourquoi ces interrogatives ont des propriétés de DPs sur le plan syntaxique (car la couche D est bel et bien présente dans la structure) et pourquoi ces DPs sont malgré tout interprétés comme des questions ('ce' est explétif, donc dénué de référentialité, donc ininterprétable sur le plan sémantique).

Précisons que le fait d'analyser des interrogatives indirectes comme des relatives sur le plan syntaxique n'est pas si surprenant si l'on pense aux propriétés communes que ces structures partagent (dépendances A-barre contenant une lacune pouvant être interprétée comme une variable) et à l'interdépendance qui existe entre ces deux types de structures. Ce n'est pas anodin qu'une structure relative se soit étendue à des interrogatives indirectes. D'ailleurs, après un stade de différenciation (emploi de 'qu'est-ce qu-' dans les interrogatives indirectes et emploi de 'ce qu-' dans les relatives), il se pourrait bien que l'on retourne vers deux systèmes interdépendants, où les interrogatives indirectes auraient cette fois-ci une

influence sur les relatives. En effet, nous avons trouvé quelques emplois de ‘qu’est-ce qu-’ dans des relatives dans FrWaC et il est probable que cette extension prenne de l’ampleur en français parlé.

Ma proposition qu’il y a en français un ‘ce’ explétif dans les complétives et dans les interrogatives indirectes crée un lien entre ces deux structures. Il est intéressant de remarquer que les deux types de structures semblent appartenir à un registre normé, à un français standard, sûrement du fait de la présence de cet explétif. Le test de jugement d’acceptabilité sur l’ambiguïté entre complétives et relatives a en effet montré que les jugements pour les complétives en P+ce+que n’étaient pas très élevés (en particulier pour les moins de 35 ans). En ce qui concerne les interrogatives indirectes en ‘ce qu-’, nous avons découvert que les enfants étaient assez réticents envers cet élément D. Les adultes eux aussi ont tendance à remplacer ‘ce qu-’ par ‘qu’est-ce qu-’ dans les interrogatives indirectes, comme l’a montré notre étude à partir du corpus FrWaC. Cet usage explétif de ‘ce’ s’est d’ailleurs développé très tard dans les interrogatives indirectes, comme le signale le changement soudain observé dans ces structures vers le milieu du 17^{ème} siècle. D’ailleurs, ce changement correspond probablement à une intervention de nature normative.

Pour aller plus loin : perspectives de recherche...

Certains points méritent à mon avis d’être étudiés dans la continuité de ma thèse sur le plan syntaxique, sur le plan diachronique, sur le plan de l’acquisition du langage, ainsi que sur le plan du traitement par les adultes.

...dans la continuité de la thèse

Sur le plan syntaxique, je souhaiterais étudier les propriétés des interrogatives indirectes introduites par ‘el que’ en catalan. En effet, comme dans le cas du français et du portugais, un élément D est présent et il semblerait que cela ne pose pas problème dans les interrogatives indirectes. S’agit-il d’un cas de grammaticalisation comme en portugais ou bien d’un cas de D explétif comme en français ou bien encore d’un troisième type d’analyse différent des deux premiers ?

Sur le plan diachronique, il serait intéressant de se pencher sur l’évolution de ‘ce’ dans les complétives et de dater son apparition. Il est clair que les relatives et les interrogatives indirectes entretiennent une relation d’interdépendance et que les complétives n’entretiennent pas de relation de ce type ni avec les relatives ni avec les interrogatives indirectes. Donc leur développement est sûrement indépendant du développement des deux autres structures. D’après PIERRARD, 1995, ‘ce que’ est obligatoire avec les prépositions ‘à’, ‘de’, ‘par’ et ‘en’ dès les stades les plus anciens de la langue. En ancien français, l’emploi de ‘ce que’ à la place de ‘que’ était très fréquent et varié, aussi bien en complétive sujet, complément

d'objet direct ou complément d'objet indirect, mais aussi en complétive adjoint (voir respectivement (78a), (78b), (78c) et (78d) données par PIERRARD, 1995 : 115).

- (78) a. Mes ce que il se desguisa en semblance de nouvel chevalier m'en toli la droite connoissance. [Mort Artu 30,38; cité par Moignet 1973 : 338]
 b. Ge li dis ce qu'il s'en alast. [Bérroul 435; *ibid.*]
 c. Se sont acordé a ce qu'il s'en alast clamer
 La ou porra le roi trover. [Ren. 8243; cité par Tobler & Lommatzsch, II, 1971 :82]
 d. Après ce que mestres Gautiers Map ot mis en escrit... [La mort le roi Artu 1,1; cité par Léard 1990 : 290]

En Moyen Français, 'ce' s'insère dans la plupart des locutions conjonctives (voir (79), PIERRARD, 1995 : 116) et apparaît même en comparative au 15ème siècle (voir (80), *ibid.*).

- (79) ...; car cil ne le vouloit pas faire pendre ne ardoir, pour ce qu'il eüst plus lonc torment. [Berinus 425; cité par Wunderli 1978 : 237]
 (80) Il n'estoit rien au monde dont le Roy eust plus grand paour que ce qu'i uy eschappast quelque mot. [Commynes 1, 322, M.; cité par Léard 1990 : 290]

'Ce' a connu par la suite un déclin aux 15ème et 16ème siècle et il se maintiendra uniquement après les prépositions 'par', 'en', 'de' et 'à'.

Ces faits semblent *a priori* problématiques pour notre analyse : il n'y aurait apparemment pas de concordance temporelle en ce qui concerne l'apparition de 'ce' dans les complétives et celle de 'ce' dans les interrogatives indirectes. Cela signifie que ce n'est pas le 'ce' explétif en soi mais le 'ce' explétif dans les interrogatives indirectes en particulier qui a mis du temps à se développer. On peut supposer que le problème ne s'est posé dans les interrogatives indirectes que lorsque le 'que' interrogatif s'est affaibli et est devenu clitique. Il faudrait donc vérifier s'il existe une concordance temporelle entre l'affaiblissement du 'que' en français et l'apparition de 'ce' dans les interrogatives indirectes.

Sur le plan de l'acquisition du langage, il serait à mon avis très intéressant de mettre en place un test de répétition similaire à ceux que j'ai mis en place en français (sur 'ce qu-') et en portugais (sur 'o que') en anglais, en me focalisant sur l'ambiguïté avec 'what', pouvant introduire soit des relatives libres soit des interrogatives indirectes. Si l'avantage des relatives est bel et bien dû à la présence d'un élément D ('ce' en français et 'o' en portugais), alors on s'attend à ne pas trouver de différence de performance entre les relatives libres et les interrogatives indirectes en anglais, du fait de l'absence d'élément D.

Sur le plan du traitement par les adultes, je souhaiterais vérifier expérimentalement l'acceptabilité de l'extraction d'arguments et d'adjoints dans les deux constructions impliquant un 'ce' explétif selon mon analyse, à savoir les complétives en P+ce+que et les interrogatives

indirectes introduites par ‘ce qu-’. En effet, cela permettrait de vérifier qu’il y a bien une couche D sur le plan syntaxique dans ces structures. Nous savons d’une part que les complétives typiques (sans élément D) n’agissent pas comme des îlots pour l’extraction et d’autre part que les interrogatives indirectes typiques introduites par un élément wh- (sans élément D) sont des îlots faibles, ce qui signifie que l’extraction d’un argument est davantage acceptable que l’extraction d’un adjectif qui elle est agrammaticale. Nous savons également que les DP complexes, tels que les relatives, sont des îlots forts, ce qui signifie que l’on ne peut extraire ni un adjectif ni un argument sans que la phrase soit agrammaticale. Idéalement, il faudrait créer deux tests de jugement d’acceptabilité différents :

1. Le premier permettrait de comparer les complétives sans ‘ce’ aux complétives avec ‘ce’. Les notes devraient être élevées pour les complétives sans ce avec extraction d’arguments ou d’adjectifs. Par contre, pour les complétives avec ‘ce’, si ‘ce’ ajoute effectivement une couche DP à la structure, alors les notes devraient être très basses dans les cas d’extraction d’adjectifs et dans les cas d’extraction d’arguments, c’est-à-dire qu’elles devraient se comporter comme un îlot fort.
2. Le deuxième permettrait de comparer les interrogatives indirectes introduites par ‘ce qu-’ aux relatives à tête allégée d’une part et aux interrogatives indirectes introduites par ‘où’ d’autre part. Dans le cas de ces dernières, les notes devraient être basses lorsque l’on extrait un adjectif et plus élevées lorsque l’on extrait un argument, c’est-à-dire qu’elles devraient se comporter comme un îlot faible. Par contre, on ne devrait pas trouver de différence entre les interrogatives indirectes introduites par ‘ce qu-’ et les relatives à tête allégée. Les deux devraient se comporter comme des îlots forts : l’extraction d’arguments devrait être aussi mauvaise que l’extraction d’adjectifs. Ces prédictions semblent confirmées par les jugements introspectifs que j’ai recueillis auprès de collègues francophones mais il faudrait vérifier l’asymétrie en question dans de réelles conditions expérimentales (avec plusieurs items pour chaque condition et des distracteurs et beaucoup plus de participants).

...au-delà de la thèse

Au-delà de la thèse maintenant, j’aimerais étudier à l’avenir d’autres cas de CPs nominalisés. Le cas des pseudo-relatives me semble de ce point de vue particulièrement intéressant.

Les pseudo-relatives sont des constructions finies que l’on trouve dans plusieurs langues et qui ressemblent, en surface, à des noms modifiés par des relatives mais qui décrivent en réalité des événements (GRILLO et MOULTON, 2016).

(81) J’ai vu Léa qui courrait.

Une des analyses classiques est de percevoir ces structures comme des petites propositions, c’est-à-dire comme des CPs (RADFORD, 1975, RADFORD, 1977, GUASTI, 1988, GUASTI, 1992, GUASTI, 1993, CAMPOS, 1994).

Dans leur article de 2015, Moulton et Grillo (qui se basent essentiellement sur l'italien) postulent que les pseudo-relatives sont des DPs et non pas des CPs (GRILLO et MOULTON, 2015). Plus particulièrement, il y a d'après eux une couche DP dans la structure sous-jacente avec un D nul au-dessus du CP, ce qui fait donc penser à une forme de CP nominalisé.

Sur le plan syntaxique, leurs arguments sont les suivants : ils montrent que les pseudo-relatives en italien peuvent être complément de prépositions (voir (82), tiré de GRILLO et MOULTON, 2015 : 3), comme les DPs et contrairement aux CPs finis standard et aux infinitives.

- (82) La vista di [Carlo che balla il tango] è da non perdere.
 la vue de Carlo COMP danser-PRS.3SG le tango est à NEG manquer
 'La vue de Carlo qui danse le tango est à ne pas manquer.'

De plus, on peut coordonner des pseudo-relatives avec d'autres DPs et cela déclenche un accord pluriel (voir (83), tiré de GRILLO et MOULTON, 2015 : 3).

- (83) [Lea che piange] e [l'evento che temi] sono
 Lea COMP pleurer-PRS.3SG et l'événement COMP craindre-PRS.2SG sont
 immagini orribili.
 images horribles
 'Lea qui pleure et l'événement que tu crains sont des images horribles.'

Il est également intéressant de noter que les pseudo-relatives sont des îlots forts, bloquant ainsi l'extraction A-barre des adjoints (84) mais aussi des arguments. Cela est un argument supplémentaire venant appuyer l'analyse de DP de Grillo & Moulton (GRILLO et MOULTON, sous révision : 34).

- (84) *Come_i hai visto [Maria che picchiava il ragazzo t_i] ?
 comment as vu [Maria COMP battre-IPF.3SG le garçon t
 '*Comment as-tu vu Maria qui battait le garçon ?'
- (85) *La persona con cui_i ho visto [tua sorella che stava parlando t_i] è
 la personne avec qui ai vu [ta soeur COMP était parler-GRDV] t est
 questa.
 celle-ci
 '*C'est la personne avec qui j'ai vu ta sœur qui était en train de parler.'

Quant à la sémantique de ces structures, ils mettent également en évidence le fait que les pseudo-relatives qui sont compléments de verbes de perception sont référentielles. Elles dénotent des événements et non pas des individus. Il s'agit donc de DPs qui décrivent des situations. Le D nul postulé correspondrait alors à un déterminant défini et convertirait la structure en une description d'une situation individuelle (voir Figure 7.1). Ce D serait donc responsable du caractère référentiel de la pseudo-relative.

engagement d'existence sous la négation. Cet engagement d'existence projette hors de la négation. C'est pour cela que l'exemple (89) paraît bizarre [GRILLO et MOULTON, 2016 : 3].

- (89) Dato che Lea non ha mai ballato...# Max non ha visto [Lea che
 donné que Lea NEG a jamais dansé...# Max NEG a vu Lea COMP
 ballava].
 danser-IPF.3SG
 'Étant donné que Lea n'a jamais dansé, # Max n'a jamais vu Léa qui dansait.'

Leur analyse est intéressante pour notre discussion car c'est un exemple de structures qui devraient être des CPs mais qui se comportent comme des DPs sur les plans syntaxique et sémantique. Le cas des interrogatives indirectes introduites par 'ce que' est toutefois un peu différent : il s'agit certes de DPs sur le plan syntaxique, mais, sur le plan sémantique, elles ne sont pas référentielles. Elles sont interprétées comme des questions, c'est-à-dire qu'elles dénotent un ensemble de propositions. De plus, le D des interrogatives indirectes en français est explicite et explétif alors que le D des pseudo-relatives postulé par Grillo & Moulton est nul et référentiel. J'envisage donc d'étudier les propriétés des pseudo-relatives en français et de déterminer si elles se comportent comme les pseudo-relatives en italien.

Bibliographie Générale

- ABEILLÉ, Anne (2008). “Tout ça, quoi. Remarques sur les mots “légers” en français”. In : *Modèles syntaxiques. La syntaxe à l’aube du XXIème siècle*. D. van Raemdonck (ed). Peter Lang, Bruxelles, p. 85–111.
- ABEILLÉ, Anne et Danièle GODARD (2004). “De la légèreté en syntaxe”. In : *Bulletin de la Société de Linguistique de Paris* 99.1, p. 69–106.
- ABNEY, Steven P (1987). “The English noun phrase in its sentential aspect”. Thèse de doct. Massachusetts Institute of Technology.
- ADANI, Flavia (2011). “Rethinking the acquisition of relative clauses in Italian : towards a grammatically based account”. In : *Journal of Child Language* 38.1, p. 141–165.
- ADANI, Flavia, Marie SEHM et Andrea ZUKOWSKI (2012). “How do German children and adults deal with their relatives”. In : *Advances in language acquisition*, p. 14–22.
- ADLI, Aria (2006). “French wh-in-situ questions and syntactic optionality : Evidence from three data types”. In : *Zeitschrift für Sprachwissenschaft* 25.2, p. 163–203.
- AKAL, Taylan (2017). “Wh-Argument/Adjunct Asymmetry in Sentence Processing”. In : *Dilbilim Araştırmaları Dergisi* 28.2, p. 43–71.
- ALEXIADOU, Artemis et al. (2000). *The syntax of relative clauses*. John Benjamins Publishing.
- AOUN, Joseph, Norbert HORNSTEIN et Dominique SPORTICHE (1981). “Some aspects of wide scope quantification”. In : *Journal of Linguistic Research* 1.3, p. 69–95.
- AUTHIER, Jean-Marc et Lisa REED (2010). “Clausal complementation and the status of French à/de ce que”. In : *Lingua* 120.9, p. 2193–2210.
- BAAYEN, R Harald, Douglas J DAVIDSON et Douglas M BATES (2008). “Mixed-effects modeling with crossed random effects for subjects and items”. In : *Journal of memory and language* 59.4, p. 390–412.
- BAIÃO, Vera et Maria LOBO (2014). “Aquisição de interrogativas preposicionadas no português europeu”. In : *Textos Seleccionados do XXIX Encontro Nacional da APL*, p. 57–70.
- BAKER, Carl Leroy (1968). “Indirect questions in English”. Thèse de doct. University of Illinois.
- BAKER, Mark (1988). *Incorporation : A theory of grammatical function changing*. University of Chicago Press.
- BÁRÁNY, András et Michelle SHEEHAN (2019). “When dependent case is not enough”. In : *Generative Linguistics in the Old World* 42 (GLOW). Communication à l’Université d’Oslo.
- BARON, Naomi S (1977). *Language acquisition and historical change*. T. 36. North Holland Publishing Company.

- BARONI, Marco et al. (2009). “The WaCky wide web : a collection of very large linguistically processed web-crawled corpora”. In : *Language resources and evaluation* 43.3, p. 209–226.
- BATES, Douglas et al. (2015). “Fitting Linear Mixed-Effects Models Using lme4”. In : *Journal of Statistical Software* 67.1, p. 1–48.
- BATES, Douglas et al. (2015). “Parsimonious mixed models”. In : *preprint arXiv :1506.04967v1 [stat.ME]*.
- BATTYE, Adrian (1989). “Free relatives, pseudo-free relatives and the syntax of CP in Italian”. In : *Rivista di linguistica* 1.2, p. 219–246.
- BAUNAZ, Lena et Eric LANDER (2018). “Syncretisms with the nominal complementizer”. In : *Studia Linguistica* 72.3, p. 537–570.
- BEAUZÉE, Nicolas (1819). *Grammaire générale, ou exposition raisonnée des éléments nécessaires du langage*. A. Delalain.
- BENȚEA, Anamaria et Stephanie DURRLEMAN (2013). “A’-dependencies in French”. In : *Romance Languages and Linguistic Theory 2011 : Selected papers from ‘Going Romance’ Utrecht 2011* 5, p. 1.
- BENȚEA, Anamaria, Stephanie DURRLEMAN et Luigi RIZZI (2016). “Refining intervention : the acquisition of featural relations in object A-bar dependencies”. In : *Lingua* 169, p. 21–41.
- BHATT, Rajesh (2002). “The raising analysis of relative clauses : Evidence from adjectival modification”. In : *Natural language semantics* 10.1, p. 43–90.
- BIANCHI, Valentina (1999). *Consequences of Antisymmetry : Headed Relative Clauses*. Berlin : Mouton de Gruyter.
- (2000). “Some issues in the syntax of relative determiners”. In : *The syntax of relative clauses*. ed. par Artemis Alexiadou, Paul Law, André Meinunger, et Chris Wilder. Amsterdam : John Benjamins., p. 53–81.
- BORSLEY, Robert D (1997). “Relative clauses and the theory of phrase structure”. In : *Linguistic Inquiry* 28, p. 629–647.
- BOUCHARD, Denis et Paul HIRSCHBÜHLER (1986). “French *quoi* and its clitic allomorph *que*”. In : *Studies in Romance languages*. ed. par Carol Neidle et Rafael Nuñez Cedeño. Dordrecht :Foris.
- BRESNAN, Joan et Jane GRIMSHAW (1978). “The syntax of free relatives in English”. In : *Linguistic inquiry* 9.3, p. 331–391.
- CAMPOS, Ana Rita Coimbra Pina de (2015a). “A Extensão Média de Enunciado como indicador de desenvolvimento linguístico no português europeu”. Mém.de mast. Escola Superior de Saúde - Instituto Politécnico de Setúbal/ FCSH- Universidade Nova de Lisboa.
- CAMPOS, Héctor (1994). “Seudo-elevación y seudo-relativas en español”. In : *Gramática del español. México : El Colegio de México*, p. 201–236.
- CAMPOS, Raquel João Pinheiro Marques de (2015b). “Rácio Type-Token e D como indicadores de desenvolvimento linguístico no português europeu”. Mém.de mast. Escola Superior de Saúde - Instituto Politécnico de Setúbal/ FCSH- Universidade Nova de Lisboa.

- CAPONIGRO, Ivano (2002). “Free relatives as DPs with a silent D and a CP complement”. In : *Proceedings of the Western Conference on Linguistics (WECOL)*. Sous la dir. de Vida SAMILAN. Fresno : California State University.
- (2019). “In defense of what (ever) free relative clauses they dismiss : A reply to Donati and Cecchetto (2011)”. In : *Linguistic Inquiry* 50.2, p. 356–371.
- CARLIER, Anne (1996). “‘Les gosses ça se lève tôt le matin’ : l’interprétation générique du syntagme nominal disloqué au moyen de ce ou ça”. In : *Journal of French Language Studies* 6.2, p. 133–162.
- CECCHETTO, Carlo et Caterina DONATI (2010). “On labeling : Principle C and head movement”. In : *Syntax* 13.3, p. 241–278.
- (2015). “Subject Intervention in Free Relatives”. In : *Structures, Strategies and Beyond : Studies in honour of Adriana Belletti* 223, p. 369.
- CEREJEIRA, Joana (2010). “Assimetrias na aquisição de interrogativas de sujeito e de objecto : dados de produção”. In : *Textos seleccionados do XXV Encontro Nacional da Associação Portuguesa de Linguística*, p. 291–306.
- CHOMSKY, Noam (1961). “Some methodological remarks on generative grammar”. In : *Word* 17.2, p. 219–239.
- (1965). *Aspects of the theory of syntax*. Cambridge, MA : MIT Press.
- (1968). *Remarks on nominalization*. Linguistics Club, Indiana University.
- (1973). “Conditions on transformations”. In : *A festschrift for Morris Halle*.
- (1981). *Lectures on government and binding*. Dordrecht : Foris.
- (1986a). *Barriers*. Cambridge, MA : MIT press.
- (1986b). *Knowledge of language : Its nature, origin, and use*. Greenwood Publishing Group.
- (1988). “Generative grammar”. In : *Studies in English Linguistics and Literature*.
- (1993). *Lectures on government and binding : The Pisa lectures*. 9. Walter de Gruyter.
- (1995). *The Minimalist Program*. Cambridge, MA : MIT Press.
- (2000). “Minimalist inquiries : The framework (mitopl 15)”. In : *Step by step : Essays on minimalist syntax in honor of Howard Lasnik*, p. 89–155.
- (2002). *Syntactic structures*. Walter de Gruyter.
- (2008). “On phases”. In : *Current Studies in Linguistics Series* 45, p. 133.
- (2011). *Current issues in linguistic theory*. Walter de Gruyter.
- (2013). *Topics in the theory of generative grammar*. Walter de Gruyter.
- (2014). *The minimalist program*. MIT press.
- CHOMSKY, Noam et Howard LASNIK (1993). “The Theory of Principles and Parameters”. In : *Syntax : An International Handbook of Contemporary Research*. Sous la dir. de Joachim JACOBS et al. T. I. Berlin/New York : Walter de Gruyter, p. 506–569.
- CHRISTENSEN, Rune Haubo Bojesen (2018). *ordinal-Regression models for ordinal data. R package version 2018.4-19*.
- CINQUE, Guglielmo (1990). *Types of \bar{A} -dependencies*. Cambridge, MA : MIT press.
- (2006). *Restructuring and Functional Heads*. Oxford Studies in Comparative Syntax : The Cartography of Syntactic Structures. Oxford University Press.

- CINQUE, Guglielmo et Luigi RIZZI (2008). “The cartography of syntactic structures”. In : *Studies in linguistics* 2, p. 42–58.
- CITKO, Barbara (1999). “Light-Headed Relatives”. In : *University of Pennsylvania Working Papers in Linguistics* 6.1, p. 8.
- (2004). “On headed, headless, and light-headed relatives”. In : *Natural Language & Linguistic Theory* 22.1, p. 95–126.
- CLIFTON JR, Charles, Shari SPEER et Steven P ABNEY (1991). “Parsing arguments : Phrase structure and argument structure as determinants of initial parsing decisions”. In : *Journal of Memory and Language* 30.2, p. 251–271.
- CLIFTON, Charles et Lyn FRAZIER (1989). “Comprehending sentences with long-distance dependencies”. In : *Linguistic structure in language processing*. Springer, p. 273–317.
- COMBETTES, Bernard (1995). “Le statut de la séquence CELUI+ expansion : aspects syntaxiques et textuels”. In : *Pratiques* 87.1, p. 93–106.
- CORBLIN, Francis (1987). ““ Ceci” et” cela” comme formes à contenu indistinct”. In : *Langue française* 75, p. 75–93.
- (1994). “Existe-t-il un «ça» impersonnel en français?” In : *L’information grammaticale* 62.1, p. 45–47.
- (1995). *Les formes de reprise dans le discours. Anaphores et chaînes de référence*. Presses universitaires de Rennes.
- CORBLIN, Francis, Jean-Marie MARANDIN et Petra SLEEMAN (2004). “Nounless determiners”. In : *Handbook of French semantics*, p. 23–40.
- COSTA, João, Nino GRILLO et Maria LOBO (2012). “Minimality beyond lexical restrictions : processing and acquisition of free wh-dependencies in European Portuguese”. In : *Revue roumaine de linguistique* 57.2, p. 143–160.
- COSTA, João, Maria LOBO et Carolina SILVA (2011). “Subject–object asymmetries in the acquisition of Portuguese relative clauses : Adults vs. children”. In : *Lingua* 121.6, p. 1083–1100.
- COSTA, João et al. (2014). “The boy that the chef cooked : Acquisition of PP relatives in European Portuguese and Hebrew”. In : *Lingua* 150. Elsevier, p. 386–409.
- COVENEY, Aidan (1995). “The use of the QU-final interrogative structure in spoken French”. In : *Journal of French language studies* 5.2, p. 143–171.
- (1997). “L’approche variationniste et la description de la grammaire du français : le cas des interrogatives”. In : *Langue française*, p. 88–100.
- DAVIES, William D et Stanley DUBINSKY (2001). “Functional architecture and the distribution of subject properties”. In : *Objects and other subjects*. Springer, p. 247–279.
- DE VINCENZI, Marica (1991). *Syntactic parsing strategies in Italian*. Dordrecht, The Netherlands : Kluwer.
- DELAVEAU, Annie, Paul CAPPEAU et Anne DAGNAC (A paraître). “Les phrases interrogatives”. In : *La grande grammaire du français*. Sous la dir. d’Anne ABEILLÉ et Danièle GODARD. Actes Sud.
- DEMONTE, Violeta (1977). *La subordinación sustantiva*. Madrid : Cátedra.
- DINDELEGAN, Gabriela Pană (2016). *The Syntax of Old Romanian*. Oxford University Press.

- DOBROVIE-SORIN, Carmen et Ion GIURGEA (2013). *A reference grammar of Romanian : Volume 1 : The noun phrase*. John Benjamins Publishing.
- DONATI, Caterina et Carlo CECCHETTO (2011). “Relabeling heads : A unified account for relativization structures”. In : *Linguistic inquiry* 42.4, p. 519–560.
- DRUMMOND, Alex (2013). “Ibex farm”. In : *Online server : <http://spellout.net/ibexfarm>*.
- ERIKSSON, Olof (1982). “Il m’a dit ce qu’il pense : interrogative ou relative ?” In : *Revue romane* 2.XVII, p. 3–20.
- FLYNN, Suzanne et Barbara LUST (1980). “Acquisition of relative clauses : Developmental changes in their heads”. In : *Cornell working papers in linguistics* 1, p. 33–45.
- FODOR, Janet Dean (1978). “Parsing strategies and constraints on transformations”. In : *Linguistic Inquiry* 9.3, p. 427–473.
- FOPPOLO, Francesca, Ingrid KONRAD, Massimo BURATTIN et al. (2018). “Avoiding gaps in Romance”. In : *Architectures and Mechanisms of Language Processing (AMLAP)*. Présentation poster à Berlin, en septembre 2018.
- FOPPOLO, Francesca et al. (2019a). “Avoiding GAPS in Romance”. In : *45ème édition de Incontro di Grammatica Generativa (IGG)*. Communication à Padoue, en février 2019.
- (2019b). “Strategies of filler-gap avoidance in Romance”. In : *Conference on Human Sentence Processing (CUNY)*. Présentation poster à l’Université de Colorado Boulder, en mars 2019.
- FOURNIER, Nathalie (1998). *Grammaire du français classique*. Belin, Paris.
- FRAGMAN, Cathy (2000). “Grammaire et pouvoir global de computation : deux sources de difficultés possibles dans la production des relatives chez les enfants francophones”. In : *Canadian Journal of Linguistics/Revue canadienne de linguistique* 45.1-2, p. 7–48.
- FRANA, Ilaria (2006). “The de re analysis of concealed questions : a unified approach to definite and indefinite concealed questions”. In : *Semantics and Linguistic Theory*. 16, p. 17–34.
- (2017). *Concealed questions*. Oxford University Press.
- FRAZIER, Lyn et Charles CLIFTON (2002). “Processing “d-linked” phrases”. In : *Journal of Psycholinguistic Research* 31.6, p. 633–659.
- FRIEDEMANN, Marc-Ariel (1990). “Le pronom interrogatif que et la montée du verbe en C°”. In : *Rivista di Grammatica Generativa* 15, p. 123–139.
- FRIEDMANN, Naama, Adriana BELLETTI et Luigi RIZZI (2009). “Relativized relatives : Types of intervention in the acquisition of A-bar dependencies”. In : *Lingua* 119.1, p. 67–88.
- FRIEDMANN, Naama et Rama NOVOGRODSKY (2004). “The acquisition of relative clause comprehension in Hebrew : A study of SLI and normal development”. In : *Journal of Child language* 31.3, p. 661–681.
- GANDON, Ophélie (2016). “La relativisation dans une perspective aréale : l’aire Caucase-Anatolie de l’Est-Iran de l’Ouest”. Thèse de doct. Université Sorbonne Nouvelle - Paris 3.
- GINZBURG, Jonathan et Ivan SAG (2000). *Interrogative investigations*. Stanford : CSLI publications.

- GODARD, Danièle (A paraître). “Les relatives sans antécédent”. In : *La grande grammaire du français*. Sous la dir. d’Anne ABEILLÉ et Danièle GODARD. Actes Sud.
- (1988). *La syntaxe des relatives en français*. Editions du Centre national de la recherche scientifique.
- GODARD, Danièle et Anne ABEILLÉ (2011). “Les interrogatives compléments en français”. In : *Cahiers de lexicologie : Revue internationale de lexicologie et lexicographie* 98.1, p. 161–176.
- GOLDSMITH, John (1978). “Que, c’est quoi? que, c’est QUOI”. In : *Recherches linguistiques à Montreal* 11, p. 1–13.
- (1981). “Complementizers and root sentences”. In : *Linguistic Inquiry* 12.4, p. 541–574.
- GOODLUCK, Helen et Susan TAVAKOLIAN (1982). “Competence and processing in children’s grammar of relative clauses”. In : *Cognition* 11.1, p. 1–27.
- GOUGENHEIM, Georges (1965). *Celui et ce aux points de vue syntaxique et fonctionnel*. C. Klincksieck.
- GRIGORE, Georges (2007). *L’Arabe parlé à Mardin. Monographie d’un parler arabe ”périphérique”*. Editura Universității din București.
- GRILLO, Antonino (2005). “Minimality effects in agrammatic comprehension”. In : *Proceedings of ConSOLE XIII*. York, p. 107–120.
- GRILLO, Antonino et Keir MOULTON (2015). “Pseudo Relatives : Big and Direct”. In : *45th North Eastern Linguistic Society*. York, p. 193–202.
- (2016). “Event kinds and the pseudo-relative”. In : *46th North Eastern Linguistic Society*. York.
- (sous révision). “Sorting out Pseudo-Relatives : Clausal Determiners and AGREE”. In : *Natural Language & Linguistic Theory*.
- GRILLO, Nino (2008). *Generalized minimality : Syntactic underspecification in Broca’s aphasia*. LOT.
- GRIMSHAW, Jane (1979). “Complement selection and the lexicon”. In : *Linguistic Inquiry* 10.2, p. 279–326.
- GRIMSHAW, Jane Barbara (1977). “English Wh-Constructions and the Theory of Grammar”. Unpublished. Thèse de doct. University of Massachusetts, Amherst.
- GROENENDIJK, Jeroen et Martin STOKHOF (1982). “Semantic analysis of wh-complements”. In : *Linguistics and philosophy* 5.2, p. 175–233.
- GROOS, Anneke et Henk VAN RIEMSDIJK (1981). “Matching effects in free relatives : A parameter of core grammar”. In : *Theory of markedness in generative grammar*, p. 171–216.
- GROSS, Maurice (1968). *Grammaire transformationnelle du français*. Larousse.
- (1975). *Méthodes en syntaxe - régime des constructions complétives*. Paris : Hermann.
- GROSU, Alexander et Fred LANDMAN (1998). “Strange relatives of the third kind”. In : *Natural language semantics* 6.2, p. 125–170.
- GUASTI, Maria (1993). *Causative and perception verbs. A comparative study*. Rosenberg & Sellier.
- GUASTI, Maria Teresa (1988). “La pseudorelative et les phénomènes d’accord”. In : *Rivista di Grammatica Generativa* 13, p. 35–57.

- GUASTI, Maria Teresa et Ur SHLONSKY (1995). "The acquisition of French relative clauses reconsidered". In : *Language Acquisition* 4.4, p. 257–276.
- GUASTI, T (1992). "Pseudorelatives and Prepositional Infinitives : A Unified Account". In : *Geneva Generative Papers* 1, p. 53–65.
- HAMANN, Cornelia (2006). "Speculations about early syntax : The production of wh-questions by normally developing French children and French children with SLI". In : *Catalan Journal of Linguistics* 5.1, p. 143–189.
- (2008). "The first language acquisition of French from a generative perspective". In : *Studies in French Applied Linguistics*. Ayoun D (ed.), Amsterdam : John Benjamins, p. 135–183.
- HEIM, Irene (1979). "Concealed questions". In : *Semantics from Different Points of View*. Sous la dir. de R. BÄUERLE, U. EGLI et A. von STECHOW. Springer, p. 51–60.
- HIRSCHBUHLER, Paul (2016). *The syntax and semantics of wh-constructions*. Routledge.
- HIRSCHBÜHLER, Paul et Marie LABELLE (1990). "Celui comme noyau de syntagme nominal". In : *Travaux de linguistique* 20, p. 109–122.
- HIRSCHBÜHLER, Paul et Maria-Luisa RIVERO (1983). "Remarks on free relatives and matching phenomena". In : *Linguistic Inquiry* 14.3, p. 505–520.
- HOFMEISTER, Philip et Ivan A SAG (2010). "Cognitive constraints on syntactic islands". In : *Language* 86.2, p. 366–415.
- HUDDLESTON, Rodney, Geoffrey PULLUM et Peter PETERSON (2002). "Cambridge grammar of the English language". In : sous la dir. de R. HUDDLESTON et G. PULLUM. Cambridge University Press. Chap. Relative constructions and unbounded dependencies, p. 1031–1096.
- HULK, Aafke (1995). "The syntax of wh-questions in Child French". In : *W. Philip, W et F. Wijnen (eds.), Connecting Children's Language and Linguistic Theory, Amsterdam Series in Child Language* 5, p. 129–172.
- HULK, Aafke et Shalom ZUCKERMAN (2000). "The interaction between input and economy : Acquiring optionality in French wh-questions". In : *Proceedings of the 24th Annual Boston University Conference on Language Development*. Sous la dir. de Sarah A. Fish S. CATHERINE HOWELL et Thea KEITH-LUCAS. Somerville, MA : Cascadilla.
- HUOT, Hélène (1981). *Constructions infinitives du français : le subordonnant de*. Librairie Droz, Genève.
- IATRIDOU, Sabine et David EMBICK (1997). "Apropos pro". In : *Language* 73, p. 58–78.
- IZVORSKI, Roumyana (1996). "(Non-) Matching Effects in Free Relatives and pro-Drop". In : *Proceedings of the 12th Eastern States Conference on Linguistics (ESCOL 1995)*. Sous la dir. de M. PRZEZDZIECKI et L. WHALEY. Dartmouth College : Hanover, New Hampshire.
- JACOBSON, Pauline (1995). "On the quantificational force of English free relatives". In : *Quantification in natural languages*. Springer, p. 451–486.
- JAEGER, T Florian (2008). "Categorical data analysis : Away from ANOVAs (transformation or not) and towards logit mixed models". In : *Journal of memory and language* 59.4, p. 434–446.

- JAKUBOWICZ, Celia (2003). "Computational complexity and the acquisition of functional categories by French-speaking children with SLI". In : *Linguistics* 41.2, p. 175–212.
- (2004). "Is movement costly? The grammar and the processor in language acquisition". In : *Comunicação apresentada na JEL*.
- JAKUBOWICZ, Celia et Lea NASH (2001). "Functional categories and syntactic operations in (ab) normal language acquisition". In : *Brain and Language* 77.3, p. 321–339.
- JAKUBOWICZ, Celia et Nelleke STRIK (2008). "Scope-marking strategies in the acquisition of long distance wh-questions in French and Dutch". In : *Language and Speech* 51.1-2, p. 101–132.
- JONES, Michael Allan (1996). *Foundations of French syntax*. Cambridge University Press.
- KATO, Mary et Jairo NUNES (2009). "A uniform raising analysis for standard and nonstandard relative clauses in Brazilian Portuguese". In : *Minimalist essays on Brazilian Portuguese syntax*, p. 93–120.
- KAYNE, Richard (1976). "French relative que". In : *Current studies in Romance linguistics*, p. 255–299.
- (1984). "Connectedness". In : *Linguistic Inquiry* 14, p. 223–49.
- KAYNE, Richard S (1994). *The Antisymmetry of Syntax : Linguistic Inquiry Monographs*. Cambridge, Mass./London, England : MIT Press.
- (1999). "Prepositional complementizers as attractors". In : *Probus* 11.1, p. 39–74.
- (2014). "Why isn't this a complementizer?" In : *Functional Structure from Top to Toe : A Festschrift for Tarald Taraldsen*, Oxford University Press, New York.
- KHOLODILOVA, Maria (2013). "Inverse attraction in Ingrian Finnish relative clauses". In : *Linguistica Uralica* 49.2, p. 96–116.
- KONRAD, Ingrid (A paraître). "At the crossroads between (semi-)free relatives and indirect questions in French". In : *Romance Languages and Linguistic Theory*. John Benjamins.
- (2016). "At the crossroads between (semi-)free relatives and indirect questions in French". In : *30ème édition de Going Romance (GR)*. Communication à l'Université de Goethe, à Francfort, en décembre 2016.
- (2017a). "L'acquisition d'une construction romane ambiguë - étude comparative entre le français et le portugais". In : *10ème Journée des Doctorants*. Communication à l'Université Paris Diderot, en octobre 2017.
- (2017b). "On (semi-)free relatives and indirect questions in French". In : *14ème édition du Workshop on Syntax, Semantics and Phonology (WoSSP)*. Communication à l'Université d'Oviedo, en juin 2017.
- (2018a). "The Acquisition of A-bar dependencies in French and Portuguese : when the nature of the head/wh-word matters regardless of the gap position". In : *9ème édition de The Romance Turn*. Communication à l'Université de Bucarest, en août 2018.
- (2018b). "When indirect questions are disguised as relative clauses - the case of French 'ce que'". In : *26ème édition de Conference of the Student Organisation of Linguistics in Europe (ConSOLE)*. Communication à UCL (University College London), à Londres, en février 2018.

- (2018c). “When nominalization pervades indirect questions - a syntactic and diachronic analysis of French ‘ce que’”. In : *45ème édition de Incontro di Grammatica Generativa (IGG)*. Communication à Padoue, en février 2018.
 - (2019). “Des questions indirectes déguisées en relatives : étude syntaxique et diachronique de ‘ce que’ en français”. In : *Rencontres lyonnaises des jeunes chercheurs en linguistique historique*. Communication à L’Université Jean Moulin Lyon 3, en juin 2019.
 - (2015 (non publié)). “Les relatives pleines, pronominales et libres en français : syntaxe et acquisition”. Mém.de mast. Université Paris Diderot.
- KONRAD, Ingrid, Massimo BURATTIN et al. (Soumis). “Avoiding gaps in Romance : Evidence for a structural parsing principle from Italian and French.” In : *Syntax*.
- KONRAD, Ingrid, Caterina DONATI et al. (2018). “Processing ambiguities in Romance : another gap to avoid”. In : *Workshop Ambigo*. Communication à l’Université de Göttingen, en juillet 2018.
- KONRAD, Ingrid, Yair HAENDLER et Caterina DONATI (2017a). “The acquisition of a French ambiguous embedded structure - a comparative study with European Portuguese”. In : *Colloque en l’honneur de Janet Fodor, Language Structure in Acquisition and Processing*. Communication à l’Université Paris Diderot, en octobre 2017.
- (2017b). “The acquisition of French ambiguous embedded structures introduced by ‘ce que’”. In : *42ème édition de Boston University Conference on Language Development (BUCLD)*. Présentation poster à l’Université de Boston, en novembre 2017.
 - (2018). “The acquisition of French ambiguous embedded structures introduced by ‘ce que’”. In : *42nd Boston University Conference on Language Development (BUCLD) Proceedings*. T. 1. Boston, MA, Cascadilla Press, p. 400–412.
- KOOPMAN, Hilda (1982). “Quelques problèmes concernant que/quoi, ce que et qu’est-ce que”. In : *La syntaxe comparée du français standard et populaire : approches formelle et fonctionnelle* 1, p. 135–170.
- KOOPMAN, Hilda et Dominique SPORTICHE (2008). “The que/qui alternation : new analytical directions”. In : *Manuscript, University of California, Los Angeles*. Available from : <http://ling.auf.net/lingBuzz/000638>.
- KUNSTMANN, Pierre (1990). *Le relatif-interrogatif en ancien français*. Librairie Droz.
- LABELLE, Marie (1990). “Predication, wh-movement, and the development of relative clauses”. In : *Language acquisition* 1.1, p. 95–119.
- (1996). “The acquisition of relative clauses : Movement or no movement?” In : *Language acquisition* 5.2, p. 65–82.
- LAMIROY, Béatrice et Michel CHAROLLES (2010). “Les clitiques accusatifs versus datifs dans les constructions causatives en faire”. In : *2ème Congrès Mondial de Linguistique Française*. Nouvelle-Orléans, p. 2141–2150.
- LARSON, Richard K (1987). “Missing prepositions and the analysis of English free relative clauses”. In : *Linguistic inquiry* 18.2, p. 239–266.
- LEARD, Jean-Marcel (1987). “Le développement de ce que en français et l’évolution du relatif-interrogatif-exclamatif”. In : *Historical linguistics*, p. 289–302.

- LEE, Kwee-Ock (1991). "On the First Language Acquisition of Relative Clauses in Korean : The Universal Structure of COMP". Thèse de doct. Cornell University.
- LEES, Robert B (1961). "The constituent structure of noun phrases". In : *American Speech* 36.3, p. 159–168.
- LIGHTFOOT, David W (2018). "12. Nothing in Syntax Makes Sense Except in the Light of Change". In : *Language, Syntax, and the Natural Sciences*, p. 224.
- LOBO, Maria, Ana Lúcia SANTOS et Carla SOARES-JESEL (2016). "Syntactic structure and information structure : the acquisition of Portuguese clefts and be-fragments". In : *Language acquisition* 23.2, p. 142–174.
- LUST, Barbara, Suzanne FLYNN et Claire FOLEY (1996). "What children know about what they say : Elicited imitation as a research method for assessing children's syntax". In : *Methods for assessing children's syntax*, p. 55–76.
- MACWHINNEY, Brian (2000). *The CHILDES project : Tools for analyzing talk : Volume I : Transcription format and programs, volume II : The database*.
- MANZINI, Maria Rita (2014). "The Romance k-complementizers". In : *Functional structure from top to toe : The cartography of syntactic structures* 9, p. 148–187.
- MANZINI, Maria Rita et Leonardo SAVOIA (2003). "The nature of complementizers". In : *Rivista di Grammatica Generativa* 28, p. 87–110.
- (2005). *I dialetti italiani e romanci. Morfosintassi generativa*. Alessandria : Ed. dell'Orso.
- (2011). *Grammatical categories : Variation in Romance languages*. Cambridge University Press.
- MATHIEU, Eric (2004). "The mapping of form and interpretation : The case of optional wh-movement in French". In : *Lingua* 114.9, p. 1090–1132.
- MATUSCHEK, Hannes et al. (2017). "Balancing Type I error and power in linear mixed models". In : *Journal of Memory and Language* 94, p. 305–315.
- MEDEIROS JUNIOR, Paulo (2016). "From [o[que]] to [o que] in Brazilian Portuguese Free Relatives : A diachronic view". In : Kato, Mary A. & Francisco Ordoñez (eds.), *The Morphosyntax of Portuguese and Spanish in Latin America*. New York : Oxford University Press, p. 308–331.
- MORGENSTERN, Aliyah et Christophe PARISSÉ (2007). "Codage et interprétation du langage spontané d'enfants de 1 à 3 ans". In : *Corpus* 6, p. 55–78.
- MOURET, François (2007). "Grammaire des constructions coordonnées. Coordinations simples et coordinations à redoublement en français contemporain". Thèse de doct. Université Paris 7.
- MULLER, Claude (1989). "Sur la syntaxe et la sémantique des relatives indépendantes et des interrogatives indirectes partielles". In : *Revue romane* 1.
- (1996). *La subordination en français : le schème corrélatif*. Armand Colin.
- MÜLLER, Gereon (1993). "On deriving movement type asymmetries". In : *Seminar fur Sprachwissenschaft, Eberhard-Karls-Universität Tübingen : Doctoral Dissertation*.
- MUNARO, Nicola (2001). "Free relatives as defective wh-elements". In : *Romance Languages and Linguistic Theory 1999 : Selected papers from Going Romance 1999, Leiden, 9 11 December 1999* 221, p. 281.

- NATHAN, Lance Edward (2006). "On the interpretation of concealed questions". Thèse de doct. Massachusetts Institute of Technology.
- NIÉGER, Monique (1975). "L'interrogation indirecte : étude diachronique". In : *Cahier de linguistique* 5, p. 1–15.
- OBENAUER, Hans-Georg (1976). "Etudes de syntaxe interrogative du français : quoi, combien et le complémenteur". Thèse de doct.
- ÖHL, Peter et Ahmad LOTFI (2007). "Nominalised CPs in Persian : a parametric account". In : *Proceedings of the 2nd International Conference on Iranian Linguistics*.
- PACKARD, Jerome L et al. (1988). "1. Constraints on NP-movement in mandarin Chinese". In : *Journal of Chinese Linguistics* 16.1.
- PESETSKY, David (1982). "Complementizer-trace phenomena and the nominative island condition". In : *The Linguistic Review* 1.3, p. 297–344.
- (1987). "Wh-in-situ : Movement and unselective binding". In : *The representation of (in) definiteness* 98, p. 98–129.
- PHILLIPS, Colin (2013). "Some arguments and nonarguments for reductionist accounts of syntactic phenomena". In : *Language and Cognitive Processes* 28.1-2, p. 156–187.
- PICALLO, M Carme (2002). "Abstract agreement and clausal arguments". In : *Syntax* 5.2, p. 116–147.
- PIERRARD, Michel (1988). *La relative sans antécédent en français moderne : essai de syntaxe propositionnelle*. Peeters Publishers.
- (1993). "Ce que l'on conçoit bien s'énonce clairement : la valeur de ce qu-dans les relatives nominalisées." In : *Revue romane* 1.
- (1995). "Système et histoire du système : l'évolution de" ce que" introducteur de subordonnées". In : *Langue française*, p. 111–124.
- PITNER, Karin (1995). "The case of German relatives". In : *The linguistic review* 12.3, p. 197–231.
- PLANN, Susan (1981). "The two *el*+ infinitive constructions in Spanish". In : *Linguistic Analysis Seattle, Wash.* 7.3, p. 203–240.
- PLUNKETT, Bernadette (1999). "Targeting complex structure in French questions". In : *Proceedings of the 23rd Annual Boston University Conference on Language Development*, p. 764–775.
- (2000). "What's 'what' in French questions". In : *Journal of linguistics* 36.3, p. 511–530.
- RADFORD, Andrew (1975). "Pseudo-relatives and the unity of subject raising". In : *Archivum linguisticum* 6.32-64.
- (1977). *Italian syntax : transformational and relational grammar*. Cambridge University Press.
- RCORETEAM (2017). *R : A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. 2016.
- REBUSCHI, Georges (2001). "Semi-Free Relative Clauses and the DP Hypothesis : Basque Evidence and Theoretical Consequences". In : *Semi-Free Relative Clauses and the DP Hy-*

- pothesis : Basque Evidence and Theoretical Consequences*. 8. Tel Aviv, Israel : Alex Grosu (ed) - Tel Aviv University, p. 55–64.
- RIEMSDIJK, Henk van (2000). “Free relatives”. In : *SynCom*. Blackwell.
- RINKE, Esther et Elisabeth ABMANN (2017). “The syntax of relative clauses in European Portuguese. Extending the Determiner Hypothesis of Relativizers to relative que”. In : *Journal of Portuguese Linguistics* 16.1.
- RIZZI, Luigi (1982). *Issues in Italian Syntax*. Foris, Dordrecht.
- (1990). *Relativized minimality*. The MIT Press.
- (1997). “The fine structure of the left periphery”. In : *Elements of grammar*. Springer, p. 281–337.
- (2001). “On the position “Int (errogative)” in the left periphery of the clause”. In : *Current studies in Italian syntax : Essays offered to Lorenzo Renzi*, p. 287–296.
- (2004). “Locality and left periphery”. In : *Structures and beyond : The cartography of syntactic structures* 3, p. 223–251.
- (2018). “Intervention effects in grammar and language acquisition”. In : *Probus* 30.2, p. 339–367.
- ROBERTS, Ian et Anna ROUSSOU (2003). *Syntactic change : A minimalist approach to grammaticalization*. Cambridge University Press.
- ROMAINE, Suzanne (1989). “The role of children in linguistic change”. In : *Breivik, LE & Jahr, EH eds*, p. 199–226.
- ROMERO, Maribel (2006). “On concealed questions”. In : *Semantics and Linguistic Theory*. T. 16, p. 208–227.
- ROSS, John Robert (1967). “Constraints on variables in syntax.” Thèse de doct. MIT, Cambridge, MA.
- ROTHSTEIN, Susan D (1992). “Case and NP licensing”. In : *Natural Language & Linguistic Theory* 10.1, p. 119–139.
- ROUQUIER, Magali (1988). “Le ce que-phrases en ancien français”. Thèse de doct. Université Paris 7.
- (2002). “Les interrogatives en ‘qui/qu’est-ce qui/que’ en ancien français et en moyen français”. In : *Cahiers de grammaire* 27, p. 97–120.
- ROUSSOU, Anna (1991). “Nominalized clauses in the syntax of Modern Greek”. In : *UCL working papers in linguistics* 3, p. 77–100.
- SAG, Ivan A et al. (1985). “Coordination and how to distinguish categories”. In : *Natural Language & Linguistic Theory* 3.2, p. 117–171.
- SALVESEN, Christine Meklenborg (2008). “Ce que and ce que are two different things”. In : *Linguistics in the making : selected papers from the second Scandinavian PhD conference in Linguistics and philology*, p. 413–435.
- SAUERLAND, Uli (2003). “Unpronounced heads in relative clauses”. In : *The interfaces : Deriving and interpreting omitted structures* 61, p. 205–226.
- SCHACHTER, Paul (1973). “Focus and relativization”. In : *Language*, p. 19–46.

- SHLONSKY, Ur (2012). “Notes on wh in situ in French”. In : *Functional heads : The cartography of syntactic structures*. L. Bruge, A. Cardinaletti, G. Giusti, N. Munaro et C. Poletto (eds). New York : Oxford University Press, p. 242–252.
- SLEEMAN, Petra (2003). “Quelques-uns, celui, chacun : pronoms ou déterminants?” In : *Langages*, p. 105–123.
- SLOBIN, D (2004). “From Ontogenesis to Phylogenesis : what can child language tell us about language evolution?” In : *Biology and Knowledge revisited : From neurogenesis to psychogenesis*. Mahwah, Nj : Lawrence Erlbaum Associates.
- SLOBIN, Dan I (2002). “Language evolution, acquisition and diachrony : Probing the parallels”. In : *The evolution of language out of pre-language*, p. 375–392.
- SOARES, Carla (2003). “The C-domain and the acquisition of European Portuguese : The case of wh-questions”. In : *Probus* 15.1, p. 147–176.
- (2004). “Computational complexity and the acquisition of the CP field in European Portuguese”. In : *Proceedings of Console XII*. Sous la dir. de L. Vicente S. BLAHO et M. de VOS, p. 125–140.
- (2006). “La syntaxe de la périphérie gauche en portugais européen et son acquisition”. Thèse de doct. Université Paris 8.
- SPORTICHE, Dominique (2011). “French relative qui”. In : *Linguistic Inquiry* 42.1, p. 83–124.
- STAUB, Adrian et al. (2018). “Relative clause avoidance : Evidence for a structural parsing principle”. In : *Journal of Memory and Language* 98, p. 26–44.
- STOWE, Laurie A (1986). “Parsing WH-constructions : Evidence for on-line gap location”. In : *Language and cognitive processes* 1.3, p. 227–245.
- STOWELL, Timothy Angus (1981). “Origins of phrase structure”. Thèse de doct. Massachusetts Institute of Technology.
- STRIK, Nelleke (2007). “L’acquisition des phrases interrogatives chez les enfants francophones”. In : *Psychologie française* 52.1, p. 27–39.
- SUÑER, Margarita (1993). “About indirect questions and semi-questions”. In : *Linguistics and Philosophy* 16.1, p. 45–77.
- SZABOLCSI, Anna (2006). “Strong vs. weak islands”. In : *The Blackwell companion to syntax* 4, p. 479–531.
- VAN RIEMSDIJK, Henk (2006). “Free relatives”. In : *The Blackwell companion to syntax*, p. 338–382.
- VELAND, Reidar (1990). “Structure fonctionnelle et interprétation lexicale du syntagme en celui. À propos d’un article de Michel Pierrard”. In : *Travaux de linguistique* 20, p. 123–134.
- VERGNAUD, Jean-Roger (1974). “French relative clauses.” Thèse de doct. Massachusetts Institute of technology.
- WASOW, Thomas (1997). “Remarks on grammatical weight”. In : *Language variation and change* 9.1, p. 81–105.
- WILLIAMSON, Janis (1987). “An indefiniteness restriction for relative clauses in Lakhota”. In : *The representation of (in) definiteness*, p. 168–190.

- WOOD, Jim, Einar Freyr SIGURÐSSON et Iris Edda NOWENSTEIN (2017). "Inverse attraction in Icelandic relative clauses". In : *Syntactic variation in Insular Scandinavian*, p. 199–232.
- YAMAKOSHI, Kyoko (2002). "Wh-drop in child languages and adult ASL". In : *Proceedings of ConSole IX*. Lund : Lund University.
- YAR-SHATER, Ehsan (1969). *A grammar of Southern Tati dialects*. The Hague, Paris : Mouton.
- ZARING, Laurie (1992). "French ce as a clausal determiner". In : *Probus* 4.1, p. 53–80.
- ZARING, Laurie et Paul HIRSCHBÜHLER (1997). "Qu'est-ce que ce que ? The diachronic evolution of a French complementizer". In : *Parameters of morphosyntactic change*, p. 353–379.
- ZRIBI-HERTZ, Anne (1996). *L'anaphore et les pronoms : une introduction à la syntaxe générative*. Villeneuve d'Ascq. Presses Univ. Septentrion.

Bibliographie des sources citées dans le Chapitre 3

En Ancien Français

- ANONYME (1155). *Eneas*. T. 1.
– (1223). *Queste del Saint Graal*. Anonyme.
– (1230). *La mort le roi Artu*.
– (1250). *Del Confortement de Philosophie*.
– (1322). *Chroniques de Morée*.
– (1325). *Donat (M2)*.
BORON, Robert de (1199). *Le Roman de l'Estoire dou Graal*.
COINCI, Gautier de (1223a). *Miracles de Notre-Dame*. T. 2.
– (1223b). *Miracles de Notre-Dame*. T. 4.
D'ARRAS, Gautier (1180). *Eracle*.
DENAIN, Wauchier de (1210). *Deuxième continuation de Perceval*.
JOINVILLE, Jean de (1307). *Mémoires ou Vie de Saint Louis*.
LA HALLE, Adam de (1285). *Le jeu de Robin et Marion*.
LORRIS, Guillaume de (1230). *Le Roman de la Rose*.
PARIS, Alexandre de (1180). *Roman d'Alexandre*.
TROYES, Chrétien de (1176). *Cligès*.
– (1177). *Lancelot ou Le Chevalier à la Charrette*.
– (1179). *Yvain ou Le Chevalier au Lion*.
– (1181a). *Conte du Graal*.
– (1181b). *Perceval*.

En Moyen Français

- ANONYME (1346). *Miracle de un pape qui vendi le basme*.
– (1365). *Miracle de Nostre Dame d'Amis et d'Amille*.
– (1390). *Les Quinze joies de mariage*.
BAGNYON, Jean (1468). *L'Histoire de Charlemagne*.
BUEIL, Jean de (1461). *Le Jouvenciel*. T. 2.

- COMMYNES, Philippe de (1490). *Mémoires*. T. 2.
 — (1497). *Mémoires*. T. 3.
 CRAPILLET, Pierre (1450). *Cur deus homo*.
 DESCHAMPS, Eustache (1394). *Le Miroir de mariage*.
 FAUQUEMBERGUE, Clément de (1433). *Journal*. T. 3.
 FROISSART, Jean (1400). *Chroniques*.
 LEFEVRE, Raoul (1460). *L'Histoire de Jason*.
 MACHAUT, Guillaume de (1341). *Remède de Fortune*.
 RÉGNIER, Jean (1432). *Les Fortunes et adversitez*.

En Français du 16ème siècle

- ANONYME (1521). *Le Violier des histoires romaines moralisées*.
 BELON, Pierre (1555). *L'Histoire de la nature des oyseaux*.
 BRUÈS, Guy de (1557). *Les Dialogues de Guy de Bruès contre les nouveaux académiciens*.
 LE CLERC, Jean (1502). *Interpolations et variantes de la Chronique scandaleuse*.
 MONTAIGNE, Michel de (1592). *Essais*. T. 1.
 PALMA-CAYET, Pierre-Victor (1598). *L'Histoire prodigieuse du Docteur Fauste*.
 RIVAudeau, André de (1566). *Aman : tragédie sainte*.
 VIENNE, Philibert de (1548). *Le Philosophe de court*.

En Français Classique

- BUSSY-RABUTIN, Roger de (1681). *Les Lettres de messire Roger de Rabutin, comte de Bussy*.
 T. 1.
 CORNEILLE, Pierre (1639). *L'Illusion comique*.
 — (1682). *La Galerie du Palais*.
 D'URFÉ, Honoré (1610). *L'Astrée*. T. 2.
 — (1631). *L'Astrée*. T. 3.
 HEROARD, Jean (1608). *Histoire particulière de Louis XIII*.
 L'HERMITE, Tristan (1639). *Panthée*.
 MALLEVILLE, Claude (1649). *Oeuvres poétiques*.
 MONTCHRESTIEN, Antoine de (1615). *Traicté de l'oeconomie politique*.
 SALES, Saint-François de (1619). *Introduction à la vie dévote*.
 SÉVIGNÉ, Madame de (1675). *Correspondances*. T. 1.
 VAUGELAS, Claude Favre de (1647). *Remarques sur la langue françoise utiles à ceux qui veulent
 bien parler et bien escrire*.
 VIAU, Théophile de (1621). *Oeuvres poétiques*.

Appendices

Annexe A

**Avis Favorable par le Conseil
d'évaluation éthique pour les recherches
en santé (CERES) de l'Université Paris
Descartes**



Conseil d'évaluation éthique pour les recherches en santé (CERES)

Présidente: Jacqueline Fagard

N° 2017 - 28

PROTOCOLE : Expériences sur l'acquisition des relatives avec tête allégée et des interrogatives indirectes en français- étude comparative avec le portugais européen.

Noms du/ des chercheur(s) : Ingrid Konrad et Caterina Donati

Email pour la correspondance :
konrad.ingrid4@gmail.com

Labo/ Service:
Université Paris Diderot
Laboratoire de Linguistique Formelle
8, rue Albert Einstein - 75013 Paris

Évalué à la séance du 25/04/2017

AVIS : Favorable

N°CERES : 201728

Jacqueline Fagard

Annexe B

Liste des items du test de répétition de phrases contenant des subordonnées introduites par *ce qu-* en français (Chapitre 4, Section 4.4)

B.1 Items d'entraînement

1. Le garçon croit que la petite fille est partie.
2. L'élève trouve que la maîtresse est très gentille.
3. Mamie pense que papi a trop mangé de soupe.

B.2 Items expérimentaux

B.2.1 Relatives

B.2.1.1 Sujet

1. Le papa mange **ce qui** est tombé du placard.
2. La directrice efface **ce qui** est au tableau.
3. L'enfant attrape **ce qui** est sorti de la boîte.
4. Le monsieur attrape **ce qui** a roulé par terre.

B.2.1.2 Objet

5. Le maître efface **ce que** la fille a dessiné.
6. L'élève attrape **ce que** la maîtresse a lancé.

7. La dame mange **ce que** la maman a cuisiné.
8. La nounou attrape **ce que** bébé a caché.

B.2.2 Interrogatives indirectes

B.2.2.1 Sujet

9. Papa sait pas **ce qui** est tombé du placard.
10. La directrice demande **ce qui** est au tableau.
11. L'enfant devine **ce qui** est sorti de la boîte.
12. Le monsieur demande **ce qui** a roulé par terre.

B.2.2.2 Objet

13. Le maître demande **ce que** la fille a dessiné.
14. L'élève demande **ce que** la maîtresse a lancé.
15. La dame sait pas **ce que** maman a cuisiné.
16. La nounou devine **ce que** bébé a caché.

B.3 Distracteurs

1. Le cahier de Léo est blanc.
2. Marie est amoureuse de Lucas.
3. Il y a beaucoup d'élèves dans la classe.
4. Ma trousse est remplie de stylos.
5. Je vais à la garderie mardi.
6. Il mange à la cantine lundi.
7. La fille adore la peinture.
8. C'est l'heure de la récréation.
9. Les règles du jeu sont très simples.
10. Le garçon fait beaucoup de sport.

Annexe C

Liste des items du test de répétition de phrases contenant des subordonnées introduites par *qui* en français (Chapitre 4, Section 4.7)

C.1 Items d'entraînement

1. Mamie est sûre que papi a mangé le pain.
2. La nounou pense que le bébé est très malade.
3. Papa pense que maman a fait de bons gâteaux.
4. La maîtresse trouve que Julie sait bien dessiner.
5. Le monsieur trouve que la petite fille est gentille.

C.2 Items expérimentaux

C.2.1 Relatives libres sujet

1. La maîtresse va punir **qui** dira un gros mot.
2. Le monsieur va punir **qui** crachera par terre.
3. L'enfant va inviter **qui** gagnera le jeu.
4. Max va applaudir **qui** trouvera le trésor.
5. Mamie va gronder **qui** se cachera dehors.
6. Papa va applaudir **qui** fera une roulade.
7. Tatie va gronder **qui** écrira sur le livre.
8. Tom va inviter **qui** portera des bonbons.

C.2.2 Interrogatives indirectes sujet

9. Le maître demande **qui** a dit un gros mot.
10. La dame demande **qui** a craché par terre.
11. La nounou ne sait pas **qui** a gagné le jeu.
12. Maman se demande **qui** a trouvé le trésor.
13. Le directeur devine **qui** est caché dehors.
14. Amélie ne sait pas **qui** a fait une roulade.
15. Papi se demande **qui** a écrit sur le livre.
16. Mélanie devine **qui** a porté des bonbons.

C.3 Distracteurs

1. Le cahier de Léa est noir.
2. Lucie est amoureuse de Dylan.
3. Il y a beaucoup de garçons dehors.
4. Ma trousse est remplie de crayons.
5. Je vais à la garderie jeudi.
6. Il mange à la cantine mardi.
7. La fille adore la danse et le chant.
8. Julie aime bien manger des pâtes.
9. Le jeu est facile et rigolo.
10. Le garçon aime jouer au foot.

Annexe D

Liste des items du test de répétition de phrases contenant des subordonnées introduites par *o que* en portugais européen (Chapitre 5, Section 5.3)

D.1 Items d'entraînement

1. O pai tem a certeza que a mãe fez um bolo.
Le papa est certain que la maman a fait un gâteau.
2. O menino pensa que a menina foi brincar.
Le garçon pense que la fille est allée jouer.
3. A mãe acha que o professor é simpático.
La maman trouve que le maître est gentil.
4. A neta pensa que o avô comeu muito pão.
La petite-fille pense que le papi a mangé beaucoup de pain.
5. A profesora acha que o menino mentiu.
La maîtresse trouve que le garçon a menti.

D.2 Items expérimentaux

D.2.1 Relatives

D.2.1.1 Sujet

1. A tia come **o que** caiu do armário.
La tante mange ce qui est tombé de l'armoire.
2. O professor apaga **o que** está no quadro.
Le maître efface ce qui est au tableau.
3. A criança apanha **o que** saiu da caixa.
L'enfant attrape ce qui est sorti de la boîte.
4. O senhor apanha **o que** caiu na sala.
Le monsieur attrape ce qui est tombé dans la salle.

D.2.1.2 Objet

5. O pai apaga **o que** a menina desenhou.
Le papa efface ce que la fille a dessiné.
6. O senhor apanha **o que** o rapaz lhe atirou.
Le monsieur attrape ce que le garçon lui a lancé.
7. A senhora come **o que** a filha cozinhou.
La dame mange ce que la fille a cuisiné.
8. O senhor apanha **o que** o menino partiu.
Le monsieur ramasse ce que le garçon a cassé.

D.2.2 Interrogatives indirectes

D.2.2.1 Sujet

9. O tio não sabe **o que** caiu do armário.
L'oncle sait ce qui est tombé de l'armoire.
10. A menina adivinha **o que** está no quadro.
La fille devine ce qui est au tableau.
11. O avô pergunta **o que** saiu da caixa.
Le papi demande ce qui est sorti de la boîte.
12. O senhor pergunta **o que** caiu na sala.
Le monsieur demande ce qui est tombé dans la salle.

D.2.2.2 Objet

13. O pai pergunta **o que** a menina desenhou.
Le papa demande ce que la fille a dessiné.
14. A tia pergunta **o que** o rapaz atirou.
La tante demande ce que le garçon a lancé.
15. O avô não sabe **o que** a prima cozinhou.
Le papi ne sait pas ce que la cousine a cuisiné.
16. A mãe adivinha **o que** o menino partiu.
La maman devine ce que le garçon a cassé.

D.3 Distracteurs

1. O lápis do João é branco.
Le crayon de Jean est blanc.
2. A Maria gosta do João.
Marie aime Jean.
3. Há muitos meninos na turma.
Il y a beaucoup de garçons dans la classe.
4. O meu caderno é castanho.
Mon cahier est marron.
5. Eu vou ao parque no domingo.
Je vais au parc le dimanche.
6. A Matilde come no café.
Mathilde mange dans le café.
7. A menina gosta de pintar.
La fille aime bien peindre.
8. Chegou a hora do recreio.
J'arrive à l'heure de la récréation.
9. As regras do jogo são fáceis.
Les règles du jeu sont faciles.
10. O menino faz ginástica.
Le garçon fait de la gymnastique.

Annexe E

Liste des items du test de répétition de phrases contenant des subordinées introduites par *quem* en portugais européen (Chapitre 5, Section 5.6)

E.1 Items d'entraînement

1. O pai acha que o menino é educado.
Le papa trouve que le garçon est poli.
2. A mãe pensa que a senhora está constipada.
La maman pense que la dame est enrhumée.
3. O professor acha que a Clara desenha bem.
Le maître trouve que Clara dessine bien.
4. A senhora pensa que a tia cozinha bem.
La dame pense que la tante cuisine bien.
5. A mãe tem a certeza que o pai comeu o pão.
La maman est certaine que le papa a mangé le pain.

E.2 Items expérimentaux

E.2.1 Relatives libres sujet

1. O avô beijará *quem* pintar o desenho.
Le papi embrassera qui coloriera le dessin.

2. A senhora beijará **quem** lavar a loiça.
La dame embrassera qui fera la vaisselle.
3. O pai abraçará **quem** trazer um presente.
Le papa serrera dans ses bras qui apportera un cadeau.
4. O professor aplaudirá **quem** comer tudo.
Le maître applaudira qui mangera tout.
5. O avô castigará **quem** partir o vaso.
Le papi punira qui cassera le vase.
6. O Duarte aplaudirá **quem** comprar um bolo.
Édouard applaudira qui achètera un gâteau.
7. O senhor castigará **quem** rasgar o livro.
Le monsieur punira qui déchirera le livre.
8. A Beatriz abraçará **quem** levar gomas.
Béatrice serrera dans ses bras qui portera des bonbons.

E.2.2 Interrogatives indirectes sujet

9. O avô perguntou **quem** pintou o desenho.
Le papi a demandé qui a colorié le dessin.
10. A senhora perguntou **quem** lavou a loiça.
La dame a demandé qui a fait la vaisselle.
11. O menino não soube **quem** trouxe o presente.
Le garçon n'a pas su qui a apporté le cadeau.
12. A professora perguntou **quem** comeu tudo.
La maitresse a demandé qui a tout mangé.
13. O senhor adivinhou **quem** partiu o vaso.
Le monsieur a deviné qui a cassé le vase.
14. O Francisco não soube **quem** comprou o bolo.
François n'a pas su qui a acheté le gâteau.
15. A menina perguntou **quem** rasgou o livro.
La fille a demandé qui a déchiré le livre.
16. O diretor adivinhou **quem** levou gomas.
Le directeur a deviné qui a porté des bonbons.

E.3 Distracteurs

1. O lápis do João é preto.
Le stylo de Jean est noir.

2. A Laura gosta do Pedro.
Laure aime Pierre.
3. Há muitos rapazes fora.
Il y a beaucoup de garçons dehors.
4. Há lápis no meu estojo.
Il y a beaucoup de crayons dans ma trousse.
5. Eu vou à praia no verão.
Je vais à la plage en été.
6. Eu hoje comi na cantina.
J'ai mangé à la cantine aujourd'hui.
7. A Catarina faz ballet.
Catherine fait du ballet.
8. A Matilde adora massa.
Mathilde adore les pâtes.
9. O jogo é divertido.
Le jeu est amusant.
10. O menino joga futebol.
Le garçon joue au football.

Annexe F

Liste des items du test de jugement d'acceptabilité ainsi que de l'expérience de lecture par présentation auto-segmentée portant sur le traitement de l'ambiguïté complétives/relatives (Chapitre 6, Section 6.1.3)

F.1 Items d'entraînement

1. Antoine a offert un magnifique bouquet de roses à sa femme.
2. Si je me souviendrais bien, ils ont annoncé de la pluie pour demain.
3. Julie aime bien les vêtements que Pauline a achetés en Italie.

F.2 Items expérimentaux

- (1) a. Pierre veille à ce que Marie enregistre la photo sur son portable. [ambigu, complétive]
b. Pierre veille à ce que Marie enregistre sur son portable. [ambigu, relative]
c. Pierre regarde ce que Marie enregistre sur son portable. [non ambigu, relative]
- (2) a. Léo tient à ce que sa sœur offre un cadeau à leur mère.
b. Léo tient à ce que sa sœur offre à leur mère.
c. Léo apprécie ce que sa sœur offre à leur mère.
- (3) a. Paul fait attention à ce que sa femme regarde le film sur l'ordinateur.

- b. Paul fait attention à ce que sa femme regarde sur l'ordinateur.
 - c. Paul contrôle ce que sa femme regarde sur l'ordinateur.
- (4)
- a. Julien s'oppose à ce que son patron organise une réunion pour le mois prochain.
 - b. Julien s'oppose à ce que son patron organise pour le mois prochain.
 - c. Julien est contre ce que son patron organise pour le mois prochain.
- (5)
- a. Théo se soucie de ce que Pauline poste un nouveau statut sur Facebook.
 - b. Théo se soucie de ce que Pauline poste sur Facebook.
 - c. Théo suit ce que Pauline poste sur Facebook.
- (6)
- a. Mickael est attentif à ce que son frère achète des aliments sains avant le marathon.
 - b. Mickael est attentif à ce que son frère achète avant le marathon.
 - c. Mickael surveille ce que son frère achète avant le marathon.
- (7)
- a. Amélie s'habitue à ce que son père prépare des sushis pour le déjeuner.
 - b. Amélie s'habitue à ce que son père prépare pour le déjeuner.
 - c. Amélie adore ce que son père prépare pour le déjeuner.
- (8)
- a. Le danger vient de ce que la pollution provoque un changement dans l'air des villes.
 - b. Le danger vient de ce que la pollution provoque dans l'air des villes.
 - c. Le danger a pour origine ce que la pollution provoque dans l'air des villes.
- (9)
- a. Luc veille à ce que Lou organise une surprise pour la fête.
 - b. Luc veille à ce que Lou organise pour la fête.
 - c. Luc filme ce que Lou organise pour la fête.
- (10)
- a. Samuel tient à ce que Léa lui donne des livres pour le remercier.
 - b. Samuel tient à ce que Léa lui donne pour le remercier.
 - c. Samuel prend en photo ce que Léa lui donne pour le remercier.
- (11)
- a. Rémi fait attention à ce que son père achète ses médicaments chaque semaine.
 - b. Rémi fait attention à ce que son père achète chaque semaine.
 - c. Rémi vérifie ce que son père achète chaque semaine.
- (12)
- a. Le député s'oppose à ce que le gouvernement propose une nouvelle réforme au Parlement.
 - b. Le député s'oppose à ce que le gouvernement propose au Parlement.
 - c. Le député approuve ce que le gouvernement propose au Parlement.
- (13)
- a. Nicolas se soucie de ce que son ex femme garde des meubles dans la cave.

- b. Nicolas se soucie de ce que son ex femme garde dans la cave.
c. Nicolas fait la liste de ce que son ex femme garde dans la cave.
- (14) a. Marie est attentive à ce que sa sœur raconte une blague à chaque repas.
b. Marie est attentive à ce que sa sœur raconte à chaque repas.
c. Marie écoute ce que sa sœur raconte à chaque repas.
- (15) a. Jonas s'habitue à ce que son frère prépare un poisson d'avril tous les ans.
b. Jonas s'habitue à ce que son frère prépare tous les ans.
c. Jonas déteste ce que son frère prépare tous les ans.
- (16) a. Le problème vient de ce que Samia réclame une augmentation pour le mois prochain.
b. Le problème vient de ce que Samia réclame pour le mois prochain.
c. Le problème concerne ce que Samia réclame pour le mois prochain.
- (17) a. Marc veille à ce que Julie raconte cette anecdote à sa mère.
b. Marc veille à ce que Julie raconte à sa mère.
c. Marc entend ce que Julie raconte à sa mère.
- (18) a. Claire tient à ce que son frère lui achète un bijou pour son anniversaire.
b. Claire tient à ce que son frère lui achète pour son anniversaire.
c. Claire revend ce que son frère lui achète pour son anniversaire.
- (19) a. Pauline fait attention à ce que son mari nettoie les fenêtres dans la chambre.
b. Pauline fait attention à ce que son mari nettoie dans la chambre.
c. Pauline salit ce que son mari nettoie dans la chambre.
- (20) a. Rémi s'oppose à ce que son enfant demande un ordinateur pour Noël.
b. Rémi s'oppose à ce que son enfant demande pour Noël.
c. Rémi mémorise ce que son enfant demande pour Noël.
- (21) a. Lucile se soucie de ce que Gabriel cache un secret à sa femme.
b. Lucile se soucie de ce que Gabriel cache à sa femme.
c. Lucile est au courant de ce que Gabriel cache à sa femme.
- (22) a. Mike est attentif à ce que Jane réserve leurs billets pour Cuba.
b. Mike est attentif à ce que Jane réserve pour Cuba.
c. Mike note ce que Jane réserve pour Cuba.
- (23) a. Yasmina s'habitue à ce que sa mère regarde la télévision chaque matin.
b. Yasmina s'habitue à ce que sa mère regarde chaque matin.
c. Yasmina voit ce que sa mère regarde chaque matin.

- (24) a. L'addiction vient de ce que la nicotine déclenche une stimulation dans le cerveau.
b. L'addiction vient de ce que la nicotine déclenche dans le cerveau.
c. L'addiction découle de ce que la nicotine déclenche dans le cerveau.

F.3 Distracteurs

F.3.1 Concordance temporelle

F.3.1.1 Distracteurs grammaticaux

1. L'université a décidé que tout le monde devait régulariser son inscription au mois d'octobre.
2. Un parieur a prévu que tous les chevaux allaient tomber avant l'arrivée.
3. Le président pensait qu'aucun ministre n'allait critiquer sa dernière déclaration.
4. Si le gouvernement démissionnait, il serait inévitable d'organiser de nouvelles élections.

F.3.1.2 Distracteurs agrammaticaux

5. Si je perdrais mon portefeuille, je ne saurais pas où porter plainte.
6. Le témoin voudrait que la mariée arrêtaît de rire.
7. Mon cousin avait dit que son professeur expliquât mal.
8. La pollution fait que le ciel deviendrait gris.

F.3.2 Îlots

F.3.2.1 Distracteurs grammaticaux

9. J'ai une télévision qui marche mal si j'utilise la télécommande.
10. Il y a un ami à moi qui éternue si j'ouvre la fenêtre.
11. Il y a des équations que tu ne peux résoudre que si tu connais les mathématiques.
12. Cette salle devient froide si tu allumes la climatisation.

F.3.2.2 Distracteurs agrammaticaux

13. Voici la photo que je vais protester si on met dans le catalogue de l'exposition.
14. Ils m'ont donné une tablette que mon professeur se plaint si j'utilise en classe.
15. J'ai une lampe que tu vas te réveiller si j'allume la nuit.
16. Voici le produit que l'usine fermera si les consommateurs n'apprécient pas.

F.3.3 Items à polarité négative

F.3.3.1 Distracteurs grammaticaux

17. Je n'ai jamais vu un arc-en-ciel comme celui de hier.
18. Si jamais j'étais pris au concours, je devrais remercier le professeur de mathématiques.
19. Je suis toujours arrivé à temps à l'école, même si je n'ai jamais utilisé de réveil.
20. Je n'ai jamais eu aucun doute que tu allais revenir pour Noël.

F.3.3.2 Distracteurs agrammaticaux

21. Je n'ai pas jamais rencontré une personne avec un caractère aussi difficile.
22. Si je gagnais à la loterie, je ne gaspillerais pas jamais mon argent.
23. Je n'ai pas protesté contre mon licenciement même si je ne me suis pas jamais absenté.
24. Je n'avais aucun espoir que tu arriverais jamais.

Annexe G

Liste des items de l'expérience de lecture par présentation auto-segmentée portant sur le traitement de l'ambiguïté interrogatives indirectes/relatives (Chapitre 6, Section 6.2.2)

G.1 Items d'entraînement

1. Samia a offert beaucoup de cadeaux à son mari pour son anniversaire.
2. Si j'ai fait une erreur, je m'en rendrais compte tôt ou tard.
3. Justine adore les gaufres maison que prépare sa mamie chaque dimanche.

G.2 Items expérimentaux

- (1)
 - a. Le fils du boulanger apprécie ce que l'apprenti cuisine pour le mariage. [**relative objet**]
 - b. Le fils du boulanger se demande ce que l'apprenti cuisine pour le mariage. [**interrogative indirecte objet**]
 - c. Le fils du boulanger se demande quand l'apprenti cuisine pour le mariage. [**interrogative indirecte adjoint**]
 - d. Le fils du boulanger apprécie ce qui a été cuisiné pour le mariage. [**relative sujet**]
 - e. Le fils du boulanger se demande ce qui a été cuisiné pour le mariage. [**interrogative indirecte sujet**]
- (2)
 - a. Le patron du postier liste ce que le nouveau trie dans le bureau.
 - b. Le patron du postier ignore ce que le nouveau trie dans le bureau.

- c. Le patron du postier ignore quand le nouveau trie dans le bureau.
 - d. Le patron du postier liste ce qui a été trié dans le bureau.
 - e. Le patron du postier ignore ce qui a été trié dans le bureau.
- (3)
- a. La gérante du circuit entend ce que le pilote négocie pour la course.
 - b. La gérante du circuit devine ce que le pilote négocie pour la course.
 - c. La gérante du circuit devine quand le pilote négocie pour la course.
 - d. La gérante du circuit entend ce qui a été négocié pour la course.
 - e. La gérante du circuit devine ce qui a été négocié pour la course.
- (4)
- a. La secrétaire de l'hôpital encaisse ce que le patient paye en espèces.
 - b. La secrétaire de l'hôpital demande ce que le patient paye en espèces.
 - c. La secrétaire de l'hôpital demande quand le patient paye en espèces.
 - d. La secrétaire de l'hôpital encaisse ce qui a été payé en espèces.
 - e. La secrétaire de l'hôpital demande ce qui a été payé en espèces.
- (5)
- a. Le gendarme du village visionne ce que son collègue filme en son absence.
 - b. Le gendarme du village ignore ce que son collègue filme en son absence.
 - c. Le gendarme du village ignore quand son collègue filme en son absence.
 - d. Le gendarme du village visionne ce qui a été filmé en son absence.
 - e. Le gendarme du village ignore ce qui a été filmé en son absence.
- (6)
- a. Le Maire du village enregistre ce que son adjoint écrit sur l'ordinateur.
 - b. Le Maire du village ignore ce que son adjoint écrit sur l'ordinateur.
 - c. Le Maire du village ignore quand son adjoint écrit sur l'ordinateur.
 - d. Le Maire du village enregistre ce qui a été écrit sur l'ordinateur.
 - e. Le Maire du village ignore ce qui a été écrit sur l'ordinateur.
- (7)
- a. Le moniteur de l'auto-école évalue ce que l'apprenti dépasse dans un virage.
 - b. Le moniteur de l'auto-école demande ce que l'apprenti dépasse dans un virage.
 - c. Le moniteur de l'auto-école demande quand l'apprenti dépasse dans un virage.
 - d. Le moniteur de l'auto-école évalue ce qui a été dépassé dans un virage.
 - e. Le moniteur de l'auto-école demande ce qui a été dépassé dans un virage.
- (8)
- a. Le chef de l'entreprise adapte ce que le comptable calcule sur le bilan annuel.
 - b. Le chef de l'entreprise devine ce que le comptable calcule sur le bilan annuel.
 - c. Le chef de l'entreprise devine quand le comptable calcule sur le bilan annuel.
 - d. Le chef de l'entreprise adapte ce qui a été calculé sur le bilan annuel.
 - e. Le chef de l'entreprise devine ce qui a été calculé sur le bilan annuel.
- (9)
- a. Le frère du buraliste affiche ce que le joueur gagne au grand loto.

- b. Le frère du buraliste ignore ce que le joueur gagne au grand loto.
 - c. Le frère du buraliste ignore quand le joueur gagne au grand loto.
 - d. Le frère du buraliste affiche ce qui a été gagné au grand loto.
 - e. Le frère du buraliste ignore ce qui a été gagné au grand loto.
- (10)
- a. La femme du traiteur déguste ce que son mari prépare pour le matin.
 - b. La femme du traiteur devine ce que son mari prépare pour le matin.
 - c. La femme du traiteur devine quand son mari prépare pour le matin.
 - d. La femme du traiteur déguste ce qui a été préparé pour le matin.
 - e. La femme du traiteur devine ce qui a été préparé pour le matin.
- (11)
- a. Le cadet des scouts goûte ce que l'aîné boit pendant le dîner.
 - b. Le cadet des scouts ignore ce que l'aîné boit pendant le dîner.
 - c. Le cadet des scouts ignore quand l'aîné boit pendant le dîner.
 - d. Le cadet des scouts goûte ce qui a été bu pendant le dîner.
 - e. Le cadet des scouts ignore ce qui a été bu pendant le dîner.
- (12)
- a. L'ouvrier de l'usine conçoit ce que la SNCF commande pour les rails.
 - b. L'ouvrier de l'usine demande ce que la SNCF commande pour les rails.
 - c. L'ouvrier de l'usine demande quand la SNCF commande pour les rails.
 - d. L'ouvrier de l'usine conçoit ce qui a été commandé pour les rails.
 - e. L'ouvrier de l'usine demande ce qui a été commandé pour les rails.
- (13)
- a. Le responsable de l'infirmierie fixe ce que la secrétaire réclame au service livraison.
 - b. Le responsable de l'infirmierie ignore ce que la secrétaire réclame au service livraison.
 - c. Le responsable de l'infirmierie ignore quand la secrétaire réclame au service livraison.
 - d. Le responsable de l'infirmierie fixe ce qui a été réclamé au service livraison.
 - e. Le responsable de l'infirmierie ignore ce qui a été réclamé au service livraison.
- (14)
- a. Le fautif de l'accident rembourse ce que la victime perd en constat amiable.
 - b. Le fautif de l'accident se demande ce que la victime perd en constat amiable.
 - c. Le fautif de l'accident se demande quand la victime perd en constat amiable.
 - d. Le fautif de l'accident rembourse ce qui a été perdu en constat amiable.
 - e. Le fautif de l'accident se demande ce qui a été perdu en constat amiable.
- (15)
- a. La vendeuse du kiosque méprise ce que le client achète dans la journée.
 - b. La vendeuse du kiosque devine ce que le client achète dans la journée.
 - c. La vendeuse du kiosque devine quand le client achète dans la journée.

- d. La vendeuse du kiosque méprise ce qui a été acheté dans la journée.
 - e. La vendeuse du kiosque devine ce qui a été acheté dans la journée.
- (16)
- a. L'assistant du cabinet liste ce que le dentiste oublie lors du paiement.
 - b. L'assistant du cabinet ignore ce que le dentiste oublie lors du paiement.
 - c. L'assistant du cabinet ignore quand le dentiste oublie lors du paiement.
 - d. L'assistant du cabinet liste ce qui a été oublié lors du paiement.
 - e. L'assistant du cabinet ignore ce qui a été oublié lors du paiement.
- (17)
- a. Le gérant de l'usine approuve ce que son employé produit dans la semaine.
 - b. Le gérant de l'usine se demande ce que son employé produit dans la semaine.
 - c. Le gérant de l'usine se demande quand son employé produit dans la semaine.
 - d. Le gérant de l'usine approuve ce qui a été produit dans la semaine.
 - e. Le gérant de l'usine se demande ce qui a été produit dans la semaine.
- (18)
- a. Le patient de l'hospice mémorise ce que l'aide-soignant lit pour le divertir.
 - b. Le patient de l'hospice redemande ce que l'aide-soignant lit pour le divertir.
 - c. Le patient de l'hospice redemande quand l'aide-soignant lit pour le divertir.
 - d. Le patient de l'hospice mémorise ce qui a été lu pour le divertir.
 - e. Le patient de l'hospice redemande ce qui a été lu pour le divertir.
- (19)
- a. L'ouvrier de l'entreprise déteste ce que son collègue mange pour le déjeuner.
 - b. L'ouvrier de l'entreprise ignore ce que son collègue mange pour le déjeuner.
 - c. L'ouvrier de l'entreprise ignore quand son collègue mange pour le déjeuner.
 - d. L'ouvrier de l'entreprise déteste ce qui a été mangé pour le déjeuner.
 - e. L'ouvrier de l'entreprise ignore ce qui a été mangé pour le déjeuner.
- (20)
- a. Le père de l'élève écoute ce que le professeur discute pendant la réunion.
 - b. Le père de l'élève devine ce que le professeur discute pendant la réunion.
 - c. Le père de l'élève devine quand le professeur discute pendant la réunion.
 - d. Le père de l'élève écoute ce qui a été discuté pendant la réunion.
 - e. Le père de l'élève devine ce qui a été discuté pendant la réunion.
- (21)
- a. La mère de l'enfant conteste ce que le gendarme sanctionne dans le parking.
 - b. La mère de l'enfant demande ce que le gendarme sanctionne dans le parking.
 - c. La mère de l'enfant demande quand le gendarme sanctionne dans le parking.
 - d. La mère de l'enfant conteste ce qui a été sanctionné dans le parking.
 - e. La mère de l'enfant demande ce qui a été sanctionné dans le parking.
- (22)
- a. Le directeur de thèse relit ce que son doctorant rédige pendant la semaine.
 - b. Le directeur de thèse ignore ce que son doctorant rédige pendant la semaine.

- c. Le directeur de thèse ignore quand son doctorant rédige pendant la semaine.
 - d. Le directeur de thèse relit ce qui a été rédigé pendant la semaine.
 - e. Le directeur de thèse ignore ce qui a été rédigé pendant la semaine.
- (23)
- a. Le meneur de l'expédition photographie ce que l'aventurier explore dans la grotte.
 - b. Le meneur de l'expédition redemande ce que l'aventurier explore dans la grotte.
 - c. Le meneur de l'expédition redemande quand l'aventurier explore dans la grotte.
 - d. Le meneur de l'expédition photographie ce qui a été exploré dans la grotte.
 - e. Le meneur de l'expédition redemande ce qui a été exploré dans la grotte.
- (24)
- a. Le standardiste de l'hôtel évalue ce que son collègue répond par téléphone.
 - b. Le standardiste de l'hôtel ignore ce que son collègue répond par téléphone.
 - c. Le standardiste de l'hôtel ignore quand son collègue répond par téléphone.
 - d. Le standardiste de l'hôtel évalue ce qui a été répondu par téléphone.
 - e. Le standardiste de l'hôtel ignore ce qui a été répondu par téléphone.
- (25)
- a. Le locataire de l'appartement aperçoit ce que son enfant fouille dans le jardin.
 - b. Le locataire de l'appartement devine ce que son enfant fouille dans le jardin.
 - c. Le locataire de l'appartement devine quand son enfant fouille dans le jardin.
 - d. Le locataire de l'appartement aperçoit ce qui a été fouillé dans le jardin.
 - e. Le locataire de l'appartement devine ce qui a été fouillé dans le jardin.
- (26)
- a. L'élève de primaire note ce que le maître écrit au tableau.
 - b. L'élève de primaire devine ce que le maître écrit au tableau.
 - c. L'élève de primaire devine quand le maître écrit au tableau.
 - d. L'élève de primaire note ce qui a été écrit au tableau.
 - e. L'élève de primaire devine ce qui a été écrit au tableau.
- (27)
- a. Le conducteur du véhicule appréhende ce que le radar calcule pendant le flash.
 - b. Le conducteur du véhicule se demande ce que le radar calcule pendant le flash.
 - c. Le conducteur du véhicule se demande quand le radar calcule pendant le flash.
 - d. Le conducteur du véhicule appréhende ce qui a été calculé pendant le flash.
 - e. Le conducteur du véhicule se demande ce qui a été calculé pendant le flash.
- (28)
- a. Le gardien de but approuve ce que l'attaquant prépare pour le match.
 - b. Le gardien de but demande ce que l'attaquant prépare pour le match.
 - c. Le gardien de but demande quand l'attaquant prépare pour le match.
 - d. Le gardien de but approuve ce qui a été préparé pour le match.
 - e. Le gardien de but demande ce qui a été préparé pour le match.

- (29) a. Le vendeur de portables sauvegarde ce que l'acheteur réclame pour la commande.
b. Le vendeur de portables redemande ce que l'acheteur réclame pour la commande.
c. Le vendeur de portables redemande quand l'acheteur réclame pour la commande.
d. Le vendeur de portables sauvegarde ce qui a été réclamé pour la commande.
e. Le vendeur de portables redemande ce qui a été réclamé pour la commande.
- (30) a. Le gardien du gîte dédommage ce que le touriste perd lors du nettoyage.
b. Le gardien du gîte redemande ce que le touriste perd lors du nettoyage.
c. Le gardien du gîte redemande quand le touriste perd lors du nettoyage.
d. Le gardien du gîte dédommage ce qui a été perdu lors du nettoyage.
e. Le gardien du gîte redemande ce qui a été perdu lors du nettoyage.
- (31) a. Le visiteur du musée réarrange ce que le vigile fouille à l'entrée.
b. Le visiteur du musée demande ce que le vigile fouille à l'entrée.
c. Le visiteur du musée demande quand le vigile fouille à l'entrée.
d. Le visiteur du musée réarrange ce qui a été fouillé à l'entrée.
e. Le visiteur du musée demande ce qui a été fouillé à l'entrée.
- (32) a. Le spectateur du cinéma critique ce que son enfant mange pendant le film.
b. Le spectateur du cinéma demande ce que son enfant mange pendant le film.
c. Le spectateur du cinéma demande quand son enfant mange pendant le film.
d. Le spectateur du cinéma critique ce qui a été mangé pendant le film.
e. Le spectateur du cinéma demande ce qui a été mangé pendant le film.
- (33) a. Le proviseur du collège conteste ce que l'élève discute avec le professeur.
b. Le proviseur du collège demande ce que l'élève discute avec le professeur.
c. Le proviseur du collège demande quand l'élève discute avec le professeur.
d. Le proviseur du collège conteste ce qui a été discuté avec le professeur.
e. Le proviseur du collège demande ce qui a été discuté avec le professeur.
- (34) a. Le caissier du supermarché inspecte ce que le responsable commande pour son assise.
b. Le caissier du supermarché se demande ce que le responsable commande pour son assise.
c. Le caissier du supermarché se demande quand le responsable commande pour son assise.
d. Le caissier du supermarché inspecte ce qui a été commandé pour son assise.
e. Le caissier du supermarché se demande ce qui a été commandé pour son assise.

- (35) a. Le chef deux étoiles inventorie ce que le commis cuisine pour la soirée.
b. Le chef deux étoiles redemande ce que le commis cuisine pour la soirée.
c. Le chef deux étoiles redemande quand le commis cuisine pour la soirée.
d. Le chef deux étoiles inventorie ce qui a été cuisiné pour la soirée.
e. Le chef deux étoiles redemande ce qui a été cuisiné pour la soirée.
- (36) a. Le batteur du groupe constate ce que le bassiste oublie pendant les répétitions.
b. Le batteur du groupe se demande ce que le bassiste oublie pendant les répétitions.
c. Le batteur du groupe se demande quand le bassiste oublie pendant les répétitions.
d. Le batteur du groupe constate ce qui a été oublié pendant les répétitions.
e. Le batteur du groupe se demande ce qui a été oublié pendant les répétitions.
- (37) a. La doctorante du cours remarque ce que l'étudiant trie sur son téléphone.
b. La doctorante du cours devine ce que l'étudiant trie sur son téléphone.
c. La doctorante du cours devine quand l'étudiant trie sur son téléphone.
d. La doctorante du cours remarque ce qui a été trié sur son téléphone.
e. La doctorante du cours devine ce qui a été trié sur son téléphone.
- (38) a. La dame du CDI critique ce que l'adolescent lit en ce moment.
b. La dame du CDI redemande ce que l'adolescent lit en ce moment.
c. La dame du CDI redemande quand l'adolescent lit en ce moment.
d. La dame du CDI critique ce qui a été lu en ce moment.
e. La dame du CDI redemande ce qui a été lu en ce moment.
- (39) a. Le chauffeur de bus autorise ce que l'enfant boit à bord.
b. Le chauffeur de bus redemande ce que l'enfant boit à bord.
c. Le chauffeur de bus redemande quand l'enfant boit à bord.
d. Le chauffeur de bus autorise ce qui a été bu à bord.
e. Le chauffeur de bus redemande ce qui a été bu à bord.
- (40) a. Le gérant de la salle réévalue ce que l'artiste produit en musique.
b. Le gérant de la salle se demande ce que l'artiste produit en musique.
c. Le gérant de la salle se demande quand l'artiste produit en musique.
d. Le gérant de la salle réévalue ce qui a été produit en musique.
e. Le gérant de la salle se demande ce qui a été produit en musique.
- (41) a. L'équipe de foot approuve ce que l'entraîneur gagne pour la saison.
b. L'équipe de foot demande ce que l'entraîneur gagne pour la saison.
c. L'équipe de foot demande quand l'entraîneur gagne pour la saison.

- d. L'équipe de foot approuve ce qui a été gagné pour la saison.
 - e. L'équipe de foot demande ce qui a été gagné pour la saison.
- (42)
- a. Le moniteur de l'auto-école repère ce que l'apprenti dépasse sur la route.
 - b. Le moniteur de l'auto-école devine ce que l'apprenti dépasse sur la route.
 - c. Le moniteur de l'auto-école devine quand l'apprenti dépasse sur la route.
 - d. Le moniteur de l'auto-école repère ce qui a été dépassé sur la route.
 - e. Le moniteur de l'auto-école devine ce qui a été dépassé sur la route.
- (43)
- a. La contrôleuse de gestion contredit ce que le directeur sanctionne dans l'open-space.
 - b. La contrôleuse de gestion devine ce que le directeur sanctionne dans l'open-space.
 - c. La contrôleuse de gestion devine quand le directeur sanctionne dans l'open-space.
 - d. La contrôleuse de gestion contredit ce qui a été sanctionné dans l'open-space.
 - e. La contrôleuse de gestion devine ce qui a été sanctionné dans l'open-space.
- (44)
- a. Le proviseur du collège désapprouve ce que l'équipe filme dans la cour.
 - b. Le proviseur du collège redemande ce que l'équipe filme dans la cour.
 - c. Le proviseur du collège redemande quand l'équipe filme dans la cour.
 - d. Le proviseur du collège désapprouve ce qui a été filmé dans la cour.
 - e. Le proviseur du collège redemande ce qui a été filmé dans la cour.
- (45)
- a. Le passager du vol considère ce que l'hôtesse répond à l'embarquement.
 - b. Le passager du vol se demande ce que l'hôtesse répond à l'embarquement.
 - c. Le passager du vol se demande quand l'hôtesse répond à l'embarquement.
 - d. Le passager du vol considère ce qui a été répondu à l'embarquement.
 - e. Le passager du vol se demande ce qui a été répondu à l'embarquement.
- (46)
- a. Le vendeur de la voiture approuve ce que son ami paye en liquide.
 - b. Le vendeur de la voiture demande ce que son ami paye en liquide.
 - c. Le vendeur de la voiture demande quand son ami paye en liquide.
 - d. Le vendeur de la voiture approuve ce qui a été payé en liquide.
 - e. Le vendeur de la voiture demande ce qui a été payé en liquide.
- (47)
- a. Le propriétaire du château surveille ce que son chat explore pendant la journée.
 - b. Le propriétaire du château se demande ce que son chat explore pendant la journée.
 - c. Le propriétaire du château se demande quand son chat explore pendant la journée.

- d. Le propriétaire du château surveille ce qui a été exploré pendant la journée.
 - e. Le propriétaire du château se demande ce qui a été exploré pendant la journée.
- (48)
- a. Le surveillant du contrôle examine ce que l'élève rédige sur son bloc-notes.
 - b. Le surveillant du contrôle se demande ce que l'élève rédige sur son bloc-notes.
 - c. Le surveillant du contrôle se demande quand l'élève rédige sur son bloc-notes.
 - d. Le surveillant du contrôle examine ce qui a été rédigé sur son bloc-notes.
 - e. Le surveillant du contrôle se demande ce qui a été rédigé sur son bloc-notes.
- (49)
- a. Le loueur de voitures déprécie ce que l'adulte négocie en vain.
 - b. Le loueur de voitures redemande ce que l'adulte négocie en vain.
 - c. Le loueur de voitures redemande quand l'adulte négocie en vain.
 - d. Le loueur de voitures déprécie ce qui a été négocié en vain.
 - e. Le loueur de voitures redemande ce qui a été négocié en vain.
- (50)
- a. L'adolescent de la famille préfère ce que sa mère achète au boucher.
 - b. L'adolescent de la famille redemande ce que sa mère achète au boucher.
 - c. L'adolescent de la famille redemande quand sa mère achète au boucher.
 - d. L'adolescent de la famille préfère ce qui a été acheté au boucher.
 - e. L'adolescent de la famille redemande ce qui a été acheté au boucher.

G.3 Distracteurs

1. Marie pense que Julie a mangé les chocolats.
2. Lucas croit que Sam a dormi chez une amie.
3. Claire trouve que Julien est très intelligent.
4. Juliette pense que son ex est parti à l'étranger.
5. Céline croit que Louis est devenu vegan.
6. Manon trouve que son père a pris du poids.
7. Fanny est sûre que John lui a menti.
8. Vanessa est sûre que son mari l'a trompée.
9. Timothée est certain que sa copine est enceinte.
10. Luc est certain que Léa est en vacances.
11. Amélie pense que Matthieu a bu toutes les bières.
12. Nina croit que Damien a déménagé en région parisienne.
13. Yann trouve que Pauline est très bien habillée pour l'occasion.
14. Kévin pense que son ami commet une grosse erreur.

15. Antoine trouve que son patron profite de ses employés.
16. Lou croit que sa fille a attrapé la varicelle.
17. Yasmina est sûre que son frère n'est pas allé au lycée.
18. Simon est sûr que Nicolas ira au festival de musique à Amsterdam.
19. Aurélie est certaine que sa sœur a rencontré quelqu'un récemment.
20. Samia est certaine que Yacine a mangé la dernière part de gâteau.
21. Thomas est sûr que son amie a dansé toute la soirée.
22. Rémi est sûr que Lucie a recommencé à fumer.
23. Nathalie pense que sa sœur a inventé toute l'histoire.
24. Jérôme trouve que son chat ne mange pas assez.
25. Rachel croit que son père veut demander le divorce.
26. Philippe est convaincu que son frère a tout mangé.
27. Ada pense qu'Emma devrait sortir plus souvent.
28. Maya trouve que sa sœur fait beaucoup de caprices.
29. Anne croit que Chloé est partie depuis longtemps.
30. Cécile pense que Barbara devrait appeler le médecin.
31. Christophe est sûr que Julien aura beaucoup de succès.
32. Pascal trouve que son cousin n'est pas très aimable.
33. Auguste pense qu'Héloïse écrit de très belles histoires.
34. Juliette est sûre que Félix a menti au professeur.
35. Louise est sûre que Alban deviendra un champion.
36. Jérémy pense que Lucie ne travaille pas assez.
37. Clara trouve que Stella devrait faire plus de sport.
38. Louna est certaine que son copain est au bureau.
39. Lisa pense qu'Oscar est allé au stade dimanche.
40. Lucas est convaincu que son père ne sait pas écouter.
41. Livia pense que Jeanne est la meilleure en maths.
42. Liam est certain que Tamara a rangé ses affaires.
43. Marc trouve que Zora n'aurait pas dû insister autant.
44. Noëlle croit qu'André veut préparer le déjeuner.
45. Caroline pense qu'Antoine fume beaucoup trop.
46. Christelle est sûre qu'Ismaël achètera ses colliers.
47. Lou pense que ses professeurs perdent leur temps.
48. Claire trouve que sa mère ne sort pas assez.
49. Vincent est sûr que le comité fera le bon choix.
50. Éric croit que sa mère ne sera pas là samedi prochain.

Annexe H

Analyses statistiques pour l'étude diachronique (Chapitre 3, Section 3.3.2.3.7)

```
> m1 <- lm(PresenceCE ~ Date*Verbe, data=data)
> summary(m1)

Call:
lm(formula = PresenceCE ~ Date * Verbe, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-1.05147 -0.13789  0.09044  0.19441  1.27038

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -2.0647972  0.1650971 -12.507 < 2e-16 ***
Date           0.0016475  0.0001081  15.237 < 2e-16 ***
VerbeFAIRE    2.1275407  0.1895733  11.223 < 2e-16 ***
VerbeSAVOIR   -0.6615936  0.1804959  -3.665 0.000251 ***
Date:VerbeFAIRE -0.0010659  0.0001241  -8.587 < 2e-16 ***
Date:VerbeSAVOIR 0.0004789  0.0001180   4.059 5.05e-05 ***
---

```


Annexe I

**Analyses statistiques pour le premier test
de répétition en français, avec ‘ce qui/ce
que’**

(Chapitre 4, Section 4.4.3)

I.1 Performance générale

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [glmerMod]
Family: binomial ( logit )
Formula: Score ~ sentence_type * gap_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: data2
Control: glmerControl(optimizer = "bobyqa")

      AIC      BIC  logLik deviance df.resid
584.2   630.6  -282.1   564.2       751

Scaled residuals:
   Min       1Q   Median       3Q      Max
-4.0458 -0.2760 -0.0597  0.2691  5.4502

Random effects:
 Groups Name  Variance Std.Dev.
 Subj (Intercept) 10.362  3.2191
 Item (Intercept)  0.155  0.3937
Number of obs: 761, groups: Subj, 49; Item, 16

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -0.5612    0.5275  -1.064  0.28735
sentence_type  -0.8659    0.3162  -2.738  0.00618 **
gap_type       0.1473    0.3112   0.473  0.63594
Age_Centered   2.4564    0.5961   4.120  3.78e-05 ***
sentence_type:gap_type -0.4996    0.6218  -0.803  0.42174
sentence_type:Age_Centered 0.2078    0.3246   0.640  0.52205
gap_type:Age_Centered  0.4018    0.3180   1.263  0.20645
sentence_type:gap_type:Age_Centered -0.0079    0.6291  -0.013  0.98998

```

I.2 Types d'erreur

I.2.1 Omission de 'ce'

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial ( logit )
Formula: omission ~ sentence_type * gap_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: data2
Control: glmerControl(optimizer = "bobyqa")

      AIC      BIC logLik deviance df.resid
573.6  619.9  -276.8  553.6      751

Scaled residuals:
   Min      1Q  Median      3Q      Max
-2.9024 -0.2908 -0.1552 -0.0426  5.2650

Random effects:
 Groups Name      Variance Std.Dev.
 Subj (Intercept) 6.461   2.5418
 Item (Intercept) 0.362   0.6016
Number of obs: 761, groups: Subj, 49; Item, 16

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -2.39541    0.46861  -5.112 3.19e-07 ***
sentence_type  0.65447    0.39277   1.666  0.0957 .
gap_type       0.62095    0.39384   1.577  0.1149
Age_Centered  -0.57112    0.42565  -1.342  0.1797
sentence_type:gap_type -0.43678    0.78094  -0.559  0.5760
sentence_type:Age_Centered 0.07371    0.27973   0.263  0.7922
gap_type:Age_Centered  0.43892    0.28437   1.543  0.1227
sentence_type:gap_type:Age_Centered 0.06161    0.55343   0.111  0.9114
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

I.2.2 Emploi de *kesk*

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [glmerMod]
Family: binomial ( logit )
Formula: KESK ~ sentence_type * gap_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: data2
Control: glmerControl(optimizer = "bobyqa")

      AIC      BIC    loglik deviance df.resid
332.4   378.8   -156.2   312.4         751

Scaled residuals:
      Min       1Q   Median       3Q      Max
-3.0289 -0.0772 -0.0454 -0.0347  4.7360

Random effects:
 Groups Name      Variance Std.Dev.
 Subj (Intercept) 14.91   3.861
 Item (Intercept)  0.00   0.000
Number of obs: 761, groups: Subj, 49; Item, 16

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -5.73502    1.44260   -3.975 7.02e-05 ***
sentence_type  0.64228    0.35129    1.828  0.0675 .
gap_type      -0.27744    0.34743   -0.799  0.4245
Age_Centered  0.45689    0.64358    0.710  0.4778
sentence_type:gap_type  0.07119    0.69301    0.103  0.9182
sentence_type:Age_Centered  0.19803    0.45857    0.432  0.6659
gap_type:Age_Centered  -0.27337    0.45870   -0.596  0.5512
sentence_type:gap_type:Age_Centered  0.14402    0.91205    0.158  0.8745

```

I.2.3 Stratégie SVO (phrase simple)

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [glmerMod]
Family: binomial ( logit )
Formula: SVO ~ sentence_type * gap_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: data2
Control: glmerControl(optimizer = "bobyqa")

      AIC      BIC logLik deviance df.resid
125.8    172.2   -52.9   105.8         751

Scaled residuals:
   Min       1Q   Median       3Q      Max
-1.9673 -0.0139 -0.0029 -0.0001  3.3045

Random effects:
 Groups Name      Variance Std.Dev.
 Subj (Intercept) 19.82916  4.4530
 Item (Intercept)  0.01816  0.1348
Number of obs: 761, groups: Subj, 49; Item, 16

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -13.1217    3.4992  -3.750 0.000177 ***
sentence_type    -1.3880    2.7923  -0.497 0.619141
gap_type        -6.5255    3.4197  -1.908 0.056367 .
Age_Centered    -5.7690    2.0169  -2.860 0.004232 **
sentence_type:gap_type  4.7065    5.7024  0.825 0.409175
sentence_type:Age_Centered -0.3437    1.9827  -0.173 0.862372
gap_type:Age_Centered  -5.7532    2.4912  -2.309 0.020921 *
sentence_type:gap_type:Age_Centered  3.3459    4.0154  0.833 0.404696

```

I.2.4 Emploi de 'ce que' ou 'ce qu'il' à la place de 'ce qui' dans les dépendances sujet

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [glmerMod]
Family: binomial ( logit )
Formula: ske_skil ~ sentence_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: subset(data2, GapType != "OBJ")
Control: glmerControl(optimizer = "bobyqa")

AIC      BIC      logLik deviance df.resid
196.9    220.6    -92.5    184.9    378

Scaled residuals:
   Min       1Q   Median       3Q      Max
-1.2076 -0.1748 -0.1228 -0.0991  3.4348

Random effects:
 Groups Name      Variance Std.Dev.
 Subj  (Intercept)  4.0951   2.0236
 Item  (Intercept)  0.2403   0.4902
Number of obs: 384, groups:  Subj, 49; Item, 8

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    -3.9853      0.7359  -5.416 6.1e-08 ***
sentence_type  -0.2796      0.5757  -0.486  0.627
Age_Centered    0.4602      0.4436  1.038  0.299
sentence_type:Age_Centered  0.3016      0.5259  0.574  0.566

```

Annexe J

**Analyses statistiques pour le deuxième
test de répétition en français, avec ‘qui’
(Chapitre 4, Section 4.7.3)**

J.1 Performance générale

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [‘glmerMod’]
Family: binomial ( logit )
Formula: Score ~ sentence_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: data
Control: glmerControl(optimizer = "bobyqa")

      AIC      BIC logLik deviance df.resid
670.2  698.8  -329.1  658.2      874

Scaled residuals:
   Min       1Q   Median       3Q      Max
-7.4527 -0.1819  0.1524  0.3424  3.9649

Random effects:
 Groups Name      Variance Std.Dev.
 Subj (Intercept) 7.92721  2.8155
 Item (Intercept) 0.08814  0.2969
Number of obs: 880, groups:  Subj, 55; Item, 16

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    1.4004    0.4335   3.231 0.001234 **
sentence_type    0.6730    0.2758   2.440 0.014701 *
Age_Centered    1.4038    0.4297   3.267 0.001086 **
sentence_type:Age_Centered  0.8693    0.2251   3.861 0.000113 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

```

J.2 Types d'erreur

J.2.1 Élimination de l'enchâssement

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [ 'glmerMod' ]
Family: binomial ( logit )
Formula: elimination_enchassement ~ sentence_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: data
Control: glmerControl(optimizer = "bobyqa")

      AIC      BIC logLik deviance df.resid
337.1  365.8  -162.6   325.1     874

Scaled residuals:
   Min      1Q  Median      3Q      Max
-3.2872 -0.1248 -0.0441 -0.0233  5.4091

Random effects:
 Groups Name      Variance Std.Dev.
 Subj (Intercept)  8.5992   2.9324
 Item (Intercept)  0.4486   0.6698
Number of obs: 880, groups: Subj, 55; Item, 16

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -5.49219    0.87399  -6.284  3.3e-10 ***
sentence_type -0.02271    0.57629  -0.039  0.968560
Age_Centered -1.97239    0.58273  -3.385  0.000712 ***
sentence_type:Age_Centered  0.59176    0.41417   1.429  0.153063
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

J.2.2 Interrogatives clivées

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial ( logit )
Formula: int_clivees ~ sentence_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: data
Control: glmerControl(optimizer = "bobyqa")

AIC      BIC      loglik deviance df.resid
249.2    277.9    -118.6    237.2    874

Scaled residuals:
   Min      1Q  Median      3Q      Max
-3.5220 -0.0837 -0.0315 -0.0137 10.0178

Random effects:
 Groups Name      Variance Std.Dev.
 Subj (Intercept) 13.16    3.627
 Item (Intercept) 0.00    0.000
Number of obs: 880, groups: Subj, 55; Item, 16

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -6.76916    1.24894  -5.420 5.96e-08 ***
sentence_type  2.04239    0.55079   3.708 0.000209 ***
Age_Centered  -1.39824    0.67011  -2.087 0.036927 *
sentence_type:Age_Centered 0.07231    0.47214   0.153 0.878270
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

J.2.3 Relativisation explicite

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [‘glmerMod’]
Family: binomial ( logit )
Formula: explicitation_rel ~ sentence_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: data
Control: glmerControl(optimizer = "bobyqa")

      AIC      BIC loglik deviance df.resid
240.2  268.9  -114.1  228.2      874

Scaled residuals:
  Min       1Q   Median       3Q      Max
-1.9258 -0.0581 -0.0277 -0.0068  11.3583

Random effects:
 Groups Name      Variance Std.Dev.
 Subj (Intercept) 17.35   4.166
 Item (Intercept)  0.00   0.000
Number of obs: 880, groups:  Subj, 55; Item, 16

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    -7.8073    1.6199  -4.820 1.44e-06 ***
sentence_type  -3.9009    0.7892  -4.943 7.70e-07 ***
Age_Centered    0.7339    0.7329    1.001  0.317
sentence_type:Age_Centered -0.5590    0.6805  -0.822  0.411
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

```

J.2.4 Autres subordonnées

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial ( logit )
Formula: adjonctivisation ~ sentence_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: data
Control: glmerControl(optimizer = "bobyqa")

AIC      BIC      logLik deviance df.resid
220.9    249.6    -104.5    208.9     874

Scaled residuals:
   Min       1Q   Median       3Q      Max
-1.2434 -0.0729 -0.0469 -0.0184  6.1680

Random effects:
 Groups Name      Variance Std.Dev.
 Subj   (Intercept) 8.6019   2.9329
 Item   (Intercept) 0.2678   0.5175
Number of obs: 880, groups: Subj, 55; Item, 16

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -6.5149    1.2828  -5.078 3.8e-07 ***
sentence_type  -2.0235    0.6879  -2.942 0.00327 **
Age_Centered  -1.0200    0.5621  -1.815 0.06959 .
sentence_type:Age_Centered -0.7915    0.5152  -1.536 0.12444
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

J.2.5 Omission du mot wh- 'qui'

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [ 'glmerMod' ]
Family: binomial ( logit )
Formula: OmissionQUI ~ sentence_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: data
Control: glmerControl(optimizer = "bobyqa")

      AIC      BIC    logLik deviance df.resid
178.5   207.2   -83.3   166.5      874

Scaled residuals:
   Min       1Q   Median       3Q      Max
-1.2575 -0.0569 -0.0241 -0.0131  5.0665

Random effects:
 Groups Name  Variance Std.Dev.
 Subj (Intercept) 9.595   3.098
 Item (Intercept) 0.000   0.000
Number of obs: 880, groups: Subj, 55; Item, 16

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    -7.0753     1.4974  -4.725  2.3e-06 ***
sentence_type    -0.5441     0.8440  -0.645  0.5192
Age_Centered    -1.6926     0.7541  -2.245  0.0248 *
sentence_type:Age_Centered -0.6509     0.6988  -0.931  0.3516
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```


Annexe K

**Analyses statistiques pour le premier test
de répétition en portugais, avec ‘o que’
(Chapitre 5, Section 5.3.3)**

K.1 Performance générale

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [ 'glmerMod' ]
Family: binomial ( logit )
Formula: Score ~ sentence_type * gap_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: data1
Control: glmerControl(optimizer = "bobyqa")

      AIC      BIC logLik deviance df.resid
513.2   560.4  -246.6   493.2      822

Scaled residuals:
      Min       1Q   Median       3Q      Max
-4.8072  0.0422  0.1236  0.2877  3.4819

Random effects:
 Groups Name      Variance Std.Dev.
 Subj (Intercept)  9.631    3.103
 Item (Intercept)  0.000    0.000
Number of obs: 832, groups: Subj, 52; Item, 16

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)      2.7208    0.5373   5.064  4.1e-07 ***
sentence_type    -1.0151    0.2830  -3.586  0.000335 ***
gap_type         -0.7036    0.2783  -2.528  0.011465 *
Age_Centered     1.3954    0.5115   2.728  0.006373 **
sentence_type:gap_type  0.3924    0.5498   0.714  0.475431
sentence_type:Age_Centered  0.0105    0.2908   0.036  0.971194
gap_type:Age_Centered -0.1396    0.2850  -0.490  0.624367
sentence_type:gap_type:Age_Centered -0.2562    0.5662  -0.453  0.650903
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

K.2 Types d'erreur

K.2.1 Omission de 'o'

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [‘glmerMod’]
Family: binomial ( logit )
Formula: omission ~ sentence_type * gap_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: data2
Control: glmerControl(optimizer = "bobyqa")

      AIC      BIC logLik deviance df.resid
318.6  365.8  -149.3  298.6      816

Scaled residuals:
   Min      1Q  Median      3Q      Max
-7.3838 -0.0531 -0.0224 -0.0095  5.3992

Random effects:
 Groups Name  Variance Std.Dev.
 Subj (Intercept) 30.09194  5.4856
 Item (Intercept)  0.02608  0.1615
Number of obs: 826, groups:  Subj, 52; Item, 16

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -7.07503    1.64838  -4.292 1.77e-05 ***
sentence_type    0.48848    0.41229    1.185  0.236
gap_type        1.94997    0.45795    4.258 2.06e-05 ***
Age_Centered   -1.35727    0.91394   -1.485  0.138
sentence_type:gap_type  0.11559    0.82165    0.141  0.888
sentence_type:Age_Centered -0.08728    0.42849   -0.204  0.839
gap_type:Age_Centered  0.06346    0.47601    0.133  0.894
sentence_type:gap_type:Age_Centered -0.87552    0.85801   -1.020  0.308
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

```

K.2.2 Emploi de 'ó que è que'

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial ( logit )
Formula: OqueQue ~ sentence_type * gap_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: data2
Control: glmerControl(optimizer = "bobyqa")

      AIC      BIC logLik deviance df.resid
136.9  184.1  -58.5  116.9      816

Scaled residuals:
      Min       1Q   Median       3Q      Max
-1.6613 -0.0183 -0.0049  0.0000  3.3983

Random effects:
 Groups Name  Variance Std.Dev.
 Subj  (Intercept) 35.87229 5.9893
 Item  (Intercept)  0.07764 0.2786
Number of obs: 826, groups:  Subj, 52; Item, 16

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -14.2441    21.7088  -0.656   0.512
sentence_type    11.7400    43.2395   0.272   0.786
gap_type       -1.7372    42.1824  -0.041   0.967
Age_Centered   -2.8566    13.0348  -0.219   0.827
sentence_type:gap_type    1.6188    84.3538   0.019   0.985
sentence_type:Age_Centered  7.1011    26.0629   0.272   0.785
gap_type:Age_Centered    0.5327    25.3970   0.021   0.983
sentence_type:gap_type:Age_Centered  1.0607    50.7945   0.021   0.983

```

Annexe L

**Analyses statistiques pour le deuxième
test de répétition en portugais, avec
'quem'
(Chapitre 5, Section 5.6.3)**

L.1 Performance générale

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial ( logit )
Formula: Score ~ sentence_type * Age_Centered + (1 | Subj) + (1 | Item)
Data: data1
Control: glmerControl(optimizer = "bobyqa")

      AIC      BIC    loglik deviance df.resid
318.7  346.0  -153.4   306.7      698

Scaled residuals:
  Min       1Q   Median       3Q      Max
-5.6232  0.1067  0.1439  0.2329  1.0955

Random effects:
 Groups Name      Variance Std.Dev.
 Subj (Intercept) 1.9031  1.3795
 Item (Intercept) 0.2079  0.4559
Number of obs: 704, groups: Subj, 44; Item, 16

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  3.60059    0.42441   8.484 <2e-16 ***
sentence_type  0.09513    0.43592   0.218  0.8273
Age_Centered  0.74075    0.29513   2.510  0.0121 *
sentence_type:Age_Centered -0.28605    0.31930  -0.896  0.3703
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Annexe M

**Analyses statistiques pour le test de
jugement d'acceptabilité sur l'ambiguïté
complétive/relative
(Chapitre 6, Section 6.1.3.1.3)**

M.1 Modèle général

M.1.1 Avec la condition a (COMPL) comme niveau de comparaison et avec les sujets comme variables aléatoires

```

> ord.french2 <- clmm(Judgment.f ~ Condition1 +(1|SubjIP)+(0+Condition2|SubjIP)+(1|Item) , data=french[french$Item!="1",])
> summary(Ord.french2)
Cumulative Link Mixed Model fitted with the Laplace approximation

formula: Judgment.f ~ Condition1 + (1 | SubjIP) + (0 + Condition2 | SubjIP) + (1 | Item)
data:    french[french$Item != "1", ]

link threshold nobs logLik AIC niter max.grad cond.H
logit flexible 1610 -3325.90 6691.80 4113(12555) 5.45e-03 5.7e+04

Random effects:
Groups Name      Variance Std.Dev. Corr
SubjIP Condition2exp-b 5.229e-01 0.723085
Condition2exp-a 8.742e-01 0.934987 0.897
Condition2exp-c 8.247e-01 0.908128 0.978 0.969
SubjIP (Intercept) 3.916e-06 0.001979
Item (Intercept) 2.401e-01 0.489972
Number of groups: SubjIP 70, Item 23

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
Condition1exp-b  -0.3384    0.1236  -2.738  0.00618 **
Condition1exp-c   0.8592    0.1179   7.287  3.17e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

M.1.2 Avec la condition b (REL.AMB) comme niveau de comparaison et avec les sujets comme variables aléatoires

```
ord.french1c <- clmm(Judgment.f ~ Condition2 +(1|SubjIP)+(0+Condition2|SubjIP)+(1|Item) ,data=french[french$Item!="1",])
summary(ord.french1c)
Cumulative Link Mixed Model fitted with the Laplace approximation

formula: Judgment.f ~ Condition2 + (1 | SubjIP) + (0 + Condition2 | SubjIP) + (1 | Item)
data: french[french$Item != "1", ]

link threshold nobs logLik AIC niter max.grad cond.H
logit flexible 1610 -3325.90 6691.80 4002(12244) 8.86e-03 7.6e+02

Random effects:
Groups Name Variance Std.Dev. Corr
SubjIP Condition2exp-b 0.5229 0.7231
Condition2exp-a 0.8743 0.9350 0.897
Condition2exp-c 0.8248 0.9082 0.978 0.969
SubjIP (Intercept) 0.0000 0.0000
Item (Intercept) 0.2401 0.4900
Number of groups: SubjIP 70, Item 23

Coefficients:
Estimate Std. Error z value Pr(>|z|)
Condition2exp-a 0.3384 0.1236 2.738 0.00618 **
Condition2exp-c 1.1976 0.1188 10.079 < 2e-16 ***
```

M.1.3 Avec la condition b (REL.AMB) comme niveau de comparaison et avec les items comme variables aléatoires

```

> ord.french1d <- clmm(Judgment.f ~ Condition2 +(1|Item)+(0+Condition2|Item), data=french[french$Item!="1",])
> summary(ord.french1d)
Cumulative Link Mixed Model fitted with the Laplace approximation

formula: Judgment.f ~ Condition2 + (1 | Item) + (0 + Condition2 | Item)
data:    french[french$Item != "1", ]

link threshold nobs logLik AIC niter max.grad cond.H
logit flexible 1610 -3245.64 6529.27 3092(12352) 5.21e-03 6.7e+05

Random effects:
Groups Name          Variance Std.Dev. Corr
Item (Intercept)    0.03627 0.1904
Item Condition2exp-b 1.40665 1.1860
Item Condition2exp-a 1.66702 1.2911 -0.784
Item Condition2exp-c 0.47237 0.6873 0.344 0.028
Number of groups: Item 23

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
Condition2exp-a  0.4100    0.5005  0.819  0.413
Condition2exp-c  1.2526    0.2656  4.715 2.41e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

M.2 Conditions COMPL et REL.AMB seulement

M.2.1 Effet de fréquence

```

> ord2 <- clmm(Judgment.f ~ Condition*Frequency + (1|Subj)+(0+Condition*Frequency|Subj)+(1|Item)+(0+Condition*Frequency|Item), data=
a=resultsmergedbis)
> summary(ord2)
Cumulative Link Mixed Model fitted with the Laplace approximation

formula: Judgment.f ~ Condition * Frequency + (1 | Subj) + (0 + Condition *
Frequency | Subj) + (1 | Item) + (0 + Condition * Frequency | Item)
data:      resultsmergedbis

link threshold nobis logLik      AIC      niter      max.grad cond.H
logit flexible 1071 -2186.02 4426.04 4961(34742) 8.40e+01 1.3e+06

Random effects:
Groups Name      Variance Std.Dev. Corr
Subj Condition  1.142e+00 1.068711
      Frequency  5.870e-05 0.007661  0.548
Condition:Frequency 1.321e-04 0.011493 -0.670 -0.988
Subj (Intercept)  8.865e-01 0.941519
Item Condition   1.566e+00 1.251292
      Frequency  8.163e-06 0.002857 -0.960
Condition:Frequency 1.854e-04 0.013614  0.988 -0.992
Item (Intercept)  3.241e-01 0.569286
Number of groups:  Subj 70,  Item 23

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
Condition    0.190644    0.482505  0.395 0.692760
Frequency    0.020056    0.006060  3.309 0.000935 ***
Condition:Frequency -0.024503  0.009501 -2.579 0.009912 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

M.2.2 Effet du groupe d'âge

```

> ord6 <- c1mm(Judgment.f ~ Condition*Group + (1|Subj)+(0+Condition|Subj)+(1|Item), data=resultsmerged[resultsmerged$Item!="1",])
> summary(ord6)
Cumulative Link Mixed Model fitted with the Laplace approximation

formula: Judgment.f ~ Condition * Group + (1 | Subj) + (0 + Condition | Subj) + (1 | Item)
data: resultsmerged[resultsmerged$Item != "1", ]

link threshold nobs logLik AIC niter max.grad cond.H
logit flexible 1071 -2361.02 4758.04 3105(9396) 2.04e-03 7.3e+02

Random effects:
Groups Name Variance Std.Dev. Corr
Subj ConditionCOMPL 6.533e-01 8.083e-01
ConditionREL_AMB 4.646e-01 6.816e-01 1.000
Subj (Intercept) 2.755e-09 5.249e-05
Item (Intercept) 1.602e-01 4.003e-01
Number of groups: Subj 70, Item 23

Coefficients:
ConditionREL_AMB Estimate Std. Error z value Pr(>|z|)
Groupplusde35ans -0.1275 0.1183 -1.077 0.2815
ConditionREL_AMB:Groupplusde35ans 0.9499 0.3719 2.554 0.0107 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

M.2.2.1 Condition COMPL *versus* condition REL.AMB chez les moins de 35 ans

```

> ord9 <- clmm(Judgment.f ~ Condition + (1|Subj)+(0+Condition|Subj)+(1|Item), data=moinsde35ans)
> summary(ord9)
Cumulative Link Mixed Model fitted with the Laplace approximation

formula: Judgment.f ~ Condition + (1 | Subj) + (0 + Condition | Subj) + (1 | Item)
data:      moinsde35ans

link threshold nobs logLik  AIC   niter   max.grad cond.H
logit flexible  934  -2071.84 4175.67 2758(8346) 2.59e-03 5.6e+02

Random effects:
Groups Name      Variance Std.Dev.  Corr
Subj      ConditionCOMPL  7.077e-01 0.8412323
          ConditionREL_AMB 5.712e-01 0.7557908 1.000
Subj (Intercept)  1.152e-07 0.0003395
Item (Intercept)  1.414e-01 0.3760317
Number of groups:  Subj 61, Item 23

Coefficients:
          Estimate Std. Error z value Pr(>|z|)
ConditionREL_AMB  -0.1277    0.1180  -1.082   0.279

```

M.2.2.2 Condition COMPL *versus* condition REL_AMB chez les plus de 35 ans

```

> ord10 <- clmm(Judgment.f ~ Condition + (1|Subj)+(0+Condition|Subj)+(1|Item), data=plusde35ans)
> summary(ord10)
Cumulative Link Mixed Model fitted with the Laplace approximation

formula: Judgment.f ~ Condition + (1 | Subj) + (0 + Condition | Subj) + (1 | Item)
data: plusde35ans

link threshold nobs logLik AIC niter max.grad cond.H
logit flexible 137 -281.13 594.26 2110(4526) 6.66e-04 5.5e+02

Random effects:
Groups Name Variance Std.Dev. Corr
Item (Intercept) 0.10546 0.3247
Subj ConditionCOMPL 0.50722 0.7122
ConditionREL_AMB 0.06761 0.2600 -1.000
Subj (Intercept) 0.00000 0.0000
Number of groups: Item 23, Subj 9

Coefficients:
ConditionREL_AMB -1.6580 0.4702 -3.526 0.000422 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Annexe N

**Analyses statistiques pour la première
expérience de lecture par présentation
auto-segmentée, sur l'ambiguïté**

COMPL/REL

(Chapitre 6, Section 6.1.3.2.2)

N.1 Au niveau du point de désambiguïisation (codé 0)

N.1.1 Avec la Condition a (COMPL) comme niveau de comparaison

```

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(TempsLecture) ~ Condition + WordLength + (Condition | Sujet) + (Condition | Item)
Data: dis.word

REML criterion at convergence: 450.2

Scaled residuals:
   Min       1Q   Median       3Q      Max
-1.9919 -0.6000 -0.1651  0.4267  6.2675

Random effects:
 Groups Name      Variance Std.Dev. Corr
Sujet  (Intercept)  0.025502  0.15969
       Conditionb 0.015203  0.12330   -0.22
       Conditionc 0.015438  0.12425   -0.14  0.64
Item   (Intercept) 0.002228  0.04721
       Conditionb 0.009344  0.09666   -1.00
       Conditionc 0.001284  0.03583   -1.00  1.00
Residual 0.089353  0.29892

Number of obs: 768, groups: Sujet, 32; Item, 24

Fixed effects:
              Estimate Std. Error      df t value Pr(>|t|)
(Intercept)  6.299930    0.043633  64.629993  144.384 <2e-16 ***
Conditionb   0.023219    0.040059  24.531233   0.580  0.567
Conditionc   0.029472    0.035804  28.192312   0.823  0.417
WordLength  -0.002901    0.009281  125.595440  -0.313  0.755
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
(Inter) Cndtnb Cndtnc
Conditionb -0.290
Conditionc -0.212  0.587
WordLength -0.591 -0.160 -0.191

```

N.1.2 Avec la Condition b (REL.AMB) comme niveau de comparaison

```

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(TempsLecture) ~ cond2 + WordLength + (cond2 | Sujet) + (cond2 | Item)
Data: dis.word

REML criterion at convergence: 450.2

Scaled residuals:
   Min       1Q   Median       3Q      Max
-1.9919 -0.6000 -0.1651  0.4267  6.2675

Random effects:
 Groups Name      Variance Std.Dev. Corr
Sujet  (Intercept)  0.032220  0.17950
      cond2a     0.015203  0.12330    -0.50
      cond2c     0.011139  0.10554    -0.21  0.42
Item   (Intercept)  0.002446  0.04946
      cond2a     0.009344  0.09666    -1.00
      cond2c     0.003701  0.06083    -1.00  1.00
Residual                    0.089353  0.29892

Number of obs: 768, groups: Sujet, 32; Item, 24

Fixed effects:
              Estimate Std. Error      df t value Pr(>|t|)
(Intercept)  6.323149   0.049949  75.582243  126.593 <2e-16 ***
cond2a      -0.023219   0.040059  24.531229  -0.580  0.567
cond2c       0.006254   0.034662  23.630431  0.180  0.858
WordLength  -0.002901   0.009281  125.595862  -0.313  0.755
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
      (Intr) cond2a cond2c
cond2a  -0.549
cond2c  -0.339  0.549
WordLength -0.644  0.160 -0.012

```

N.2 Au niveau de la région 'post.disamb' (codée +1)

N.2.1 Avec la Condition a (COMPL) comme niveau de comparaison

```

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(TempsLecture) ~ Condition + WordLength + (Condition | Sujet) + (Condition | Item)
Data: after.dis

REML criterion at convergence: 679.3

Scaled residuals:
   Min       1Q   Median       3Q      Max
-2.8899 -0.5914 -0.1672  0.4353  4.9018

Random effects:
 Groups Name Variance Std.Dev. Corr
Sujet   (Intercept) 0.055474 0.23553
        Conditionb 0.010257 0.10128 -0.55
        Conditionc 0.017440 0.13206 -0.88 0.88
Item    (Intercept) 0.007551 0.08690
        Conditionb 0.012806 0.11316  0.03
        Conditionc 0.006980 0.08354 -0.15 0.99
Residual 0.118670 0.34448

Number of obs: 768, groups: Sujet, 32; Item, 24

Fixed effects:
              Estimate Std. Error   df t value Pr(>|t|)
(Intercept)  6.195978   0.080606 59.389640  76.867 < 2e-16 ***
Conditionb   0.076455   0.052405 31.803845   1.459 0.15439
Conditionc   0.001707   0.052453 38.382418   0.033 0.97420
WordLength   0.026614   0.008238 37.982800   3.231 0.00255 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
(Intr) Cndtnb Cndtnc
Conditionb  -0.668
Conditionc  -0.792  0.797
WordLength  -0.783  0.591  0.598

```

N.2.2 Avec la Condition b (REL.AMB) comme niveau de comparaison

```

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(TempsLecture) ~ condZ + WordLength + (condZ | Sujet) + (condZ | Item)
Data: after.dis

REML criterion at convergence: 679.3

Scaled residuals:
  Min       1Q   Median       3Q      Max
-2.8899 -0.5914 -0.1672  0.4353  4.9018

Random effects:
 Groups Name      Variance Std.Dev. Corr
Sujet  (Intercept)  0.039446  0.19861
      condZa     0.010257  0.10128   0.14
      condZc     0.004231  0.06505  -1.00 -0.22
Item   (Intercept)  0.020859  0.14443
      condZa     0.012806  0.11316  -0.80
      condZc     0.001160  0.03407  -0.98  0.91
Residual                    0.118670  0.34448
Number of obs: 768, groups: Sujet, 32; Item, 24

Fixed effects:
              Estimate Std. Error      df t value Pr(>|t|)
(Intercept)  6.272433    0.060019  53.941761 104.507 < 2e-16 ***
condZa      -0.076455    0.052405  31.803850  -1.459  0.15439
condZc      -0.074747    0.033401  75.840952  -2.238  0.02816 *
WordLength  0.026614    0.008238  37.982794   3.231  0.00255 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
      (Intr) condZa condZc
condZa  0.024
condZc -0.541  0.317
WordLength -0.536 -0.591  0.012

```

N.3 Au niveau de la région 'post.post.disamb' (codée +2)

N.3.1 Avec la Condition a (COMPL) comme niveau de comparaison

```

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(TempsLecture) ~ Condition + WordLength + (Condition | Sujet) + (Condition | Item)
Data: after.after.dis

REML criterion at convergence: 543.2

Scaled residuals:
   Min       1Q   Median       3Q      Max
-3.6565 -0.6025 -0.1192  0.4873  4.1385

Random effects:
 Groups Name          Variance Std.Dev. Corr
Sujet   (Intercept) 0.021077 0.14518
        Conditionb 0.053604 0.23153 0.79
        Conditionc 0.011898 0.10908 1.00 0.77
Item     (Intercept) 0.002131 0.04616
        Conditionb 0.027487 0.16579 -0.76
        Conditionc 0.012416 0.11143 0.45 0.23
Residual                    0.112199 0.33496

Number of obs: 618, groups:  Sujet, 32; Item, 24

Fixed effects:
              Estimate Std. Error   df t value Pr(>|t|)
(Intercept)  6.235686   0.050348 71.704484 123.852 <2e-16 ***
Conditionb   0.157378   0.066965 35.230950  2.350  0.0245 *
Conditionc   0.088371   0.051574 31.492695  1.713  0.0965 .
WordLength   0.019187   0.009419 43.104241  2.037  0.0478 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
      (Intr) Cndtnb Cndtnc
Conditionb  0.207
Conditionc  0.340 0.462
WordLength -0.729 -0.226 -0.387

```

N.3.2 Avec la Condition b (REL.AMB) comme niveau de comparaison

```

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(TempsLecture) ~ cond2 + WordLength + (cond2 | Sujet) + (cond2 | Item)
Data: after.after.dis

REML criterion at convergence: 543.2

Scaled residuals:
   Min       1Q   Median       3Q      Max
-3.6565 -0.6025 -0.1192  0.4873  4.1385

Random effects:
 Groups Name      Variance Std.Dev. Corr
Sujet  (Intercept)  0.12751  0.3571
      cond2a    0.05360  0.2315  -0.97
      cond2c    0.02656  0.1630  -0.77  0.90
Item   (Intercept) 0.01795  0.1340
      cond2a    0.02749  0.1658  -0.97
      cond2c    0.03138  0.1771  -0.63  0.79
Residual                    0.11220  0.3350
Number of obs: 618, groups: Sujet, 32; Item, 24

Fixed effects:
Estimate Std. Error   df t value Pr(>|t|)
(Intercept)  6.393063  0.091741 55.440985  69.686 <2e-16 ***
cond2a     -0.157377  0.066965 35.231017  -2.350  0.0245 *
cond2c     -0.069007  0.062882 22.395208  -1.097  0.2841
WordLength  0.019187  0.009419 43.104261  2.037  0.0478 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
(Intr) cond2a cond2c
cond2a   -0.844
cond2c   -0.469  0.686
WordLength -0.565  0.226 -0.077

```

Annexe O

Analyses statistiques pour la deuxième expérience de lecture par présentation auto-segmentée, sur l'ambiguïté REL/INT (Chapitre 6, Section 6.2.1.3)

O.1 Première comparaison : REL-ce qui/ce que versus INT-ce qui/ce que

O.1.1 Au niveau de la région critique, SUB word (codée 0)

```

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(TempsLecture) ~ PhrseType * GapType + longueur + (PhrseType *
GapType | Sujet) + (PhrseType * GapType | Item)
Data: crit.point

REML criterion at convergence: 539.1

Scaled residuals:
  Min       1Q   Median       3Q      Max
-2.7208 -0.6176 -0.1037  0.4567  7.4828

Random effects:
 Groups Name                            Variance Std.Dev. Corr
 Item   (Intercept)                        0.001263 0.03554
        PhrseType                       0.003630 0.06025  0.36
        GapType                          0.006634 0.08145 -0.54 -0.29
 Sujet  PhrseType:GapType                0.030823 0.17556  0.12 -0.23  0.00
        (Intercept)                    0.106671 0.32661
        PhrseType                       0.004251 0.06520  0.06
        GapType                          0.003135 0.05599 -0.50 -0.08
        PhrseType:GapType                0.016175 0.12718 -0.13  0.09  0.33

Residual                               0.065477 0.25588
Number of obs: 2080, groups: Item, 50; Sujet, 28

Fixed effects:
              Estimate Std. Error      df t value Pr(>|t|)
(Intercept)  6.085458   0.062326 27.405050  97.639 <2e-16 ***
PhrseType    0.007341   0.020021 33.731622   0.367  0.716
GapType     -0.009904   0.020404 33.577160  -0.485  0.631
PhrseType:GapType -0.021064   0.043558 38.461928  -0.484  0.631
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
(Intr) PhrsTy GapTyp
PhrseType  0.048
GapType   -0.288 -0.098
PhrsTyp:GpT -0.066 -0.053  0.105
fit warnings:
fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
    
```

O.1.2 Au niveau de la région critique+1, DET/AUX

```

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(TempsLecture) ~ PhraseType * GapType + longueur + (PhraseType *
      GapType | Sujet) + (PhraseType * GapType | Item)
Data: after.crit

REML criterion at convergence: 837.1

Scaled residuals:
  Min       1Q   Median       3Q      Max
-4.4046 -0.5653 -0.1113  0.4401  6.3983

Random effects:
 Groups Name                    Variance Std.Dev. Corr
 Item   (Intercept)                0.002340  0.04837
        PhraseType              0.007221  0.08498      0.00
        GapType                 0.017499  0.13228     -0.54  0.25
Sujet  PhraseType:GapType       0.006108  0.07815     0.62 -0.70 -0.79
        (Intercept)            0.126205  0.35525
        PhraseType              0.002537  0.05037     -0.09
        GapType                 0.024110  0.15527     -0.63 -0.32
        PhraseType:GapType      0.009644  0.09820     0.36 -0.14 -0.21
Residual                    0.074592  0.27312

Number of obs: 2080, groups: Item, 50; Sujet, 28

Fixed effects:
              Estimate Std. Error      df t value Pr(>|t|)
(Intercept)    6.163195    0.092638  66.940572  66.530 <2e-16 ***
PhraseType     -0.010214    0.021143  39.328935  -0.483  0.6317
GapType        -0.094800    0.056310  73.578866  -1.684  0.0965 .
longueur       -0.006905    0.008419  49.915982  -0.820  0.4160
PhraseType:GapType 0.022770    0.036562  38.871412  0.623  0.5371
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
      (Intr) PhrsTy GapTyp Longur
PhraseType -0.075
GapType    -0.758  0.017
longueur   -0.680  0.048  0.741
PhrsTyp:gpT 0.182 -0.221 -0.214 -0.045

```

O.1.3 Au niveau de la région critique+2, N3/AUX

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
 Formula: log(TempsLecture) ~ PhraseType * GapType + longueur + (PhraseType *
 GapType | Sujet) + (PhraseType * GapType | Item)
 Data: after.after.crit

REML criterion at convergence: 835.6

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.1728	-0.5337	-0.1053	0.4065	6.4717

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
Item	(Intercept)	0.0008639	0.02939	
	PhraseType	0.0068142	0.08255	0.72
	GapType	0.0091547	0.09568	-0.80
	PhraseType:GapType	0.0030698	0.05541	0.59
Sujet	(Intercept)	0.0945610	0.30751	
	PhraseType	0.0052597	0.07252	0.36
	GapType	0.0142633	0.11943	0.11
	PhraseType:GapType	0.0111053	0.10538	0.28
Residual		0.0759797	0.27564	

Number of obs: 2080, groups: Item, 50; Sujet, 28

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	6.017293	0.064064	38.560716	93.926	< 2e-16 ***
PhraseType	-0.009145	0.022758	44.314548	-0.402	0.689749
GapType	-0.046363	0.029951	33.967667	-1.548	0.130899
longueur	0.016772	0.004237	226.042273	3.958	0.000101 ***
PhraseType:GapType	-0.006856	0.035565	39.365531	-0.193	0.848129

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

	(Intr)	PhrsTy	GapTyp	longur
PhraseType	0.205			
GapType	0.042	-0.003		
longueur	-0.399	0.011	0.014	
PhrsTyp:GpT	0.169	0.283	0.065	-0.049

O.1.4 Au niveau de la région critique+3, SUB V

```

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(TempsLecture) ~ PhraseType * GapType + longueur + (PhraseType *
      GapType | Sujet) + (PhraseType * GapType | Item)
Data: three.after.crit

REML criterion at convergence: 63.3

Scaled residuals:
   Min       1Q   Median       3Q      Max
-5.4821 -0.5451 -0.0994  0.4615  6.8483

Random effects:
 Groups Name                    Variance Std.Dev. Corr
 Item   (Intercept)                0.0007604 0.02758
        PhraseType              0.0028667 0.05354  0.41
        GapType                 0.0063648 0.07978  -0.04 -0.24
 Sujet  PhraseType:GapType       0.0052458 0.07243  -0.50 -0.66 -0.55
        (Intercept)            0.0755088 0.27479
        PhraseType             0.0042881 0.06548  0.12
        GapType                0.0024054 0.04904  0.16 -0.22
        PhraseType:GapType     0.0164088 0.12810  0.00 -0.27 -0.79
Residual                    0.0525445 0.22923

Number of obs: 2080, groups: Item, 50; Sujet, 28

Fixed effects:
              Estimate Std. Error    df t value Pr(>|t|)
(Intercept)    6.066571    0.054611 32.010914 111.088 <2e-16 ***
PhraseType    -0.002397    0.018657 34.926744  -0.128  0.8985
GapType       -0.032697    0.018497 34.863379  -1.768  0.0859 .
longueur      0.005067    0.004165 46.516530  1.216  0.2299
PhraseType:GapType -0.016076    0.035309 34.144252  -0.455  0.6518
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

O.1.5 Au niveau de la région critique+4, PREP

```

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(Tempslecture) ~ PhraseType * GapType + longueur + (PhraseType *
      GapType | Sujet) + (PhraseType * GapType | Item)
Data: four.after.crit

REML criterion at convergence: 578.4

Scaled residuals:
   Min       1Q   Median       3Q      Max
-5.1691 -0.5446 -0.1026  0.4124  6.0906

Random effects:
 Groups Name                    Variance Std.Dev. Corr
 Item   (Intercept)                0.001776  0.04214
        PhraseType              0.010872  0.10427   -0.16
        GapType                 0.007847  0.08858    0.23  0.46
 Sujet  PhraseType:GapType      0.001012  0.03182   -0.58  0.23  0.60
        (Intercept)            0.079862  0.28260
        PhraseType             0.004594  0.06778    0.05
        GapType                0.005574  0.07466    0.56  0.48
        PhraseType:GapType     0.017920  0.13386   -0.16  0.52 -0.11
Residual                    0.066521  0.25792

Number of obs: 2080, groups: Item, 50; Sujet, 28

Fixed effects:
              Estimate Std. Error   df t value Pr(>|t|)
(Intercept)  5.986059   0.055217 29.852801 108.409 < 2e-16 ***
PhraseType   -0.024403   0.023654 45.523528  -1.032  0.308
GapType      0.020702   0.023139 44.311932  0.895  0.376
longueur     0.023090   0.002523 44.541088  9.152 8.52e-12 ***
PhraseType:GapType 0.032805   0.037337 34.264410  0.879  0.386
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
    
```

O.2 Deuxième comparaison : INT-ce qui/ce que *versus* INT-quand

O.2.1 Au niveau de la région critique, SUB word (codée 0)

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
 Formula: log(TempsLecture) ~ CEQUE + CEQUI + longueur + (CEQUE + CEQUI | Sujet) +
 (CEQUE + CEQUI | Item)
 Data: crit.point

REML criterion at convergence: 343.7

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.4192	-0.5814	-0.0985	0.4810	5.4869

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
Item	(Intercept)	0.001516	0.03893	
	CEQUE	0.014806	0.12168	-0.17
	CEQUI	0.008055	0.08975	-0.49
Sujet	(Intercept)	0.106738	0.32671	
	CEQUE	0.005033	0.07094	0.14
	CEQUI	0.005257	0.07251	-0.15
Residual		0.062164	0.24933	

Number of obs: 1560, groups: Item, 50; Sujet, 28

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	6.56572	0.58966	29.99809	11.135	3.54e-12 ***
CEQUE	0.06828	0.10442	32.15963	0.654	0.518
CEQUI	0.06849	0.10372	31.40049	0.660	0.514
longueur	-0.08399	0.10335	29.35111	-0.813	0.423

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

O.2.2 Au niveau de la région critique+1, DET/AUX

```

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(TempsLecture) ~ CEQUE + CEQUI + longueur + (CEQUE + CEQUI | Sujet) +
(CEQUE + CEQUI | Item)
Data: after.crit

REML criterion at convergence: 557.7

Scaled residuals:
   Min       1Q   Median       3Q      Max
-3.5947 -0.5781 -0.1180  0.4579  7.9106

Random effects:
 Groups Name      Variance Std.Dev. Corr
 Item   (Intercept) 0.002935 0.05417
        CEQUE     0.020029 0.14153      0.53
        CEQUI     0.013347 0.11553     -0.31 0.37
 Sujet  (Intercept) 0.126487 0.35565
        CEQUE     0.009178 0.09580      0.69
        CEQUI     0.005119 0.07155     -0.60 -0.98
Residual                    0.070354 0.26524
Number of obs: 1560, groups: Item, 50; Sujet, 28

Fixed effects:
              Estimate Std. Error   df t value Pr(>|t|)
(Intercept)  5.9304949  0.0872859 58.2406251  67.943 < 2e-16 ***
CEQUE       -0.0003403  0.0341266 37.2849131  -0.010 0.992096
CEQUI        0.0500245  0.0426677 58.4671681  1.172 0.245783
longueur     0.0241167  0.0065735 44.4669032  3.669 0.000649 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

O.2.3 Au niveau de la région critique+2, N3/Aux

```

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(TempsLecture) ~ CEQUE + CEQUI + longueur + (CEQUE + CEQUI | Sujet) +
(CEQUE + CEQUI | Item)
Data: after.after.crit

REML criterion at convergence: 500.2

Scaled residuals:
   Min       1Q   Median       3Q      Max
-3.1797 -0.6082 -0.1123  0.4770  6.2708

Random effects:
 Groups Name Variance Std.Dev. Corr
 Item   (Intercept) 0.0001853 0.01361
        CEQUE      0.0096925 0.09845  0.98
        CEQUI      0.0007145 0.02673 -0.99 -1.00
 Sujet  (Intercept) 0.0927233 0.30451
        CEQUE      0.0038982 0.06244 -0.22
        CEQUI      0.0046767 0.06839 -0.11 -0.12
Residual                    0.0708259 0.26613
Number of obs: 1560, groups: Item, 50; Sujet, 28

Fixed effects:
Estimate Std. Error   df t value Pr(>|t|)
(Intercept)  5.994905   0.063200  37.095628  94.855 < 2e-16 ***
CEQUE       0.030163   0.025749  24.315145   1.171  0.253
CEQUI      -0.014221   0.022020  26.519647  -0.646  0.524
longueur    0.020651   0.004064  45.982923   5.081 6.69e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

O.2.4 Au niveau de la région critique+3, SUB V

```

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(TempsLecture) ~ CEQUE + CEQUI + longueur + (CEQUE + CEQUI | Sujet) +
(CEQUE + CEQUI | Item)
Data: three.after.crit

REML criterion at convergence: 350

Scaled residuals:
  Min       1Q   Median       3Q      Max
-5.0989 -0.5589 -0.0941  0.4372  5.1940

Random effects:
 Groups Name      Variance Std.Dev. Corr
 Item   (Intercept) 0.0006391 0.02528
        CEQUE     0.0108675 0.10425
        CEQUI     0.0060322 0.07767
 Sujet  (Intercept) 0.0792864 0.28158
        CEQUE     0.0030045 0.05481
        CEQUI     0.0084449 0.09190
Residual                    0.0635777 0.25215
Number of obs: 1560, groups: Item, 50; Sujet, 28

Fixed effects:
Estimate Std. Error      df t value Pr(>|t|)
(Intercept) 6.072733  0.057054 33.983862 106.439 < 2e-16 ***
CEQUE      -0.055956  0.024885 34.101398  -2.249 0.03111 *
CEQUI      -0.080692  0.027385 29.606267  -2.947 0.00621 **
longueur    0.010303  0.004991 56.239301  2.064 0.04362 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

O.2.5 Au niveau de la région critique+4, PREP

```

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(TempsLecture) ~ CEQUE + CEQUI + longueur + (CEQUE + CEQUI | Sujet) +
(CEQUE + CEQUI | Item)
Data: four.after.crit

REML criterion at convergence: 576.7

Scaled residuals:
  Min      1Q  Median      3Q      Max
-4.8481 -0.5831 -0.1183  0.4104  5.6111

Random effects:
 Groups Name      Variance Std.Dev. Corr
 Item   (Intercept)  0.002358  0.04856
        CEQUE     0.018003  0.13418  -0.27
        CEQUI     0.021089  0.14522  -0.23  0.82
 Sujet  (Intercept)  0.072145  0.26860
        CEQUE     0.010751  0.10369  -0.03
        CEQUI     0.024429  0.15630  0.26  0.81
Residual                    0.071440  0.26728
Number of obs: 1560, groups: Item, 50; Sujet, 28

Fixed effects:
Estimate Std. Error      df t value Pr(>|t|)
(Intercept)  6.052956    0.053390  31.343321 113.373 < 2e-16 ***
CEQUE      -0.094203    0.033815  42.936764  -2.786  0.00791 **
CEQUI      -0.084987    0.041511  46.998350  -2.047  0.04624 *
longueur    0.016203    0.003011  37.231430   5.381  4.25e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```