



HAL
open science

Acquisition et maîtrise des interrogatives partielles en français : La variation sociolinguistique comme outil interactionnel

Gabriel Thiberge

► **To cite this version:**

Gabriel Thiberge. Acquisition et maîtrise des interrogatives partielles en français : La variation sociolinguistique comme outil interactionnel. Linguistique. Université Paris Cité, 2020. Français. NNT : . tel-03784454

HAL Id: tel-03784454

<https://shs.hal.science/tel-03784454>

Submitted on 23 Sep 2022

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

Spécialité doctorale “Sciences du Langage”

Acquisition et maîtrise des interrogatives partielles en français : La variation sociolinguistique comme outil interactionnel

présentée et soutenue publiquement par

Gabriel THIBERGE

le 14 décembre 2020

Directrice de thèse : **Barbara HEMFORTH**

Jury

Aliyah MORGENSTERN, Professeure, Université Sorbonne Nouvelle, Paris 3, Rapportrice

Sarah SCHIMKE, Professeure, Technische Universität Dortmund, Rapportrice

Olivier BONAMI, Professeur, Université de Paris, Examineur

Heather BURNETT, Directrice de Recherche, Université de Paris - CNRS, Examinatrice

Christophe PARISSÉ, Chargé de Recherche, Université Paris Nanterre - CNRS, INSERM, Examineur

Barbara HEMFORTH, Directrice de Recherche, Université de Paris - CNRS, Directrice de Thèse

Université de Paris

École doctorale n° 622 : Sciences du Langage

Laboratoire de Linguistique Formelle (LLF)

UMR CNRS 7110, F-75013 Paris, France



**Acquiring and mastering French partial interrogatives :
Sociolinguistic variation as a tool in interaction**

0.1 Summary

This work investigates the great variability in the production and comprehension of French partial interrogatives, as well as their acquisition. These sentences can be produced with multiple forms and, most notably, the interrogative element can be either after the verb (in situ) or before the verb (fronting, with or without verb-subject inversion).

This variability has been linked to constraints of a syntactic, phonotactic or pragmatic origins, as well as to sociolinguistic factors. Some syntactic theories argue for instance that there is a complexity gradient among interrogative forms, which could explain why French speaking children often seem to first produce in situ forms : under this view, it is the least costly.

After a short introduction in Chapter 1, Chapter 2 shows the need for an update of the available data, for children and adults alike, to develop a multifactorial account of the variation phenomenon. It is also argued that there is a need to refine the very notion of sociolinguistic variation. Alternative linguistic forms allow the user to build and modify their social identity (*persona*) throughout interaction. Children have to acquire this ability at the same time as they acquire language.

The approach here is data-driven, as explained in Chapter 3, and the analyses presented build upon corpus and experimental results. Chapter 4 is an analysis of modern-day spontaneous and spoken French from the 1960s-2000s (ESLO and EPAC projects). The results show that the use of interrogative forms is specialised, with influences from context formality and from the highly social and goal-oriented nature of language. The various interrogative sentence types are not used interchangeably among all contexts, and some variants are used according to speaker strategies.

These results are then refined with experimental studies. Chapter 5 presents seven experiments with adults. Acceptability judgments are used to assess French speakers' preferences with respect to the different interrogative forms. With a refined notion of acceptability, they allow for a distinction between "good French", and constructions that should be used in a given context. Three matched-guise tasks further investigate the social stereotypes associated with interrogative forms by adults. From there, a "social sketch" emerges which defines the stereotype of a *persona* using these variants.

Chapter 6 describes two studies run with French speaking children. The first study is an adapted version of a study previously run with adults. In a group of 136 children (aged 3-11), it appears some associations between linguistic forms and social stereotypes (richness, hobbies, education) are active from very early on, and evolve over time. The second study is a production task where 68 children (aged 4-9) played a role-playing game. It shows an early capacity to actively adapt the production of interrogatives depending on the context. Even young children display different linguistic behavior in formal or informal contexts. They refine this ability with increasing age.

Data from adults and children, when confronted, argue for a broader perspective in the study of French partial interrogatives and their acquisition. The coexistence of different interrogative forms in French is tied to social aspects, and to stereotypical representations of the people using them. Children are confronted with and impacted by these representations from early childhood, and their linguistic production is impacted as well.

Keywords : experimental, sociolinguistics, French, partial interrogation, acquisition, corpus

0.2 Résumé

Ce travail explore la variation en matière de production et de réception d'interrogatives partielles en français, et certains aspects de leur acquisition. Les phrases interrogatives partielles du français présentent une grande variété de formes et l'élément interrogatif peut notamment être après le verbe (in situ) ou avant (antéposition, accompagnée ou non d'une inversion verbe-sujet).

Cette variation a été reliée à des contraintes syntaxiques, phonotactiques, pragmatiques, ou encore à des enjeux sociolinguistiques. Des théories syntaxiques établissent notamment une différence de complexité entre les formes interrogatives, qui expliquerait la production précoce générale de formes in situ par l'enfant francophone : ce type serait le moins coûteux à produire.

Après l'introduction au Chapitre 1, le Chapitre 2 montre la nécessité d'enrichir les données disponibles, chez l'enfant et l'adulte, pour une vision multifactorielle du phénomène variationnel. Le besoin de considérer la variation sociolinguistique comme un outil est avancé. La concurrence de formes linguistiques permet à l'individu de construire et modifier son masque social (*persona*) dans l'interaction. Les enfants doivent développer cette capacité pendant leur acquisition du langage.

L'approche est empirique, comme expliqué dans le Chapitre 3, sur base de données de corpus et expérimentales. Le Chapitre 4 analyse des données de français oral spontané des années 1960-2000 (projets ESLO, EPAC). En intégrant la formalité du contexte interactionnel et le caractère fondamentalement social et utilitaire du langage, les données témoignent d'une utilisation spécialisée des formes interrogatives. Les variantes ne sont pas utilisées de la même façon dans tous les contextes, et certaines sont utilisées en fonction de la stratégie argumentative de la personne qui les produit.

Ces observations de corpus sont croisées et systématisées par des protocoles expérimentaux. Le Chapitre 5 présente sept expériences menées chez l'adulte. Des jugements d'acceptabilité sondent les préférences francophones en matière d'interrogatives, et nuancent la notion d'acceptabilité pour distinguer entre « bon français » et français adapté au contexte. Des tâches de *matched-guise* approfondissent les représentations et stéréotypes sociaux liés chez l'adulte aux différentes variantes interrogatives. Il s'en dégage un profil social type des personnes les utilisant.

Le Chapitre 6 décrit deux expériences menées auprès d'enfants francophones. La première suit un protocole parallèle aux études menées chez l'adulte mais adapté à l'enfant (N=136, âge=3-11 ans). Des associations entre formes linguistiques et stéréotypes sociaux (richesse, loisirs, études) se révèlent actives dès le plus jeune âge, puis se modifient avec le développement de l'enfant. Une activité de production spontanée met ensuite les enfants dans une situation de jeux de rôles (N=68, âge=4-9 ans). Elle montre une capacité active d'adaptation au contexte dans leurs productions interrogatives. Les enfants changent très tôt leur comportement linguistique selon que l'interaction est formelle ou qu'elle est informelle, puis perfectionnent leur maîtrise de cette capacité.

La confrontation des données des adultes et des enfants plaident pour un élargissement général du cadre d'analyse des interrogatives partielles et de leur acquisition. L'alternance de formes en français est régie par des considérations sociales, liées aux représentations stéréotypées des personnes qui les utilisent. Les enfants subissent l'impact dès leur plus jeune âge de ces représentations, et leurs productions précoces en sont nécessairement touchées.

Mots-Clefs : expérimental, sociolinguistique, français, interrogation partielle, acquisition, corpus

0.3 Déclaration

En tant que chercheur doctorant, j'ai travaillé sous la direction de Barbara Hemforth, Directrice de Recherches au CNRS. Ses apports à toutes les différentes facettes du présent travail sont indéniables et difficilement quantifiables.

Je reste néanmoins seul responsable du propos développé tout au long des chapitres suivants et des analyses qui y sont présentées. Notamment, toutes les erreurs, omissions ou contre-vérités qui s'y trouveraient encore ne sont que les miennes.

L'intégralité de ce texte a été rédigé par mes soins, sous la relecture attentive de bonnes âmes, et bien évidemment de Barbara. Autant que possible, il a été indiqué l'origine des idées, théories, protocoles, matériels et autres faits présentés. Tout manquement à cette règle n'est que le fruit de mon erreur. Je ferai tout mon possible pour y remédier, sur simple contact.

Pour contribuer à une science totalement ouverte, accessible et transparente, j'ai cherché à inclure dans le corps du texte aussi bien que dans les annexes la liste exhaustive des fruits de ce travail. Il s'agit de résultats, de matériels, de protocoles expérimentaux ou de méthodes d'analyse. Je fournirai plus d'informations sur chacun de ces points sur simple contact.

Un accès libre aux données (anonymisées), scripts et analyses principales, ainsi qu'à une copie numérique de ce document et des annexes, est ou sera possible via mon répertoire personnel sur le site du projet OSF [FOSTER et DEARDORFF, 2017], par le lien osf.io/ud4ks.

Comme illustré en couverture, la réutilisation de tout ou partie de ce travail est entièrement libre et gratuite, et ne nécessite pas d'autorisation préalable de ma part, tant :

- qu'elle sert l'avancement de la connaissance humaine,
- qu'elle est faite dans le respect de la vie et avec éthique, sous un quelconque contrôle extérieur,
- qu'elle ne génère aucun profit pécunier.

La mention de mon nom et la référence de ce travail est un plus.

0.4 Crédits

Beaucoup de gens et d'institutions ont permis à ce travail de commencer et d'aboutir.

0.4.1 Soutiens institutionnels

Ce travail a bénéficié partiellement d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du programme "Investissements d'Avenir" portant la référence ANR-10-LABX-0083. Il contribue à l'IdEx Université de Paris - ANR-18-IDEX-0001.

Il a été financé par un contrat doctoral de l'École Doctorale 622 (anciennement 132), liée à l'Université de Paris.

Il a été mené au sein de l'Unité Mixte de Recherche 7110 du CNRS, le Laboratoire de Linguistique Formelle, qui a soutenu financièrement la diffusion des résultats, par exemple à des colloques et conférences, en France mais aussi à l'international, ainsi que la collecte de données et diverses missions et formations.

Des mobilités à l'étranger et la rémunération de certaines expériences ont aussi été financées par le Labex EFL, Empirical Foundations of Linguistics. D'autres mobilités ont été financées par la région Île-de-France (Aide à la Mobilité Internationale, AMI, à destination des doctorants et doctorantes). Ces mobilités ont toutes contribué à la réalisation de ce travail et à la maturation des idées qui y sont présentées.

0.4.2 Logiciels et *packages* utilisés

Le template de ce document de thèse est dérivé de celui développé par Dorian Depriester (blog.dorian-depriester.fr), pour les thèses rédigées en français sous LaTeX (latex-project.org).

Ce document a été rédigé notamment en ligne via le projet Overleaf (overleaf.com).

L'intégralité des descriptions statistiques et des visualisations présentées a été développée grâce au logiciel R [[RSTUDIO TEAM, 2015](#)]. Plusieurs modules de ce logiciel ont été utilisés en plus des modules de base, dont :

- pour l'import, et plus généralement le traitement des données : *dplyr* [[WICKHAM et al., 2020](#)], *plyr* [[WICKHAM, 2011](#)], *readr* [[WICKHAM et al., 2018](#)], *reshape2* [[WICKHAM, 2007](#)];
- pour le calcul des corrélations : *corrplot* [[WEI et SIMKO, 2017](#)], *hmisc* [[HARRELL JR et al., 2020](#)];
- pour la visualisation des données : *ggplot2* [[WICKHAM, 2016](#)], *scales* [[WICKHAM, 2018](#)];
- pour le calcul des modèles bayésiens, la constitution des tableaux en découlant et leur représentation graphique : *brms* [[BÜRKNER, 2018](#); [BÜRKNER et al., 2017](#)], *domc* [[ANALYTICS et WESTON, 2019](#)], *ggridges* [[WILKE, 2020](#)], *ggstance* [[HENRY et al., 2019](#)], *shinystan* [[GABRY, 2018](#)], *sjPlot* [[LÜDECKE, 2019](#)], *stan* [[CARPENTER et al., 2017](#)], *tidyverse* [[WICKHAM et al., 2019](#)];
- pour faciliter l'import des tableaux dans LaTeX : *xtable* [[DAHL et al., 2019](#)].

Le calcul des nombreux modèles statistiques présentés dans ce travail est une activité qui demande une mobilisation de ressources informatiques importantes. La licence d'utilisation de Rstudio Web fournie par la TGIR HUMANUM (UMS 3598, <https://r-tools.huma-num.fr/>, <https://www.huma-num.fr/>) a énormément facilité la menée de ces analyses.

0.5 Remerciements

Cette section est dans les thèses celle que je préfère lire. Elle est la trace de ce que ces travaux sont réalisés par des humains. J'en suis un également et beaucoup d'autres m'ont aidé.

Barbara Hemforth est bien sûr la première personne à avoir permis à ce travail de voir le jour. C'est à elle, à sa perspective globale et ouverte à travers les domaines linguistiques, et à son expérience, que je dois d'avoir pu conceptualiser, organiser, mener, traiter, écarter, réorganiser, retraiter, et enfin condenser la myriade d'idées qui me sont venues avant et pendant la menée de cette thèse. C'est aussi grâce à elle si les phrases de ce travail sont d'une longueur qui les rend lisibles. Merci.

Les membres de mon comité de suivi, Heather Burnett et Christoph Scheepers, m'ont apporté une foule d'idées et de commentaires tout au long de ces trois dernières années, pendant et en dehors de nos réunions annuelles. Je les en remercie. Je voudrais aussi remercier Carla Soares-Jesel pour m'avoir permis de réintégrer un cursus de Sciences du Langage quand la chose me paraissait inenvisageable, et pour avoir guidé mes premiers pas dans la voie des études en acquisition. Je remercie aussi Caterina Donati pour m'avoir permis de « mettre le nez dans les données » et d'interroger mes premiers corpus d'interrogatives en Master. Je remercie aussi Anne Abeillé pour m'avoir donné mes premières nouvelles données à mouder en début de doctorat avec le corpus EPAC, pour sa curiosité permanente sur mon travail et les liens, idées et travaux reliés aux miens qu'elle m'a régulièrement transmis, et encore pour avoir relu une partie de cette thèse.

Je remercie par ailleurs les membres de mon jury pour leur intérêt sur mon travail : Olivier Bonami, Heather Burnett, Christophe Parisse, Aliyah Morgenstern et Sarah Schimke.

La majorité de mon temps ayant été passé au Laboratoire de Linguistique Formelle, c'est en fait l'intégralité de ses membres qui m'a aidé au quotidien. Je voudrais remercier en particulier Jeff Alexandre, Olivier Bonami, Ali Dhibi, Doriane Gras, Achille Falaise, Sophie Lempérière, Viviane Makougni et Alexandre Roulois pour leur disponibilité et leur assistance aux niveaux les plus divers de mon évolution dans le laboratoire. J'en profite pour remercier Ioana Chitoran pour son aide au niveau de l'école doctorale 132/622 et pour son implication sans faille, ainsi que Claire Saillard puis Benoît Crabbé pour leur bienveillance au sein de l'UFR Linguistique. Je remercie aussi Yair Haendler pour la clarté et la patience de ses explications statistiques, ainsi que Amelia Kimball pour nos discussions bayésiennes. Bien sûr, je remercie aussi Loïc Liégeois pour toutes nos discussions et pour sa collaboration dans l'exploitation de corpus, mais aussi pour être à l'origine de mes premières réflexions sur une tâche de jeu de rôle avec les enfants, qui permette un recueil de production spontanée mais contrôlée.

J'ai aussi bénéficié de discussions de fond avec Marzena Watorek quand je commençais à concevoir des études impliquant des enfants. Je remercie aussi tout le département de Philosophie, Psychologie et de Sciences du Langage de l'Université d'Édimbourg pour son accueil, et les premiers regards étrangers sur mon travail pendant les trois mois que j'ai passés en Écosse. Merci en particulier à Holly Branigan, Hannah Rohde, Kevin Smith, Anita Tobar et Alisdair Tullo.

Dans un autre registre, je veux aussi exprimer ma gratitude à Marc Pasquet pour son aide dans l'annotation de corpus ainsi qu'aux quatre voix anonymes qui m'ont permis de créer et de mener

trois de mes expériences, celles avec des stimuli audio. Mes expériences n'auraient bien sûr rien été non plus sans les volontaires qui les ont passées. Elles n'auraient rien été non plus, pour celles passées auprès d'enfants, sans l'implication de Caroline Fargeton, Vincent Brunie et Marie-Hélène Pargnien, que je remercie. Je remercie encore Élisabeth Houel et toute l'équipe pédagogique de son école pour m'avoir ouvert la porte de leurs classes, ainsi que les enfants qui ont accepté de jouer le jeu. Maxime Bouret a aussi été un marionnettiste formidable.

Dans des cadres plus ou moins formels, mais plutôt moins que plus, je remercie aussi pour leurs aides diverses, que ce soit sur une question éthique, un point expérimental ou une aide informatique, juste une idée au détour d'une conversation plus triviale, et même pour ces conversations complètement triviales : Aixiu An, Sandro Capo Chichi, Sacha Beniamine, Maximin Coavoux, Quentin Dénigot, Patty Garret, Jiaying Huang, Ingrid Konrad, Saïda Loucif, Vincent Ségonne, Olga Seminck, Yiqin Qiu et Wei Xiang. Merci à Suzanne Lesage, Céline Pozniak et Élodie Winckel pour beaucoup d'autres choses mais notamment pour l'éco-support au sein du laboratoire. Un deuxième merci à Céline pour avoir relu tant de passages de ma thèse. Merci à Olaf Mikkelsen, Anisia Popescu et Danilo Lombardo pour nos échanges, scientifiques mais pas que, en extérieur.

Encore plus à l'extérieur, je remercie toutes les personnes qui ne sont pas linguistes et qui m'ont soutenu sans vraiment s'en rendre compte : ma famille, mes proches, mes moins proches mais proches quand même. Ce sont ces gens qui ont fait tourner le monde pendant que la thèse m'engloutissait et qui, par leur distance vis-à-vis de mon domaine, ont permis d'ancrer un tant soit peu cette thèse dans le réel. C'est grâce à eux si le langage est un fait social, malgré les tentatives des linguistes pour le déconstruire.

Enfin enfin, merci à toi, *e*.

Les souvenirs resteront.

0.6 Notes de lecture

Derniers préalables à ce travail, voici quelques éléments de la grille de lecture avec laquelle ce document doit être parcouru :

1. **Typographie**

- une écriture inclusive a été utilisée dans ce travail, avec pour principal axe l'évacuation des contraintes de genre et d'accord grammaticaux. Au point médian, aux coordinations, et aux autres procédés disponibles, c'est l'utilisation de noms épïcènes qui a été préférée quand il était possible.

2. **Renvois internes**

- la version .pdf de ce document comprend un nombre important de liens internes qui permettent de naviguer entre des sous-sections, des chapitres, ou d'aller voir la bibliographie en cliquant sur une référence. Pour ce qui est des références, un lien vers la page d'origine est explicitement disponible pour chaque item de la bibliographie, et un retour en arrière est donc facile. Pour les autres liens, il faut cliquer pour revenir en arrière sur le bouton PRÉCÉDENT de votre logiciel de lecture de document .pdf (localisation variable selon le logiciel). Sinon, il faut retenir le numéro de la page de départ.

3. **Figures et Tableaux**

- Au sein d'un même chapitre, les tableaux et les figures suivent deux numérotations distinctes. La numérotation des figures et des tableaux repart à 1 pour chaque nouveau chapitre ou nouvelle annexe. Le numéro du chapitre ou de l'annexe précède le numéro de la figure ou du tableau. Une liste des figures et une liste des tableaux sont disponibles après la Table des matières.

4. **Présentation des données et statistiques**

- Un court chapitre est consacré aux méthodologies de recueil de données, qui présente également comment lire les visualisations et résultats des analyses statistiques présentées : il s'agit du Chapitre 3.

LINGUIST, *n.* A person more learned in the languages of others than wise in his own.
(Ambrose Bierce – *The Devil's Dictionary*, 1906)

Table des matières

0.1	Summary	iii
0.2	Résumé	v
0.3	Déclaration	vii
0.4	Crédits	ix
0.5	Remerciements	xi
0.6	Notes de lecture	xiii
	Table des matières	xvii
	Liste des figures	xxi
	Liste des tableaux	xxvii
1	Introduction Générale	1
1.1	Origines de la réflexion	3
1.2	Objectifs	4
1.3	Cadre de la réflexion et moyens d’investigation	5
1.4	Résultats généraux	5
1.5	Plan	7
2	L’interrogation partielle et son acquisition, des faits sociaux	9
2.1	L’interrogation comme objet d’étude sociolinguistique	11
2.2	Analyses de l’interrogation partielle en français	17
2.3	Pour une vision plus utilitaire de la variation linguistique	25
2.4	Pour une étude sociale de l’acquisition des interrogatives partielles	30
2.5	Bilan et résumé des axes de recherche	36
3	Méthodologies retenues et analyses effectuées	39
3.1	Méthodologies de recherche retenues	41
3.2	Représentations graphiques	44
3.3	Analyses bayésiennes et distributions postérieures	47
4	Nouvelles analyses de corpus	51
4.1	Un corpus adulte en contexte spécifique : EPAC	53
4.2	Un corpus adulte diachronique avec des contextes variés : ESLO	68
4.3	Conclusion intermédiaire	82
4.4	Apports majeurs	83

5	Nouvelles données expérimentales chez l'adulte francophone	85
5.1	Démarche : Protocoles expérimentaux retenus	87
5.2	Jugements d'acceptabilité et biais sociolinguistique	87
5.3	Matched-Guise et projection sociolinguistique	128
5.4	Conclusion intermédiaire	171
5.5	Apports majeurs	173
6	Nouvelles données expérimentales chez l'enfant	175
6.1	Démarche et choix des protocoles retenus	177
6.2	Matched-Guise enfantin (MG-C) et stéréotypes sociolinguistiques	178
6.3	Adaptation sociolinguistique en production : Activité de jeux de rôles (JDR-C)	193
6.4	Conclusion intermédiaire	212
6.5	Apports majeurs	213
7	Conclusion Générale	215
7.1	Une réanalyse des usages interrogatifs en français	217
7.2	L'acquisition des interrogatives en français : un phénomène social	220
7.3	Réflexions sur la démarche et les méthodes utilisées	222
7.4	Perspectives	223
A	Corpus EPAC : Extractions	I
B	Corpus EPAC : Statistiques	XIII
C	Corpus ESLO : Extractions	XXIX
D	Corpus ESLO : Modèles statistiques	LXXXV
E	AJT2 : Norming préalable à l'étude	XCIII
F	AJT2 : Items et modèles statistiques	CI
G	AJT3 : Norming des voix préalables à l'étude	CXVII
H	AJT3 : Matériels et modèles statistiques	CXXI
I	AJT4 : Matériels et modèles statistiques	CXXVII
J	MG2 : Création et norming des échelles utilisées dans l'étude	CXXXVII
K	MG2 : Matériel et modèles statistiques	CLIX
L	MG3 : Matériels et modèles statistiques	CCXI
M	MG enfants : Matériel et modèles statistiques	CCXXI
N	Production Enfantine : Dossier Éthique	CCXLVII
O	Production Enfantine : Corpus	CCLXIII
P	Production Enfantine : Statistiques	CCLXXIII

TABLE DES MATIÈRES

Q Liste des acronymes	CCLXXIX
R Liste des symboles	CCLXXXI
Références	a

Liste des figures

3.1	Exemple de graphique à barres représentant des proportions	44
3.2	Exemple de visualisation de la distribution de données ordinales	45
3.3	Exemple de matrice de corrélation entre variables indépendantes	46
3.4	Exemple de graphique illustrant les distributions postérieures d'un modèle brms	48
4.1	Interrogatives partielles dans EPAC par type et par thématique	58
4.2	Interrogatives partielles dans EPAC par type et rôle interactionnel	59
4.3	Distributions postérieures pour le modèle brms d'analyse de EPAC, effet de RÔLE	60
4.4	Corrélations entre critères interactionnels (EPAC)	62
4.5	Interrogatives partielles dans EPAC par type et critère de "positionnement" (EPAC)	63
4.6	Interrogatives partielles dans EPAC par type et critère de "construction" (EPAC)	65
4.7	Interrogatives partielles dans EPAC par type et critère d'"argumentation" (EPAC)	66
4.8	Interrogatives partielles dans ESLO1 vs ESLO2 (15-25 et 35-55 ans confondus)	71
4.9	Distributions postérieures pour le modèle brms (ESLO, effet de CORPUS)	72
4.10	Interrogatives partielles par tranche d'âge (ESLO 1&2 confondus)	73
4.11	Interrogatives partielles par corpus et par tranche d'âge (ESLO)	74
4.12	Distributions postérieures pour le modèle brms (ESLO, effets CORPUS*ÂGE)	75
4.13	Interrogatives partielles par contexte, ESLO 1&2 confondus	76
4.14	Distributions postérieures pour le modèle brms (ESLO, effet de CONTEXTE)	77
4.15	Interrogatives partielles par corpus et par tranche d'âge (ESLO, interviews)	79
4.16	Interrogatives partielles par tranche d'âge et contexte (ESLO 2)	80
4.17	Distributions postérieures pour le modèle brms (ESLO2, effets CONTEXTE*ÂGE)	80
5.1	Note moyenne en fonction du type interrogatif (AJT2)	92
5.2	Note moyenne en fonction du type interrogatif et de l'âge (AJT2)	93
5.3	Note moyenne en fonction du type interrogatif et du contexte (AJT2))	94
5.4	Note moyenne en fonction du type interrogatif, du contexte, et de l'âge (AJT2)	95
5.5	Distributions postérieures pour le modèle brms (AJT2)	96
5.6	Note moyenne en fonction du type interrogatif, du contexte, et de l'âge (AJT2)	98
5.7	Capture d'écran sur un item expérimental (AJT3)	100
5.8	Note moyenne en fonction du type interrogatif (AJT3, <i>bon français</i>)	102
5.9	Note moyenne en fonction du type interrogatif et de l'âge (AJT3, <i>bon français</i>)	103
5.10	Note moyenne en fonction du type interrogatif et du contexte (AJT3, <i>bon français</i>)	104
5.11	Note moyenne relative aux type interrogatif, contexte et âge (AJT3, <i>bon français</i>)	105
5.12	Distributions postérieures pour le modèle brms (AJT3, <i>bon français</i>)	106

5.13	Note moyenne en fonction du type interrogatif (AJT3, <i>français adapté</i>)	107
5.14	Note moyenne en fonction du type interrogatif et de l'âge (AJT3, <i>français adapté</i>)	109
5.15	Note moyenne en fonction du type interrogatif et du contexte (AJT3, <i>français adapté</i>)	110
5.16	Note moyenne relative aux type interrogatif, contexte et âge (AJT3, <i>français adapté</i>)	111
5.17	Distributions postérieures pour le modèle brms (AJT3, <i>français adapté</i>)	111
5.18	Capture d'écran sur un item expérimental (AJT4)	115
5.19	Note moyenne en fonction du type interrogatif (AJT4, <i>bon français</i>)	117
5.20	Note moyenne en fonction du type interrogatif et de l'âge (AJT4, <i>bon français</i>) . .	118
5.21	Note moyenne en fonction du type interrogatif et du contexte (AJT4, <i>bon français</i>)	119
5.22	Note moyenne relative aux type interrogatif, contexte et âge (AJT4, <i>bon français</i>) .	120
5.23	Distributions postérieures pour le modèle brms (AJT4, <i>bon français</i>)	121
5.24	Note moyenne en fonction du type interrogatif (AJT4, <i>français adapté</i>)	122
5.25	Note moyenne en fonction du type interrogatif et de l'âge (AJT4, <i>français adapté</i>)	123
5.26	Note moyenne en fonction du type interrogatif et du contexte (AJT4, <i>français adapté</i>)	124
5.27	Note moyenne relative aux type interrogatif, contexte et âge (AJT4, <i>français adapté</i>)	125
5.28	Distributions postérieures pour le modèle brms (AJT4, <i>français adapté</i>)	125
5.29	Échelles d'évaluation présentée aux volontaires (MG2, RICHESSE)	130
5.30	Captures d'écran en cours de passation (MG2)	131
5.31	Corrélations entre échelles (Spearman) (MG2)	133
5.32	Note moyenne en fonction du type interrogatif (MG2, ÂGE)	134
5.33	Note moyenne en fonction du type interrogatif (MG2, RICHESSE)	135
5.34	Note moyenne en fonction du type interrogatif (MG2, SOCIABILITÉ)	136
5.35	Note moyenne en fonction du type interrogatif (MG2, ORIGINE)	137
5.36	Note moyenne en fonction du type interrogatif (MG2, ÉTUDES)	138
5.37	Note moyenne en fonction du type interrogatif (MG2, LOISIRS)	139
5.38	Résultats généraux moyens pour toutes les échelles (MG2)	140
5.39	Résultats moyens pour toutes les échelles en fonction de l'âge (MG2)	141
5.40	Résultats moyens pour toutes les échelles en fonction du contexte (MG2)	142
5.41	Résultats moyens pour toutes les échelles en fonction du contexte et de l'âge (MG2)	144
5.42	Distributions postérieures pour le modèle brms (MG2, ÂGE)	146
5.43	Distributions postérieures pour le modèle brms (MG2, RICHESSE)	147
5.44	Distributions postérieures pour le modèle brms (MG2, SOCIABILITÉ)	148
5.45	Distributions postérieures pour le modèle brms (MG2, ORIGINE)	149
5.46	Distributions postérieures pour le modèle brms (MG2, ÉTUDES)	150
5.47	Distributions postérieures pour le modèle brms (MG2, LOISIRS)	151
5.48	Capture d'écran sur un item expérimental (MG3)	154
5.49	Corrélations entre échelles (Spearman) (MG3)	156
5.50	Jugement moyen en fonction du type interrogatif (MG3, ÂGE)	157
5.51	Jugement moyen en fonction du type interrogatif et de l'âge (MG3, ÂGE)	158
5.52	Jugement moyen en fonction du type interrogatif et du contexte (MG3, ÂGE) . . .	159
5.53	Distributions postérieures pour le modèle brms (MG3, ÂGE)	160
5.54	Jugement moyen en fonction du type interrogatif (MG3, ÉTUDES)	161
5.55	Jugement moyen en fonction du type interrogatif et de l'âge (MG3, ÉTUDES) . . .	162

5.56	Jugement moyen relatif aux types interrogatif, contexte et âge (MG3, ÉTUDES) . . .	163
5.57	Distributions postérieures pour le modèle brms (MG3, ÉTUDES)	164
5.58	Jugement moyen en fonction du type interrogatif (MG3, ORIGINE)	165
5.59	Jugement moyen en fonction du type interrogatif et de l'âge (MG3, ORIGINE) . . .	166
5.60	Jugement moyen en fonction du type interrogatif et du contexte (MG3, ORIGINE)	167
5.61	Jugement moyen relatif aux types interrogatif, contexte et âge (MG3, ORIGINE) . .	168
5.62	Distributions postérieures pour le modèle brms (MG3, ORIGINE)	168
5.63	Résultats généraux moyens pour toutes les échelles (MG3)	170
6.1	Captures d'écrans (Étude MG-C, Stéréotype : RICHESSE)	179
6.2	Stéréotypes choisis en fonction du type phrastique (MG-C)	181
6.3	Distributions postérieures pour le modèle brms (MG-C, effet TYPE)	182
6.4	Évaluation de la RICHESSE en fonction du type phrastique (MG-C)	183
6.5	Évaluation des LOISIRS en fonction du type phrastique (MG-C)	184
6.6	Évaluation de l'ÉDUCATION en fonction du type phrastique (MG-C)	185
6.7	Évaluation de la RICHESSE, par TYPE et CLASSE (MG-C)	187
6.8	Évaluation des LOISIRS, par TYPE et CLASSE (MG-C)	189
6.9	Évaluation de l'ÉDUCATION, par TYPE et CLASSE (MG-C)	190
6.10	Évolution des associations forme-stéréotype (MG-C)	191
6.11	Évolution du profil stéréotypique des interrogatives partielles (MG-C)	192
6.12	Marionnettes utilisées (JDR-C)	194
6.13	Environnement de jeu de rôle, marionnettes et matériel (JDR-C)	195
6.14	Exemple de transcription, JDR-C - CHI33, phases 0-32	197
6.15	Interrogatives partielles produites selon le contexte (JDR-C)	200
6.16	Interrogatives partielles produites selon la classe (JDR-C)	202
6.17	Interrogatives partielles par classe et contexte (Informel/Formel) (JDR-C)	205
6.18	Interrogatives partielles produites selon le genre de l'enfant (JDR-C)	206
6.19	Interrogatives partielles produites selon les destinataires (JDR-C)	207
6.20	Productions avec tutoiement/vouvoiement selon la marionnette (JDR-C)	208
6.21	Type interrogatif produit selon le tutoiement/vouvoiement (JDR-C)	209
6.22	Productions interrogatives par enfant (JDR-C)	211
7.1	Profils sociaux types liés à l'utilisation des variantes interrogatives (MG2-3)	219
7.2	Évolution du profil stéréotypique des interrogatives partielles (MG-C)	221
B.1	Interrogatives partielles dans EPAC par type et par thématique (tout)	XIII
B.2	Distributions postérieures pour le modèle brms d'analyse de EPAC, effet de THÉ- MATIQUE	XVI
B.3	Distributions postérieures pour le modèle brms (EPAC, effet MARQUE D'INFORMA- LITÉ)	XIX
B.4	Distributions postérieures pour le modèle brms (EPAC, effet PROXIMITÉ DESTINA- TAIRE)	XX
B.5	Distributions postérieures pour le modèle brms (EPAC, effet POSTURE)	XXI
B.6	Distributions postérieures pour le modèle brms (EPAC, effet ANIMATION)	XXII

B.7	Distributions postérieures pour le modèle brms (EPAC, effet REFORMULATION)	XXIII
B.8	Distributions postérieures pour le modèle brms (EPAC, effet AUTO-RÉPÉTITION)	XXIV
B.9	Distributions postérieures pour le modèle brms (EPAC, effet RÉPÉTITION AUTRE)	XXV
B.10	Distributions postérieures pour le modèle brms (EPAC, effet FIN ARGUMENTAIRE)	XXVI
B.11	Distributions postérieures pour le modèle brms (EPAC, effet RÉPONSE AUTO-FOURNIE)	XXVII
B.12	Distributions postérieures pour le modèle brms (EPAC, effet OPINION IMPLICITE)	XXVIII
D.1	Distributions postérieures pour le modèle brms (ESLO, effet d'ÂGE)	LXXXVII
D.2	Distributions postérieures pour le modèle brms (ESLO, interviews, CORPUS*ÂGE)	XCI
E.1	Consigne (AJT2p - Norming)	XCIV
E.2	Capture d'écran (AJT2p - Norming)	XCV
E.3	Résultats du norming (AJT2p, Entraînements)	XCVIII
E.4	Résultats du norming (AJT2p, items 1-15)	XCIX
E.5	Résultats du norming (AJT2p, items 16-30)	C
F.1	Consigne (AJT2)	CXIII
G.1	Consigne (AJT3p - Norming)	CXVIII
G.2	Capture d'écran (AJT3p - Norming)	CXIX
G.3	Résultats du norming - Genre des voix (AJT3p)	CXX
G.4	Résultats du norming - Familiarité des voix (AJT3p)	CXX
H.1	Consigne (AJT3)	CXXII
I.1	Écran de consentement (AJ4)	CXXVIII
I.2	Consigne (AJT4)	CXXIX
J.1	Consigne (MG2p - Norming)	CXXXVIII
J.2	Captures d'écran (MG2p - Norming)	CXXXIX
J.3	Item 1 (MG2p - Norming)	CXL
J.4	Résultats du norming - Item 1 (MG2p)	CXLI
J.5	Item 2 (MG2p - Norming)	CXLII
J.6	Résultats du norming - Item 2 (MG2p)	CXLIII
J.7	Item 3 (MG2p - Norming)	CXLIV
J.8	Résultats du norming - Item 3 (MG2p)	CXLV
J.9	Item 4 (MG2p - Norming)	CXLVI
J.10	Résultats du norming - Item 4 (MG2p)	CXLVII
J.11	Item 5 (MG2p - Norming)	CXLVIII
J.12	Résultats du norming - Item 5 (MG2p)	CXLIX
J.13	Item 6 (MG2p - Norming)	CL
J.14	Résultats du norming - Item 6 (MG2p)	CLI
J.15	Item 7 (MG2p - Norming)	CLII
J.16	Résultats du norming - Item 7 (MG2p)	CLIII
J.17	Item 8 (MG2p - Norming)	CLIV
J.18	Résultats du norming - Item 8 (MG2p)	CLV

J.19	Item 9 (MG2p - Norming)	CLVI
J.20	Résultats du norming - Item 9 (MG2p)	CLVII
K.1	Consigne (MG2)	CLX
K.2	Résultats moyens combinés entre toutes les échelles, par participant (MG2)	CLXXX
K.3	Note moyenne en fonction du type interrogatif et de l'âge (MG2, ÂGE)	CLXXXI
K.4	Note moyenne en fonction du type interrogatif et de l'âge (MG2, RICHESSE)	CLXXXII
K.5	Note moyenne en fonction du type interrogatif et de l'âge (MG2, SOCIABILITÉ) . .	CLXXXIII
K.6	Note moyenne en fonction du type interrogatif et de l'âge (MG2, ORIGINE)	CLXXXIV
K.7	Note moyenne en fonction du type interrogatif et de l'âge (MG2, ÉTUDES)	CLXXXV
K.8	Note moyenne en fonction du type interrogatif et de l'âge (MG2, LOISIRS)	CLXXXVI
K.9	Note moyenne en fonction du type interrogatif et du contexte (MG2, ÂGE)	CLXXXVII
K.10	Note moyenne en fonction du type interrogatif et du contexte (MG2, RICHESSE) .	CLXXXVIII
K.11	Note moyenne en fonction du type interrogatif et du contexte (MG2, SOCIABILITÉ)	CLXXXIX
K.12	Note moyenne en fonction du type interrogatif et du contexte (MG2, ORIGINE) . .	CXC
K.13	Note moyenne en fonction du type interrogatif et du contexte (MG2, ÉTUDES) . .	CXCI
K.14	Note moyenne en fonction du type interrogatif et du contexte (MG2, LOISIRS) . .	CXCII
K.15	Note moyenne en fonction des type interrogatif, âge et contexte (MG2, ÂGE)	CXCIII
K.16	Note moyenne en fonction des type interrogatif, âge et contexte (MG2, RICHESSE)	CXCIV
K.17	Note moyenne en fonction des type interrogatif, âge et contexte (MG2, SOCIABI- LITÉ)	CXCV
K.18	Note moyenne en fonction des type interrogatif, âge et contexte (MG2, ORIGINE) .	CXCVI
K.19	Note moyenne en fonction des type interrogatif, âge et contexte (MG2, ÉTUDES) .	CXCVII
K.20	Note moyenne en fonction des type interrogatif, âge et contexte (MG2, LOISIRS) .	CXCVIII
L.1	Écran de consentement (MG3)	CCXII
L.2	Consigne (MG3)	CCXIII
M.1	Échelle de l'étude MG-C. Stéréotype : RICHESSE	CCXXVIII
M.2	Échelle de l'étude MG-C. Stéréotype : ÉDUCATION	CCXXIX
M.3	Échelle de l'étude MG-C. Stéréotype : LOISIRS	CCXXX
M.4	Écrans de progression (Étude MG-C.)	CCXXXI
M.5	Diplôme de participation à l'étude MG-C	CCXXXII
M.6	Distributions postérieures pour le modèle brms (MG-C, effet TYPE)	CCXXXIII
M.7	Distributions postérieures pour le modèle brms (MG-C, stéréotype RICHESSE, effet TYPE)	CCXXXIV
M.8	Distributions postérieures pour le modèle brms (MG-C, stéréotype LOISIRS, effet TYPE)	CCXXXV
M.9	Distributions postérieures pour le modèle brms (MG-C, stéréotype ÉDUCATION, ef- fet TYPE)	CCXXXVI
M.10	Distributions postérieures pour le modèle brms (MG-C, stéréotype RICHESSE, effet TYPE*CLASSE)	CCXXXVIII
M.11	Distributions postérieures pour le modèle brms (MG-C, stéréotype LOISIRS, effet TYPE*CLASSE)	CCXL

M.12	Distributions postérieures pour le modèle brms (MG-C, stéréotype ÉDUCATION, effet TYPE*CLASSE)	CCXLII
M.13	Distributions postérieures pour un modèle brms global (MG-C)	CCXLIV
O.1	Diplôme de participation au jeu de rôle (F)	CCLXIII
O.2	Diplôme de participation au jeu de rôle (M)	CCLXIV
P.1	Distributions postérieures pour le modèle brms (JDR-C, effet de CONTEXTE)	CCLXXIV
P.2	Distributions postérieures pour le modèle brms (JDR-C, effet de CONTEXTE*CLASSE)	CCLXXVII
P.3	Distributions postérieures pour le modèle brms (JDR-C, effet de GENRE)	CCLXXVIII

Liste des tableaux

2.1	Stratégies principales d'interrogation partielle en français	14
2.2	Types d'interrogatives partielles dans des études du 20 ^{ème} siècle	15
2.3	Types d'interrogatives partielles dans le corpus <i>Là-Bas</i> , Hamlaoui (2009)	16
2.4	Types d'interrogatives partielles dans la base <i>sgs</i> , Adli (2015)	17
3.1	Exemple de tableau de valeurs brutes	44
3.2	Exemple de tableau de moyennes brutes	45
4.1	Émissions dont les transcriptions constituent le sous-corpus EPAC analysé	53
4.2	Interrogatives partielles dans EPAC par type	57
4.3	Interrogatives partielles dans EPAC par type et par thématique	57
4.4	Interrogatives partielles dans EPAC par type et rôle interactionnel	59
4.5	Interrogatives partielles par type et par critère de "positionnement" (EPAC)	62
4.6	Interrogatives partielles par type et par critère de "construction" (EPAC)	64
4.7	Interrogatives partielles par type et par critère d'"argumentation" (EPAC)	66
4.8	Interrogatives partielles dans ESLO par type (15-25 et 35-55 ans confondus)	71
4.9	Interrogatives partielles dans ESLO1 vs ESLO2 (15-25 et 35-55 ans confondus)	71
4.10	Interrogatives partielles par tranche d'âge (ESLO 1&2 confondus)	73
4.11	Interrogatives partielles par corpus et par tranche d'âge (ESLO)	74
4.12	Interrogatives partielles par contexte, ESLO 1&2 confondus	76
4.13	Interrogatives partielles par corpus et par tranche d'âge (ESLO, interviews)	78
4.14	Interrogatives partielles par corpus et par tranche d'âge (ESLO, interviews)	79
5.1	Exemple d'item expérimental (AJT1)	88
5.2	Exemple d'item expérimental (AJT2)	90
5.3	Note moyenne en fonction du type interrogatif (AJT2)	91
5.4	Note moyenne en fonction du type interrogatif et de l'âge (AJT2)	92
5.5	Note moyenne en fonction du type interrogatif et du contexte (AJT2)	93
5.6	Note moyenne en fonction du type interrogatif, du contexte, et de l'âge (AJT2)	94
5.7	Exemple d'item expérimental (AJT2-3)	99
5.8	Combinaisons de voix créées (AJT3)	100
5.9	Note moyenne en fonction du type interrogatif (AJT3, <i>bon français</i>)	102
5.10	Note moyenne en fonction du type interrogatif et de l'âge (AJT3, <i>bon français</i>)	103
5.11	Note moyenne en fonction du type interrogatif et du contexte (AJT3, <i>bon français</i>)	104
5.12	Note moyenne relative aux type interrogatif, contexte et âge (AJT3, <i>bon français</i>)	104

5.13	Note moyenne en fonction du type interrogatif (AJT3, <i>français adapté</i>)	107
5.14	Note moyenne en fonction du type interrogatif et de l'âge (AJT3, <i>français adapté</i>)	109
5.15	Note moyenne en fonction du type interrogatif et du contexte (AJT3, <i>français adapté</i>)	110
5.16	Note moyenne relative aux type interrogatif, contexte et âge (AJT3, <i>français adapté</i>)	110
5.17	Rappel d'un item expérimental (AJT2)	114
5.18	Exemple d'item expérimental (AJT4)	115
5.19	Note moyenne en fonction du type interrogatif (AJT4, <i>bon français</i>)	117
5.20	Note moyenne en fonction du type interrogatif et de l'âge (AJT4, <i>bon français</i>)	117
5.21	Note moyenne en fonction du type interrogatif et du contexte (AJT4, <i>bon français</i>)	118
5.22	Note moyenne relative aux type interrogatif, contexte et âge (AJT4, <i>bon français</i>)	119
5.23	Note moyenne en fonction du type interrogatif (AJT4, <i>français adapté</i>)	122
5.24	Note moyenne en fonction du type interrogatif et de l'âge (AJT4, <i>français adapté</i>)	123
5.25	Note moyenne en fonction du type interrogatif et du contexte (AJT4, <i>français adapté</i>)	124
5.26	Note moyenne relative aux type interrogatif, contexte et âge (AJT4, <i>français adapté</i>)	124
5.27	Récapitulatif (Jugements d'acceptabilité)	127
5.28	Exemple de différence entre interviews (MG1)	128
5.29	Exemple d'item expérimental (AJT3 - MG2)	131
5.30	Note moyenne en fonction du type interrogatif (MG2, ÂGE)	134
5.31	Note moyenne en fonction du type interrogatif (MG2, RICHESSE)	135
5.32	Note moyenne en fonction du type interrogatif (MG2, SOCIABILITÉ)	136
5.33	Note moyenne en fonction du type interrogatif (MG2, ORIGINE)	137
5.34	Note moyenne en fonction du type interrogatif (MG2, ÉTUDES)	138
5.35	Note moyenne en fonction du type interrogatif (MG2, LOISIRS)	139
5.36	Exemple d'item expérimental (AJT4-MG3)	154
5.37	Jugement moyen en fonction du type interrogatif (MG3, ÂGE)	157
5.38	Jugement moyen en fonction du type interrogatif et de l'âge (MG3, ÂGE)	158
5.39	Jugement moyen en fonction du type interrogatif et du contexte (MG3, ÂGE)	158
5.40	Jugement moyen relatif aux types interrogatif, contexte et âge (MG3, ÂGE)	159
5.41	Jugement moyen en fonction du type interrogatif (MG3, ÉTUDES)	161
5.42	Jugement moyen en fonction du type interrogatif et de l'âge (MG3, ÉTUDES)	162
5.43	Jugement moyen en fonction du type interrogatif et du contexte (MG3, ÉTUDES)	162
5.44	Jugement moyen relatif aux types interrogatif, contexte et âge (MG3, ÉTUDES)	163
5.45	Jugement moyen en fonction du type interrogatif (MG3, ORIGINE)	165
5.46	Jugement moyen en fonction du type interrogatif et de l'âge (MG3, ORIGINE)	166
5.47	Jugement moyen en fonction du type interrogatif et du contexte (MG3, ORIGINE)	166
5.48	Jugement moyen relatif aux types interrogatif, contexte et âge (MG3, ORIGINE)	167
6.1	Répartition des enfants par classes (MG-C)	180
6.2	Répartition des enfants par niveaux (MG-C)	180
6.3	Répartition des enfants par âges (MG-C)	180
6.4	Répartition des réponses retenues pour analyse par groupes-classe (MG-C)	186
6.5	Répartition des enfants par classes (JDR-C)	196
6.6	Répartition des enfants par niveaux (JDR-C)	196
6.7	Répartition des enfants par âges (JDR-C)	196

6.8	Interrogatives partielles par type (JDR-C)	199
6.9	Interrogatives partielles produites selon le contexte (JDR-C)	200
6.10	Répartition des interrogatives par âge (JDR-C)	201
6.11	Répartition des interrogatives par classe (JDR-C)	201
6.12	Interrogatives partielles produites selon la classe (JDR-C)	201
6.13	Répartition des éléments interrogatifs selon le type phrastique (JDR-C)	203
6.14	Productions de l'élément "qu'/que/quoi" selon la classe des enfants (JDR-C)	204
6.15	Interrogatives partielles par classe et contexte (<i>Informel/Formel</i>) (JDR-C)	204
6.16	Interrogatives partielles produites selon le genre de l'enfant (JDR-C)	206
6.17	Interrogatives partielles produites selon les destinataires (JDR-C)	207
6.18	Productions avec tutoiement/vouvoiement selon la marionnette (JDR-C)	208
6.19	Type interrogatif produit selon le tutoiement/vouvoiement (JDR-C)	209
A.1	Extractions du corpus EPAC retenues pour analyse	I
A.1	Extractions du corpus EPAC retenues pour analyse	II
A.1	Extractions du corpus EPAC retenues pour analyse	III
A.1	Extractions du corpus EPAC retenues pour analyse	IV
A.1	Extractions du corpus EPAC retenues pour analyse	V
A.1	Extractions du corpus EPAC retenues pour analyse	VI
A.1	Extractions du corpus EPAC retenues pour analyse	VII
A.1	Extractions du corpus EPAC retenues pour analyse	VIII
A.1	Extractions du corpus EPAC retenues pour analyse	IX
A.1	Extractions du corpus EPAC retenues pour analyse	X
A.1	Extractions du corpus EPAC retenues pour analyse	XI
A.1	Extractions du corpus EPAC retenues pour analyse	XII
B.1	Interrogatives partielles dans EPAC par type et par thématique (tout)	XIII
B.2	Résumé du modèle brms (EPAC, effet THÉMATIQUE)	XIV
B.3	Distributions postérieures pour le modèle brms (EPAC, effet THÉMATIQUE)	XV
B.4	Résumé du modèle brms (EPAC, effet RÔLE)	XVII
B.5	Distributions postérieures pour le modèle brms d'analyse pour EPAC, effet de RÔLE	XVII
B.6	Coefficients de corrélations entre critères interactionnels, méthode Spearman (EPAC)	XVIII
B.7	P-values pour les corrélations entre critères (EPAC)	XVIII
B.8	Résumé du modèle brms (EPAC, effet MARQUE D'INFORMALITÉ)	XIX
B.9	Distributions postérieures pour le modèle brms (EPAC, effet MARQUE D'INFORMALITÉ)	XIX
B.10	Résumé du modèle brms (EPAC, effet PROXIMITÉ DESTINATAIRE)	XX
B.11	Distributions postérieures pour le modèle brms (EPAC, effet PROXIMITÉ DESTINATAIRE)	XX
B.12	Résumé du modèle brms (EPAC, effet POSTURE)	XXI
B.13	Distributions postérieures pour le modèle brms (EPAC, effet POSTURE)	XXI
B.14	Résumé du modèle brms (EPAC, effet ANIMATION)	XXII
B.15	Distributions postérieures pour le modèle brms (EPAC, effet ANIMATION)	XXII
B.16	Résumé du modèle brms (EPAC, effet REFORMULATION)	XXIII

B.17 Distributions postérieures pour le modèle brms (EPAC, effet REFORMULATION) . . .	XXIII
B.18 Résumé du modèle brms (EPAC, effet AUTO-RÉPÉTITION)	XXIV
B.19 Distributions postérieures pour le modèle brms (EPAC, effet AUTO-RÉPÉTITION) . .	XXIV
B.20 Résumé du modèle brms (EPAC, effet RÉPÉTITION AUTRE)	XXV
B.21 Distributions postérieures pour le modèle brms (EPAC, effet RÉPÉTITION AUTRE) . .	XXV
B.22 Résumé du modèle brms (EPAC, effet FIN ARGUMENTAIRE)	XXVI
B.23 Distributions postérieures pour le modèle brms (EPAC, effet FIN ARGUMENTAIRE) . .	XXVI
B.24 Résumé du modèle brms (EPAC, effet RÉPONSE AUTO-FOURNIE)	XXVII
B.25 Distributions postérieures pour le modèle brms (EPAC, effet RÉPONSE AUTO-FOURNIE)	XXVII
B.26 Résumé du modèle brms (EPAC, effet OPINION IMPLICITE)	XXVIII
B.27 Distributions postérieures pour le modèle brms (EPAC, effet OPINION IMPLICITE) . .	XXVIII
C.1 Extractions du corpus ESLO retenues pour analyse	XXIX
C.1 Extractions du corpus ESLO retenues pour analyse	XXX
C.1 Extractions du corpus ESLO retenues pour analyse	XXXI
C.1 Extractions du corpus ESLO retenues pour analyse	XXXII
C.1 Extractions du corpus ESLO retenues pour analyse	XXXIII
C.1 Extractions du corpus ESLO retenues pour analyse	XXXIV
C.1 Extractions du corpus ESLO retenues pour analyse	XXXV
C.1 Extractions du corpus ESLO retenues pour analyse	XXXVI
C.1 Extractions du corpus ESLO retenues pour analyse	XXXVII
C.1 Extractions du corpus ESLO retenues pour analyse	XXXVIII
C.1 Extractions du corpus ESLO retenues pour analyse	XXXIX
C.1 Extractions du corpus ESLO retenues pour analyse	XL
C.1 Extractions du corpus ESLO retenues pour analyse	XLI
C.1 Extractions du corpus ESLO retenues pour analyse	XLII
C.1 Extractions du corpus ESLO retenues pour analyse	XLIII
C.1 Extractions du corpus ESLO retenues pour analyse	XLIV
C.1 Extractions du corpus ESLO retenues pour analyse	XLV
C.1 Extractions du corpus ESLO retenues pour analyse	XLVI
C.1 Extractions du corpus ESLO retenues pour analyse	XLVII
C.1 Extractions du corpus ESLO retenues pour analyse	XLVIII
C.1 Extractions du corpus ESLO retenues pour analyse	XLIX
C.1 Extractions du corpus ESLO retenues pour analyse	L
C.1 Extractions du corpus ESLO retenues pour analyse	LI
C.1 Extractions du corpus ESLO retenues pour analyse	LII
C.1 Extractions du corpus ESLO retenues pour analyse	LIII
C.1 Extractions du corpus ESLO retenues pour analyse	LIV
C.1 Extractions du corpus ESLO retenues pour analyse	LV
C.1 Extractions du corpus ESLO retenues pour analyse	LVI
C.1 Extractions du corpus ESLO retenues pour analyse	LVII
C.1 Extractions du corpus ESLO retenues pour analyse	LVIII
C.1 Extractions du corpus ESLO retenues pour analyse	LIX
C.1 Extractions du corpus ESLO retenues pour analyse	LX

C.1 Extractions du corpus ESLO retenues pour analyse	LXI
C.1 Extractions du corpus ESLO retenues pour analyse	LXII
C.1 Extractions du corpus ESLO retenues pour analyse	LXIII
C.1 Extractions du corpus ESLO retenues pour analyse	LXIV
C.1 Extractions du corpus ESLO retenues pour analyse	LXV
C.1 Extractions du corpus ESLO retenues pour analyse	LXVI
C.1 Extractions du corpus ESLO retenues pour analyse	LXVII
C.1 Extractions du corpus ESLO retenues pour analyse	LXVIII
C.1 Extractions du corpus ESLO retenues pour analyse	LXIX
C.1 Extractions du corpus ESLO retenues pour analyse	LXX
C.1 Extractions du corpus ESLO retenues pour analyse	LXXI
C.1 Extractions du corpus ESLO retenues pour analyse	LXXII
C.1 Extractions du corpus ESLO retenues pour analyse	LXXIII
C.1 Extractions du corpus ESLO retenues pour analyse	LXXIV
C.1 Extractions du corpus ESLO retenues pour analyse	LXXV
C.1 Extractions du corpus ESLO retenues pour analyse	LXXVI
C.1 Extractions du corpus ESLO retenues pour analyse	LXXVII
C.1 Extractions du corpus ESLO retenues pour analyse	LXXVIII
C.1 Extractions du corpus ESLO retenues pour analyse	LXXIX
C.1 Extractions du corpus ESLO retenues pour analyse	LXXX
C.1 Extractions du corpus ESLO retenues pour analyse	LXXXI
C.1 Extractions du corpus ESLO retenues pour analyse	LXXXII
C.1 Extractions du corpus ESLO retenues pour analyse	LXXXIII
C.1 Extractions du corpus ESLO retenues pour analyse	LXXXIV
D.1 Résumé du modèle brms (ESLO, effet CORPUS)	LXXXVI
D.2 Distributions postérieures pour le modèle brms (ESLO, effet CORPUS)	LXXXVI
D.3 Résumé du modèle brms (ESLO, effet ÂGE)	LXXXVII
D.4 Distributions postérieures pour le modèle brms (ESLO, ÂGE)	LXXXVII
D.5 Résumé du modèle brms (ESLO, effets CORPUS*ÂGE)	LXXXVIII
D.6 Distributions postérieures pour le modèle brms (ESLO, CORPUS*ÂGE)	LXXXVIII
D.7 Résumé du modèle brms (ESLO, effet CONTEXTE)	LXXXIX
D.8 Distributions postérieures pour le modèle brms (ESLO, CONTEXTE)	LXXXIX
D.9 Résumé du modèle brms (ESLO, interviews, effets CORPUS*ÂGE)	XC
D.10 Distributions postérieures pour le modèle brms (ESLO, interviews, CORPUS*ÂGE)	XC
D.11 Résumé du modèle brms (ESLO 2, effets ÂGE*CONTEXTE)	XCII
D.12 Distributions postérieures pour le modèle brms (ESLO2, ÂGE*CONTEXTE)	XCII
E.1 Items expérimentaux pour le norming préalable à l'étude AJT2 (contrôles 101-103)	XCV
E.2 Items expérimentaux pour le norming préalable à l'étude AJT2 (cibles 1-10)	XCV
E.3 Items expérimentaux pour le norming préalable à l'étude AJT2 (cibles 11-30)	XCVI
E.4 Résultats du norming (AJT2p, items 1-20)	XCVII
E.5 Résultats du norming (AJT2p, items 21-30 - Entraînements)	XCVIII

E.1	Items expérimentaux pour l'étude AJT2 (cibles 001-003)	CI
E.2	Items expérimentaux pour l'étude AJT2 (cibles 003-006)	CII
E.3	Items expérimentaux pour l'étude AJT2 (cibles 007-009)	CIII
E.4	Items expérimentaux pour l'étude AJT2 (cibles 0010-012)	CIV
E.5	Items expérimentaux pour l'étude AJT2 (cibles 013-015)	CV
E.6	Items expérimentaux pour l'étude AJT2 (cibles 016-018)	CVI
E.7	Items expérimentaux pour l'étude AJT2 (cibles 019-021)	CVII
E.8	Items expérimentaux pour l'étude AJT2 (cibles 022-024)	CVIII
E.9	Items expérimentaux pour l'étude AJT2 (cibles 025-027)	CIX
E.10	Items expérimentaux pour l'étude AJT2 (cibles 028-030)	CX
E.11	Items expérimentaux pour l'étude AJT2 (distracteurs 041-060)	CXI
E.12	Items expérimentaux pour l'étude AJT2 (contrôles 031-040 et entraînements 101-103)	CXII
E.13	Résumé du modèle brms d'analyse de AJT2	CXIV
E.14	Distributions postérieures pour le modèle brms maximal d'analyse de AJT2	CXV
G.1	Résultats du norming - Genre des voix (AJT3p)	CXIX
G.2	Résultats du norming - Familiarité des voix (AJT3p)	CXX
H.1	Résumé du modèle brms d'analyse de AJT3 - Bon français	CXXIII
H.2	Distributions postérieures pour le modèle brms d'analyse de AJT3 - Bon français	CXXIV
H.3	Résumé du modèle brms d'analyse de AJT3 - Français adapté	CXXV
H.4	Distributions postérieures pour le modèle brms d'analyse de AJT3 - Français adapté	CXXVI
I.1	Distracteurs pour l'étude AJT4 (031-040)	CXXX
I.2	Distracteurs pour l'étude AJT4 (051-062)	CXXXI
I.3	Distracteurs pour l'étude AJT4 (063-074)	CXXXII
I.4	Résumé du modèle brms (AJT4, <i>bon français</i>)	CXXXIII
I.5	Distributions postérieures pour le modèle brms (AJT4, <i>bon français</i>)	CXXXIV
I.6	Résumé du modèle brms (AJT4, <i>français adapté</i>)	CXXXV
I.7	Distributions postérieures pour le modèle brms (AJT4, <i>français adapté</i>)	CXXXVI
J.1	Résultats du norming - Item 1 (MG2p)	CXL
J.2	Résultats du norming - Item 2 (MG2p)	CXLII
J.3	Résultats du norming - Item 3 (MG2p)	CXLIV
J.4	Résultats du norming - Item 4 (MG2p)	CXLVI
J.5	Résultats du norming - Item 5 (MG2p)	CXLVIII
J.6	Résultats du norming - Item 6 (MG2p)	CL
J.7	Résultats du norming - Item 7 (MG2p)	CLII
J.8	Résultats du norming - Item 8 (MG2p)	CLIV
J.9	Résultats du norming - Item 9 (MG2p)	CLVI
K.1	Coefficients de corrélations entre échelles, méthode Spearman (MG2)	CLXXIX
K.2	P-values pour les corrélations entre échelles, méthode Spearman (MG2)	CLXXX
K.3	Note moyenne en fonction du type interrogatif et de l'âge (MG2, ÂGE)	CLXXXI

K.4	Note moyenne en fonction du type interrogatif et de l'âge (MG2, RICHESSE)	CLXXXII
K.5	Note moyenne en fonction du type interrogatif et de l'âge (MG2, SOCIABILITÉ) . .	CLXXXIII
K.6	Note moyenne en fonction du type interrogatif et de l'âge (MG2, ORIGINE)	CLXXXIV
K.7	Note moyenne en fonction du type interrogatif et de l'âge (MG2, ÉTUDES)	CLXXXV
K.8	Note moyenne en fonction du type interrogatif et de l'âge (MG2, LOISIRS)	CLXXXVI
K.9	Note moyenne en fonction du type interrogatif et du contexte (MG2, ÂGE)	CLXXXVII
K.10	Note moyenne en fonction du type interrogatif et du contexte (MG2, RICHESSE) .	CLXXXVIII
K.11	Note moyenne en fonction du type interrogatif et du contexte (MG2, SOCIABILITÉ)	CLXXXIX
K.12	Note moyenne en fonction du type interrogatif et du contexte (MG2, ORIGINE) . .	CXC
K.13	Note moyenne en fonction du type interrogatif et du contexte (MG2, ÉTUDES) . .	CXCI
K.14	Note moyenne en fonction du type interrogatif et du contexte (MG2, LOISIRS) . .	CXCII
K.15	Note moyenne en fonction des type interrogatif, âge et contexte (MG2, ÂGE)	CXCIII
K.16	Note moyenne en fonction des type interrogatif, âge et contexte (MG2, RICHESSE)	CXCIV
K.17	Note moyenne en fonction des type interrogatif, âge et contexte (MG2, SOCIABI- LITÉ)	CXCV
K.18	Note moyenne en fonction des type interrogatif, âge et contexte (MG2, ORIGINE) .	CXCVI
K.19	Note moyenne en fonction des type interrogatif, âge et contexte (MG2, ÉTUDES) .	CXCVII
K.20	Note moyenne en fonction des type interrogatif, âge et contexte (MG2, LOISIRS) .	CXCVIII
K.21	Résumé du modèle brms (MG2, ÂGE)	CXCIX
K.22	Distributions postérieures pour le modèle brms (MG2, ÂGE)	CC
K.23	Résumé du modèle brms (MG2, RICHESSE)	CCI
K.24	Distributions postérieures pour le modèle brms (MG2, RICHESSE)	CCII
K.25	Résumé du modèle brms (MG2, SOCIABILITÉ)	CCIII
K.26	Distributions postérieures pour le modèle brms (MG2, SOCIABILITÉ)	CCIV
K.27	Résumé du modèle brms (MG2, ORIGINE)	CCV
K.28	Distributions postérieures pour le modèle brms (MG2, ORIGINE)	CCVI
K.29	Résumé du modèle brms (MG2, ÉTUDES)	CCVII
K.30	Distributions postérieures pour le modèle brms (MG2, ÉTUDES)	CCVIII
K.31	Résumé du modèle brms (MG2, LOISIRS)	CCIX
K.32	Distributions postérieures pour le modèle brms (MG2, LOISIRS)	CCX
L.1	Coefficients de corrélations entre échelles, méthode Spearman (MG3)	CCXIII
L.2	P-values pour les corrélations entre échelles, méthode Spearman (MG3)	CCXIII
L.3	Résumé du modèle brms (MG3, ÂGE)	CCXIV
L.4	Distributions postérieures pour le modèle brms (MG3, ÂGE)	CCXV
L.5	Résumé du modèle brms (MG3, ÉTUDES)	CCXVI
L.6	Distributions postérieures pour le modèle brms (MG3, ÉTUDES)	CCXVII
L.7	Résumé du modèle brms (MG3, ORIGINE)	CCXVIII
L.8	Distributions postérieures pour le modèle brms (MG3, ORIGINE)	CCXIX
M.1	Nombre de passations par liste par classe (MG-C)	CCXXI
M.2	Item d'entraînement pour toutes les listes (MG-C)	CCXXI
M.3	Items expérimentaux - Liste 1 (MG-C)	CCXXII
M.4	Items expérimentaux - Liste 2 (MG-C)	CCXXIII

M.5	Items expérimentaux - Liste 3 (MG-C)	CCXXIV
M.6	Items expérimentaux - Liste 4 (MG-C)	CCXXV
M.7	Items expérimentaux - Liste 5 (MG-C)	CCXXVI
M.8	Items expérimentaux - Liste 6 (MG-C)	CCXXVII
M.9	Résumé du modèle brms (MG-C, effet TYPE)	CCXXXIII
M.10	Distributions postérieures pour le modèle brms (MG-C, effet TYPE)	CCXXXIII
M.11	Résumé du modèle brms (MG-C, stéréotype RICHESSE, effet TYPE)	CCXXXIV
M.12	Distributions postérieures pour le modèle brms (MG-C, stéréotype RICHESSE, effet TYPE)	CCXXXIV
M.13	Résumé du modèle brms (MG-C, stéréotype LOISIRS, effet TYPE)	CCXXXV
M.14	Distributions postérieures pour le modèle brms (MG-C, stéréotype LOISIRS, effet TYPE)	CCXXXV
M.15	Résumé du modèle brms (MG-C, stéréotype ÉDUCATION, effet TYPE)	CCXXXVI
M.16	Distributions postérieures pour le modèle brms (MG-C, stéréotype ÉDUCATION, effet TYPE)	CCXXXVI
M.17	Résumé du modèle brms (MG-C, stéréotype RICHESSE, effet TYPE*CLASSE)	CCXXXVII
M.18	Distributions postérieures du modèle brms (MG-C, stéréotype RICHESSE, effet TYPE*CLASSE)	CCXXXVII
M.19	Résumé du modèle brms (MG-C, stéréotype LOISIRS, effet TYPE*CLASSE)	CCXXXIX
M.20	Distributions postérieures du modèle brms (MG-C, stéréotype LOISIRS, effet TYPE*CLASSE)	CCXXXIX
M.21	Résumé du modèle brms (MG-C, stéréotype ÉDUCATION, effet TYPE*CLASSE)	CCXLI
M.22	Distributions postérieures du modèle brms (MG-C, stéréotype ÉDUCATION, effet TYPE*CLASSE)	CCXLI
M.23	Résumé du modèle brms global (MG-C)	CCXLIII
M.24	Distributions postérieures pour un modèle brms global (MG-C)	CCXLIV
O.1	Métadonnées du corpus de Jeux de rôles - CHI 01-40	CCLXV
O.2	Métadonnées du corpus de Jeux de rôles - CHI 41-69	CCLXVI
O.3	Métadonnées du corpus de Jeux de rôles - Global	CCLXVI
O.4	Interrogatives du Jeu de Rôle enfantin retenues pour analyse	CCLXVII
O.4	Interrogatives du Jeu de Rôle enfantin retenues pour analyse	CCLXVIII
O.4	Interrogatives du Jeu de Rôle enfantin retenues pour analyse	CCLXIX
O.4	Interrogatives du Jeu de Rôle enfantin retenues pour analyse	CCLXX
O.4	Interrogatives du Jeu de Rôle enfantin retenues pour analyse	CCLXXI
O.4	Interrogatives du Jeu de Rôle enfantin retenues pour analyse	CCLXXII
P.1	Résumé du modèle brms (JDR-C, effet de CONTEXTE)	CCLXXIV
P.2	Distributions postérieures pour le modèle brms (JDR-C, effet de CONTEXTE)	CCLXXIV
P.3	Résumé du modèle brms (JDR-C, effet de CLASSE*CONTEXTE)	CCLXXV
P.4	Distributions postérieures pour le modèle brms (JDR-C, effet de CONTEXTE*CLASSE)	CCLXXVI
P.5	Résumé du modèle brms (JDR-C, effet de GENRE)	CCLXXVIII
P.6	Distributions postérieures pour le modèle brms (JDR-C, effet de GENRE)	CCLXXVIII

Chapitre 1

Introduction Générale

Sommaire

1.1 Origines de la réflexion	3
1.2 Objectifs	4
1.3 Cadre de la réflexion et moyens d'investigation	5
1.4 Résultats généraux	5
1.4.1 Chez l'adulte	6
1.4.2 Chez l'enfant	6
1.5 Plan	7

Ce chapitre pose les bases générales de la réflexion présentée dans ce travail.

1.1 Origines de la réflexion

Des quatre actes de langage principaux (assertion, interrogation, exclamation, injonction) dégagés par exemple dans [BEYSSADE et MARANDIN, 2006; GINZBURG et SAG, 2000; MARANDIN, 2018], l'interrogation est celui qui illustre peut-être mieux la dimension dialogique du langage puisqu'il invite en général à une réponse. Pour l'interrogation partielle en particulier, l'éventail offert aux francophones est large et les exemples (1-7) illustrent la diversité formelle des phrases interrogatives en français, sans même prétendre à l'exhaustivité (voir COVENEY [2011]; DELAVEAU et al. [2021] pour des panoramas plus complets). Le présent travail se base sur l'hypothèse qu'il serait particulièrement étonnant de pouvoir réaliser un même acte de langage avec autant de variété sans que certaines nuances de sens ne soient convoyées.

- (1) Tu viens **quand**?
- (2) **Quand** tu viens?
- (3) **Quand** viens-tu?
- (4) **Quand** est-ce que tu viens?
- (5) C'est **quand** que tu viens?
- (6) **Quand** que tu viens?
- (7) **Quand** c'est-y que tu viens?

Chez l'adulte, cette alternance de formes a été décrite et explorée sous plusieurs angles théoriques. La notion de situation informationnelle, et de position de focus ou de topic de l'élément interrogatif, a par exemple été invoquée [ABEILLÉ et al., 2012; BEYSSADE, 2006; BOECKX, 1999; DÉPREZ et al., 2013; HAMLAOUI, 2011]. Des analyses par la phonotaxe ont aussi été avancées [HAMLAOUI, 2009b], avec des corrélations entre le ratio en nombre de syllabes du syntagme interrogatif et du reste de la phrase pour expliquer le positionnement de l'élément interrogatif (antéposition ou non). Des approches sociolinguistiques du phénomène l'ont par ailleurs présenté comme un phénomène sociolectal, avec des groupes sociaux utilisant plus telle ou telle variante [ASHBY, 1977; BEHNSTEDT, 1973; POHL et STRAKA, 1965; QUILLARD, 2001; SÖLL, 1983; TERRY, 1970]. Enfin, des analyses évoquent une évolution diachronique de l'emploi des formes d'interrogation partielle et un changement de leur(s) valeur(s) pragmatiques [LARRIVÉE, 2016, 2019]. Ce bref tableau laisse déjà envisager le besoin d'approcher le phénomène d'alternance sous un angle multi-factoriel.

Il faut cependant noter que peu de ces travaux évitent le prisme de l'équivalence sémantique entre les différentes interrogatives partielles, avec un seul contenu propositionnel, partiellement spécifié mais dont la variable à spécifier est unique et identique à travers toutes les formes. (Dans les exemples (1-7) ci-dessus, le but est ainsi toujours de recueillir l'*information temporelle* spécifiant le moment de l'arrivée du ou de la destinataire).

Il est notable aussi que la plupart des travaux menés en acquisition du langage, constatant certes l'optionnalité linguistique présente dans l'input interrogatif que les enfants francophones reçoivent, et traitant chaque structure interrogative distinctement les unes des autres, ne le font généralement que pour observer et tenter d'établir un ordre d'acquisition [HAMANN, 2000; HULK,

1996; JAKUBOWICZ et STRIK, 2008; SCHEIDNES et TULLER, 2010; ZUCKERMAN et HULK, 2001]. Les explications sont alors souvent d'ordre syntaxique et computationnel, notamment sous l'influence des grammaires génératives et des théories du mouvement de constituants : certaines formes étant dérivées d'une forme canonique, plus le processus de dérivation comprend d'étapes et plus l'enfant aura du mal à acquérir la forme interrogative en question (voir par exemple JAKUBOWICZ [2011] et l'échelle de complexité dérivationnelle). Ce type d'analyse pèse autant sur les études de corpus que sur les travaux plus expérimentaux. Ici, non seulement les différents facteurs influant sur les productions adultes sont rarement mentionnés, mais l'hypothèse plus large que des sens différents puissent être associés aux différentes formes disponibles n'est également pas intégrée aux théories.

1.2 Objectifs

Il apparaît donc nécessaire de nuancer la vision générale actuelle sur l'interrogation partielle en français et sur son acquisition, et l'objectif est double avec deux questions principales :

1. À quel point les francophones adultes projettent-ils des sens différents sur chacune des variantes interrogatives auxquelles ils ont accès?
2. Ces considérations peuvent-elles aussi influencer sur les productions enfantines précoces?

Au niveau le plus général, il s'agit donc de resituer dans une perspective multifactorielle les travaux analysant la variation dans les interrogatives partielles du français. Des objectifs plus immédiats sont :

- explorer cette variation par une méthodologie alliant données empiriques naturelles et systématisation expérimentale.
- étudier l'évolution de cette variation, notamment en réactualisant les analyses de corpus des années 1960, par une confrontation avec des corpus plus contemporains (années 2000 et plus).
- analyser cette variation en intégrant la donnée CONTEXTE, au sens le plus large (confrontation des usages interrogatifs dans différents types interactionnels comme l'entretien semi-contrôlé ou l'interaction libre), et dans un sens plus restreint (comparaison de contextes formels et informels).
- élargir la vision sociolinguistique de cette variation, en intégrant des paramètres liés à une vision du langage très utilitariste : la langue est un outil social et la coexistence de formes linguistiques différentes permet ou répond à un besoin de véhiculer des sens différents.
- intégrer ces données à l'étude de l'acquisition des interrogatives partielles par l'enfant francophone, notamment en développant des protocoles expérimentaux adaptés.

Une étude sur l'acquisition des phrases interrogatives prépositionnelles [THIBERGE, 2017], menée sur le sous-corpus « parisien » du projet CHILDES [MORGENSTERN et PARISSÉ, 2007], illustre déjà la possible influence de la complexité syntaxique interne des éléments interrogatifs sur leur position dans la phrase. Plus précisément, les interrogatives avec préposition (*Tu veux que je donne ça à qui?*) ont un comportement un peu différent des interrogatives sans préposition, avec moins d'antéposition de l'élément interrogatif autant chez l'adulte que chez l'enfant. En parallèle, une première série d'expériences menée en 2017-2018 [THIBERGE, 2018] a aussi mis en lumière des influences sociolinguistiques sur les préférences des adultes francophones en matière

d'interrogative partielle. Des effets d'âge et de contexte ont ainsi été mis au jour : différentes personnes utilisent différentes stratégies interrogatives et l'âge peut être un bon prédicteur de ces différences. Par ailleurs la thématique discutée dans l'interaction influe aussi sur les formes interrogatives dont on attend l'utilisation. En outre, les francophones projettent des caractéristiques d'ordre sociologique (richesse, éducation...) sur les personnes utilisant telle structure interrogative plutôt qu'une autre. Ces travaux servent de base aux études de corpus menées dans le cadre de cette thèse, ainsi que pour les protocoles expérimentaux mis en place pour explorer la variation de façon parallèle chez l'adulte et chez l'enfant.

1.3 Cadre de la réflexion et moyens d'investigation

L'approche privilégiée est résolument empirique. L'échange linguistique est pris comme nécessairement marqué par le caractère interactionnel des énoncés produits et des personnes en interaction, leurs enjeux sociaux et leurs biais respectifs. Resituer le fait linguistique dans sa dimension utilitariste implique d'intégrer à celui-ci des éléments comme le statut des personnes intervenant, leurs enjeux au moment de l'interaction, leurs buts, ou encore leurs stratégies communicationnelles et sociales. Les théories mentionnées seront confrontées aux données, et non l'inverse, où les données seraient présentées sous une perspective théorique plutôt que sous une autre, par exemple une perspective syntaxique, du développement cognitif, etc.

Cette perspective implique des considérations sociolinguistiques aussi bien de type « Première » que « Seconde » ou même « Troisième » vague [ECKERT, 2008]. Elle s'accorde donc aussi bien sur l'existence de sociolectes propres à des groupes sociaux que sur l'existence de stéréotypes justifiant le recours par un ou des individus à une variante linguistique pour modifier leur propre persona ou masque social [OCHS, 1992], au cours d'une même interaction. Ces modifications ont pour origine les enjeux stratégiques de l'interaction sociale [BURNETT, 2017; GOODMAN et LASSITER, 2015; LEWIS, 1969].

Pour servir au mieux cette démarche, il est nécessaire de combiner plusieurs méthodologies, et les résultats présentés sont aussi bien des données empiriques disponibles en corpus (majoritairement spontanés) que des données tirées de protocoles expérimentaux. Une systématisation crédible des résultats obtenus repose sur des statistiques inférentielles robustes. Le choix s'est ici porté sur des analyses bayésiennes, lesquelles permettent de tester directement l'hypothèse en jeu et non de raisonner par contraste avec une *hypothèse nulle*, mais aussi de s'abstraire de la notion de *significativité* pour une caractérisation plus fine des effets décrits.

1.4 Résultats généraux

Il ressort de l'ensemble des analyses présentées que la production d'interrogatives partielles sous des formes variables en français résulte d'une interaction de facteurs linguistiques, comprenant des facteurs phonotactiques, syntaxiques, pragmatiques, mais aussi des facteurs socio-pragmatiques comme le montrera ce travail de thèse. Cette interaction vaut chez l'adulte comme chez l'enfant.

1.4.1 Chez l'adulte

La confrontation de données de multiples corpus et des résultats de protocoles expérimentaux variés (jugements d'acceptabilité, *matched-guise*, sous la modalité écrite autant qu'orale) a illustré l'influence de trois types de facteurs sur les choix faits par les adultes francophones :

- *interactionnels* : des stratégies argumentatives ou oratoires, ou encore de positionnement réciproque entre personnes intervenant dans une situation de communication, modifient les préférences en matière de production d'interrogatives partielles. Les francophones emploient ainsi plus d'interrogatives *in situ* (exemple (1) ci-dessus) lorsque la réponse à la question est explicitement donnée immédiatement après celle-ci (par la personne qui pose la question elle-même). Plus d'interrogatives *in situ* sont aussi utilisées en complément de stratégies de rapprochement du ou de la destinataire dans l'échange.
- *sociaux* : une interaction des facteurs diachroniques (effet d'époques) et générationnels (groupes d'âge au sein d'une même époque) affecte les préférences des francophones autant en production spontanée que dans leur appréciation des variantes interrogatives. Les interrogatives avec antéposition et inversion (exemple (3) ci-dessus) sont ainsi moins employées dans les années 2010 qu'elles ne l'étaient dans les années 1960, et globalement les formes *in situ* sont plus fréquentes aujourd'hui que les différents types d'antéposition. Les antépositions avec inversion sont aussi, de manière générale, plus systématiquement jugées un meilleur français par les personnes âgées de plus de 30 ans que par celles de moins de 30 ans.
- *sociopragmatiques* : la formalité du contexte joue un rôle non-négligeable sur la production d'interrogatives, mais aussi sur la vision qu'en ont les francophones, et module l'influence des facteurs précédents. En particulier, l'antéposition avec inversion est par exemple associée à des contextes plus formels que les autres variantes.

Non seulement des groupes différents au sein de la population francophone posent des regards différents sur les variantes interrogatives qui leur sont disponibles, mais il y a aussi une forte projection d'indices sociolinguistiques sur les personnes utilisant chacune des variantes (richesse, éducation, fréquente lecture, par exemple, pour les antépositions avec inversion). Ces stéréotypes sociaux doivent être pris en compte dans la réanalyse des travaux déjà menés sur l'interrogation en français, et dans l'analyse des résultats obtenus à l'avenir sur ce sujet. Cette observation pourrait sans doute être élargie à d'autres travaux portant sur des phénomènes d'apparente optionnalité linguistique.

1.4.2 Chez l'enfant

En matière d'acquisition de la variation sociolinguistique, les protocoles mis en oeuvre pour ce travail (*matched-guise*, jeu de rôles) ont permis d'explorer le développement des stéréotypes sociolinguistiques associés à chacune des principales variantes utilisées par les francophones adultes, chez des enfants de 3 à 11 ans. Ils ont aussi permis de récolter des données de production spontanée, mais contrôlée, d'interrogatives partielles par des enfants de 4 à 9 ans.

Un développement précoce des stéréotypes sociaux est visible, avec une variabilité entre les stéréotypes considérés (richesse, éducation, loisirs) et une évolution de ceux-ci au cours de l'enfance, en lien avec l'input linguistique reçu. Les personnes utilisant telle ou telle variante syntaxique

pour l'interrogation partielle ont un profil social différent pour des enfants de 3 ans ou pour des enfants de 11 ans. Les données montrent cependant que des associations entre forme linguistique et stéréotype social existent déjà même chez les plus jeunes (par exemple l'antéposition avec inversion correspond plus, de façon constante, au parler d'une personne médecin qu'à celui d'une personne caissière). C'est une donnée importante qui peut influencer sur les productions précoces d'interrogatives.

En production justement, les données recueillies illustrent avant tout une très grande variabilité interindividuelle entre les enfants sur le plan du développement linguistique et psychosocial. Le corpus constitué montre néanmoins que cette variabilité se lisse avec le temps. Les enfants de 8-9 ans montrent ainsi une intégration plus grande des critères de formalité du contexte sur le choix des formes linguistiques (tels qu'observés dans la grammaire adulte), que les enfants plus jeunes.

Ces résultats plaident pour une forte influence du facteur – ou des facteurs – sociolinguistique.s dans les comportements linguistiques en matière d'optionnalité, et ce très tôt dans le développement linguistique des individus.

1.5 Plan

Le Chapitre 2 synthétise les théories actuelles sur la variation présente dans les interrogatives partielles en français, ainsi que les données déjà dégagées sur son acquisition par les enfants francophones. Après une brève présentation des méthodes d'exploration linguistique et d'analyse statistiques retenues (Chapitre 3), les chapitres suivants complètent ce panorama. Il s'agit notamment de l'apport de nouvelles données de corpus adultes oraux (Chapitre 4), mais aussi d'autres données issues de protocoles expérimentaux et récoltées auprès d'adultes francophones (Chapitre 5). Le Chapitre 6 présente des données d'acquisition récoltées également par l'expérience, avec des tâches d'interprétation et de production, auprès d'enfants francophones (de 3 à 11 ans). Le Chapitre 7, conclusif, dresse un bilan du chemin parcouru et encore à parcourir sur le sujet.

Chapitre 2

L'interrogation partielle et son acquisition, des faits sociaux

Sommaire

2.1 L'interrogation comme objet d'étude sociolinguistique	11
2.1.1 Interrogation totale et interrogation partielle	11
2.1.1.1 Les interrogatives alternatives	12
2.1.1.2 L'interrogation en subordonnée	12
2.1.2 L'interrogation partielle en français	13
2.1.2.1 Nomenclatures sur le français	13
2.1.2.2 Fréquence d'usage et données de corpus	15
2.2 Analyses de l'interrogation partielle en français	17
2.2.1 Les théories en syntaxe formelle	18
2.2.2 Les théories liées à la structure informationnelle	19
2.2.2.1 Structure de l'information et oppositions entre phrases in/ex situ	19
2.2.2.2 Une influence de la phonotaxe?	21
2.2.3 Les théories sociolinguistiques	22
2.2.3.1 Une vision diachronique	22
2.2.3.2 Une vision diatopique	23
2.2.3.3 Une vision diastratique	24
2.3 Pour une vision plus utilitaire de la variation linguistique	25
2.3.1 Enjeux sociaux du langage et vagues de la sociolinguistique	26
2.3.2 Dégager les aspects sociaux de l'interrogation partielle en français	29
2.3.2.1 Interrogation partielle et contexte interactionnel	29
2.3.2.2 Interrogation partielle et stéréotypes sociaux	29
2.4 Pour une étude sociale de l'acquisition des interrogatives partielles	30
2.4.1 Des théories syntaxiques dominées par la notion de complexité structurelle	30
2.4.1.1 Variation dans les interrogatives prépositionnelles du français	32
2.4.2 L'hypothèse de la diglossie	32
2.4.3 Un développement parallèle des capacités linguistiques et sociales	33
2.4.3.1 Acquisition du langage, théorie de l'usage et théorie de l'esprit	33
2.4.3.2 Développement sociolinguistique, saillance et stéréotypes	34

2.4.4 Un appariement progressif entre forme interrogative et charge sociale? . . .	36
2.5 Bilan et résumé des axes de recherche	36

Ce chapitre argumente pour une analyse du langage et de la variation linguistique en tant qu'outils sociaux. En particulier, il est pertinent de tenir compte des aspects utilitaires du langage dans l'étude d'un point d'alternance linguistique comme l'interrogation partielle en français. Ce phénomène a souvent été analysé sous le prisme de formalismes linguistiques qui n'intégraient pas l'impact de la dimension sociale du langage, autant sur les productions linguistiques qu'en réception et en compréhension des énoncés.

Il ne s'agit pas de nier la valeur des facteurs déjà dégagés, qui influencent les choix et ressentis des francophones sur les phrases interrogatives. Pour autant, il semble nécessaire d'élargir la liste de ces facteurs pour construire une vision plus globale de ce phénomène. En outre, la réanalyse des données adultes sur cette thématique doit mener à un enrichissement des études conduites sur l'acquisition des interrogatives chez l'enfant.

2.1 L'interrogation comme objet d'étude sociolinguistique

Comme mentionné dans le Chapitre 1, l'interrogation est l'acte de langage par excellence pour approcher la dimension sociale du langage. Des quatre principaux (assertion, interrogation, exclamation, injonction) dégagés par exemple dans [BEYSSADE et MARANDIN, 2006; GINZBURG et SAG, 2000; MARANDIN, 2018], c'est le seul qui invite en effet explicitement à une réponse. Ce sont au moins deux individus qui sont nécessaires, de fait, à la menée complète d'un acte interrogatif : la personne posant la question d'une part, et son ou sa destinataire d'autre part, qui doit y répondre. Même dans le cas particulier de l'interrogation rhétorique, ou une réponse de l'allocutaire n'est en principe pas nécessaire, la possibilité en est en fait toujours ouverte comme en (8).

(8) A : Tu ne trouves pas que tu exagères?!

B : Si.

Le présent travail est plus précisément centré sur l'interrogation partielle, dont il convient de préciser la spécificité à la fois en tant qu'acte de langage mais à la fois comme type phrastique.

2.1.1 Interrogation totale et interrogation partielle

Au niveau de l'acte de langage, deux grands types d'interrogation peuvent être distingués.

L'interrogation peut être *totale*. La personne posant une question totale énonce un fait, une proposition logique complète, et l'acte interrogatif porte sur la valeur de vérité de cette proposition logique : l'auditoire peut répondre que cette valeur de vérité est soit « Vrai » soit « Faux », ou encore qu'elle est inconnue. En (9) les seules réponses possibles sont : « oui », « non », « peut-être » ou « je ne sais pas », ou leurs équivalents.

(9) Tu viens demain?

L'interrogation peut aussi être *partielle*. La personne posant une question partielle ne peut pas énoncer une proposition logique complète : il lui manque une information sur le monde pour la compléter et l'acte interrogatif vise ici à combler ce manque. L'auditoire peut remplir ce manque, de façon pertinente ou erronée, ou dire que l'information est inconnue. L'information manquante est remplacée dans l'énoncé linguistique par un mot interrogatif (dit aussi *mot QU-* ou *élément Wh-*), dont les exemples (10-13) ne sont que quelques illustrations en français.

- (10) Tu viens **quand**?
- (11) **Où** pars-tu?
- (12) **Qui** est-ce que tu as vu?
- (13) **Comment** on fait?

Les informations manquantes gravitent en général autour d'un verbe. Au niveau de la phrase, le syntagme correspondant à cette information manquante peut avoir des fonctions variées, indiquant un lien de dépendance essentiel avec celui-ci comme SUJET (14), COMPLÉMENT DIRECT (15) ou COMPLÉMENT OBLIQUE (16)), ou bien encore être un AJOUT (17).

- (14) **Qui** vient?
- (15) Tu vois **quoi**?
- (16) Tu vas **où**?
- (17) Tu vis ici **depuis combien de temps**?

2.1.1.1 Les interrogatives alternatives

Un troisième type d'interrogation doit être mentionné, qui consiste à proposer plusieurs alternatives dans une seule phrase. On parle d'interrogative « alternative ». Ces phrases se construisent comme les interrogatives totales, mais comportent une conjonction. À la différence des interrogatives totales cependant, ces phrases n'attendent pas de réponse en termes de valeur de vérité (« oui » ou « non »), mais elles demandent à leur desintataire de choisir parmi les options proposées (18-19) (exemples repris de DELAVEAU et al. [2021]). En ce sens, on peut aussi les rapprocher des interrogatives partielles : on peut d'ailleurs les reformuler sous cette forme avec une proposition explicite des réponses (20).

- (18) Est-ce qu'on part ou est-ce qu'on ne part pas?
- (19) Viendrez-vous à la remise du prix, ou préférez-vous rester chez vous?
- (20) Que préférez-vous : venir à la remise du prix ou rester chez vous?

Leur usage est cependant plus limité puisque l'ensemble des réponses possibles est fourni par la personne même qui recourt à ce type d'interrogation. L'information manquante est donc d'une nature un peu différente que pour l'interrogation partielle.

2.1.1.2 L'interrogation en subordonnée

Les interrogatives totales, partielles et alternatives peuvent être insérées dans d'autres phrases dans un rapport syntaxique de subordination (21-23) (exemples repris de DELAVEAU et al. [2021]).

- (21) Je me demande si Paul viendra.
- (22) Je me demande quand Paul viendra.
- (23) Je me demande si Paul prend du thé ou s'il préfère du café?

Ces phrases sont des subordonnées complément du verbe qui les introduit : on peut les pronominaliser par une proforme à l'accusatif (24). Elles servent parfois à rapporter un discours et on parle alors souvent d'interrogatives « indirectes » (25). En ce qu'elles sont un complément du verbe de la phrase racine, elles ne constituent pas par elles-mêmes un acte de langage.

(24) Je me **le** demande.

(25) Il m'a dit quand il partait.

Pour ce qui est de leur forme syntaxique, ces interrogatives subordonnées sont généralement introduites dans la phrase racine par un mot interrogatif (dans les exemples (21-23) : *si, quand*). L'ordre des mots y est donc la plupart du temps contraint même si, comme pour les interrogatives racines (non-subordonnées), d'autres formes sont attestées (26-27) (exemples repris de DELAVEAU et al. [2021], le symbole « ! » indique une forme marquée, divergente de la norme linguistique).

(26) !Je me suis demandé *est-ce qu'*ils auront le temps de finir?

(27) !Je me demande *c'est quoi*, son problème.

Le recours à ces formes réputées *non-standard* [DELAVEAU et al., 2021], s'il est attesté, ne constitue pas un acte de langage interrogatif en lui-même. Pour cette raison, les données présentées par la suite (corpus ou données expérimentales) se focaliseront uniquement sur les interrogatives partielles racines.

2.1.2 L'interrogation partielle en français

Au niveau du type syntaxique de phrase, les nomenclatures établissant les formes interrogatives présentes en français sont nombreuses et plus ou moins complètes. L'exploration de corpus permet de clarifier et d'en actualiser les usages par les francophones.

2.1.2.1 Nomenclatures sur le français

COVENEY [2011] dresse un panorama assez exhaustif des travaux menés au 20^{ème} siècle sur l'interrogation en français, avec pour thématique générale l'interrogation directe et non pas seulement partielle. Pour ce faire, il compile notamment les données d'un certain nombre d'études de corpus, qui seront reprises en Section 2.1.2.2. DELAVEAU et al. [2021] présentent pour leur part, dans le projet de la Grande Grammaire du Français [ABEILLÉ et GODARD, 2012], une vision réactualisée des différentes formes d'interrogation attestées sur corpus, en français hexagonal aussi bien que dans d'autres variétés comme le français du Québec. Il ressort de la combinaison de ces travaux que les francophones ont à leur disposition un large éventail de types phrastiques différents pour produire des interrogatives.

Pour ce qui est des *questions totales* tout d'abord, trois formes principales sont disponibles en français, exemplifiées en (28-30) avec, respectivement : une déclarative questionnante, une interrogative avec une forme clitique du sujet postposée au verbe, et une interrogative introduite par la locution figée « est-ce que ».

(28) Tu viens demain?

(29) Viens-tu demain?

(30) Est-ce que tu viens demain ?

ABEILLÉ et al. [2012] étudient la valeur pragmatique de ces trois différentes variantes d'interrogation totale et déduisent d'une exploration de corpus que seules l'inversion et la question en « est-ce que » constituent de réels actes interrogatifs. La déclarative questionnante, dont la valeur de question repose sur la seule prosodie, ne serait en fait pas une interrogation mais plutôt un appel déguisé pour que la personne recevant l'énoncé valide la proposition sémantique que celui-ci transmet.

Pour poser une *question partielle*, l'éventail offert aux francophones est plus large, comme illustré dans le Tableau 2.1 par les exemples (31-37), sans même prétendre à l'exhaustivité. Les stratégies qui seront principalement étudiées dans ce travail sont les quatre principalement utilisées en français hexagonal d'après les données de corpus (voir 2.1.2.2), à savoir l'*in situ* (IS), l'*antéposition simple* (*Fronting*, F), l'*antéposition avec inversion* (*Fronting with INversion*, FINV), et l'*antéposition en est-ce que* (*Fronting with ESK*, FESK). Les autres formes semblent tenir soit d'un usage limité à des aires géographiques plus réduites ou diffuses, soit d'un usage plus sociolectal, dont il faut encore définir plus précisément les contours.

Exemple	Label
(31) Tu viens quand ?	IS
(32) Quand tu viens?	F
(33) Quand viens-tu?	FINV
(34) Quand est-ce que tu viens?	FESK
(35) C'est quand que tu viens?	clivée
(36) Quand que tu viens?	antéposition + complémenteur
(37) Quand c'est-y que tu viens?	antéposition + complémenteur + c'est-y

TABLEAU 2.1 – Stratégies principales d'interrogation partielle en français

En termes de sémantique stricte, il est généralement admis que le contenu propositionnel d'une interrogative partielle est « *incomplet et comporte une variable qui demande à être spécifiée* » [MARANDIN, 2018, p. 12], et ce indépendamment du type phrastique. Dans les exemples (31-37), le but est ainsi toujours de recueillir l'information temporelle spécifiant le moment de l'arrivée du ou de la destinataire.

Il est déjà intéressant de constater que, selon leur visée (grammaires de référence, grammaires pour l'apprentissage, catalogue scientifique des formes utilisées par les membres d'un groupe linguistique), les nomenclatures de l'interrogation en français ne sont pas les mêmes. En comparaison avec COVENEY [2011]; DELAVEAU et al. [2021], un ouvrage de référence tel que RIEGEL et al. [2014, pp. 668-683] ne consacre ainsi qu'une dizaine de lignes, sur les quinze pages dédiées aux phrases interrogatives, à la présentation des phrases *in situ*. Dans cette description, le type *in situ* est classé parmi les « *Formes familières de l'interrogation partielle* » [RIEGEL et al., 2014, p. 678]. Ce type phrastique et d'autres (clivée, antéposition + complémenteur) seraient réservés à des conditions d'usage spécifiques (langue parlée notamment) et permettraient notamment d'éviter l'inversion. Les structures inversées sont en effet ici vues comme « mal acceptées » à l'oral, car trop éloignées

dans leur complexité de la phrase déclarative canonique.

Ainsi, la langue parlée familière simplifie les structures pour aboutir à une certaine unité de l'interrogation : l'intonation joue un rôle essentiel et la phrase garde l'ordre habituel de la phrase déclarative. (RIEGEL et al. [2014, p .680])

2.1.2.2 Fréquence d'usage et données de corpus

RIEGEL et al. [2014] évoquent donc un usage limité de certains types phrastiques à des conditions linguistiques particulières. En particulier, le type *in situ* tiendrait d'une langue familière et parlée, et serait utilisé pour « simplifier » les phrases et les rapprocher des structures déclaratives. De nombreuses études de corpus ont déjà cherché à caractériser les emplois des différents types d'interrogatives partielles par les francophones, et il est intéressant en particulier d'y observer les proportions d'usage des phrases *in situ*.

Corpus du 20^{ème} siècle

COVENEY [2011] compile entre autres les résultats de multiples études menées pendant la seconde moitié du 20^{ème} siècle sur l'interrogation en français. Le tableau qui synthétise ce bilan est reproduit, un peu adapté, dans le Tableau 2.2. Y sont présentés, par colonne, les pourcentages d'emploi des types interrogatifs majeurs relevés dans les multiples corpus qui servent de base à toutes ces études.

Type	Corpus d'origine et Fréquences (%)				
	Radio	Théâtre	Bourgeois familial	Enfants (9ans)	Ouvriers
In situ <i>Jean vient quand?</i>	25	–	33	12.9	12
Antéposition simple <i>Quand Jean vient?</i>	10	12.5	46	31.6	36
Antéposition + complémenteur <i>Quand que Jean vient?</i>	0	–	0	6.9	26
Antéposition + ESK <i>Quand est-ce que Jean vient?</i>	3	38.9	12	41.5	8
Antéposition et inversion (clitique) <i>Quand vient-il?</i>	50		3		
Antéposition et inversion (nom) <i>Quand vient Jean?</i>	12	48.6	2	3.6	9

Adapté de [COVENEY, 2011]

TABEAU 2.2 – Types d'interrogatives partielles dans des études du 20^{ème} siècle

Si la somme par colonne n'est pas toujours de 100, ceci est dû au fait qu'il s'agit là des structures principales employées dans les corpus analysés et non de toutes celles relevées. Voici pour chaque corpus (colonne) la taille de l'échantillon et sa source :

- RADIO : N=4367, chiffres issus de BEHNSTEDT [1973].
- THÉÂTRE : N=1515, chiffres issus de TERRY [1970].
- BOURGEOIS FAMILIER : N=446, chiffres issus de BEHNSTEDT [1973].
- ENFANTS : N=364, chiffres issus de SÖLL [1983].
- OUVRIERS : N=587, chiffres issus de BEHNSTEDT [1973].

De manière générale, même en intégrant la spécificité du corpus de THÉÂTRE et de celui de RADIO (textes lus), même dans les corpus plus spontanés (les trois derniers), et même enfin dans la variété de français a priori la plus « populaire » (corpus OUVRIERS), la préférence semblait aller, ici, aux éléments interrogatifs réalisés frontalement, avec ou sans inversion sujet/verbe. La proportion des éléments interrogatifs réalisés in situ ne s'élevait, elle, pas au-delà de 33%.

Par ailleurs, dans une autre étude, ASHBY [1977] relevait quant à lui 38.8 % d'interrogations in situ dans un corpus – réduit (N=85) – de parler parisien.

Si ces nombres sont intéressants, ils présentent un inconvénient majeur. Ces données datent en effet, pour les plus récentes de plus de trente ans et pour certaines de près de cinquante ans. Or, les langues évoluent et la représentativité de ces données est donc limitée pour avoir une vision sur le ou les français du 21^{ème} siècle.

Données plus contemporaines

Une analyse du corpus établi par HAMLAOUI [2009a], sur des données radiophoniques de 2005-2009, offre un contraste avec les précédentes (Tableau 2.3). Les interrogatives de ce corpus (N=200) proviennent d'environ 20 heures d'interviews menées par une journaliste de France Inter pour l'émission *Là-Bas si j'y suis*. Selon l'auteure, il s'agit d'un français démotique, représentatif d'une langue standard utilisée au quotidien par des francophones en interaction non-formelle. Les entretiens se déroulent en effet sur le lieu de travail ou de vie des personnes enregistrées. La journaliste s'y rend par exemple pour demander à ces personnes d'expliquer leur métier ou l'histoire qui les a conduites à l'exercer.

Type	Fréquence (%)
In situ	61.24
<i>Tu fais comment?</i>	
Antéposition simple	23.45
<i>Combien t'as de budget?</i>	
Antéposition + ESK	11
<i>Quand est-ce que Jean vient?</i>	
Clivée	4.31
<i>C'est quoi qui est nouveau?</i>	

TABLEAU 2.3 – Types d'interrogatives partielles dans le corpus *Là-Bas*, Hamlaoui (2009)

Les proportions sont ici quasiment inverses à celles établies par les études citées par Coveney, même les plus généreuses, en matière de réalisation in situ. Il y a ici une nette prépondérance des phrases in situ, à près de 65 % si les clivées sont comptées. L'absence d'interrogatives avec inversion y est notable. Ces données sont cependant issues d'un corpus relativement restreint (20 heures de parole, environ 200 interrogatives), et ne proviennent que d'une seule journaliste, ce qui est une limite importante.

D'autres données récentes confirment néanmoins cette proportion importante de phrases in situ avec par exemple ADLI [2015], sur des données de productions orales (27 heures d'enregistrement, pour 1721 interrogatives produites au total par 101 volontaires). Le Tableau 2.4 donne ainsi la proportion des types interrogatifs majeurs produits lors d'une tâche de jeu, par des volontaires qui devaient enquêter sur un meurtre fictif en dialoguant avec un ou une membre de l'équipe de recherche.

Type	Fréquence (%)
In situ	57.6
Antéposition simple	15.6
Antéposition + ESK	16.7
Antéposition + Inversion	9.9

TABLEAU 2.4 – Types d'interrogatives partielles dans la base sgs, Adli (2015)

Ces données mettent donc ici aussi en évidence un recours massif au type phrastique in situ par les francophones, plus important en proportion que la somme des différents types antéposés. Ces chiffres contrastent de nouveau avec les données plus anciennes compilées par [COVENEY, 2011].

Il est à noter toutefois que ces données récentes reposent sur des données de français oral, ce qui est aussi un point de contraste avec certaines des données plus anciennes (corpus de THÉÂTRE par exemple, où aucune phrase in situ n'était présente). D'autres données de la fin du 20^{ème} siècle [KAISER et QUAGLIA, 2015], issues de 3 romans écrits entre 1995-1998 pour environ 480 interrogatives, laissent ainsi déjà à penser que la différence de modalité entre oral et écrit joue sans doute un rôle. En effet, dans cette étude, 70% d'antépositions (tous types confondus) étaient relevés pour seulement 30% de phrases in situ.

L'évolution des chiffres semble un peu différente pour la forme antéposée avec inversion, mentionnée comme « standard » par [RIEGEL et al., 2014]. Alors qu'elle était employée dans 50 à 60 % des productions dans les corpus de RADIO et THÉÂTRE mentionnés par COVENEY [2011], ADLI [2015] n'en trouve pas plus de 10 % dans son corpus, ce qui était déjà le cas pour les autres corpus mentionnés par COVENEY [2011]. HAMLAOUI [2009a], pour sa part, n'en trouve aucune trace. Ce déclin de la forme avec antéposition et inversion coïncide pour [ELSIG, 2009] avec une restriction de son usage à des formules figées comme « *Comment allez-vous ?* », à des verbes précis (*vouloir, voir, avoir, être*) et à la deuxième personne.

2.2 Analyses de l'interrogation partielle en français

De nombreux éléments ont déjà été avancés pour caractériser la pluralité des formes disponibles aux francophones en matière d'interrogation directe partielle. En premier lieu, la distinction est souvent effectuée entre les phrases dites *in situ* et les types antéposés, dans une analyse plutôt syntaxique où l'ordre des constituants est primordial. Des facteurs comme la phonotaxe, la structure informationnelle, ou même des éléments de type sociolinguistique, ont été avancés pour distinguer les conditions d'usage des différents types phrastiques disponibles.

L'objectif de la présente section n'est pas une présentation exhaustive des différentes théories avancées pour expliquer la variation du français sur les interrogatives partielles. Il s'agit plutôt de montrer qu'une intégration plus fine des enjeux sociaux de l'interaction linguistique est cruciale à une meilleure compréhension des choix faits par les francophones en la matière. Ces enjeux sociaux requièrent d'aller plus loin qu'une intégration du facteur temporel (évolution diachronique des usages linguistiques) ou du facteur sociolectal (le groupe social de l'individu définissant ses préférences linguistiques).

2.2.1 Les théories en syntaxe formelle

Un premier champ d'investigation est donc l'élaboration de travaux syntaxiques pouvant rendre compte de manière unifiée de la diversité de formes phrastiques utilisées pour convoquer une interrogation partielle. Les théories syntaxiques s'accordent généralement sur l'existence d'une construction *canonique*, dans laquelle le mot interrogatif est à la position habituelle de l'argument du verbe en français (après celui-ci). La comparaison de (38) et de (39) illustre cette comparabilité des structures avec, ici, un ajout temporel.

(38) Tu viens **quand** ?

(39) Tu viens **demain**.

Les autres constructions interrogatives sont par comparaison réputées « non-canoniques » ou marquées, avec le mot interrogatif en périphérie gauche de la phrase. Si des théories syntaxiques sont plus directement descriptives de ce phénomène, comme les grammaires HPSG (*Head-driven Phrase Structure Grammar* [SAG, 2010], mais voir plus largement les grammaires d'unification [ABEILLÉ, 2007]), les grammaires génératives et proches des cadres minimalistes, par exemple, proposent une analyse en termes de *mouvement* en profondeur dans la phrase et de *catégorie vide*.

La plupart des travaux s'inscrivant dans ces cadres théoriques font ainsi de la forme interrogative canonique (38) la construction fondamentale : littéralement, le label « in situ » qui désigne souvent ce type phrastique signifie « en place ». Par contraste avec cette position in situ, la plupart des autres formes interrogatives, parfois appelées *ex situ*, présentent donc une « antéposition » de l'élément interrogatif. Le point de comparaison est ici le verbe.

De ce point de vue, il est notable que les questions partielles portant sur le sujet d'une phrase présentent en français une certaine particularité. Le mot interrogatif est en effet alors à la fois *in situ* et avant le verbe (40-41). Au sens des théories du mouvement, il n'y a cependant aucun mouvement visible (même si des mouvements sous-jacents peuvent être postulés) par rapport à la phrase déclarative. De manière tout aussi notable, il y a moins de variation sur les interrogatives partielles sujet du français, avec tout au plus la possibilité de « renforcer » l'interrogation par le recours à la locution *est-ce que* (42) ou au clivage (43).

(40) **Qui** vient demain ?

(41) **Jean** vient demain.

(42) **Qui est-ce qui** vient demain ?

(43) *C'est qui qui* vient demain ?

L'antéposition est donc ici conceptualisée comme une dérivation depuis la forme originelle, incluant une opération de déplacement (« *move* ») de l'élément interrogatif (« mouvement Wh- ») depuis le syntagme verbal (SV ou VP, *Verbal Phrase*) jusqu'à la périphérie gauche de la phrase, et plus spécifiquement jusqu'au domaine du complémenteur (CP, *Complementizer Phrase*). À cette place de la structure hiérarchique de la phrase, il y a une opération de fusion (« *merge* ») de l'élément déplacé.

Ce phénomène de déplacement a été source de nombreux travaux syntaxiques. Il s'est agi autant de le caractériser généralement [CHOMSKY, 1981; KAYNE, 1981; ROSS, 1967; STOCKWELL, 1977] que sous des perspectives plus précises, comme [DONATI, 2006] tentant de discriminer entre

mouvement de tête et simple mouvement de mot, ou [FALK, 2012] tentant de caractériser les limites d'un tel mouvement, ou encore [BOŠKOVIĆ, 2002; RUDIN, 1988] se penchant sur les langues qui présentent une antéposition de multiples éléments interrogatifs dans une même phrase.

Les théories intégrant ce concept de mouvement vont généralement de pair avec la notion de trace, et postulent qu'au sein de la phrase dérivée demeure, à l'emplacement d'origine de l'élément interrogatif, une trace invisible mais toujours opérante, à laquelle il est relié (44).

(44) Comment_i Sarah paie ses courses en espèces *t_i* ?

C'est sur cette base que repose par exemple l'impossibilité d'extraire certains éléments interrogatifs en cas d'éléments interrogatifs multiples dans une phrase comme en (45) pour le français. C'est également sur la base de ces contraintes qu'il a pu être réputé impossible d'extraire un constituant dans des subordinées ou hors du syntagme sujet (46) (exemple tiré de [ABEILLÉ et WINCKEL, 2019], et voir WINCKEL et ABEILLÉ [accepté] pour d'autres attestations de corpus).

(45) *Pour qui_j pourquoi_k Jean fait les courses pour sa mère *t_j* car elle est malade *t_k* ?

(46) De qui_i l'anniversaire de qui *t_i* tombe-t-il le 27 ?

Quoi qu'il en soit, la théorie syntaxique du mouvement de constituants depuis une position fondamentale in situ analyse le phénomène variationnel au niveau de la phrase interrogative elle-même. Dans les limites de ce cadre, elle se prononce par définition peu sur les usages qui sont faits des différentes autres variantes possibles.

2.2.2 Les théories liées à la structure informationnelle

Plus globalement, ces phrases interrogatives sont cependant utilisées au cours de séquences discursives impliquant généralement plusieurs personnes et plusieurs énoncés successifs. C'est dans ce cadre plus large que la notion pragmatique du statut informationnel des éléments de l'énoncé peut intervenir. La présence d'éléments au statut informationnel différent dans la phrase a dans cette approche un impact sur le type phrastique sélectionné par les francophones pour produire une interrogative partielle.

2.2.2.1 Structure de l'information et oppositions entre phrases in/ex situ

Sous cette perspective, il a ainsi été argumenté que les interrogatives in situ étaient notamment à distinguer des formes antéposées, en ce qu'elles étaient plus fortement présuppositionnelles [BOECKX, 1999; CHANG, 1997; LAI-SHEN CHENG et ROORYCK, 2000]. L'argument serait ici qu'une phrase comme (47) est légitimée uniquement par un contexte où Marie est présupposée avoir acheté quelque chose, et où la réponse ne pourrait qu'être une information positive nouvelle : l'élément acheté par Marie.

(47) Marie a acheté quoi ?

Cette conception de l'interrogation in situ est nuancée par exemple dans MATHIEU [2004] qui montre qu'une telle question peut avoir pour réponse quelque chose comme « Rien. », faisant alors tomber la présupposition. BEYSSADE [2006] apporte pour continuation l'hypothèse que, s'il doit y avoir une différence informationnelle entre une interrogative in situ et une interrogative avec antéposition, celle-ci est plutôt liée au thème du discours. Dans cette conception, les interrogatives

avec antéposition permettent d'introduire un nouveau thème de discours, alors que les phrases in situ continuent simplement le thème en cours. Pour illustrer cette théorie, l'auteure évoque les titres de romans, journaux ou articles scientifiques qui, lorsqu'ils sont interrogatifs, prennent en général la forme antéposée (48-49), et jamais ou presque jamais la forme in situ (50-51) [BEYSSADE, 2006, p. 15] (le symbole # indique ici une incongruité de la phrase dans le contexte supposé d'une une de journal). Le titre est alors vu comme introduisant un nouveau thème de discours et donc incompatible *a priori* avec une forme in situ.

(48) **Où** va la France?

(49) **À quoi** rêvent les jeunes filles?

(50) # La France va **où**?

(51) # Les jeunes filles rêvent **à quoi**?

BOUCHER [2010] reprend quant à lui l'idée d'une question in situ plus fortement présuppositionnelle, ce qu'appuierait son emploi possible en question écho comme en (52) (repris de ENGBAHL [2006], cité par BOUCHER [2010]). L'interrogative s'appuie ici directement sur l'énoncé préalable.

(52) A : Mon fils, il lit [inaudible]

B : Il lit **quoi**?

Par contraste, BOUCHER [2010] propose une lecture des questions avec antéposition simple comme remettant en cause la présupposition à l'origine de cette question, alors que les antépositions avec inversion appuieraient au contraire cette présupposition (53-54).

(53) **Pourquoi** voulez-vous l'appeler?

Présupposition : Vous voulez l'appeler pour quelque raison.

(54) **Pourquoi** vous voulez l'appeler?

Présupposition : Il n'y a pas de raison pour l'appeler.

Se basant sur des données de corpus, HAMLAOUI [2011] argumente quant à elle pour une distinction au niveau de la portée du focus dans l'interrogative partielle selon que celle-ci présente une antéposition ou non. Pour l'auteure, le focus d'une interrogative avec antéposition peut inclure au moins une partie des éléments non-interrogatifs de la phrase interrogative, alors que le focus d'une interrogative in situ doit inclure seulement le syntagme interrogatif. C'est une théorie assez proche de celle de BEYSSADE [2006] dans le sens où les interrogatives in situ sont ici vues comme moins appropriées lorsque la question impose un changement de focus (55) (extrait du corpus *Là-Bas* [HAMLAOUI, 2009a]).

(55) *Contexte : R et ses allocutaires se tiennent devant une école pendant la récréation. L'un des allocutaires raconte que cette école était avant une école pour filles et qu'il n'y est pour sa part arrivé que quand il était déjà adolescent.*

R : Et vous, vous jouez **à quoi** par exemple? [pause : 372ms] Dans la cour, puisque là on est devant la cour et que c'est la récréation, vous jouez **à quoi**?

Ici, l'auteure s'intéresse à la pause marquée et au besoin pour R de répéter l'interrogation en précisant le cheminement de pensée qui a conduit à la poser. Ces éléments indiquent que ce que R prenait pour acquis ne l'était pas pour ses destinataires. Dans cette vision, une phrase comme 56 aurait sans doute assuré une transition moins abrupte pour les allocutaires.

(56) Et à **quoi** vous jou*(i)ez*, vous, par exemple (dans cette cour)?

Une autre opposition entre interrogatives antéposées et interrogatives in situ est par exemple proposée par MYERS [2007], qu'elle explique entre autres par la *possibilité de répondre (answerability)*. Son analyse est construite sur la base de données déjà anciennes (Corpus Barnes-Blyth de français parlé, BARNES [1984]). Pour elle, les phrases in situ seraient plus susceptibles d'être produites dans des contextes où il est possible de répondre directement et, au contraire, les phrases avec antéposition seraient liées à des situations où les réponses peuvent être plus ouvertes et moins prédictibles par la personne posant la question. C'est ainsi qu'elle explique notamment la différence entre des interrogatifs tels « comment » ou « pourquoi » d'une part, qui sont plus souvent antéposés, et le diptyque « qu' » (antéposition)/ « quoi » (in situ), ou encore « où », plus équitablement répartis entre position in situ et antéposition. La notion d'*answerability* repose aussi sur une analyse syntaxique, avec une réponse qui exige souvent l'introduction d'une phrase subordonnée pour « comment » et « pourquoi » par exemple, et moins souvent pour « quoi » et « où » (voir les séquences (57-58) ci-dessous, extraites de MYERS [2007, p. 128].

(57) Tu prends **quoi**? (*possibilité de répondre directement*)

Je prends du bacon...

(58) Et **pourquoi** tu en as pas toi? (*utilisation d'une subordonnée pour répondre*)

Ben parce que je suis pas sage

2.2.2.2 Une influence de la phonotaxe?

Dans une perspective proche, HAMLAOUI [2009b] évoque pour sa part des stratégies phonotactiques qui pourraient favoriser le recours à l'une ou l'autre variante disponible pour l'interrogation partielle en français, à partir d'un corpus constitué pour sa thèse (corpus *Là-Bas*, voir Section 2.1.2.2).

La longueur de la partie non-interrogative de la phrase serait ainsi corrélée à la position de l'élément interrogatif. D'après ses recherches, il est ainsi plus probable d'avoir une interrogative avec antéposition (sans inversion en l'occurrence) lorsque le nombre de syllabes de la partie non-interrogative augmente. Pour des parties non-interrogatives constituée de moins de 3 syllabes, il y a ainsi 86% de phrases in situ dans ses données, mais cette proportion tombe à 40 % quand la partie non-interrogative fait entre 7-9 ou syllabes (29 % quand la partie non-interrogative fait entre 10-12 syllabes, et 0 occurrence quand la partie non-interrogative dépasse 13 syllabes).

Si les méthodes d'extraction des interrogatives et la méthodologie du découpage en syllabes sont décrites dans l'étude, il conviendrait néanmoins de comparer des données exemplifiées pour mesurer la portée de cette proposition.

Il ne s'agit sans doute pas néanmoins que du nombre de syllabes et l'auteure envisage dans le même temps que c'est en fait l'informativité de la partie non-interrogative de la phrase qui serait aussi en jeu. Ceci s'inscrit dans la lignée d'autres travaux plus psycholinguistiques explorant

l'influence de la charge informationnelle sur l'ordre et les relations entre constituants au sein de la phrase (voir [ALMOR \[1999\]](#) sur la thématique de la référence, ou encore par exemple [HEMFORTH et al. \[2013\]](#), où les préférences d'attachement du syntagme prépositionnel changent, non pas en fonction de sa longueur en syllabes, mais de sa charge informationnelle).

Et pour ce qui est des interrogatives partielles du corpus *Là-Bas*, en effet, il y a une différence entre les verbes sémantiquement « légers » (être, avoir, faire), qui sont supposés être moins informatifs, et les verbes *lexicaux* (terminologie employée par [HAMLAOUI \[2009b\]](#) pour désigner les verbes sémantiquement « pleins »). Ces verbes sémantiquement légers semblent favoriser le recours à une phrase in situ (à 87%, dans le corpus analysé). Une telle opposition entre phrases in situ et phrases avec antéposition ne se trouve pas pour les verbes *lexicaux*, où la charge informationnelle est plus équitablement répartie entre l'élément interrogatif et la partie non-interrogative de la phrase. Il s'agit donc, dans cette perspective aussi, de mettre en contraste les formes antéposées et les formes in situ, cette fois sur la base d'une localisation différente de la charge informationnelle au sein de la phrase interrogative.

Ici encore, au-delà des pourcentages résumant les données qui sont fournis dans l'étude, un accès à des exemples permettrait une meilleure vision.

2.2.3 Les théories sociolinguistiques

Les théories syntaxiques analysent les différents types de phrases interrogatives avec pour cadre la phrase prise en tant qu'élément fini. Les théories pragmatiques dépassent le domaine de la phrase et étendent le fait linguistique au « cotexte » pour trouver des facteurs propres à l'énoncé qui pourraient influencer sur le type de phrase interrogative. Une troisième catégorie de facteurs relève de la sociolinguistique. Dans ces analyses, des éléments extérieurs de « contexte » (pris au sens large) sont intégrés. Il peut s'agir d'éléments temporels (les usages interrogatifs ont évolué au cours des époques). Il peut aussi plus directement s'agir d'éléments sociaux, avec des emplois interrogatifs différents selon le groupe social des personnes impliquées dans l'acte de langage (que la distinction soit effectuée sur des critères socio-économiques ou géographiques).

2.2.3.1 Une vision diachronique

[LARRIVÉE \[2016, 2019\]](#) dresse un tableau évolutif des usages interrogatifs en français à travers les époques temporelles. Remontant jusque des sources du 15^{ème} siècle, l'auteur converge conceptuellement avec les théories faisant des phrases in situ un type phrastique à la valeur pragmatique particulière (en lien avec une information ancienne déjà présente dans le discours). Sa proposition est néanmoins que ces théories ont « *deux siècles de retard* » [[LARRIVÉE, 2019](#), p.128].

Pour Larrivée, l'augmentation de la proportion des phrases in situ dans les données les plus récentes, des 20^{ème} et 21^{ème} siècles, coïncide avec une perte de la valeur pragmatique de ce type phrastique. Cette analyse repose notamment sur des données tirées du corpus ESLO [[ESHKOL-TARAVELLA et al., 2011](#); [LLL, 2017](#)] qui sera utilisé et décrit plus en détail dans le Chapitre 4. Il est d'ores et déjà utile de savoir que ce projet compile à la fois des données récoltées dans les années 1960 (ESLO1) et des données récoltées dans les années 2000 (ESLO2).

En comparant les extractions du mot *comment* dans ESLO1 et ESLO2, l'auteur note d'abord une apparente augmentation des usages des interrogatives in situ pour cet élément interrogatif (70 interrogatives sur 1113 extractions, soit 6.8% dans ESLO2, contre 13 interrogatives sur 1434

extractions, soit 0.98% dans ESLO1). Cette augmentation est à prendre avec nuances, de l'aveu même de Larrivée, puisque les extractions totales ne sont pas forcément toutes des interrogatives.

Ensuite, sur la base du relevé de ces 70 réalisations in situ de l'interrogatif « Comment » dans les données des années 2000, l'auteur montre qu'il n'est plus besoin à cette période d'une activation explicite dans le discours précédent pour licencier le type in situ. L'exemple (59) donne une séquence où le type in situ est explicitement activé dans le contexte, et l'exemple (60) illustre une réalisation qui n'a pas été activée au préalable.

(59) OW26 : Dans les jeux antiques euh *ils se dopaient quand même avec des méthodes un peu bizarres* mais

ch_PP6 : ils se dopaient **comment** ?

OW26 : ils prenaient euh des plantes

(ESLO2_ENT_1026)

(60) finalement tu trouves **comment** la vie à Orléans ?

(ESLO2_ENT_1021)

Si elle reste possible, l'activation explicite des interrogatifs *comment* réalisés in situ ne correspond plus, dans ESLO2 (années 2000), qu'à 2 occurrences parmi les 70 relevées par l'auteur (2.8%). Dans ESLO1, cette proportion se situe à 30.8% (4 occurrences sur les 13 interrogatives in situ relevées). Dans des données francophones plus anciennes encore, d'abord des 19^{ème} et 18^{ème} (issues du corpus Frantext [ATILF-CNRS, 1998-2020]), et même du 15^{ème} siècle, la réalisation du type in situ d'un élément interrogatif et de sa valeur pragmatique nécessitent, dans la quasi-totalité des cas, une activation explicite dans le discours précédent. Il est à noter que dans ces sources, les occurrences in situ correspondent à moins de 1% des interrogatives attestées.

Malgré les limites quantitatives de ces données, l'auteur en fait une illustration du changement linguistique. Il relie ainsi la multiplication du type in situ à une perte de sa valeur spécifique originelle, à mesure que son usage s'est répandu. Il est par ailleurs déjà à noter que l'augmentation de proportion des usages in situ de l'interrogatif « comment » dans ESLO corrèle avec les données contemporaines de HAMLAOUI [2009a] ou ADLI [2015]. En contraste avec les données plus anciennes compilées par COVENEY [2011], la variante in situ est devenue plus fréquente. Le Chapitre 4 reviendra sur cette évolution diachronique.

2.2.3.2 Une vision diatopique

Un autre type d'études sociolinguistiques relie classiquement les phénomènes variationnels à des usages préférentiels de certaines variantes selon le groupe social. Un premier niveau de séparation du groupe linguistique est l'approche géographique et un pan de la littérature sociolinguistique a caractérisé quelques variétés du français, de type géolectales, dans leur emploi des types de phrases interrogatives.

C'est ainsi par exemple que GURYEV et DELAFONTAINE [2015] se penchent sur un corpus suisse d'écrits sms. L'objectif des auteurs est dans ce travail d'analyser des spécificités de l'écrit sms, et de plaider pour une analyse multifactorielle de la variation dans les interrogatives au sens large (totales et partielles). Cette analyse devrait pour eux tenir compte de la morphosyntaxe autant que d'influences sociolinguistiques, mais aussi du style personnel des individus. Cette étude s'insère dans un projet beaucoup plus vaste analysant plus de 25947 écrits sms collecté par les universités

de Zurich et Neufchâtel dans plusieurs langues. Parmi les 4619 sms écrits en langue française dans cette masse, par 2784 personnes différentes âgées de 12 à 77 ans (46% entre 20 et 29 ans), il se trouve 427 interrogatives partielles. Les chiffres portant spécifiquement sur ces phrases montrent une relative proximité avec les données de [HAMLAOUI \[2009a\]](#) et [ADLI \[2015\]](#), au moins pour ce qui est des phrases in situ : 53% de phrases in situ, 16.6% d'antépositions simples, 14.8% d'antépositions avec inversion et 5.9% d'antépositions en « est-ce que ».

Plus loin de la France, [MATHIEU \[2009\]](#) évoque pour sa part des différences importantes entre le français « de France » et le français laurentien (Québec), avec par exemple l'inexistence de formes inversées pour les questions partielles (61), dont l'équivalent serait (62) (exemples proposés par l'auteur lui-même dans [MATHIEU \[2009, p. 13\]](#)).

(61) *Quand es-tu parti?

(62) Quand c'est qu't'es parti?

Il est à noter que, en parallèle, les interrogatives totales à sujet suffixé existent au Québec, avec un suffixe *-tu* qui semble réservé à ce type d'interrogation (63) (exemple de [DELAVEAU et al. \[2021\]](#), tiré du Corpus de Français Parlé au Québec, le CFPQ [[DOSTIE, 2016](#)]).

(63) Après je lui dis bon, on **va-tu** prendre un café (CFPQ, 19, segment 7, P.63).

En contradiction apparente avec [MATHIEU \[2009\]](#), [DELAVEAU et al. \[2021\]](#) mentionnent pour les interrogatives partielles que « *dans l'oral spontané, le verbe à sujet suffixé est très minoritaire, y compris au Québec, et pour toutes les classes sociales : il semble se cantonner essentiellement à des questions de routine (Comment fait-on... ?)* ».

Pour intéressantes qu'elles sont, ces données diatopiques ne sont pas les plus fournies, et c'est sans doute l'approche diastratique de la variation qui a donné lieu au plus de travaux sociolinguistiques sur les interrogatives du français.

2.2.3.3 Une vision diastratique

Une autre distinction à l'échelle du groupe consiste en effet à établir, au sein d'un groupe social concentré sur un même territoire et parlant a priori la même variété linguistique, des sous-ensembles. Cette perspective conduit à relier les phénomènes variationnels à des usages préférentiels de certaines variantes par un sous-ensemble du groupe social, et à d'autres usages par d'autres sous-groupes. Le postulat est alors que l'appartenance à un sous-ensemble du groupe social façonne les habitudes, notamment linguistiques, de l'individu (dans une application de la maxime de conformité de [KELLER \[1994\]](#) et [HASPELMATH \[1999\]](#) : « Talk like others talk. »).

Cette approche est celle des études compilées par [COVENEY \[2011\]](#), à savoir [BEHNSTEDT \[1973\]](#); [TERRY \[1970\]](#) et [SÖLL \[1983\]](#) (voir Tableau 2.2, p. 15). Dans ces données, en comparant les chiffres de Behnstedt et les corpus BOURGEOIS FAMILIER et OUVRIERS, la proportion de phrases in situ est ainsi assez différente (33% contre 12%), ainsi que la proportion d'antépositions en « est-ce que » (42% contre 8%) et celle d'antépositions avec complémentateur (0 occurrence contre 26%). Ce genre d'opposition pourrait permettre, dans une vision sociolinguistique classique, de caractériser, et éventuellement de mettre en contraste, une « langue des bourgeois » qui serait différente d'une « langue des ouvriers ».

(Il faut noter au passage que dans le même Tableau, les corpus de RADIO de Behnstedt et celui de THÉÂTRE de Terry font apparaître une représentation importante des phrases avec antéposition et inversion. Cette spécificité peut s'attribuer à une langue encore différente, qui serait celle « de la culture et des intellectuels », mais elle peut aussi s'attribuer à un contexte interactionnel très différent (absence de public, texte lu, etc.). La section 4.1 et plus largement le Chapitre 4 viendront explorer cette thématique.)

C'est aussi cette approche diastratique qui gouvernait déjà le travail de POHL et STRAKA [1965], établissant un lien entre niveau d'études et types interrogatifs produits, ou encore celui de ASHBY [1977] venu prolonger cette thèse. Ce dernier montre en effet que l'inversion est certes la plus rare des formes interrogatives en production francophone (classe moyenne à Paris), à l'inverse des questions in situ notamment, mais que malgré cela les personnes issues du milieu académique (étudiant.e.s et corps professoral) produisent en proportion plus d'inversions que les autres.

C'est encore cette approche de la variation qui guide les conclusions de QUILLARD [2001] dont les données sont plus récentes. À partir d'un corpus de 1658 questions issues de contextes variés (repas, réunions, formations professionnelles, cours et examens universitaires, émissions-débat politique), l'auteure trouve une l'influence sur les productions interrogatives des facteurs

- « âge » : 60% environ de phrases in situ chez les moins de trente-cinq ans, 40% au-delà,
- « catégorie socio-professionnelle » : environ 45% seulement d'in situ pour la catégorie la plus haute contre 50 à 65% pour les catégories moins élevées.

Dans une vision affinée du sous-groupe social, ADLI [2006, 2013] sépare les individus en fonction de leur mode de vie sur la base de données expérimentales dont il reconnaît qu'elles sont à compléter. Alors que dans ADLI [2006] la vision de la variation se fait plutôt en termes de registre (antéposition avec inversion = registre soutenu), les conclusions de ADLI [2013] établissent deux groupes distincts de personnes, par le biais de jugements d'acceptabilité. Il distingue ainsi dans les volontaires de son étude, d'une part une majorité de gens pour qui les formes antéposées ou in situ sont relativement équivalentes, et d'autre part deux groupes minoritaires d'individus préférant soit l'antéposition, soit les formes in situ. La première minorité est composée de personnes au profil orienté vers les domaines culturels ou politiques, ou encore qui étudient le langage ou l'art, et qui auraient un comportement linguistique plus marqué par la norme (selon laquelle l'antéposition avec inversion serait meilleure). La seconde minorité est constituée par contraste de gens au profil plus tourné vers le sport et plus éloigné au quotidien des contraintes normatives.

Enfin, un travail comme FARMER [2013] mêle étude diachronique et étude sociolectale de la variation, puisque l'auteure y note des évolutions dans les usages interrogatifs d'une époque à l'autre, selon la classe et le groupe sociaux des individus. Sur un corpus filmographique allant des années 1930 aux années 2000, il apparaît ainsi que l'utilisation des formes non-standard d'interrogation partielle (pour Farmer, les antépositions simples, les phrases in situ, et les clivées) devient ainsi plus fréquente dans les groupes qui traditionnellement les utilisaient moins. Ces groupes dont les habitudes semblent avoir changé sont les classes sociales hautes d'une part (par opposition aux classes moins favorisées), et les femmes d'autre part (par opposition aux hommes).

2.3 Pour une vision plus utilitaire de la variation linguistique

ADLI [2013] propose une vision affinée du groupe sociolinguistique en le définissant entre

autres par son mode de vie et sa confrontation habituelle à la norme linguistique. Autrement dit, l'utilisation plus fréquente d'une variante linguistique par tel sous-groupe social, en raison de contraintes par exemple professionnelles, a une influence sur les préférences linguistiques des membres de ce groupe. C'est une piste à prolonger.

Le présent travail propose de réorienter la réflexion sur la variation en replaçant les phénomènes d'alternance linguistique au coeur du phénomène langagier en lui-même. Le langage étant un outil majeur pour la communication humaine, la variation est elle-même à considérer sous un angle utilitaire. Il s'agit donc d'élargir encore le cadre d'analyse, et d'intégrer à la notion de « contexte » linguistique, non plus seulement le statut social des individus qui participent à l'interaction, mais également leurs enjeux respectifs au moment circonstancié où se déroule l'interaction.

2.3.1 Enjeux sociaux du langage et vagues de la sociolinguistique

Cette prise de recul sur le fait linguistique n'est pas neuve en soi, et s'inscrit dans ce que [ECKERT \[2012\]](#) appelle la « Troisième vague » de la sociolinguistique.

La « première » vague d'une sociolinguistique à la [LABOV \[1966\]](#) décrivait la variation sociolinguistique comme un phénomène principalement passif et examinait les rapports entre un *vernaculaire* individuel et la langue standard, sous un prisme éminemment socioéconomique et presque déterministe. Ce vernaculaire y est classiquement défini, non plus comme le système linguistique d'une communauté locale, mais plutôt comme le premier système linguistique développé, et donc celui qui est le mieux intégré, par l'individu [[LABOV, 1972](#)].

L'agentivité sociolinguistique de l'individu, et la notion de *style* personnel, se limite sous cette perspective à une capacité d'auto-correction, permettant d'étouffer les formes linguistiques les plus marquées du vernaculaire à mesure que l'individu monte les strates sociales. Des méthodologies de recueil de données de type entretiens, assorties d'analyses des productions linguistiques intra- et inter- individuelles, permettent une approche empirique renouvelée et plus systématique de la variation. Cette dernière met en lumière que les formes linguistiques marquées vis-à-vis du standard [[BOURDIEU et BOLTANSKI, 1975](#)] sont plus rares dans les productions des classes socioéconomiques les plus aisées. Par exemple, [LABOV \[1966\]](#) étudie les productions du */th/-stopping* américain (réalisation phonétique marquée en [t] dans des mots comme *thing*) dans trois contextes (informel, entretien, lecture de texte). Une comparaison inter-individuelle montre que les personnes de classes moyennes, même en contexte informel, produisent moins cette variante que les gens de classes ouvrières ou moins aisées.

Dans cette vision initiale de la sociolinguistique moderne, le changement linguistique est conceptualisé comme résultant de pressions au sein du système linguistique. Ces pressions affectent d'abord les populations les moins soumises à l'influence de la langue standard, c'est-à-dire les classes sociales les moins favorisées. Une des limites de la méthode d'enquête est cependant le besoin de prédéfinir les classes sociales supposées déterminantes de la capacité à masquer ou à jouer avec la variation sociolinguistique. Cette détermination en amont des critères de sélection de l'appartenance à telle ou telle classe sociale a pu masquer l'attitude plus dynamique des classes sociales intermédiaires vis-à-vis de la variation [[LABOV, 2001](#)]. Parce qu'elles sont souvent les plus impliquées au niveau local, l'emploi de formes vernaculaires par ces populations rend en fait ces formes des marqueurs linguistiquement socialement acceptables, au moins dans certains contextes. Cette contradiction avec le présupposé initial invite déjà à revoir la notion de variation, et à ne

plus considérer celle-ci seulement comme « une déviation depuis la langue standard qui doit être masquée », mais bien comme un outil plus actif.

À partir des années 1980 déjà [ECKERT, 1989, 2000; MILROY, 1980], la variation sociolinguistique a pris une connotation plus agentive et le vernaculaire est théorisé comme l'expression d'une identité locale, ou encore de classe, en parallèle d'autres modes d'expression comme la musique ou le style vestimentaire. Cette approche du langage en tant que fait intrinsèquement social donne aux travaux qui s'en inspirent une dimension presque « ethnographique ». Ainsi, MILROY [1980] met en relation une étude de la variation phonologique à Belfast et les réseaux sociaux denses et multiplexes de sa population ouvrière. Elle y trouve une corrélation entre la production de variantes phonologiques et les différents réseaux communautaires auxquels appartiennent les personnes interviewées. La variation sociolinguistique apparaît ici comme une marque d'adhésion à ces réseaux.

Dans son étude de lycées américains, c'est aussi ce qui ressort des travaux de ECKERT [1989, 2000]. Des comportements linguistiques différents y sont en effet manifestes entre les élèves selon leur groupe d'appartenance, « jocks » (athlètes populaires dont l'intégration à la vie de la micro-communauté du lycée est pleine et entière) ou « burnouts » (personnes plus en marge, de façon plus ou moins volontaire, mais finalement assumée dans leur langage). Cette opposition, typique des établissements scolaires américains d'après l'auteure, se traduit en effet par des usages linguistiques de formes de négation marquées par les membres de la catégories de *burnouts*, et beaucoup moins chez les *jocks*.

Si la classe sociale d'origine des *jocks* est plutôt aisée à l'échelle locale et celle des *burnouts* moins, il y a dans les données suffisamment de variation pour qu'il ne s'agisse pas d'un simple déterminisme socioéconomique. Comme chez Milroy, il semble y avoir ici une véritable affirmation de l'identité sociale par le recours à telle ou telle variante linguistique. Ces mécanismes s'inscrivent en parallèle d'autres processus d'identification au groupe comme les vêtements, le partage de goûts communs, des postures similaires de résistance ou non à l'autorité, ou encore le fait de porter des armes [CHESHIRE, 1982].

Dans cette perspective, la notion de « classe » demeure conceptuellement assez statique, et la variation permet de mettre en opposition ces classes socioéconomiques par le langage. Le recours à différentes variantes linguistiques est cependant associé à une *identité de classe* (sentiment d'appartenance), et non simplement au fait d'appartenir à une classe. Ceci invite à revoir encore la notion de variation, en ce que ces variantes linguistiques indexent des caractéristiques, sociales ou psychologiques, projetées sur l'individu qui y a recours, et qui peut choisir de les utiliser ou non.

Dans les années 2000 [BUCHOLTZ et HALL, 2005; IRVINE, 2001], la variation devient ainsi une propriété essentielle et expressive du langage à part entière, qui n'est plus la simple conséquence de la nature ou des origines sociales de l'individu. La variation est étudiée dans une approche dynamique, où l'individu présente des alternances successives dans son usage des diverses variantes sociolinguistiques qui lui sont disponibles, en fonction du contexte d'expression et de ses stratégies du moment. Ici, dans cette vision de pragmatique enrichie influencée par la théorie des jeux [GOODMAN et LASSITER, 2015; LEWIS, 1969], l'individu se construit une identité sociale adaptée à chaque instance de communication, la *persona* [OCHS, 1992]. L'énoncé acquiert alors un « sens social » [BURNETT, 2017], qui s'analyse au regard de l'interaction en cours et des enjeux qui pesaient sur la personne qui a produit cet énoncé au moment où elle l'a produit.

Cette influence contextuelle de l'interaction sur les choix faits par l'individu, en matière de production linguistique, est illustrée par exemple par LABOV [2012]. En étudiant les productions du président américain Barack Obama dans différents contextes, ce travail montre ainsi que l'homme politique utilise la variante ING sous une forme courte (*in-* dans *grillin', eatin', drinkin'*)

- dans près de 75% des cas en contexte relâché (barbecue organisé à la Maison Blanche, où le but est de donner au président l'image d'un homme ordinaire et accessible),
- seulement dans un tiers de ses productions lors d'interviews politiques (exercice exigeant une certaine formalité tout en continuant de montrer une certaine accessibilité à l'audience),
- et quasiment jamais lors de la lecture d'un discours préparé d'investiture dans une convention démocrate (exercice hautement formel, pendant une campagne politique aux enjeux très importants pour le pays).

La variation devient ici un concept nécessairement évolutif, reflet des usages qu'en font les individus dans leur construction identitaire permanente. Et ce sont ces usages qui sont en eux-mêmes le moteur du changement linguistique, selon un processus que décrit Eckert :

at some initial stage, a population may become salient, and a distinguishing feature of that population's speech may attract attention. Once recognized, that feature can be extracted from its linguistic surroundings and come, on its own, to index membership in that population. It can then be called up in ideological moves with respect to the population, invoking ways of belonging to, or characteristics or stances associated with, that population. Such an index can be used by outsiders to call up stereotypes associated with the population.

(ECKERT [2012, p.94])

Ce processus d'association entre un stéréotype et une population est visible notamment dans des expériences de *matched-guise* [LAMBERT et al., 1960]. Ces tâches de « locuteur.ice masqué.e », dont le protocole sera repris plus loin dans ce travail (Chapitre 5, Chapitre 6), permet d'explorer les associations entre une forme linguistique et des caractéristiques sociales. Pour cela, il s'agit de faire écouter ou lire des productions linguistiques (minimalement différentes, par exemple sur un point phonétique) d'une même personne à des volontaires. Ces volontaires pensent écouter plusieurs personnes différentes et doivent déterminer plusieurs caractéristiques sociales les concernant.

C'est par ce biais que CAMPBELL-KIBLER [2007] a ainsi montré un lien entre la syllabe verbale finale ING réalisée *-ing* en anglais américain et le stéréotype d'une personne mieux éduquée et plus éloquente, par comparaison avec une personne qui utiliserait la réalisation *-in'*. PODESVA et al. [2015] ont réexploré la sphère politique américaine avec un protocole similaire en comparant la réalisation de la variable /t/ selon qu'elle est prononcée battue ou relâchée par des personnalités politiques. Parmi beaucoup d'autres, les versions battues du /t/ prononcées par Barack Obama le faisaient passer auprès des volontaires de l'étude pour plus passionné dans ses discours. En parallèle, la sénatrice américaine Nancy Pelosi leur paraissait moins amicale et moins sincère lorsqu'elle produit des versions relâchées de cette variable.

C'est cette vision renouvelée et actuelle du langage qui orientera les travaux présentés dans la suite de cette thèse. Qu'il s'agisse d'études de corpus ou études expérimentales, pour affiner la vision des formes interrogatives en français, il est en effet nécessaire de considérer ces productions linguistiques particulières en les repositionnant au coeur de l'acte interactionnel (voir Section 2.3.2 ci-après).

Une autre implication majeure de ce renouvellement de l'analyse sociolinguistique concerne l'acquisition des phénomènes d'alternance linguistique. Si la variation est une propriété essentielle et expressive du langage, par laquelle l'individu se projette et construit son identité à chaque échange, l'acquisition et la maîtrise de phénomènes variationnels par l'enfant en est nécessairement marquée. Comme il sera détaillé en Section 2.4, un enfant est un individu en construction, qui doit à la fois développer sa compétence langagière de manière générale, mais aussi explorer son rapport au monde et aux codes sociaux qui domine celui-ci. Le développement de la première ne peut se faire en totale indépendance du développement des seconds.

2.3.2 Dégager les aspects sociaux de l'interrogation partielle en français

Pour ce qui est des interrogatives partielles du français, la problématique générale (chez l'adulte) dépasse celle de savoir si un groupe social francophone produit ou préfère majoritairement l'une ou l'autre variante disponible. Pour comprendre d'où viennent ces préférences de groupe si elles existent, et ce qui les maintient, il faut développer une meilleure vision des représentations mentales associées à chacune des variantes disponibles.

2.3.2.1 Interrogation partielle et contexte interactionnel

Un premier axe pour affiner ces représentations mentales doit consister à caractériser à quel point l'individu linguistique, qui construit et modifie son identité sociale, sa *persona*, à chaque instant de l'interaction linguistique, varie dans son usage des variantes interrogatives principales. Pour cela, une analyse de l'influence du contexte interactionnel sur les préférences des francophones est primordiale.

Cette notion de contexte interactionnel doit s'entendre au sens large, et peut donc recouvrir le fait de ne pas avoir le même public d'une situation à l'autre, ou d'avoir différents types de destinataires, ou encore d'avoir des buts communicationnels différents, à l'image de Barack Obama dans les différents contextes analysés par LABOV [2012]. La notion de contexte peut aussi s'entendre dans un sens plus réduit et se limiter à la notion, plus intuitive mais moins précise, de « formalité ». Il s'agit alors d'opposer des contextes informels et des contextes formels, les seconds exigeant un respect de la norme linguistique plus rigoureux (ceci, à cause d'un public, de destinataires, ou d'enjeux communicationnels plus prestigieux, plus sérieux, ou en tout cas considérés comme tels).

2.3.2.2 Interrogation partielle et stéréotypes sociaux

Derrière cette problématique de la nécessité et de la possibilité pour l'individu de s'adapter au contexte interactionnel, il faut essayer d'atteindre les représentations mentales associées à chaque variante. Si une variante linguistique est appropriée à un contexte (par exemple informel) mais inappropriée pour un autre (par exemple formel), c'est car elle véhicule des indices sociaux concordants avec le premier et discordants avec le second. Cette discordance pourrait modifier la *persona* de la personne qui l'utilise dans une direction non-souhaitée, ou incompatible avec la poursuite de l'échange dans de bonnes conditions vis-à-vis des enjeux communicationnels de cette personne. Ou encore, elle pourrait déséquilibrer la balance des relations entre les partenaires linguistiques jusqu'au point de rupture. À l'inverse, l'utilisation d'une variante appropriée peut

transmettre les éléments d'une *persona* compatible avec les choix de la personne, et favoriser le maintien d'une balance propice à la poursuite de l'échange [HEIDER, 1946; NEWCOMB, 1961].

Cette modification possible de la *persona* est à associer avec celle de *stéréotype*, et ce sont donc les stéréotypes sociaux associés à telle ou telle variante interrogative qu'il faut également dégager, pour mieux comprendre en quoi chacune est à privilégier dans tel ou tel contexte. Cette notion de stéréotype est elle-même auto-entretenu par les usages répétés, et éventuellement majoritaires, d'une variante par un sous-ensemble du groupe social, selon le processus décrit par ECKERT [2012]. Si par exemple l'usage a fait, diachroniquement, que les interrogatives partielles avec inversion sont plus utilisées dans la langue des élites intellectuelles, il pourra être associé à cette variante l'image d'une variante sophistiquée et appropriée à des contextes où il est important de se conformer à la norme linguistique. L'utilisation de cette variante dans un contexte de relâchement linguistique (informel) sera discordant, de même que son utilisation par une personne d'une classe socioéconomique incompatible *a priori* avec la maîtrise de cette « langue des élites »

2.4 Pour une étude sociale de l'acquisition des interrogatives partielles

Affiner les dimensions sociales pesant sur l'étude de la variation dans les interrogatives partielles a des conséquences importantes sur l'étude de son acquisition par les enfants francophones. Les études en acquisition du langage sont en effet nombreuses à se pencher sur les interrogatives du français, sur la base notamment des théories chomskyennes sur la grammaire universelle et la grammaire générative développées depuis CHOMSKY [1965, 1977].

De la même façon qu'il ne s'agissait pas plus tôt d'établir une présentation exhaustive des travaux analysant les interrogatives partielles en français chez l'adulte, il ne s'agit pas ici de retracer toute l'histoire des études sur leur acquisition en langue maternelle. Mais ces travaux sont généralement guidés par un même principe de complexité graduelle des phrases interrogatives et, comme chez l'adulte, cette perspective peut gagner à être combinée avec une approche plus sociale et utilitariste du langage. À nouveau, il s'agit d'intégrer la réalité des usages des différentes variantes interrogatives en français courant et de leurs implications sociales.

2.4.1 Des théories syntaxiques dominées par la notion de complexité structurelle

Selon PRÉVOST [2009], les premiers éléments interrogatifs intégrés dans une phrase apparaissent chez l'enfant francophone vers 2 ans, avec de très fortes variations interindividuelles. Sur cette thématique, un grand terrain d'investigation a été de chercher à comprendre pourquoi les enfants francophones semblaient produire plus souvent des questions in situ au début de leur développement linguistique [HULK, 1996]. Ce questionnement a donné lieu à de multiples travaux (voir par exemple HAMANN [2006]; JAKUBOWICZ et STRIK [2008]; STRIK [2007]; ZUCKERMAN et HULK [2001], ou encore PRÉVOST [2009] pour un panorama). Cette recherche s'est même intéressée à d'autres langues présentant de multiples variantes interrogatives se distinguant par la position de l'élément interrogatif (par exemple en portugais, voir SOARES [2003, 2004, 2006] ou encore BAIÃO et GONÇALVES [2014]; BAIÃO [2012]).

Dans les théories syntaxiques du mouvement de constituants, et notamment celles du *mouvement Wh-* pour ce qui est des interrogatives, les variantes interrogatives du français sont souvent vues comme dérivées d'une forme originelle. Celle-ci, comme présenté en Section 2.2.1, serait

celle de l'ordre canonique du français pour les phrases déclaratives (ordre Sujet-Verbe-Objet). L'interrogative partielle originelle serait donc la phrase in situ, et par exemple une métrique de complexité dérivationnelle (Derivational Complexity Metrics, DCM, voir JAKUBOWICZ [2005, 2011]; JAKUBOWICZ et STRIK [2008]) établit une échelle classant les autres variantes par leur éloignement depuis la forme in situ en fonction des opérations syntaxiques qu'elles imposent pour les produire, avec pour base de « calcul » les principes en (64), repris de JAKUBOWICZ [2011].

- (64) a. Merging α_i n times gives rise to a less complex derivation than merging α_i $(n + 1)$ times.
 b. Internal Merge of α gives rise to a less complex derivation than Internal Merge of $\alpha + \beta$.

Suivant ces principes, l'enfant subit dans ses productions la double influence du nombre de fois qu'un élément doit être déplacé au sein d'une question, et du nombre d'éléments qui doivent être déplacés dans la phrase. Pour reprendre les exemples déjà donnés, une telle échelle pose que (65) est plus simple, à produire comme à interpréter, que (66) qui exige un mouvement de l'élément interrogatif vers la périphérie gauche de la phrase. À son tour, (66) est elle-même plus simple que (67), qui requiert non seulement ce mouvement de l'élément interrogatif, mais aussi une inversion du verbe et du sujet (analysée dans ce cadre comme une montée parallèle du Verbe à la gauche de la phrase, dans le domaine du complémenteur).

- (65) Tu viens **quand**?
 (66) **Quand** _{z} tu viens t_z ?
 (67) **Quand** _{z} viens _{w} -tu t_w t_z ?

Cette prédiction a été testée par exemple dans JAKUBOWICZ [2011] par un protocole d'élicitation de questions. Dans cette étude, l'enfant a pour consigne explicite d'interroger Nina, un personnage présent sur un écran d'ordinateur, dans le contexte d'une histoire, avec pour résultat attendu des séquences comme (68). Les productions d'enfants au développement typique de groupes d'âges différents (3, 4 et 6 ans), mais aussi de deux groupes d'enfants avec trouble du langage (8 et 11 ans), sont comparées à celles d'adultes.

- (68) ADU : *Demande à Nina ce que Grenouille boit.*
 CHI : Nina, Grenouille boit quoi?

Dans les résultats, il apparaît que les questions locatives sont peu produites in situ (69) quel que soit le groupe considéré, mais que le phénomène est visiblement différent pour les questions objets (70). En effet, pour ces dernières, les enfants avec trouble du langage produisent plus d'in situ que les groupes des adultes et des enfants au développement typique, ce qui est pris comme une préférence de leur part pour les types phrastiques les plus simples.

- (69) Le poisson nage où?
 (70) Le lapin boit quoi?

Cette échelle de complexité a influencé non seulement des travaux en acquisition du français langue maternelle chez des enfants typiques, mais également chez des enfants au développement atypique [PRÉVOST et al., 2017] ou encore chez des apprenants du français L2 [PRÉVOST et al., 2014; SCHEIDNES et TULLER, 2010; STRIK, 2012]. Dans les données de ces travaux, il semble y avoir une préférence plus immédiate pour la structure in situ, qui n'est complétée que plus tard (ou pas), par les autres types phrastiques.

2.4.1.1 Variation dans les interrogatives prépositionnelles du français

Cette hypothèse que la complexité syntaxique d'une phrase serait liée à sa maîtrise plus tardive par les enfants pourrait trouver un écho dans une analyse de l'acquisition des interrogatives prépositionnelles du français [THIBERGE, 2017]. Ce travail est basé sur des données de corpus du projet CHILDES [MACWHINNEY, 2009], et plus spécifiquement des données de quatre enfants enregistré.e.s dans les années 2000 à Paris [MORGENSTERN et PARISSÉ, 2007; MORGENSTERN et SEKALI, 2009].

En intégrant un degré de complexité supplémentaire dans l'analyse de l'élément interrogatif, le coût de dérivation serait en effet augmenté en présence d'une préposition (71) par rapport à un interrogatif simple (72) selon l'échelle de complexité dérivationnelle de JAKUBOWICZ [2011] (principe (64b)).

(71) **À qui**_m tu penses *t*_m ?

(72) **Qui**_n tu vois *t*_n ?

En corpus, il apparaît bien que les éléments interrogatifs à préposition sont moins souvent antéposés que les éléments simples. Ce phénomène est visible chez l'adulte, dans les données par exemple de HAMLAOUI [2009a] (mais voir aussi [REINHARDT, 2019, p. 161] pour des données allant dans le même sens avec 80% de phrases in situ quand l'élément interrogatif est prépositionnel, contre un peu moins de 40% pour les éléments non-prépositionnels). Il est aussi relevé dans THIBERGE [2017] pour des données partielles des parents des enfants du projet CHILDES.

Il est intéressant de voir qu'une telle différence entre phrases interrogatives avec ou sans préposition est encore plus importante pour les enfants, chez qui la proportion d'éléments prépositionnels produits in situ équivaut quasiment à 100% des productions (73). Ce, alors que leurs productions in situ d'éléments interrogatifs sans préposition se situent plutôt entre 60 et 85 %. (74) illustre une stratégie utilisée par les enfants autre que l'in situ, avec une antéposition simple.

(73) il est **dans quelle** classe

(ANA - 3;04.27)

(74) **où** il est le bébé qui est un garçon

(ANA - 3;01.07)

Cette différence entre enfants et adultes, associée au fait que les premières productions d'interrogatives prépositionnelles par l'enfant sont plus tardives que les premières productions d'interrogatives sans préposition, pourrait corroborer l'hypothèse d'un coût plus important des structures phrastiques plus complexes.

2.4.2 L'hypothèse de la diglossie

Pour en revenir aux interrogatives partielles de manière générale, si la notion de complexité structurelle ou dérivationnelle peut expliquer les données observées en acquisition, ce n'est pas la seule hypothèse qui ait été proposée. Il arrive par exemple que des enfants produisent des interrogatives in situ plus tardivement que d'autres types (par exemple un enfant chez qui l'antéposition simple semble apparaître d'abord [CRISMA, 1992; HULK, 1996]). HAMANN [2000] envisage ainsi le rôle de l'input parental sur les premières productions enfantines.

FAURE et PALASIS [2020]; PALASIS [2013]; PALASIS et al. [2019] vont plus loin en argumentant pour l'existence de deux grammaires françaises. Sur le plan des interrogatives partielles, la grammaire de la première langue, G1, serait une « langue in situ » et serait la langue première apprise par les enfants, celles dont l'usage leur est quotidien. La période de scolarisation coïnciderait avec le développement d'une autre grammaire, G2, utilisée en contexte normé. Cette seconde grammaire comprendrait, entre autres, la possibilité d'antéposer le mot interrogatif dans les questions partielles (mais aussi par exemple le fait d'utiliser la négation double en « ne... pas » ou encore l'utilisation préférentielle du « nous » au lieu de « on »).

Cette hypothèse n'explique pas tout. Notamment, elle n'explique pas directement le cas des enfants qui ne semblent pas produire de phrases in situ en première phase de développement (sauf à considérer que les parents utilisent le système G2 avec leurs enfants). Cependant, elle n'est pas si éloignée des arguments de ADLI [2006] distinguant divers registres de langue pour les différentes variantes interrogatives. Et, plus largement, l'hypothèse d'une diglossie et de la cohabitation chez l'enfant d'une grammaire vernaculaire des interrogatives, conforme aux usages linguistiques du groupe social des proches, et d'une grammaire normative, conforme aux usages linguistiques d'un groupe plus large introduit à l'enfant en premier lieu par l'école, est à mettre dans la lignée des études sociolinguistiques menées chez l'adulte.

2.4.3 Un développement parallèle des capacités linguistiques et sociales

Il est en effet concevable que l'enfant, en même temps qu'il acquiert et développe les mécanismes fondamentaux qui lui permettent d'appréhender le langage, acquiert et développe plus largement les mécanismes fondamentaux qui lui permettent de sociabiliser et de se confronter à la réalité du monde concret. Ce développement parallèle des capacités langagières, cognitives et sociales, a déjà été exploré. Comme chez l'adulte, il est cependant utile d'affiner l'impact que peuvent avoir les enjeux sociaux du langage sur l'acquisition puis la maîtrise d'un phénomène d'alternance linguistique tel que les interrogatives partielles en français.

2.4.3.1 Acquisition du langage, théorie de l'usage et théorie de l'esprit

Développée en réaction aux théories chomskyennes faisant du langage un module inné, l'idée d'une grammaire basée sur l'usage n'est pas neuve [LANGACKER, 1987]. Elle a été appliquée dans le domaine de l'acquisition par exemple par TOMASELLO [2000, 2009a,b], qui argumente pour une acquisition de la grammaire par analogie entre énoncés autour de « pivots » permettant l'émergence d'une syntaxe à proprement parler. Dans cette perspective, c'est bien l'input, et entre autres l'input parental, qui entraînerait l'intégration par l'enfant de tel ou tel type interrogatif comme variante « principale » pour poser une question partielle.

En parallèle, si les enjeux sociaux du langage s'appliquent à l'adulte comme vu plus haut (Section 2.4), il est inévitable qu'ils entrent aussi en jeu dans le développement langagier de l'enfant. Or, pour interagir avec autrui et avoir une idée des enjeux sociaux du langage, pour avoir la conscience du besoin ou de la possibilité de maintenir et modifier sa *persona* au cours d'une interaction linguistique, il faut avoir acquis la capacité à imaginer que la personne en face, dans l'interaction, n'ait pas le même cadre mental, et notamment le même cadre social que soi. C'est ici que les travaux menés sur le développement d'une Théorie de l'Esprit par l'enfant (*Theory of Mind*, [WIMMER et PERNER, 1983]) prennent leur importance.

Classiquement, il était considéré que c'est vers 4 ans que l'enfant devient capable de falsifier son propre état d'esprit pour en imaginer un différent chez autrui (étape dite du *false belief* ou de la *fausse croyance*). Certains travaux plus récents [BUTTELMANN et al., 2009] laissent entrevoir que l'enfant pourrait être capable bien plus tôt, dès 18 mois, de projeter un état mental différent du sien sur des adultes pour déterminer leurs buts et leur venir en aide. TOMASELLO [2018] établit toutefois un pont théorique original entre ces travaux récents et le consensus ancien sur une période critique située vers 4-5 ans.

Pour lui, les enfants les plus jeunes peuvent certes projeter un état mental sur autrui (comme certains grands singes), mais ce ne serait que plus tard, vers 4-5 ans qu'un processus « d'intentionnalité conjointe » (« *shared intentionality* ») se développerait. Ce processus serait celui qui permettrait à l'enfant d'opérer une « triangulation » entre son propre état mental, celui de la personne lui faisant face, et la réalité de la situation d'interaction. Le développement de certaines capacités exécutives, ajouté à une spécificité de l'espèce humaine qui la distinguerait des grands singes sur ce point, permettrait l'apparition de cette pleine capacité de fausse croyance vers 4 ou 5 ans. C'est cette capacité pleine à trianguler qui permettrait une prise en compte et une réflexion consciente de l'enfant sur les différentes perspectives subjectives concurrentes chez de multiples individus.

En parallèle de ce processus, l'enfant est également en train d'apprendre à adapter sa production linguistique pour tenir compte d'états d'esprits supposés chez le ou la destinataire [VENEZIANO, 2009, 2015]. Si l'influence de facteurs comme le contexte interactionnel et l'association de stéréotypes sociaux à certaines formes linguistiques est à envisager dans les productions interrogatives de l'adulte, leur portée doit être envisagée chez l'enfant différemment avant et après le plein développement de la capacité de *false belief*. Avant, ces facteurs ne peuvent être qu'en cours d'intégration, et ce de façon hautement passive et non-consciente. Par la suite, ces facteurs ne sont pas encore nécessairement intégrés dans leur totalité et une phase d'adaptation doit encore être nécessaire, ne serait-ce que pour des raisons pratiques. Les contextes sociaux auxquels les enfants se confrontent ne se diversifient en effet qu'avec le temps, par exemple grâce et au cours de la phase scolaire, durant laquelle la proportion d'adultes augmente ainsi que les exigences de formalité et de respect des normes sociales par l'individu.

2.4.3.2 Développement sociolinguistique, saillance et stéréotypes

Le développement d'une conscience des enjeux sociaux du langage chez l'enfant n'est pas un domaine délaissé par la littérature sociolinguistique. LABOV [1964] envisageait ainsi déjà un développement progressif de la sensibilité et de l'adaptabilité des enfants au contexte social de l'interaction. Dans ce cadre, avant 5 ans, le stade de la « grammaire de base » postulait seulement une acquisition puis une connaissance simple par l'enfant des règles de base et du lexique de sa langue. C'est entre 5 et 12 ans que la notion de « vernaculaire », dialecte local ou des proches, prend son sens, et ce sont les variantes de cette langue que les enfants apprennent à maîtriser à ce stade. C'est ensuite au cours de l'adolescence, en deux phases successives (12-14 ans puis après 14 ou 15 ans), que l'individu développerait et forgerait sa capacité de perception sociale (différence entre le vernaculaire internalisé et d'autres dialectes) puis sa maîtrise de la variation stylistique (traduite par une capacité à se conformer à la langue « standard » dans des situations formelles).

Les points de langue sur lesquelles ont été menées les études sociolinguistiques chez l'enfant ont souvent été d'ordre phonologique, et ce dès LABOV [1964]. Plus récemment et pour l'anglais,

ROBERTS [1994] note que, en entretien, des enfants dont l'âge varie entre 3;2 ans et 4;11 ans semblent moduler leur usage différents des variantes de la terminaison ING (*in'* ou *ing*) en fonction de si leurs destinataires sont des adultes ou s'il s'agit aussi d'enfants. Quand il s'agit d'adultes, les enfants sélectionnent plus souvent la variante standard - *ing*.

En français, NARDY [2008] et CHEVROT et al. [2011] s'intéressent au recours à la liaison facultative en français à partir de productions de 185 enfants dont l'âge s'étend de 2;3 ans et 6;0 ans. Par une tâche de description d'images, ces auteurs montrent que la production de liaisons facultatives est plus fréquente chez les enfants de classe sociale aisée que chez les enfants de classe sociale moins aisée. Ceci semble donc en partie le reflet d'une tendance sociolectale (les enfants se comportent selon un patron similaire aux adultes), mais il y a aussi une part développementale. En effet, un autre résultat de ces études est que ces différences entre classes sociales d'origine augmentent avec l'âge. Les enfants de 5-6 ans des classes plus aisées finissent ainsi par produire jusqu'à deux fois plus fréquemment des liaisons facultatives que les enfants de classes moins aisées. NARDY et al. [2013] donne un panorama plus général de la confrontation du domaine de la sociolinguistique et des études en acquisition du langage, et des perspectives en la matière.

Il reste relativement rare d'étudier en quoi ce développement d'une compétence sociolinguistique se fait en parallèle du développement social de l'enfant et par un processus d'alignement avec son milieu social et les autres enfants notamment [MARTIN et al., 2010; NARDY et al., 2014]. L'extrême variabilité interindividuelle entre les enfants rend la tâche très ardue, que cette variabilité soit due à leurs origines socioéconomiques très diverses, ou qu'elle résulte de différences d'accès à la langue (lesquelles sont souvent liées aux origines socioéconomiques du foyer) [LAHIRE, 2019].

BUSON [2009]; BUSON et al. [2014] sont des travaux qui, au-delà d'un point de langue précis, essaient de modéliser les processus par lesquels l'enfant intègre la variation (qualifiée de « stylistique » dans ces travaux). L'idée est qu'il est certes pour l'enfant besoin d'apprendre à s'adapter sociolinguistiquement, mais que cette adaptation passe par un jeu constant de statistiques : plus une forme est fréquente, et plus l'enfant peut noter, répéter et tester les variantes éventuelles. Sur un modèle à la Bybee et sa théorie des exemplaires (BYBEE [2013]; BYBEE et HOPPER [2001], et voir AMBRIDGE [2019] pour une vision encore plus radicale), Buson propose que l'enfant mémorise non pas seulement une forme linguistique, mais aussi son contexte social d'utilisation.

Par un effet combiné de la saillance (de parties plus visibles que d'autres des énoncés) et de stéréotypes (associations récurrentes entre formes et contextes interactionnels), l'enfant actualise en permanence ses schémas stéréotypiques. Cette actualisation se fait par une mise en ligne des nouveaux éléments avec les éléments déjà stockés en mémoire (les exemplaires). De ce fait, il se construit peu à peu chez l'enfant une représentation du langage qui tient compte de ses utilisations contextuellement appropriées. Cette activation régulière des schémas/stéréotypes permet l'émergence d'un méta-discours sur le message et d'une conscience sociolinguistique, laquelle permet au besoin une restauration stylistique quand une forme ne convient pas à un contexte social. Dès 9-11 ans, cette conscience métalinguistique peut même être explicitement exprimée par l'enfant, qui devient capable de dire des choses comme (exemples tirés de BUSON [2009, pp. 141-144]) :

- « *"salut les gars" ça se dit pas à tout le monde* »,
- « *le [premier] il parle bien français et le jeune il parle pas vraiment correctement* ».

2.4.4 Un appariement progressif entre forme interrogative et charge sociale?

D'autres exemples de méta-discours produits par des enfants [BUSON, 2009, p. 144] attirent l'attention, en ce qu'ils révèlent cette association plus ou moins consciente entre une forme et un contexte social précis d'interaction :

- « le [troisième] il parle un peu comment dire un peu **comme une racaille** »,
- « le [premier] il est poli et l'autre **il parle voyou** »,
- « ils ont pas dit les mêmes mots, parce que il y en a un qui est plus, par exemple qui travaille dans un métier et il parle correctement et tout. **si il serait pas pauvre il aurait mieux parlé** »

En dehors des connotations négatives particulières à ces exemples, il en ressort qu'au moins deux charges sociolinguistiques différentes sont appariées aux énoncés linguistiques (en l'occurrence, des messages de répondeurs téléphoniques que les enfants écoutaient). La première est qu'une variante linguistique peut faire penser à un groupe social particulier (à mettre en lien avec les études faisant de la variation le reflet d'un sociolecte propre à un sous-groupe social), ici marginalisé. La seconde charge sociolinguistique va dans la même direction mais dépasse cette idée de groupe social, dans le sens où elle apparie directement une forme linguistique à une caractéristique sociale, ici la pauvreté.

Dans les travaux de Buson, cet appariement est progressif, et plus conscient chez les enfants de 9-11 ans que chez les plus jeunes (aux variations interindividuelles près). Il s'en déduit que c'est bien progressivement que se cristallisent les stéréotypes, sans doute par cette intégration toujours plus importante d'exemplaires « sociolinguistiques », à mesure que l'enfant se confronte à des contextes d'interaction variés.

Sur le point des interrogatives partielles en particulier, et en fonction des charges sociales véhiculées par les différentes formes interrogatives chez l'adulte (Section 2.4), il faut s'attendre à un même phénomène, tout aussi progressif. Il est nécessaire d'explorer quels stéréotypes présents chez l'adulte sont construits par l'enfant et quand. Si ces stéréotypes sont présents mais ignorés par les enfants plus jeunes, il est important de voir aussi suivant quelle progression ils sont assimilés ou modifiés, jusqu'à peser sur les préférences par exemple en production spontanée. De même, la notion de contexte interactionnel étant visiblement un élément auquel les enfants savent qu'il est important de s'adapter dès 10 ans (« *"salut les gars" ça se dit pas à tout le monde* », BUSON [2009, p. 141]), il faut observer quand et dans quelle mesure ceci s'opérationnalise chez l'enfant.

2.5 Bilan et résumé des axes de recherche

La littérature linguistique sur les phrases interrogatives partielles et leur diversité en français est déjà très étayée. De nombreuses théories ont été formulées pour tenter d'expliquer ce phénomène variationnel, tant au niveau de la phrase elle-même qu'au niveau du discours, et même au niveau plus large de la situation sociale de l'interaction linguistique. Ces travaux ont différemment intégré les facteurs potentiels à l'oeuvre (complexité structurelle, structure informationnelle, changement linguistique, particularités dialectales), selon qu'il était question d'usage adulte ou d'acquisition des interrogatives par l'enfant.

Ces études n'ont cependant pas intégré complètement la dimension utilitaire du langage et le fait que la disponibilité de multiples formes linguistiques permet aussi de faire passer différentes

nuances de sens social. Ce dernier repose sur des associations de type stéréotypiques, par lesquelles différentes charges sociales finissent par être attribuées à différentes variantes (par une combinaison des effets temporels d'évolution de la langue et sociolectaux, sans doute).

Dans les études empiriques actuelles, il existe par ailleurs une certaine variabilité quant aux chiffres avancés sur les usages réels faits, notamment par l'adulte, des variantes principales du français (in situ, antéposition simple, antéposition en « est-ce que », antéposition avec inversion verbe-sujet).

Pour affiner encore une vision multifactorielle régissant l'usage des interrogatives partielles en français, il est donc nécessaire :

1. **d'actualiser les données empiriques sur les usages et les préférences francophones quant aux principales variantes d'interrogation partielle**, par des études de corpus et les méthodes expérimentales (Chapitres 4 et 5).
2. **d'observer les associations forme-contexte et forme-stéréotype chez l'adulte**, par des études de corpus mais aussi par des méthodes expérimentales (Chapitres 4 et 5).
3. **de confronter les résultats de ces investigations chez l'adulte au domaine de l'acquisition, pour voir comment ces associations se forment chez l'individu au cours de son développement linguistique**, par des méthodes expérimentales adaptées permettant la comparabilité des données (Chapitre 6).

Chapitre 3

Méthodologies retenues et analyses effectuées

Sommaire

3.1 Méthodologies de recherche retenues	41
3.1.1 Études de corpus	41
3.1.2 Études expérimentales	41
3.1.2.1 Jugements d'acceptabilité (adultes)	42
3.1.2.2 Matched-Guise (adultes et enfants)	42
3.1.2.3 Production spontanée (enfants)	43
3.1.3 Données et éthique de la recherche	43
3.2 Représentations graphiques	44
3.2.1 Graphiques à barre	44
3.2.2 Diagrammes en violon	45
3.2.3 Matrices de corrélation	46
3.3 Analyses bayésiennes et distributions postérieures	47

Ce chapitre présente brièvement les choix faits en matière méthodologique pour répondre aux questions de recherche soulevées au Chapitre 2. Il s'agit notamment d'expliquer les parti-pris méthodologiques (corpus et études expérimentales) et comment lire les visualisations proposées pour les données, ainsi que de décrire le choix fait de recourir aux statistiques inférentielles bayésiennes.

3.1 Méthodologies de recherche retenues

Comme précisé en introduction, l'approche scientifique choisie se veut résolument empirique, et seront présentées aussi bien des données de corpus que des données expérimentales.

3.1.1 Études de corpus

Dans le Chapitre 4, deux corpus de français contemporain et oral principaux seront analysés, résultats des projets EPAC [BAZILLON et al., 2011; ESTÈVE et al., 2010] et ESLO [ESHKOL-TARAVELLA et al., 2011; LLL, 2017]. Pour le premier corpus, ces données ont pour intérêt principal d'apporter un éclairage sur les usages en interaction des interrogatives partielles en français, permettant une étude des stratégies discursives et interactionnelles à contexte constant. La nature du second permet une analyse des interrogatives sous trois perspectives combinées : diachronique, sociolinguistique et interactionnelle.

L'analyse d'autres corpus aurait bien sûr été d'un intérêt indéniable, mais le temps et la méthodologie nécessaire ont conduit à remettre à plus tard l'exploration, notamment, des projets :

- French Question Bank (FQB) [BONFANTE et al., 2018; JUDGE et al., 2006; SEDDAH et CANDITO, 2016] : l'origine multiple de ce corpus écrit dédié à l'entraînement d'un annotateur automatique a rendu son analyse complexe pour une première approche ;
- Corpus de Français parlé parisien des années 2000 (CFPP) [BRANCA-ROSOFF et al., 2000; CLESTHIA, 2018] : ce corpus étant spécialement centré autour de la région parisienne, il n'aurait pas permis, pour une première approche, d'avoir une vision vraiment générale de l'usage des interrogatives en français ;
- Acquisition de la Liaison et Interactions Parents Enfants (ALIPE) [LIEGEOIS, 2014; LIÉGEOIS et al., 2014; LRL, 2017] : ce corpus a été constitué dans le cadre d'une étude sur la variation et l'emploi des liaisons par les enfants francophones. Il était donc un complément tentant aux études sur les interrogatives partielles prépositionnelles mentionnées en Section 2.4.1.1 et menées sur le corpus CHILDES *parisien* [MORGENSTERN, 2006; MORGENSTERN et PARISSÉ, 2007]. Si les tâches utilisées pour le constituer n'ont pas permis une production nombreuse des questions dans ALIPE, ce sont pourtant les quelques exemples qui s'y trouvent de questions partielles, produites par les enfants en situation de jeu de rôle, qui ont inspiré la tâche de production décrite au chapitre 6.3.

3.1.2 Études expérimentales

Les Chapitres 5 et 6 apportent des données expérimentales nouvelles, à confronter aux données de corpus. La nécessité de recourir aux travaux expérimentaux tient en trois éléments principaux :

- *systematisation* des données recueillies, notamment par des designs contrebalancés et randomisés à travers les participantes et participants, par le recours généralisé aux principes du

carré latin, ou encore par la limitation des points de variation entre items à un seul point de variation entre conditions.

- *analyses inférentielles* rendues possibles par ce caractère systématique et contrôlé du recueil de données, notamment par le recours à des statistiques bayésiennes (voir infra section 3.3).
- *réplicabilité* des résultats présentés, notamment par leur mise à disposition de la communauté scientifique (via notamment le projet Open Science Framework (OSF, FOSTER et DEARDORFF [2017])); les résultats bruts présentés dans la suite de cette thèse et les analyses sont ou seront accessibles via le lien osf.io/ud4ks).

Les différentes méthodologies expérimentales retenues sont décrites ci-dessous.

3.1.2.1 Jugements d’acceptabilité (adultes)

Cette méthode a été déclinée selon diverses versions, qui seront décrites plus précisément dans les Sections concernées. Globalement, l’emploi de ce protocole visait à explorer la représentation générale qu’ont les francophones (adultes) sur les différents types d’interrogatives partielles. Dans la lignée de travaux comme HARRIS [1968]; ZLATKA et DESCLÉS [1991], cette représentation couvre moins la notion de *grammaticalité stricte* (« Est-ce que cette phrase appartient à la grammaire du français? ») que celle de *préférence* (« Toutes ces phrases appartenant au français, quel regard ont sur elles les francophones et par quels facteurs ce regard est-il modulé? »).

Les différentes études menées chez l’adulte, décrites dans le Chapitre 5, ont ainsi d’abord visé à établir la « meilleure forme interrogative », ou celle qui serait la plus « acceptable ». Les résultats étaient discordants avec les données de corpus antérieures (Section 2.1.2.2) et celles présentées au Chapitre 4. Il a donc fallu affiner la notion d’acceptabilité. Les études suivantes ont donc tenté de distinguer ce qui, pour les francophones, ressort du *Bon Français* et ce qui n’en est pas nécessairement mais qui est adapté malgré tout au contexte illocutoire. La notion de contexte (formel ou informel) est par exemple une variable qui a alors été manipulée. Tous ces travaux sont décrits dans les Sections 5.2.2-5.2.4.

3.1.2.2 Matched-Guise (adultes et enfants)

En dehors des préférences pour telle ou telle variante par les francophones, un des buts de cette thèse était de dégager les caractéristiques sociales associées à celles-ci. Pour ceci, le concept des expériences de *matched-guise* de LAMBERT et al. [1960] a été repris. Ce type d’études est désormais fréquemment utilisé en sociolinguistique contemporaine dans une perspective de type « Troisième vague » (par exemple par CAMPBELL-KIBLER [2007]; PODESVA et al. [2015], voir Section 2.3.1).

Ici, il s’agissait de confronter des francophones à diverses variantes interrogatives, soit à l’écrit soit à l’oral, puis de leur demander de situer la personne à l’origine de ces phrases sur différentes échelles sociales. Ces échelles, graduées, portaient sur des critères tels l’âge, la richesse ou encore l’éducation. Les résultats obtenus permettent de se faire une idée des associations entre stéréotypes et formes linguistiques qui existent, consciemment ou non, chez les francophones. Ainsi, il devient possible d’esquisser un « portrait social » *type* des personnes utilisant telle ou telle variante.

Pour les adultes, ces études sont décrites dans les Sections 5.3.2-5.3.3. Une version allégée de cette méthodologie a été créée, à destination d’un public d’enfants francophones âgés de 3 à 11 ans, pour explorer le parallélisme de ces associations entre enfants et adultes. L’évolution de ces

associations entre forme linguistique et stéréotype social chez l'enfant a aussi pu être observée, puisque cette tâche a été passée auprès d'un nombre important d'enfants d'âges très variés. Les résultats sont décrits en Section 6.2.

3.1.2.3 Production spontanée (enfants)

Les corpus de production enfantine sont nombreux en français, avec le corpus *parisien* du projet CHILDES par exemple [MACWHINNEY, 2009; MORGENSTERN, 2006; MORGENSTERN et PARISSÉ, 2007] mais aussi le corpus ALIPE [LIEGEOIS, 2014], déjà mentionnés. Il reste cependant difficile d'avoir une vision précise de la production vraiment spontanée d'interrogatives par l'enfant. La tâche d'élicitation de JAKUBOWICZ [2011], mentionnée en Section 2.4.1, illustre ces difficultés. L'équipe de recherche y sollicite ainsi explicitement l'enfant en lui demandant de produire une interrogative, tout en produisant déjà une interrogative (certes indirecte) (75).

(75) ADU : *Demande à Nina ce que Grenouille boit.*

CHI : Nina, Grenouille boit quoi?

Inspirée par des extraits du corpus ALIPE, une tâche de jeu de rôle élicitant de façon systématique et contrôlée la production de cet acte de langage particulier a été développée. Les enfants y incarnaient le patron ou la patronne d'un restaurant, et de manière très ludique et sans contrainte spécifique, il leur était demandé d'accueillir divers personnages (sous forme de marionnettes), dont il leur fallait s'occuper. Cette tâche a permis une exploration des capacités actives des enfants à adapter leur production linguistique à une variation du contexte interactionnel (formel ou informel). Les données obtenues sont décrites en Section 6.3.

3.1.3 Données et éthique de la recherche

Les données recueillies et traitées dans toutes ces études l'ont été dans un respect strict des règles de l'éthique, avec plusieurs grands principes :

- *anonymisation* : l'ensemble des productions recueillies l'a été d'une manière empêchant de remonter jusqu'aux personnes qui les ont produites. Aucune donnée nominale ou personnelle permettant l'identification des personnes n'est ou n'a été fournie ni stockée sur les serveurs de dépôt ou de travail en ligne (OSF, Overleaf, R Studio Server Pro, communications mail).
- *consentement éclairé* : les volontaires aux études expérimentales ont lu puis signé un formulaire de consentement après une information exhaustive de leurs droits et des limites du recueil des données. Ceci comprenait le droit de retirer leur participation, de consulter ou demander la suppression de leurs données, pour les adultes comme pour les enfants (via leurs parents au préalable, et le jour des passations en leur demandant directement).
- *contrôle éthique* : les études impliquant des enfants et décrites au Chapitre 6 ont été menées après consultation des autorités compétentes (Direction Académique de Caen, Université de Paris, direction et équipe pédagogique de l'école de passation), ainsi que d'un Comité d'Éthique de la Recherche (CER de l'Université Paris Descartes). Les documents afférents à cette procédure sont compilés dans l'Annexe N.

3.2 Représentations graphiques

Pour une visualisation optimale des données présentées, qu'elles proviennent d'analyses de corpus ou d'études expérimentales, il est important de définir ce que chaque type de visualisation choisi permet et comment les lire.

3.2.1 Graphiques à barre

Pour une présentation des données de corpus, en parallèle de tableaux présentant les valeurs brutes, des graphiques à barre permettront une visualisation des proportions sur le modèle suivant (repris de la Section 4.1.3.1, Tableau 4.3 et Figure 4.1) :

	Antéposition simple	Antéposition + ESK	Antéposition et inversion	In situ
Société	6	1	8	1
Santé	6	3	10	3
Économie	4	4	16	2
Culture	7	2	22	5
International	9	3	24	3
Politique	10	5	27	4

TABLEAU 3.1 – Exemple de tableau de valeurs brutes

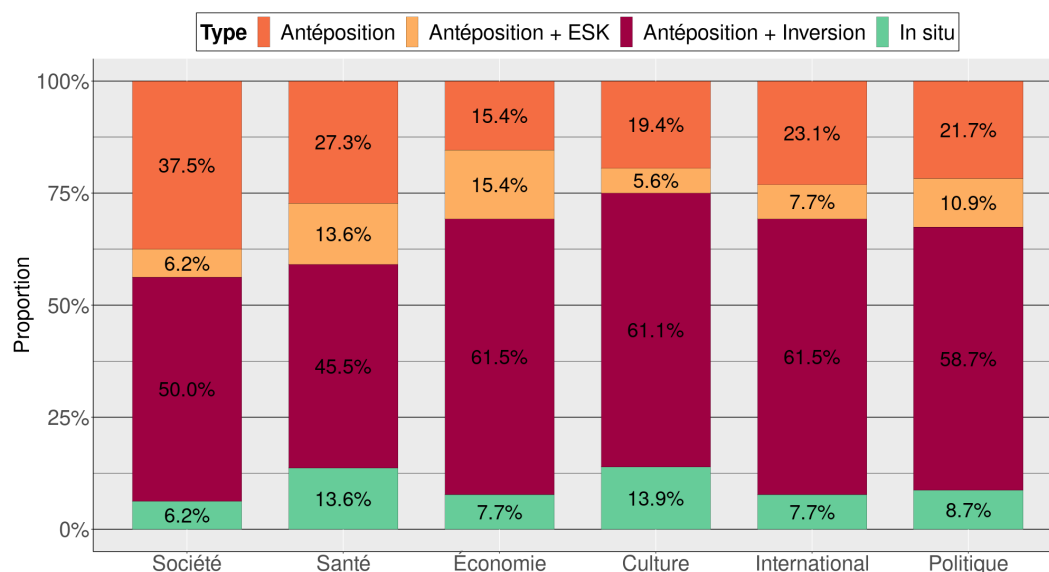


FIGURE 3.1 – Exemple de graphique à barres représentant des proportions

Le tableau permet une lecture des valeurs numériques brutes issues des extractions menées sur le corpus, en croisant les différents niveaux de la variable d'intérêt principal (variable dépendante, ici, le type syntaxique d'interrogative partielle) et les différents niveaux d'une autre variable (en l'occurrence des sous-ensembles du corpus principal, définis par la thématique abordée dans les transcriptions).

Le graphique permet une visualisation des proportions calculées à partir de tableau pour chacun des niveaux de la variable analysée (ici, le type syntaxique des interrogatives, en couleur, dont la légende est donnée en tête de figure). En abscisse du graphique se trouvent les sous-ensembles d'intérêt, divisions du corpus principal, et en ordonnée sont visibles les proportions

calculées, avec un rappel en milieu de colonne pour chaque niveau de la variable dépendante. Il est important de noter que les proportions calculées le sont par sous-ensemble, ce qui signifie que chaque colonne a pour somme 100%. Ceci permet une comparaison globale des sous-ensembles les uns avec les autres. Sur le graphique d'exemple ici, il y a en proportion plus d'*Antépositions* + *ESK* pour le sous-ensemble « Société » que pour les autres sous-ensembles.

3.2.2 Diagrammes en violon

Pour la représentation des données ordinales issues des études expérimentales (notes rendues sur des échelles graduées, qu'il s'agisse de jugements d'acceptabilité, de jugements sociaux, etc.), c'est sur des diagrammes en « violon » (*violin-plot*) que le choix s'est porté. Ces graphiques permettent une visualisation de la distribution des données. Trois lignes blanches indiqueront sur les diagrammes les quartiles (respectivement, en partant du bas, 25% des données, médiane, 75% des données), et la moyenne sera indiquée en noir, marquée par un point affublé de barres d'erreurs (erreur-type sauf mention contraire). Ci-dessous se trouve un modèle, repris de la Section 5.2.2, Tableau 5.3 et Figure 5.1.

Antéposition simple	Antéposition et inversion	In situ
6.23	8.01	6.79

TABLEAU 3.2 – Exemple de tableau de moyennes brutes

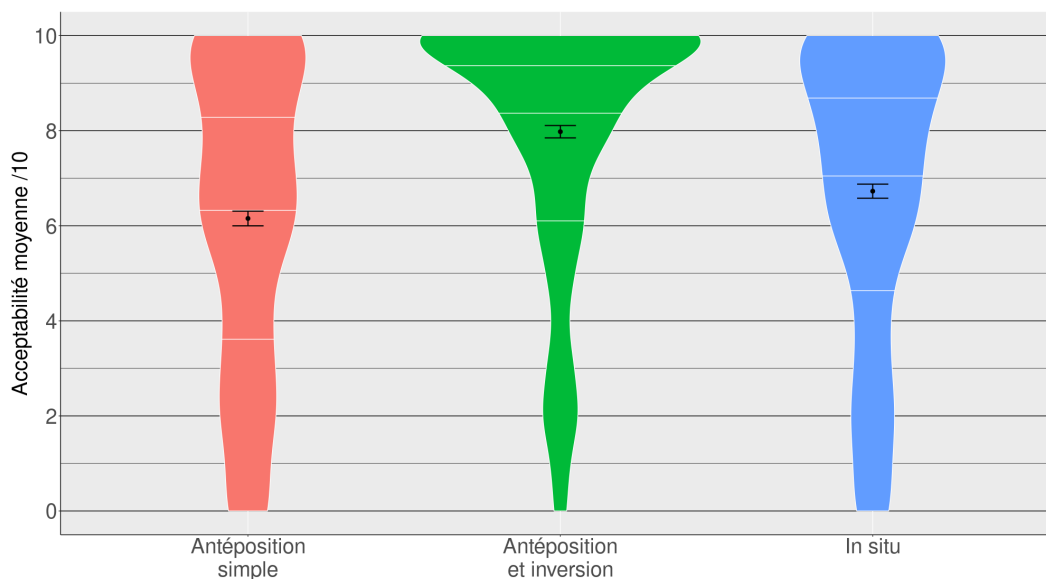


FIGURE 3.2 – Exemple de visualisation de la distribution de données ordinales

Le tableau présente une valeur unique par niveau de la variable dépendante (ici le type syntaxique des interrogatives), correspondant à la valeur moyenne des réponses fournies par l'intégralité des personnes ayant participé à l'étude et dont les réponses ont été conservées pour analyse. Les moyennes sont une information typiquement donnée dans ce type d'étude, mais cependant insuffisante pour caractériser la distribution des réponses sur une échelle ordinale comme des jugements d'acceptabilité. Le graphique permet ainsi une visualisation plus précise des valeurs à l'origine de ces moyennes, avec en abscisse les sous-ensembles des données (ici les différents niveaux de la variable dépendante, à savoir le type syntaxique), et en ordonnée les différents degrés

de l'échelle sur laquelle les volontaires pouvaient répondre (ici un jugement d'acceptabilité entre 0 et 10). Sur cette représentation, plus de la moitié des volontaires a noté les *antépositions avec inversion*, en vert, au-dessus de 8/10 (médiane), et même près des trois quarts ont donné une note supérieure à 6/10 (premier quartile) pour ce sous-ensemble. Cette distribution très décentrée des données contraste avec les deux autres niveaux de la variable dépendante (en rouge et bleu, respectivement *antéposition simple* et *in situ*), pour lesquels les notes sont réparties de façon beaucoup plus homogène sur l'ensemble de l'échelle.

3.2.3 Matrices de corrélation

Il est souvent important de calculer la corrélation existant, le cas échéant, entre plusieurs variables indépendantes utilisées comme prédicteurs de modèles statistiques (pour savoir s'il est utile ou souhaitable de calculer une interaction entre elles par exemple, ou les limites d'un tel calcul). Autant que possible, des matrices de corrélations seront présentées pour illustrer la portée de ces corrélations. La Figure 3.3 illustre ces matrices, reprise de la Section 4.1.4, Figure 4.4.

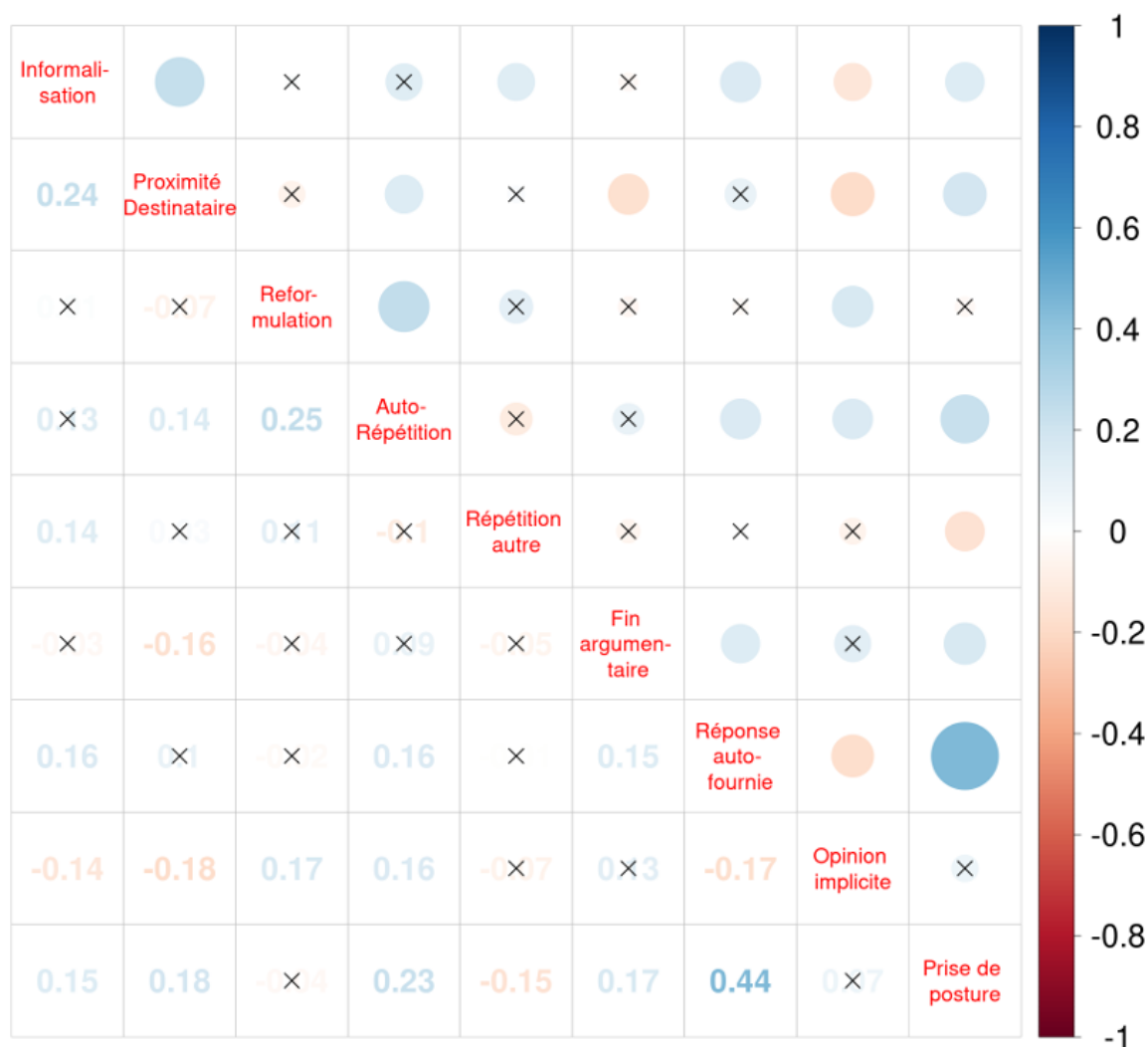


FIGURE 3.3 – Exemple de matrice de corrélation entre variables indépendantes

Les informations fournies ici sont nombreuses :

- en diagonale, en rouge, se trouvent les différentes variables pour lesquelles une corrélation

a été calculée (sauf précision contraire, par la méthode Spearman, adaptée aux variables ordinales).

- dans la partie inférieure gauche se trouvent les coefficients de corrélation (r), dont la valeur entre -1 et +1 indique la corrélation entre les deux variables croisées. -1 ou +1 indique une relation entre elles (respectivement, négative ou positive). Une valeur de 0 indique l'absence de relation entre elles.
- dans la partie supérieure droite, une visualisation de ces coefficients est présentée, avec des cercles de couleur dont la taille dépend de la valeur du coefficient, et la couleur de son signe (voir légende à droite, avec le rouge pour les valeurs négatives et le bleu pour les valeurs positives). Les coefficients les moins élevés sont faiblement colorés pour laisser ressortir les plus importants.
- les croix sur le graphique représentent les corrélations pour lesquelles la p.value est supérieure à, pour ce travail, 0.05 (α ou valeur *alpha*), indiquant que le coefficient calculé n'est pas statistiquement significatif. Les valeurs numériques associées à cette significativité sont reportées, soit dans le texte soit en annexe.

Sur ce graphique d'exemple, une corrélation intéressante (parmi d'autres) est ainsi celle entre *Réponse auto-fournie* et *Prise de posture* ($r = .44$, $p < 0.01$). Il y a une relation linéaire entre ces deux variables, et cette relation est positive.

3.3 Analyses bayésiennes et distributions postérieures

Il est possible d'inférer des généralisations statistiques à partir d'échantillons réduits de données, que ceux-ci soient extraits de corpus ou d'études expérimentales. Ces généralisations doivent néanmoins être faites de façon robuste et c'est ici le raisonnement bayésien qui a été privilégié. Représenter graphiquement ces inférences aide à les comprendre, ce qui ensuite permet de les rapprocher des valeurs numériques relevées (chiffres bruts et proportions, ou moyennes simples, calculées sur ces chiffres).

Une première raison au choix de modèles inférentiels bayésiens est que ceux-ci sont calculés directement sur l'hypothèse testée, par contraste avec des raisonnements visant à invalider une *hypothèse nulle* (inverse). Ce choix résulte par ailleurs d'une volonté de se détacher de la notion de significativité statistique propre aux analyses fréquentistes, ce qui permet une caractérisation plus fine des effets visibles (ou non) dans les données, et notamment pour des échantillons de données dont la taille est relativement limitée et, ou au sein desquels la variation interindividuelle est élevée. SORENSEN et al. [2016] donnent une vision plus large et bien plus détaillée des avantages du raisonnement bayésien dans l'analyse de données pour les sciences cognitives, y compris la linguistique.

Les modèles bayésiens calculés sur les données renvoient une distribution postérieure pour chacun des paramètres ajoutés au modèle, avec ou sans interaction entre ceux-ci, pour les effets principaux comme pour les effets aléatoires. Dans le corps du texte ne seront reportés que les effets principaux et en annexe une description complète du modèle sera fournie, avec :

- la formule utilisée pour le modèle (telle que fournie au package *brms* [BÜRKNER, 2018; BÜRKNER et al., 2017; CARPENTER et al., 2017]), la famille du modèle (ordinaire, catégorielle ou

binomiale, en fonction de la variable dépendante), le nombre d'observations et le nombre de chaînes et d'itérations utilisées pour le calcul (4 chaînes sauf mention contraire, et 3000 ou 6000 itérations par chaîne, dont la moitié pour l'entraînement du modèle; le nombre requis varie selon les données pour que le modèle converge). Pour une meilleure lecture des résultats [BREHM et ALDAY, 2020], les variables indépendantes sont centrées autour de la moyenne (*mean-centered coding*).

- un tableau récapitulant : les coefficients estimés pour le modèle ($\hat{\beta}$), l'erreur estimée, un intervalle de crédibilité (95%) donnant les valeurs entre lesquelles il y a 95% de chances de trouver la valeur réelle du coefficient (voir MOREY et al. [2016] pour une description des avantages des intervalles de crédibilité sur les intervalles de confiance), et enfin des informations sur la bonne convergence du modèle pour chaque paramètre (Bulk_ESS, Tail_ESS et Rhat; pour les données reportées ici, le critère de convergence est que ce dernier soit toujours égal à 1).
- un second tableau présentant les distributions postérieures issues du modèle avec : le coefficient estimé ($\hat{\beta}$), la probabilité P que ce paramètre (β) soit supérieur ou inférieur à 0 pour l'ensemble des valeurs calculées par le modèle, et divers intervalles de crédibilité (successivement, 95%, 90% et 85%, centrés autour du coefficient estimé $\hat{\beta}$).
- un graphique (dans le corps du texte quand c'est approprié, en annexe sinon) permettant une visualisation du second tableau et présentant ces distributions postérieures autour de 0, avec en bleu les valeurs comprises dans l'intervalle de crédibilité (95%) et en rose les traînes (*tails*) (voir par exemple la Figure 3.4, reprise de la Section 4.1.3.2, Figure 4.2).

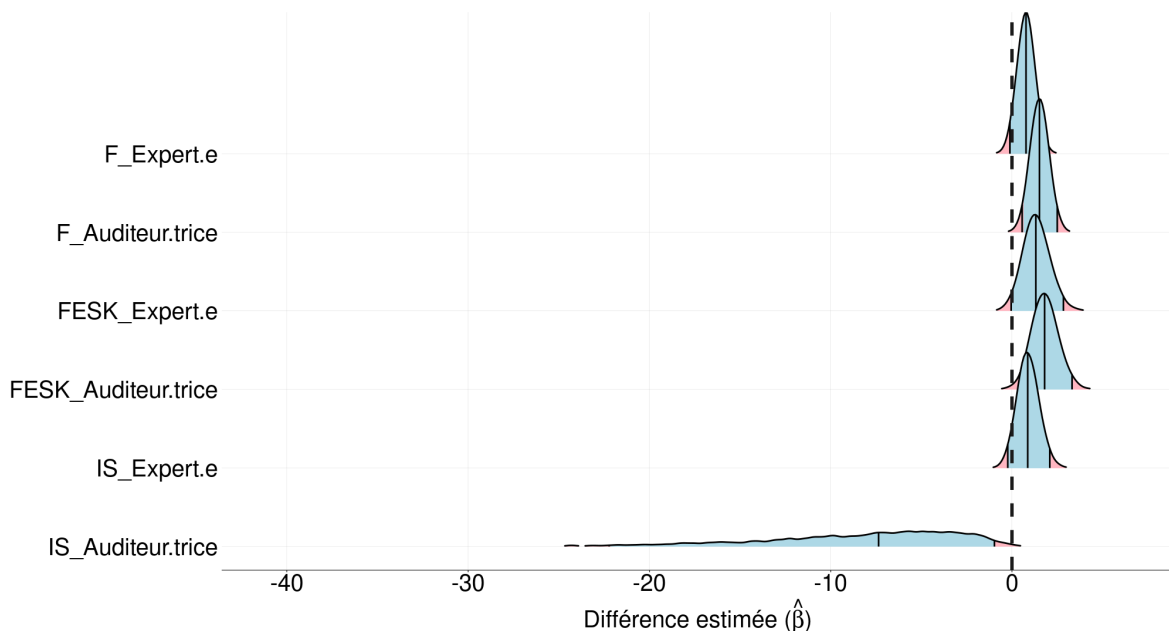


FIGURE 3.4 – Exemple de graphique illustrant les distributions postérieures d'un modèle brms

Pour les paramètres (β), une probabilité différente de 0 illustre l'effet, négatif ou positif, du paramètre (variable indépendante) introduit sur la variable dépendante. Le rôle d'un facteur dans le modèle statistique dépend ici de la combinaison entre la probabilité que β soit supérieur ou inférieur à 0, l'inclusion ou non du 0 dans l'intervalle de crédibilité retenu, mais aussi la largeur

globale de la distribution postérieure. Contrairement aux analyses fréquentistes et au consensus sur la valeur significative d'une *p.value* inférieure à 0.05, il n'existe pas actuellement de consensus sur quoi rapporter pour les inférences bayésiennes [ENGELMANN et al., 2019].

Un des avantages de l'analyse bayésienne est justement qu'il reste permis d'apprécier la valeur des effets en fonction des divers paramètres et de ne plus raisonner de manière binaire (en rejetant par défaut ce qui n'est pas significatif sur la base d'une valeur *alpha* seuil, comme une *p.value* > 0.05). Contrairement aux modèles fréquentistes, les modèles *maximaux* [BARR et al., 2013] bayésiens, calculés en intégrant une plus grande variabilité des données, présentent par ailleurs moins de risques de « dégénérer » (*convergence failure*).

Ce type d'analyse permet donc de tenir compte des effectifs relativement réduits des expériences linguistiques (généralement autour de 40 volontaires) et de la variabilité des données (dues à la nature exploratoire de cette recherche). En conséquence, ici, les effets qui seront rapportés seront ceux pour lesquels la probabilité $P(\beta)$ d'être différent de 0 est supérieure à 0.7. Ces valeurs pourraient servir de référence pour de futures expériences, au titre de priors pour les modèles. Il est nécessaire pour lire ces valeurs de garder en tête que :

- en dessous de 0.7 cette valeur indique que l'existence d'un effet n'est pas corroborée par les données ;
- de 0.7 à 0.8 cette valeur est indicative d'une tendance générale, qui peut donner une indication sur l'existence d'un effet mais incite à la prudence (chiffres inspirés de l'étude de VASISHTH et al. [2013] sur une méta-analyse du traitement des phrases relatives avec extraction du sujet et de l'objet en chinois mandarin) ;
- de 0.8 à 0.9 cette valeur indique que les données supportent modérément l'existence d'un effet ;
- au-delà de 0.9 cette valeur indique que les données supportent fortement l'existence d'un effet.

Chapitre 4

Nouvelles analyses de corpus

Sommaire

4.1 Un corpus adulte en contexte spécifique : EPAC	53
4.1.1 Présentation	53
4.1.2 Annotation	54
4.1.3 Observations : Résultats généraux	57
4.1.3.1 Un effet de thématique?	57
4.1.3.2 Un effet de personne?	59
4.1.4 Observations : Des stratégies interactionnelles	61
4.1.4.1 Des stratégies liées au positionnement des personnes intervenant	62
4.1.4.2 Des stratégies liées à la construction de l'interaction	64
4.1.4.3 Des stratégies liées à l'argumentation	66
4.1.5 Bilan provisoire	68
4.2 Un corpus adulte diachronique avec des contextes variés : ESLO	68
4.2.1 Présentation et préparation pour l'analyse syntaxique	68
4.2.2 Annotation syntaxique	69
4.2.3 Observations : Résultats généraux	70
4.2.4 Observations : Une évolution diachronique	71
4.2.5 Observations : Des effets générationnels?	72
4.2.6 Confrontation des facteurs d'époque et d'âge	73
4.2.7 Observations : Effets de contexte et stratégies interactionnelles	75
4.2.8 Confrontation des facteurs d'époque, d'âge et de contexte	78
4.2.8.1 Confrontation des facteurs d'époque et d'âge (interviews)	78
4.2.8.2 Confrontation des facteurs d'âge et de contexte, ESLO 2	79
4.2.9 Bilan provisoire	81
4.3 Conclusion intermédiaire	82
4.4 Apports majeurs	83

Ce chapitre présente les résultats d’exploration de plusieurs corpus francophones oraux déjà existants, dérivant des projets de recherche EPAC [BAZILLON et al., 2011; ESTÈVE et al., 2010] et ESLO [ESHKOL-TARAVELLA et al., 2011; LLL, 2017]. Ces corpus ont été retenus pour analyse, au détriment par exemple des autres sources potentielles de données mentionnées au Chapitre 3, bien sûr pour leur disponibilité et leur taille, mais aussi pour des raisons plus liées à l’analyse menée.

Le projet EPAC permet ainsi une analyse en profondeur de l’usage de l’interrogation partielle dans un contexte interactionnel très contrôlé, celui d’interviews de radio par un journaliste et les personnes invitées dans son émission. Le projet ESLO permet pour sa part, par un procédé technique qui sera décrit, d’injecter sur chaque énoncé les métadonnées sociolinguistiques des personnes les produisant (âge en l’occurrence mais également niveau d’études, catégorie socioprofessionnelle, etc.), pour conduire ensuite une analyse fine sur ce plan.

4.1 Un corpus adulte en contexte spécifique : EPAC

Le premier corpus étudié est un sous-corpus du projet EPAC [BAZILLON et al., 2011; ESTÈVE et al., 2010]. Il présente l’avantage de couvrir un contexte interactionnel relativement constant (émissions publiques de radio), et d’avoir été annoté spécifiquement pour relever les phrases interrogatives. En ceci, ce sous-corpus était une porte d’entrée appropriée pour une réactualisation des données disponibles sur les interrogatives partielles du français et leur usage, en particulier ici dans leur dimension pragmatique.

4.1.1 Présentation

Le sous-corpus EPAC analysé ici consiste en les transcriptions écrites d’environ 30 heures d’émissions de radio diffusées en 2004 sur la station publique France Inter, sous le nom *Le téléphone sonne*. Ces émissions sont majoritairement des débats, avec un journaliste animant les séances et des experts et expertes dont le rôle est de commenter et de discuter le sujet de la séance (de la politique intérieure à des sujets culturels, en passant par l’actualité internationale, l’économie ou bien d’autre). Ces discussions impliquent aussi des réponses à des questions posées par des membres de l’audience, qui soit les posent en direct à l’antenne, soit les ont envoyées au journaliste par Internet. Les dates de diffusion des 30 émissions dont les transcriptions ont été utilisées sont résumées dans le Tableau 4.1.

Mois	Jour de diffusion
Mars	19, 22, 25, 26, 29, 30
Avril	2
Août	4, 5, 6, 9, 12, 13, 16, 17, 18, 19, 29, 23, 25, 26, 30, 31
Septembre	2, 6, 7, 8, 9, 10, 17

TABLEAU 4.1 – Émissions dont les transcriptions constituent le sous-corpus EPAC analysé

Ce corpus donne un aperçu de production spontanée d’interrogatives partielles entre adultes. Cette production est parfois préparée ou semi-préparée, pour les productions d’interrogatives de l’animateur, recevant les questions de l’auditoire puis les reformulant à l’antenne. Parfois néanmoins les usages sont plus spontanées (personnalités expertes invitées, débat houleux, etc.) et ces situations permettent une analyse de la variation en français et de ses usages en interaction.

4.1.2 Annotation

Comme il a été mentionné, les interrogatives du sous-corpus étant déjà annotées, leur extraction en a été facilitée. Cinq catégories d'annotation relevées dans BAZILLON et al. [2011] ont été utilisées : *Question directe partielle complexe*, *Question directe partielle déterminant*, *Question directe partielle pronom*, *Question directe partielle adverbe* et *Question directe rhétorique*. Les quatre premières catégories étaient d'un intérêt explicite, mais une rapide revue des interrogatives dites rhétoriques a aussi fait apparaître que de nombreuses interrogatives partielles étaient recensées sous ce label.

Les phrases annotées sous ces cinq labels ont été automatiquement extraites, avec l'entière du tour de parole correspondant ainsi que le tour précédent et le tour suivant. L'idée était de pouvoir resituer la production d'interrogative dans le contexte interactionnel et argumentationnel. De toutes les phrases ainsi extraites, 495 ont été conservées après évacuation notamment des interrogatives rhétoriques non-partielles (totales). Ces 495 formes d'interrogation partielle ont ensuite été annotées sur deux catégories de critères bien distincts : des critères d'abord syntaxiques et des critères plus pragmatiques ou interactionnels.

Les critères syntaxiques sont décrits ci-dessous. Pour chaque critère, un exemple d'extraction du corpus est fourni où le critère était rempli (le plus souvent, valeur = 1). Ces exemples ont pour label un numéro de fichier (correspondant à une date d'émission) et un numéro d'item (précédé du symbole #), permettant leur identification unique parmi les 495 extractions.

- *Phrase* : l'interrogatif est-il seul ou dans une phrase non verbale (76), dans une phrase verbale mais infinitive (77), dans une phrase finie mais enchâssée (78), ou dans une phrase finie racine (79).

(76) J'ai dû l'aider à payer son loyer () Puis-je mentionner tout ou partie de ces aides sur ma propre déclaration? **Sous quelles formes?**

(EPAC 20040326 - #37)

(77) **que** penser également du front national implanté dans la plupart (...)

(EPAC 20040322 - #20)

(78) mais il reste à savoir sur le fond **dans quelle mesure** le rythme et surtout la teneur et l'orientation des réformes sera infléchi ou renforcé

(EPAC 20040330 - #62)

(79) **jusqu'à quand** l'administration peut-elle procéder à un redressement

(EPAC 20040326 - #47)

- *Sujet* : l'interrogatif est-il le sujet syntaxique de la phrase racine, oui (1) ou non (0) : (80).

(80) **qui** veut répondre?

(EPAC 20040805 - #100)

- *Antéposition* : l'élément interrogatif est-il au début du syntagme verbal dont il est un argument ou un adjectif (1), ou non (0) : (81).

(81) (...) n'ont pas arrêté pendant toute la campagne de dénoncer l'absence de projet de l'un et de l'autre : **comment** peuvent-ils se retrouver entre les deux tours et surtout autour de quel projet concret pour l'aquitaine

(EPAC 20040322 - #18)

- *Inversion* : y a-t-il une inversion verbe-sujet (1) ou non (0) dans le syntagme verbal dont l'élément interrogatif est un argument ou un adjectif : (82). La différence n'a pas été faite ici entre sujet plein et sujet clitique.

(82) que peut-on dire à ceux par exemple dont le père (...)?

(EPAC 20040325 - #25)

- *ESK* : la locution interrogative *est-ce que* relie-t-elle l'élément interrogatif au reste de la phrase interrogative (1) ou non (0) : (83).

(83) (...) ou **qu'**est-ce que j'en fais exactement?

(EPAC 20040326 - #51)

- *Négation* : le verbe dominant directement l'élément interrogatif sur le plan syntaxique a-t-il une négation explicite (1) ou non (0) : (84). La négation double (*ne... V... pas*) et la négation simple (\emptyset ... *V... pas*) ont ici été codées de façon similaire.

(84) **pourquoi** n'essaie-t-on pas au moins de le faire pour les salariés?

(EPAC 20040326 - #44)

Les critères plus relatifs à l'interaction ont été conçus de façon binaire, pour être annotés le moins subjectivement possible. Ils comprenaient notamment :

- *Personne* : quelle personne a produit l'interrogative (journaliste, personnalité experte, auditeur ou auditrice).
- *Informalisation* : l'interrogative se trouve-t-elle dans un contexte linguistique informel (marque d'oralité, négation sans *ne*, vocabulaire familier), oui (1) ou non (0) : (85).

(85) **qu'**est-ce qu'il va devenir en collège ce gamin?

(EPAC 20040902 - #374)

- *Proximité destinataire* : l'interrogative ou son contexte immédiat contiennent-ils un élément indiquant que la personne qui la produit cherche à réduire la distance relationnelle avec son ou sa destinataire, ou avec l'audience (pour des destinataires explicites, il s'agit par exemple de leur nom), oui (1) ou non (0) : (86).

(86) **qu'**en pense notre confrère ivan xxxx?

(EPAC 20040402 - #83)

- *Posture* : l'interrogative permet-elle à la personne qui la produit de se positionner dans une posture sociale dogmatique ou experte dans son domaine, qui aurait par exemple la solution immédiate à un problème sociétal complexe, oui (1) ou non (0) : (87).

(87) je vous rappelle quand même que la libération du fils de () c'est la libération obtenue dans les délais les plus rapides grâce à **qui**? Eh bah grâce à cette collaboration que nous avons avec (...)

(EPAC 20040809 - #131)

- *Opinion implicite* : l'interrogative convoie-t-elle, en même temps que l'acte langage interrogatif, une idée de ce que devrait être la réponse pour la personne qui l'a posée, oui (1) ou non (0) : (88).

(88) bon patrimoine euh euh très bien mais **pourquoi** ces objets sont-ils pas tout le temps visibles?

(EPAC 20040917 - #483)

— *Réponse auto-fournie* : une réponse à l’interrogative est-elle fournie immédiatement après par la personne qui l’a posée, oui (1) ou non (0) : (89).

(89) si demain ce cinéma disparaît (...) **qu**’est-ce qu’on fait? On prend la voiture
on va à paris on va à lyon

(EPAC 20040819 - #245)

— *Fin d’argumentaire* : l’interrogative est-elle posée après une longue tirade de la personne qui la pose, sur le sujet qu’elle concerne, oui (1) ou non (0) : (90).

(90) vous imaginez pour les directeurs des ressources humaines pour les pdg
quand euh les professions les plus modestes ont ont souffert euh **comment**
expliquer ce décalage alors?

(EPAC 20040824 - #294)

— *Animation* : l’interrogative est-elle pour la personne qui la pose un moyen de stimuler la réflexion, le débat ou de demander la poursuite d’une argumentation, oui (1) ou non (0) : (91). Ce critère concerne notamment les interrogatives posées par le journaliste animant l’émission.

(91) beaucoup de questions comme ça... **qui** veut là-dessus commencer euh
oui euh (...)

(EPAC 20040319 - #13)

— *Reformulation* : l’interrogative a-t-elle pour but de reprendre et/ou modifier un propos déjà tenu, oui (1) ou non (0) : (92).

(92) c’est pas franchement la question de xxxx elle dit quand même **comment**
dire non à à la constitution sans dire non à l’europe?

(EPAC 20040825 - #311)

— *Auto-Répétition* : l’interrogative ou une partie de l’interrogative est-elle une répétition d’un propos tenu par la personne qui la produit, oui (1) ou non (0) : (93).

(93) **que** peut-on dire à ceux par exemple dont le père a été victime de de ce
type de cancer un père ou un oncle euh **que** peut-on leur dire?

(EPAC 20040325 - #25)

— *Répétition autre* : l’interrogative ou une partie de l’interrogative est-elle une répétition d’un propos tenu précédemment par une autre personne que celle qui la produit, oui (1) ou non (0) : (94).

(94) on parle beaucoup de YYYY on a raison () mais ce que je constate on n’a
presque pas parlé du programme de YYYY **qu**’est-ce qu’il cherche YYYY? ()

(EPAC 20040813 - #156)

Certains de ces critères étant assez subjectifs (*posture*, par exemple), il était important de s’assurer de la viabilité des annotations et une seconde personne a ainsi annoté un sous-corpus d’une centaine de questions puis ses annotations ont été comparées aux premières. Le taux global d’accord sur les critères interactionnels (hors *Personne*, qui n’est pas subjectif) était de 79.5%. Il a été considéré que cette proportion de correspondance entre les annotations indiquait une fiabilité raisonnable des premières.

Pour les raisons évoquées dans le Chapitre 2 (Sections 2.1.1 et 2.2.1), l'analyse des interrogatives sujet et subordonnées n'est pas l'objectif du présent travail. Sur les 495 extractions, 211 interrogatives partielles racines finies et non-sujet ont donc été conservées pour analyse. Elles sont présentées en Annexe A. L'intégralité des extractions est ou sera disponible dans le dépôt OSF (osf.io/ud4ks).

4.1.3 Observations : Résultats généraux

Le Tableau 4.2 présente la répartition des 211 phrases isolées selon la procédure décrite plus haut, selon leur type. Il s'y voit une répartition inégale des différentes variantes, et une prépondérance des interrogatives avec antéposition et inversion.

Antéposition simple	Antéposition + ESK	Antéposition et inversion	In situ
45	24	124	18
21.3 %	11.4%	58.8 %	8.5 %

TABLEAU 4.2 – Interrogatives partielles dans EPAC par type

Ces proportions rappellent certains des chiffres rapportés par COVENEY [2011] et tirés de TERRY [1970] (corpus de THÉÂTRE) et BEHNSTEDT [1973] (corpus de RADIO), pour lesquels l'antéposition avec inversion était majoritaire. Ces proportions sont par contre en contradiction avec, d'une part les autres données des années 1970-80 rapportées par COVENEY [2011] mais aussi avec les données plus récentes de HAMLAOUI [2009a], QUILLARD [2001] ou ADLI [2015], où les formes in situ représentent 40 à 50 % des données (Section 2.2 pour plus de détails). Ceci illustre que, à époque égale, le contexte d'interaction joue sans doute un rôle sur la proportion d'utilisation des différentes variantes disponibles. En l'occurrence, un contexte d'émissions radiophoniques de débat entre personnalités publiques, dans les années 1970 comme dans les années 2000, semble propice à l'utilisation massive d'antépositions avec inversion.

4.1.3.1 Un effet de thématique?

Une analyse des différents thèmes du sous-corpus EPAC permet de creuser un peu plus cette dimension de contexte. Le Tableau 4.3 et la Figure 4.1 montrent la distribution des quatre variantes interrogatives principales en fonction du sujet de l'émission. Les chiffres présentés concernent les thématiques pour lesquelles plus de 10 interrogatives ont été produites dans le sous-corpus ou qui ont été abordées à travers plusieurs émissions différentes (N=185/211, une version complète pour les 211 interrogatives est présentée en Annexe B.0.1).

	Antéposition simple	Antéposition + ESK	Antéposition et inversion	In situ
Société	6	1	8	1
Santé	6	3	10	3
Économie	4	4	16	2
Culture	7	2	22	5
International	9	3	24	3
Politique	10	5	27	4

TABLEAU 4.3 – Interrogatives partielles dans EPAC par type et par thématique

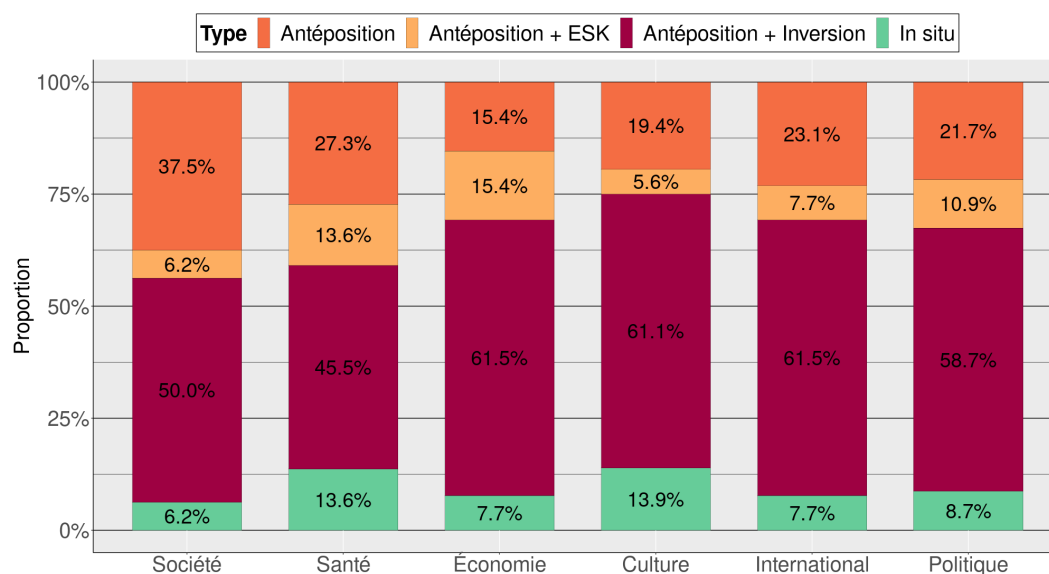


FIGURE 4.1 – Interrogatives partielles dans EPAC par type et par thématique

Un modèle statistique de régression bayésienne a été calculé pour les 185 interrogatives recensées pour ces six thématiques. Le modèle lancé avait pour variable dépendante le TYPE de l'interrogative (antéposition simple, en est-ce que, avec inversion et in situ, avec pour niveau de référence l'antéposition avec inversion) et pour variable indépendante la THÉMATIQUE (société, santé, économie, culture, international et politique, avec pour niveau de référence la politique pour des raisons de nombres d'occurrences). Un effet avec intercept aléatoire a été ajouté pour tenir compte de la variabilité entre les émissions ("fichier"), ce qui sera le cas pour tous les modèles présentés par la suite. La variable dépendante étant catégorielle, c'est la famille *categorical* qui a été sélectionnée pour faire tourner ce modèle, avec 4 chaînes et 6000 itérations par chaîne. La convergence du modèle a été atteinte et vérifiée en contrôlant que chaque Rhat soit à 1.00 pour tous les paramètres. Le détail de ce modèle est présenté en Annexe B.0.2.

Les analyses confirment quelques effets légers de thématique, directement visibles en Figure 4.1. Ainsi, la proportion d'antépositions simples par rapport aux antépositions avec inversion augmente, en comparant les émissions avec un thème sociétal ou sanitaire à celles avec un thème plus directement politique (respectivement, $\hat{\beta} = 0.73$, 95%CrI=[-1.47,2.99], $P(\beta)>0=0.76$ et $\hat{\beta} = 0.71$, 95%CrI=[-1.34,2.93], $P(\beta)>0=0.76$). La proportion d'antéposition en *est-ce que* par rapport à celles avec inversion augmente aussi en comparant les émissions avec un thème sanitaire ($\hat{\beta} = 1.48$, 95%CrI=[-2.85,6.72], $P(\beta)>0=0.75$) ou encore un thème économique ($\hat{\beta} = 1.36$, 95%CrI=[-2.56,6.30], $P(\beta)>0=0.75$) avec les émissions politiques. Enfin, la proportion de phrases in situ augmente par rapport aux antépositions avec inversion augmente également quand la thématique est sanitaire ($\hat{\beta} = 0.63$, 95%CrI=[-1.62,2.80], $P(\beta)>0=0.73$), par comparaison avec la thématique politique.

L'étalement des distributions postérieures et le nombre réduit d'occurrences d'interrogatives partielles par thématique incite néanmoins à prendre ces tendances pour ce qu'elles sont et à ne pas en tirer de conclusions autres qu'indicatives. La prépondérance des antépositions avec inversion par rapport aux autres variantes, si elle est de degré variable, reste présente quel que soit le thème abordé dans l'émission. Si une notion de contexte importe ici, c'est sans doute bien plus celle prise au sens large de « entretien entre personnalités publiques, diffusé par une radio publique vers une audience potentielle de milliers de personnes ».

4.1.3.2 Un effet de personne?

Les effets de la contrainte de contexte, si elle existe, peuvent différer entre individus. Une analyse des interrogatives relevées dans le sous-corpus EPAC, permet d’avoir un aperçu de ceci. Le Tableau 4.4 et la Figure 4.2 montrent la distribution des quatre variantes interrogatives principales en fonction du rôle des personnes qui les ont produites dans l’interaction. Les chiffres présentés concernent le journaliste animateur de l’émission (toujours la même personne), les personnalités expertes invitées dans les émissions, et les personnes reçues à l’antenne pour y poser leur question en direct. Les questions de journalistes qui parfois co-interviennent dans l’émission (N=3/211) ont été retirées des analyses correspondantes.

	Antéposition simple	Antéposition + ESK	Antéposition et inversion	In situ
Auditeur ou Auditrice	17	9	21	0
Personnalité experte	16	9	40	10
Journaliste	11	5	62	8

TABLEAU 4.4 – Interrogatives partielles dans EPAC par type et rôle interactionnel

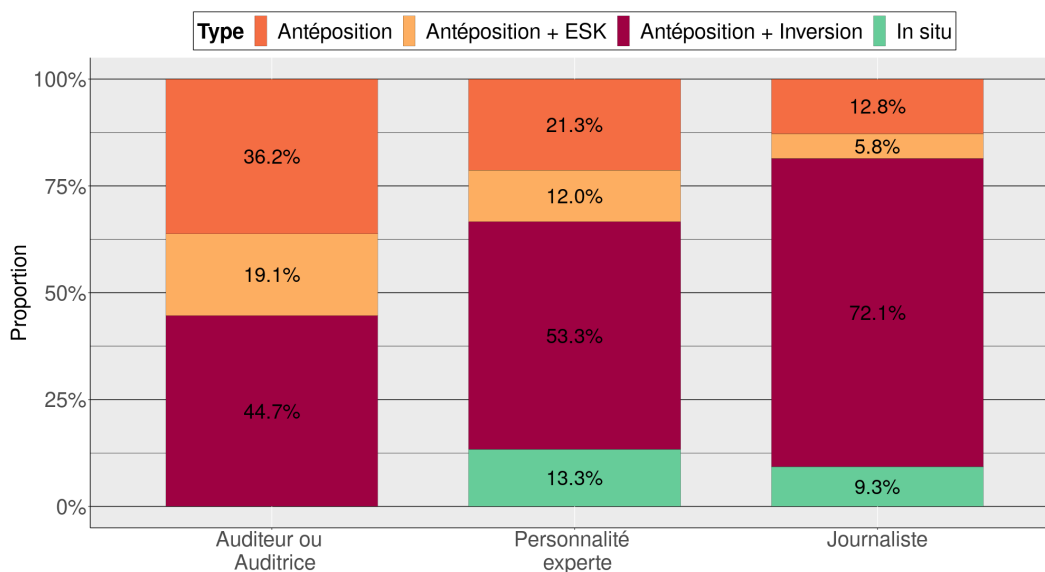


FIGURE 4.2 – Interrogatives partielles dans EPAC par type et rôle interactionnel

Un modèle statistique de régression bayésienne a été calculé pour les 208 interrogatives recensées pour ces trois rôles. Le modèle lancé avait pour variable dépendante le TYPE de l’interrogative (antéposition simple, en est-ce que, avec inversion et in situ, avec pour niveau de référence l’antéposition avec inversion) et pour variable indépendante le RÔLE (journaliste, personnalité experte, auditeur ou auditrice, avec pour niveau de référence le journaliste). Le même effet avec intercept aléatoire a été ajouté ("fichier"). Ce modèle a été calculé avec 4 chaînes et 3000 itérations par chaîne. La convergence du modèle a été atteinte et vérifiée. Le détail de ce modèle est présenté en Annexe B.0.3 et la Figure 4.3 en donne les distributions postérieures.

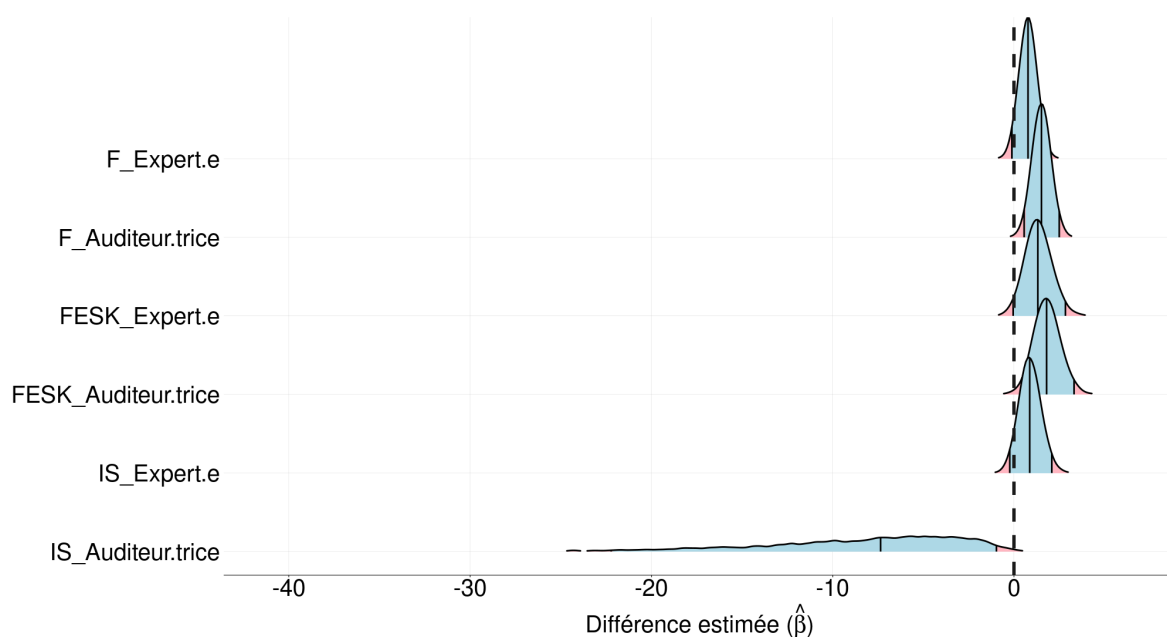


FIGURE 4.3 – Distributions postérieures pour le modèle brms d'analyse de EPAC, effet de RÔLE

Le modèle statistique confirme les grandes différences visibles en Figure 4.2. Avec pour point de repère les usages du journaliste, les personnalités expertes utilisent une proportion plus importante d'antépositions simples ($\hat{\beta} = 0.78$, 95%CrI=[-0.11,1.69], $P(\beta)>0=0.96$), d'antépositions en *est-ce que* ($\hat{\beta} = 1.35$, 95%CrI=[-0.04,2.84], $P(\beta)>0=0.97$) et de phrases in situ ($\hat{\beta} = 0.89$, 95%CrI=[-0.23,2.09], $P(\beta)>0=0.94$), comparativement à leur utilisation des antépositions avec inversion. Pour leur part et toujours en prenant comme référence le journaliste, les personnes qui posent leurs questions à l'antenne en direct utilisent également une proportion plus importante d'antépositions simples ($\hat{\beta} = 1.52$, 95%CrI=[0.56,2.50], $P(\beta)>0=1$) et d'antépositions en *est-ce que* ($\hat{\beta} = 1.82$, 95%CrI=[0.46,3.32], $P(\beta)>0=0.99$) que d'antépositions avec inversion, mais par contre une proportion moins importante de phrases in situ ($\hat{\beta} = -8.56$, 95%CrI=[-22.22,-0.97], $P(\beta)<0=1$). Ce dernier résultat est le reflet du fait qu'il n'y a, dans les données disponibles, aucune phrase in situ pour le groupe des auditeurs et auditrices.

Les résultats importants ici sont multiples. D'une part, la diminution progressive du recours à l'antéposition avec inversion (journaliste > personnalité experte > audience) pourrait être reliée à une moindre implication directe dans la bonne marche de l'émission et à un moindre besoin de s'exprimer en français parfaitement académique ou littéraire. Ceci vaut aussi pour la proportion supérieure de phrases in situ chez les personnalités expertes, par rapport au journaliste. L'absence totale ou quasi-totale de phrases in situ chez les personnes de l'audience qui appellent la radio vient néanmoins en conflit avec cette idée.

Il peut s'envisager que pèsent sur ces personnes un plus fort poids normatif, les incitant à ne pas parler un français trop peu soutenu lorsque virtuellement n'importe qui peut les entendre. Les personnalités expertes, plus rôdées à ce type d'exercice, seraient plus libres. Une autre hypothèse envisageable est que les personnes appelant une chaîne de radio ont préparé leur question avant de la poser, à l'inverse des productions de personnalités expertes, plus spontanées et liées au débat en cours et pour lesquelles elles ont été invitées. Enfin, il faut noter que chez les personnes de l'audience, cette absence de productions in situ est en quelque sorte compensée par une proportion bien supérieure d'antépositions simples et d'antépositions en *est-ce que*.

4.1.4 Observations : Des stratégies interactionnelles

Un intérêt majeur du corpus EPAC est la stabilité de son contexte interactionnel, pris au sens large d'entretien radiodiffusé, malgré une certaine diversité dans les personnes enregistrées (journaliste, personnalités expertes, personnes de l'audience). Un autre atout est l'exercice même qui est au coeur de ces interactions : des entretiens sur des sujets contemporains, avec des points de vue parfois contradictoires quand les réflexions tournent au débat de société entre personnalités expertes. Le tout permet, grâce aux dix critères plus interactionnels d'annotation présentés en Section 4.1.2 (hors le critère de *Personne*), d'explorer l'usage des différentes variantes interrogatives à un niveau très concret, et de relier telle variante à tel moment interactionnel ou à telle visée stratégique au cours de l'interaction.

Les Sections 4.1.4.1-4.1.4.3 distinguent trois groupes de critères : ceux plutôt liés à des stratégies de positionnement des personnes intervenant dans l'interaction vis-à-vis des autres (Informalisation, Proximité Destinataire, Prise de Posture), ceux plutôt liés à la construction de l'interaction (Animation, Reformulation, Auto-Répétition, Répétition autre) et enfin des critères plus directement liés au développement d'une argumentation par les personnes qui les produisent (Fin d'argumentaire, Réponse Auto-fournie, Opinion implicite). Dans la suite de la présentation de ces données, les extractions correspondant aux critères considérés comme remplis sont associées aux réponses « Oui » (« 1 » lors de l'annotation) et les autres aux réponses « Non » (« 0 » pour l'annotation).

Un modèle statistique de régression bayésienne a été calculé pour chacun des dix critères présenté, pour les 211 interrogatives du sous-corpus, avec pour variable dépendante le TYPE de l'interrogative (antéposition simple, en est-ce que, avec inversion et in situ, avec pour niveau de référence l'antéposition avec inversion) et pour variable indépendante chacun des CRITÈRES (et "fichier" en effet avec intercept aléatoire). Ces modèles ont été calculés avec 4 chaînes et 3000 itérations par chaîne. La convergence des modèles a été atteinte et vérifiée. Le détail de ces modèles est présenté en Annexe B.0.5 et les résultats principaux en sont rapportés ci-après en parallèle des données descriptives (chiffres bruts et proportions de variantes interrogatives pour chacun des critères).

Avant cela, il importe d'observer les relations entre tous ces critères. Une matrice des corrélations entre critères (méthode = Spearman, niveau de significativité = 0.05) est présentée en Figure 4.4. Le critère *animation* n'y a pas été reporté car il est très spécifique au journaliste animateur de l'émission, ce qui le distingue des autres critères. Pour rappel (Chapitre 3), la couleur des cercles sur la partie supérieure droite du graphique indique si la corrélation est positive ou négative, et la taille de ces cercles est proportionnelle au coefficient de corrélation. Les corrélations non significatives sont indiquées par une croix. Le coefficient de corrélation est la valeur chiffrée indiquée en vis-à-vis des cercles sur la partie inférieure gauche.

La matrice indique et illustre que, si un certain nombre de corrélations sont statistiquement significatives, leurs coefficients en sont pour la plupart assez peu élevés. Le lien entre *Prise de posture* et *Réponse-autofournie* ($r = .44$, $p < 0.01$) traduit une certaine corrélation, positive, entre les stratégies linguistiques mises en oeuvre quand chacun de ces critères est rempli.

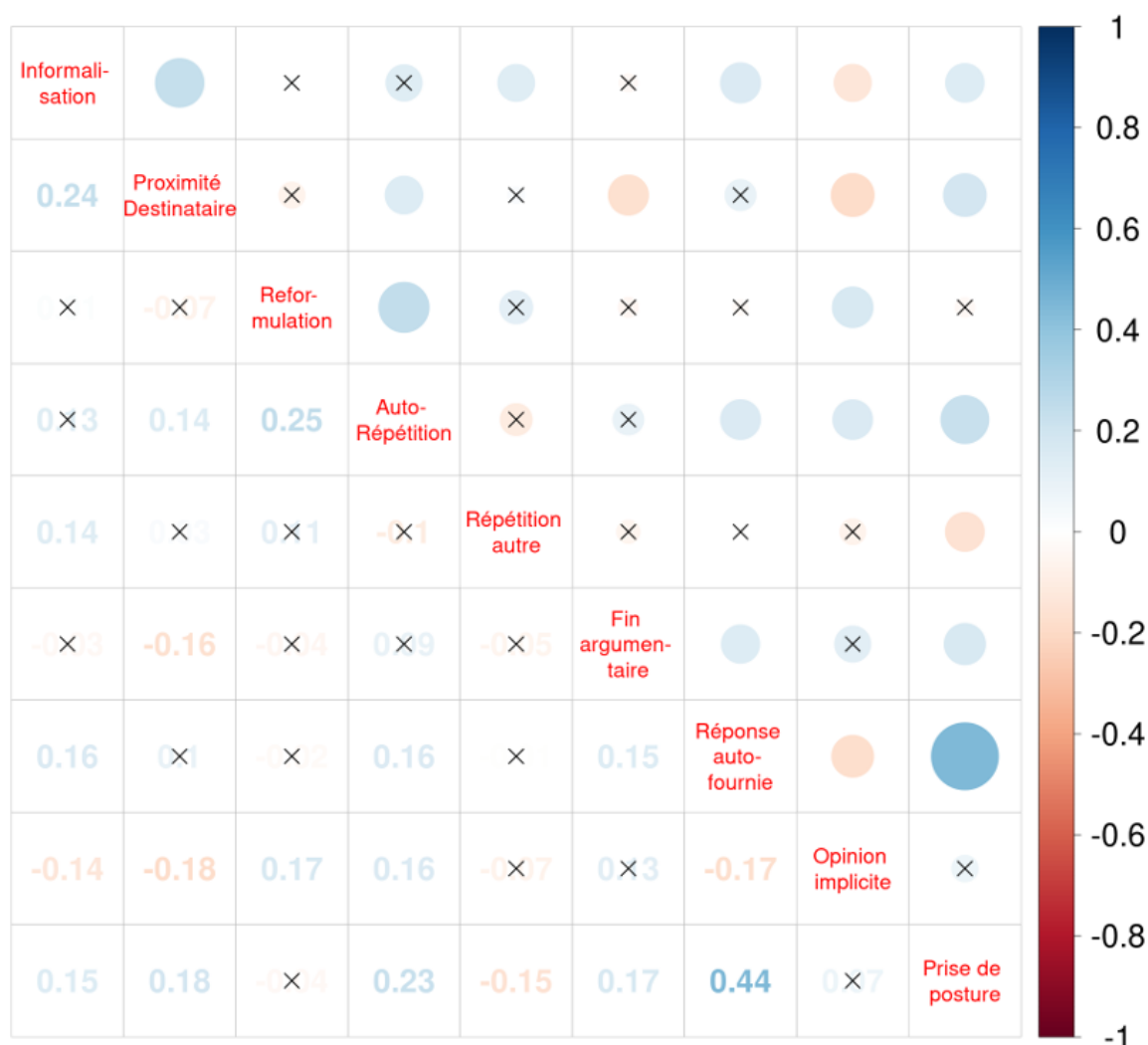


FIGURE 4.4 – Corrélations entre critères interactionnels (EPAC)

Plus de détails statistiques sur les corrélations sont disponibles dans l'Annexe B.0.4.

4.1.4.1 Des stratégies liées au positionnement des personnes intervenant

Le premier bloc de critères interactionnels analysés comprend ceux relatifs aux liens entre les personnes impliquées dans la situation de communication (*informalisation*, *proximité destinataire*, *prise de posture*). Le Tableau 4.5 et la Figure 4.5 donnent la distribution des variantes interrogatives dans le sous-corpus selon que ces critères sont remplis ou non.

	Informalisation		Proximité Destinataire		Prise de Posture	
	Oui	Non	Oui	Non	Oui	Non
Antéposition simple	30	15	7	38	15	30
Antéposition + ESK	16	8	3	21	6	18
Antéposition et inversion	14	110	7	117	39	85
In situ	18	-	10	8	10	8
TOTAL	78	133	27	184	70	141

TABEAU 4.5 – Interrogatives partielles par type et par critère de "positionnement" (EPAC)

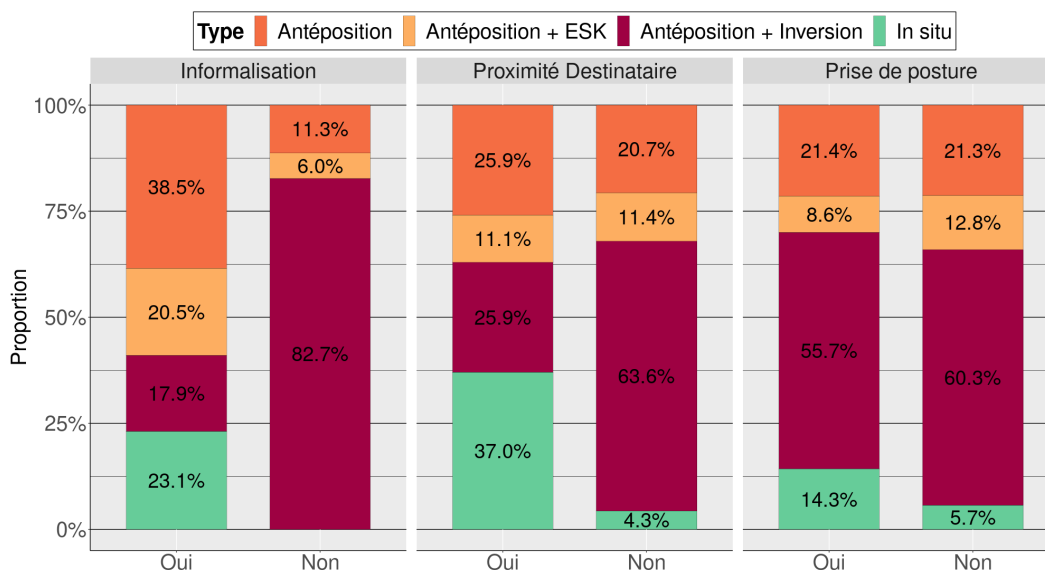


FIGURE 4.5 – Interrogatives partielles dans EPAC par type et critère de "positionnement" (EPAC)

Le critère *informalisation* et celui de *proximité destinataire*, bien que peu corrélés, sont similaires en ce qu'ils correspondent à des situations où la personne qui produit une interrogative est dans une démarche de réduction de la distance entre elle-même et le ou la destinataire de sa question. Ceci se traduit par une variation similaire pour les deux critères de la proportion des trois autres types interrogatifs par rapport à celle des antépositions avec inversion, quand les critères sont remplis. La proportion d'antépositions simples par rapport à celle d'antépositions avec inversion augmente autant en cas d'*informalisation* ($\hat{\beta}= 2.92$, 95%CrI=[2.03,3.85], $P(\beta)>0=1$) que pour la *proximité destinataire* ($\hat{\beta}= 1.24$, 95%CrI=[0.01,2.47], $P(\beta)>0=0.98$), ce qui se retrouve pour les antépositions en *est-ce que* (respectivement, $\hat{\beta}= 2.97$, 95%CrI=[1.87,4.18], $P(\beta)>0=1$ et $\hat{\beta}= 0.86$, 95%CrI=[-0.88,2.52], $P(\beta)>0=0.86$) et pour les phrases in situ (respectivement, $\hat{\beta}= 10.49$, 95%CrI=[4.79,21.15], $P(\beta)>0=1$ et $\hat{\beta}= 3.22$, 95%CrI=[1.94,4.64], $P(\beta)>0=1$).

Les critères *informalisation* et *proximité destinataire* diffèrent toutefois en ce que le premier traduit un contexte linguistique dé-formalisé (par le lexique, la structure syntaxique, des marqueurs d'oralité), alors que le second rend compte des situations où une apostrophe explicite au ou à la destinataire est insérée dans la phrase interrogative elle-même. Cette différence se voit dans l'augmentation plus importante de la proportion d'antépositions simples et en *est-ce que* en cas d'*informalisation* qu'en cas d'adresse explicite au destinataire. Il faut par ailleurs constater que virtuellement aucune phrase in situ ne correspond aux contextes où aucune *informalisation* n'a été détectée pendant l'annotation. Enfin, il est bien sûr à relever que le faible nombre d'occurrences où le critère *proximité destinataire* est rempli incite à la prudence.

Ces deux premiers critères contrastent avec celui de *posture*, lequel traduit pour sa part une volonté de la personne produisant l'interrogative de se positionner dans un rôle ou une position supérieure de légitimité quant au sujet en discussion (qu'il s'agisse de légitimité à contredire ou à avancer un argument, par exemple). Il est notable ici que seule la proportion de phrases in situ semble augmenter par rapport à celle des antépositions avec inversion quand le critère est rempli ($\hat{\beta}= 1.32$, 95%CrI=[0.17,2.61], $P(\beta)>0=0.99$). La proportion des antépositions simples reste stable pour sa part, et si la proportion des antépositions en *est-ce que* évolue, c'est plutôt à la baisse ($\hat{\beta}= -0.39$, 95%CrI=[-1.66,0.80], $P(\beta)<0=0.73$).

L'antéposition avec inversion semble donc ici une variante interrogative relativement incompatible avec un marquage informel du propos ((95) pour une exception) ou un rapprochement avec le ou la destinataire dans un échange ((96) pour une exception, où le recours stylistique à la troisième personne nuance la stratégie de rapprochement). En revers, l'antéposition avec inversion semble tout à fait compatible avec une prise de posture dans l'échange (97).

(95) parce que finalement où est l'ascenc() où est l'ascenseur social dans cette histoire?

(EPAC 20040817 - #199)

(96) qu'en pense notre confrère ivan xxxx?

(EPAC 20040805 - #105)

(97) () et pourquoi les lit-on? Il y a un élément totalement irrationnel là-dedans ()

(EPAC 20040805 - #105)

L'autre fait à constater, en miroir, est la plus grande compatibilité des phrases in situ avec des contextes informels ((98), avec une dislocation) ou de rapprochement avec le ou la destinataire (99), mais aussi dans les cas de prise de posture (100).

(98) et la prévention ça repose sur quels principes essentiels qu'i(l) faudrait qu'on retienne ce soir (...)

(EPAC 20040805 - #105)

(99) vous nous appelez d'où bernard?

(EPAC 20040817 - #226)

(100) je vous rappelle quand même que la libération du fils de () c'est la libération obtenue dans les délais les plus rapides grâce à qui? Eh bah grâce à cette collaboration que nous avons avec (...)

(EPAC 20040809 - #131)

4.1.4.2 Des stratégies liées à la construction de l'interaction

Le second bloc de critères interactionnels comprend ceux relatifs à des processus de construction de l'interaction au sens de suite discursive (*reformulation, auto-répétition, répétition autre*). Les chiffres bruts et les proportions pour le critère d'*animation* sont aussi présentés ici, mais il est lié principalement au journaliste dont c'est le rôle d'animer les émissions. Le Tableau 4.6 et la Figure 4.6 donnent la distribution des variantes interrogatives dans le sous-corpus selon que ces critères sont remplis ou non.

	Animation		Reformulation		Auto-Répétition		Répétition Autre	
	Oui	Non	Oui	Non	Oui	Non	Oui	Non
Antéposition simple	12	33	9	36	8	37	7	38
Antéposition + ESK	6	18	10	14	13	11	2	22
Antéposition et inversion	54	70	25	99	23	101	9	115
In situ	9	9	3	15	8	10	1	17
TOTAL	81	130	47	164	52	159	19	192

TABLEAU 4.6 – Interrogatives partielles par type et par critère de "construction" (EPAC)

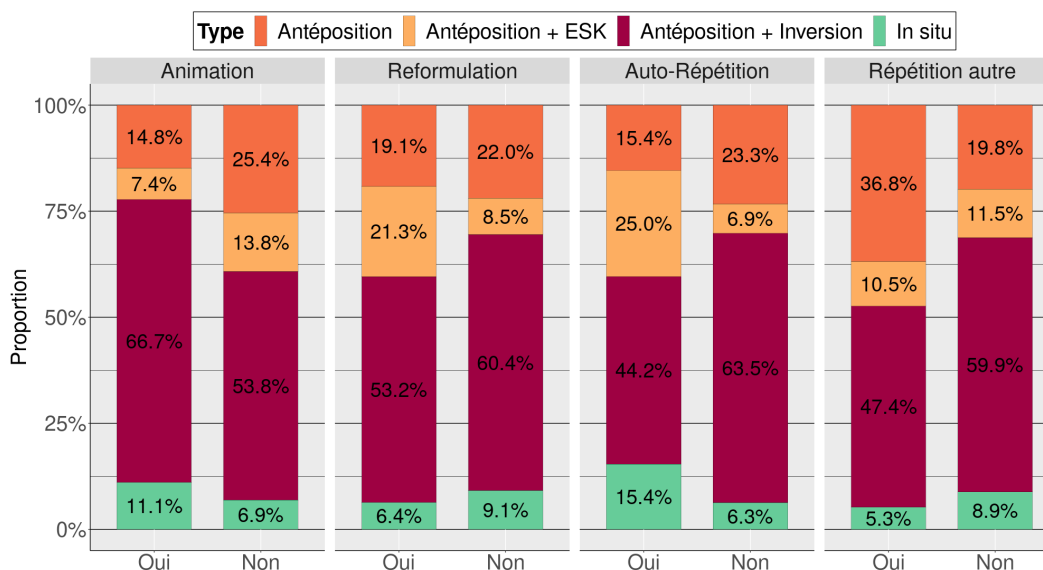


FIGURE 4.6 – Interrogatives partielles dans EPAC par type et critère de "construction" (EPAC)

La comparaison de ces données et des modèles statistiques associés permet de voir que les stratégies discursives semblent différentes selon le type d’apport à l’interaction qui est effectué. Par exemple, lorsqu’une interrogative est associée à un processus de répétition, la variante syntaxique choisie ne sera pas la même selon que cette répétition concerne un élément prononcé par-soi-même ou par quelqu’un d’autre. Si dans les deux cas l’antéposition avec inversion semble moins utilisée qu’en absence de répétition, dans le premier cas c’est une augmentation du recours aux phrases in situ ($\hat{\beta}= 1.60$, 95%CrI=[0.40,2.94], $P(\beta)>0=1$) ou aux antépositions en *est-ce que* ($\hat{\beta}= 2.11$, 95%CrI=[0.97,3.33], $P(\beta)>0=1$) qui semble compenser (101-102). Dans le second cas, c’est au contraire une augmentation du recours aux phrases avec antéposition simple (103) ($\hat{\beta}= 1.02$, 95%CrI=[-0.15,2.21], $P(\beta)>0=0.95$).

- (101) d’abord i(l) faut qu’on parle du cancer dans sa globalité et quand même le cancer c’est **quoi**? c’est l’annonce quand même d’une terrible nouvelle ()
(EPAC 20040325 - #30)
- (102) on parle beaucoup de YYYY on a raison () mais ce que je constate on n’a presque pas parlé du programme de YYYY **qu’est-ce** qu’il cherche YYYY? ()
(EPAC 20040813 - #156)
- (103) la question de l’auditrice était très intéressante **pourquoi** la tchéchénie qui est si pauvre représente tant d’importance pour la russie?
(EPAC 20040906 - #390)

Il est en outre notable que les séquences impliquant une reformulation sont associées, comparativement à la baisse visible de la proportion d’antépositions avec inversion, à une augmentation du recours aux antépositions en *est-ce que* (104) ($\hat{\beta}= 1.13$, 95%CrI=[0.06,2.21], $P(\beta)>0=0.98$).

- (104) la question que je voudrais poser à () s’il peut répondre ce serait important je crois c’est de savoir pour quelle raison compte tenu de la campagne intense qu’a menée () **pourquoi** est-ce qu’il n’a pas pris le risque d’annoncer qu’il était candidat
(EPAC 20040817 - #208)

Pour ce qui est des interrogatives servant une visée d'animation, surtout associées au journaliste animateur, il est intéressant de constater qu'elles sont associées à une plus forte proportion d'antépositions avec inversion ((105), en ligne avec les chiffres présentés dans la comparaison des rôles, voir Section 4.1.3.2). Cette plus forte proportion a pour conséquence première une baisse de l'usage, dans ces circonstances, des deux autres formes d'antépositions (simples, $\hat{\beta} = -0.74$, 95%CrI=[-1.56,0.03], $P(\hat{\beta}) < 0.001$; en *est-ce que*, $\hat{\beta} = -1.15$, 95%CrI=[-2.58,0.05], $P(\hat{\beta}) < 0.001$).

(105) **que** pouvez-vous répondre à cette question?

(EPAC 20040813 - #165)

4.1.4.3 Des stratégies liées à l'argumentation

Le dernier bloc de critères interactionnels comprend ceux, plus subjectifs par nature, qui sont relatifs à des processus de type oratoire (*fin d'argumentaire*, *réponse auto-fournie*, *opinion implicite*). Le Tableau 4.7 et la Figure 4.7 donnent la distribution des variantes interrogatives dans le sous-corpus selon que ces critères sont remplis ou non.

	Fin d'argumentaire		Réponse Auto-fournie		Opinion implicite	
	Oui	Non	Oui	Non	Oui	Non
Antéposition simple	27	18	6	39	21	24
Antéposition + ESK	19	5	5	19	17	7
Antéposition et inversion	76	48	15	109	50	74
In situ	7	11	9	9	1	17
TOTAL	129	82	35	176	89	122

TABLEAU 4.7 – Interrogatives partielles par type et par critère d'"argumentation" (EPAC)

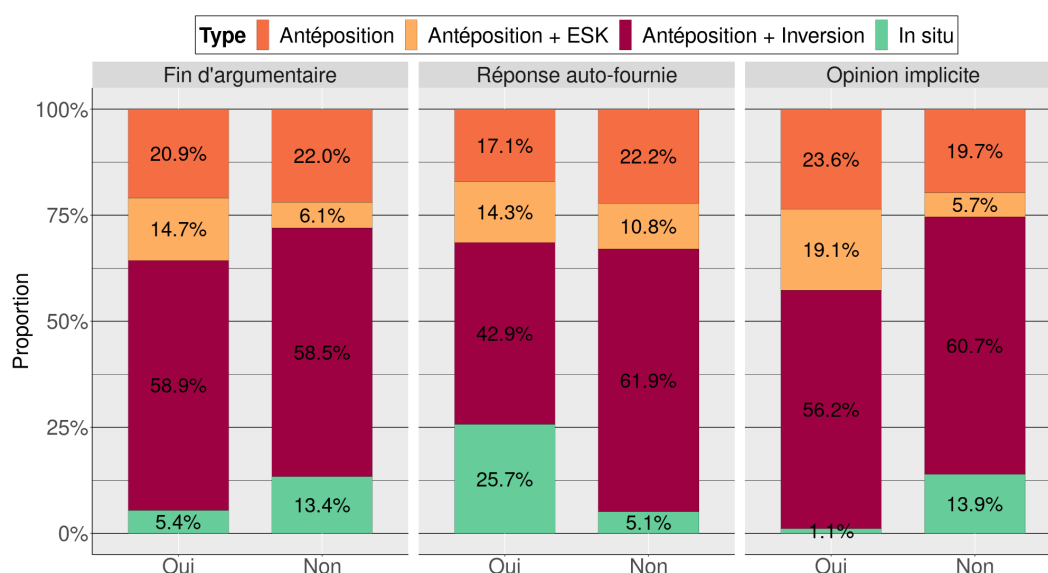


FIGURE 4.7 – Interrogatives partielles dans EPAC par type et critère d'"argumentation" (EPAC)

La comparaison des proportions brutes et les modèles calculés pour ces critères mettent à nouveau ici en relief que des stratégies interactionnelles différentes sont associables à un comportement linguistique différent. Le premier critère permet ainsi de voir que les interrogatives

produites en fin d'argumentation, ou après le développement appuyé d'un thème de réflexion, sont plus susceptibles d'être des antépositions en *est-ce que* (106) ($\hat{\beta} = 0.87$, 95%CrI=[-0.22,2.10], $P(\beta) > 0 = .94$) et moins susceptibles d'être des phrases in situ ((107) pour une exception) ($\hat{\beta} = -0.87$, 95%CrI=[-1.96,0.20], $P(\beta) < 0 = 0.94$).

- (106) () les entreprises en fait ne pouvant pas ou () ont été hyper sélectives et **qu'**est-ce qu'elles ont fait? Elles ont augmenté (...)

(EPAC 20040824 - #295)

- (107) c'est vrai que tout récemment le président bush a déclaré qu'il était pas question de toucher () ces réserves c'est **quoi** exactement?

(EPAC 20040806 - #125)

Par ailleurs, il faut comparer les deux procédés les plus proches de questions rhétoriques (*réponse auto-fournie*, où la réponse à la question est immédiatement fournie par la personne qui l'a posée, et *opinion implicite*, où la réponse que la personne croit être la bonne est facilement perceptible à travers la question qu'elle pose).

Le premier (corrélé par ailleurs à la prise de posture mentionné en 4.1.4.1) se traduit par une baisse de la proportion d'antépositions avec inversion, directement compensée par une augmentation des phrases in situ (108) ($\hat{\beta} = 2.42$, 95%CrI=[1.13,3.92], $P(\beta) > 0 = 1$), mais aussi des antépositions en *est-ce que* (109) ($\hat{\beta} = 1.03$, 95%CrI=[-0.39,2.46], $P(\beta) > 0 = 0.93$).

- (108) quand vous rencontrez un libraire i(l) vous parle **de quoi?** i(l) vous parle des livres de ()

(EPAC 20040826 - #321)

- (109) on lui a confié les con() les fonctions de numéro deux du gouvernement avec celles de ministre de l'économie () **qu'**est-ce que cela veut dire au-delà du sigle du ministère? Ça veut dire qu'on lui a confié

(EPAC 20040817 - #218)

Le second se traduit au contraire par une proportion stable d'antépositions avec inversion mais une disparition quasi-totale, en comparaison, des phrases in situ ((110) pour une exception) ($\hat{\beta} = -2.95$, 95%CrI=[-5.99,-0.97], $P(\beta) < 0 = 1$), et une augmentation des deux autres types d'antépositions, notamment de celles en *est-ce que* (111) ($\hat{\beta} = 1.53$, 95%CrI=[0.47,2.73], $P(\beta) > 0 = 1$).

- (110) () l'euro c'est **pour quand?** ()

(EPAC 20040402 - #76)

- (111) **pourquoi** est-ce que les autorités les pouvoirs publics ne euh demandent pas des dépistages beaucoup plus généralisés de cette maladie dont le traitement est simple?

(EPAC 20040910 - #472)

En confrontant l'observation de ces deux critères, *réponse auto-fournie* et *opinion implicite*, c'est ainsi une sorte de spécification de la structure in situ en termes rhétorique ou oratoire qui pourrait être illustrée. Une question qui n'attend pas de réponse parce que celle-ci est *fournie* juste après sera ainsi plus probablement sous forme in situ (108). Au contraire, une question qui n'attend pas de réponse car celle-ci est *sous-entendue* dans la question elle-même aura moins de chances

d'être une phrase in situ (sauf exceptions comme (110), pour laquelle la réponse sous-entendue, *il faudrait que l'euro arrive bientôt* (?), demeure qui plus est très subjective).

4.1.5 Bilan provisoire

Malgré sa taille relativement réduite en termes d'interrogatives partielles (N=211), ce sous-corpus du projet EPAC apporte déjà de nombreux éléments sur la variation en français des interrogatives partielles. Dans un contexte interactionnel spécifique et relativement homogène (peu de variation à travers différents thèmes de discussion), il apparaît que peuvent avoir une influence autant le rôle des personnes impliquées dans la communication (animateur, personne interrogée, personne tierce) que leurs enjeux et leurs stratégies discursives et argumentatives personnelles au cours de la communication. La gestion des relations inter-personnelles et l'évolution de celles-ci au cours de l'interaction changent les stratégies interrogatives, tout autant que le type d'apport à l'échange (reformulation, répétitions) ou la stratégie oratoire choisie pour faire passer son message (implicite ou explicite). C'est donc, plus largement, toute la balance des rôles dans l'interaction (argumentation ou animation, contribution ou contradiction, prise de posture ou rapprochement des destinataires) qui peut avoir une influence.

Cette balance varie au cours même d'une interaction, comme l'illustre le sous-corpus EPAC, mais elle est aussi plus largement définie par le type de contexte (social, mais aussi temporel) qui entoure l'interaction linguistique. C'est pour élargir la réflexion à ces autres dimensions qu'une étude du corpus ESLO a été menée.

4.2 Un corpus adulte diachronique avec des contextes variés : ESLO

Le corpus ESLO [ESHKOL-TARAVELLA et al., 2011; LLL, 2017] est un ensemble de deux corpus (ESLO1, créé dans les années 1960, et ESLO2, lancé dans les années 2010) de français spontané oral, récoltés à 50 ans d'intervalle selon des protocoles identiques, ce qui le rend assez unique. Une projection des métadonnées sociolinguistiques des individus intervenant dans le corpus sur les transcriptions [BADIN et al., *Accepté*], couplée à une annotation syntaxique des énoncés interrogatifs extraits par la suite [THIBERGE et al., *en révision*], a permis une analyse particulièrement enrichissante de la variation présente dans les interrogatives partielles du français.

4.2.1 Présentation et préparation pour l'analyse syntaxique

ESLO1 comporte 280 heures d'enregistrement pour 3.1 millions de mots et ESLO2 142 heures d'enregistrement pour 1.8 million de mots. Par enregistrement, ces données sont associées à de riches métadonnées concernant à la fois le type de l'interaction en elle-même (interview, conférence, interaction libre) mais aussi les personnes enregistrées (âge, éducation, catégorie socio-professionnelle). Un outil proposé par l'équipe du projet ESLO permet normalement de trier les enregistrements en fonction de ces critères, mais ne permet pas une exploration systématique et ne permet pas une recherche en partie du discours, primordiale pour l'étude des interrogatives partielles.

Il a donc fallu développer une chaîne de traitement pour projeter sur chaque énoncé, et même chaque mot du corpus, les informations des locuteurs et locutrices les ayant produits. C'est ce qui a été fait via l'outil TXM [HEIDEN et al., 2010], qui permet une importation combinant lemmatisation

et annotation en partie du discours avec TreeTagger [SCHMID, 1994, 1995]. Après avoir téléchargé la version Transcriber [BARRAS et al., 2001] du corpus, il a été converti dans un format XML compatible avec TXM par l’outil TEI-CORPO [LIÉGEOIS et al., 2015; MODYCO, 2016]. Une attention particulière a été portée, dans la projections des métadonnées sociolinguistiques, aux situations de chevauchement, dans lesquelles plusieurs personnes appartenaient à un même tour de parole dans le fichier originel.

Les métadonnées sociolinguistiques étaient regroupées dans un tableur avec une ligne pour chaque enregistrement et, en colonnes, le corpus d’appartenance (ESLO1 ou ESLO2) et le type d’interaction, puis les informations sur les personnes y intervenant : identifiant unique, genre, âge, profession, catégorie socio-professionnelle et niveau d’études. C’est à ce moment de la fusion des données qu’a aussi été opéré un découpage en parties du discours en utilisant TreeTagger et un modèle développé précisément pour les données orales francophones [ATILF et al., 2012; BENZITOUN et al., 2012]. Le corpus ainsi compilé a pu être interrogé par le moteur de recherche interne de l’outil TXM, avec des recherches en *Corpus Query Language (CQL)* classiques, au niveau général du corpus aussi bien que dans des sous-corpus déterminés sur la base des métadonnées projetées (tranche d’âge, profession, etc.).

Le détail de ces opérations techniques est rapporté dans BADIN et al. [Accepté].

4.2.2 Annotation syntaxique

Une extraction automatique des interrogatives partielles dans les sous-corpus générés par la procédure décrite ci-dessus, pour les personnes des groupes 15-25 ans et 35-55 ans, a été effectuée par expression régulière. Cette extraction a renvoyé un grand nombre de réponses (N=1715). De ces réponses, une partie a dû être supprimée car elles correspondaient en fait à des productions de membres appartenant à l’équipe scientifique de collecte du projet ESLO2, dont les identifiants uniques suivent un patron de codage commun avec les personnes interrogées. 1399 réponses à la requête provenaient de personnes naïves et étaient d’intérêt.

Ces 1399 éléments ont été annotés manuellement par deux annotateurs, pour vérifier la nature des extractions et pour exclure les phrases relatives en *qui* et *que*, homophones de mots interrogatifs utilisés pour la requête en expression régulière, ainsi que d’autres réponses non-pertinentes (circonstancielle en *quand*, etc.). Les critères utilisés pour cette extraction étaient de nature binaire et codés 1 ou 0 selon que le critère était rempli ou non. La liste des critères se trouve ci-dessous.

- *Partielle* : est-ce que l’extraction correspond à la définition d’une interrogative partielle en demandant une information manquante sur le monde (1) ou non (0). Ce critère permet à la fois d’exclure les phrases relatives mais aussi les actes de langage comme les demandes de répétitions etc.
- *Solo* : est-ce qu’un verbe domine syntaxiquement l’élément interrogatif présent (0) ou non (1). Ce critère, s’il est rempli, permet d’exclure les mots interrogatifs en isolation.
- *Racine* : est-ce que l’extraction est une phrase racine (1) ou non (0). Ce critère permet d’exclure les interrogatives enchâssées, au fonctionnement un peu différent des racines.
- *Enchâssé* : est-ce que l’élément interrogatif de l’extraction est bien présent dans la partie racine de l’extraction, et non pas dans une phrase subordonnée (redondant avec le précédent critère).

- *Infinitive* : est-ce que l'élément interrogatif est un argument d'un verbe fini (0) ou non (1). Ce critère, si rempli, exclut les phrases du type « Que faire? ».
- *Antéposition* : l'élément interrogatif est-il au début du syntagme verbal dont il est un argument ou un adjectif (1), ou non (0).
- *Inversion* : y a-t-il une inversion verbe-sujet (1) ou non (0) dans le syntagme verbal dont l'élément interrogatif est un argument ou un adjectif.
- *ESK* : la locution interrogative *est-ce que* relie-t-elle l'élément interrogatif au reste de la phrase interrogative (1) ou non (0).
- *Négation* : le verbe dominant directement l'élément interrogatif sur le plan syntaxique a-t-il une négation explicite (1) ou non (0). La négation double (*ne... V... pas*) et la négation simple (\emptyset ... *V... pas*) ont ici été codées de façon similaire.

Le type d'élément interrogatif utilisé dans les extractions a aussi été relevé. Pour les raisons déjà décrites (notamment la confusion entre forme « in situ » et position préverbale, voir 2.2.1), les interrogatives sujet ont été également relevées, pour pouvoir être exclues par la suite.

Au total, pour les groupes des 15-25 ans et des 35-55 ans, dans les deux corpus publiquement accessibles ESLO1 et ESLO2, 612 extractions parmi les 1399 ont été conservées pour analyse et correspondaient au critère « *interrogative partielle formant une phrase racine finie et complète, qui ne soit pas une interrogative sujet* ». Elles sont disponibles en Annexe C.

L'annotation étant ici uniquement syntaxique, elle est très différente de celle effectuée pour le corpus EPAC. Elle a pour but d'étudier les usages faits des différents types interrogatifs par des populations d'âge différent, à des époques différentes, dans des contextes interactionnels différents. Voici quatre exemples tirés des 612 extraits. Comme pour les exemples tirés du corpus EPAC, ils ont pour label un identifiant de fichier (repris tel quel du projet ESLO, auquel la tranche d'âge a été adjointe) et un numéro d'item (précédé du symbole #), permettant leur identification unique parmi les 612 extractions. Les quatre types syntaxiques relevés sont illustrés : antéposition avec inversion (112), antéposition simple (113), antéposition en *est-ce que* (114) et in situ (115).

(112) **que** voulez-vous? j'ai du raisin de l'ananas du de l'abricot ou du pamplemousse
et de l'orange

(ESLO1ENT086C (35-55) - #22)

(113) ah bah maintenant vous savez c'est qu'ils ont tous euh le plus jeune **quel âge** il
a? hum il a dix-sept ans

(ESLO1ENT092C (15-25) - #22)

(114) **où** est-ce que je l'ai mise ma copie d'histoire?

(ESLO2REPAS1254C (15-25) - #593)

(115) et le but c'est **quoi** c'est d'étudier la façon de parler des gens?

(ESLO2ENT1001C (35-55) - #575)

4.2.3 Observations : Résultats généraux

La répartition générale des 612 extractions selon les quatre différentes variantes considérées est présentée dans le Tableau 4.8. Ces données ainsi que les résultats et analyses statistiques présentées

par la suite dans les Sections 4.2.4-4.2.8 reprennent, globalement, celles présentées dans THIBERGE et al. [en révision].

Antéposition simple	Antéposition + ESK	Antéposition et inversion	In situ
179	77	85	271
29.2 %	12.6 %	13.9 %	44.3 %

TABLEAU 4.8 – Interrogatives partielles dans ESLO par type (15-25 et 35-55 ans confondus)

Il y a déjà ici une différence de répartition avec les données présentées pour le sous-corpus EPAC (Section 4.1.3). La proportion importante de phrases in situ et la présence relativement limitée d’antépositions avec inversion rappelle en revanche les données de HAMLAOUI [2009a], de QUILLARD [2001] ou encore de ADLI [2015]. Une explication à cette différence avec le sous-corpus EPAC est la variété des contextes interactionnels présents dans ESLO, qui n’existe par nature pas dans EPAC. Cette variété sera explorée dans la Section 4.2.7.

4.2.4 Observations : Une évolution diachronique

Le premier atout du projet ESLO est sa nature diachronique, avec des données récoltées sur deux périodes espacées d’une cinquantaine d’années. Le Tableau 4.9 et la Figure 4.8 présentent la répartition des interrogatives partielles extraites selon leur type et en fonction du corpus dont elles proviennent.

	ESLO1	ESLO2	TOTAL
Antéposition simple	91	88	189
Antéposition + ESK	43	34	77
Antéposition et inversion	55	30	85
In situ	61	210	271
TOTAL	250	362	612

TABLEAU 4.9 – Interrogatives partielles dans ESLO1 vs ESLO2 (15-25 et 35-55 ans confondus)

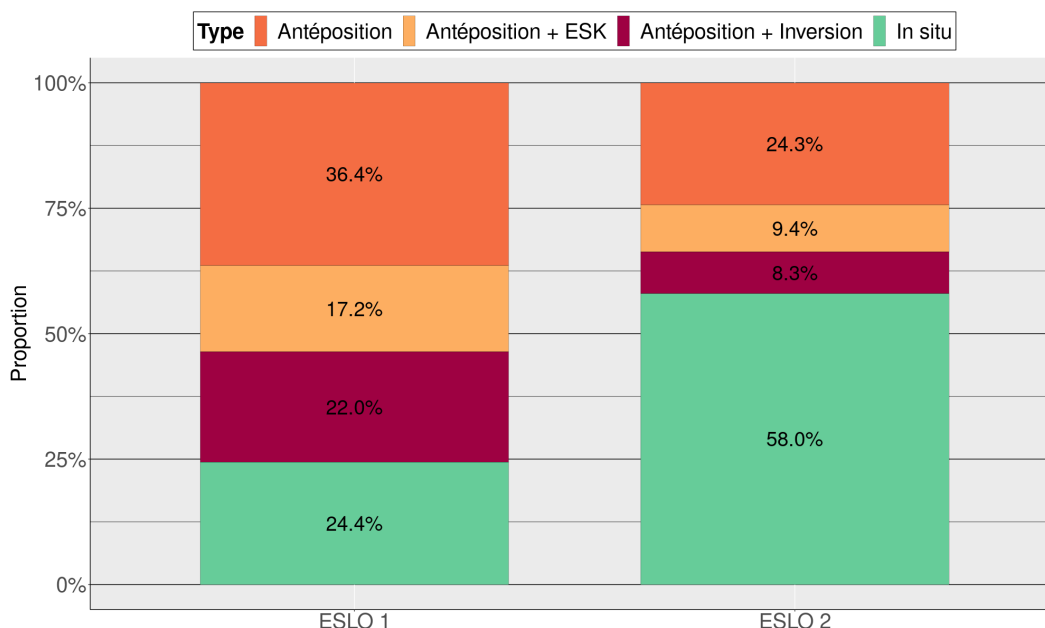


FIGURE 4.8 – Interrogatives partielles dans ESLO1 vs ESLO2 (15-25 et 35-55 ans confondus)

Un modèle de régression bayésienne a été calculé pour les 612 interrogatives recensées pour ces deux périodes, en confondant les données des deux groupes d'âge considérés. Le modèle lancé avait pour variable dépendante le TYPE de l'interrogative (antéposition simple, en est-ce que, avec inversion et in situ, avec pour niveau de référence cette fois les phrases in situ, plus nombreuses) et pour variable indépendante le CORPUS (ESLO1 ou ESLO2, avec pour niveau de référence ESLO1). Un effet avec intercept aléatoire a été ajouté pour tenir compte de la variabilité entre les transcriptions ("fichier"). Ce modèle a été calculé avec 4 chaînes et 3000 itérations par chaîne. La convergence du modèle a été atteinte et vérifiée. Le détail de ce modèle est présenté en Annexe D.0.1 et la Figure 4.9 en donne les distributions postérieures.

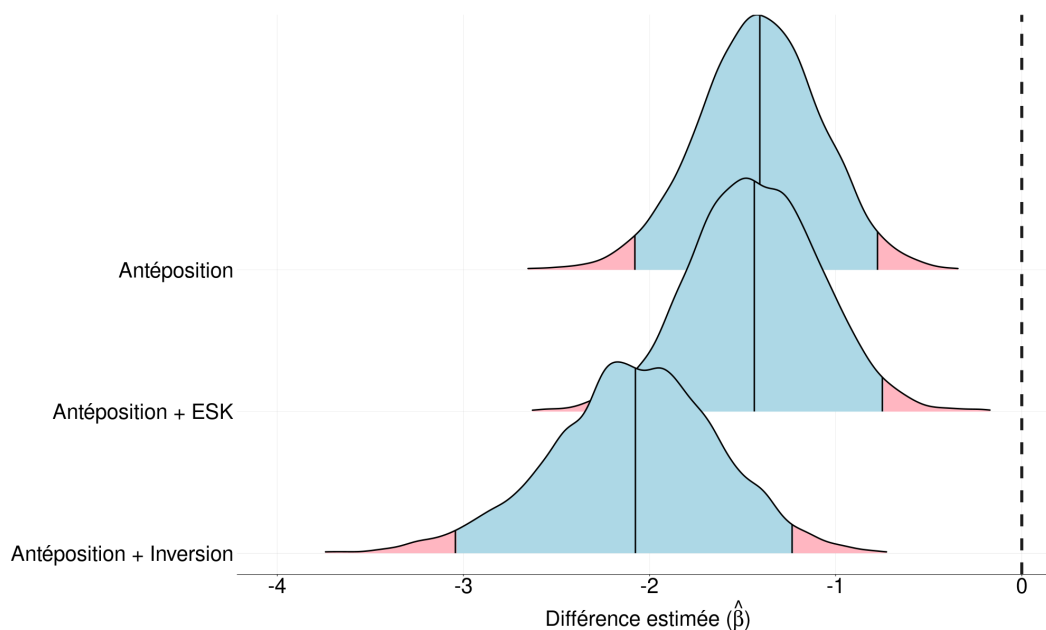


FIGURE 4.9 – Distributions postérieures pour le modèle brms (ESLO, effet de CORPUS)

Il ressort de ces données qu'une importante évolution diachronique semble avoir eu lieu entre les années 1960 et les années 2010, avec une diminution globale de la présence des formes antéposées prises comme un tout, au profit de productions de phrases interrogatives in situ. Plus précisément, la multiplication par plus de deux de la proportion de phrases in situ s'accompagne d'une diminution aussi bien de la proportion d'antépositions simples ($\hat{\beta} = -1.41$, 95%CrI=[-2.08,-0.78], $P(\beta) < 0=1$) que de celle d'antépositions en *est-ce que* ($\hat{\beta} = -1.43$, 95%CrI=[-2.10,-0.75], $P(\beta) < 0=1$) ou encore de celle d'antépositions avec inversion ($\hat{\beta} = -2.09$, 95%CrI=[-3.04,-1.23], $P(\beta) < 0=1$). Ces résultats peuvent être mis en lien avec le parcours diachronique plus large de la construction in situ (LARRIVÉE [2016, 2019]) présenté en Section 2.2.3.1. Il convient néanmoins d'approfondir ces données.

4.2.5 Observations : Des effets générationnels ?

La chaîne de traitement décrite en Section 4.2.1 et appliquée aux transcriptions du projet ESLO permet en effet de comparer les productions extraites selon qu'elles sont produites par des personnes de la tranche des 15-25 ans ou par des personnes de 35-55 ans. Le Tableau 4.10 et la Figure 4.10 montrent la répartition des interrogatives partielles extraites selon leur type et en fonction de ces groupes d'âge.

	15-25 ans	35-55 ans	TOTAL
Antéposition simple	37	142	179
Antéposition + ESK	19	58	77
Antéposition et inversion	14	71	85
In situ	73	198	271
TOTAL	143	469	612

TABLEAU 4.10 – Interrogatives partielles par tranche d’âge (ESL0 1&2 confondus)

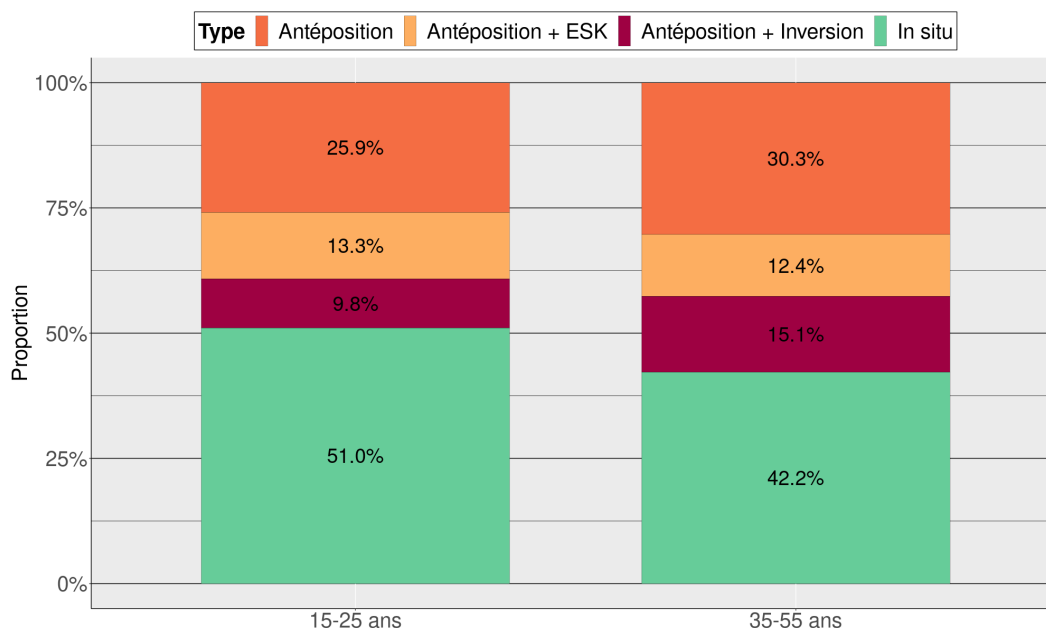


FIGURE 4.10 – Interrogatives partielles par tranche d’âge (ESL0 1&2 confondus)

Un modèle de régression bayésienne a été calculé pour les 612 interrogatives recensées pour ces deux groupes d’âge, en confondant les données des deux corpus ESLO1&2. Le modèle lancé sur le modèle du précédent, avec la même variable dépendante (TYPE syntaxique) mais pour variable indépendante le groupe d’ÂGE (avec pour niveau de référence le groupe des 15-25 ans). Le "fichier" a de nouveau été intégré comme effet avec intercept aléatoire. Ce modèle a été calculé avec 4 chaînes et 3000 itérations par chaîne; la convergence a été atteinte et vérifiée. Le détail de ce modèle est présenté en Annexe D.0.2. Il montre que, en prenant pour niveau de référence les productions de phrase in situ, il n’y a en fait pas de forte différence de proportions entre les groupes d’âge pour les trois types d’antépositions. Dit autrement, s’il y a certes moins de phrases in situ chez les 35-55 ans généralement, cela ne se traduit pas par une préférence supérieure pour un autre type interrogatif particulier, à l’échelle de l’ensemble du projet ESLO pris dans sa dualité.

4.2.6 Confrontation des facteurs d’époque et d’âge

C’est en traitant en interaction les deux types de données précédemment mentionnées, l’époque et le groupe d’âge, qu’une perspective plus complexe apparaît. Le Tableau 4.11 et la Figure 4.11 montrent la répartition des interrogatives partielles extraites selon leur type et en fonction de ces groupes d’âge, pour le projet ESLO1 d’une part et pour le projet ESLO2 d’autre part.

	ESLO 1		ESLO 2	
	15-25 ans	35-55 ans	15-25 ans	35-55 ans
Antéposition simple	23	68	14	74
Antéposition + ESK	13	30	6	28
Antéposition et inversion	8	47	6	24
In situ	14	47	59	151
TOTAL	58	85	192	277

TABLEAU 4.11 – Interrogatives partielles par corpus et par tranche d’âge (ESLO)

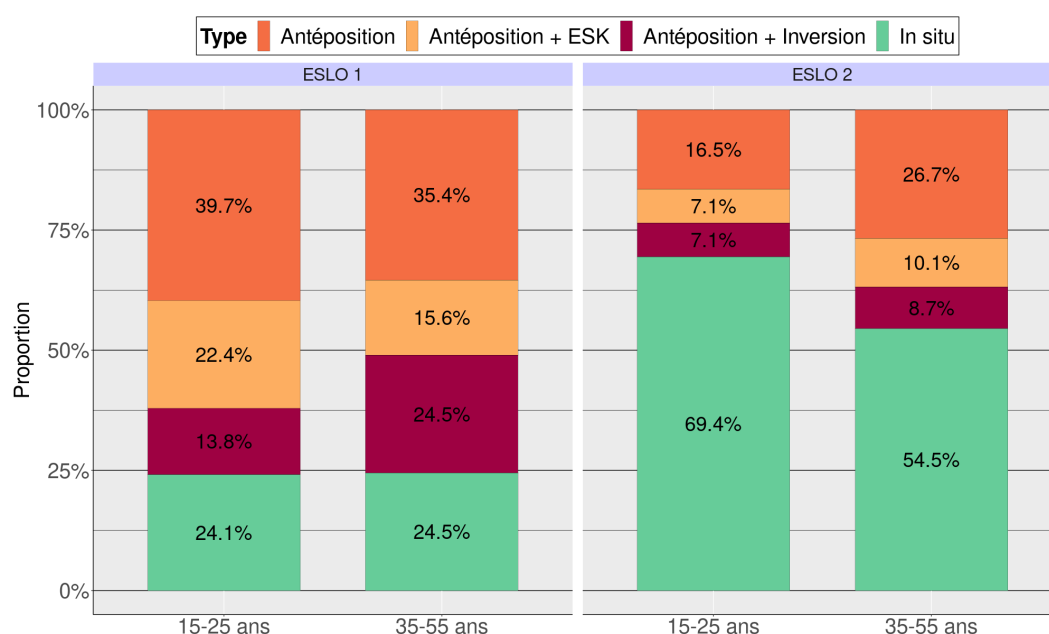


FIGURE 4.11 – Interrogatives partielles par corpus et par tranche d’âge (ESLO)

Un nouveau modèle de régression bayésienne a été calculé pour les 612 interrogatives extraites, toujours avec la même variable dépendante (TYPE syntaxique) mais avec cette fois pour variables indépendantes le CORPUS et le groupe d’ÂGE mises en interaction (mêmes niveaux de référence qu’auparavant, et même effet avec intercept aléatoire). Ce modèle a été calculé avec 4 chaînes et 3000 itérations par chaîne. La convergence du modèle a été atteinte et vérifiée. Le détail de ce modèle est présenté en Annexe D.0.3 et la Figure 4.12 en montre les distributions postérieures.

Les analyses conduisent à une confirmation de l’effet simple de CORPUS et donc de l’influence de l’époque sur la production des trois antépositions (proportions moindres de ces trois types dans les années 2010 que dans les années 1960). L’absence de fort effet de groupe d’ÂGE est également confirmée. La mise en interaction des deux effets néanmoins fait apparaître une influence combinée, au moins pour les antépositions simples et les antépositions en *est-ce que*. D’une génération à l’autre les proportions d’emploi de ces deux types phrastiques évoluent et s’inversent, quand leur usage est comparé à celui des phrases in situ.

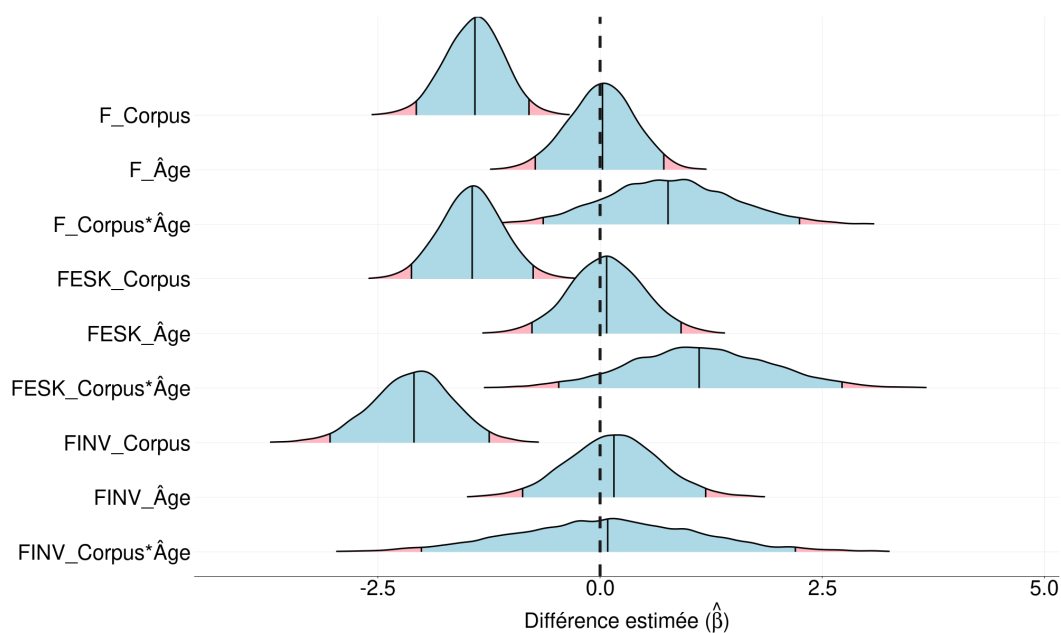


FIGURE 4.12 – Distributions postérieures pour le modèle brms (ESLO, effets CORPUS*ÂGE)

Si les 35-55 ans utilisent moins d’antépositions simples et en *est-ce que* que les 15-25 ans dans ESLO1, c’est en effet l’inverse dans ESLO2 : la proportion d’emploi devient supérieure chez les 35-55 ans autant pour les premières ($\hat{\beta} = 0.78$, 95%CrI=[-0.64,2.24], $P(\beta) > 0 = 0.86$) que pour les secondes ($\hat{\beta} = 1.13$, 95%CrI=[-0.46,2.72], $P(\beta) > 0 = 0.92$). De façon notable, la différence entre générations dans ESLO1 quant aux antépositions avec inversion (proportions d’usage supérieure chez les 35-55 ans) est quasiment neutralisée dans ESLO2.

Au final, si les deux groupes d’âge se distinguent dans ESLO1 par leurs usages différents des trois types interrogatifs avec antéposition, dans ESLO2 la distinction s’opère donc plutôt au niveau de leur usage des phrases in situ et, principalement, des antépositions simples et en *est-ce que*. Une autre façon de décrire ce changement est de dire que les habitudes linguistiques de ces deux groupes générationnels semblent bien avoir changé d’une époque à l’autre (augmentation générale des emplois in situ), mais que l’antéposition simple, notamment, a moins reculé pour le groupe des 35-55 ans. Il est au passage important de souligner ici que le groupe des 35-55 ans du projet ESLO2 ne sont pas les personnes qui appartenaient au groupe des 15-25 ans dans ESLO1.

4.2.7 Observations : Effets de contexte et stratégies interactionnelles

Un autre atout important du projet ESLO, pour la présente analyse, est la grande variété de contextes d’interaction dans lesquels les productions de francophones ont été recueillies. Le Tableau 4.12 donne la distribution des 612 interrogatives extraites, ici par type syntaxique mais aussi par contexte, tels que définis dans la méthodologie de récolte du projet. Dans l’ordre, le contexte 24H correspond à des enregistrements sur 24h d’une même personne, le contexte CINÉMA à des discussions dans la rue avec des gens attendant aux abords de salle de projection dans Orléans, le contexte ÉCOLE à des enregistrements d’enseignant.e.s de primaire en salle de classe, le contexte ENTRETIEN à des interviews de personnes chez elles selon un questionnaire pré-scripté, le contexte ITINÉRAIRE à des demandes d’indications d’itinéraire à des personnes interpellées dans les rues d’Orléans, le contexte REPAS à des enregistrements de familles aux moments des repas sans obser-

vation extérieure, le contexte RÉUNION à des réunions professionnelles, et le contexte TÉLÉPHONE à des appels téléphoniques. Le label DIVERS regroupe des récoltes plus éparses.

Comme il est visible, tous les contextes de récolte de données ne sont pas aussi propices à la production d'interrogatives partielles par les consultants et consultantes linguistiques (il est à préciser qu'aucune interrogative produite par les équipes de récolte ne se trouve dans les 612 extractions). La Figure 4.13 donne la proportion des interrogatives partielles produites dans les trois contextes les plus favorables que sont les ENTRETIENS, l'ÉCOLE et les REPAS (N=556).

	Antéposition simple	Antéposition + ESK	Antéposition et inversion	In situ
24h	-	-	1	2
Cinéma	1	-	1	14
Divers	3	-	-	2
École	39	7	19	65
Entretien	112	59	60	103
Itinéraire	1	1	-	10
Repas	16	6	3	67
Réunion	7	4	1	7
Téléphone	-	-	-	1
TOTAL	179	77	85	271

TABLEAU 4.12 – Interrogatives partielles par contexte, ESLO 1&2 confondus

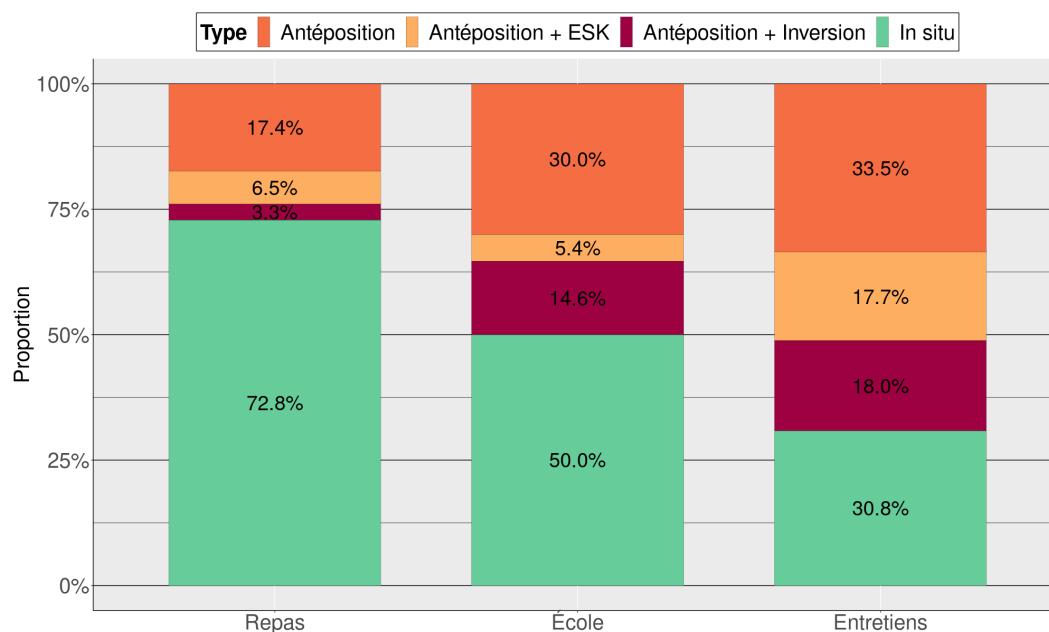


FIGURE 4.13 – Interrogatives partielles par contexte, ESLO 1&2 confondus

Un nouveau modèle de régression bayésienne a été calculé pour ces 556 interrogatives recensées pour les trois contextes principaux, en confondant les données des deux groupes d'âge considérés et des deux périodes de récolte. Le modèle lancé avait pour variable dépendante le TYPE de l'interrogative (trois types d'antéposition, avec pour niveau de référence les phrases in situ) et pour variable indépendante le CONTEXTE (repas, école ou entretiens, avec pour niveau de référence l'école, qui semble visuellement un état intermédiaire entre les deux autres). L'effet avec intercept aléatoire "fichier" est toujours intégré. Ce modèle a été calculé avec 4 chaînes et 3000 itérations par chaîne. La convergence du modèle a été atteinte et vérifiée en contrôlant que chaque Rhat soit à

1.00 pour tous les paramètres. Le détail de ce modèle est présenté en Annexe D.0.4 et la Figure 4.14 en donne les distributions postérieures.

Le premier phénomène visible est la différence de proportions des phrases in situ dans les trois contextes, avec une sorte de gradient allant des ENTRETIENS (un peu moins d'un tiers) aux REPAS (presque les trois quarts), en passant par l'ÉCOLE (environ la moitié). En comparaison avec cette évolution, le modèle statistique confirme plusieurs autres différences. Notamment, il confirme les moindres proportions d'antépositions simples et d'antépositions avec inversion dans les contextes de REPAS (respectivement $\hat{\beta} = -1.38$, 95%CrI=[-2.80,-0.03], $P(\beta) < 0 = 0.98$ et $\hat{\beta} = -2.37$, 95%CrI=[-4.58,-0.38], $P(\beta) < 0 = 0.99$) que dans les contextes scolaires. De l'autre côté, il y a plus d'antépositions simples et plus d'antépositions en *est-ce que* en ENTRETIENS que dans les contextes SCOLAIRES (respectivement $\hat{\beta} = 0.59$, 95%CrI=[-0.31,1.53], $P(\beta) > 0 = 0.91$, $\hat{\beta} = 1.82$, 95%CrI=[0.78,2.96], $P(\beta) > 0 = 1$). Il y aussi, mais c'est une tendance moins visible, plus d'antépositions avec inversion en ENTRETIENS qu'à l'ÉCOLE ($\hat{\beta} = 0.35$, 95%CrI=[-0.95,1.60], $P(\beta) > 0 = 0.72$).

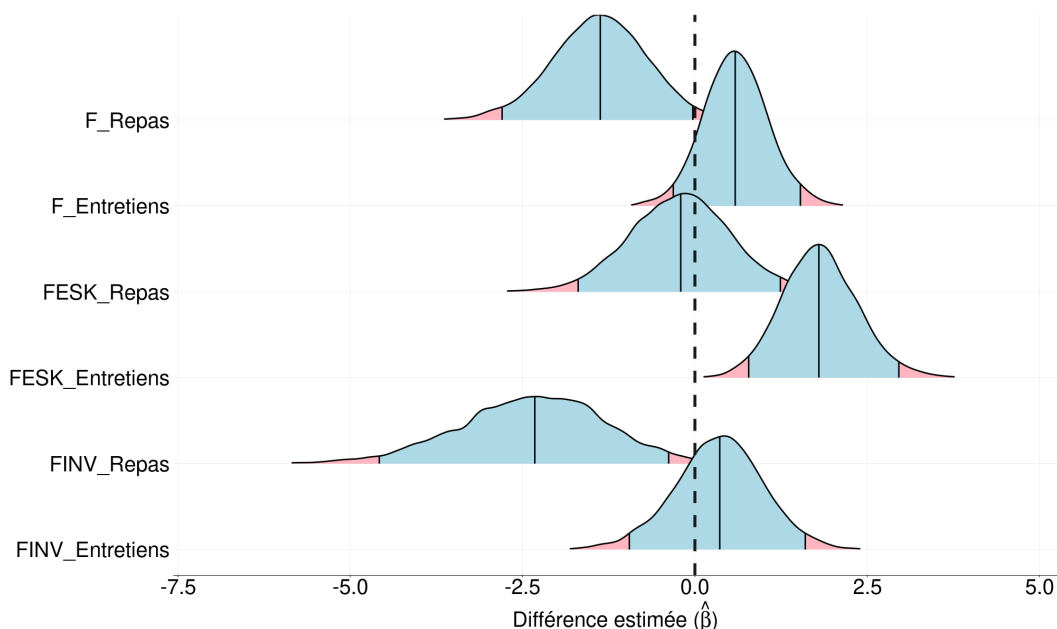


FIGURE 4.14 – Distributions postérieures pour le modèle brms (ESLO, effet de CONTEXTE)

Ces observations font bien ressortir un aspect graduel de l'influence du contexte interactionnel sur les proportions des variantes interrogatives utilisées. Une hypothèse qui peut expliquer cette variation inter-contextes tient aux caractéristiques de ceux-ci. Parmi ces caractéristiques, le public de l'interaction et la formalité requise en fonction de ce public sont en effet différents.

Lors d'un repas en famille, l'échange est par nature généralement plus relâché, se déroulant entre proches d'un même foyer. Au contraire, lors d'un échange avec une équipe de recherche, la personne à qui l'on s'adresse est étrangère à ce cercle proche, et ce même si l'échange se déroule dans un endroit familial et touche à des thèmes courants (les entretiens se déroulaient chez les personnes interrogées et tournaient autour de leur vie quotidienne). Qui plus est, malgré les précautions prises, lors des entretiens, discuter de soi et de ses habitudes (alimentaires, linguistiques, culturelles) installe la personne interrogée dans une posture où son intimité est dévoilée, en partie, à un tiers. C'est un contexte qui pousse sans doute l'individu à plus surveiller ses choix linguistiques et à plus les soumettre à la pression normative, par comparaison avec un échange lors d'une activité

de réunion familiale, autour d'un repas.

Les contextes scolaires seraient, dans cette perspective, une sorte d'intermédiaire. Sur ces contextes pèse une certaine exigence de formalité pour l'équipe pédagogique (poids de l'institution scolaire, apprentissage du « français standard » en classe). En contrepartie cependant, le public y est cette fois constitué d'enfants, dont il faut s'assurer l'attention permanente. Ceci, notamment, peut se faire en créant une certaine proximité linguistique, par des marques d'informalité, mais aussi donc potentiellement par des phrases *in situ*, plus souvent employées dans des contextes informels.

Il pourrait s'agir d'une influence spécifique d'un « langage dirigé vers l'enfant » dans ce contexte scolaire. Sans nier cette possibilité, le contraste entre (116) et (117) illustre néanmoins la tension chez les pédagogues. Le premier exemple provient d'une situation où l'antéposition avec inversion semble répondre à un impératif normatif (exercice de calcul avec une consigne écrite). Le second exemple, toujours dans un exercice, semble plus spontané et la subordonnée hypothétique « *si tu avais deux kilos de bonbons* » permet un rapprochement avec l'enfant.

(116) le vendeur lui reprend son ancien appareil pour cent quatre euros et les accessoires qui le complètent pour trente-quatre euros **quelle somme** Monsieur Portrait devra-t-il ajouter? le vendeur accepte un paiement en trois versements (...)

(ESLO2ECOLE1278C (35-55) - #267)

(117) si tu avais deux kilos de bonbons tu ferais **quoi** comme opération? réfléchis à cette question et ça te donnera la réponse

(ESLO2ECOLE1280C (35-55) - #297)

4.2.8 Confrontation des facteurs d'époque, d'âge et de contexte

Vu la disparité entre contextes en termes de productions d'interrogatives, il est nécessaire d'intégrer ce paramètre aux données précédemment considérées, notamment les effets de CORPUS et de groupe d'ÂGE.

4.2.8.1 Confrontation des facteurs d'époque et d'âge (interviews)

Le seul contexte permettant une analyse des effets d'ÂGE et de CORPUS en interaction sont les entretiens, avec suffisamment de données pour les deux tranches d'âge et sur les deux époques de récolte. Le Tableau 4.13 et la Figure 4.13 présentent la distribution des interrogatives par corpus et par tranche d'âge pour ces données (N=334 au total).

	ESLO 1		ESLO 2	
	15-25 ans	35-55 ans	15-25 ans	35-55 ans
Antéposition simple	16	65	6	25
Antéposition + ESK	9	30	5	15
Antéposition et inversion	7	47	3	3
In situ	7	39	21	36
TOTAL	39	181	35	79

TABLEAU 4.13 – Interrogatives partielles par corpus et par tranche d'âge (ESLO, interviews)

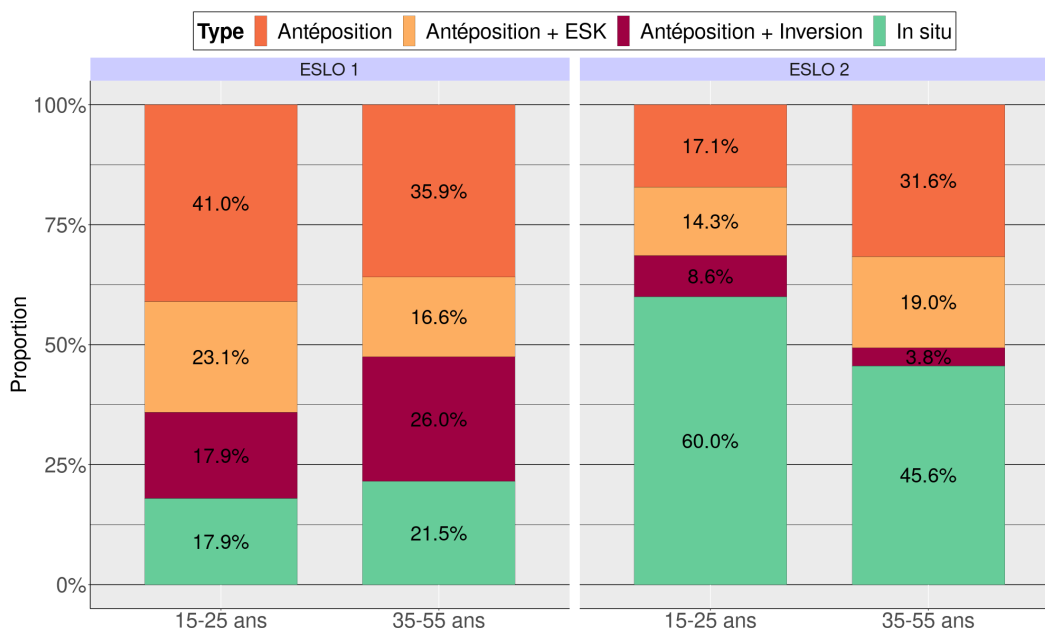


FIGURE 4.15 – Interrogatives partielles par corpus et par tranche d'âge (ESLO, interviews)

Les données sont moins fournies d'une part pour le groupe des 15-25 ans que pour les 35-55 ans (N=74 contre 260), et d'autre part pour ESLO1 que pour ESLO2. Pour en voir la portée généralisable, un modèle de régression bayésienne, en tout point semblable à celui calculé pour l'ensemble des extractions ESLO (N=612) et intégrant les facteurs de CORPUS et de groupe d'ÂGE, a été calculé sur ce sous-ensemble de données. Les résultats en sont relativement semblables (détails en Annexe D.0.5).

Il semble y avoir un changement de comportement linguistique entre générations, avec plus de productions in situ pour les deux groupes d'âge dans les années 2010 et donc une réduction globale de chacun des trois types d'antéposition (effet simple de CORPUS), mais ceci s'accompagne d'une plus grande persistance chez les 35-55 ans des productions avec antépositions simples ($\hat{\beta} = 1.31$, 95%CrI=[-0.85,3.55], $P(\beta) > 0 = 0.89$) et en *est-ce que* ($\hat{\beta} = 1.15$, 95%CrI=[-0.75,3.07], $P(\beta) > 0 = 0.88$).

4.2.8.2 Confrontation des facteurs d'âge et de contexte, ESLO 2

L'influence du contexte peut également être observée en la croisant avec les données d'âge, cette fois-ci au sein du seul corpus ESLO2, en comparant au sein de sous-ensemble les contextes d'ENTRETIENS et de REPAS. Le Tableau 4.14 et la Figure 4.16 donnent les proportions d'interrogatives par contexte et par groupe d'âge, au sein du sous-corpus ESLO2 (N=201 au total).

	Entretiens		Repas	
	15-25 ans	35-55 ans	15-25 ans	35-55 ans
Antéposition simple	6	25	7	9
Antéposition + ESK	5	15	1	5
Antéposition et inversion	3	3	1	2
In situ	21	36	21	41
TOTAL	35	79	30	57

TABEAU 4.14 – Interrogatives partielles par corpus et par tranche d'âge (ESLO, interviews)

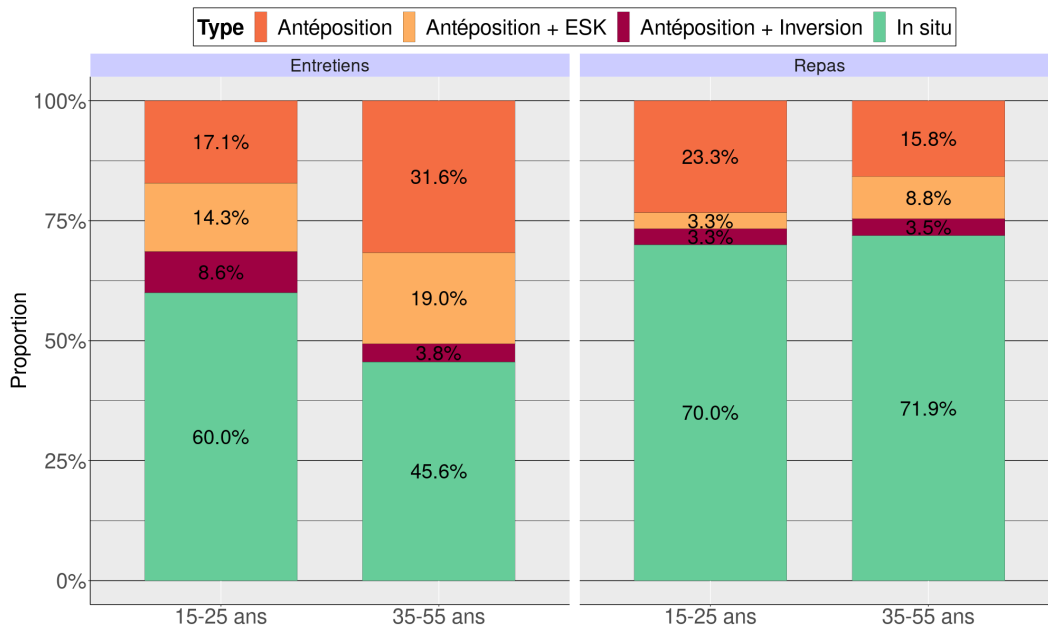


FIGURE 4.16 – Interrogatives partielles par tranche d'âge et contexte (ESLO 2)

Ici encore les données sont moins fournies pour les 15-25 ans que pour les 35-55 ans ($N=65$ contre 136), mais un dernier modèle de régression bayésienne permet d'évaluer la possibilité de généraliser ces données. Le TYPE syntaxique était la variable dépendante (phrases in situ comme niveau de référence) et les variables indépendantes étaient l'interaction entre groupe d'ÂGE (avec les 15-25 ans pour niveau de référence) et CONTEXTE (repas ou entretiens, avec ces derniers en niveau de référence). Avec "fichier" en effet à intercept aléatoire, le modèle a été calculé à nouveau avec 4 chaînes et 3000 itérations par chaîne, et la convergence en a été vérifiée. Le détail du modèle est disponible en Annexe D.0.6 et la Figure 4.17 en illustre les distributions postérieures.

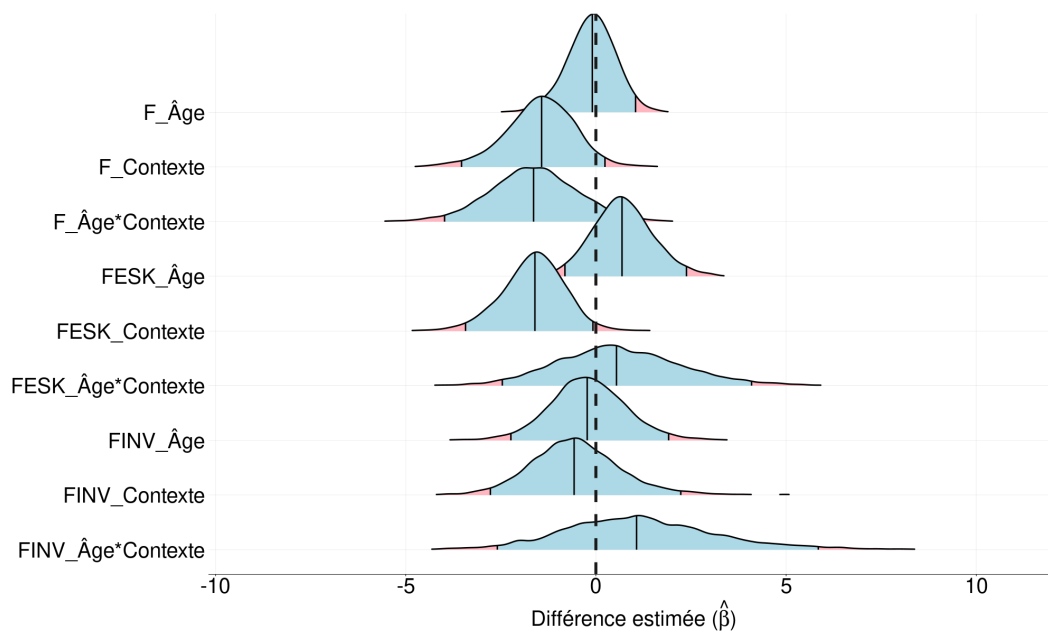


FIGURE 4.17 – Distributions postérieures pour le modèle brms (ESLO2, effets CONTEXTE*ÂGE)

Ces analyses font cette fois apparaître un effet simple du groupe d'ÂGE pour les interrogatives en *est-ce que*. Celui-ci signifie que dans les années 2010, en les comparant aux interrogatives in

situ et indépendamment du contexte, ce type d'antépositions est employé plus souvent par les 35-55 ans que par les 15-25 ans ($\hat{\beta} = 0.71$, 95%CrI=[-0.81,2.38], $P(\beta) > 0 = 0.82$). Un effet simple de CONTEXTE est par ailleurs visible pour les trois types d'antéposition, avec une signification proche de ce qui a déjà été montré en Section 4.2.7 : des contextes différents favorisent des stratégies interrogatives différentes. En l'occurrence, en comparant les REPAS aux ENTRETIENS, il s'observe que l'augmentation des productions in situ entraîne une baisse des proportions d'antépositions simples ($\hat{\beta} = -1.47$, 95%CrI=[-3.53,0.24], $P(\beta) < 0 = 0.96$), d'antépositions en *est-ce que* ($\hat{\beta} = -1.65$, 95%CrI=[-3.43,-0.08], $P(\beta) < 0 = 0.98$) mais aussi, dans une moindre mesure, d'antépositions avec inversion ($\hat{\beta} = -0.51$, 95%CrI=[-2.77,2.23], $P(\beta) < 0 = 0.70$).

Une autre donnée intéressante concerne l'interaction des facteurs CONTEXTE et groupe d'ÂGE, négative pour les antépositions simples ($\hat{\beta} = -1.63$, 95%CrI=[-3.98,0.68], $P(\beta) < 0 = 0.92$). Elle traduit le fait que, alors que les 35-55 ans utilisent plus d'antépositions simples que les 15-25 ans dans un contexte d'ENTRETIEN, c'est l'inverse qui peut s'observer dans le contexte des REPAS. Cette interaction est positive (et moins forte) pour les antépositions avec inversion ($\hat{\beta} = 1.20$, 95%CrI=[-2.59,5.85], $P(\beta) > 0 = 0.71$). Elle traduit ici le fait que les 35-55 ans utiliseraient plus d'antépositions avec inversion dans le contexte des REPAS alors que c'est l'inverse au cours des ENTRETIENS. Le très faible nombre d'occurrences d'antépositions avec inversion dans ce sous-ensemble de données incite néanmoins à traiter cette interaction avec prudence.

En prenant les analyses précédentes à l'envers, un fait émerge néanmoins. Les stratégies interrogatives semblent certes plutôt homogènes entre les deux groupes d'âge en cours de repas (proportion importante d'in situ, notamment). En contraste, il y a plus de différences entre générations durant les entretiens (moins de phrases in situ pour les 35-55 ans, et comparativement plus d'antépositions simples et en *est-ce que*). Cette observation montre que, si des comportements linguistiques sont différents d'une époque à une autre et d'un contexte à l'autre alors qu'aucune influence du facteur générationnel ne semble perceptible à l'échelle globale, ce n'est pas pour autant qu'à un niveau plus précis d'analyse, il en est de même.

Ici, une disparité est visible entre les 15-25 ans et les 35-55 ans, au sein de la même fenêtre temporelle, en ce qui concerne leur appréciation des contraintes pesant sur tel ou tel contexte. Le fait que les proportions d'emploi des quatre variantes soient semblables lors des repas montre un accord sur la notion de contexte relâché ou informel, avec un public connu. La différence de comportement linguistique en entretiens (plus de phrases in situ pour les 15-25 ans et plus d'antépositions simples et en *est-ce que* pour les 35-55 ans) plaide par contre pour une stratégie différente de conformation à la norme. La contrainte normative, qui rend globalement le type in situ moins acceptable en contexte plus formel, pèse peut-être moins sur les 15-25 ans que sur les 35-55 ans. Il peut aussi s'envisager –et c'est peut-être plus probable– que leur définition même du standard linguistique est différente de celle des 35-55 ans. Ce standard serait alors, pour les 15-25 ans, quelque chose comme « *Les phrases in situ ne sont pas si inacceptables en contexte formel* » ou, au contraire, comme « *Les antépositions simples sont moins acceptables en contexte formel* ».

4.2.9 Bilan provisoire

Le corpus EPAC avait pour intérêt d'illustrer combien la variation dans l'utilisation des interrogatives partielles du français pouvait, à contexte égal, correspondre à des stratégies interactionnelles différentes, selon les personnes et leur but communicatif. L'exploitation du projet ESLO donne une

perspective plus globale encore sur les productions d'adultes francophones, cette fois dans des contextes variés et à travers des époques différentes, mais aussi en considérant leur groupe d'âge. Il en ressort d'abord que l'utilisation des interrogatives partielles varie à travers les époques, avec un développement apparent de l'emploi des phrases *in situ* au détriment des productions antéposées notamment, au cours de la seconde moitié du vingtième siècle. Ceci rejoint les observations de LARRIVÉE [2016, 2019] (Section 2.2.3.1), sur le même corpus ESLO mais étendu à d'autres éléments interrogatifs que *comment*.

Une autre observation est que le contexte semble avoir une forte influence sur les stratégies interrogatives, avec notamment une diminution du recours aux phrases *in situ* à mesure que le contexte devient plus formel. Enfin, l'âge d'une personne semble avoir une influence non-négligeable sur son recours à telle ou telle variante interrogative, avec des préférences différentes en la matière entre les générations, que ce soit en comparant une époque à une autre dans le même contexte, ou bien un contexte à un autre au sein d'une même époque.

4.3 Conclusion intermédiaire

La combinaison des résultats issus aussi bien du sous-corpus du projet EPAC que de celui du projet ESLO plaide donc pour une analyse revisitée des interrogatives partielles du français et de leur usage par les adultes francophones, en conversation spontanée. Sans nier la pertinence des autres facteurs dégagés par la littérature, il faut donc en intégrer de nouveaux ou les réintégrer en coordination les uns avec les autres. Les théories liant situation informationnelle et forme interrogative peuvent ainsi gagner à prendre en compte les séquences oratoires dans lesquelles celle-ci apparaissent, et la façon dont l'argumentation y est menée. Les stratégies de distanciation ou de rapprochement entre personnes participant à une même conversation sont aussi à y relier.

Par ailleurs, le croisement des facteurs d'époque et de génération permet d'inscrire les pratiques en matière d'interrogation partielle dans une réflexion plus large sur le changement linguistique. Ces données rejoignent en effet les théories présentées par exemple dans KELLER [1994] et HASPEL-MATH [1999] faisant de l'individu linguistique le moteur de l'évolution linguistique, par un cercle auto-entretenu d'innovation (maxime d'*extravagance*) et de conformisme (maxime de *conformité*) à des phénomènes de « mode » linguistique. Les différences entre générations en matière d'interrogation partielle, autant à travers les époques que dans leur appréciation différente des contraintes pesant sur un même contexte, en sont une illustration.

Il est néanmoins besoin, pour intégrer ces facteurs déduits d'observations de corpus, de s'assurer avant tout de leur valeur généralisable. Dans une perspective sociolinguistique, ce sont notamment les thématiques du positionnement social (posture, proximité avec le ou la destinataire) mais aussi du contexte qui seront approfondies dans les chapitres suivants, par les méthodologies expérimentales. Ceci passera par une exploration systématisée du rôle de la formalité du contexte mais aussi des stéréotypes sociaux associés aux diverses formes interrogatives, autant chez les adultes (Chapitre 5) que chez les enfants (Chapitre 6).

4.4 Apports majeurs

1. **L'emploi des interrogatives partielles du français varie en fonction de stratégies discursives, oratoires, ou plus largement interactionnelles, dans un même contexte.**
2. **Les interrogatives in situ sont plus utilisées dans les années 2010 qu'elles ne l'étaient dans les années 1960.**
3. **Des groupes d'âge différents ont des préférences différentes en matière d'emploi des interrogatives partielles.**
4. **Des contextes plus formels emportent une moindre proportion d'emploi des phrases in situ.**
5. **Il faut explorer plus systématiquement ces résultats, via les méthodologies expérimentales.**

Chapitre 5

Nouvelles données expérimentales chez l'adulte francophone

Sommaire

5.1 Démarche : Protocoles expérimentaux retenus	87
5.2 Jugements d'acceptabilité et biais sociolinguistique	87
5.2.1 Préférence pour l'antéposition avec inversion et effet d'âge (AJT1)	88
5.2.2 Lien entre antéposition avec inversion et contexte formel (AJT2)	88
5.2.2.1 Hypothèse et prédiction	89
5.2.2.2 Protocole et items expérimentaux	89
5.2.2.3 Participantes et participants	91
5.2.2.4 Résultats	91
5.2.3 Effets de modalité écrit/oral et notion d'acceptabilité (AJT3)	98
5.2.3.1 Hypothèses et prédictions	99
5.2.3.2 Protocole et items expérimentaux	99
5.2.3.3 Participantes et participants	101
5.2.3.4 Résultats	102
5.2.3.5 Bilan : L'antéposition avec inversion, un <i>standard mal-adapté</i>	112
5.2.4 Le cas des interrogatives en <i>est-ce que</i> (AJT4)	113
5.2.4.1 Hypothèse et prédictions	113
5.2.4.2 Protocole et items expérimentaux	114
5.2.4.3 Participantes et participants	115
5.2.4.4 Résultats	116
5.2.4.5 Bilan : L'antéposition en <i>est-ce que</i> , bien formée mais peu adaptée	126
5.2.5 Jugements d'acceptabilité : Récapitulatif	126
5.3 Matched-Guise et projection sociolinguistique	128
5.3.1 Interrogatives et projection d'indices sociaux à l'écrit (MG1)	128
5.3.2 Interrogatives et projection d'indices sociaux à l'oral (MG2)	129
5.3.2.1 Hypothèses et prédictions	129
5.3.2.2 Protocole et items expérimentaux	130
5.3.2.3 Participantes et participants	132
5.3.2.4 Résultats	133

5.3.2.5	Bilan : Variantes interrogatives et faisceau de stéréotypes	152
5.3.3	Le cas des interrogatives en <i>est-ce que</i> (MG3)	153
5.3.3.1	Hypothèse et prédictions	153
5.3.3.2	Protocole et items expérimentaux	153
5.3.3.3	Participant·es et participants	155
5.3.3.4	Résultats	156
5.3.3.5	Bilan : L'antéposition en <i>est-ce que</i> , des stéréotypes <i>intermédiaires</i>	170
5.3.4	Études de matched-Guise : Récapitulatif	170
5.4	Conclusion intermédiaire	171
5.5	Apports majeurs	173

Ce chapitre présente les résultats de plusieurs études expérimentales menées auprès d'adultes francophones pour dégager leurs préférences en matière d'interrogative partielle et les différentes projections, en particulier sociales, qui sont associées aux principaux types phrastiques disponibles. Après un exposé de la démarche globale dans laquelle s'inscrivent ces études, des expériences de jugements d'acceptabilité et de *matched-guise* sont présentées.

5.1 Démarche : Protocoles expérimentaux retenus

Comme il a été précisé dans le Chapitre 3, le recours aux expériences de jugements d'acceptabilité (AJT) s'est justifié pour plusieurs raisons. C'est d'abord un protocole simple à mettre en place. C'est aussi un protocole adapté pour sonder superficiellement les préférences des adultes francophones en matière de constructions interrogatives (« *Quelle est la meilleure façon de poser une question en français ?* »). Ce type d'expériences a petit à petit été amélioré pour aller plus en détail, notamment au regard d'expériences de *matched-guise* (MG) menées en parallèle. Il est en effet apparu nécessaire d'affiner la notion d'acceptabilité et notamment de distinguer ce qui ressort du *Bon français* d'une part, et d'un français *adapté* (au contexte interactionnel) d'autre part.

Pour rappel également, les tâches de *matched-guise* s'inspirent du modèle de [LAMBERT et al., 1960] et sont un protocole de plus en plus souvent utilisé en sociolinguistique contemporaine (CAMPBELL-KIBLER [2007]; PODESVA et al. [2015], voir Section 2.3.2). Elles consistent à demander à des volontaires de situer sur différentes échelles sociales une personne en ne se basant que sur un énoncé linguistique qu'elle aurait produit. Concrètement, il peut s'agir de déterminer si la personne ayant écrit/prononcé la phrase X est plutôt riche, plutôt de Paris, si elle aime plutôt lire, etc. Ces tâches ont quant à elles appuyé la nécessité d'affiner l'analyse des interrogatives partielles en français sous un angle sociolinguistique, et les critères à explorer sous cette nouvelle perspective, en parallèle des données de corpus précédemment évoquées (Chapitre 4).

Ces deux types de tâches ont été menées à la fois en laboratoire et sur Internet, à la fois sur des stimuli écrits et sur des stimuli audio, pour des raisons qui seront présentées plus bas. Parce que les études de corpus menées au Chapitre 3 ont appuyé l'hypothèse d'un changement diachronique sur les usages interrogatifs en français, les époques de passation des diverses expériences seront systématiquement indiquées. Pour cette thèse, à peine étalée sur quelques années, une évolution diachronique importante n'est pas à envisager, mais ces informations deviendront peut-être pertinentes plus tard.

5.2 Jugements d'acceptabilité et biais sociolinguistique

Quatre tâches de jugements d'acceptabilité ont été menées de 2016 à 2020 et explorent les préférences francophones en matière d'interrogatives.

Ces études ont été menées sur ordinateur, avec des matériels hébergés sur la plateforme Ibex-Farm [DRUMMOND, 2016]. La plupart des volontaires ont participé par Internet mais certaines passations se sont effectuées en laboratoire (pour l'étude AJT3, voir Section 5.2.3). Le recrutement s'est fait majoritairement par l'intermédiaire du Relais d'Information sur les Sciences de la Cognition (RISC, CNRS, UMR 3352) et les réseaux sociaux. Notamment pour les expériences réalisées en laboratoire, il s'agissait aussi de personnes étudiant à l'Université de Paris ou dans les environs.

5.2.1 Préférence pour l'antéposition avec inversion et effet d'âge (AJT1)

Les premières données présentes dans ce chapitre sont le résultat d'un travail réalisé avant cette thèse. Elle est présentée ici comme début et première inspiration du projet de la recherche réalisée. Elle est décrite en détail dans [THIBERGE, 2018]. 57 francophones (âge = 18-72 ans, $\mu = 28$, $m = 33$) ont noté des phrases qui apparaissaient sur leur écran, précédée d'une courte première phrase de contexte. L'échelle de notation était une échelle d'acceptabilité, et allait de 1 (pas du tout acceptable) à 10 (parfaitement acceptable). L'acceptabilité était définie comme le degré auquel la phrase était compréhensible et la justesse de sa formation dans le court contexte fourni. 15 items cibles avaient été mélangés à 30 distracteurs. Il s'agissait de phrases interrogatives précédées d'un court contexte et déclinées sous trois types syntaxiques. Le Tableau 5.1 donne un exemple d'item dans toutes ses conditions.

Contexte	Phrase Cible	Condition
Jean arrive demain.	Il arrive à quelle heure?	In situ
	À quelle heure il arrive?	Antéposition simple
	À quelle heure arrive-t-il?	Antéposition avec inversion

TABLEAU 5.1 – Exemple d'item expérimental (AJT1)

Les phrases interrogatives avec antéposition de l'élément interrogatif et inversion du sujet et du verbe ont reçu, en données brutes et pour les données prises dans leur ensemble, la note moyenne de 9.1/10, contre 8.1 pour les phrases avec l'élément interrogatif in situ et 7.6 pour les phrases avec antéposition sans inversion. Soumises à des analyses de type fréquentiste, les données confirment une préférence significative pour la structure antéposée avec inversion. Même si toutes les variantes sont globalement acceptables, il semble donc y avoir une relative stigmatisation des phrases in situ et avec antéposition simple, du moins en contraste avec les phrases avec antéposition et inversion.

Fait intéressant, un phénomène d'âge apparaît déjà dans cette étude. Les personnes de plus de 30 ans ($N=33$) ont en effet donné une note significativement plus élevée aux interrogatives avec antéposition et inversion qu'aux interrogatives in situ, alors que cette différence n'apparaît pas chez les moins de 30 ans. À l'échelle de toutes les personnes ayant passé l'expérience, cette préférence observée pour les phrases avec antéposition et inversion est ainsi uniquement due à un sous-groupe de l'échantillon.

Ces résultats peuvent se voir comme une confirmation expérimentale de l'existence de différences générationnelles, dégagées dans l'analyse du corpus ESLO [ESHKOL-TARAVELLA et al., 2011; LLL, 2017], et notamment dans sa composante ESLO2 portant sur les années 2000 (Section 4.2.5).

5.2.2 Lien entre antéposition avec inversion et contexte formel (AJT2)

L'étude des corpus EPAC [BAZILLON et al., 2011; ESHKOL-TARAVELLA et al., 2011] et ESLO (Chapitre 4) a par ailleurs fait émerger la nécessité de prendre en compte le contexte dans l'analyse des usages et des préférences francophones en matière d'interrogation partielle. Ici, la notion de contexte doit s'entendre autant en termes linguistiques (« cotexte ») qu'en termes de contexte interactionnel. Les analyses du corpus EPAC ont mis en lumière que, à contexte interactionnel constant, des phases discursives d'*informalisation* se traduisent par une augmentation des productions in situ (Section 4.1.4.1). Les analyses du corpus ESLO ont montré que la proportion d'emploi des

différents types phrastiques varient selon le contexte interactionnel (REPAS, ENTRETIENS, ÉCOLE, Section 4.2.7).

Une nouvelle tâche de jugements d'acceptabilité a donc été conçue pour intégrer ce facteur. Elle a été construite sur le modèle de la précédente (AJT1), mais avec de nouveaux items et en rajoutant une condition de formalité sur la phrase de contexte.

5.2.2.1 Hypothèse et prédiction

En partant des données de corpus préexistantes ou issues des études présentées plus tôt dans ce travail, le contexte semble donc particulièrement influencer la production d'interrogatives par les francophones. Ceci, indépendamment de leur âge (même si dans une moindre mesure pour les 15-25 ans dans le sous-corpus ESLO2). Les interviews semblent plus favoriser l'emploi d'interrogatives avec antéposition et inversion que les interactions à l'école ou les repas par exemple, car la proportion de ces interrogatives par rapport aux trois autres types phrastiques considérés augmente dans ce contexte. Les contextes scolaires semblent aussi favoriser l'emploi de l'antéposition avec inversion, par comparaison avec les autres types phrastiques, par rapport à un contexte de repas familial.

Dans [THIBERGE et al., en révision], qui reprend le détail de ces résultats de corpus, il est proposé que ces trois contextes se distinguent sur plusieurs axes. La personne à qui l'interrogative est adressée n'est pas la même (enfants vs. adultes, scientifique externe au cercle intime vs. personnes proches) d'une part. D'autre part l'environnement où l'interaction se produit n'est pas de la même nature (lieu public vs. maison). La combinaison de ces critères conduit à une sorte de hiérarchisation des interactions, avec les interviews qui paraissent perçues comme les interactions les plus formelles (ou les moins informelles) et les repas comme les plus informelles (ou les moins formelles). Les interactions scolaires seraient entre les deux.

Cette influence possible de la formalité est visible aussi dans le corpus EPAC, avec plus d'interrogatives dans les phases d'informalisation, mais aussi quand une personne cherche à se « rapprocher » de la personne à qui elle parle. Ces indices ont conduit à l'ajout d'une condition de contexte dans les items d'une tâche de jugements d'acceptabilité. En lien avec ces données de corpus, la prédiction était qu'une phrase cible rigoureusement identique serait notée différemment en termes d'acceptabilité, dépendamment du degré de formalité du contexte linguistique qui la précéderait.

Cette formalité pourrait être définie en opposant « informel » à « non-informel » aussi bien qu'en opposant « formel » à « informel ». Le label est moins important que la différence relative de formalité entre les contextes qui seront opposés. Pour cette raison et dans un souci de lisibilité, ce seront par la suite les labels FORMEL et INFORMEL qui seront utilisés, et ce même si les contextes dits « formels » manipulés peuvent parfois plutôt sembler neutres.

5.2.2.2 Protocole et items expérimentaux

30 très courts dialogues ont été construits, entre deux individus A et B. Ces dialogues respectaient toujours la structure montrée en (118).

(118) A : *Courte phrase de contexte, soit informelle soit formelle.*

B : *Interrogative partielle avec antéposition et inversion, antéposition simple, ou in situ.*

Le Tableau 5.2 présente un exemple d'item sous toutes ses conditions. L'intégralité des items construits est visible en Annexe F0.1.

Type	Contexte	
	INFORMEL	FORMEL
Antéposition simple	A : Ouais, moi je me barre demain. B : Où tu vas?	A : Je pars demain. B : Où tu vas?
Antéposition avec inversion	A : Ouais, moi je me barre demain. B : Où vas-tu?	A : Je pars demain. B : Où vas-tu?
In situ	A : Ouais, moi je me barre demain. B : Tu vas où?	A : Je pars demain. B : Tu vas où?

TABLEAU 5.2 – Exemple d'item expérimental (AJT2)

Sous chaque item se trouvait une échelle d'acceptabilité numérique (0-10, de *Pas du tout acceptable* à *Très acceptable*). L'idée d'inclure un 0 était de voir si la présence d'une valeur centrale changeait les résultats (11 crans sur l'échelle contre 10 précédemment), ce qui n'a pas été le cas. Les volontaires pouvaient cliquer sur les crans de l'échelle ou presser une touche de clavier numérique qui correspondait à leur jugement. L'action lançait alors automatiquement la présentation de l'item suivant. Aucune limite de temps n'était imposée, il était seulement indiqué en consigne préalable qu'aucune bonne ou mauvaise réponse n'était attendue. L'acceptabilité était définie, comme dans [THIBERGE, 2018], comme le degré auquel la phrase était compréhensible et la justesse de sa formation dans le court contexte fourni.

Les 30 items cibles ont été présentés selon un carré latin de design 2x3, avec deux niveaux de la variable CONTEXTE (*informel* vs. *formel*) et trois niveaux de la variable TYPE interrogatif (*antéposition simple*, *antéposition avec inversion*, *in situ*). Le degré de formalité des contextes avait préalablement été normé au cours d'une précédente étude (Annexe E).

Chaque personne effectuant la tâche ne voyait un item qu'une seule fois parmi les six combinaisons possibles (Tableau 5.2), tout en voyant chacune des six combinaisons possible un nombre égal de fois au cours de la tâche. Les 30 items-cible étaient mélangés à 30 distracteurs construits sur la même structure (dialogue entre A et B, items tirés de [RIOU et HEMFORTH, 2015], sans interrogation partielle, sans variation de formalité, et portant à la base sur la dislocation en français). Trois items précédaient la tâche pour une familiarisation avec les consignes.

Après une courte page de consigne (Annexe F0.2) puis une seconde de recueil de données personnelles basiques (âge, niveau d'étude, domaine professionnel ou d'études, langues maîtrisées et degré de maîtrise), l'étude commençait. Elle comportait donc 63 items, pour un temps total estimé après étalonnage à environ 15 minutes.

5.2.2.3 Participantes et participants

44 personnes déclarant avoir pour langue maternelle le français ont terminé l'étude. Elles ont été recrutées sur Internet par le biais du RISC et de divers réseaux sociaux en février/mars 2018. Ces personnes avaient entre 19 et 80 ans ($\mu = 39.8$, $m = 27.5$); 30 ont déclaré être de genre féminin, 13 de genre masculin, et une personne n'a pas précisé. Les plus hauts diplômes obtenus par ces personnes étaient :

- en dessous du baccalauréat, N=2
- le baccalauréat, N=7
- un diplôme équivalent à Bac+2, N=9
- un diplôme équivalent à Bac+3, N=5
- un diplôme équivalent à Bac+5, N=16
- un diplôme équivalent à Bac+8, N=2

18 personnes ont indiqué avoir pour région de naissance l'Île de France, 3 un pays étranger (Corée du sud, Maroc, non-précisé), et 23 une autre circonscription administrative du territoire français métropolitain. 4 personnes ont précisé travailler, étudier ou être à la retraite après avoir travaillé dans un domaine touchant aux langues, au langage ou à la linguistique, 5 dans un domaine lié à l'éducation ou l'enseignement, 1 en psycho-pédagogie; le reste des domaines d'activité indiqués couvre l'administration, la santé, l'ingénierie, la mécanique, l'informatique, la vente, la communication, le dessin, le cinéma, l'assurance, et le droit. Les données sur les langues maîtrisées n'ont pas été traitées, à part une vérification de la langue maternelle indiquée. Les personnes n'ont pas été rémunérées pour leur participation.

Au total, 2272 jugements d'acceptabilité ont été reçus, entraînements et distracteurs compris.

5.2.2.4 Résultats

Seuls les 1320 jugements rendus sur les items cibles ont été conservés pour analyse. Après une analyse visuelle des résultats, les notes attribuées et les métadonnées des personnes ayant participé ont été analysées avec des modèles de régression bayésienne.

5.2.2.4.1 Résultats généraux

Le Tableau 5.3 et la Figure 5.1 donnent les moyennes brutes et la distribution des jugements portés sur les trois formes interrogatives analysées. Pour rappel (Chapitre 3), les diagrammes en violon illustrent à la fois la moyenne (en noir) et la distribution des réponses (quartiles en blanc).

Antéposition simple	Antéposition et inversion	In situ
6.23	8.01	6.79

TABLEAU 5.3 – Note moyenne en fonction du type interrogatif (AJT2)

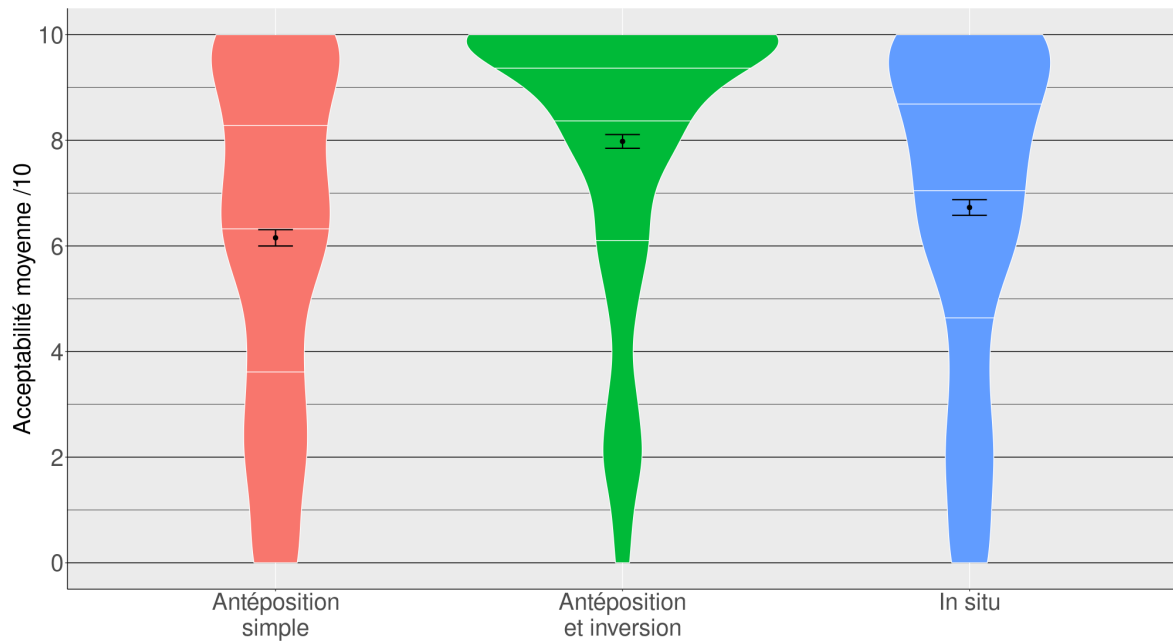


FIGURE 5.1 – Note moyenne en fonction du type interrogatif (AJT2)

Ces résultats illustrent une tendance générale des francophones, déjà observée dans la tâche AJT1 [THIBERGE, 2018], à préférer les interrogatives partielles avec antéposition et inversion aux deux autres types. Comme dans l'étude AJT1, les trois types sont jugés relativement acceptables, mais la distribution des réponses pour l'antéposition avec inversion se démarque de celles pour l'antéposition simple et pour l'in situ.

5.2.2.4.2 Un effet d'âge

Toujours comme dans [THIBERGE, 2018], l'influence de l'âge des francophones ayant participé est aussi observable. Le Tableau 5.4 et la Figure 5.2 donnent les moyennes brutes et la distribution des jugements associés aux différents types phrastiques évalués, en fonction du groupe d'âge des personnes. Le groupe d'âge sépare les moins de 30 ans et les plus de 30 ans, parallèlement au découpage effectué pour l'étude AJT1.

	Antéposition simple	Antéposition et inversion	In situ
-30 ans	6.97	8.02	7.56
+30 ans	5.26	8.00	5.77

TABEAU 5.4 – Note moyenne en fonction du type interrogatif et de l'âge (AJT2)

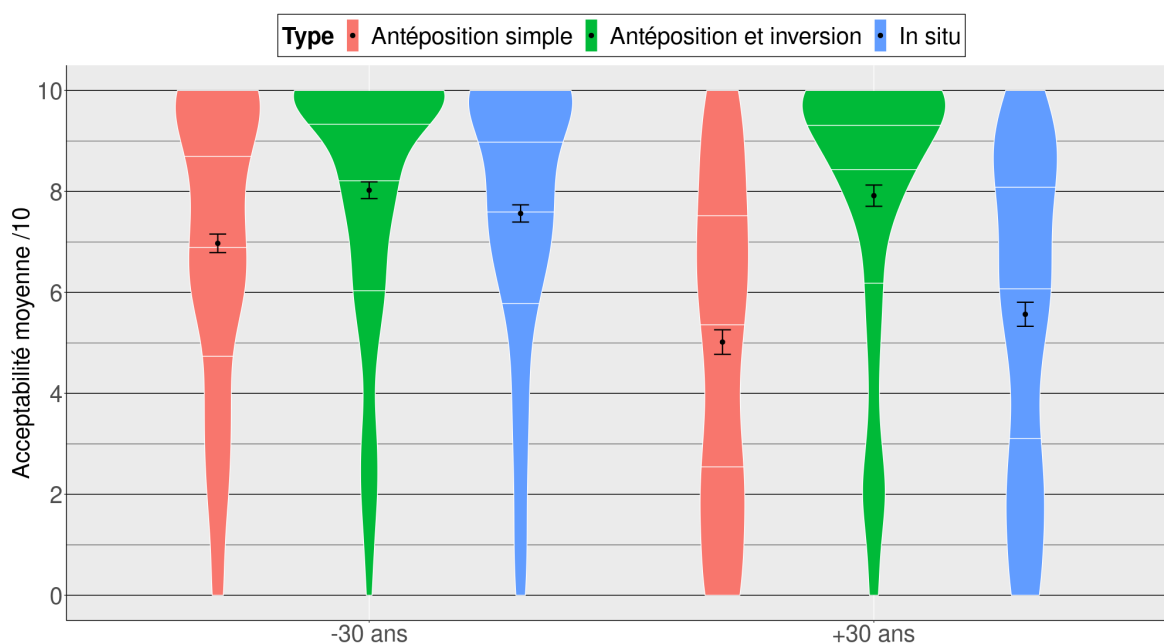


FIGURE 5.2 – Note moyenne en fonction du type interrogatif et de l'âge (AJT2)

Le patron pour chacun des groupes est ici semblable au patron général : l'antéposition avec inversion obtient dans les deux cas une note moyenne supérieure à celle obtenue par les deux autres types phrastiques. Cependant, l'écart entre ces trois types semble plus important chez les plus de 30 ans que chez les moins de 30 ans. En particulier, chez les plus de 30 ans, la note moyenne associée à l'antéposition simple frôle même la moyenne d'acceptabilité (5.26/10). Le jugement des moins de 30 ans pour différencier les types phrastiques semble donc moins définitif que celui des plus de 30 ans, où l'antéposition avec inversion se démarque plus particulièrement comme le type « le plus acceptable ».

5.2.2.4.3 Un effet de contexte

Condition nouvellement introduite dans cette étude, la formalité du contexte semble avoir un effet sur le jugement associé à une interrogative. Le Tableau 5.4 et la Figure 5.3 compilent les moyennes des jugements rendus par les volontaires, en fonction du type syntaxique et de la formalité de la phrase de contexte.

	Antéposition simple	Antéposition et inversion	In situ
Contexte formel	6.26	8.26	6.81
Contexte informel	6.21	7.77	6.76

TABLEAU 5.5 – Note moyenne en fonction du type interrogatif et du contexte (AJT2)

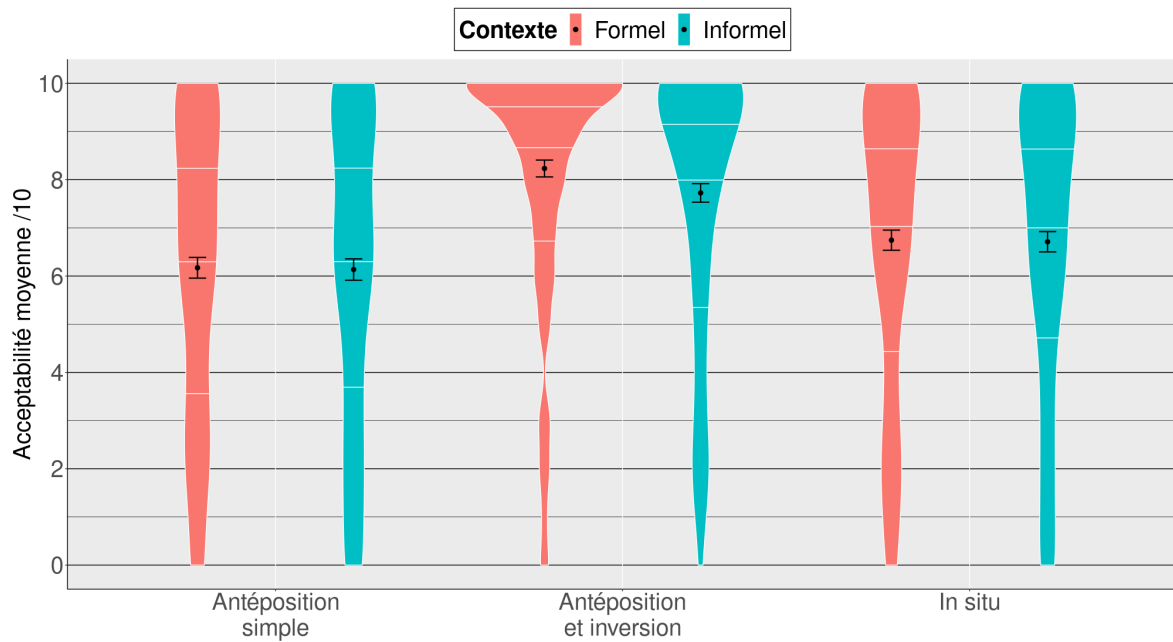


FIGURE 5.3 – Note moyenne en fonction du type interrogatif et du contexte (AJT2)

Ces données indiquent une stabilité globale des jugements rendus sur les interrogatives in situ et avec antéposition simple à travers les contextes. S’il semble y avoir un effet de contexte, il porte principalement sur les interrogatives avec antéposition et inversion, moins bien jugées quand elles s’insèrent dans un contexte informel.

5.2.2.4.4 Une interaction entre âge et contexte

L’analyse peut être encore affinée cependant, en confrontant les données de contexte et celles de groupe d’âge. Le Tableau 5.6 et la Figure 5.4 synthétisent les jugements moyens rendus en fonction du type syntaxique, de la formalité du contexte, en distinguant parmi les volontaires les moins de 30 ans des plus de 30 ans.

	-30 ans		+30 ans	
	Formel	Informel	Formel	Informel
Antéposition simple	7.08	6.86	5.18	5.35
Antéposition et inversion	8.13	7.92	8.42	7.57
In situ	7.54	7.59	5.86	5.67

TABLEAU 5.6 – Note moyenne en fonction du type interrogatif, du contexte, et de l’âge (AJT2)

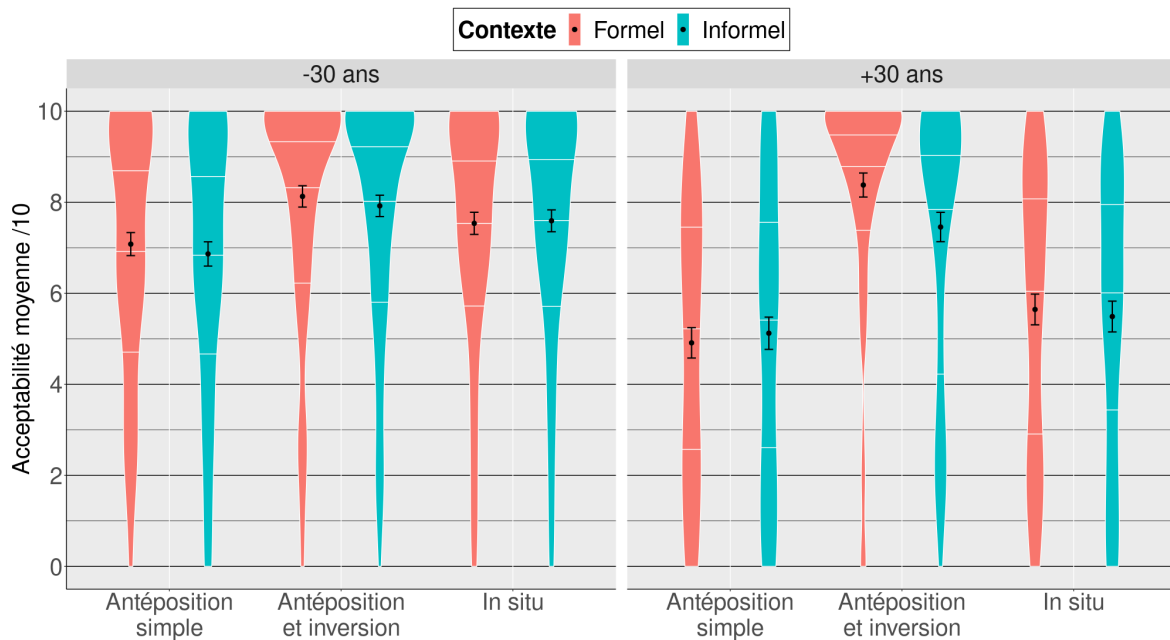


FIGURE 5.4 – Note moyenne en fonction du type interrogatif, du contexte, et de l'âge (AJT2)

Les moins de 30 ans semblent ici aussi se distinguer des plus de 30 ans, puisqu'il ne s'observe pas pour ce groupe de grande différence entre contextes. À l'inverse, pour les plus de 30 ans, la différence de notation des phrases avec antéposition et inversion est amplifiée entre contexte formel et informel. Ce sont donc seulement les personnes qui appartiennent à ce groupe, qui est aussi le groupe le plus sensible à la différence entre types syntaxiques de manière générale, qui ont une appréciation différente selon le contexte.

5.2.2.4.5 Analyses inférentielles et perspectives

Un modèle de régression bayésienne général a été calculé sur l'ensemble de ces données (voir Annexe E.0.3). Le modèle lancé avait pour variable dépendante la note attribuée, et pour variable indépendante l'interaction triple entre trois effets fixes :

- le TYPE syntaxique (antéposition simple, in situ, antéposition avec inversion, avec cette dernière comme niveau de référence),
- le groupe d'ÂGE (moins ou plus de 30 ans, avec le groupe des moins de 30 ans pour référence),
- le CONTEXTE (formel ou informel, avec informel pour niveau de référence).

Des effets avec intercepts aléatoires ont été ajoutés pour tenir compte de la variabilité entre les personnes ayant passé l'expérience ("ID", avec pour pente l'interaction *type*contexte*), et pour tenir compte de la variabilité potentielle entre les items ("Item", avec pour pente l'interaction triple *type*âge*contexte*). La variable dépendante étant ordinaire (échelle de notation aux échelons équidistants), c'est la famille "cumulative" qui a été sélectionnée pour faire tourner ce modèle, avec 4 chaînes et 6000 itérations par chaîne. La convergence du modèle a été atteinte et vérifiée en contrôlant que chaque Rhat soit à 1.00 pour tous les paramètres. La Figure 5.5 rend compte des distributions postérieures de ce modèle pour chacun des paramètres, en interaction ou non.

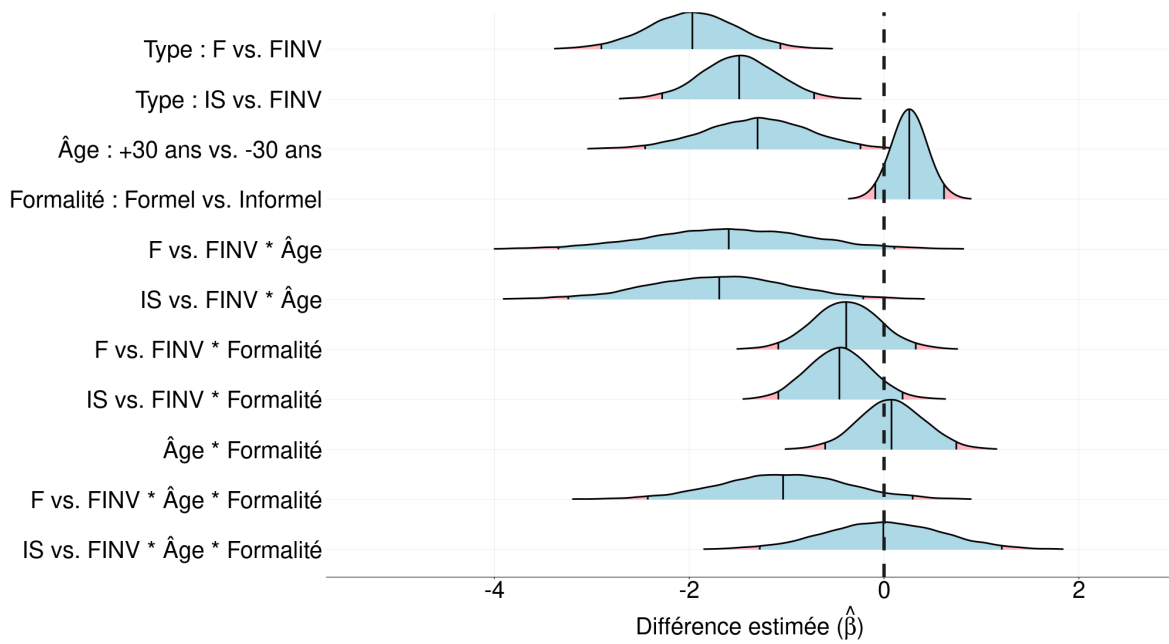


FIGURE 5.5 – Distributions postérieures pour le modèle brms (AJT2)

Ce modèle appuie les observations déjà établies. L'effet principal de TYPE est en effet confirmé : comparées aux antépositions avec inversion, la note donnée est de manière générale plus basse pour les phrases avec antéposition simple ($\hat{\beta} = -1.97$, 95%CrI=[-2.90,-1.06], $P(\beta) < 0=1$) et pour les phrases in situ ($\hat{\beta} = -1.49$, 95%CrI=[-2.28,-0.72], $P(\beta) < 0=1$).

En dehors d'un effet simple d'ÂGE (les plus de 30 ans ont eu tendance à rendre des notes plus basses que les moins de 30 ans), ce prédicteur est impliqué dans des interactions TYPE*ÂGE. Chez les plus de 30 ans, par comparaison au moins de 30 ans, les différences respectives de notation se creusent entre antéposition simple et antéposition avec inversion d'une part ($\hat{\beta} = -1.60$, 95%CrI=[-3.34,0.11], $P(\beta) < 0=0.97$), et entre in situ et antéposition avec inversion d'autre part ($\hat{\beta} = -1.70$, 95%CrI=[-3.24,-0.21], $P(\beta) < 0=0.99$).

Ce phénomène rappelle les observations du corpus ESLO (Section 4.2.5). À l'échelle générale du corpus, une différence de préférences existait en effet entre le groupe des 15-25 ans et le groupe des 35-55 ans, quelle que soit l'époque considérée. Pour ESLO2 et les années 2000, les 15-25 ans avaient un usage plus fréquent des phrases in situ que les 35-55 ans, ce qui pourrait expliquer qu'ici les moins de 30 ans les notent mieux. En revanche, pour ESLO2 toujours, les 35-55 ans faisaient un usage important des interrogatives avec antéposition simple (plus de 25%, contre 16.5 % pour les 15-25 ans). Sur ce point, les jugements portés ici par les plus de 30 ans sont donc discordants. De manière générale, il y a aussi discordance entre la note portée sur les antépositions avec inversion et la fréquence limitée de leur emploi dans ESLO2 (de 7.1 à 8.7%).

Par ailleurs, en dehors d'un effet simple de CONTEXTE (les notes sont plus élevées de manière générale en contexte formel), il y a aussi des interactions TYPE*CONTEXTE dans les jugements rendus expérimentalement. Quand le contexte est formel, par comparaison avec les contextes informels, les différences respectives de notation se creusent entre antéposition simple et antéposition avec inversion d'une part ($\hat{\beta} = -0.39$, 95%CrI=[-1.09,0.33], $P(\beta) < 0=0.87$), et entre in situ et antéposition avec inversion d'autre part ($\hat{\beta} = -0.46$, 95%CrI=[-1.08,0.19], $P(\beta) < 0=0.92$). L'interaction est moindre pour les antépositions simples, mais supporte quand même les observations : puisque les antépositions

avec inversion sont mieux notées en contexte formel, l'écart se creuse avec les deux autres types dont la notation moyenne varie peu.

Ces résultats rappellent ici aussi les observations du corpus ESLO (Section 4.2.7) où l'usage des antépositions avec inversion était plus limité en contexte de repas qu'à l'école ou en entretiens. De même, dans le corpus EPAC (Section 4.5), les phases d'informalisation du discours étaient peu compatibles avec les antépositions avec inversion.

Enfin, une interaction triple TYPE*ÂGE*CONTEXTE est constatée pour les questions avec antéposition simple. Pour les plus de 30 ans, en contexte formel, l'écart de notation entre antéposition simple et antéposition avec inversion se creuse ($\hat{\beta} = -1.04$, 95%CrI=[-2.43,0.29], $P(\beta) < 0.94$). Si cette interaction n'est pas corroborée pour les questions in situ, c'est sans doute dû à la moindre stigmatisation des questions in situ par rapport à celles avec antéposition et inversion.

Ces données-ci sont en tout cas à confronter aux observations du corpus ESLO selon lesquelles les 15-25 ans semblaient avoir une appréciation différente des 35-55 ans de la notion de formalité du contexte (Section 4.2.8.2). La sensibilité à la formalité du contexte qui est cette fois montrée par le groupe des plus de 30 ans dans son appréciation des interrogatives, et pas par le groupe des moins de 30 ans, pourrait y faire écho.

L'importance du contexte social de l'interaction, défini par sa formalité, semble quoi qu'il en soit confirmée par cette étude AJT2. Il est néanmoins nécessaire de rappeler une différence importante entre celle-ci et les données par exemple des corpus ESLO ou EPAC : la modalité. En dehors du côté toujours un peu artificiel de la tâche expérimentale, l'intérêt des corpus ESLO et EPAC réside dans le caractère oral de leurs données, récoltées en situation spontanée. Or, les stimuli soumis au jugement des volontaires étaient ici exclusivement écrits. L'étude suivante (AJT3, Section 5.2.3) est une tentative de généraliser à l'oral les résultats obtenus.

5.2.2.4.6 Une autre vue sur l'effet d'âge

Avant cela, il faut néanmoins remarquer que l'âge a été considéré ici comme un facteur de répartition de l'échantillon des francophones ayant participé à l'expérience, entre les personnes de plus de 30 ans et les personnes de moins de 30 ans. Cette démarche s'inscrit dans la continuité des résultats de [THIBERGE, 2018] et des analyses menées sur le corpus ESLO (Section 4.2.5), qui illustrent l'influence d'un facteur « générationnel » dans la variation. Cependant, l'âge est de fait une variable continue. La Figure 5.6 donne le jugement moyen des individus ayant participé à l'étude par type syntaxique, en fonction de leur âge (abscisse) et du contexte (formel vs. informel).

La partie gauche de cette figure renvoie à d'autres constats de la sociolinguistique sur la norme linguistique. Le phénomène normatif est une considération éminemment dépendante de l'individu. L'appréciation notamment des interrogatives in situ (en bleu) et, dans une moindre mesure, celle des avec antéposition et inversion (en vert), évoluent en effet ici en « fer à cheval » à travers l'échantillon des volontaires. Les personnes entre 40 et 60 ans sont ainsi celles qui jugent le plus sévèrement les interrogatives in situ dans un contexte plutôt formel, et sont dans le même temps celles qui notent les phrases avec antéposition et inversion le plus haut dans ce même type de contexte.

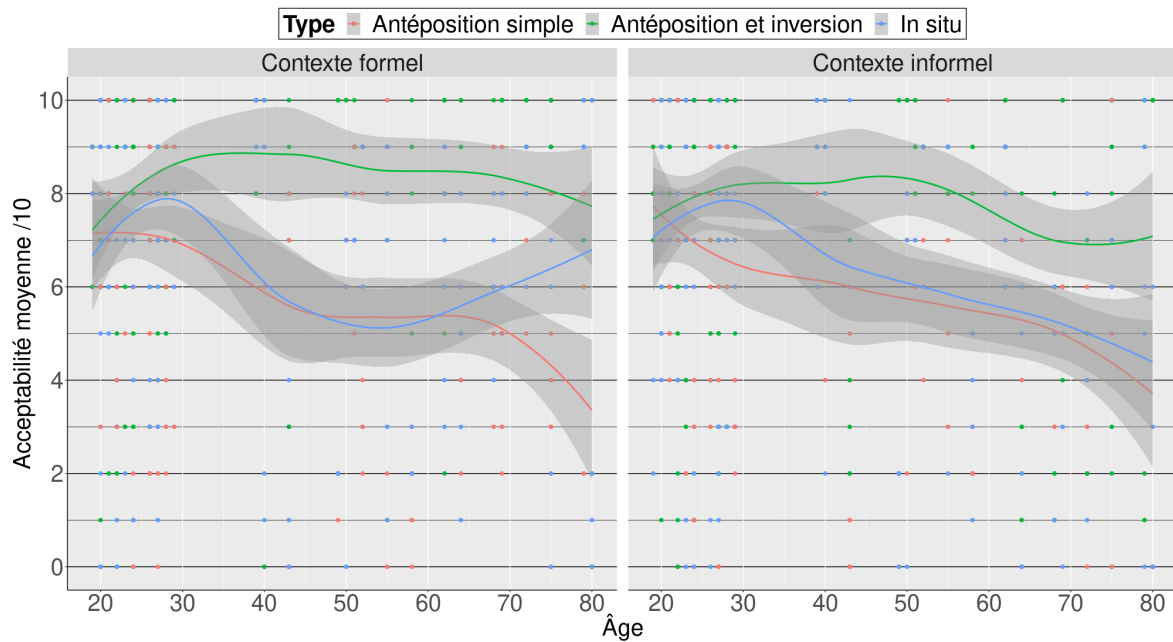


FIGURE 5.6 – Note moyenne en fonction du type interrogatif, du contexte, et de l'âge (AJT2)

Ce type de courbe rappelle la notion d'*age grading* [WAGNER, 2012] et illustre le poids de la norme linguistique, qui fait des interrogatives avec antéposition et inversion la variante « standard » d'interrogation partielle [RIEGEL et al., 2014]. Ce poids de la norme est plus important chez les personnes les plus souvent confrontées à des situations où respecter la norme est socialement important, typiquement via leur travail. Il est notable qu'une telle répartition n'est pas visible pour les interrogatives avec antéposition simple (en rouge), non plus que pour les trois types phrastiques lorsque le contexte est informel, lesquels sont par définition moins concernés par la norme linguistique.

La variable d'âge n'ayant pas été calibrée *a priori* pour une analyse de ce type (via un recrutement ciblé des volontaires par exemple), des statistiques inférentielles n'auraient pas beaucoup de portée. C'est cependant un résultat dont de futurs protocoles expérimentaux pourraient bénéficier.

5.2.3 Effets de modalité écrit/oral et notion d'acceptabilité (AJT3)

Les résultats des études AJT1 et AJT2 (Sections 5.2.1 et 5.2.2), en parallèle d'analyses de corpus (EPAC et ESLO, Chapitre 4), ont montré l'influence de l'âge mais aussi du contexte interactionnel sur le jugement des interrogatives partielles en français. La différence de modalité entre les données des études AJT1/AJT2 (écrites) et les données issues de corpus oraux comme ESLO ou EPAC doit néanmoins être intégrée. En ce sens, il est possible que la modalité orale ou écrite de réception d'une interrogative puisse aussi influencer sur le jugement rendu dessus.

Il est par ailleurs apparu dans les données de l'étude AJT2 une discordance entre les préférences affichée dans les résultats par les francophones et leur emploi réel des différents types interrogatifs, observé sur corpus. La meilleure acceptabilité des phrases avec antéposition et inversion ne semble pas cohérente avec leur emploi marginal de cette construction observé dans les données contemporaines du sous-corpus ESLO2, mais aussi des travaux comme [ADLI, 2015; HAMLAOUI, 2009a]. L'étude AJT3 a été menée pour explorer cette différence entre modalité écrite et orale, et pour en parallèle affiner la notion d'acceptabilité.

5.2.3.1 Hypothèses et prédictions

L'objectif n'était pas de comparer la modalité écrite à la modalité orale en considérant que l'une allait par exemple emporter des jugements généralement plus hauts que l'autre. L'hypothèse de travail était que la modalité orale, plus proche d'une situation d'interaction réelle et quotidienne de deux francophones, pouvait affecter la différence de jugement entre les différentes variantes interrogatives analysées. Notamment, il était envisagé qu'elle interfère avec la notion de formalité ou d'informalité du contexte, laquelle était une condition de l'expérience AJT2.

Le statut relativement privilégié des formes avec antéposition et inversion dans les études AJT1 et AJT2 est à associer à une langue française « standard », qui en fait la « bonne » façon de poser une question en français. Dans ce *standard*, tout au plus la langue « *familière parlée* » peut-elle s'en affranchir pour « *simplifier les structures et [faire que] la phrase garde l'ordre habituel de la phrase déclarative* » [RIEGEL et al., 2014, p. 680]. Il était prévu que s'éloigner de la modalité écrite reflète ceci, et lisse la différence de notation entre formes avec antéposition et inversion d'une part et, d'autre part, formes avec antéposition simple et formes in situ.

Par ailleurs, pour encore affiner cette question de recherche, il n'était plus demandé aux volontaires de juger l'*acceptabilité* des phrases qui leur étaient soumises, mais d'évaluer tour à tour à quel point la personne prononçant la phrase cible : 1°/ *parle un bon français*, et 2°/ *parle un français adapté à la situation*.

Il était attendu que la première échelle enregistre des jugements au patron similaire à l'étude AJT2 : des phrases avec antéposition et inversion mieux jugées qu'avec antéposition simple ou in situ. Il était attendu que la seconde échelle enregistre des résultats plus équilibrés quelle que soit la condition syntaxique de l'interrogative. Une hypothèse plus avancée attendait même que sur cette seconde échelle les résultats soient inversés, surtout dans un contexte informel, avec par exemple un jugement négatif des phrases avec antéposition et inversion. Ceci serait cohérent avec leur faible fréquence d'emploi en corpus.

5.2.3.2 Protocole et items expérimentaux

L'intégralité des items utilisés dans l'étude AJT2 (Annexe F.0.1) a été réutilisée (cibles, distracteurs, contrôles et entraînements). Le Tableau 5.7 rappelle les différentes conditions d'un item.

Type	Contexte	
	INFORMEL	FORMEL
Antéposition simple	A : Elle, elle y sera pas toute seule. B : Qui elle connaîtra?	A : Elle n'y sera pas toute seule. B : Qui elle connaîtra?
Antéposition avec inversion	A : Elle, elle y sera pas toute seule. B : Qui connaîtra-t-elle?	A : Elle n'y sera pas toute seule. B : Qui connaîtra-t-elle?
In situ	A : Elle, elle y sera pas toute seule. B : Elle connaîtra qui?	A : Elle n'y sera pas toute seule. B : Elle connaîtra qui?

TABLEAU 5.7 – Exemple d'item expérimental (AJT2-3)

Les deux lignes de chaque dialogue (A et B) ont été lues et enregistrées par deux femmes (FEM1 et FEM2) et deux hommes (HOM1 et HOM2) francophones. Quatre exemplaires de chaque dialogue ont ainsi été récoltés. Toutes les combinaisons de ces quatre voix ont ensuite été créées à l'aide du logiciel Audacity [AUDACITYTEAM, 1999-2020], avec pour résultats de petits dialogues lus

entre une personne A et une personne B, exactement sur le même principe que dans AJT2. Les 12 combinaisons obtenues correspondent aux successions de voix possibles (Tableau 5.8).

FEM1 puis FEM2	FEM1 puis HOM1	FEM2 puis HOM2
FEM2 puis FEM1	FEM2 puis HOM1	FEM2 puis HOM2
HOM1 puis FEM1	HOM1 puis FEM2	HOM1 puis HOM2
HOM2 puis FEM1	HOM2 puis FEM2	HOM2 puis HOM1

TABLEAU 5.8 – **Combinaisons de voix créées (AJT3)**

Après un norming de la formalité et du genre des voix utilisées (Annexe G), quatre listes d'items ont été créées. Chaque liste comptait 30 items cibles, 30 distracteurs et 3 entraînements. Le même carré latin utilisé pour l'étude AJT2 a été utilisé : un design 3x2 avec une variable CONTEXTE (*formel* vs. *informel*) et une variable TYPE syntaxique (*antéposition simple*, *antéposition avec inversion*, *in situ*). Pour équilibrer les voix dans les dialogues, deux premières listes alternaient des dialogues de structure HOM-HOM ou FEM-FEM en contrebalançant les items, alors que deux autres listes alternaient des dialogues de structure HOM-FEM et FEM-HOM, en contrebalançant également.

L'étude commençait après une courte page de consigne (Annexe H.0.1), puis une seconde de recueil de données personnelles basiques (âge, niveau d'étude, domaine professionnel ou d'études, langues maîtrisées et degré de maîtrise). Elle comportait donc, comme l'étude AJT2, 63 items, pour un temps total estimé après étalonnage à environ 20 minutes.

Sous chaque item se trouvaient deux échelles. La première permettait aux volontaires de juger si la dernière personne s'exprimant parlait un *bon français* (défini comme « grammaticalement correct » dans les consignes). La seconde échelle leur permettait de juger si la personne parlait un français *adapté à la situation* (défini comme « vous pourriez vous exprimer ainsi dans le contexte » dans les consignes). Chaque échelle possédait six niveaux, de "Pas du tout d'accord" à "Totalement d'accord". La Figure 5.7 illustre ce qui était visible sur l'écran.

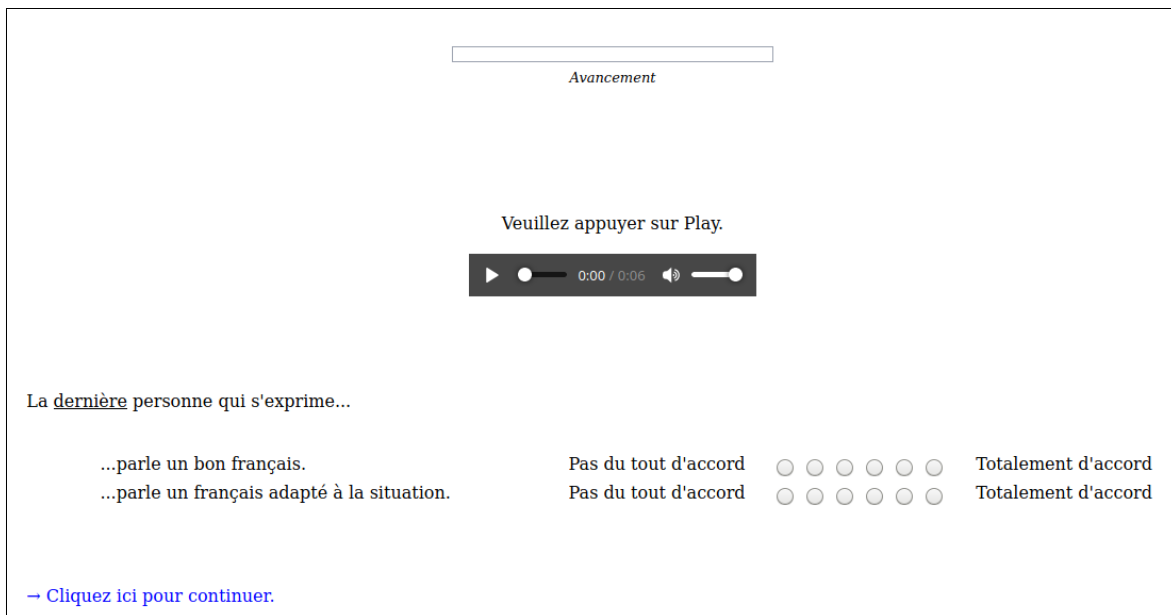


FIGURE 5.7 – **Capture d'écran sur un item expérimental (AJT3)**

Les personnes faisant l'exercice devaient cliquer sur le lecteur audio pour lancer le dialogue,

puis cliquer sur un cran de chacune des deux échelles. Il leur était possible de modifier leurs choix, puis il fallait cliquer sur un lien explicite pour passer au dialogue suivant. Aucune limite de temps n'était imposée. Il était indiqué dans les consignes qu'aucune bonne ou mauvaise réponse n'était attendue. Il était possible de réécouter les extraits audio autant de fois que nécessaire.

5.2.3.3 Participantes et participants

66 personnes ont pris part à l'expérience au printemps 2018, via la plateforme IbexFarm [DRUMMOND, 2016]. Parmi elles, 27 ont été recrutées dans une première phase par le RISC et ont effectué la tâche sans compensation, de chez elles. L'étude étant longue et la contrainte de devoir écouter les sons sont deux raisons pour lesquelles, dans une seconde phase, des passations compensées pécuniairement ont été organisées au laboratoire dans l'enceinte de l'université. Les volontaires provenaient alors des listes de diffusion du RISC ainsi que de l'université même (affichage public) et venaient pour 30 minutes de passation, avec une compensation de 5 euros.

Ces personnes avaient entre 19 et 75 ans ($\mu = 37.9$, $m = 31.0$). 57 ont déclaré être de genre féminin, 8 ont déclaré être de genre masculin, et une personne a choisi de ne pas préciser ce point. Les plus hauts diplômes obtenus par ces personnes étaient :

- en dessous du baccalauréat, N=4
- le baccalauréat, N=14
- un diplôme équivalent à Bac+2, N=9
- un diplôme équivalent à Bac+3, N=15
- un diplôme équivalent à Bac+5, N=22
- un diplôme équivalent à Bac+8, N=2

26 personnes ont indiqué avoir pour région de naissance l'Île de France, 3 un pays étranger (Algérie, Cameroun, Chili), et 35 une autre circonscription administrative du territoire français métropolitain. 7 personnes ont précisé travailler, étudier ou être à la retraite après avoir travaillé dans un domaine touchant aux langues, au langage ou à la linguistique (orthophonie comprise), 7 dans un domaine lié à l'éducation ou l'enseignement, 7 en psychologie ou psychopédagogie. Le reste des domaines d'activité indiqués couvre l'administration, la santé, l'ingénierie, la mécanique, la biochimie, la vente, la communication, le cinéma, la culture, le transport, la restauration, l'industrie, le légal ou para-légal, le management, ou encore le commerce. Les données sur les langues maîtrisées n'ont pas été traitées, à part une vérification de la langue maternelle indiquée.

Au total, 3936 paires de jugements (notes sur les deux échelles prises ensemble) ont été enregistrées, sur 4158 possibles en comptant les distracteurs et les entraînements. Le différentiel s'explique, notamment lors des passations en laboratoire, par des bugs de serveurs informatiques qui ont empêché certains items de passer. Les volontaires pouvaient aussi pour cette expérience ne pas voter sur une des deux échelles, ou sur les deux. L'obligatorité d'un choix n'avait pas été incluse dans la programmation du template expérimental.

Il était rappelé au début de chaque passation effectuée en laboratoire, et avant signature du formulaire de consentement, qu'un retrait de la tâche à tout moment, sans raison et sans perte de contrepartie, était possible.

5.2.3.4 Résultats

Seules les 1855 paires de jugements reçues sur les items cibles ont été conservées pour analyse. Les réponses fournies sur les deux échelles disponibles ont été codées numériquement entre 1 (Pas du tout d'accord) et 6 (Totalement d'accord).

Les résultats seront décrits en trois temps. D'abord seront présentées les données concernant le premier jugement : « *La personne qui s'exprime (...) parle un bon français* ». Ensuite seront présentées les données concernant le second jugement : « *La personne qui s'exprime (...) parle un français adapté à la situation.* ». Enfin, un bilan de ces résultats sera effectué.

Avant cela, il est important de reconnaître une corrélation entre les réponses données sur les deux échelles. Un calcul rapide rapporte un coefficient de corrélation $r = .32$ ($p < 0.01$, méthode = Spearman). Le protocole employé, avec notamment la présentation simultanée des deux échelles sur un même écran, explique sans doute en partie cette corrélation. Celle-ci n'empêche pourtant pas un certain nombre d'observations intéressantes à partir des données.

5.2.3.4.1 Résultats généraux (*bon français*)

Le Tableau 5.9 et la Figure 5.8 donnent les moyennes globales brutes des jugements rendus sur chacune des trois formes interrogatives analysées, pour la première échelle proposée.

Antéposition simple	Antéposition et inversion	In situ
4.35	5.52	4.58

TABEAU 5.9 – Note moyenne en fonction du type interrogatif (AJT3, *bon français*)

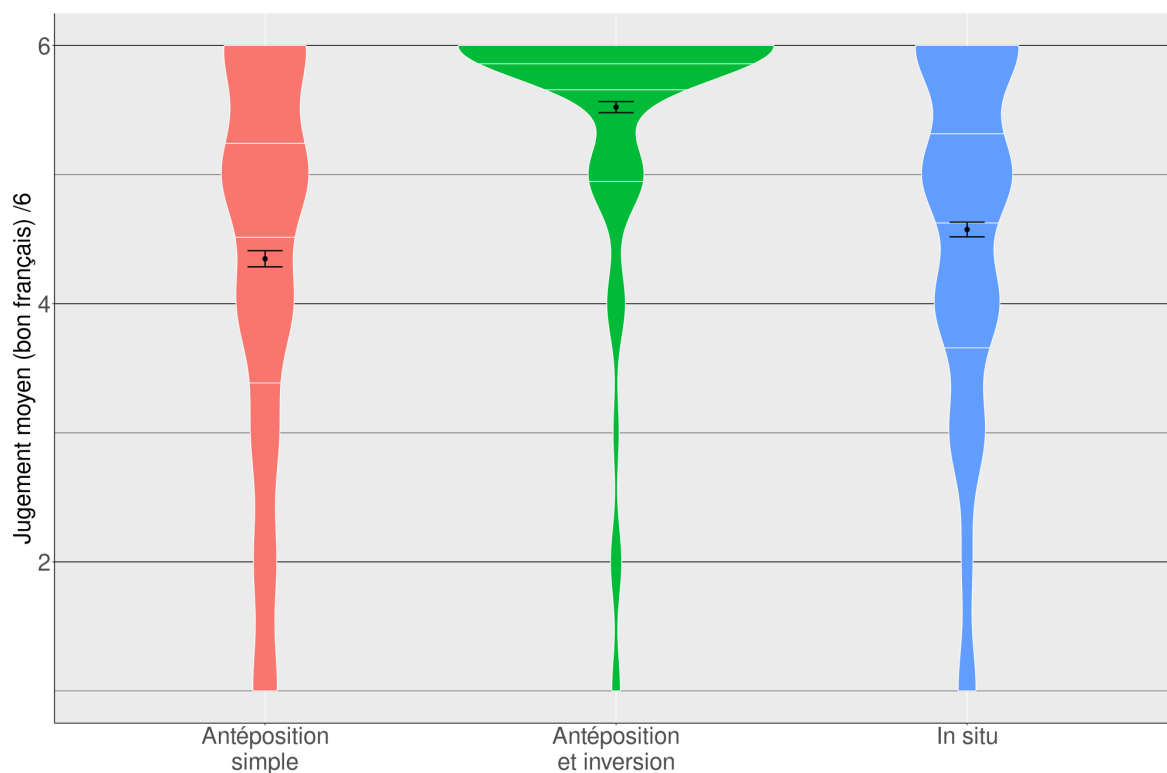


FIGURE 5.8 – Note moyenne en fonction du type interrogatif (AJT3, *bon français*)

Comme dans les études AJT1 et AJT2, la tendance générale qui apparaît est une préférence pour

les phrases avec antéposition et inversion, par opposition aux deux autres types. Ce type interrogatif semble le plus acceptable et, dans la définition donnée de cette notion, le « mieux formé ».

5.2.3.4.2 Un effet d'âge (*bon français*)

Le Tableau 5.10 et la Figure 5.9 donnent les moyennes globales brutes des jugements pour chacune des trois formes interrogatives, pour la première échelle à nouveau. Dans un souci de continuité avec les expériences passées, la séparation en âge a de nouveau été établie au seuil de 30 ans.

	Antéposition simple	Antéposition et inversion	In situ
-30 ans	4.38	5.60	4.53
+30 ans	4.32	5.46	4.62

TABLEAU 5.10 – Note moyenne en fonction du type interrogatif et de l'âge (AJT3, *bon français*)

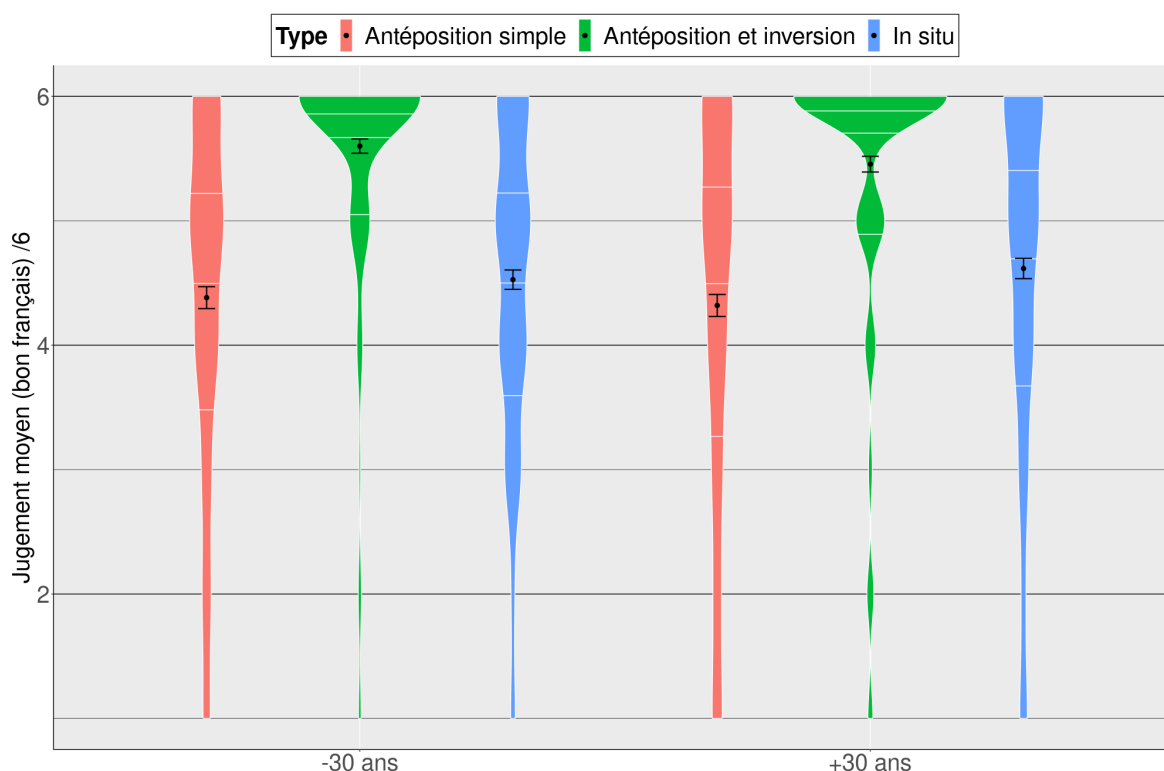


FIGURE 5.9 – Note moyenne en fonction du type interrogatif et de l'âge (AJT3, *bon français*)

Le patron apparaît ici semblable entre les deux groupes, avec toujours cette préférence générale pour les phrases avec antéposition et inversion. Ceci tranche avec les résultats des études AJT1 et AJT2, où le groupe des plus de 30 ans affichait cette préférence mais pas le groupe des moins de 30 ans, chez qui aucune préférence forte pour un type particulier n'apparaissait.

5.2.3.4.3 Un effet de contexte (*bon français*)

Le Tableau 5.11 et la Figure 5.11 donnent les moyennes brutes des jugements pour chacune des trois formes interrogatives, pour la première échelle, en fonction de la formalité du contexte (*formel* vs. *informel*).

	Antéposition simple	Antéposition et inversion	In situ
Contexte formel	4.34	5.63	4.63
Contexte informel	4.35	5.42	4.52

TABLEAU 5.11 – Note moyenne en fonction du type interrogatif et du contexte (AJT3, *bon français*)

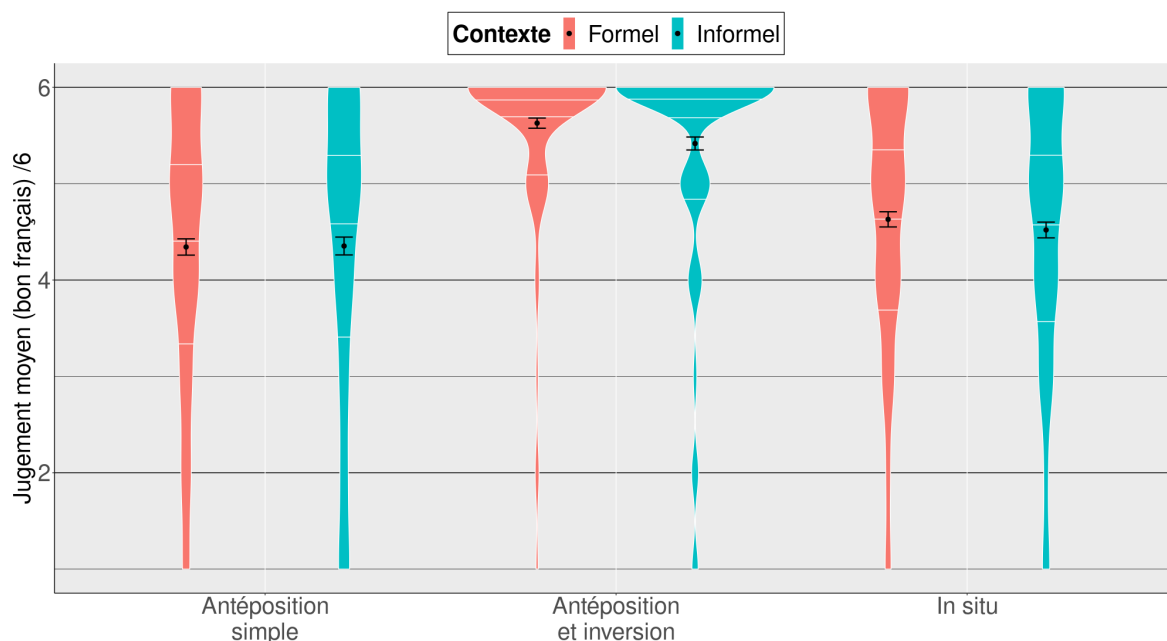


FIGURE 5.10 – Note moyenne en fonction du type interrogatif et du contexte (AJT3, *bon français*)

Les jugements supérieurs sont ici toujours supérieurs quand les interrogatives sont antéposées avec inversion. Il apparaît aussi une légère différence selon que le contexte est informel ou non. Il semble que les interrogatives antéposées avec inversion soient jugées un peu moins bonnes en contexte informel, ce qui rappelle les données de l'étude AJT2.

5.2.3.4.4 Une interaction entre âge et contexte (*bon français*)

Le Tableau 5.12 et la Figure 5.11 synthétisent les moyennes des jugements donnés sur chaque type phrastique en fonction du contexte mais aussi du groupe d'âge.

	-30 ans		+30 ans	
	Formel	Informel	Formel	Informel
Antéposition simple	4.31	4.46	4.38	4.26
Antéposition et inversion	5.67	5.52	5.59	5.33
In situ	4.55	4.51	4.70	4.53

TABLEAU 5.12 – Note moyenne relative aux type interrogatif, contexte et âge (AJT3, *bon français*)

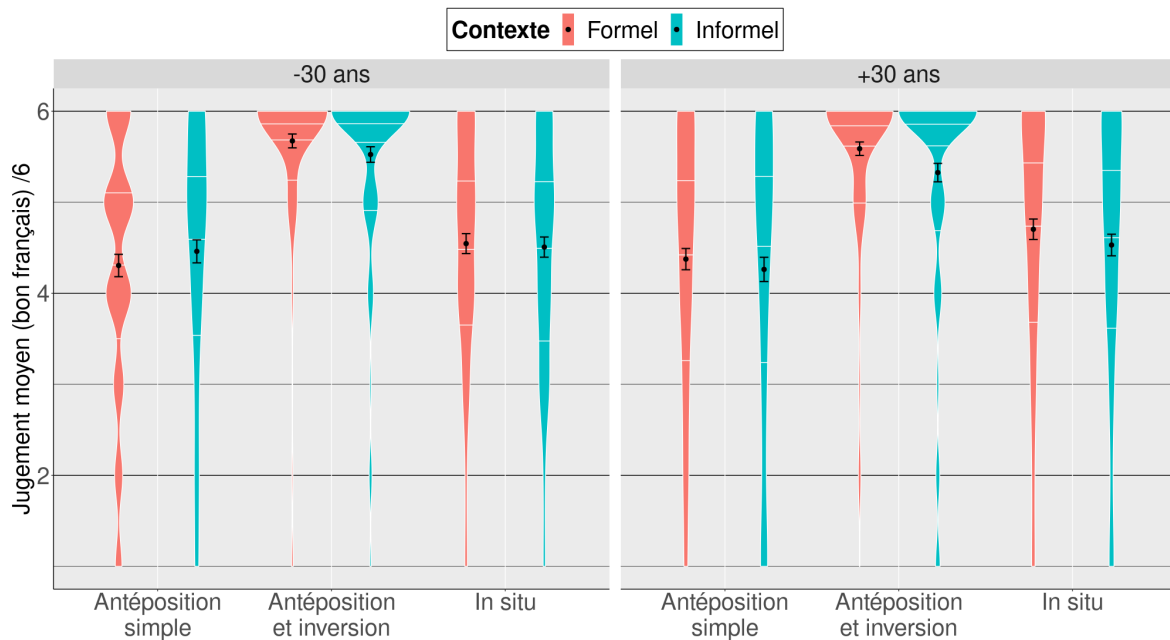


FIGURE 5.11 – Note moyenne relative aux type interrogatif, contexte et âge (AJT3, *bon français*)

Dans un patron toujours similaire à ce qui ressortait de l'étude AJT2, il apparaît que ce sont les plus de 30 ans qui font le plus la différence entre contexte informel et contexte formel, notamment pour les interrogatives avec antéposition et inversion. Cependant cette fois, les moins de 30 ans semblent aussi trouver que les antépositions avec inversion sont un peu moins bonnes lorsque le contexte est informel.

5.2.3.4.5 Analyses bayésiennes (*bon français*)

Un modèle de régression bayésienne a été calculé sur la base des réponses à cette question du bon français (Annexe H.0.2), strictement parallèle au modèle calculé pour l'étude AJT2. Le modèle avait donc pour variable dépendante la note attribuée, et en variable indépendante l'interaction triple entre trois effets fixes : le TYPE syntaxique (*antéposition simple*, *antéposition avec inversion*, *in situ*, avec l'antéposition avec inversion comme niveau de référence), le groupe d'ÂGE (moins ou plus de 30 ans, avec le groupe des moins de 30 ans pour référence), et le CONTEXTE (*formel* ou *informel*, avec informel pour niveau de référence). Les effets aléatoires étaient les individus ("ID"), avec pour pente l'interaction *type*contexte* et les items ("Item"), avec pour pente l'interaction triple *type*âge*contexte*. La variable dépendante étant ordinale (échelle de notation aux échelons équidistants), c'est la famille "cumulative" qui a été sélectionnée pour faire tourner ces modèles, avec 4 chaînes et 6000 itérations par chaîne. La convergence des modèles a été atteinte et vérifiée (Rhats = 1.00 pour tous les paramètres).

La Figure 5.12 illustre les distributions postérieures de ce modèle pour chacun des paramètres, en interaction ou non, et confirme la valeur généralisable des principaux résultats observés plus haut. Pour cette question du « bon français », il y a une certaine correspondance entre le patron général de ces résultats et celui de l'étude AJT2.

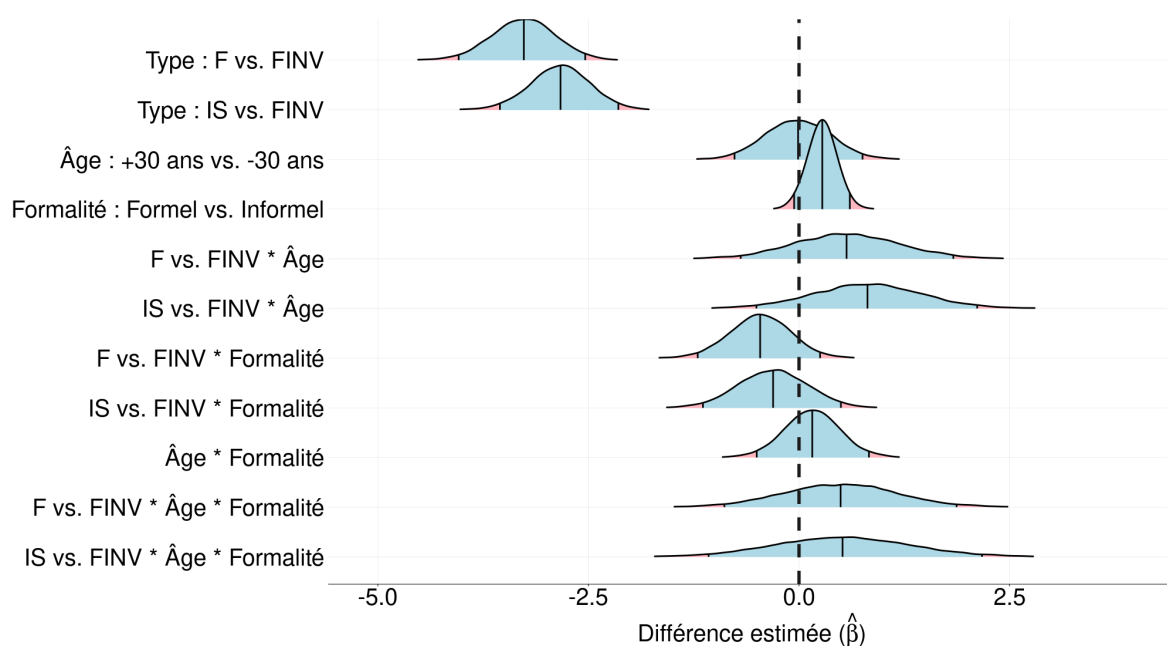


FIGURE 5.12 – Distributions postérieures pour le modèle brms (AJT3, bon français)

L'effet simple du TYPE interrogatif est ainsi toujours présent. Comparées aux antépositions avec inversion, la note donnée est de manière générale plus basse pour les phrases avec antéposition simple ($\hat{\beta} = -3.27$, 95%CrI=[-4.04,-2.54], $P(\beta) < 0 = 1$) comme pour les phrases in situ ($\hat{\beta} = -2.84$, 95%CrI=[-3.55,-2.14], $P(\beta) < 0 = 1$).

L'effet d'ÂGE constaté dans l'étude AJT2 ne peut cependant ici pas être confirmé. Comme il était visible sur les visualisations brutes, les moins de 30 ans ont ici un comportement général similaire aux plus de 30 ans.

En parallèle, le même effet simple de CONTEXTE est constaté que dans l'étude AJT2 (les notes étaient plus élevées de manière générale en contexte formel). Il y a par ailleurs aussi des indices d'interactions négatives TYPE*CONTEXTE, même si elles sont moins fortes voire seulement tendancielles. Quand le contexte est formel, par comparaison avec les contextes informels, les différences respectives de notation se creusent entre antéposition simple et antéposition avec inversion d'une part ($\hat{\beta} = -0.46$, 95%CrI=[-1.20,0.25], $P(\beta) < 0 = 0.90$), et entre in situ et antéposition avec inversion d'autre part ($\hat{\beta} = -0.31$, 95%CrI=[-1.14,0.50], $P(\beta) < 0 = 0.77$). Là encore, ceci est dû aux plus hauts jugements qui sont rendus en contexte formels sur les antépositions avec inversion.

Point notable, les deux interactions TYPE*ÂGE constatées dans l'étude AJT2 semblent ici renversées, et moins fortes. Il semble ainsi que la différence entre les phrases avec antéposition simple et celles avec antéposition et inversion se réduise ici dans le groupe des plus de 30 ans par rapport à celui des moins de 30 ans ($\hat{\beta} = 0.57$, 95%CrI=[-0.69,1.83], $P(\beta) > 0 = 0.81$). Il en est de même pour la différence entre les phrases in situ et celles avec antéposition et inversion ($\hat{\beta} = 0.81$, 95%CrI=[-0.51,2.12], $P(\beta) > 0 = 0.89$). La raison en est sans doute que les moins de 30 ans ont en moyenne, sur ces études, rendu des jugements plus hauts pour les interrogatives avec antéposition et inversion que les plus de 30 ans.

Pour ce qui est des interactions triple TYPE*ÂGE*CONTEXTE, il faut mentionner les tendances qu'elles évoquent. Il semble que pour les plus de 30 ans, en contexte formel, l'écart de notation entre antéposition simple et antéposition avec inversion tende à se réduire par rapport aux moins

de 30 ans ($\hat{\beta} = 0.49$, 95%CrI=[-0.89,1.87], $P(\beta)>0=0.76$). Il semble en aller de même pour l'écart entre phrases in situ et phrases avec antéposition et inversion ($\hat{\beta} = 0.52$, 95%CrI=[-1.07,2.17], $P(\beta)>0=0.74$). De nouveau, ceci s'explique par le jugement moyen brut plus élevé que les moins de 30 ans ont rendu en contexte formel pour les antépositions avec inversion. En dehors de son sens, positif ici et négatif pour l'étude AJT2, il est intéressant de constater que cette différence de préférences entre groupe d'âge, en fonction du contexte, est toujours active.

5.2.3.4.6 Résultats généraux (français adapté)

Si les jugements rendus sur le premier point du « bon français » vont dans un sens similaire à ceux rendus dans l'étude AJT2, il en va autrement pour ceux rendus sur le second point : « *La personne qui s'exprime (...) parle un français adapté à la situation* ».

Le Tableau 5.13 et la Figure 5.13 donnent les moyennes globales brutes des jugements associés aux différents types interrogatifs par les volontaires.

Antéposition simple	Antéposition et inversion	In situ
4.91	4.87	5.10

TABLEAU 5.13 – Note moyenne en fonction du type interrogatif (AJT3, français adapté)

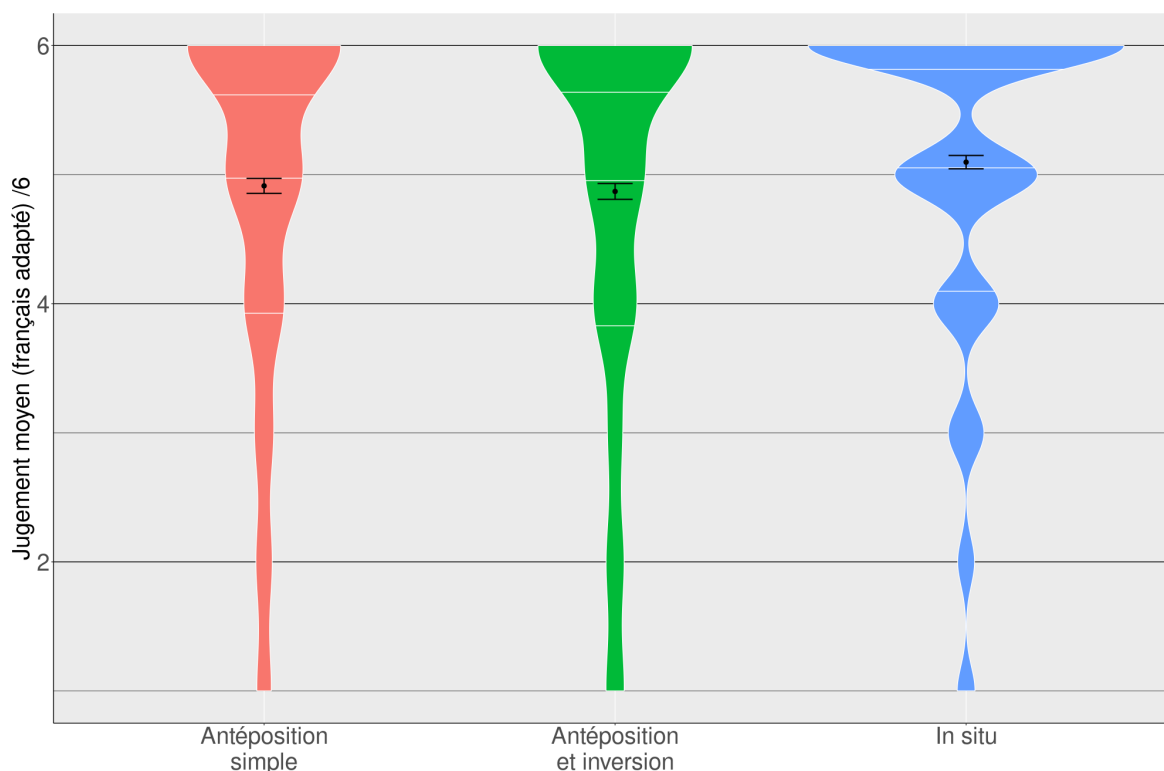


FIGURE 5.13 – Note moyenne en fonction du type interrogatif (AJT3, français adapté)

Trois faits sont ici d'importance. Premièrement, les jugements moyens rendus sont tous relativement hauts, en comparaison avec ceux rendus sur la question du « bon français ». L'écart moyen entre le type phrastique ayant reçu les jugements les plus hauts et ceux ayant reçu les jugements les plus bas est ici seulement de 0.23, contre 1.17 pour la question du « bon français ». Il ne semble pas y avoir de type interrogatif jugé vraiment « inadapté » dans l'interaction. Deuxièmement, et en conséquence, l'opposition forte s'efface entre les phrases avec antéposition et inversion d'une part,

et les phrases avec antéposition simple ou in situ d'autre part. Troisièmement, si cette première opposition s'efface, une autre apparaît et le type in situ semble ici préféré aux deux autres.

La première observation pourrait appuyer une relative indifférence des francophones quant aux formes interrogatives utilisées en conversation réelle. La troisième observation va, elle, dans un sens cohérent avec un emploi qui serait devenu majoritaire des phrases in situ. Celui-ci a été observé par exemple dans le corpus ESLO2 (près de 60% de phrases in situ), mais aussi déjà par [HAMLAOUI \[2009a\]](#) ou [ADLI \[2015\]](#) avec également des proportions d'emploi voisines de 60%.

La confrontation de ce résultat général à celui observé pour la première question sur le « bon français » appuie empiriquement l'idée d'un biais et d'une discordance chez les francophones, entre prescription normative et usage réel en situation ordinaire. Pour produire une interrogative, le français « bien construit », celui présenté comme *standard* par les grammaires de référence, demanderait une antéposition avec inversion. Cependant cette forme de phrase n'est pas celle qui est généralement la plus adaptée à la situation. Le type phrastique qui est généralement le plus adapté est la phrase in situ. Ceci, alors que la modalité de cette expérience est orale, est cohérent avec le fait qu'il se trouve plus de phrases in situ dans les corpus oraux (ESLO, EPAC) que dans d'autres études sur des corpus écrits (voir [KAISER et QUAGLIA \[2015\]](#) sur une analyse de romans avec 70% d'antépositions, Chapitre 2, Section 2.1.2.2).

5.2.3.4.7 Un effet d'âge (*français adapté*)

Le Tableau 5.14 et la Figure 5.14 donnent les moyennes brutes des jugements rendus sur les types phrastiques en fonction du groupe d'âge (plus ou moins de 30 ans) des volontaires.

	Antéposition simple	Antéposition et inversion	In situ
-30 ans	5.00	4.82	5.14
+30 ans	4.83	4.92	5.06

TABLEAU 5.14 – Note moyenne en fonction du type interrogatif et de l'âge (AJT3, *français adapté*)

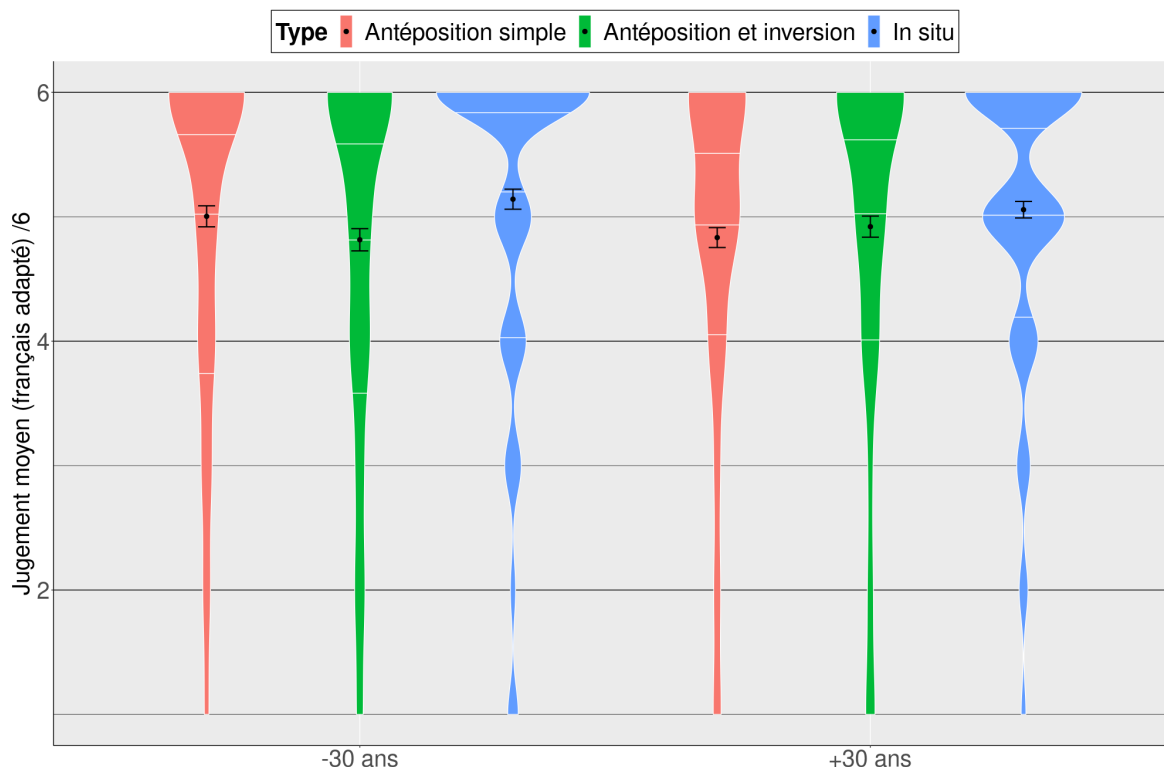


FIGURE 5.14 – Note moyenne en fonction du type interrogatif et de l'âge (AJT3, *français adapté*)

Les patrons sont ici différents d'un groupe à l'autre. Les moins de 30 ans semblent estimer que les interrogatives avec inversion sont de manière générale moins adaptées que les deux autres types. Chez les plus de 30 ans, en contraste, un gradient semble apparaître, avec les interrogatives in situ qui seraient plus « adaptées » que les interrogatives avec antéposition et inversion, elles-mêmes plus adaptées que les interrogatives avec antéposition simple.

En fait, l'écart entre antéposition avec inversion et type in situ est réduit chez les plus de 30 ans, par rapport à ce qu'il est chez les moins de 30 ans. Ceci conduit à considérer que, en gardant l'antéposition avec inversion comme référence, les phrases in situ sont proportionnellement moins adaptées chez les plus de 30 ans qu'elles ne le sont chez les moins de 30 ans. De même, alors que les antépositions simples sont plus adaptées que celles avec inversion chez les moins de 30 ans, elles sont moins adaptées chez les plus de 30 ans.

5.2.3.4.8 Un effet de contexte (*français adapté*)

Le Tableau 5.15 et la Figure 5.15 permettent de visualiser les jugements moyens rendus sur chaque type interrogatif en fonction du contexte, formel ou informel, qui précédait chaque phrase.

	Antéposition simple	Antéposition et inversion	In situ
Contexte formel	4.90	4.99	5.17
Contexte informel	4.93	4.75	5.02

TABLEAU 5.15 – Note moyenne en fonction du type interrogatif et du contexte (AJT3, français adapté)

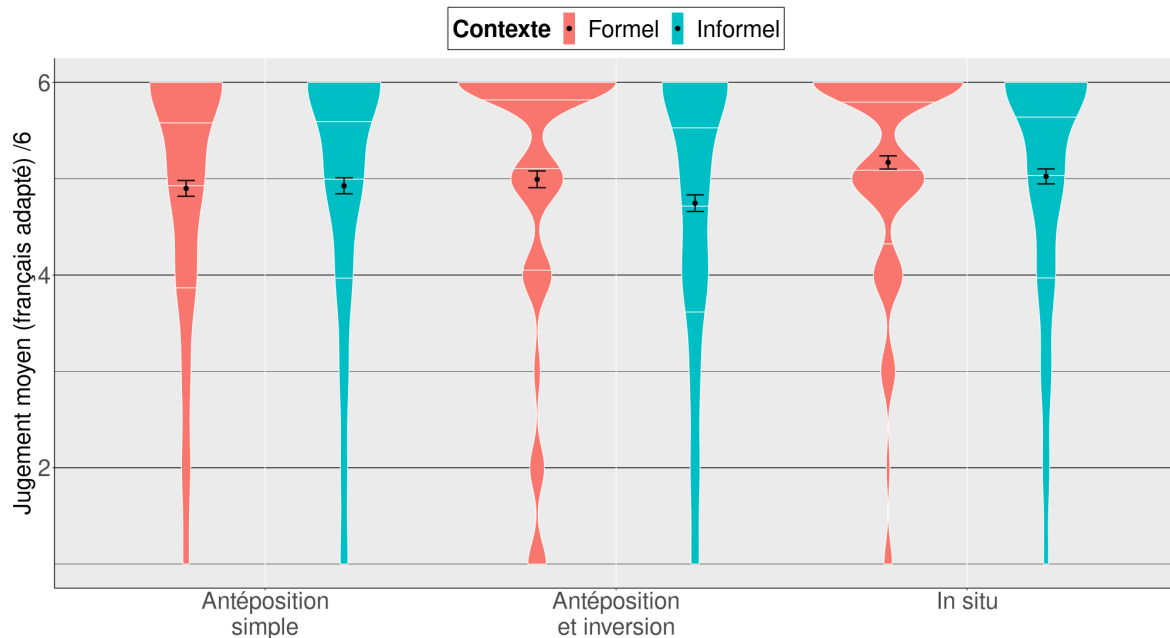


FIGURE 5.15 – Note moyenne en fonction du type interrogatif et du contexte (AJT3, français adapté)

Il se trouve ici un effet de contexte en ce sens que les phrases avec antéposition et inversion, comme les phrases in situ, donnent lieu à des jugements plus hauts en contexte formel. Cette donnée peut se lire avec deux orientations. Premièrement, il se peut que l'antéposition avec inversion soit jugée plus adaptée aux contextes formels qu'aux contextes informels. À l'inverse, elle est peut-être jugée moins adaptée aux contextes informels. Dans les deux interprétations, ce fait est cohérent avec les observations du corpus EPAC, où les phases d'informalisation du discours restreignent l'emploi des interrogatives avec antéposition et inversion (Section 4.5). À l'inverse, la différence d'adaptation au contexte est surprenante pour les phrases in situ. Dans le corpus EPAC les phrases in situ apparaissent en effet uniquement dans ces phases d'informalisation.

5.2.3.4.9 Une interaction entre âge et contexte (français adapté)

Le Tableau 5.16 et la Figure 5.16 permettent une visualisation des jugements moyens rendus par les volontaires, en fonction de leur âge mais aussi des variations de contexte.

	-30 ans		+30 ans	
	Formel	Informel	Formel	Informel
Antéposition simple	5.01	4.99	4.80	4.87
Antéposition et inversion	4.92	4.71	5.06	4.78
In situ	5.18	5.11	5.16	4.95

TABLEAU 5.16 – Note moyenne relative aux type interrogatif, contexte et âge (AJT3, français adapté)

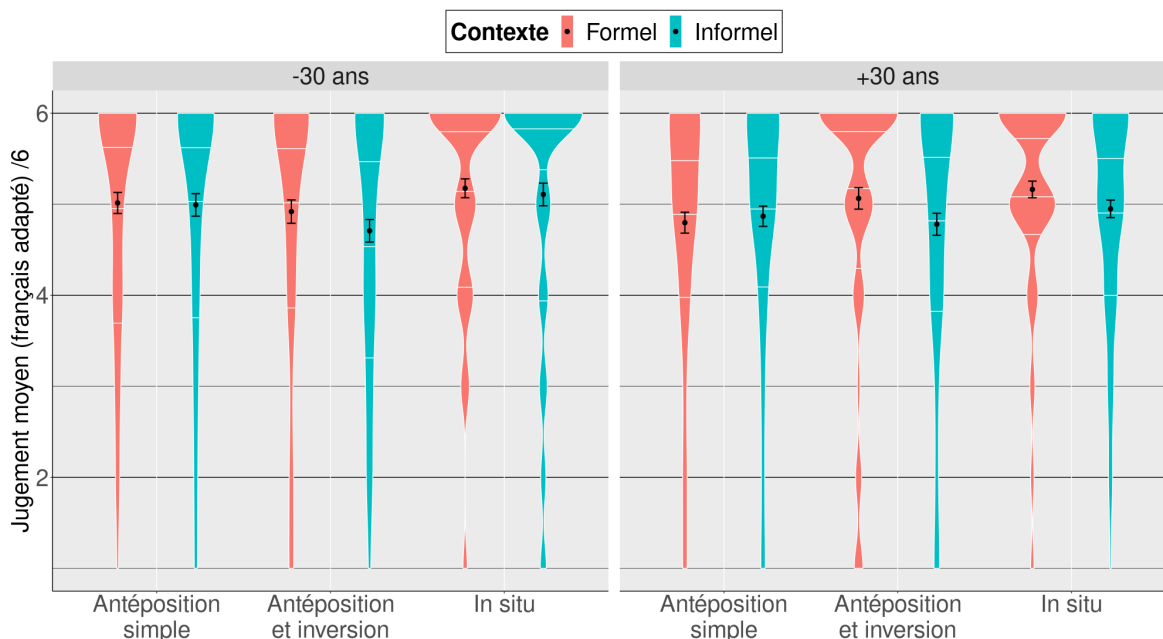


FIGURE 5.16 – Note moyenne relative aux type interrogatif, contexte et âge (AJT3, français adapté)

Une sensibilité plus générale des plus de 30 ans aux variations de contexte est encore une fois apparente. Ce groupe semble trouver moins heureuses, dans les contextes informels, aussi bien les interrogatives avec antéposition et inversion que les interrogatives in situ. En parallèle, chez les moins de 30 ans, une différence liée au contexte ne semble perceptible que pour les antépositions avec inversion.

5.2.3.4.10 Analyses bayésiennes (français adapté)

Un modèle en tout point similaire à celui calculé pour le « bon français » a été calculé sur la question du français « adapté ». L'Annexe H.0.3 en compile les détails, et la Figure 5.17 en illustre les distributions postérieures.

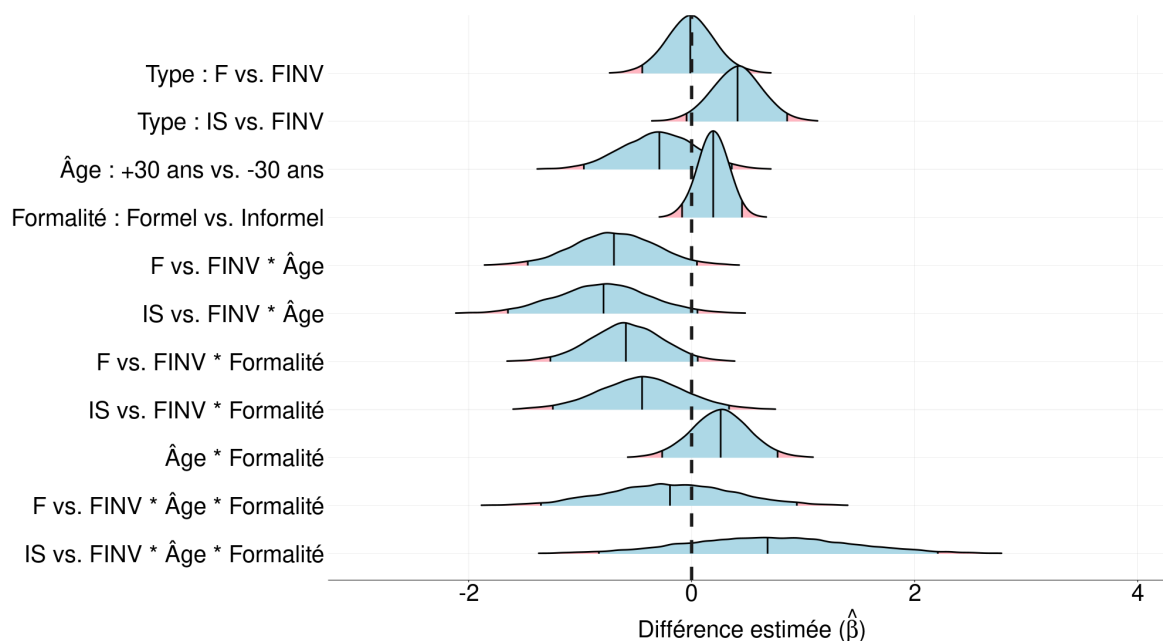


FIGURE 5.17 – Distributions postérieures pour le modèle brms (AJT3, français adapté)

Le patron général diffère de celui pour la première échelle en plusieurs points.

D'abord, l'effet simple de TYPE n'est plus le même. Comme il était visible à partir des données brutes, la différence entre interrogatives avec antéposition simple et celles avec antéposition et inversion a disparu. Au contraire, celle entre phrases in situ et phrases avec antéposition et inversion est désormais positive ($\hat{\beta} = 0.41$, 95%CrI=[-0.05,0.86], $P(\beta) > 0 = 0.96$). Elle traduit le fait que les interrogatives in situ sont bien liées à des jugements plus hauts en termes de « bonne adaptation » au contexte.

Par ailleurs, si aucun effet simple d'ÂGE n'apparaît, les interactions TYPE*ÂGE sont elles marquées, avec un renversement par rapport à l'échelle du « bon français ». Elles confirment que, par rapport au groupe des moins de 30 ans, le groupe des plus de 30 ans fait moins de différence, en termes d'adaptation au contexte, entre antéposition simple et antéposition avec inversion d'une part ($\hat{\beta} = -0.70$, 95%CrI=[-1.47,0.05], $P(\beta) < 0 = 0.97$), mais aussi d'autre part entre in situ et antéposition avec inversion ($\hat{\beta} = -0.80$, 95%CrI=[-1.65,0.05], $P(\beta) < 0 = 0.97$).

Pour ce qui est du CONTEXTE, un effet simple est présent (les interrogatives en contexte formel sont de manière générale liées à des jugements plus hauts sur l'adaptation au contexte ($\hat{\beta} = 0.19$, 95%CrI=[-0.09,0.45], $P(\beta) > 0 = 0.92$). Ceci est dû aux meilleurs jugements associés aux interrogatives avec antéposition et inversion, ainsi qu'aux phrases in situ, en contexte formel. Dans le même ordre d'idée, des interactions TYPE*CONTEXTE sont observées. En contexte formel, l'écart entre les jugements rendus sur des antépositions simples et ceux rendus des antépositions avec inversion se réduit ($\hat{\beta} = -0.59$, 95%CrI=[-1.27,0.05], $P(\beta) < 0 = 0.96$). Ceci est lié au fait que les interrogatives avec antéposition et inversion sont associées à de moins bon jugements que quand l'antéposition est simple, en contexte informel. Il en va de même pour la comparaison des jugements associés à des phrases in situ et ceux associés à des phrases avec antéposition et inversion ($\hat{\beta} = -0.45$, 95%CrI=[-1.24,0.34], $P(\beta) < 0 = 0.88$). Ici encore, ceci est le reflet des jugements moyens plus bas associés aux antépositions avec inversion en contexte informel. Les phrases avec antéposition et inversion sont, de fait, généralement moins bien adaptées aux contextes informels.

En outre, l'interaction positive ÂGE*CONTEXTE est ici plus forte qu'elle ne l'était pour la première question sur le « bon français », et pour l'étude AJT2. En contexte formel, les plus de 30 ans ont tendance à rendre des jugements plus élevés que les moins de 30 ans ($\hat{\beta} = 0.26$, 95%CrI=[-0.26,0.77], $P(\beta) > 0 = 0.84$). L'interaction triple TYPE*ÂGE*CONTEXTE pour les interrogatives in situ ($\hat{\beta} = 0.68$, 95%CrI=[-0.83,2.21], $P(\beta) > 0 = 0.81$) y est liée. Elle est indicatrice de cette tendance observée, chez les plus de 30 ans par opposition, à faire la distinction entre contextes aussi bien pour les antépositions avec inversion que pour les phrases in situ. Là où les moins de 30 ans trouvent les phrases in situ aussi bien adaptées aux contextes informels qu'aux contextes formels, les plus de 30 ans trouvent ces phrases plus adaptées en contexte formel.

5.2.3.5 Bilan : L'antéposition avec inversion, un *standard mal-adapté*

Deux hypothèses guidaient cette tâche de jugements d'acceptabilité. D'abord, il était attendu que la modalité orale influencerait sur l'écart de notation entre la structure préférée dans les études AJT1 et AJT2, l'antéposition avec inversion, et les deux autres types considérés. Cet écart devait être lissé par une modalité de soumission des items plus spontanée. En parallèle, la prévision était que la double échelle permettrait une vision affinée de l'*acceptabilité* des différents types interrogatifs en français.

Pour la première hypothèse, il semble que la modalité orale de soumission des items n'a pas lissé, dans l'absolu, les différences de préférence entre types syntaxiques. Notamment, la préférence générale pour l'antéposition avec inversion n'a pas disparu en ce qui concerne le « bon français ». Ici, les deux échelles ont bien fonctionné et la notion d'acceptabilité a pu être explorée. En ceci la seconde hypothèse est confirmée. Bien que leurs réponses aient été en partie corrélées, les francophones font bien la différence entre des critères de « bonne formation » de la phrase, et le fait qu'une phrase soit adaptée ou non à une situation d'interaction.

La première échelle a donné des résultats conformes à ce qui était attendu, avec un patron de résultats similaire à celui observé dans les études AJT1 et AJT2 : l'antéposition avec inversion ressort d'un *meilleur* français que l'antéposition simple ou l'in situ. La seconde échelle a quant à elle confirmé que malgré cela, malgré ce « standard » intégré, les francophones ont conscience de la réalité de leurs usages interrogatifs. Notamment, les phrases in situ, si elles ne sont pas du *meilleur* français, sont le type phrastique le plus généralement « adapté » aux situations d'interaction. Ceci est cohérent avec les observations de corpus contemporaines, qu'il s'agisse de travaux comme ADLI [2015] ou HAMLAOUI [2009a] ou bien de l'analyse présentée au Chapitre 4.

En plus de ces faits généraux, une influence du contexte est confirmée, avec des jugements différents concernant l'adaptation à la situation selon que le contexte est formel ou non (notamment chez les plus de 30 ans). Par ailleurs, il semble encore un peu plus confirmé qu'un facteur générationnel est à prendre en compte dans l'analyse des usages interrogatifs en français. Si l'existence d'un standard linguistique (l'antéposition avec inversion) est intégrée quel que soit le groupe d'âge, les moins de 30 ans considèrent l'antéposition avec inversion comme généralement moins heureuse que les autres types interrogatifs. C'est notamment le cas en contexte informel.

5.2.4 Le cas des interrogatives en *est-ce que* (AJT4)

Les études AJT2 et AJT3, bâties sur le modèle de l'étude AJT1, n'ont comparé les préférences francophones que pour trois types interrogatifs : l'antéposition avec inversion, l'antéposition simple, et le type in situ. Les études de corpus menées en parallèle, et présentées dans le Chapitre 4, ont imposé l'étude expérimentale d'un quatrième type phrastique couramment employé en français : l'antéposition en *est-ce que*. C'est ce que permet l'étude AJT4.

5.2.4.1 Hypothèse et prédictions

Dans le Chapitre 4, le corpus EPAC [BAZILLON et al., 2011; ESTÈVE et al., 2010] indiquait une proportion d'emploi des phrases avec antéposition en *est-ce que* de 11.4%. Sur l'intégralité du corpus ESLO [ESHKOL-TARAVELLA et al., 2011; LLL, 2017], cette proportion s'élève à 12.6%, et à 9.4% pour le sous-corpus ESLO2 centré sur les années 2000. Les proportions données par HAMLAOUI [2009a] et ADLI [2015] sont d'un ordre de grandeur comparable, avec respectivement 11% et 16.7%. Les chiffres plus anciens rapportés par COVENEY [2011] sont un peu plus variables (3, 12 et 8% pour respectivement les corpus de RADIO, BOURGEOIS FAMILIER et OUVRIERS, contre 38.9 et 41.5% pour les corpus assez spécifiques THÉÂTRE et ENFANTS (textes lus pour le premier, individus encore en phase d'acquisition linguistique pour le second)).

Toutes ces données, à part les deux dernières, donnent de l'antéposition en *est-ce que* l'image d'un type phrastique un peu « en retrait », utilisé de façon constante mais toujours avec une fréquence relativement basse. Les observations du Chapitre 4 font par ailleurs entrevoir un type

phrastique relativement « neutre », utilisé aussi bien en contextes formels qu'en contextes informels. Les données du corpus ESLO (Section 4.2.7) pourraient même laisser penser que c'est un type phrastique « refuge » dans les contextes un peu formels (entretiens), au détriment par exemple du type in situ peut-être trop associé à l'informalité. C'est également ce que peuvent laisser penser les données du corpus EPAC (Section 4.1.3.2), dans lequel les auditeurs et auditrices, qui sont les seules personnes à n'utiliser aucune phrase in situ dans ce corpus, sont par ailleurs celles qui y utilisent le plus d'antépositions en *est-ce que* (19.1%).

Pour ne pas trop complexifier le protocole expérimental, la comparaison de l'antéposition en *est-ce que* ne s'est faite ici que vis-à-vis des deux types principalement opposés dans ce travail : l'antéposition avec inversion et l'in situ. En intégrant les données précédentes, il était prévu que l'antéposition en *est-ce que* soit jugée comme un intermédiaire entre les deux autres, emportant des jugements intermédiaires, quel que soit le contexte linguistique ou le groupe d'âge considéré.

5.2.4.2 Protocole et items expérimentaux

Le protocole de cette étude est à mi-chemin entre celui de l'étude AJT2 et celui de l'étude AJT3. L'intégralité des items-cibles de l'étude AJT2 a été réutilisée (F0.1), à la différence près que la condition « *antéposition simple* » a été remplacée par une condition « *antéposition en est-ce que* ». Les Tableaux 5.17 et 5.18 permettent de voir le point de variation d'une étude à l'autre. 10 distracteurs de l'étude AJT2 ont été repris, auxquels ont été adjoints 24 nouveaux distracteurs issus d'une autre étude sur l'ordre des clitiques en français. L'Annexe I.0.3 liste ces distracteurs.

Comme pour l'étude AJT2, la modalité de soumission des items a été écrite. Comme pour l'étude AJT3, deux échelles d'*acceptabilité* ont été utilisées. Cette décision découle de deux éléments. D'abord, l'étude AJT3 a montré que les volontaires, malgré une corrélation entre réponses, avaient su utiliser ces deux échelles différemment. Ensuite, la récolte des données s'en est trouvée facilitée : avec des stimuli écrits, l'expérience est en effet facilement réalisable sur un ordinateur personnel, sans avoir besoin de s'isoler pour écouter des sons au calme.

Type	Contexte	
	INFORMEL	FORMEL
Antéposition simple	A : Non, elle est pas là, ma soeur. B : Quand elle est partie?	A : Non, ma soeur n'est pas là. B : Quand elle est partie?
Antéposition avec inversion	A : Non, elle est pas là, ma soeur. B : Quand est-elle partie?	A : Non, ma soeur n'est pas là. B : Quand est-elle partie?
In situ	A : Non, elle est pas là, ma soeur. B : Elle est partie quand?	A : Non, ma soeur n'est pas là. B : Elle est partie quand?

TABLEAU 5.17 – Rappel d'un item expérimental (AJT2)

Type	Contexte	
	INFORMEL	FORMEL
Antéposition + ESK	A : Non, elle est pas là, ma soeur. B : Quand est-ce qu' elle est partie?	A : Non, ma soeur n'est pas là. B : Quand est-ce qu' elle est partie?
Antéposition avec inversion	A : Non, elle est pas là, ma soeur. B : Quand est-elle partie?	A : Non, ma soeur n'est pas là. B : Quand est-elle partie?
In situ	A : Non, elle est pas là, ma soeur. B : Elle est partie quand?	A : Non, ma soeur n'est pas là. B : Elle est partie quand?

TABLEAU 5.18 – Exemple d'item expérimental (AJT4)

Chaque liste comptait 30 items cibles, 34 distracteurs et 3 entraînements. Le carré latin utilisé pour les études AJT2 et AJT3 a été repris : un design 3x2 avec les variables CONTEXTE (*formel* vs. *informel*) et TYPE syntaxique (*antéposition en est-ce que*, *antéposition et inversion*, *in situ*).

L'étude commençait après une page de consentement (Annexe I.0.1) puis une page de consignes (Annexe I.0.2). Elle comportait 67 items, pour un temps total étalonné à environ 20 minutes. Cette fois-ci, c'est après la passation que les volontaires remplissaient une dernière page de recueil de données personnelles basiques (âge, niveau d'étude, domaine professionnel ou d'études, langues).

Sous chaque item se trouvaient deux échelles : l'une pour le *bon français*, l'autre seconde pour juger du caractère *adapté* de ce français à la situation. Ces échelles prenaient cette fois la forme de *sliders*, des curseurs à glisser sur une échelle aux crans non-numérotés, entre deux extrémités (pas du tout d'accord – totalement d'accord). La Figure 5.18 illustre ce qui était visible sur l'écran.

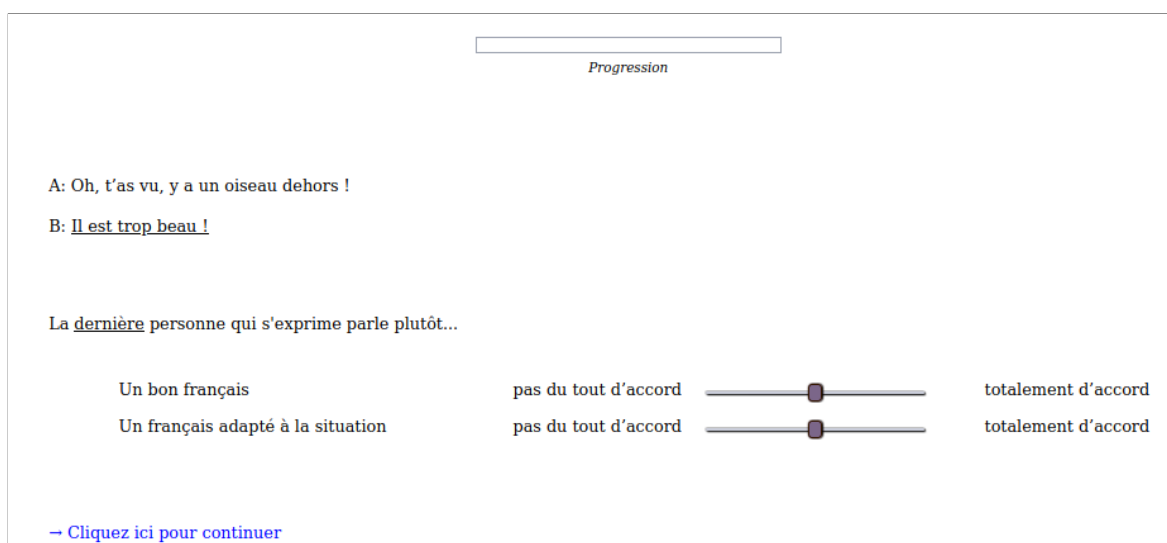


FIGURE 5.18 – Capture d'écran sur un item expérimental (AJT4)

Le programme informatique enregistrait la position du curseur comme une réponse entre 0 et 10 (il y avait donc 11 crans sur chaque échelle). Les volontaires devaient cliquer sur un lien pour passer à l'item suivant. Aucune limite de temps n'était imposée. Il était indiqué dans les consignes qu'aucune bonne ou mauvaise réponse n'était attendue.

5.2.4.3 Participantes et participants

92 personnes ont pris part à l'expérience au printemps 2020, via la plateforme IbexFarm [DRUMMOND, 2016]. Les résultats d'une personne ont été supprimés car l'âge indiqué était 8 ans, ce qui

ne correspondait pas aux conditions de recrutement. Celui-ci s'est fait via le RISC et les réseaux sociaux. Comme indiqué dans la page de consentement, il n'y avait pas de compensation.

Les 91 personnes dont les résultats ont été analysés avaient entre 18 et 84 ans ($\mu = 39.9$, $m = 30.5$). 64 ont déclaré être de genre féminin, 26 ont déclaré être de genre masculin, et 1 personne a opté pour la désignation *non-binaire*. Les plus hauts diplômes obtenus par ces personnes étaient :

- en dessous du baccalauréat, N=5
- le baccalauréat, N=17
- un diplôme équivalent à Bac+3, N=32
- un diplôme équivalent à Bac+5, N=31
- un diplôme équivalent à Bac+8, N=6

24 personnes ont indiqué avoir pour région de naissance l'Île de France, 61 une autre circonscription administrative du territoire français métropolitain, et 6 ont indiqué venir de France sans préciser la région.

13 personnes ont précisé avoir mené des études dans un domaine lié aux lettres et langues, 19 dans les sciences humaines et sociales, 20 dans les sciences et techniques, 5 dans les arts, 6 dans le droit et les sciences politiques, 6 dans l'économie et gestion, 6 dans la santé, 13 dans un autre domaine, et 3 n'étaient pas concernées. 17 personnes ont précisé travailler dans un domaine lié à l'enseignement, 3 dans le journalisme ou les professions littéraires, 9 dans une profession de type administratif, 5 dans le commerce et la vente, 1 dans l'agriculture, 10 dans un emploi de direction, 5 dans l'industrie, 7 dans la santé, 1 dans le service à la personne, et 10 estimaient relever d'une profession scientifique : 23 n'ont pas précisé. Les données sur les langues maîtrisées n'ont pas été traitées, à part une vérification de la langue maternelle indiquée.

Au total, 6097 paires de jugements (notes sur les deux échelles prises ensemble) ont été enregistrées.

5.2.4.4 Résultats

Seules les 2730 paires de jugements reçues sur les items cibles ont été conservés pour analyse. Les réponses fournies sur les deux échelles disponibles ont été codées numériquement entre 0 (Pas du tout d'accord) et 10 (Totalemment d'accord).

Comme pour l'étude AJT3, les résultats seront décrits en trois temps. D'abord seront présentées les données concernant le premier jugement : « *La personne qui s'exprime (...) parle un bon français* ». Ensuite seront présentées les données concernant le second jugement : « *La personne qui s'exprime (...) parle un français adapté à la situation* ». Enfin, un bilan de ces résultats sera effectué.

Comme pour l'étude AJT3, une corrélation entre les réponses données sur les deux échelles existe, avec un coefficient de corrélation $r = .29$ ($p < 0.01$, méthode = Spearman). Comme pour l'étude AJT3 néanmoins, cette corrélation, peu surprenante vu le protocole employé, ne masque pas un certain nombre d'observations intéressantes.

5.2.4.4.1 Résultats généraux (*bon français*)

Le Tableau 5.19 et la Figure 5.19 donnent les moyennes globales brutes des jugements rendus sur chacune des trois formes interrogatives analysées, pour la première échelle proposée aux volontaires.

Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
6.60	8.67	5.32

TABLEAU 5.19 – Note moyenne en fonction du type interrogatif (AJT4, *bon français*)

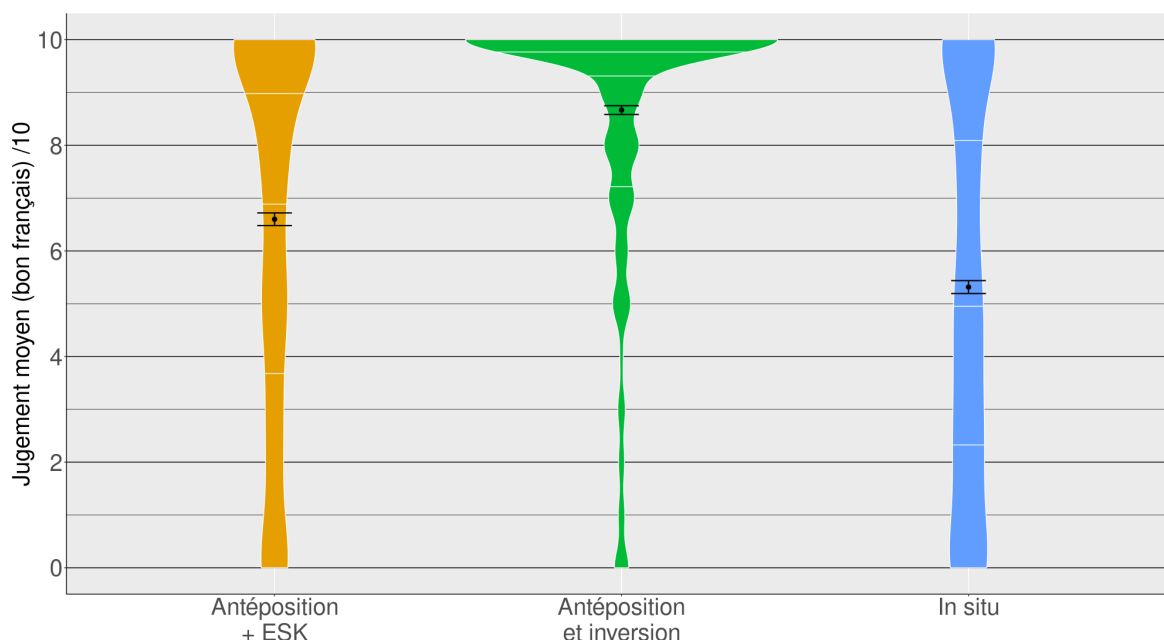


FIGURE 5.19 – Note moyenne en fonction du type interrogatif (AJT4, *bon français*)

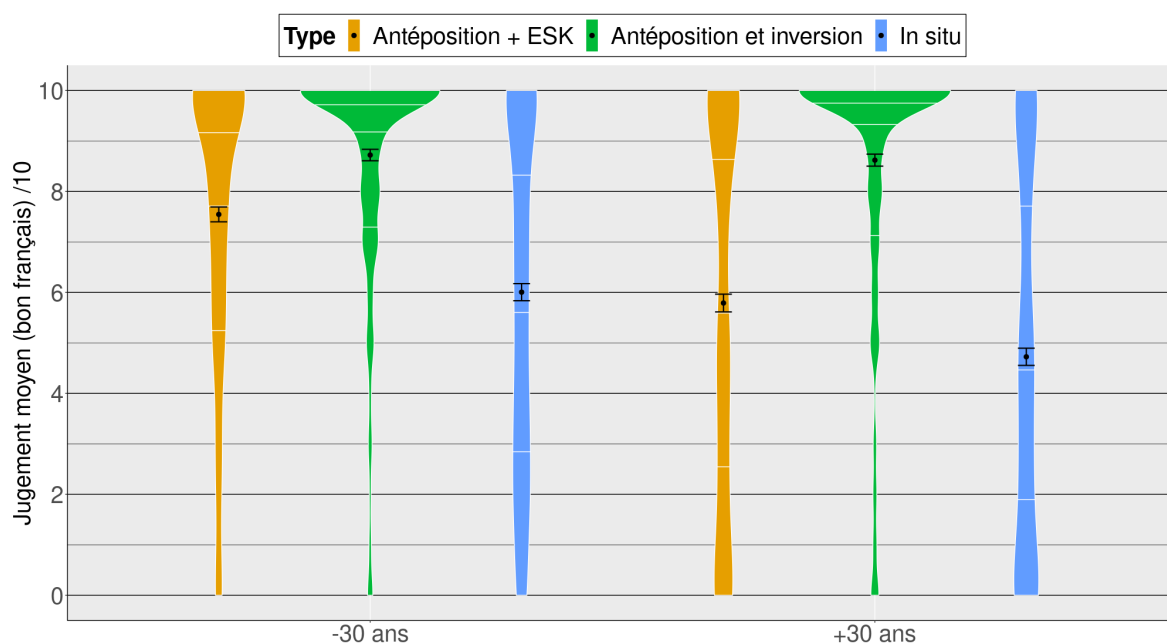
Comme dans les études AJT1, AJT2 et AJT3, il émerge de ces données une préférence pour les phrases avec antéposition et inversion, au détriment des deux autres types phrastiques. À la différence des résultats précédents cependant, ce sont les phrases in situ qui sont associées aux jugements les plus bas ici, quasiment à la limite de l'*acceptabilité* (5.32/10). Pour la même question (*bon français*), les jugements associés aux phrases in situ dans l'étude AJT3 étaient loin de cette limite (4.58/6). Plusieurs hypothèses peuvent expliquer ceci. La première est un effet de modalité : les phrases in situ peuvent être plus marquées à l'écrit qu'elles ne le sont à l'oral, par l'éloignement vis-à-vis d'une situation d'interaction spontanée. Une autre hypothèse est un effet de contraste avec la structure avec antéposition en *est-ce que*. Ce type phrastique viendrait dans cette perspective s'insérer en second, après l'antéposition avec inversion, dans « l'échelle des phrases les mieux formées du français ». La présence de cette alternative dans l'expérience pourrait avoir rendu plus sévères les jugements portés sur les phrases in situ.

5.2.4.4.2 Un effet d'âge (*bon français*)

Le Tableau 5.20 et la Figure 5.20 donnent les moyennes globales brutes des jugements pour chacune des trois formes interrogatives, pour la première échelle à nouveau. Par continuité avec les expériences passées, la séparation en âge à de nouveau été établie au seuil de 30 ans.

	Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
-30 ans	7.55	8.72	6.00
+30 ans	5.79	8.62	4.72

TABLEAU 5.20 – Note moyenne en fonction du type interrogatif et de l'âge (AJT4, *bon français*)


 FIGURE 5.20 – Note moyenne en fonction du type interrogatif et de l'âge (AJT4, *bon français*)

Le schéma qui émerge est ici conforme à ce qui était apparu dans l'étude AJT3. Quand la question est explicitement celle de la bonne formation, à l'inverse de l'étude AJT2, les moins de 30 ans ont un comportement général similaire aux plus de 30 ans. Le contraste entre structures est ainsi visible chez les deux groupes, avec la même hiérarchie. L'antéposition avec inversion emporte de meilleurs jugements que l'antéposition en *est-ce que*, laquelle est à son tour associée à de meilleurs jugements que le type in situ.

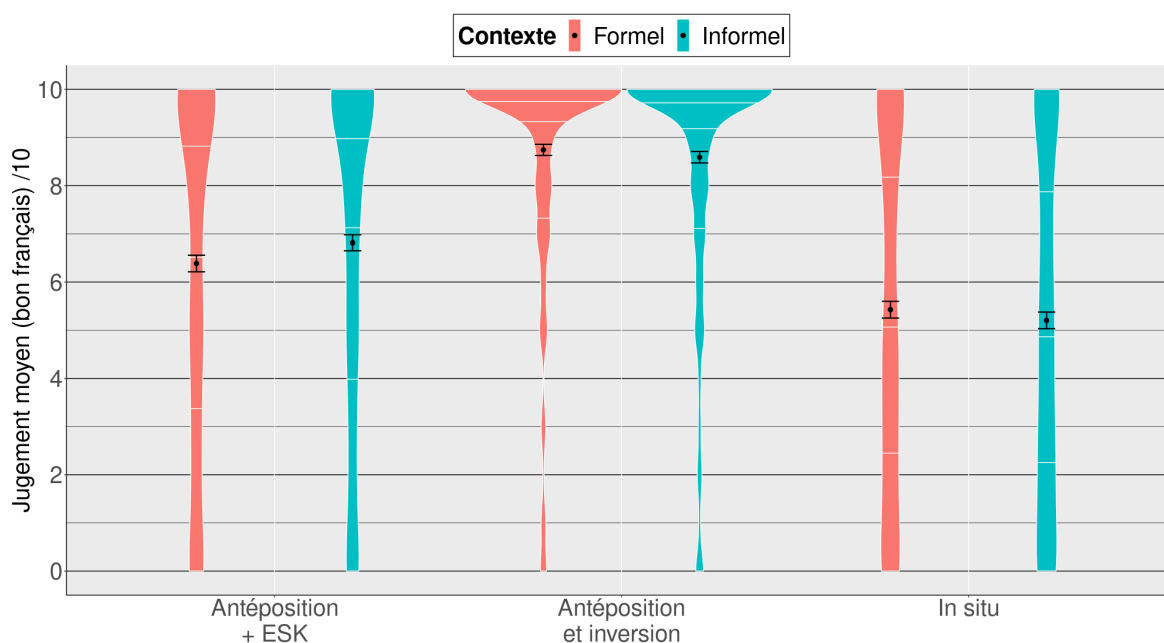
Si le comportement est le même à travers les groupes d'âge, la préférence affichée pour l'antéposition avec inversion est plus forte chez les plus de 30 ans, comme elle l'était dans l'étude AJT2. En effet, dans ce groupe, les antépositions en *est-ce que* se rapprochent de la limite d'acceptabilité (5.79/10), alors que les phrases in situ vont jusqu'à le franchir et devenir relativement *mauvaises*, au sens de mal-formées (4.72/10). Ici s'observe une conception proche de ce que présentent RIEGEL et al. [2014], pour lesquels le type in situ est une structure de la langue parlée et *familière*.

5.2.4.4.3 Un effet de contexte (*bon français*)

Le Tableau 5.21 et la Figure 5.21 donnent les moyennes brutes des jugements pour chacune des trois formes interrogatives, pour la première échelle, en fonction de la formalité du contexte (*formel* vs. *informel*).

	Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
Contexte formel	6.38	8.74	5.43
Contexte informel	6.82	8.59	5.20

 TABLEAU 5.21 – Note moyenne en fonction du type interrogatif et du contexte (AJT4, *bon français*)


 FIGURE 5.21 – Note moyenne en fonction du type interrogatif et du contexte (AJT4, *bon français*)

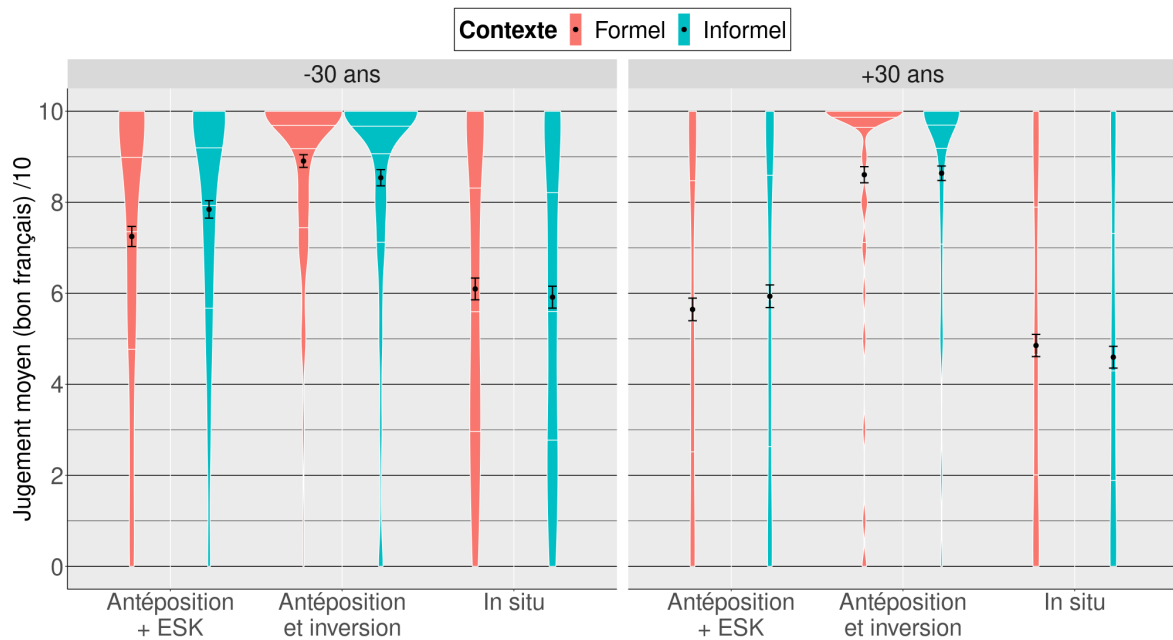
Le même effet de contexte que dans l'étude AJT3 se retrouve ici : les antépositions avec inversion sont associées à des jugements un peu plus hauts quand elles s'inscrivent dans un contexte formel. Un peu paradoxalement, il en va de même pour les phrases in situ. Il y a ici une certaine discordance, puisque ce type phrastique est celui qui est jugé le français le moins « bon » parmi les phrases proposées. Fait intéressant, les antépositions en *est-ce que* sont elles moins bien jugées en contexte formel qu'en contexte informel. Ceci entre en relative contradiction avec l'hypothèse formulée, selon laquelle ce type phrastique « remplacerait » le type in situ en contextes formels dans les données de corpus. Il faut peut-être y voir que, en contrastant l'antéposition avec inversion à l'antéposition en *est-ce que*, cette dernière est vue comme la version informelle à utiliser au lieu de l'in situ. Les phrases avec antéposition en *est-ce que* demeurent jugées, quoi qu'il en soit, comme « mieux formées » dans ces contextes formels que les phrases in situ.

5.2.4.4.4 Une interaction entre âge et contexte (*bon français*)

Le Tableau 5.22 et la Figure 5.22 approfondissent les résultats en synthétisant les moyennes des jugements donnés sur chaque type phrastique, en fonction du contexte mais aussi du groupe d'âge.

	-30 ans		+30 ans	
	Formel	Informel	Formel	Informel
Antéposition en <i>est-ce que</i>	7.25	7.84	5.64	5.93
Antéposition et inversion	8.90	8.54	8.60	8.64
In situ	6.10	5.91	4.85	4.60

 TABLEAU 5.22 – Note moyenne relative aux type interrogatif, contexte et âge (AJT4, *bon français*)


 FIGURE 5.22 – Note moyenne relative aux type interrogatif, contexte et âge (AJT4, *bon français*)

La sensibilité au contexte précédemment relevée semble ici, pour cette tâche, ressortir plus chez les moins de 30 ans. C'est un contraste avec les études AJT2 et AJT3. Notamment, pour ce groupe, les antépositions avec inversion semblent associées à de meilleurs jugements en contexte formel qu'en contexte informel. Il semble aussi pour ce groupe y avoir une différence pour les antépositions en *est-ce que*, qui sont moins bien jugées en contexte formel.

5.2.4.4.5 Analyses bayésiennes (*bon français*)

Un modèle de régression bayésienne a été calculé sur la base des réponses à cette question du bon français (Annexe I.0.4), strictement parallèle aux modèles calculés pour l'étude AJT3. Le modèle avait donc pour variable dépendante la note attribuée, et en variable indépendante l'interaction triple entre trois effets fixes : le TYPE syntaxique (*antéposition en est-ce que*, *antéposition avec inversion*, *in situ*, avec l'antéposition avec inversion comme niveau de référence), le groupe d'ÂGE (moins ou plus de 30 ans, avec le groupe des moins de 30 ans pour référence), et le CONTEXTE (*formel* ou *informel*, avec *informel* pour niveau de référence). Les effets aléatoires étaient les individus ("ID"), avec pour pente l'interaction *type*contexte*) et les items ("Item"), avec pour pente l'interaction triple *type*âge*contexte*). C'est toujours la famille "cumulative" qui a été sélectionnée pour faire tourner ces modèles, avec 4 chaînes et 6000 itérations par chaîne. La convergence des modèles a été atteinte et vérifiée.

La Figure 5.23 illustre les distributions postérieures de ce modèle pour chacun des paramètres et de leurs interactions. Elle permet une visualisation de la valeur généralisable des principaux résultats observés sur les données brutes.

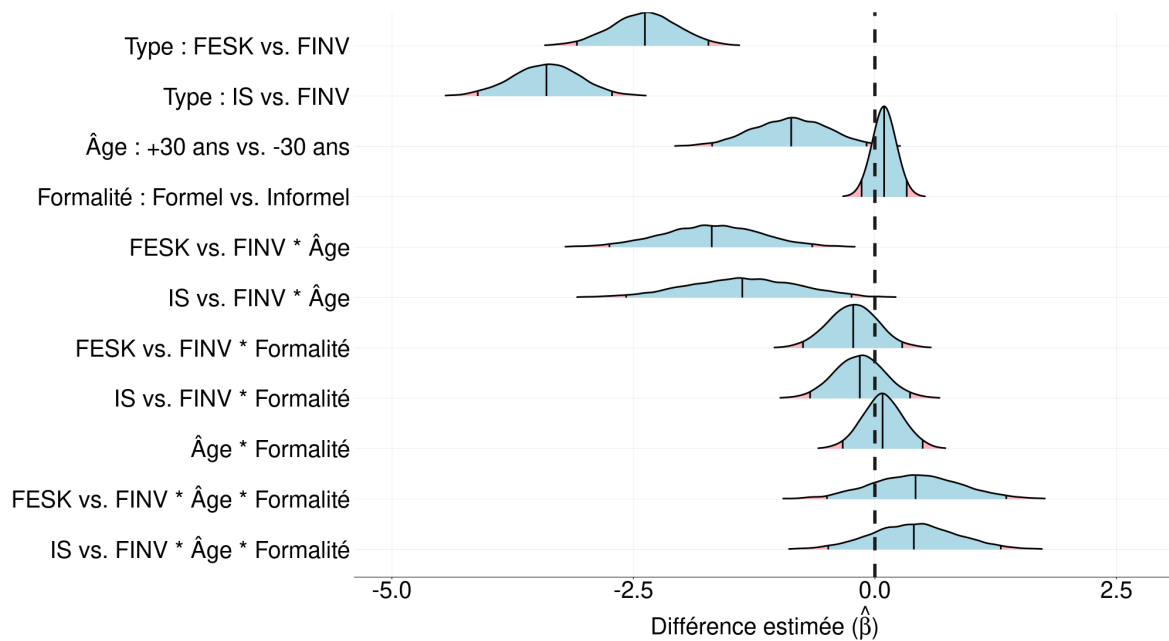


FIGURE 5.23 – Distributions postérieures pour le modèle brms (AJT4, bon français)

L'effet simple de TYPE syntaxique est confirmé. Les antépositions en *est-ce que* sont de manière générale moins bien jugées que les antépositions avec inversion ($\hat{\beta} = -2.39$, 95%CrI=[-3.08,-1.72], $P(\beta) < 0=1$), tout comme les interrogatives in situ ($\hat{\beta} = -3.41$, 95%CrI=[-4.11,-2.72], $P(\beta) < 0=1$).

Un effet simple de groupe d'ÂGE est également visible (les plus de 30 ans ont globalement donné des notes plus basses que les moins de 30 ans, $\hat{\beta} = -0.87$, 95%CrI=[-1.68,-0.09], $P(\beta) < 0=1$). Ceci est notamment dû aux jugements plus bas attribués par ce groupe, en comparaison avec celui des moins de 30 ans, aux interrogatives avec antéposition en *est-ce que* et aux interrogatives in situ. Cela se retrouve dans les interactions TYPE*ÂGE autant pour les premières ($\hat{\beta} = -1.69$, 95%CrI=[-2.75,-0.65], $P(\beta) < 0=1$) que pour les secondes ($\hat{\beta} = -1.38$, 95%CrI=[-2.57,-0.24], $P(\beta) < 0=0.99$).

Un effet principal de FORMALITÉ est également supporté par les données, de manière modérée ($\hat{\beta} = 0.10$, 95%CrI=[-0.14,0.33], $P(\beta) > 0=0.8$), qui traduit que des jugements plus hauts ont globalement été rendus en contexte formel. Ceci est à relier aux interactions TYPE*CONTEXTE. La première supporte, modérément, l'idée qu'en contexte formel l'écart se creuse entre les antépositions avec inversion et les antépositions en *est-ce que* ($\hat{\beta} = -0.23$, 95%CrI=[-0.74,0.28], $P(\beta) < 0=0.81$). La seconde relève plus de la tendance, quand les phrases avec antéposition et inversion sont comparées aux phrases in situ ($\hat{\beta} = -0.16$, 95%CrI=[-0.67,0.36], $P(\beta) < 0=0.73$).

Enfin, s'il n'y a pas d'interaction générale ÂGE*CONTEXTE, des différences entre les plus de 30 ans et les moins de 30 ans sont néanmoins visibles dans les interactions triples TYPE*ÂGE*CONTEXTE. Celles-ci corroborent ce qui ressortait des données brutes. Par comparaison avec les moins de 30 ans, il semble en effet que l'écart de jugements entre les antépositions en *est-ce que* et les antépositions avec inversion se réduisent chez les plus de 30 ans ($\hat{\beta} = 0.42$, 95%CrI=[-0.50,1.36], $P(\beta) > 0=0.81$), tout comme l'écart entre les phrases in situ et celles avec antéposition et inversion ($\hat{\beta} = 0.41$, 95%CrI=[-0.48,1.30], $P(\beta) > 0=0.82$). Il faut relier cela par exemple aux jugements plus bas rendus en contexte formel par les moins de 30 ans pour les antépositions en *est-ce que*, alors que dans le même temps leurs jugements étaient plus hauts pour les antépositions avec inversion. Les jugements des plus de 30 ans restent malgré ceci plus bas en valeur absolue, quel que soit le contexte et pour ces deux

types phrastiques, que ceux des moins de 30 ans. C'est leur sensibilité au contexte qui semble moindre, d'après ces données, notamment dans leur appréciation des antépositions avec inversion (contrairement à l'étude AJT3 par exemple). À nouveau, une explication à ce phénomène pourrait se trouver dans la modalité de l'expérience. Les formes antéposées en *est-ce que* et les formes in situ apparaissent peut-être comme plus marquées à l'écrit, ce qui entraîne des jugements très bas chez les plus de trente ans, de manière stable à travers les conditions de formalité.

5.2.4.4.6 Résultats généraux (*français adapté*)

Les interrogatives avec antéposition en *est-ce que* ne concurrencent pas les phrases avec antéposition simple en termes de « bon français ». Elles participent peut-être, cependant, à rendre plus sévères les jugements portés sur les phrases in situ. Il faut maintenant observer les données récoltées sur la seconde échelle, celle qui permet aux volontaires de juger de la bonne adaptation au contexte. Le Tableau 5.23 et la Figure 5.24 donnent les moyennes globales brutes des jugements associés aux différents types interrogatifs sur ce point.

Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
6.75	7.39	7.27

TABLEAU 5.23 – Note moyenne en fonction du type interrogatif (AJT4, *français adapté*)

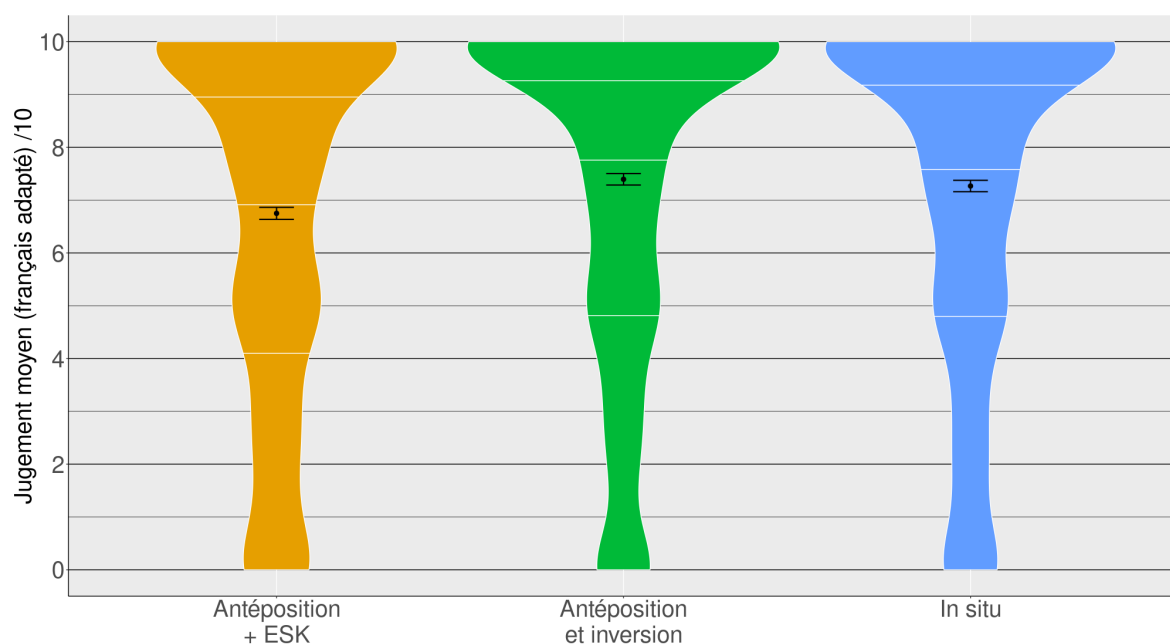


FIGURE 5.24 – Note moyenne en fonction du type interrogatif (AJT4, *français adapté*)

Comme c'était le cas pour l'étude AJT3, les jugements semblent ici différents de ceux rendus sur la première échelle. L'écart entre les trois types syntaxiques s'est ici réduit (0.64/10 au maximum, contre 3.35 sur l'échelle du « bon français »). Et, d'une façon encore parallèle avec l'étude AJT3, il semble notamment que la différence entre l'antéposition avec inversion et le type in situ ait virtuellement disparu ici : les deux semblent jugés tout autant « adaptés » l'un que l'autre. Un autre élément à noter est la stabilité des jugements rendus sur les phrases avec antéposition en *est-ce que*, par rapport à la première échelle (6.75/10 ici, contre 6.60/10 pour le « bon français »). Une dernière observation est que, alors que la moyenne la plus haute sur la première échelle était à

8.67/10 (antéposition avec inversion), ici aucune ne dépasse 7.39/10. Alors qu'il y a un « meilleur français » qui se détache nettement des autres types phrastiques, il semble y avoir, ici encore (c'était aussi le cas pour l'étude AJT3), un rapprochement entre les différents types interrogatifs au regard des usages réels qui peuvent en être faits.

5.2.4.4.7 Un effet d'âge (*français adapté*)

Le Tableau 5.24 et la Figure 5.25 donnent les moyennes brutes des jugements rendus sur les types phrastiques en fonction du groupe d'âge (de nouveau, plus ou moins de 30 ans) des volontaires.

	Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
-30 ans	7.11	7.40	7.33
+30 ans	6.44	7.39	7.21

TABLEAU 5.24 – Note moyenne en fonction du type interrogatif et de l'âge (AJT4, *français adapté*)

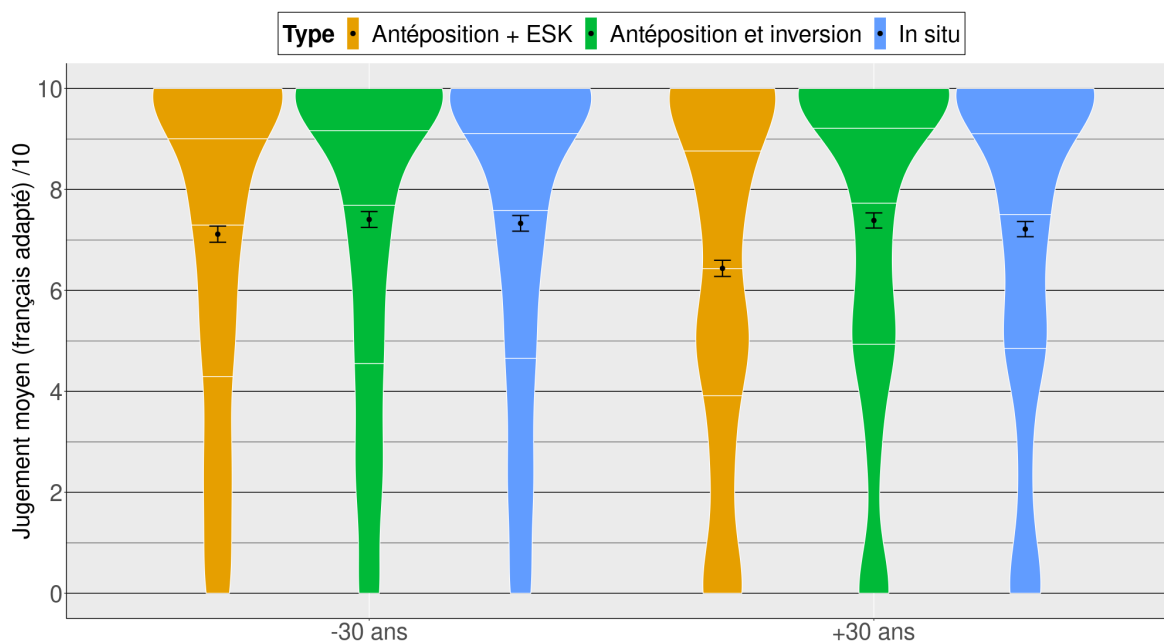


FIGURE 5.25 – Note moyenne en fonction du type interrogatif et de l'âge (AJT4, *français adapté*)

Le schéma visible ici est conforme à celui déjà observé : l'écart global entre types phrastiques semble davantage résulter des jugements du groupe de plus de 30 ans. La différence entre antéposition avec inversion et in situ semble inactive chez les deux groupes. À l'inverse, la différence entre antéposition en *est-ce que* et antéposition avec inversion est visible chez les deux groupes, mais elle est plus manifeste pour celui des plus de 30 ans. Il est à noter que, contrairement aux données de l'étude AJT3, les moins de 30 ans ne semblent pas trouver l'antéposition avec inversion « inadaptée ». Peut-être y a-t-il aussi une influence de la modalité écrite.

5.2.4.4.8 Un effet de contexte (*français adapté*)

Le Tableau 5.25 et la Figure 5.26 permettent de visualiser les jugements moyens rendus sur chaque type interrogatif en fonction du contexte, formel ou informel, qui précédait chaque phrase.

	Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
Contexte formel	6.57	7.76	7.40
Contexte informel	6.93	7.03	7.14

TABLEAU 5.25 – Note moyenne en fonction du type interrogatif et du contexte (AJT4, français adapté)

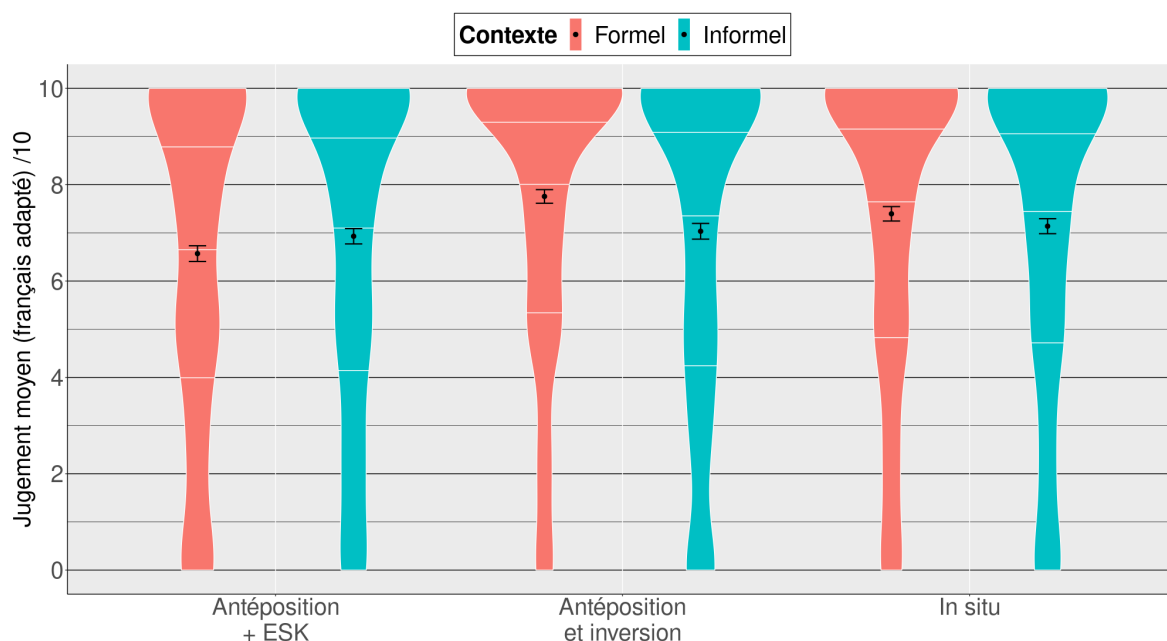


FIGURE 5.26 – Note moyenne en fonction du type interrogatif et du contexte (AJT4, français adapté)

Les effets de formalité visibles ici sont plus apparents que ceux notés sur la première échelle. Le contexte formel rend plus adaptées les interrogatives avec antéposition et inversion, ou au contraire le contexte informel les rend moins adaptées. En parallèle, les interrogatives avec antéposition en *est-ce que* sont de manière générale plus adaptées en contexte informel. Les interrogatives in situ suivent un patron observé dans l'étude AJT3 : elles semblent légèrement plus adaptées en contexte formel.

5.2.4.4.9 Une interaction entre âge et contexte? (français adapté)

Le Tableau 5.26 et la Figure 5.27 croisent à nouveau les données, et offrent une visualisation des jugements moyens rendus par les volontaires, en fonction de leur âge mais aussi des variations de contexte.

	-30 ans		+30 ans	
	Formel	Informel	Formel	Informel
Antéposition en <i>est-ce que</i>	6.97	7.26	6.22	6.65
Antéposition et inversion	7.72	7.09	7.78	6.99
In situ	7.50	7.15	7.30	7.13

TABLEAU 5.26 – Note moyenne relative aux type interrogatif, contexte et âge (AJT4, français adapté)

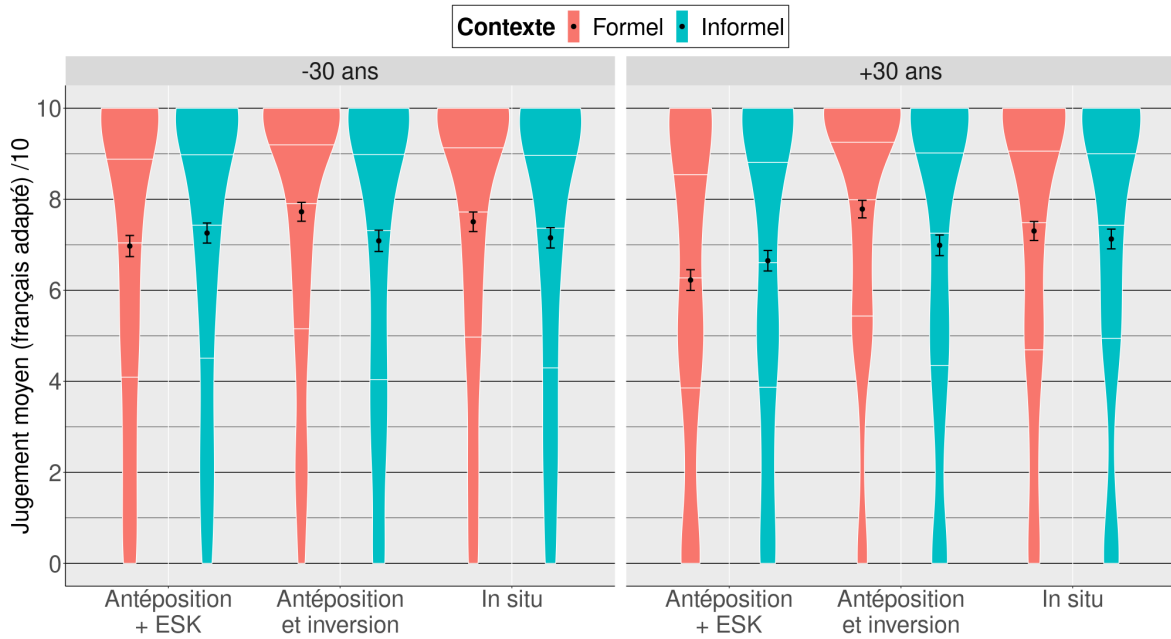


FIGURE 5.27 – Note moyenne relative aux type interrogatif, contexte et âge (AJT4, français adapté)

Le contexte affecte les jugements rendus mais semble ne pas les affecter différemment selon le groupe d'âge. Certes, une *neutralisation* des structures s'observe chez les moins de 30 ans, pour lequel les antépositions avec inversion deviennent le type le moins adapté en contexte informel. Ceci se retrouve aussi chez les plus de 30 ans. De plus, les deux groupes jugent les antépositions en *est-ce que* moins adaptées en contexte formel alors que celles avec inversion y sont plus adaptées, tout comme les phrases in situ.

5.2.4.4.10 Analyses bayésiennes (français adapté)

Un modèle similaire à celui calculé pour le « bon français » a été calculé pour le français « adapté ». L'Annexe 1.0.5 en compile les détails, et la Figure 5.28 en illustre les distributions postérieures.

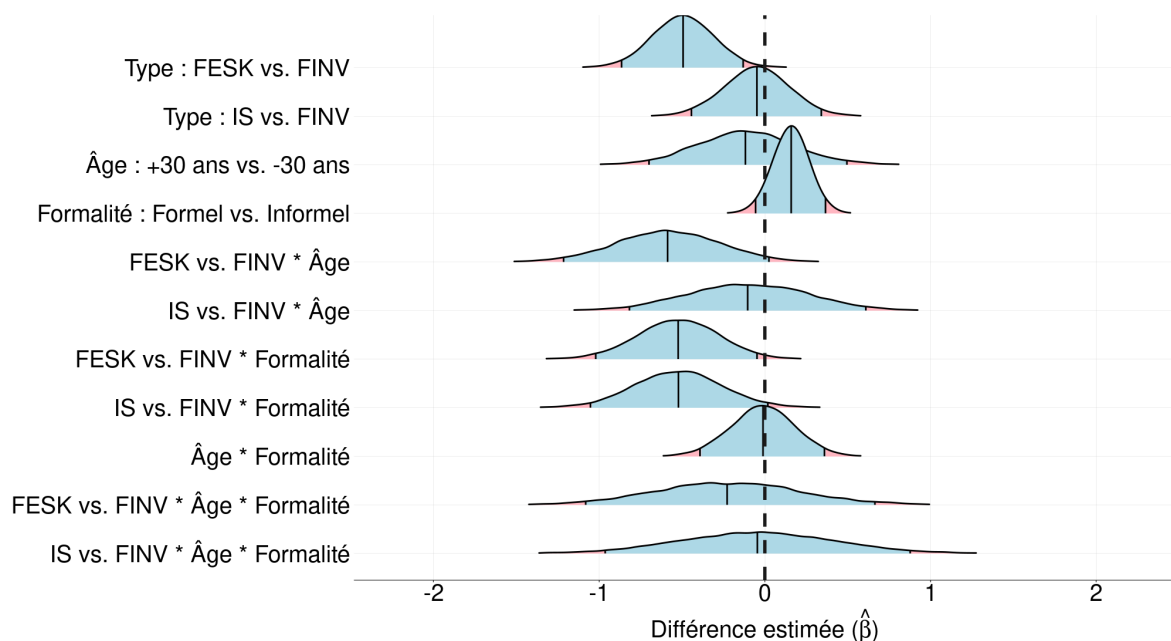


FIGURE 5.28 – Distributions postérieures pour le modèle brms (AJT4, français adapté)

Ces statistiques confirment les résultats dégagés des données brutes.

L'effet simple de TYPE n'est ainsi opérant que pour les antépositions en *est-ce que* : elles sont associées à des jugements moins hauts que les phrases avec antéposition et inversion ($\hat{\beta} = -0.49$, 95%CrI=[-0.86,-0.13], $P(\beta) < 0=0.99$), mais ce n'est pas le cas pour les phrases in situ. De même, cet effet n'entre dans une interaction TYPE*ÂGE que pour ces mêmes antépositions en *est-ce que*. Il n'y a que pour celles-ci que l'écart de jugement se creuse, vis-à-vis des antépositions avec inversion, chez les plus de 30 ans ($\hat{\beta} = -0.59$, 95%CrI=[-1.21,0.02], $P(\beta) < 0=0.98$).

Les effets liés au CONTEXTE sont par ailleurs bien présents. C'est le cas de façon générale, avec un effet simple qui traduit des jugements globalement plus hauts rendus en contexte formels ($\hat{\beta} = 0.16$, 95%CrI=[-0.06,0.37], $P(\beta) > 0=0.93$). Ceci devient intéressant dans les interactions TYPE*CONTEXTE. En contexte formel, l'écart de jugement se creuse en effet entre les antépositions avec inversion et les antépositions en *est-ce que* ($\hat{\beta} = -0.52$, 95%CrI=[-1.02,-0.05], $P(\beta) < 0=0.98$). C'est aussi le cas pour l'écart entre les antépositions avec inversion et les phrases in situ ($\hat{\beta} = -0.52$, 95%CrI=[-1.05,0.02], $P(\beta) < 0=0.97$). Ceci est à relier avec l'augmentation des jugements rendus pour les antépositions avec inversion en contexte formel, qui dépasse en intensité l'augmentation constatée pour les phrases in situ. C'est aussi à relier aux moins bons jugements associés aux antépositions en *est-ce que* en contexte formel, qui augmentent d'autant l'écart entre les deux types phrastiques.

Aucune sensibilité particulière d'un des deux groupes d'âge n'est par ailleurs corroborée par une interaction TYPE*ÂGE*CONTEXTE ou TYPE*ÂGE*CONTEXTE. Les deux groupes apparaissent donc unis dans leur appréciation de l'adaptation au contexte des différents types phrastiques. La modalité écrite peut, à nouveau, être un facteur provoquant ceci. Cette observation peut aussi se relier au statut particulier des antépositions en *est-ce que*, dont il était attendu qu'elles soient un type phrastique peu saillant par rapport notamment aux types plus marqués que sont l'in situ ou à l'antéposition avec inversion. En tant que tel, il y aurait peu de divergences liées à l'âge sur ses emplois possibles.

5.2.4.5 Bilan : L'antéposition en *est-ce que*, bien formée mais peu adaptée

L'hypothèse principale à l'origine de cette étude était que l'antéposition en *est-ce que* serait un type de phrase « neutre », voire « refuge ». Dans cette perspective, il apparaissait comme un intermédiaire entre les phrases in situ et les phrases avec antéposition et inversion. Les données semblent corroborer cette analyse, avec un type phrastique relevant globalement d'un français « correct », loin du standard qu'est l'antéposition avec inversion mais aussi loin de l'appréciation portée sur les phrases in situ (presque jugées ici comme étant du mauvais français). Quant à leur emploi, les francophones semblent les considérer comme généralement plutôt adaptées aux divers contextes proposés (notamment les moins de 30 ans). Elles sont cependant moins adaptées que les phrases in situ et les antépositions avec inversion. Cette différence se creuse dans les contextes formels, pour lesquels le caractère *adapté* de ce type phrastique se dégrade. Par ailleurs, la relative constance des jugements portés en fonction de la formalité du contexte, à travers les deux groupes d'âge, appuient un peu plus cette nature intermédiaire de l'antéposition en *est-ce que*.

5.2.5 Jugements d'acceptabilité : Récapitulatif

Le Tableau 5.27 fournit un résumé global des différents travaux décrits ci-dessus. Étude par étude, il compile le type de question posée aux volontaires sur les phrases qui leur étaient soumises,

la modalité de soumission des items, et les différents types phrastiques analysés. La préférence générale, hors effets d'âge et de contexte est indiquée par un chiffre « hiérarchisant » les préférences visibles dans les données. Les colonnes CONTEXTE et ÂGE montrent non pas l'évolution de cette hiérarchie, mais l'évolution du *jugement moyen* porté, soit dans un contexte formel par rapport à un contexte informel, soit d'un groupe d'âge à l'autre autour du seuil de 30 ans.

ÉTUDE	JUGEMENT	MODALITÉ	TYPE	PRÉFÉRENCE GÉNÉRALE	CONTEXTE informel → formel	ÂGE -30 → +30
AJT1	acceptabilité	écrit	FINV	1		+
			F	3		
			IS	2		-
AJT2	acceptabilité	écrit	FINV	1	+	=
			F	3	=	-
			IS	2	=	-
AJT3	bon français	oral	FINV	1	+	=
			F	3	=	=
			IS	2	=	=
	français adapté		FINV	2	+	+
			F	3	=	-
			IS	1	+	=
AJT4	bon français	écrit	FINV	1	+	=
			FESK	2	-	-
			IS	3	+	-
	français adapté		FINV	1	+	+
			FESK	2	-	-
			IS	1	+	=

F = antéposition simple, FINV = antéposition avec inversion, FESK = antéposition en *est-ce que*

IS = in situ

= → pas d'évolution; - → diminue; + → augmente

TABLEAU 5.27 – Récapitulatif (Jugements d'acceptabilité)

Une préférence générale s'observe pour les phrases avec antéposition et inversion, mais cette préférence est une préférence de forme, en référence à un « bon français » liée à la notion de standard et au poids de la norme linguistique. Quand il est question d'usage, et d'adaptation aux situations d'interaction, la préférence pour les antépositions avec inversion diminue, et celles pour les phrases in situ augmente. Les influences contextuelles sont cependant importantes (la formalité favorise les antépositions avec inversion comme les phrases in situ, et défavorise les antépositions en *est-ce que*). Les considérations sont aussi variables selon le groupe d'âge, avec pour tendance générale que les plus de 30 ans favorisent souvent dans leurs jugements l'antéposition avec inversion, au détriment des autres types phrastiques. La modalité orale, par ailleurs, semble réduire le biais général en faveur de l'antéposition avec inversion, puisque le type in situ y est même jugé plus « adapté » (AJT3).

Ces quatre études ont ainsi permis de sonder explicitement les préférences francophones en matière d'interrogatives partielles. Elles n'ont pas permis d'approcher les représentations mentales qui y étaient associées, notamment en termes de stéréotypes. Les tâches de *matched-guise* présentées ci-après viennent compléter cette perspective.

5.3 Matched-Guise et projection sociolinguistique

La tâche expérimental de matched-guise (MG, [LAMBERT et al., 1960]) est précieuse en ce qu'elle permet d'explorer les ramifications sociologiques ou sociétales d'une variable linguistique. C'est ainsi qu'il est possible de regarder pourquoi et comment une personnalité politique peut changer sa manière d'être perçue par son électorat [PODESVA et al., 2015]. Pour généraliser, il s'agit donc de soumettre à des personnes naïves linguistiquement des stimuli – écrits ou oraux – qui ne présentent qu'un seul point de variation linguistique, sans évidemment que ce point soit explicite. En parallèle, ces personnes doivent se prononcer sur un trait de caractère ou une caractéristique sociale décrivant la personne supposée à l'origine de ces stimuli : est-ce que c'est une personne aimable ou non, est-ce que c'est une personne riche ou non, etc.

Entre 2017 et 2020, trois tâches de ce type ont été menées sur les interrogatives partielles du français. Deux ont été menées sur internet et à l'écrit, par facilité de récolte des données, et l'une a été passée en laboratoire, sur le modèle de l'étude AJT3 présentée plus haut (Section 5.2.3, avec des stimuli audio. Les personnes interrogées sur Internet viennent majoritairement du Relais d'Information sur les Sciences de la Cognition (RISC, CNRS, UMR 3352) et des réseaux sociaux, mais sont aussi des personnes étudiant à l'Université de Paris ou dans les environs.

5.3.1 Interrogatives et projection d'indices sociaux à l'écrit (MG1)

La première étude avec ce design est un travail réalisé avant cette thèse (en complément de l'étude AJT1). Elle est décrite en détail dans [THIBERGE, 2018]. 58 francophones (âge= 18-69 ans, $\mu = 38$, $m = 34$) lisaient des interviews construites pour l'expérience, avec des personnalités telles qu'un sportif, un maire, un économiste. Les différentes versions de ces entretiens ne variaient que par le type d'interrogatives partielles utilisées par le journaliste (antéposition simple, antéposition avec inversion, ou in situ). Le Tableau 5.28 présente un exemple d'échange entre le journaliste (A) et un personnage de maire (B) dans les différentes conditions.

	Condition
A : Les gens peuvent venir vous voir à propos de quels genres de problèmes par exemple?	in situ
A : À propos de quels genres de problèmes les gens peuvent venir vous voir, par exemple?	antéposition simple
A : À propos de quels genres de problèmes peuvent venir vous voir les gens, par exemple?	antéposition avec inversion
B : Hier matin (...) un groupe d'habitants des faubourgs nord est par exemple venu me poser des questions sur l'avancement des travaux en cours dans leur quartier. La réunion d'information qui avait été programmée à ce sujet ne les avait pas complètement satisfaits et ils voulaient des réponses plus précises.	

D'après [THIBERGE, 2018]

TABLEAU 5.28 – Exemple de différence entre interviews (MG1)

Après avoir lu ces textes, les volontaires devaient porter plusieurs jugements sociaux sur le journaliste et estimer par exemple sa richesse, son éducation, ses habitudes de lecture, son âge, sa connaissance du sujet de l'interview, ou encore le degré auquel il semblait détendu pendant l'interaction (9 propositions au total).

Pour toutes les propositions, les interrogatives avec antéposition et inversion ont été liées à des hypothèses très différentes des deux autres types phrastiques. Ainsi, il était supposé que le journaliste était plus riche, plus éduqué, ou encore qu'il lisait plus souvent, quand des interrogatives avec antéposition et inversion étaient utilisées. À l'inverse, il était décrit comme moins « décontracté ».

Ces résultats étaient, déjà à l'époque, nuancés et affectés par l'âge des personnes ayant participé à l'étude. Les volontaires au-delà de 55 ans faisaient les contrastes les plus forts entre antéposition avec inversion d'une part, et d'autre part antéposition simple et in situ. Les résultats étaient aussi nuancés, dans une moindre mesure, par le niveau d'études des volontaires et par leur origine géographique. Ainsi, l'antéposition avec inversion était liée à une probabilité plus haute que le journaliste soit riche chez les volontaires qui avaient écrit être originaires de Paris, comparés aux autres personnes.

Les deux études suivantes (Sections 5.3.2 et 5.3.3) visent à élargir ces résultats. Pour la première (MG2), comme pour l'étude AJT3, c'est une modalité orale de présentation des stimuli qui a été utilisée. Cette étude compare les antépositions avec inversion, les antépositions simples et les phrases in situ. Pour la seconde (MG3), comme pour l'étude AJT4, les interrogatives avec antépositions en *est-ce que* ont été comparées aux antépositions avec inversion et aux phrases in situ.

5.3.2 Interrogatives et projection d'indices sociaux à l'oral (MG2)

Cette étude vient prolonger à la fois l'étude MG1, menée à l'écrit avant le présent travail de thèse, et les études AJT2 (Section 5.2.2) et AJT3 5.2.3. L'idée était de voir si, au-delà de leur conscience linguistique explicite (bonne formation de l'énoncé, adaptation de l'énoncé au contexte), les francophones projettent des considérations sociales sur les personnes utilisant différentes variantes interrogatives. Les matériels utilisés sont parallèles à ceux utilisés dans les jugements d'acceptabilité. Le protocole employé permet en outre une meilleure lecture des résultats obtenus que pour l'étude MG1 (pour laquelle les matériels, des interviews assez longues, étaient plus complexes).

5.3.2.1 Hypothèses et prédictions

L'hypothèse au coeur de cette étude est que la forme d'une interrogative partielle est associée, par les francophones, à différents indices sociaux. Derrière cette hypothèse se trouve l'idée que ces associations entre formes linguistiques et stéréotypes se créent petit à petit au sein d'une communauté linguistique (par un processus décrit par ECKERT [2012] et présenté en Section 2.3.2). Une hypothèse liée est que ces associations sont plus opérantes en situation spontanée qu'à l'écrit, d'où la modalité orale choisie pour présenter les stimuli.

Plus spécifiquement, les interrogatives avec antéposition et inversion sont ici conçues comme adaptées à des contextes d'interaction plus formels que les autres types. Ceci découle des observations de corpus présentées au Chapitre 4 et des études de jugements d'acceptabilité récoltés précédemment (Sections 5.2.1 à 5.2.5).

En conséquence, il était prévu que les interrogatives avec antéposition et inversion soient associées à un profil « sociologique » différent de celui associé aux deux autres (antéposition simple et in situ). La personne utilisant des antépositions avec inversion serait réputée, sous cette perspective, être plus riche, avoir une meilleure éducation ou un plus grand prestige, mais aussi des habitudes

plus littéraires. Le comportement des groupes d'âge observés dans les études de jugements d'acceptabilité laisse aussi penser que les antépositions avec inversion peuvent être associées à un âge plus avancé de la personne qui les emploie.

L'étude MG1 a aussi montré que les interrogatives avec antéposition et inversion sont associées à une moins grande décontraction dans l'échange. Ceci est relié à une moins grande adaptation de ce type phrastique à des situations d'interaction spontanées, comme vu dans l'étude AJT2. Il était donc envisagé qu'en terme de profil, elles soient associées à une personnalité moins sociable.

Enfin, il était attendu une association positive entre l'emploi d'interrogative avec inversion et une origine citadine des personnes les utilisant. Cette prédiction est inspirée par les jugements émis dans l'étude MG1 par les volontaires se décrivant « de Paris ».

Ces préférences générales peuvent être modulées. La condition de CONTEXTE (*formel vs. informel*), intégrée aux études de jugements d'acceptabilité, a été conservée ici. De même, il était envisagé qu'au sein même du groupe des volontaires, une différence pourrait être observé dans les jugements sociaux, en fonction de paramètres internes tel que l'âge.

5.3.2.2 Protocole et items expérimentaux

L'étude a été menée sur le second semestre 2019. Elle a consisté pour les volontaires, francophones, à écouter de courts dialogues, intégralement repris de l'étude AJT3 (5.2.3) (30 items cibles, 30 distracteurs et les entraînements). Contrairement aux jugements d'acceptabilité, les personnes venues passer l'expérience n'ont pas eu à formuler un jugement explicite, métalinguistique, sur la langue utilisée dans les dialogues qu'elles entendaient. Les volontaires devaient essayer de dresser un portrait social d'une des personnes impliquées dans chaque dialogue.

Pour ce faire, six échelles avec chacune 7 niveaux étaient à leur disposition, interrogeant différents aspects en partie repris des échelles utilisées dans l'étude MG1 de THIBERGE [2018]. À propos de la personne entendue dans le dialogue, ces échelles évaluaient : l'âge, la richesse, la sociabilité, l'origine géographique, le niveau d'études, et les loisirs.

La particularité de ces échelles de « localisation sociale » était de ne pas être textuelles (de "jeune" à "agé" par exemple), mais dessinées. Ce choix s'est imposé pour que ces échelles, ou une partie d'entre elles, puissent être réutilisées avec les enfants (voir Chapitre 6, Section 6.2). La Figure 5.29 illustre l'échelle de RICHESSE dans ses versions masculines et féminines. Les dessins des échelles changeaient en effet en fonction des voix utilisées.

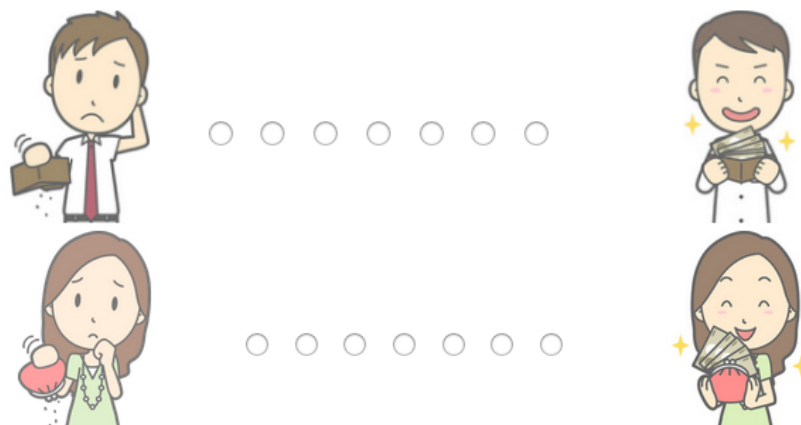


FIGURE 5.29 – Échelles d'évaluation présentée aux volontaires (MG2, RICHESSE)

Ces échelles et les stéréotypes associés aux dessins ont fait l'objet d'un norming préalable pour s'assurer de leur cohésion et de leur valeur interchangeable notamment entre hommes et femmes (Annexe J).

Comme les items sont les mêmes que pour l'étude AJT3, les variables manipulées sont identiques : TYPE syntaxique (*antéposition simple, antéposition avec inversion, in situ*) et CONTEXTE (*formel vs. informel*). Le Tableau 5.29 présente pour rappel la transcription d'un item décliné sous toutes ses conditions.

Type	Contexte	
	INFORMEL	FORMEL
Antéposition simple	A : Elle, elle y sera pas toute seule. B : Qui elle connaîtra ?	A : Elle n'y sera pas toute seule. B : Qui elle connaîtra ?
Antéposition avec inversion	A : Elle, elle y sera pas toute seule. B : Qui connaîtra-t-elle ?	A : Elle n'y sera pas toute seule. B : Qui connaîtra-t-elle ?
In situ	A : Elle, elle y sera pas toute seule. B : Elle connaîtra qui ?	A : Elle n'y sera pas toute seule. B : Elle connaîtra qui ?

TABLEAU 5.29 – Exemple d'item expérimental (AJT3 - MG2)

Chaque volontaire ne voyait qu'une et une seule version de chacun des 30 items cibles, randomisés selon les principes du carré latin (3x2 au sein de chaque liste). La Figure 5.30 montre à quoi ressemblait l'écran d'un ou d'une volontaire pendant la passation, pour un stimulus à voix féminine et pour un stimulus à voix masculine.

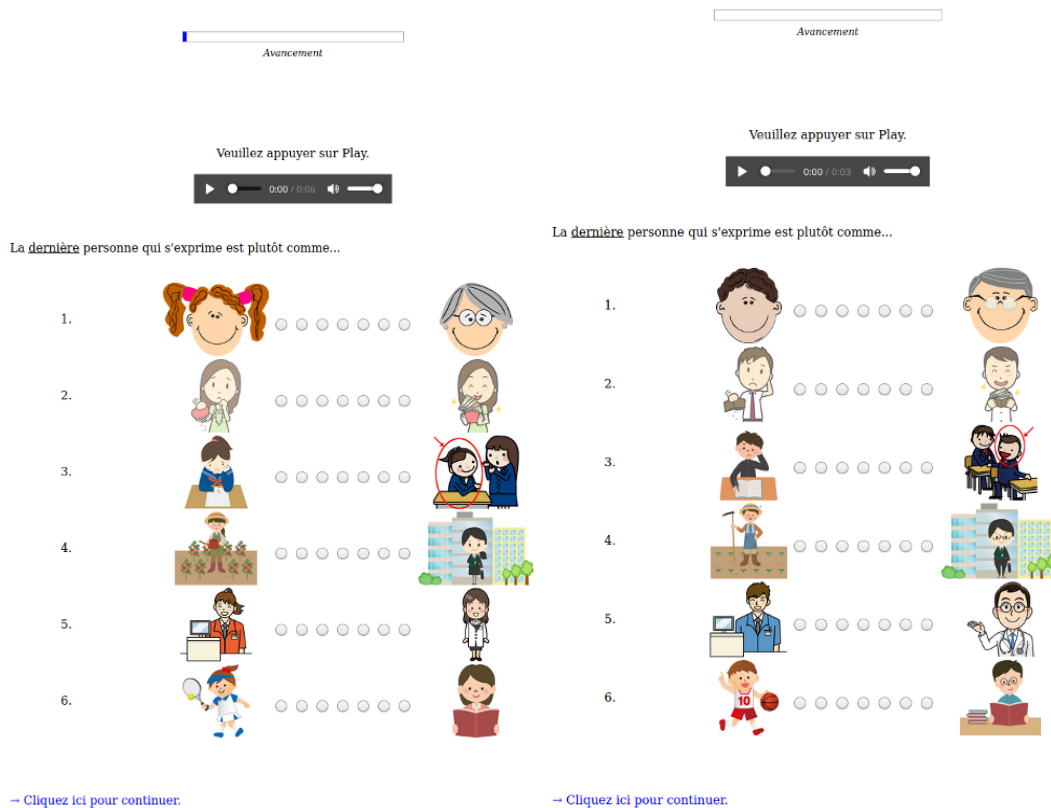


FIGURE 5.30 – Captures d'écran en cours de passation (MG2)

L'étude commençait après une phase de présentation où le but et les consignes de l'expérience étaient rappelées aux volontaires, à la fois verbalement et sur un écran (Annexe K.0.1). Avant de

démarrer l'étude proprement dite, un court diaporama présentant les personnages de chacune des échelles était montré et discuté (voir Annexe K.0.1). Ensuite, il était proposé aux volontaires de se situer eux-mêmes ou elles-mêmes sur un ensemble de six échelles adaptées à leur genre déclaré, sur une feuille de papier anonyme. Cette feuille pouvait être rapprochée des résultats d'une session expérimentale mais pas de l'identité des volontaires. En plus d'un entraînement à la manipulation des échelles, cette étape avait pour but de les laisser s'approprier les graduations de celles-ci.

Il faut ici préciser que, de manière générale et par opposition aux passations en laboratoire menées pour l'étude AJT2, la présente tâche a très souvent déstabilisé les participantes et les participants. Le fait de porter un jugement aussi explicitement sur quelqu'un qui leur était invisible, mais rendu.e réel.le par sa voix, était très difficile. Les phases préalable de présentation de l'étude, d'appropriation des échelles et d'entraînement ont eu une importance majeure. De même, les phases de discussion en fin de passation ont révélé que, si l'enjeu linguistique précis n'avait pas été perçu, le côté inhabituel de la tâche a confronté les volontaires à un comportement quotidien. Il est naturel et automatique d'extrapoler, depuis des informations minimales comme le style vestimentaire ou la manière de parler, des éléments personnels sur les interlocuteurs et les interlocutrices dans une interaction sociale. Ce trouble des personnes ayant participé à l'étude a pu créer du bruit dans les données, et masquer des effets généraux, mais il n'a pas pu créer des effets entre les différents stimuli linguistiques.

5.3.2.3 Participantes et participants

52 personnes ont pris part à l'étude, montée et accessible sur ordinateur via la plateforme IbexFarm [DRUMMOND, 2016], mais passée exclusivement dans une chambre sourde de l'Université de Paris. Ces volontaires provenaient des listes de diffusion du RISC ainsi que de l'université même ou de son voisinage physique ou social (diffusion par affichage public, bouche-à-oreilles, réseaux sociaux). Comme pour l'étude AJT2, la tâche étant longue et la contrainte audio sont deux raisons pour lesquelles les passations ont été organisées au laboratoire dans l'enceinte de l'université. Les volontaires venaient pour environ 30 minutes et recevaient une compensation de 5 euros. 7 passations parmi les 52 ont duré 45 minutes, pour 8 euros de compensation.

Ces personnes avaient entre 18 et 77 ans ($\mu = 26.8$, $m = 22.5$). 36 ont déclaré être de genre féminin et 16 ont déclaré être de genre masculin. Les plus hauts diplômes obtenus par ces personnes étaient :

- en dessous du baccalauréat, N=2
- le baccalauréat, N=18
- un diplôme équivalent à Bac+2, N=8
- un diplôme équivalent à Bac+3, N=15
- un diplôme équivalent à Bac+5, N=8
- un diplôme équivalent à Bac+8, N=1

33 personnes ont indiqué avoir pour région de naissance l'Île de France, 2 un pays étranger (Côte d'Ivoire, Suisse), et 17 une autre circonscription administrative du territoire français métropolitain. 12 personnes ont précisé travailler, étudier ou être à la retraite après avoir travaillé dans un domaine lié aux langues, au langage ou à la linguistique (orthophonie et traduction comprises), 2 dans un domaine lié à l'éducation ou l'enseignement, 2 en psychologie ou psychopédagogie. 3 personnes n'ont pas précisé et 1 personne s'est déclarée sans activité. Le reste des domaines indiqués couvre

l'administration, la santé, la communication, les lettres, le légal ou para-légal, la philosophie, la géologie, la physique-chimie, les mathématiques, ou encore le commerce. Les données sur les langues maîtrisées n'ont pas été traitées, à part pour vérifier la langue maternelle indiquée.

Il y avait six échelles et pour chaque échelle le ou la volontaire devait cliquer sur un niveau, ce qui était rendu obligatoire par la programmation informatique de l'expérience. Les personnes pouvaient cliquer sur la valeur centrale des échelles pour ne pas exprimer d'avis. C'est aussi ce choix qui leur était préconisé en cas d'erreur de chargement du son à écouter. Au total, ce sont donc $52 \times 63 = 3276$ jugements qui ont été recueillis, items cibles, distracteurs et entraînements compris.

Il était rappelé au début de chaque passation et avant signature du formulaire de consentement qu'un retrait de la tâche à tout moment, sans raison et sans perte de contrepartie, était possible.

5.3.2.4 Résultats

Seuls les 1560 jugements reçus sur les items cibles ont été conservés pour analyse ($=52 \times 30$). Ces jugements ont été codés et centrés autour de 0 (valeur médiane des échelles), de -3 (extrémités gauches des échelles : âge-, argent-, sociable-, campagne, éducation-, sport) à +3 (extrémités droites des échelles : âge+, argent+, sociable+, ville, éducation+, lecture). Ces labels sont une description du stéréotype correspondant à la paire de dessins utilisés, sans aucune considération morale.

5.3.2.4.1 Résultats généraux

Une matrice des corrélations (méthode = Spearman, niveau de significativité = 0.05) est visible sur la Figure 5.31. Pour rappel, la couleur des cercles indique si une corrélation positive ou négative, et leur taille est proportionnelle au coefficient de corrélation. Les corrélations non significatives sont indiquées par une croix. Le coefficient de corrélation est la valeur chiffrée indiquée à gauche.



FIGURE 5.31 – Corrélations entre échelles (Spearman) (MG2)

La matrice indique et illustre que, comme attendu, les stéréotypes projetés par des tiers sur

une personne en fonction de ses productions linguistiques sont inter-reliés. C'est un réseau de stéréotypes qui est projeté, et sans doute modifié tout au long de l'interaction en fonction de multiples variables (y compris para- ou extra-linguistiques). Plus de détails statistiques sur les corrélations sont disponibles dans l'Annexe K.0.20.

Il est notamment visible que la plus grande corrélation est celle entre ÉTUDES (ici *longues*) et LOISIRS (ici *lecture*) ($r = .57, p < 0.01$). Dans la lignée de ce résultats, les réponses sur la RICHESSE (*haute*) corrélient avec les ÉTUDES (*longues*) ($r = .46, p < 0.01$) et les LOISIRS (*lecture*) ($r = .40, p < 0.01$). Les réponses sur l'ÂGE (*élevé*) corrélient elle-mêmes avec les réponses sur la RICHESSE (*haute*) ($r = .42, p < 0.01$). Le réseau de stéréotypes semblerait donc fonctionner ainsi : « plus on est âgé, plus on a de chances d'avoir fait des études longues », « plus on a fait des études longues plus on a de chances d'être riche », « plus on a fait des études longues plus on passe de temps à lire plutôt qu'à faire du sport », ou une quelconque combinaison des réciproques de ces relations. Il est intéressant de constater que l'échelle de SOCIABILITÉ, au sens large où ont pu l'intégrer les volontaires, semble difficile à relier aux autres valeurs.

Réponses sur l'ÂGE

Le Tableau 5.30 et la Figure 5.32 donnent les moyennes et les distributions des réponses sur l'échelle d'ÂGE, rendus par le volontaires en fonction des interrogatives présentes dans les phrases écoutées.

Antéposition simple	Antéposition et inversion	In situ
-0.47	-0.10	-0.43

TABLEAU 5.30 – Note moyenne en fonction du type interrogatif (MG2, ÂGE)

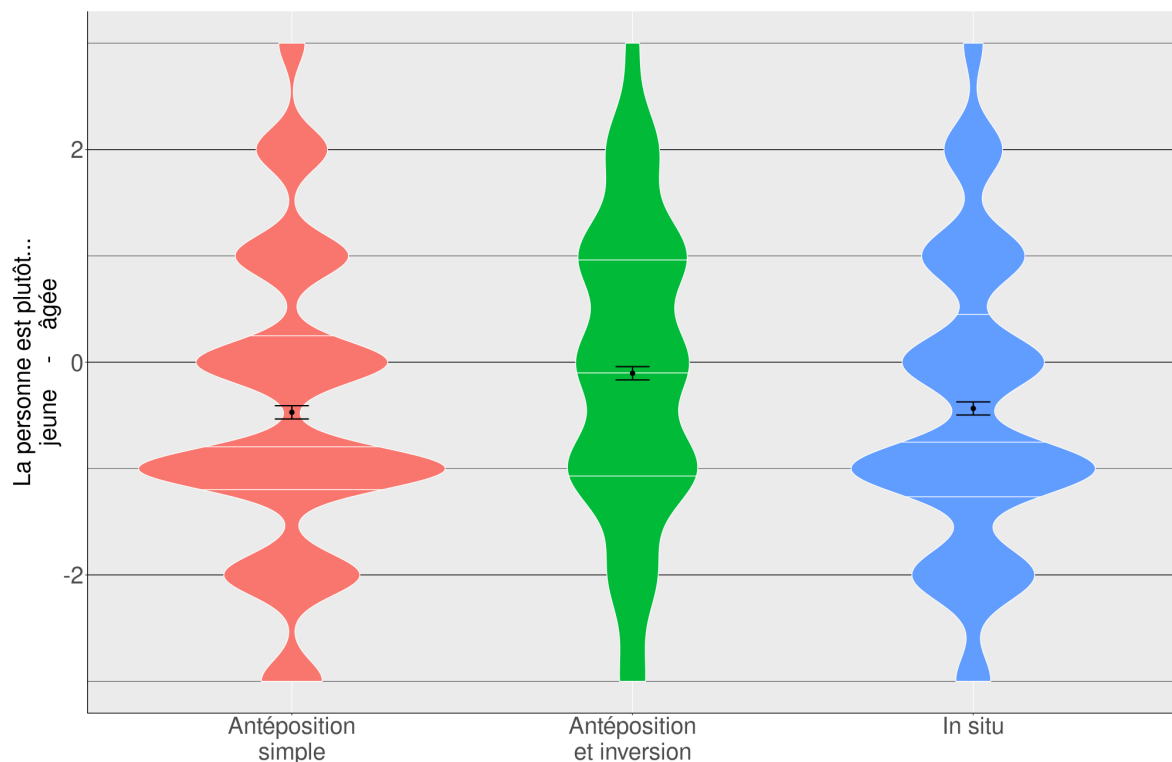


FIGURE 5.32 – Note moyenne en fonction du type interrogatif (MG2, ÂGE)

Une première observation est ici que les scores sont globalement assez bas. C'est cohérent puisque les voix qui ont été utilisées pour construire les stimuli étaient âgées de moins de 30 ans au moment des enregistrements. Néanmoins une différence est déjà visible entre les trois types d'interrogatives et le patron observé dans les jugements d'acceptabilité est ici encore opérant. Les interrogatives avec antéposition et inversion contrastent avec les deux autres types phrastiques, dans des proportions assez semblables. Une personne utilisant les interrogatives avec antéposition simple ou in situ paraît plus jeune que si elle utilise une antéposition avec inversion.

Réponses sur la RICHESSE

Le Tableau 5.31 et la Figure 5.33 donnent les moyennes et la distribution des jugements rendus sur l'échelle de RICHESSE, en fonction du type syntaxique des interrogatives.

Antéposition simple	Antéposition et inversion	In situ
-0.08	0.27	-0.02

TABLEAU 5.31 – Note moyenne en fonction du type interrogatif (MG2, RICHESSE)

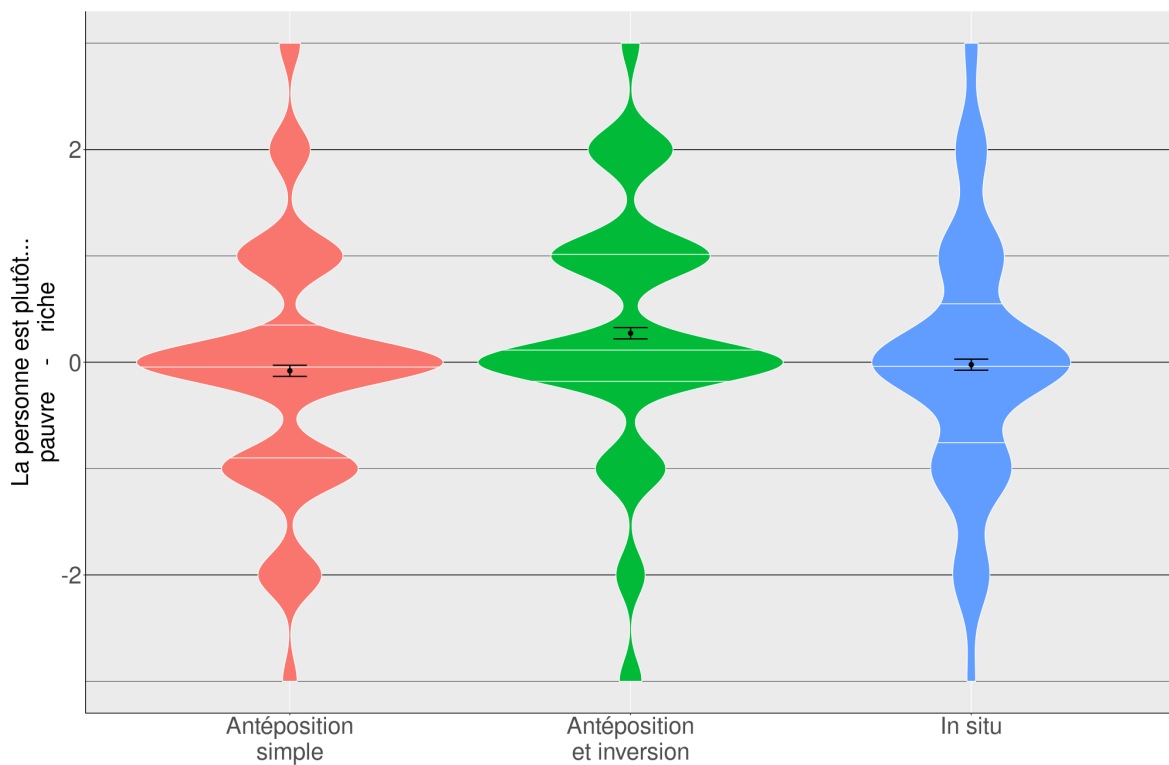


FIGURE 5.33 – Note moyenne en fonction du type interrogatif (MG2, RICHESSE)

L'observation principale est ce même patron qui sépare les interrogatives avec antéposition et inversion des deux autres types phrastiques. En moyenne, une présomption de richesse un peu plus haute pèse sur les personnes qui les utilisent. La visualisation par *violin plot* est intéressante ici : elle met en lumière le fait que, malgré une moyenne voisine autour de zéro, les interrogatives in situ et avec antéposition simple diffèrent. La distribution des jugements est beaucoup plus étalée pour les phrases in situ. Ce peut être l'illustration d'une valeur plus « passe-partout », socialement et en termes de formalité de contexte, de ce type d'interrogative.

Réponses sur la SOCIABILITÉ

Le Tableau 5.32 et la Figure 5.34 donnent les moyennes et la distribution des jugements rendus sur l'échelle de SOCIABILITÉ, en fonction du type interrogatif.

Antéposition simple	Antéposition et inversion	In situ
0.64	0.53	0.86

TABLEAU 5.32 – Note moyenne en fonction du type interrogatif (MG2, SOCIABILITÉ)

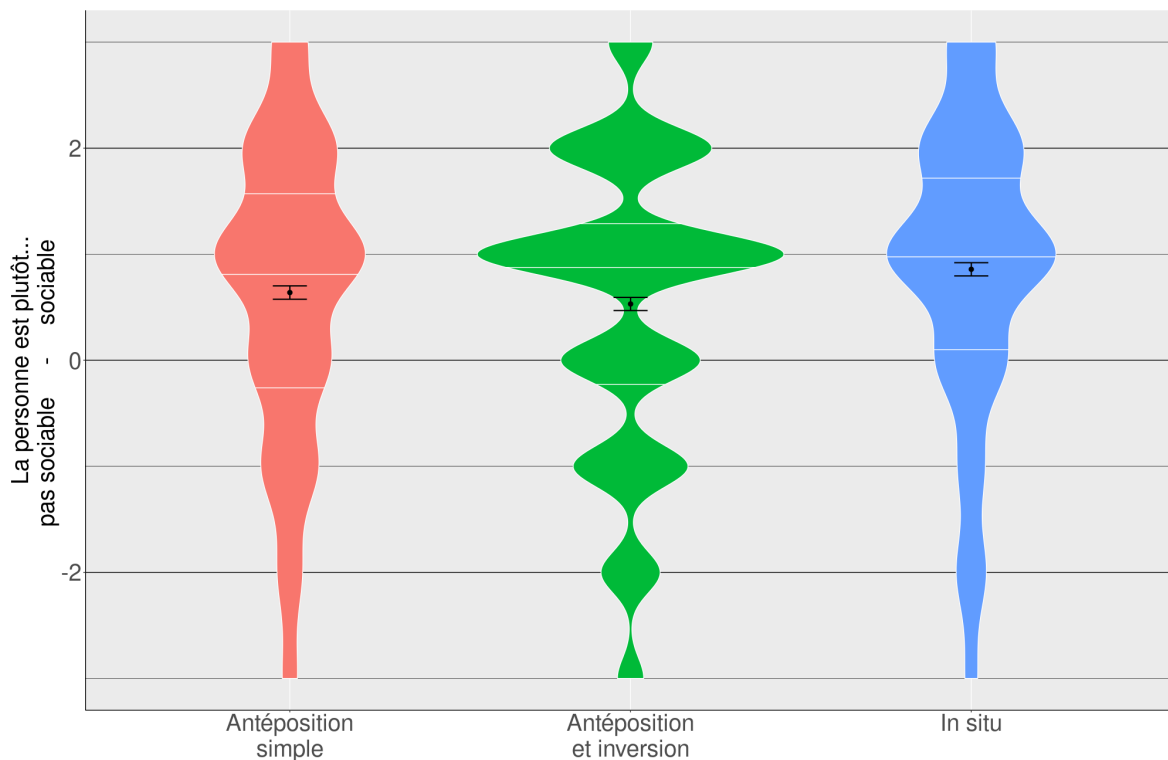


FIGURE 5.34 – Note moyenne en fonction du type interrogatif (MG2, SOCIABILITÉ)

Le patron est un peu différent des deux précédents ici, en ce qu'il n'y a pas vraiment d'opposition bipartite mais plutôt une gradation entre les trois types de phrase. Toutes les interrogatives emportent en moyenne une présomption de *sociabilité* positive. Mais cette présomption est la moins importante pour les antépositions avec inversion, un peu plus haute pour les antépositions simples, et encore plus haute comparativement pour les interrogatives in situ. L'in situ semble donc ici un type associé avec le fait d'avoir une plus grande capacité d'interaction sociale. Ces résultats rappellent ceux de l'étude AJT3 (Section 5.2.3), dans lesquels la forme in situ était la plus adaptée au contexte d'interaction de manière générale, devant les antépositions simples et avec inversion.

Réponses sur l'ORIGINE géographique

Le Tableau 5.33 et la Figure 5.35 donnent les moyennes et la distribution des jugements sur l'échelle de l'ORIGINE GÉOGRAPHIQUE, rendus par les volontaires en fonction des interrogatives écoutées.

Antéposition simple	Antéposition et inversion	In situ
0.51	0.80	0.71

TABLEAU 5.33 – Note moyenne en fonction du type interrogatif (MG2, ORIGINE)

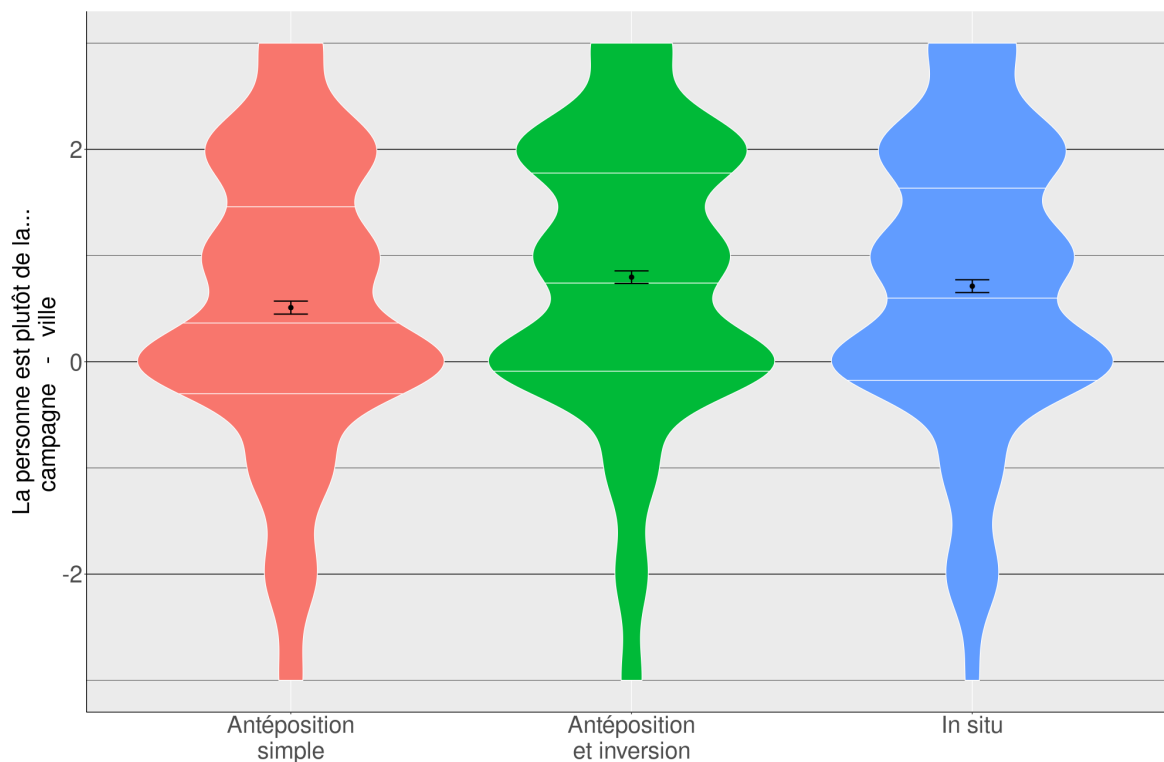


FIGURE 5.35 – Note moyenne en fonction du type interrogatif (MG2, ORIGINE)

Un patron binaire, qui opposerait les antépositions avec inversion aux antépositions simple et aux in situ confondues, n'est ici pas confirmé. Cette fois, s'il y a une opposition, elle semble être entre les antépositions simples d'une part, moins associées en moyenne avec une origine urbaine, et les deux autres types interrogatifs d'autre part.

Réponses sur les ÉTUDES

Le Tableau 5.34 et la Figure 5.36 donnent les moyennes et la distribution des jugements rendus sur l'échelle des ÉTUDES, en fonction des types interrogatifs.

Antéposition simple	Antéposition et inversion	In situ
-0.07	0.44	0.03

TABLEAU 5.34 – Note moyenne en fonction du type interrogatif (MG2, ÉTUDES)

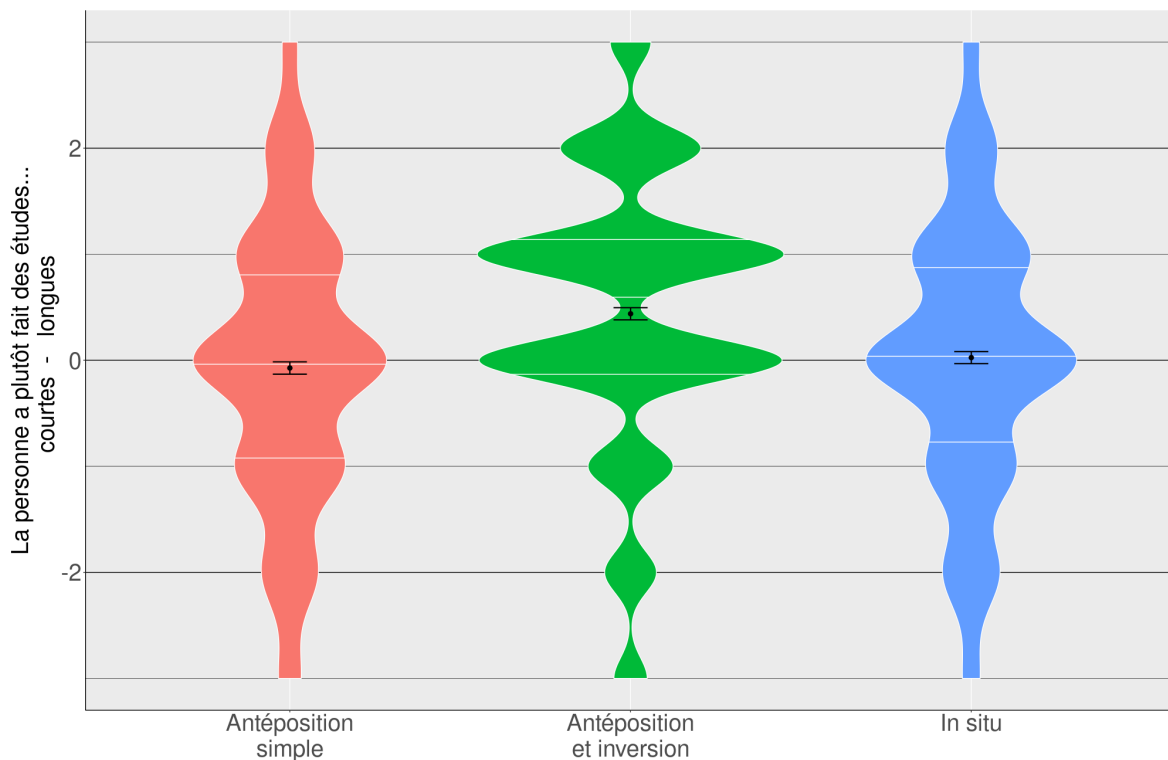


FIGURE 5.36 – Note moyenne en fonction du type interrogatif (MG2, ÉTUDES)

L'opposition est ici franche, et de nouveau binaire. Les interrogatives in situ et avec antéposition simple ne semblent pas vraiment liées à un stéréotype précis, avec une moyenne proche de zéro et une distribution des réponses bien étalée de part et d'autre. En contraste, les interrogatives avec antéposition et inversion ont une distribution plus haute. Ce type phrastique semble associé au stéréotype d'une personne ayant fait de longues études.

Réponses sur les LOISIRS

Le Tableau 5.35 et la Figure 5.37 donnent la distribution des jugements rendus sur l'échelle de LOISIRS, en fonction du type phrastique.

Antéposition simple	Antéposition et inversion	In situ
-0.05	0.61	0.04

TABLEAU 5.35 – Note moyenne en fonction du type interrogatif (MG2, LOISIRS)

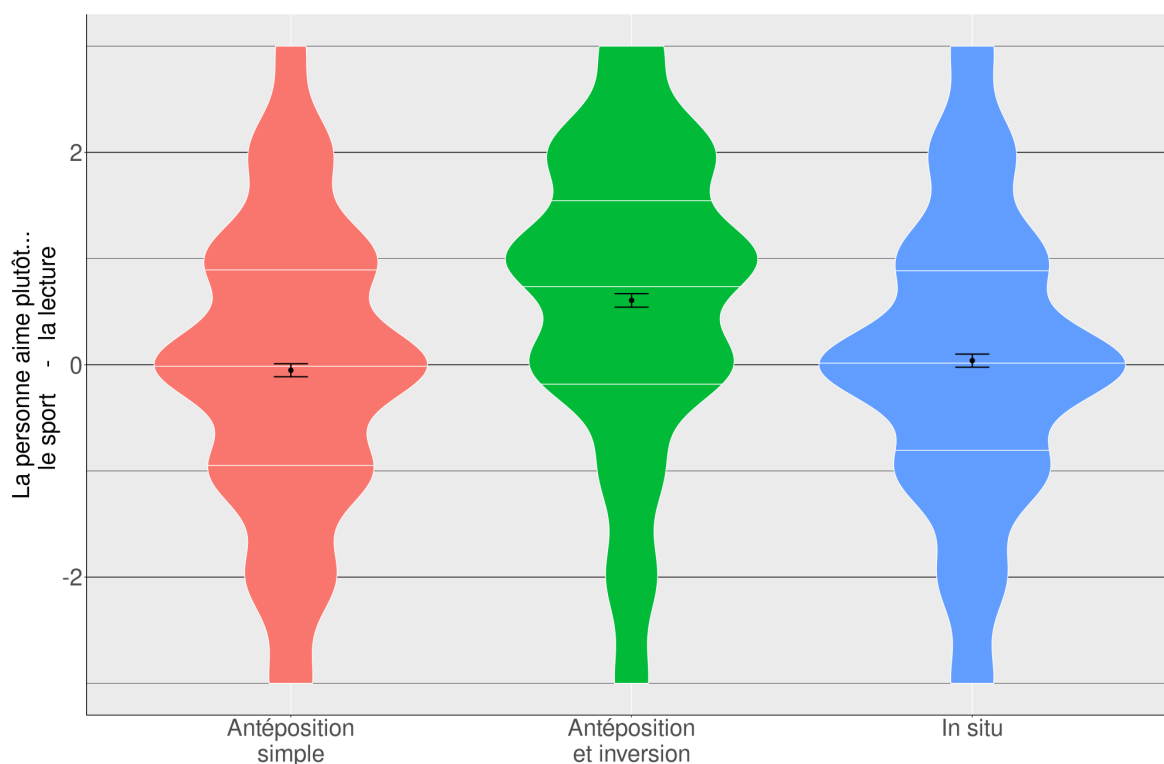


FIGURE 5.37 – Note moyenne en fonction du type interrogatif (MG2, LOISIRS)

Les distributions semblent visuellement un peu différentes, mais les résultats sont très proches du stéréotype précédent. Les interrogatives avec antéposition et inversion se démarquent ici des deux autres types : elles sont en moyenne associées à une pratique plus régulière de la lecture que les deux autres. Ceci traduit la corrélation existant entre les deux échelles (pour rappel, $r = .57$, $p < 0.01$). Il semble cohérent par ailleurs qu'un type phrastique jugé comme le « meilleur français » (AJT2/3/4), et décrit comme le standard linguistique [RIEGEL et al., 2014], soit associé à une fréquence de lecture plus importante.

Bilan

La Figure 5.38 propose une combinaison des résultats précédents et permet une visualisation du « profil stéréotypique » d'une personne employant exclusivement l'un ou l'autre des trois types analysés ici. Les résultats sont projetés sur une échelle allant du négatif au positif et centrée autour de zéro. Les valeurs négatives correspondent aux extrémités gauches des échelles (âge-, argent-, sociable-, campagne, éducation-, sport), les valeurs positives aux extrémités droites (âge+, argent+, sociable+, ville, éducation+, lecture).

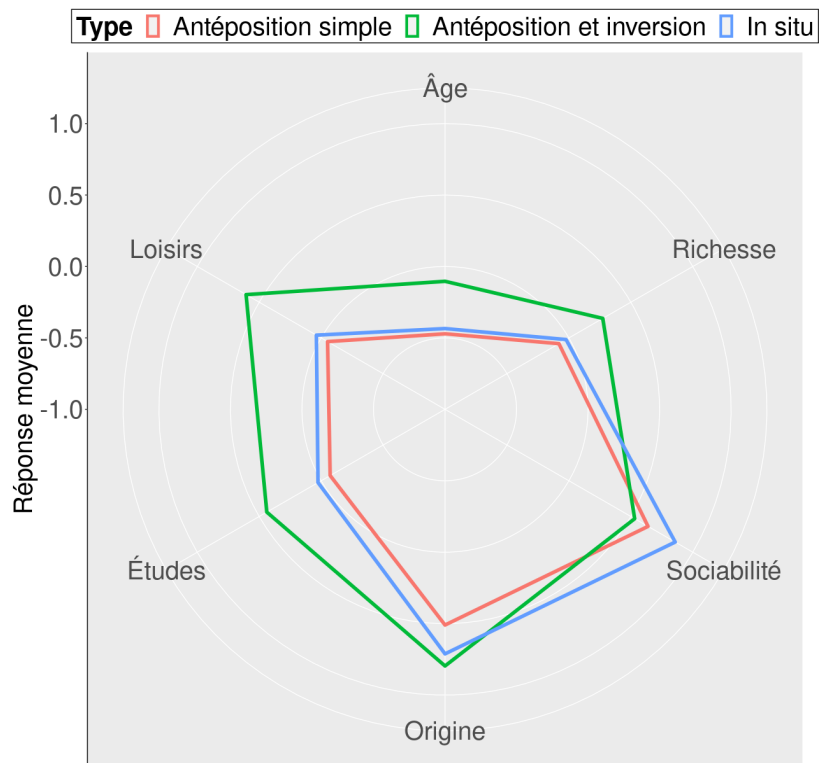


FIGURE 5.38 – Résultats généraux moyens pour toutes les échelles (MG2)

Plus que les valeurs absolues sur chacune des échelles, c'est la comparaison globale des lignes correspondant aux trois types interrogatifs qui est d'intérêt. Elle fait apparaître que les antépositions avec inversion sont associées avec le profil social d'une personne plus âgée, plus riche, ayant fait de plus longues études et lisant plus que les deux autres types. Sur l'échelle de l'origine géographique, les antépositions avec inversion et les phrases in situ sont proches les unes des autres et plus associées avec une origine urbaine que les antépositions simples. Sur la sociabilité, la différence entre types interrogatifs est presque inverse, avec les phrases in situ, et même les antépositions simples, associées à plus de sociabilité que pour les antépositions avec inversion.

Au-delà de ces résultats agrégés pour toutes les personnes ayant participé à l'étude, il faut garder en tête que les profils individuels sont parfois très différents, ce qu'illustre l'Annexe K.0.21.

5.3.2.4.2 Effets d'âge

Le détail des effets liés à l'âge des volontaires dans leur appréciation sociale des stimuli est disponible, pour les six échelles, dans l'Annexe K.0.22. Leur portée est plus limitée que dans les études précédentes de par le petit nombre d'individus à l'âge supérieur à 30 ans ($N=10/52$). Par cohérence, c'est néanmoins un facteur qui a été analysé et intégré en tant que variable indépendante dans les modèles statistiques bayésiens décrits plus bas.

De manière générale, sur les données recueillies, le groupe d'âge a un effet principal réduit dans le sens où, indépendamment du type interrogatif, les jugements des plus de 30 ans diffèrent peu en moyenne des jugements des moins de 30 ans, excepté peut-être pour les jugements sur les loisirs (où le choix des plus de 30 ans semble plus souvent orienté vers le stéréotype de la lecture).

Considéré en interaction avec le type d'interrogative, les influences de l'âge des volontaires sont plus visibles. La Figure 5.39 est une compilation de ces résultats modulés par l'âge des volontaires. De nouveau, le « profil stéréotypique » associé aux trois types interrogatifs examinés est illustré,

pour chacun des deux groupes d'âge. Il s'en dégage principalement l'idée que, chez les moins de 30 ans, les antépositions avec inversion sont associées à des stéréotypes bien distincts des deux autres types (âge+, richesse+, études+, loisirs=lecture). C'est un peu moins distinct pour le stéréotype de l'origine urbaine. C'est l'inverse pour la sociabilité, avec les phrases in situ qui se distinguent des deux formes avec antéposition. Chez les plus de 30 ans, la figure apparaît plus nuancée dans l'ensemble. Malgré cela, pour ce groupe, sur l'échelle des loisirs, l'antéposition simple est bien moins associée à la lecture que les deux autres types. Sur l'échelle de sociabilité, l'antéposition avec inversion s'oppose aux deux autres types, plutôt proches l'un de l'autre. Et enfin, sur l'échelle de richesse, ce sont les phrases in situ qui semblent les plus associées à une haute richesse.

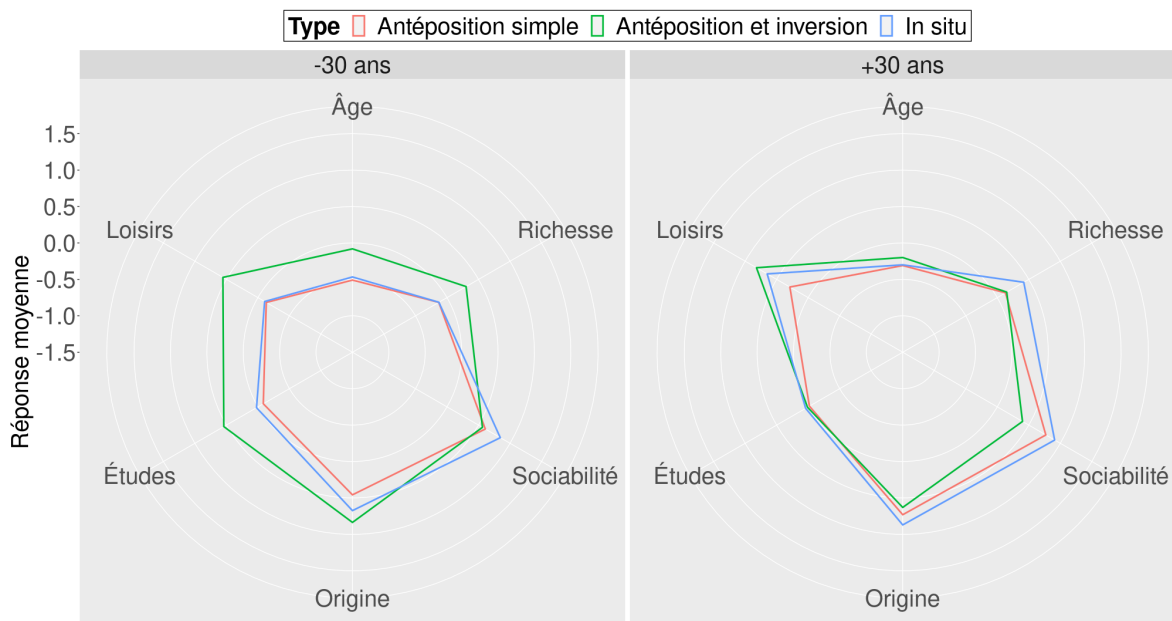


FIGURE 5.39 – Résultats moyens pour toutes les échelles en fonction de l'âge (MG2)

Sur le stéréotype de la RICHESSE par exemple, les moins de 30 ans associent beaucoup plus les antépositions avec inversion que les deux autres types à une richesse élevée (sur une échelle de -3 à +3, moyenne = +0.30 contre -0.13 pour les antépositions simples comme pour les phrases in situ). Les plus de 30 ans associent principalement l'in situ à la richesse haute (+0.42 en moyenne, contre +0.15 pour les antépositions avec inversion et +0.13 pour les antépositions simples).

Sur l'échelle de la SOCIABILITÉ, là encore les groupes d'âge se distinguent. Les moins de 30 ans semblent associer seulement l'in situ à une plus forte sociabilité (sur une échelle de -3 à +3, moyenne = +0.85, contre +0.61 pour les antépositions simples et +0.56 pour les antépositions avec inversion). Pour les plus de 30 ans, l'opposition se fait entre les antépositions avec inversion (moyenne = +0.40) et les deux autres types (+0.77 pour les antépositions simples et +0.91 pour les in situ, soit une plus forte sociabilité).

Sur l'échelle de l'ORIGINE géographique, les moins de 30 ans semblent faire une distinction entre in situ et antépositions avec inversion, qui seraient associées à une plus probable origine urbaine (sur une échelle de -3 à +3, moyennes respectives à +0.67 et +0.84, avec une distribution haute des réponses, contre +0.46 pour les antépositions simples, avec une distribution des réponses centrée autour de 0). Les plus de 30 ans sont plus nuancés avec des moyennes respectives à +0.87 pour les in situ, +0.63 pour les antépositions avec inversion, et +0.73 pour les antépositions simples.

Sur les échelles d'ÂGE et des ÉTUDES, une distinction binaire se fait chez les moins de 30 ans. L'antéposition avec inversion est associée à des jugements plus hauts que pour les deux autres types, soit un âge plus élevé et des études plus longues. Chez les plus de 30 ans cependant, cette distinction se lisse pour n'avoir plus qu'une distinction réduite entre les trois types.

Enfin, concernant l'échelle des LOISIRS, il a déjà été mentionné que les jugements y sont en moyenne plus hauts pour les plus de 30 ans que pour les moins de 30 ans, tous types interrogatifs confondus. Il y a aussi un changement dans le patron observé, avec un creusement de l'écart entre les antépositions simples et les autres structures. Les moins de 30 ans font en effet une distinction entre d'une part les antépositions avec inversion, plus associées à la lecture (sur une échelle de -3 à +3, moyenne à +0.55), et d'autre part les antépositions simples (-0.13) et les in situ (-0.11). Si les plus de 30 ans associent eux aussi le plus l'antéposition avec inversion à la lecture (moyenne = +0.82), c'est surtout l'antéposition simple qui semble vraiment à l'écart et y être moins associée (+0.29), alors que la forme in situ (moyenne = +0.65) serait plus proche de l'antéposition avec inversion.

5.3.2.4.3 Effets de contexte

Les potentiels effets du contexte (formel vs. informel) sont plus robustes que les effets d'âge puisque c'était une variable aux niveaux contrebalancés à travers les items pour chaque participant. De manière générale (voir Annexe K.0.23 pour le détail), ce n'est pas un prédicteur qui a eu un effet seul. C'est peut-être le cas pour les échelles d'ÂGE et des ÉTUDES, sur lesquelles il semble que les jugements ont globalement été plus élevés en contexte formel qu'en contexte informel. Considéré en interaction avec le type d'interrogative, quelques faits sont notables. Le contexte d'interaction semble pouvoir influencer sur les contrastes entre types interrogatifs.

La figure 5.40 trace ainsi le profil stéréotypique associé aux variations de contexte (gauche = formel; droite = informel). Ici encore, ce sont plus les écarts entre les tracés de chacun des types interrogatifs qui sont d'intérêt principal.

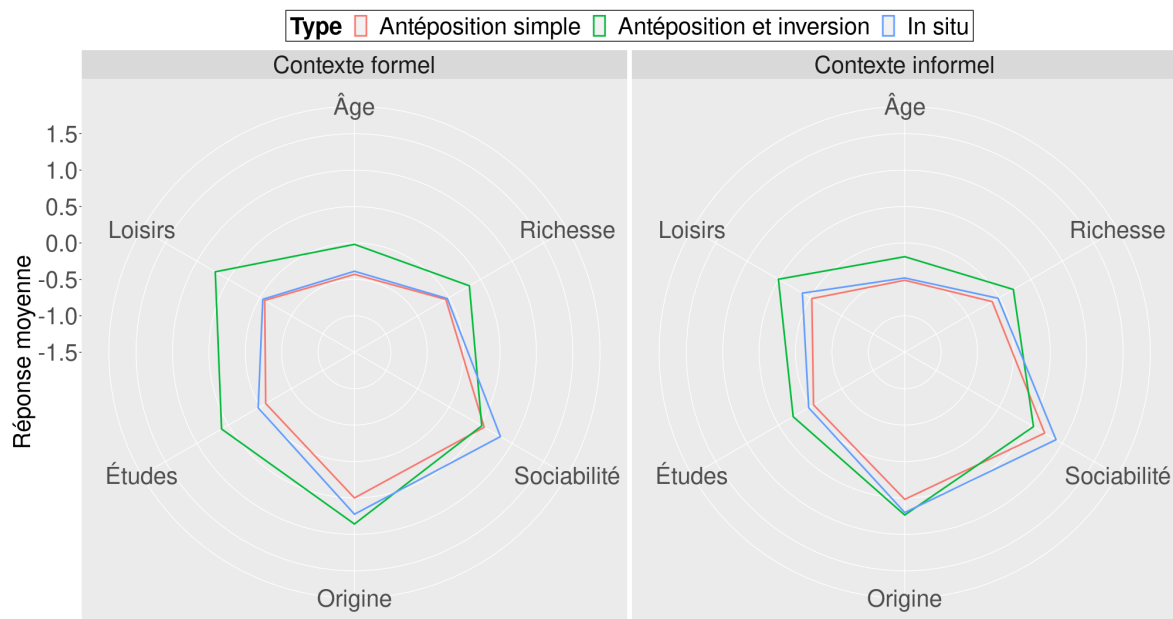


FIGURE 5.40 – Résultats moyens pour toutes les échelles en fonction du contexte (MG2)

De manière générale, il ressort de ces visualisations que l'antéposition avec inversion est plus

distincte des autres types phrastiques en contexte formel. Elle y est alors associée à des jugements plus hauts notamment sur la question de l'âge, de la richesse, des études et des loisirs. En contexte informel, l'écart est moindre avec l'antéposition simple et le type in situ.

Pour l'échelle de la *SOCIABILITÉ*, l'écart entre antéposition simple et antéposition avec inversion est quasiment neutralisé en contexte formel (sur une échelle de -3 à +3, moyennes respectives à +0.56 et +0.52), alors qu'il se creuse en contexte informel (respectivement +0.72 et +0.54). Ce sont plutôt les contextes informels qui semblent améliorer les jugements associés aux interrogatives avec antéposition simple.

Sur trois autres échelles, à savoir l'*ORIGINE*, les *ÉTUDES* et les *LOISIRS*, un contexte plus formel tend à amplifier les contrastes entre types. Par exemple pour l'origine géographique, l'écart entre antéposition simple et antéposition avec inversion se creuse (sur une échelle de -3 à +3, moyennes pour les antépositions simples = +0.52 en contexte informel, +0.50 en contexte formel; moyennes pour les antépositions avec inversion = +0.73 en contexte informel, +0.86 en contexte formel). La différence est ici due à une augmentation des jugements moyens associés à l'antéposition et inversion. C'est donc sur celle-ci que semble plutôt peser une présomption plus haute d'origine urbaine.

Pour ce qui est des échelles des *ÉTUDES* et des *LOISIRS*, le patron est assez similaire dans les deux cas. Il y a un contraste amplifié en contexte formel entre, d'une part, l'antéposition simple et l'in situ, et d'autre part l'antéposition avec inversion. Ceci traduit une augmentation des jugements rendus sur ces deux échelles en contexte plus formel. sur les antépositions avec inversion. Pour les études, sur une échelle de -3 à +3, la moyenne en contexte informel est à -0.05 pour les antépositions simples (-0.09 en contexte formel) et +0.02 pour les in situ (+0.03 en contexte formel). Par contraste, la moyenne pour les antépositions avec inversion est de +0.27 en contexte informel, contre +0.61 en contexte plus formel. Pour les loisirs, sur une échelle de -3 (sport) à +3 (lecture), la moyenne en contexte informel est à -0.03 pour les antépositions simples (-0.07 en contexte formel) et +0.12 pour les in situ (-0.05 en contexte formel). Par contraste, la moyenne pour les antépositions avec inversion sur cette échelle est de +0.50 en contexte informel, contre +0.71 en contexte plus formel. Ici, ce sont donc les antépositions avec inversion qui sont plus associées avec des études longues et une tendance à la lecture plus importante, en contexte formel.

5.3.2.4.4 Effets concomitants d'âge et de contexte?

L'influence combinée des variables groupe d'ÂGE des volontaires et formalité du CONTEXTE doit elle aussi être considérée avec prudence. Ceci tient, encore, au faible nombre d'individus dans le groupe des plus de 30 ans (N=10/52). Le détail numérique des résultats la concernant est disponible en Annexe K.0.24, où les barres d'erreur plus grandes pour les plus de 30 ans illustrent à nouveau leur moindre robustesse. C'est néanmoins une interaction qui a elle aussi été intégrée dans les modèles bayésiens décrits plus bas.

La figure 5.41 trace les profils stéréotypiques associés à la fois au groupe d'âge mais aussi aux deux niveaux de la variable CONTEXTE. Alors que le profil type ne semble pas considérablement varier entre contextes pour le groupe des moins de 30 ans, les différences sont plus apparentes pour le groupe des plus de 30 ans.

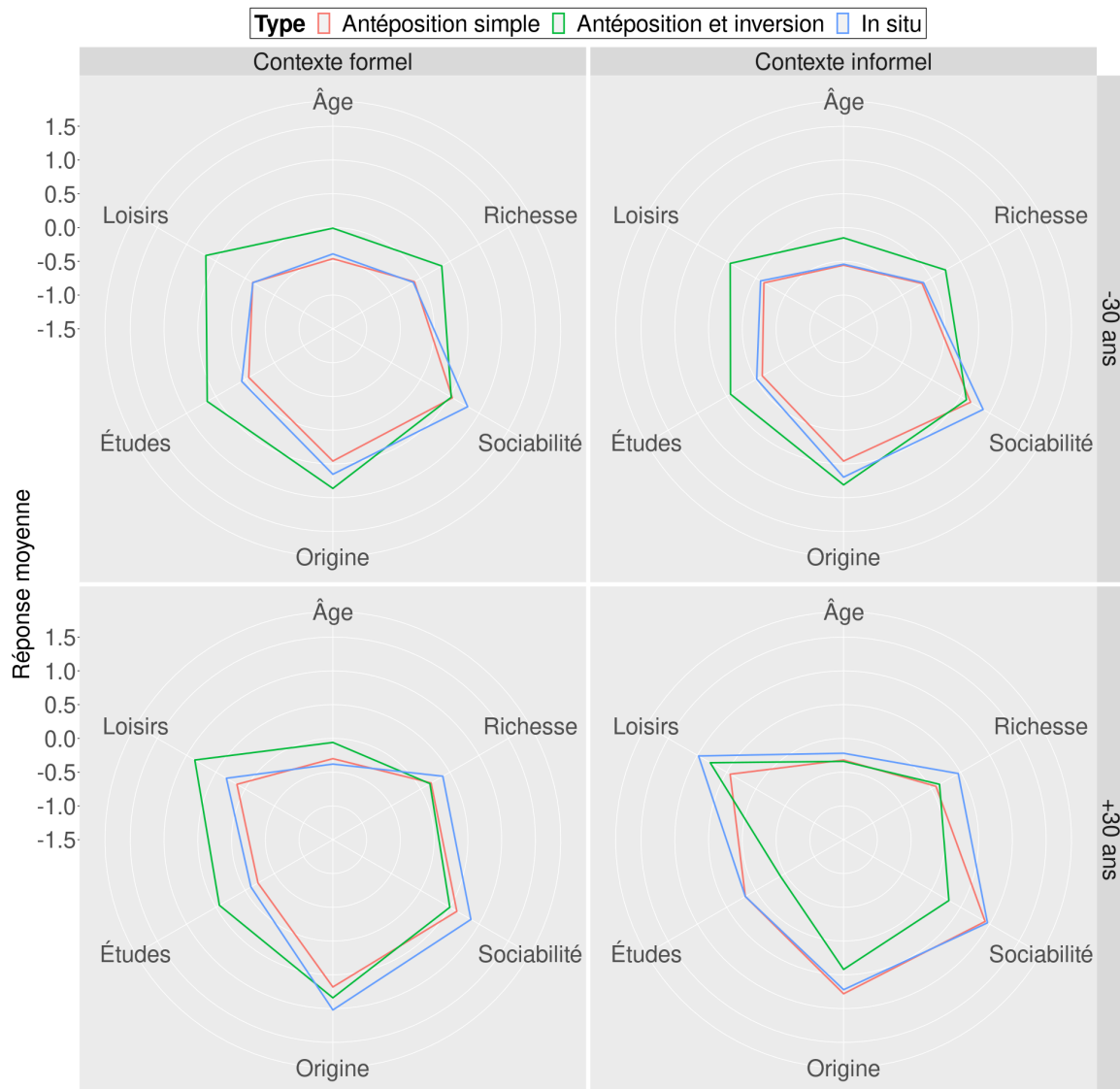


FIGURE 5.41 – Résultats moyens pour toutes les échelles en fonction du contexte et de l'âge (MG2)

Les résultats montrent une relative indifférence générale des moins de 30 ans à la formalité du contexte, sauf pour les antépositions avec inversion sur l'échelle des ÉTUDES (sur une échelle de -3 à +3, moyenne = +0.43 en contexte informel contre +0.65 en contexte formel), et sur l'échelle des LOISIRS (moyenne = +0.44 en contexte informel contre +0.67 en contexte formel).

Chez les plus de 30 ans, les choses semblent un peu différentes. Par exemple, les résultats sur l'échelle d'ÂGE semble bien différer entre contextes. Les phrases in situ y sont liées à des scores plus hauts en contexte informel que les phrases avec antéposition et inversion (sur une échelle de -3 à +3, moyennes respectives à -0.22 et -0.34), alors que c'est l'inverse en contexte formel (respectivement, moyennes à -0.38 et -0.06).

Sur l'échelle de SOCIABILITÉ, les plus de 30 ans semblent contraster antépositions simples et antépositions avec inversion différemment selon le contexte. Les phrases à antéposition simple sont liées à des scores plus hauts en contexte informel que les phrases avec antéposition et inversion (sur une échelle de -3 à +3, moyennes respectives à +0.92 et +0.30). En contexte formel, la différence est bien plus nuancée (respectivement, moyennes à +0.62 et +0.50).

Sur l'échelle de l'ORIGINE géographique, les plus de 30 ans contrastent aussi antéposition

simple et antéposition avec inversion différemment selon le contexte. Ici, ce sont les antépositions avec inversion qui sont liées à des scores bien plus bas en contexte informel (sur une échelle de -3 à +3, moyenne à +0.42) qu'en contexte formel (moyenne = +0.84). En parallèle, les antépositions simples sont liées à des scores légèrement plus élevés en contexte informel (moyenne = +0.78) qu'en contexte formel (+ 0.68).

Sur l'échelle des ÉTUDES, l'antéposition avec inversion semble chez les plus de 30 ans radicalement changer la perception d'un individu selon que le contexte soit informel (sur une échelle de -3 à +3, moyenne = -0.42) ou formel (moyenne = +0.44).

Enfin, un écart entre contextes existe aussi chez les plus de 30 ans sur l'échelle des LOISIRS. Pour ce groupe, les phrases in situ semblent emporter une présomption de lecture beaucoup plus importante en contexte informel (sur une échelle de -3 à +3, moyenne = +0.98) qu'en contexte plus formel (moyenne = +0.32). Il semble donc que pour ce groupe d'âge il y ait une sorte d'inadéquation entre phrases in situ et lecture régulière, qui ne serait révélée que dans les contextes formels.

5.3.2.4.5 Analyses statistiques bayésiennes

Des modèles de régression bayésienne ont été calculés pour étudier le caractère généralisable des données. Un modèle différent a été lancé pour chacune des six échelles de notation. À chaque fois le modèle avait pour variable dépendante la note attribuée, et en variable indépendante l'interaction triple entre trois effets fixes : le TYPE syntaxique (*antéposition simple, antéposition avec inversion, in situ*, avec l'antéposition avec inversion comme niveau de référence), le groupe d'ÂGE (moins ou plus de 30 ans, avec le groupe de -30 ans pour référence), et le CONTEXTE (*formel vs. informel*, avec informel pour niveau de référence). Des effets aléatoires ont été ajoutés pour tenir compte de la variabilité entre individus ("ID", avec pour pente l'interaction *type*contexte*) et de la variabilité entre items ("Item", avec pour pente l'interaction triple *type*âge*contexte*). La variable dépendante étant ordinale, c'est la famille "cumulative" qui a été sélectionnée pour faire tourner ces modèles, avec 4 chaînes et 6000 itérations par chaîne. La convergence des modèles a été atteinte et vérifiée (Rhats=1.00 pour tous les paramètres).

La description complète des modèles calculés est disponible en Annexe [K.0.25](#) et suivantes.

Modèle pour l'échelle d'ÂGE

La Figure [5.42](#) rend compte des distributions postérieures du premier modèle, dont la variable dépendante est ici le score obtenu sur l'échelle ÂGE. Le détail en est disponible dans l'Annexe [K.0.25](#).

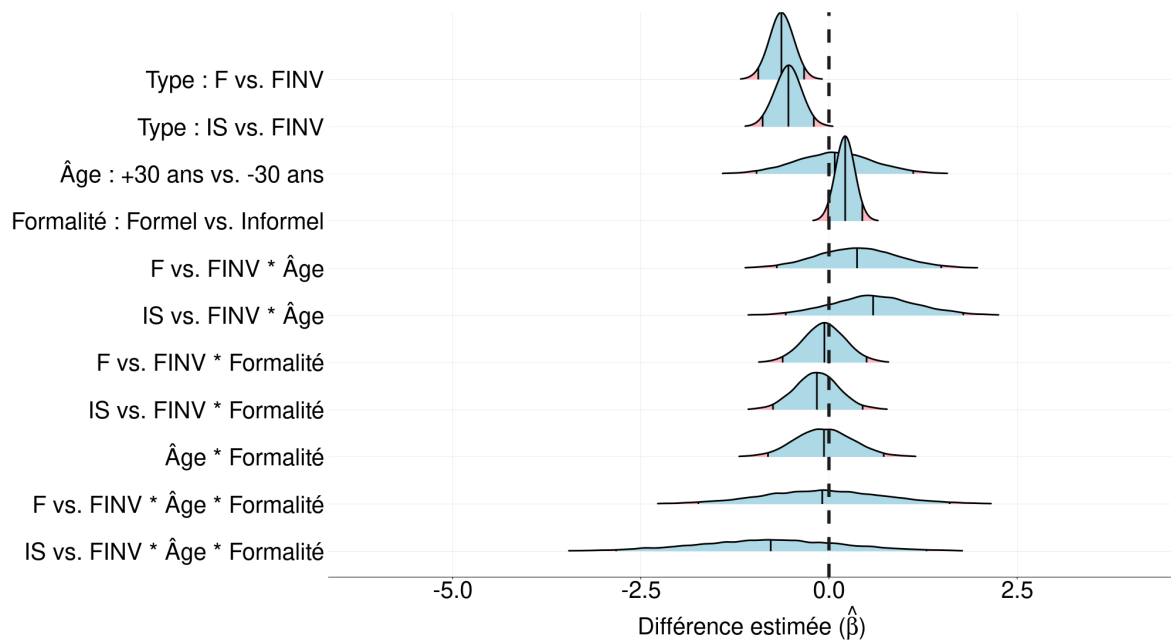


FIGURE 5.42 – Distributions postérieures pour le modèle brms (MG2, ÂGE)

Les principales observations sont ici une confirmation des effets principaux de TYPE et de CONTEXTE. Comparés aux jugements associés aux antépositions avec inversion, les jugements sont en moyenne moins élevés pour les antépositions simples ($\hat{\beta} = -0.63$, 95%CrI=[-0.94,-0.33], $P(\beta) < 0=1$) et pour les phrases in situ ($\hat{\beta} = -0.54$, 95%CrI=[-0.88,-0.20], $P(\beta) < 0=1$). Par ailleurs, les jugements rendus en contexte formel sont plus élevés que les jugements rendus en contexte informel ($\hat{\beta} = 0.08$, 95%CrI=[-0.01,0.44], $P(\beta) > 0=0.97$).

De légères interactions TYPE*ÂGE sont aussi visibles. Elles traduisent le fait que, chez les plus de 30 ans, l'écart avec les phrases à antéposition et inversion tend à se réduire, autant pour les antépositions simples ($\hat{\beta} = 0.38$, 95%CrI=[-0.69,1.49], $P(\beta) > 0=0.75$) que pour les phrases in situ ($\hat{\beta} = 0.60$, 95%CrI=[-0.57,1.78], $P(\beta) > 0=0.85$).

Enfin, une légère interaction triple *type*âge*contexte* est visible pour les phrases in situ ($\hat{\beta} = -0.77$, 95%CrI=[-2.83,1.30], $P(\beta) < 0=0.78$). Ceci est lié au fait que, chez les plus de 30 ans, les phrases in situ sont liées à des jugements plus élevés que les phrases avec antéposition et inversion en contexte informel, alors qu'en contexte formel la situation est inverse. Chez les moins de 30 ans, ce renversement ne se produit pas.

Modèle pour l'échelle de RICHESSE

La Figure 5.43 rend compte des distributions postérieures pour le deuxième modèle lancé, avec pour variable dépendante les jugements rendus sur l'échelle de RICHESSE. Le détail en est disponible dans l'Annexe K.0.26.

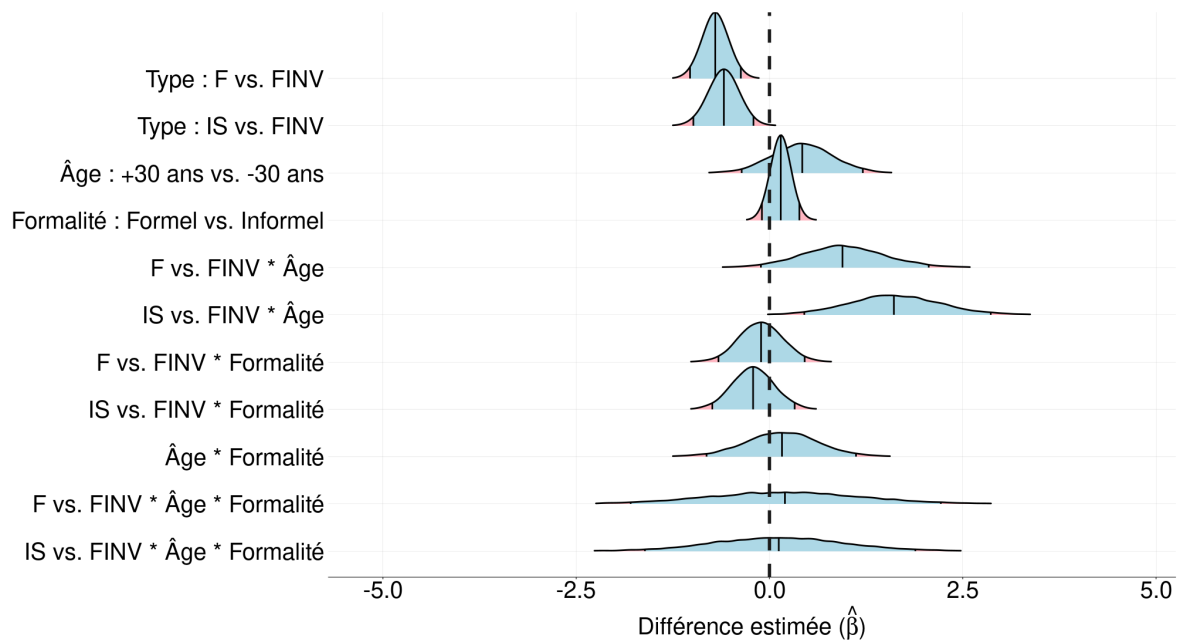


FIGURE 5.43 – Distributions postérieures pour le modèle brms (MG2, RICHESSE)

Sur cette échelle, les effets simples des variables TYPE et CONTEXTE sont assez apparents. Comparés aux jugements associés aux antépositions avec inversion, les jugements sont en moyenne moins élevés pour les antépositions simples ($\hat{\beta} = -0.70$, 95%CrI=[-1.03,-0.37], $P(\beta) < 0=1$) comme pour les phrases in situ ($\hat{\beta} = -0.59$, 95%CrI=[-0.98,-0.21], $P(\beta) < 0=1$). Par ailleurs, des jugements plus élevés sont de manière générale rendus en contexte formel ($\hat{\beta} = 0.15$, 95%CrI=[-0.10,0.39], $P(\beta) > 0=0.88$). Un effet simple du facteur groupe d'ÂGE des volontaires est aussi corroboré, avec des jugements globalement plus hauts rendus par les plus de 30 ans ($\hat{\beta} = 0.42$, 95%CrI=[-0.36,1.21], $P(\beta) > 0=0.86$).

Cet effet simple de la variable ÂGE se traduit en fait par des interactions TYPE*ÂGE. En les comparant aux antépositions avec inversion, c'est à la fois le cas pour les phrases avec antéposition simple ($\hat{\beta} = 0.96$, 95%CrI=[-0.11,2.06], $P(\beta) > 0=0.96$) et pour les phrases in situ ($\hat{\beta} = 1.62$, 95%CrI=[0.45,2.86], $P(\beta) > 0=1$). Ceci traduit la réduction de l'écart, chez les plus de 30 ans par rapport aux moins de 30 ans, entre les jugements associés aux antépositions simples, mais aussi aux phrases in situ, et ceux associés aux antépositions avec inversion.

Peu visible graphiquement, une légère interaction TYPE*CONTEXTE est à noter pour les interrogatives in situ ($\hat{\beta} = -0.21$, 95%CrI=[-0.74,0.33], $P(\beta) < 0=0.79$). Elle traduit le fait que, lorsque le contexte est formel, les jugements associés aux interrogatives in situ ont tendance à être moins proches de ceux associés à des interrogatives avec antéposition et inversion.

Modèle pour l'échelle de SOCIABILITÉ

Pour ce qui est de l'échelle de SOCIABILITÉ, les effets décrits plus hauts étaient assez contrastés et le modèle calculé le confirme. La Figure 5.44 en illustre les distributions postérieures. Le détail en est disponible dans l'Annexe K.0.27.

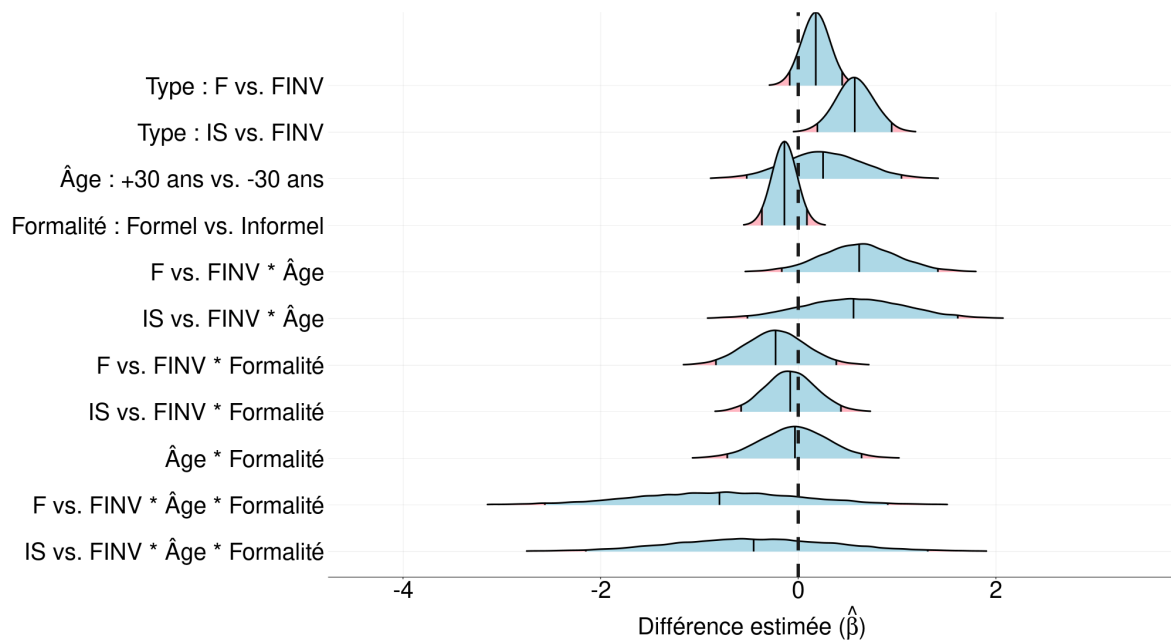


FIGURE 5.44 – Distributions postérieures pour le modèle brms (MG2, SOCIABILITÉ)

Le TYPE syntaxique est toujours un effet principal même pris en isolation. Il est cette fois positif autant pour les antépositions simples ($\hat{\beta} = 0.18$, 95%CrI=[-0.09,0.44], $P(\beta) > 0 = 0.90$) que pour les phrases in situ ($\hat{\beta} = 0.57$, 95%CrI=[0.19,0.95], $P(\beta) > 0 = 1$). Ceci traduit que les jugements rendus sur la sociabilité des personnes sont globalement plus élevés pour les deux autres types interrogatifs que pour les antépositions avec inversion.

Un autre effet simple assez visible est celui de CONTEXTE, avec des jugements moins élevés en contexte formel qu'en contexte informel ($\hat{\beta} = -0.14$, 95%CrI=[-0.37,0.09], $P(\beta) < 0 = 0.89$). L'effet simple du groupe d'âge est plus une tendance, avec des jugements globalement un peu plus hauts rendus par les plus de 30 ans ($\hat{\beta} = 0.26$, 95%CrI=[-0.52,1.04], $P(\beta) > 0 = 0.74$).

Ces effets se retrouvent dans des interactions TYPE*ÂGE. Elles traduisent le fait que, chez les plus de 30 ans, l'écart avec les phrases à antéposition et inversion tend à se réduire, autant pour les antépositions simples ($\hat{\beta} = 0.62$, 95%CrI=[-0.17,1.51], $P(\beta) > 0 = 0.94$) que pour les phrases in situ ($\hat{\beta} = 0.56$, 95%CrI=[-0.51,1.61], $P(\beta) > 0 = 0.85$).

Une interaction TYPE*CONTEXTE et une interaction triple TYPE*CONTEXTE*ÂGE sont aussi visibles pour les phrases à antéposition simple. La première traduit la tendance pour les antépositions simples, lorsque le contexte est formel, à être liées à des jugements moins éloignés de ceux rendus pour les antépositions avec inversion ($\hat{\beta} = -0.23$, 95%CrI=[-0.83,0.39], $P(\beta) < 0 = 0.77$). La seconde traduit que ce contraste entre antéposition avec inversion et antéposition simple, dépendamment de la formalité du contexte, est en fait plus manifeste chez les plus de 30 ans que chez les moins de 30 ans ($\hat{\beta} = -0.80$, 95%CrI=[-2.56,0.91], $P(\beta) < 0 = 0.82$).

Modèle pour l'échelle d'ORIGINE

La Figure 5.45 illustre les distributions postérieures du quatrième modèle calculé, avec pour variable dépendante les jugements rendus sur l'échelle visant l'origine géographique des personnes écoutées. Le détail en est disponible dans l'Annexe K.0.28.

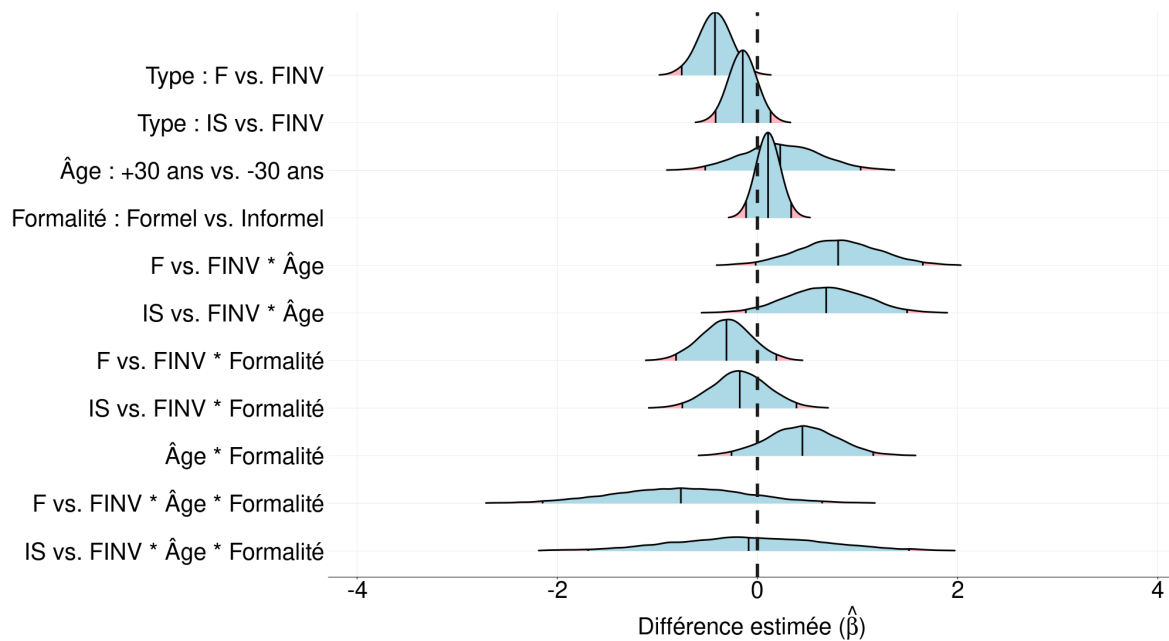


FIGURE 5.45 – Distributions postérieures pour le modèle brms (MG2, ORIGINE)

L'effet simple du TYPE est ici bien marqué pour les phrases à antéposition simple. Comparées aux phrases à antéposition et inversion, elles reçoivent des jugements plus bas ($\hat{\beta} = -0.43$, 95%CrI=[-0.76,-0.10], $P(\beta) < 0 = 0.99$). L'effet est aussi présent mais plus modéré pour les phrases in situ ($\hat{\beta} = -0.15$, 95%CrI=[-0.42,0.13], $P(\beta) < 0 = 0.86$). Par ailleurs, un léger effet simple de CONTEXTE est observable, avec des jugements globalement plus élevés rendus en contexte formel ($\hat{\beta} = 0.11$, 95%CrI=[-0.52,1.03], $P(\beta) > 0 = 0.83$).

Ces effets se retrouvent dans des interactions TYPE*ÂGE. Chez les plus de 30 ans, l'antéposition avec inversion n'est pas vraiment le type associé à des jugements d'origine urbaine, à l'inverse des moins de 30 ans. En proportion, chez les plus de 30 ans, les types avec antéposition simple ($\hat{\beta} = 0.81$, 95%CrI=[-0.02,1.65], $P(\beta) > 0 = 0.97$) et in situ ($\hat{\beta} = 0.69$, 95%CrI=[-0.12,1.50], $P(\beta) > 0 = 0.95$) sont donc liés à des jugements plus orientés vers une origine urbaine que dans l'autre groupe.

Une interaction TYPE*CONTEXTE se trouve pour les antépositions simples. Elle est liée aux jugements plus hauts qui sont associés au type avec antéposition et inversion en contexte formel, lesquels creusent l'écart avec l'antéposition simple par rapport aux contextes informels ($\hat{\beta} = -0.31$, 95%CrI=[-0.81,0.19], $P(\beta) < 0 = 0.89$).

Une interaction ÂGE*CONTEXTE doit être mentionnée ($\hat{\beta} = 0.45$, 95%CrI=[-0.26,1.16], $P(\beta) > 0 = 0.90$). Elle traduit le fait que les plus de 30 ans portent des jugements proportionnellement indiquant plus une origine urbaine quand le contexte devient formel, par rapport à ce que font les moins de 30 ans. Ceci tient au fait que sur cette échelle les plus de 30 ans ont rendu des jugements bien plus élevés pour les antépositions avec inversion et les in situ en contexte formel qu'en contexte informel.

Une interaction triple TYPE*CONTEXTE*ÂGE prolonge ceci pour les antépositions simples ($\hat{\beta} = -0.77$, 95%CrI=[-2.15,0.65], $P(\beta) < 0 = 0.86$). Elle traduit le fait que l'opposition entre antéposition simple et antéposition avec inversion sur le plan de l'origine géographique n'est quasiment pas modulée par la formalité du contexte chez les moins de 30 ans, alors qu'elle l'est visiblement chez les plus de 30 ans. Les antépositions avec inversion sont liées à des jugements bien moins sur l'origine urbaine en contexte informel pour ce groupe, et l'écart entre types se réduit donc. Ce n'est

pas autant le cas en ce qui concerne les phrases in situ.

Modèle pour l'échelle des ÉTUDES

La Figure 5.46 donne les distributions postérieures pour le cinquième modèle lancé, avec pour variable dépendante les jugements rendus sur l'échelle des ÉTUDES. Le détail en est disponible dans l'Annexe K.0.29.

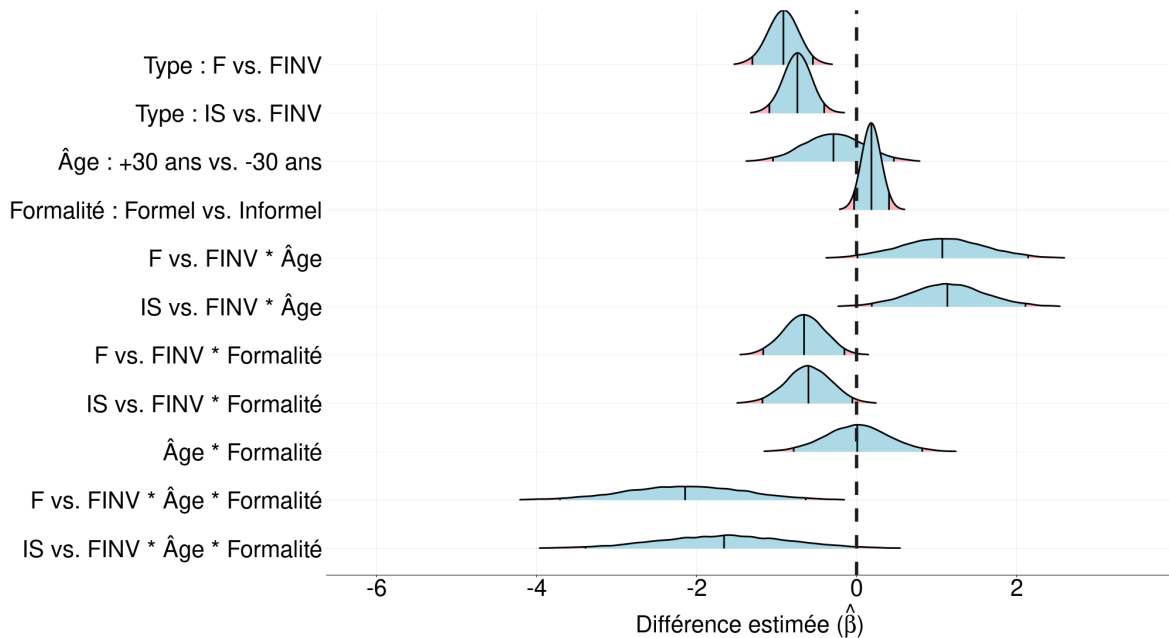


FIGURE 5.46 – Distributions postérieures pour le modèle brms (MG2, ÉTUDES)

Un effet simple du TYPE est visible. Comparativement aux antépositions avec inversion, les antépositions simples sont associées à des jugements plus bas en termes de longueur d'études ($\hat{\beta} = -0.92$, 95%CrI=[-1.30,-0.55], $P(\beta) < 0=1$). C'est également le cas pour les phrases in situ ($\hat{\beta} = -0.74$, 95%CrI=[-1.09,-0.40], $P(\beta) < 0=1$).

Un effet du CONTEXTE est également constaté ($\hat{\beta} = 0.19$, 95%CrI=[-0.03,0.41], $P(\beta) > 0=0.96$). Il traduit des jugements globalement plus hauts rendus dans des contextes formels que dans les contextes informels. Par ailleurs, un effet simple de groupe d'ÂGE, plutôt tendanciel, est notable ($\hat{\beta} = -0.29$, 95%CrI=[-1.04,0.47], $P(\beta) < 0=0.77$). Il est à mettre en lien avec le fait que la distribution des jugements rendus par les plus de 30 ans sur cette échelle est très centrée sur 0, alors que les jugements rendus par les moins de 30 ans sont plus élevés.

Ces résultats se prolongent par des interactions TYPE*ÂGE. Elles traduisent chez les plus de trente ans la disparition de l'écart entre les jugements associés aux antépositions avec inversion avec, d'une part, ceux associés aux antépositions simples ($\hat{\beta} = 1.07$, 95%CrI=[0.01,2.15], $P(\beta) > 0=0.98$) et, d'autre part, les phrases in situ ($\hat{\beta} = 1.14$, 95%CrI=[0.19,2.11], $P(\beta) > 0=0.99$).

Deux interactions TYPE*CONTEXTE sont présentes aussi. Elles sont liés à l'augmentation des jugements rendus en contexte formel sur les antépositions avec inversion, ce qui creuse leur écart avec les antépositions simples ($\hat{\beta} = -0.66$, 95%CrI=[-1.17,-0.15], $P(\beta) < 0=0.99$) mais aussi les phrases in situ ($\hat{\beta} = -0.60$, 95%CrI=[-1.18,-0.05], $P(\beta) < 0=0.98$).

Enfin, une interaction triple TYPE*CONTEXTE*ÂGE existe autant pour les phrases avec antéposition simple ($\hat{\beta} = -2.15$, 95%CrI=[-3.71,-0.64], $P(\beta) < 0=1$) que pour les phrases in situ ($\hat{\beta} = -1.67$,

95%CrI=[-3.39,0.00], $P(\beta)>0=0.97$). Elles sont la confirmation statistique du renversement total du rapport entre antéposition avec inversion et les deux autres types chez les plus de 30 ans en fonction du contexte. Pour ce groupe, l'antéposition avec inversion est liée en contexte informel à des études présumées plus courtes que les autres types phrastiques, alors qu'elle est liée à des études présumées plus longues en contexte formel. Ce, alors que ce phénomène n'est pas constaté chez les moins de 30 ans, pour qui les antépositions avec inversion sont toujours liées à un jugement plus haut sur la longueur des études, quelle que soit la formalité du contexte.

Modèle pour l'échelle de LOISIRS

La Figure 5.47 donne les distributions postérieures pour le dernier modèle calculé, avec pour variable dépendante les jugements reçus sur l'échelle des LOISIRS, avec pour extrémités le sport et la lecture. Le détail du modèle est disponible dans l'Annexe K.0.30.

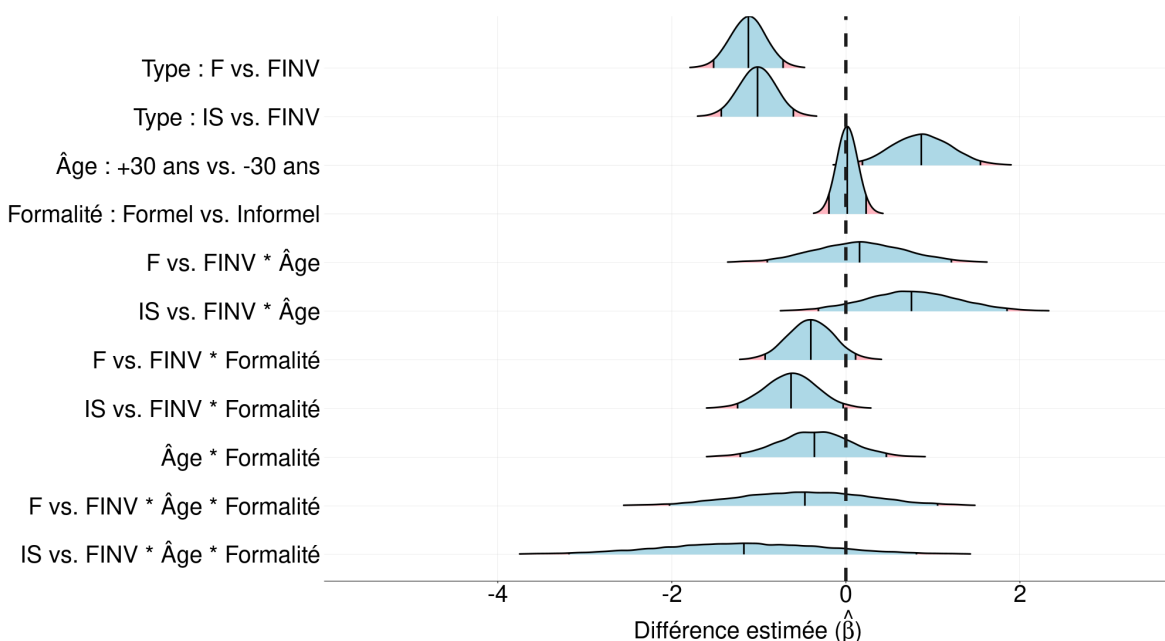


FIGURE 5.47 – Distributions postérieures pour le modèle brms (MG2, LOISIRS)

Il se trouve dans ces données un effet simple de TYPE. Comparativement aux antépositions avec inversion, les antépositions simples sont associées à des jugements plus bas ($\hat{\beta} = -1.12$, 95%CrI=[-1.52,-0.72], $P(\beta)<0=1$), comme les phrases in situ ($\hat{\beta} = -1.02$, 95%CrI=[-1.43,-0.60], $P(\beta)<0=1$). Ceci rend compte d'une association générale entre l'antéposition avec inversion et l'activité de lecture, par opposition aux deux autres types.

Un effet simple de groupe d'ÂGE est par ailleurs constaté ($\hat{\beta} = 0.87$, 95%CrI=[0.19,1.55], $P(\beta)>0=0.99$). Il illustre que les plus de 30 ans ont globalement rendu des jugements plus proches de la lecture, indépendamment du contexte et du type interrogatif, que les moins de 30 ans.

Une interaction TYPE*ÂGE est présente, seulement pour les phrases in situ. Elle est liée au fait que les interrogatives in situ sont liées à des jugements bien plus hauts chez les plus de 30 ans que chez les moins de 30 ans, effaçant presque l'écart avec les phrases à antéposition et inversion ($\hat{\beta} = 0.76$, 95%CrI=[-0.32,1.85], $P(\beta)>0=92$).

Une interaction TYPE*CONTEXTE est visible. Elle confirme une augmentation de l'écart, en contexte formel, entre les jugements rendus sur des phrases avec antéposition et inversion et les

phrases avec antéposition simple ($\hat{\beta} = -0.40$, 95%CrI=[-0.93,0.11], $P(\beta) < 0.94$). Cet écart se creuse aussi avec les phrases in situ ($\hat{\beta} = -0.63$, 95%CrI=[-1.24,-0.03], $P(\beta) < 0.98$).

Une interaction ÂGE*CONTEXTE doit être mentionnée ($\hat{\beta} = -0.36$, 95%CrI=[-1.21,0.47], $P(\beta) < 0.81$). Elle traduit des jugements globalement plus bas rendus quand le contexte devient formel par les plus de 30 ans, en comparaison avec les moins de 30 ans.

Ce dernier effet se prolonge dans une interaction triple TYPE*CONTEXTE*ÂGE pour les phrases in situ ($\hat{\beta} = -1.17$, 95%CrI=[-3.18,0.81], $P(\beta) < 0.88$). Cette interaction est liée au fait que, pour les plus de 30 ans, les phrases in situ en contexte informel sont autant liées à l'activité de lecture que les antépositions avec inversion, alors qu'elles le sont moins en contexte formel. Ce, alors que chez les moins de 30 ans il n'y a pas ce changement brusque.

5.3.2.5 Bilan : Variantes interrogatives et faisceau de stéréotypes

L'étude MG2 a confirmé les résultats de l'étude MG1 en allant plus loin. D'abord, d'un point de vue méthodologique, le protocole développé pour cette tâche, plus facilement systématisable que celui de l'étude MG1, a bien fonctionné. L'exercice a déstabilisé nombre de volontaires, mais les échelles à base de dessins stéréotypés ont été bien prises en main. Les différences de résultats entre conditions, pour des items semblables, montre la validité du design.

Sur le plan des résultats, l'hypothèse principale est à nouveau confirmée : il y a bien un faisceau d'indices sociaux qui sont associés aux variantes interrogatives principales du français. Les indices de ce faisceau sont autant des caractéristiques générationnelles (ici, l'âge), sociales (richesse, éducation, origine géographique, loisirs), ou encore psychologiques (ici, la sociabilité). Il est probable que ces indices, déjà inter-reliés comme l'a montré le calcul de leurs corrélations statistiques, soient reliés à d'autres indices encore.

Ce tissu de stéréotypes est par ailleurs conforme à ce qui était attendu, sur la base des études de corpus (Chapitre 4 et des études de jugements d'acceptabilité décrites plus tôt (Sections 5.2.1-5.2.5)). Globalement, les personnes utilisant l'antéposition avec inversion sont associées à un âge plus avancé, à une plus haute richesse, à des études plus longues, et à une origine plus urbaine. Ces trois stéréotypes sont liés à l'idée que ce sont des groupes sociaux précis qui, au sein de la communauté francophone, utilisent cette structure définie comme « standard ». Le type avec antéposition et inversion est par ailleurs aussi associé à une activité de lecture plus régulière, ce qui est cohérent avec l'idée sous-jacente d'une structure plutôt lié à l'écrit, et mal adaptée à l'oral (voir résultats de l'étude AJT3, Section 5.2.3). Enfin, les personnes recourant à cette forme interrogative sont jugées relativement moins « sociables ». Il faut sans doute, ici aussi, y voir une conséquence de la mauvaise adéquation de l'antéposition avec inversion avec des situations d'interaction spontanée, simulées dans les items expérimentaux.

Il est par ailleurs important de conserver à l'esprit que, comme pour les études de jugements d'acceptabilité, la formalité du contexte influe sur ce réseau de stéréotypes. En contexte informel, et sauf pour la notion de sociabilité où c'est l'inverse, le réseau des stéréotypes se « ressert » et les écarts de perception entre types phrastiques diminuent généralement. Comme dans les études précédentes, il faut y voir une perte relative de la spécificité de l'antéposition avec inversion dans les contextes informels. En miroir, ce type avec antéposition et inversion emporte au contraire des jugements encore moins hauts sur l'échelle de sociabilité en contextes informels. Tout ceci est cohérent avec les effets de contexte déjà observés, autant sur corpus que expérimentalement.

Il faut enfin noter que le critère d'âge et le seuil de 30 ans, déjà opérant dans les études précédemment décrites, a ici aussi un impact. Il faut le considérer avec prudence, du fait du déséquilibre entre groupes de volontaires sur ce point. Notamment sur les loisirs (lecture), les études et la richesse, les plus de 30 ans semblent avoir intégré des associations forme-stéréotype un peu différentes des moins de 30 ans, que les variations de contexte modulent plus visiblement. Peut-être faut-il y voir la trace d'une évolution des valeurs sociales des différentes variantes interrogatives, en lien avec les changements diachroniques qui affectent la langue de manière beaucoup plus générale.

5.3.3 Le cas des interrogatives en *est-ce que* (MG3)

L'étude AJT4 (Section 5.2.4) a complété les études de jugements d'acceptabilité précédemment menées en étudiant les interrogatives avec antéposition en *est-ce que*. Notamment, les préférences des francophones à leur sujet ont été précisées, en comparaison avec les antépositions avec inversion et les phrases in situ. L'étude MG3 permet, de façon parallèle, de relier ces antépositions en *est-ce que* au réseau de stéréotypes dégagé dans l'étude MG2. Avec un matériel plus « léger » et une modalité écrite, elle est un complément à l'étude MG2, mais aussi à l'étude AJT4, puisqu'elle explore les stéréotypes associés à la forme avec antéposition en *est-ce que*.

5.3.3.1 Hypothèse et prédictions

L'hypothèse générale est toujours celle d'une projection d'indices sociaux différents sur les personnes utilisant des variantes interrogatives différentes. Plus précisément, à partir des données de l'étude AJT4, l'hypothèse est que les antépositions en *est-ce que* seront associées à des jugements intermédiaires, entre ceux associés aux antépositions avec inversion et ceux associés aux phrases in situ. Pour le vérifier, seules trois échelles de « localisation sociale » ont cette fois été utilisées : celle de l'ÂGE, des ÉTUDES, et de l'ORIGINE géographique. Pour les deux premières, les antépositions avec inversion ont été associées dans l'étude MG2 à un âge plus avancé et à un niveau d'études plus haut que les phrases in situ. Il était prédit que l'antéposition en *est-ce que* serait associée à des jugements intermédiaires sur ces deux échelles. Pour la troisième échelle, le patron dégagé dans l'étude MG2 étant plus nuancé, il était prédit qu'il en serait de même ici, avec éventuellement une moins grande différence entre les présomptions d'origine urbaine pour les antépositions avec inversion et pour les autres types.

5.3.3.2 Protocole et items expérimentaux

De manière générale, le protocole de cette étude est assez proche de celui de l'étude AJT4 (Section 5.2.4). Les volontaires lisaient de courts dialogues entre deux individus (A et B), puis estimaient trois caractéristiques sociales sur B (son âge, sa richesse, son origine géographique).

L'intégralité des items écrits utilisés pour l'étude AJT4 a été reprise ici. Pour rappel, il s'agissait d'une version modifiée des 30 items-cibles de l'étude AJT2 (Annexe F0.1), dont la condition *antéposition simple* était remplacée par une condition *antéposition en est-ce que*. À ces items étaient rajoutés 10 distracteurs de l'étude AJT2, et 24 distracteurs d'une autre étude sur l'ordre des cliques en français (voir Annexe I.0.3 pour ces 34 distracteurs). Le Tableau 5.36 rappelle un exemple d'item-cible.

Type	Contexte	
	INFORMEL	FORMEL
Antéposition + ESK	A : Non, elle est pas là, ma soeur. B : Quand est-ce qu' elle est partie?	A : Non, ma soeur n'est pas là. B : Quand est-ce qu' elle est partie?
Antéposition avec inversion	A : Non, elle est pas là, ma soeur. B : Quand est-elle partie?	A : Non, ma soeur n'est pas là. B : Quand est-elle partie?
In situ	A : Non, elle est pas là, ma soeur. B : Elle est partie quand?	A : Non, ma soeur n'est pas là. B : Elle est partie quand?

TABEAU 5.36 – Exemple d'item expérimental (AJT4-MG3)

Chaque liste comportait donc 30 items cibles, 34 distracteurs et 3 entraînements. Le carré latin utilisé pour les études AJT2/3/4 a été repris, avec un designe 3x2. Les variables étaient le TYPE syntaxique (*antéposition en est-ce que*, *antéposition avec inversion*, *in situ*) et le CONTEXTE (*formel* vs. *informel*).

L'étude commençait après une page de consentement (Annexe L.0.1) puis une page de consignes (Annexe L.0.2). Elle comportait donc 67 items, pour un temps étalonné à 25 minutes. Après la passation, les volontaires remplissaient une dernière page de recueil de données personnelles basiques (âge, niveau d'étude, domaine professionnel ou d'études, langues).

Sous chaque item se trouvaient trois échelles (âge, éducation, origine). Ces échelles étaient des *sliders*, des curseurs à faire glisser une échelle aux crans non-numérotés, entre deux extrémités (gauche : jeune, peu éduquée, de la ville – droite : âgée, très éduquée, de la campagne). La Figure 5.48 est une capture d'écran pendant l'expérience.

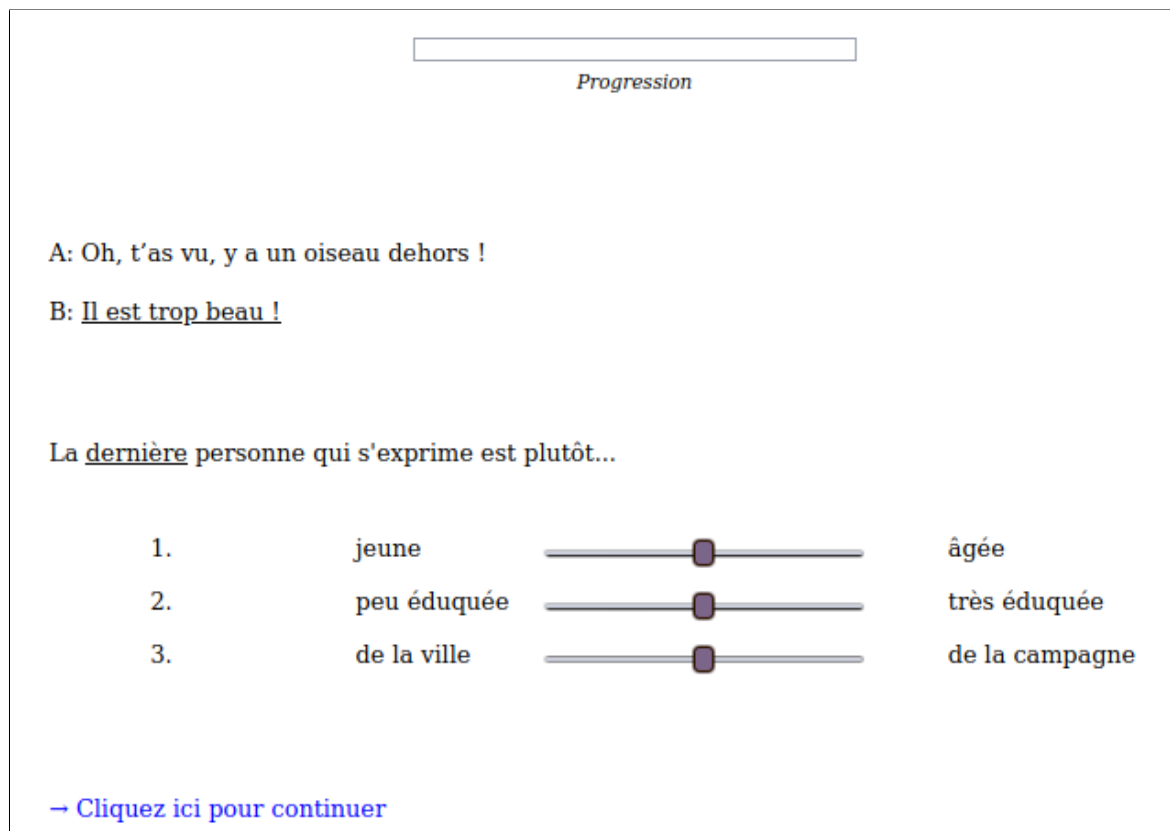


FIGURE 5.48 – Capture d'écran sur un item expérimental (MG3)

Le programme informatique enregistrerait la position du curseur comme une réponse entre 0 et

10 (il y avait donc 11 crans sur chaque échelle). Les volontaires devaient cliquer sur un lien pour passer à l'item suivant. Aucune limite de temps n'était imposée. Il était indiqué dans les consignes qu'aucune bonne ou mauvaise réponse n'était attendue.

5.3.3.3 Participantes et participants

42 volontaires ont pris part à l'étude au printemps 2020, via la plateforme IbexFarm [DRUMMOND, 2016]. Le recrutement s'est fait uniquement via les réseaux sociaux pour limiter le risque d'avoir les mêmes personnes que pour l'étude AJT4, même si quelques mois s'étaient écoulés. Comme indiqué dans la page de consentement, il n'y avait pas de compensation.

Les 42 personnes dont les résultats ont été analysés avaient entre 18 et 77 ans ($\mu = 46.7$, $m = 30$). 38 ont déclaré être de genre féminin, 4 ont déclaré être de genre masculin. Les plus hauts diplômes obtenus par ces personnes étaient :

- en dessous du baccalauréat, N=3
- le baccalauréat, N=12
- un diplôme équivalent à Bac+3, N=19
- un diplôme équivalent à Bac+5, N=7
- un diplôme équivalent à Bac+8, N=1

9 personnes ont indiqué avoir pour région de naissance l'Île de France, 30 une autre circonscription administrative du territoire français métropolitain, et 2 ont indiqué venir de France sans préciser la région. 1 personne a précisé venir de Belgique francophone.

9 personnes ont précisé avoir mené des études dans un domaine lié aux lettres et langues, 10 dans les sciences humaines et sociales, 3 dans les sciences et techniques, 4 dans les arts, 2 dans le droit et les sciences politiques, 2 dans l'économie et gestion, 2 dans la santé, et 10 dans un autre domaine. 3 personnes ont précisé travailler dans un domaine lié à l'enseignement, 1 dans le journalisme ou les professions littéraires, 9 dans une profession de type administratif, 4 dans un emploi de direction, 1 dans l'industrie, 3 dans la santé, 1 dans le service à la personne, et 1 estimait relever d'une profession scientifique : 19 n'ont pas précisé. Les données sur les langues maîtrisées n'ont pas été traitées, à part pour une vérification de la langue maternelle indiquée.

Au total, 2814 (=67*42) triplets de jugements (notes sur les trois échelles prises ensemble) ont été enregistrés.

5.3.3.4 Résultats

Seuls les 1260 (=30*42) triplets de jugements reçus sur les items cibles ont été conservés pour analyse. Les réponses fournies sur les trois échelles disponibles ont été codées numériquement entre -5 (jeune/peu éduquée/de la ville) et 5 (âgée/très éduquée/de la campagne) pour être centrées autour de 0.

Une matrice des corrélations entre échelles (méthode = Spearman, niveau de significativité = 0.05) est visible sur la Figure 5.49. Pour rappel, la couleur des cercles indique si une corrélation positive ou négative, et leur taille est proportionnelle au coefficient de corrélation. Les corrélations non significatives sont indiquées par une croix. Le coefficient de corrélation est la valeur chiffrée indiquée à gauche.

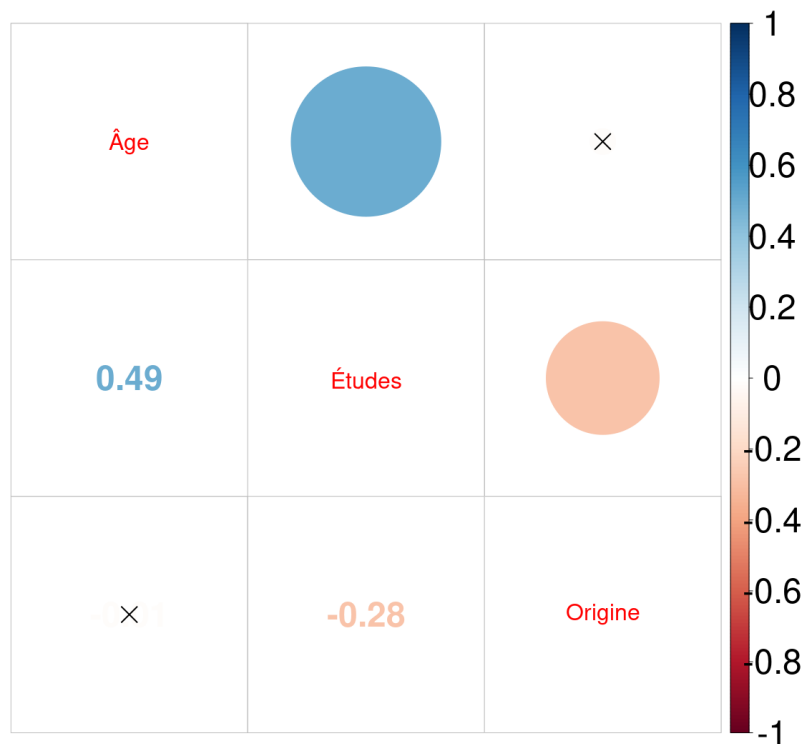


FIGURE 5.49 – Corrélations entre échelles (Spearman) (MG3)

Une corrélation existe donc entre certaines des échelles, comme il en existait dans l'étude MG2. La corrélation positive Âge/Études ($r = .49$, $p < 0.01$) était déjà présente dans l'étude MG2 ($r = .32$, $p < 0.01$), et atteste d'un certain lien entre le degré d'éducation et l'âge. La corrélation Études/Origine ($r = -.28$, $p < 0.01$) n'est guère plus surprenante. Sa valeur négative tient ici à ce que l'échelle a été présentée et codée dans un sens inverse à celle utilisée pour l'étude MG2 : une valeur négative des réponses sur cette échelle est liée à une présomption d'origine urbaine et une valeur positive indique que la personne est supposée venir de la campagne. Cette corrélation est donc en phase avec celle trouvée dans l'étude MG2 ($r = .41$, $p < 0.01$). Les personnes qui *viennent de la ville* sont associées à des études plus longues.

Par la suite, les données seront présentées échelle par échelle avant de procéder à un bilan général.

5.3.3.4.1 Résultats généraux (échelle ÂGE)

Le Tableau 5.37 et la Figure 5.50 donnent les moyennes globales brutes des jugements rendus sur chacune des trois formes interrogatives analysées, pour la première échelle proposée aux volontaires.

Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
0.38	1.32	-0.46

TABLEAU 5.37 – Jugement moyen en fonction du type interrogatif (MG3, ÂGE)

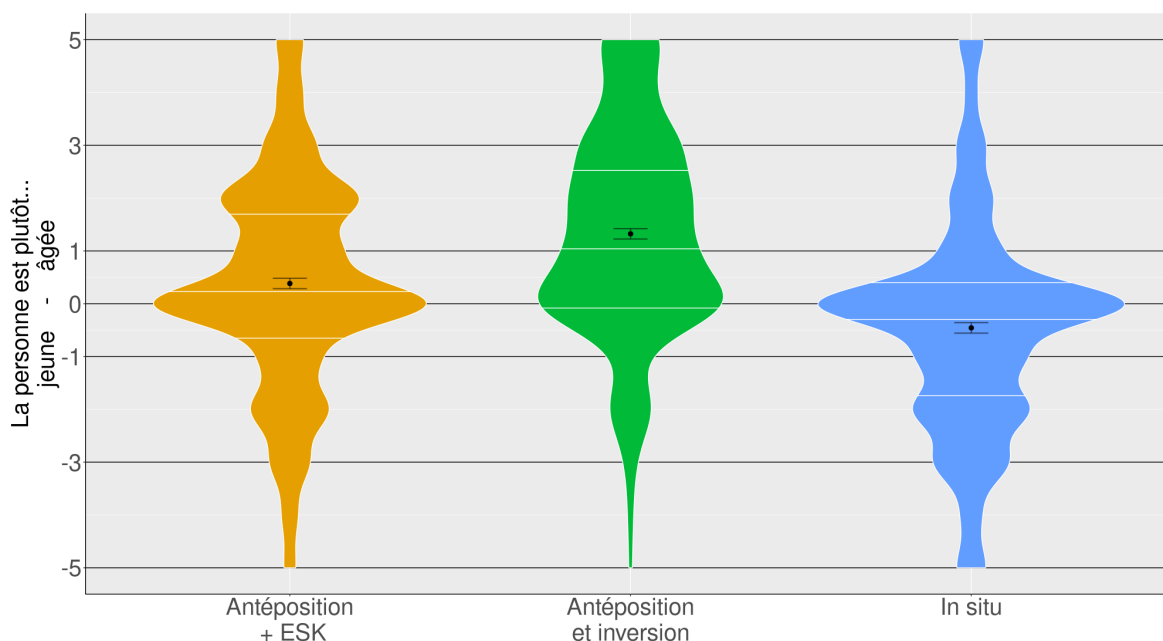


FIGURE 5.50 – Jugement moyen en fonction du type interrogatif (MG3, ÂGE)

Ces données suivent le patron attendu, avec une présomption plus haute concernant l'âge de la personne, lorsque c'est une antéposition avec inversion qui est produite. La distribution des jugements est plutôt haute. À l'autre bout de ces estimations, les phrases in situ sont elles comparativement associées à plus de jeunesse, avec une distribution ceci dit assez centrée autour de 0. Comme il était attendu, les antépositions en *est-ce que* sont pour leur part associées à des jugements intermédiaires, avec là aussi une distribution assez concentrée vers le centre.

L'écart entre les deux extrémités de ces jugements moyens, sur cette échelle d'âge, semble plus important que dans l'étude MG2. Un facteur pouvant l'expliquer est que cette fois, par nature puisqu'ils sont écrits, aucune voix n'est associée aux phrases. Alors que toutes les voix de l'étude MG2 étaient relativement jeunes (moins de 30 ans à la date des enregistrements des matériels), ici les volontaires ont eu plus de « liberté » dans leur estimation. Une autre explication peut tenir, comme il avait été proposé pour l'expérience AJT4 (5.2.4), à la présence de ce type intermédiaire que sont les antépositions en *est-ce que*. Leur présence pourrait avoir eu pour effet de marquer un peu plus la construction in situ. Enfin, une dernière explication peut être justement cette modalité écrite : les phrases in situ ressortent peut-être plus en tant que « non-standard » quand elles sont lues que quand elles sont entendues.

5.3.3.4.2 Un effet d'âge (échelle ÂGE)

Le Tableau 5.38 et la Figure 5.51 donnent les moyennes globales brutes des jugements pour chacune des trois formes interrogatives, pour la première échelle à nouveau. Par continuité avec les expériences passées, la séparation en âge a de nouveau été établie au seuil de 30 ans.

	Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
-30 ans	0.29	0.96	-0.58
+30 ans	0.51	1.86	-0.28

TABLEAU 5.38 – Jugement moyen en fonction du type interrogatif et de l'âge (MG3, ÂGE)

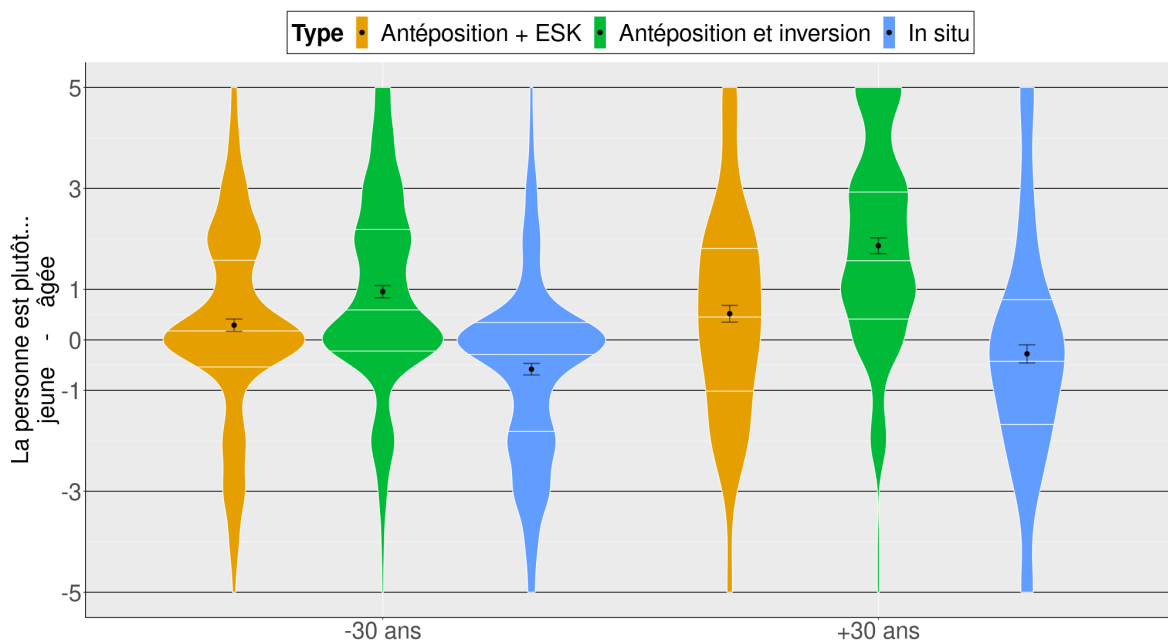


FIGURE 5.51 – Jugement moyen en fonction du type interrogatif et de l'âge (MG3, ÂGE)

Un effet d'âge est ici visible, et prend deux formes. D'abord, l'antéposition avec inversion est encore plus associée à un âge avancé par le groupe des plus de 30 ans qu'elle ne l'est par les moins de 30 ans. C'est pour ce type que la différence de jugements est la plus forte entre groupes.

Un deuxième élément, lié, est que les écarts entre types sont moins importants chez les moins de 30 ans que chez les plus de 30 ans. Ce, alors par exemple dans un patron similaire à ce qui a déjà été observé, notamment dans les jugements d'acceptabilité, les types phrastiques sont plus distingués les uns des autres par le groupe des plus de 30 ans.

5.3.3.4.3 Un effet de contexte? (échelle ÂGE)

Le Tableau 5.39 et la Figure 5.39 donnent les moyennes brutes des jugements pour chacune des trois formes interrogatives, pour la première échelle, en fonction de la formalité du contexte.

	Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
Contexte formel	0.34	1.32	-0.39
Contexte informel	0.43	1.32	-0.53

TABLEAU 5.39 – Jugement moyen en fonction du type interrogatif et du contexte (MG3, ÂGE)

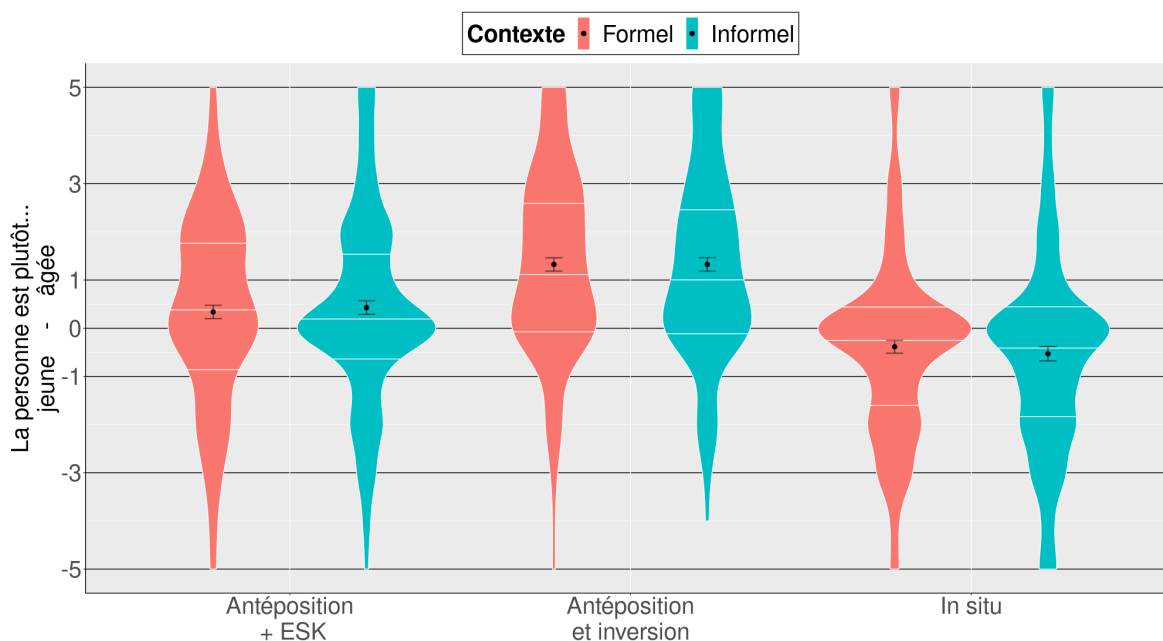


FIGURE 5.52 – Jugement moyen en fonction du type interrogatif et du contexte (MG3, ÂGE)

Il semble ici n’y avoir aucun effet important du contexte, en dehors d’une possible baisse des jugements rendus sur les interrogatives in situ en contexte informel. Ceci est assez conforme aux observations issues de l’étude MG2. L’équilibre parfait des jugements associés aux interrogatives avec antéposition et inversion, d’un contexte à l’autre, est notable. Pour rechercher une éventuelle interaction masquée entre les effets de groupe d’âge et de contexte, le Tableau 5.40 synthétise les moyennes des jugements donnés sur chaque type phrastique, en fonction du contexte mais aussi du groupe d’âge.

	-30 ans		+30 ans	
	Formel	Informel	Formel	Informel
Antéposition en <i>est-ce que</i>	0.26	0.33	0.46	0.58
Antéposition et inversion	0.98	0.93	1.82	1.91
In situ	-0.56	-0.60	-0.13	-0.42

TABLEAU 5.40 – Jugement moyen relatif aux types interrogatif, contexte et âge (MG3, ÂGE)

Aucune influence du contexte sur les jugements rendus ne semblent ici modulée par le groupe d’âge, sauf très marginalement. Les antépositions avec inversion sont liées à des jugements un tout petit plus haut en contexte informel chez les plus de 30 ans, et les phrases in situ sont liées à des jugements plus bas dans ce même contexte pour le même groupe. En conséquence de quoi l’écart entre ces deux types phrastiques, en contexte informel, pourrait sembler se réduire pour ce groupe d’âge.

5.3.3.4.4 Analyses bayésiennes (échelle ÂGE)

Un modèle de régression bayésienne a en effet été calculé sur la base des jugements rendus sur l’échelle d’ÂGE (Annexe L.0.4). Ce modèle est strictement parallèle aux modèles calculés pour l’étude AJT4, excepté pour la variable dépendante. Ici, celle-ci était le jugement rendu sur l’échelle d’âge, pris comme valeur ordinaire. La variable indépendante est l’interaction triple entre trois effets fixes : le TYPE syntaxique (*antéposition en est-ce que*, *antéposition avec inversion*, *in situ*,

avec l'antéposition avec inversion comme niveau de référence), le groupe d'ÂGE (moins ou plus de 30 ans, avec le groupe des moins de 30 ans pour référence), et le CONTEXTE (*formel* ou *informel*, avec *informel* pour niveau de référence). Les effets aléatoires étaient les individus ("ID"), avec pour pente l'interaction *type*contexte*) et les items ("Item"), avec pour pente l'interaction triple *type*âge*contexte*). C'est toujours la famille "cumulative" qui a été sélectionnée pour faire tourner ces modèles, avec 4 chaînes et 6000 itérations par chaîne. La convergence des modèles a été atteinte et vérifiée.

La Figure 5.53 illustre les distributions postérieures de ce modèle pour chacun des paramètres et de leurs interactions. Elle permet une visualisation de la valeur généralisable des principaux résultats observés sur les données brutes.

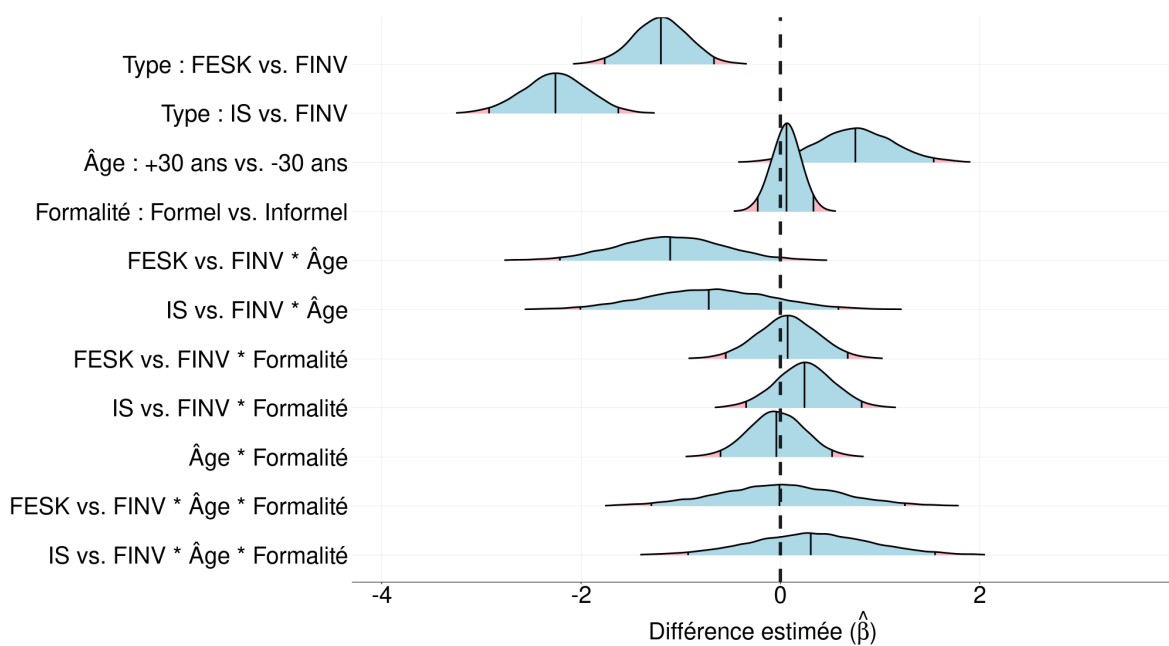


FIGURE 5.53 – Distributions postérieures pour le modèle brms (MG3, ÂGE)

L'effet simple de TYPE syntaxique est confirmé. Les antépositions en *est-ce que* sont de manière générale moins bien jugées que les antépositions avec inversion ($\hat{\beta} = -1.20$, 95%CrI=[-1.76,-0.67], $P(\beta) < 0=1$). C'est aussi le cas pour les interrogatives in situ ($\hat{\beta} = -2.26$, 95%CrI=[-2.92,-1.63], $P(\beta) < 0=1$).

Un effet simple d'ÂGE est également corroboré ($\hat{\beta} = 0.75$, 95%CrI=[-0.02,1.54], $P(\beta) > 0=0.97$). Il témoigne de ce que les jugements rendus par les plus de 30 ans sont globalement plus hauts que ceux des moins de 30 ans, ce qui est dû principalement aux hauts jugements associés aux formes avec antéposition et inversion. Cet effet simple se prolonge dans des interactions TYPE*ÂGE. Celles-ci corroborent, chez les plus de 30 ans, l'augmentation de l'écart entre les jugements rendus sur les antépositions avec inversion et, d'une part les antépositions simples ($\hat{\beta} = -1.11$, 95%CrI=[-2.21,-0.01], $P(\beta) < 0=0.98$), et d'autre part les phrases in situ ($\hat{\beta} = -0.72$, 95%CrI=[-2.01,-0.58], $P(\beta) < 0=0.87$).

Une interaction TYPE*CONTEXTE supporte modérément une influence du contexte sur les jugements rendus pour les interrogatives in situ ($\hat{\beta} = 0.24$, 95%CrI=[-0.34,0.82], $P(\beta) > 0=0.80$). Elle traduit le fait qu'en contexte informel les jugements rendus sur ces interrogatives sont plus bas encore qu'en contexte formel. En conséquence, l'écart avec les antépositions et inversion est un peu réduit en contexte formel.

Aucune interaction triple TYPE*ÂGE*CONTEXTE ne corrobore l'existence d'une sensibilité diffé-

rente des groupes d'âge à l'influence du contexte.

5.3.3.4.5 Résultats généraux (échelle ÉTUDES)

Le Tableau 5.41 et la Figure 5.54 donnent les moyennes globales brutes des jugements rendus sur chacune des trois formes interrogatives analysées, pour la deuxième échelle proposée aux volontaires.

Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
0.46	1.69	-0.11

TABLEAU 5.41 – Jugement moyen en fonction du type interrogatif (MG3, ÉTUDES)

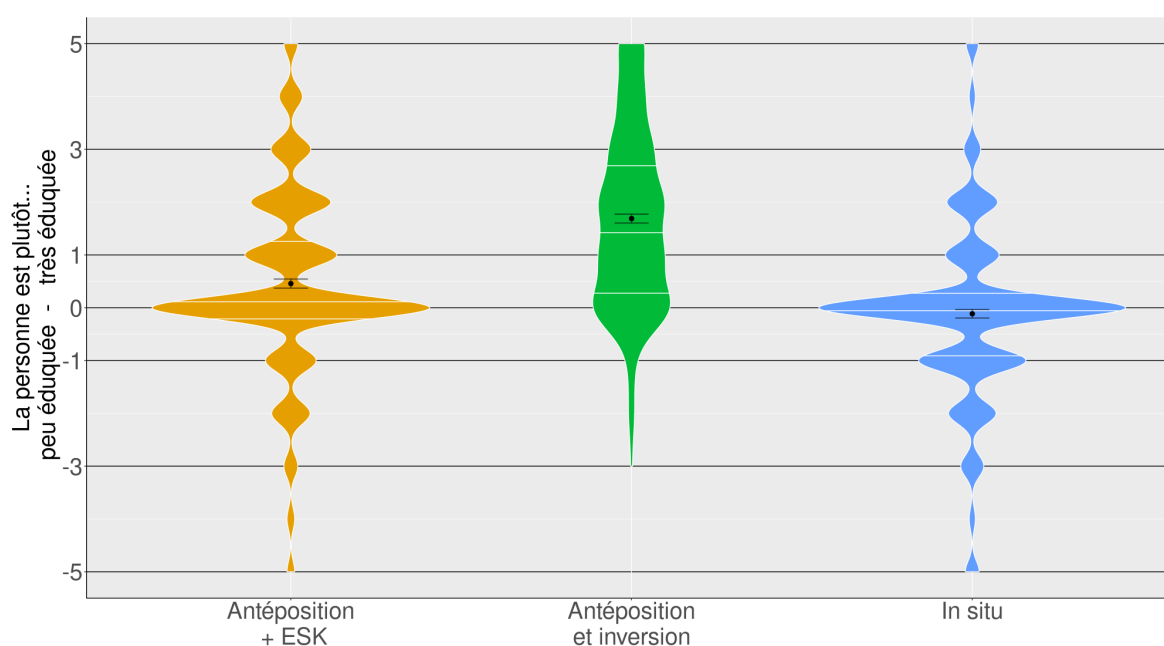


FIGURE 5.54 – Jugement moyen en fonction du type interrogatif (MG3, ÉTUDES)

Sur les données globales pour cette échelle, un patron semblable à celui observé dans l'étude MG2 émerge à nouveau. Les antépositions avec inversion sont associées à des jugements plus hauts alors que les phrases in situ sont associées à des jugements plus bas. Comme pour l'échelle d'ÂGE, les jugements associés aux antépositions en *est-ce que* semblent un intermédiaire entre les deux autres types.

Les distributions globalement centrées autour du 0 pour les antépositions en *est-ce que* comme pour les phrases in situ, néanmoins, tranchent avec la distribution majoritairement positive des réponses pour l'antéposition avec inversion. Il semble y avoir une association forte de ce type avec de longues études, qui n'existe pas pour les deux autres.

5.3.3.4.6 Un effet d'âge (échelle ÉTUDES)

Le Tableau 5.42 et la Figure 5.55 donnent les moyennes globales brutes des jugements pour chacune des trois formes interrogatives, pour la deuxième échelle.

	Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
-30 ans	0.50	1.41	-0.07
+30 ans	0.39	2.09	-0.18

TABLEAU 5.42 – Jugement moyen en fonction du type interrogatif et de l'âge (MG3, ÉTUDES)

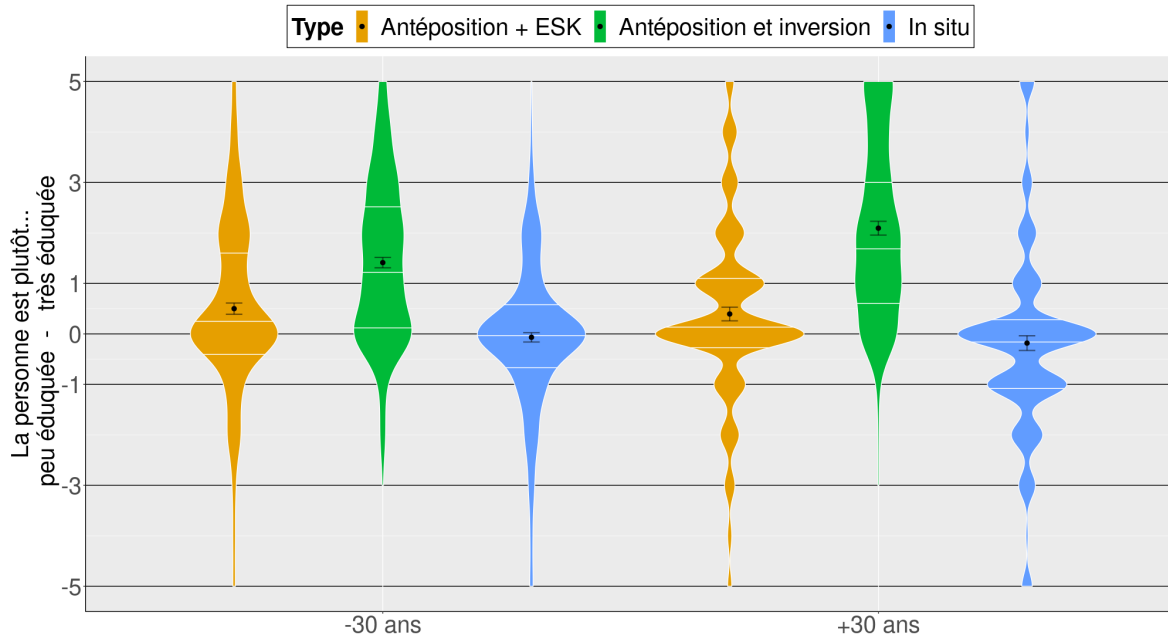


FIGURE 5.55 – Jugement moyen en fonction du type interrogatif et de l'âge (MG3, ÉTUDES)

Les données font apparaître une différence entre groupes similaires à celle notée pour la première échelle. Au premier abord, les jugements rendus sont conformes au patron général pour cette échelle, avec une association des antépositions avec inversion à des jugements plus hauts. L'antéposition en *est-ce que* est par ailleurs toujours un intermédiaire entre celles-ci et les phrases in situ.

Cependant, le groupe des plus de 30 ans rend des jugements plus hauts sur les antépositions avec inversion que celui des moins de 30 ans. Par ailleurs, les jugements rendus par ce groupe sur les deux autres types sont plus bas que ceux rendus par les moins de 30 ans. En conséquence, l'écart est amplifié entre, d'une part, les antépositions avec inversion, et d'autre part les deux autres types phrastiques.

Une dernière observation ici est que les distributions des jugements sont plutôt centrées autour du 0 pour les deux types alternatifs chez les plus de 30 ans. Les moins de 30 ans semblent manifester plus de variabilité dans leurs jugements.

5.3.3.4.7 Un effet de contexte ou une interaction avec l'âge? (échelle ÉTUDES)

Le Tableau 5.43 et la Figure 5.43 donnent les moyennes brutes des jugements pour chacune des trois formes interrogatives, pour la deuxième échelle, en fonction de la formalité du contexte.

	Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
Contexte formel	0.49	1.67	-0.07
Contexte informel	0.43	1.71	-0.16

TABLEAU 5.43 – Jugement moyen en fonction du type interrogatif et du contexte (MG3, ÉTUDES)

Il semble ici n'y avoir qu'une diminution marginale des jugements rendus en contexte informel, en association avec les antépositions en *est-ce que* et les phrases in situ.

Le Tableau 5.44 et la Figure 5.56 approfondissent ces données en synthétisant les moyennes des jugements donnés sur chaque type phrastique, en fonction du contexte mais aussi du groupe d'âge.

	-30 ans		+30 ans	
	Formel	Informel	Formel	Informel
Antéposition en <i>est-ce que</i>	0.59	0.41	0.33	0.46
Antéposition et inversion	1.34	1.49	2.15	2.04
In situ	-0.02	-0.12	-0.15	-0.21

TABLEAU 5.44 – Jugement moyen relatif aux types interrogatif, contexte et âge (MG3, ÉTUDES)

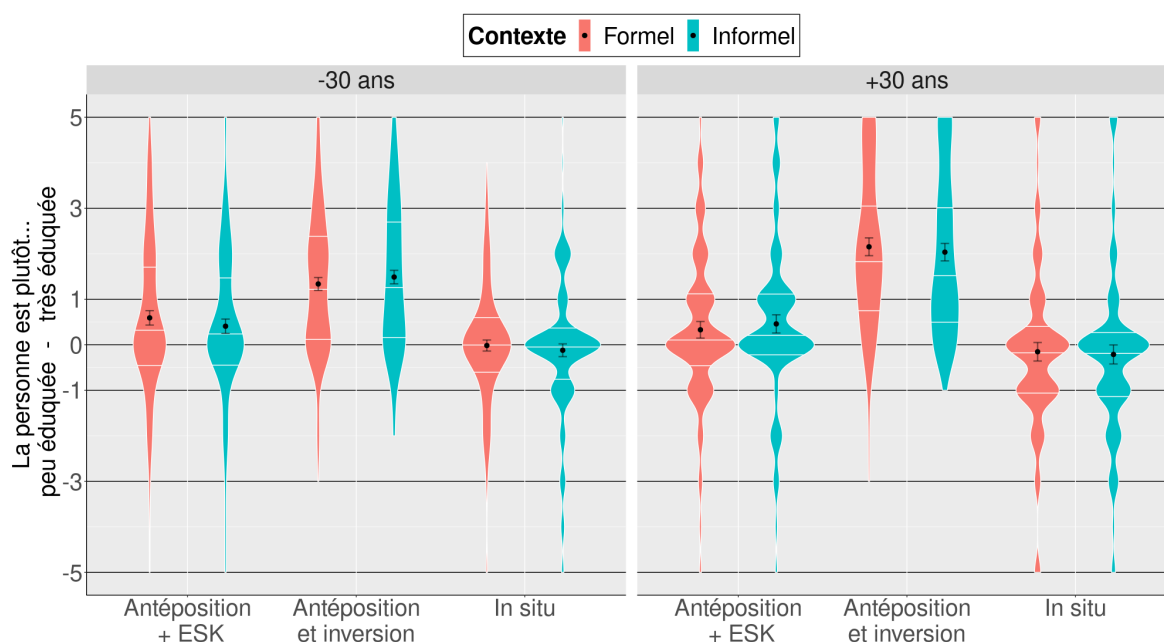


FIGURE 5.56 – Jugement moyen relatif aux types interrogatif, contexte et âge (MG3, ÉTUDES)

Ici, les antépositions avec inversion sont associées à des jugements un peu plus *hauts* par les plus de 30 ans en contexte formel, alors que c'est l'inverse chez les moins de 30 ans. En parallèle, les antépositions en *est-ce que* sont associées à des jugements un peu plus *bas* par les plus de 30 ans en contexte formel, alors que c'est l'inverse chez les moins de 30 ans. De fait, ceci entraîne une augmentation de l'écart entre antépositions avec inversion et antépositions en *est-ce que* chez les plus de 30 ans, alors que cet écart diminue chez les moins de 30 ans.

5.3.3.4.8 Analyses bayésiennes (échelle ÉTUDES)

Un modèle de régression bayésienne a été calculé sur la base des réponses à cette échelle des ÉTUDES, strictement parallèle au modèle calculé pour la précédente échelle (Annexe L.0.5). La Figure 5.57 illustre les distributions postérieures de ce modèle pour chacun des paramètres et de leurs interactions.

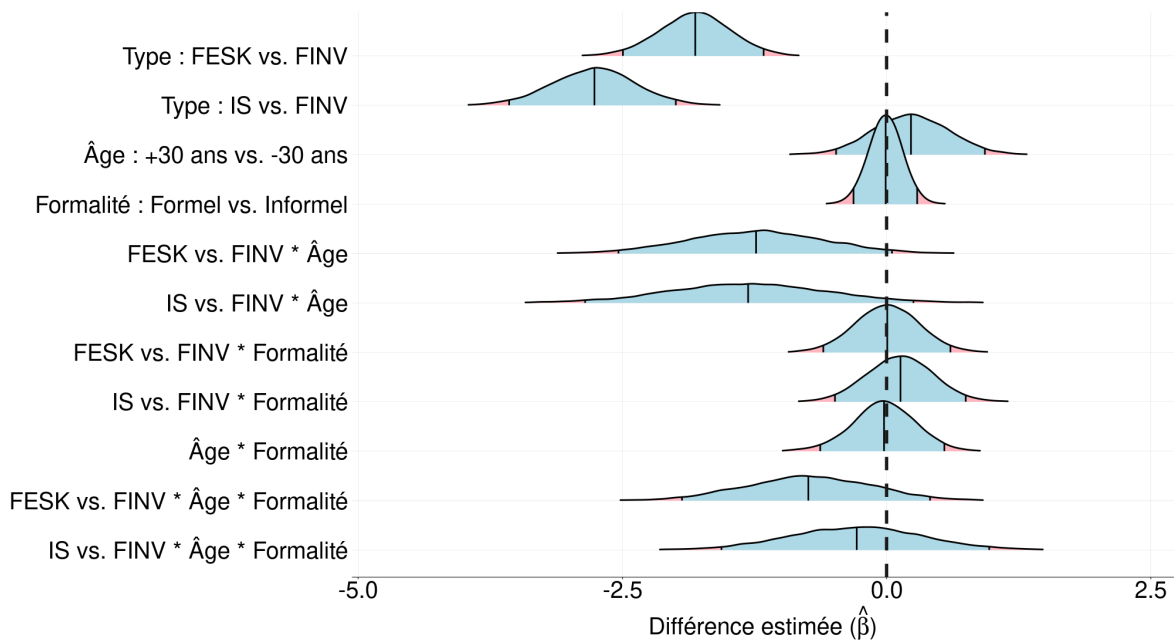


FIGURE 5.57 – Distributions postérieures pour le modèle brms (MG3, ÉTUDES)

Ce modèle confirme les effets simples de TYPE syntaxique. Les antépositions en *est-ce que* sont de manière générale associées à des jugements moins hauts que les antépositions avec inversion ($\hat{\beta} = -1.82$, 95%CrI=[-2.49,-1.16], $P(\beta) < 0=1$). C'est également le cas pour les interrogatives in situ ($\hat{\beta} = -2.77$, 95%CrI=[-3.57,-1.99], $P(\beta) < 0=1$).

Un effet simple d'ÂGE corrobore une tendance du groupe des plus de 30 ans à donner des jugements plus hauts que les moins de 30 ans ($\hat{\beta} = 0.23$, 95%CrI=[-0.48,0.93], $P(\beta) > 0=0.74$). Cet effet a une influence plus importante dans des interactions TYPE*ÂGE. Celles-ci traduisent une augmentation de l'écart entre les jugements rendus sur les antépositions avec inversion et les deux autres types phrastiques, chez les plus de 30 ans par rapport aux moins de 30 ans. C'est autant valable pour la différence entre antéposition avec inversion et antéposition en *est-ce que* ($\hat{\beta} = -1.24$, 95%CrI=[-2.54,0.05], $P(\beta) < 0=0.97$) que pour la différence entre antéposition avec inversion et in situ ($\hat{\beta} = -1.31$, 95%CrI=[-2.85,0.25], $P(\beta) < 0=0.95$).

L'effet observé sur les données brutes en croisant l'âge des volontaires et la formalité du contexte, pour les antépositions en *est-ce que*, est corroboré par une interaction triple TYPE*ÂGE*CONTEXTE ($\hat{\beta} = -0.74$, 95%CrI=[-1.94,0.41], $P(\beta) < 0=0.89$). Il semble donc bien y avoir conception différente, chez les plus de 30 ans par rapport aux moins de 30 ans, de la relation entre la longueur des études et le fait d'utiliser une interrogative en *est-ce que* dans un contexte formel.

Il est possible que ceci soit le reflet de projections sociales un peu différentes sur ces deux types phrastiques, liées à une considération différente au regard de la langue française standard. Dans l'étude AJT4 (Section 5.2.4), le groupe marquait l'antéposition en *est-ce que* comme un « moins bon français » que l'antéposition avec inversion, et en faisait aussi une forme de phrase généralement moins « adaptée » à l'interaction (moins adaptée même que le type in situ). Ce marquage était plus accentué en contexte formel. Une hypothèse peut être que cette évolution différente de jugements entre groupes en fonction du contexte, ici, soit reliée. Pour les plus de 30 ans, l'utilisation par une personne de l'antéposition en *est-ce que* en contexte formel, alors que c'est une forme de phrase que ce groupe juge éloignée du standard linguistique, conduirait à considérer cette personne comme

moins éduquée. Par contraste, il y aurait une relative indifférence chez les moins de 30 ans à ces enjeux.

5.3.3.4.9 Résultats généraux (échelle ORIGINE)

Le Tableau 5.45 et la Figure 5.58 donnent les moyennes globales brutes des jugements rendus sur chacune des trois formes interrogatives analysées, pour la dernière échelle proposée aux volontaires.

À nouveau, il est à bien noter que cette échelle est inversée par rapport à celle utilisée dans l'étude MG2. Parce que les volontaires devaient cliquer à gauche pour marquer une origine urbaine de la personne, ce sont les jugements numériquement négatifs qui y sont ici associés.

Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
-0.35	-0.59	-0.18

TABLEAU 5.45 – Jugement moyen en fonction du type interrogatif (MG3, ORIGINE)

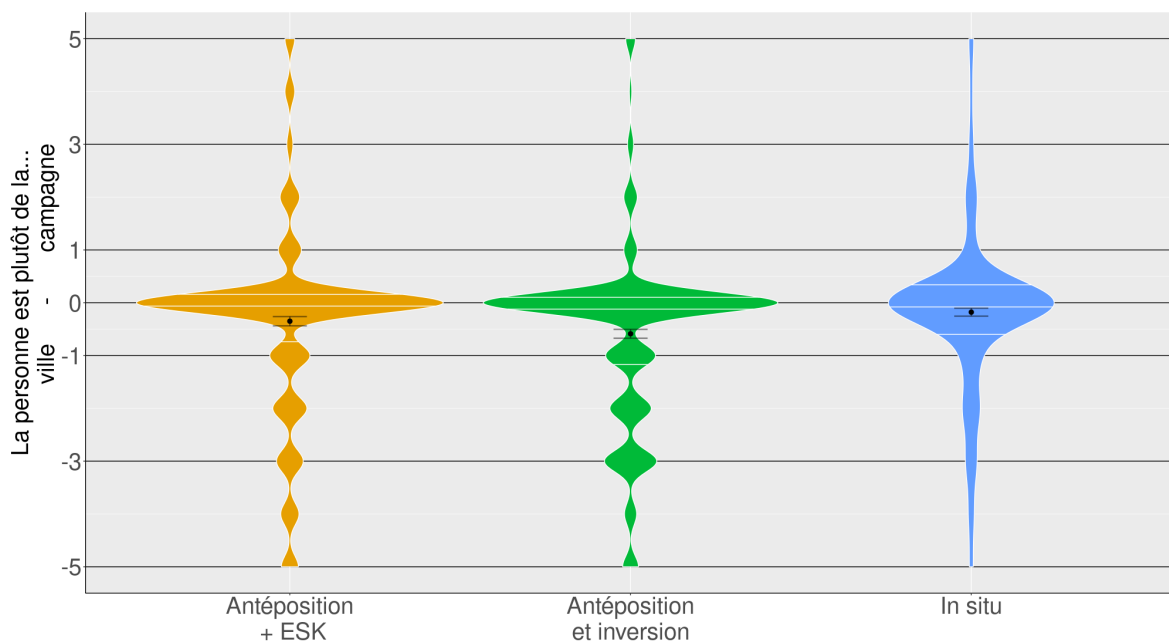


FIGURE 5.58 – Jugement moyen en fonction du type interrogatif (MG3, ORIGINE)

Le patron suivi par ces données est différent des deux précédents. Il y a moins d'écart entre les échelles, et les jugements ont une distribution très centrale, quel que soit le type interrogatif étudié. Cependant, ici encore, l'antéposition avec inversion semble se détacher en tant que type le plus associé avec une origine urbaine. C'est un résultat conforme à ce qui était observé dans l'étude MG2. Point de différence cependant, dans l'étude MG2 les phrases in situ étaient également associées à une origine urbaine, par contraste avec l'antéposition simple. Ici, le type in situ est celui qui est le plus neutre sur cette échelle, et à nouveau l'antéposition en *est-ce que* semble en position intermédiaire, entre les deux.

5.3.3.4.10 Un effet d'âge? (échelle ORIGINE)

Le Tableau 5.46 et la Figure 5.59 donnent les moyennes globales brutes des jugements pour chacune des trois formes interrogatives, pour la troisième échelle à nouveau.

	Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
-30 ans	-0.47	-0.65	-0.15
+30 ans	-0.17	-0.50	-0.22

TABLEAU 5.46 – Jugement moyen en fonction du type interrogatif et de l'âge (MG3, ORIGINE)

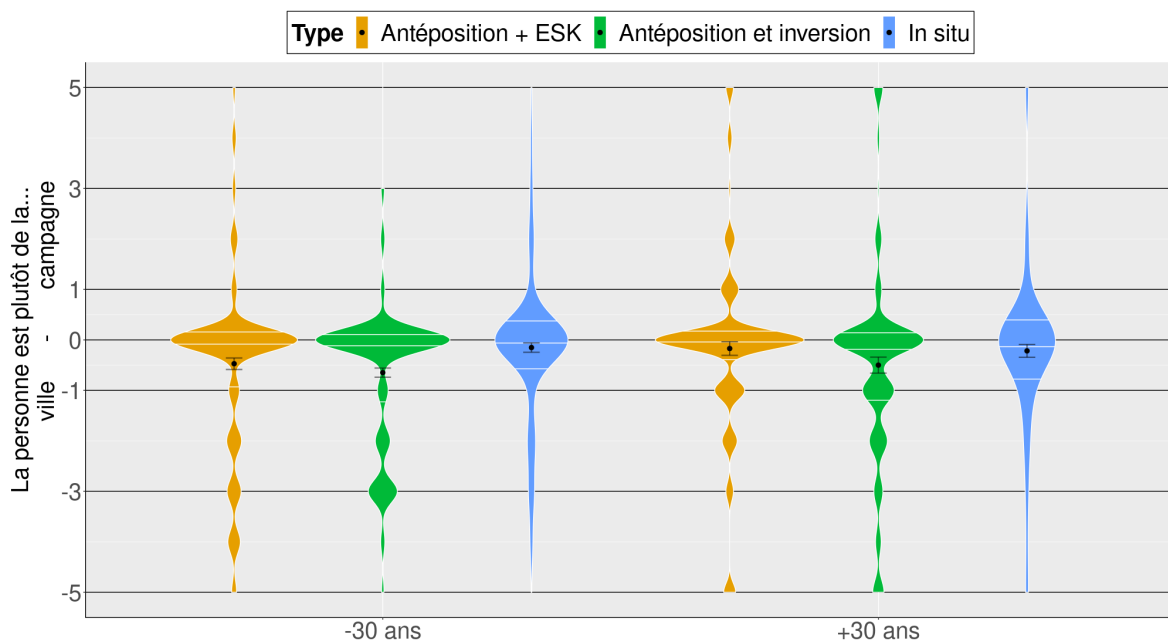


FIGURE 5.59 – Jugement moyen en fonction du type interrogatif et de l'âge (MG3, ORIGINE)

La différence de jugements entre groupes d'âge apparaît ici marginale. Ici aussi les distributions sont très centrées autour de 0, pour les deux groupes, pour presque tous les types. Seuls semblent s'en détacher les jugements des moins de 30 ans liés aux antépositions en *est-ce que*.

5.3.3.4.11 Un effet de contexte? (échelle ORIGINE)

Le Tableau 5.47 et la Figure 5.47 donnent les moyennes brutes des jugements pour chacune des trois formes interrogatives, pour la troisième échelle, en fonction du contexte.

	Antéposition en <i>est-ce que</i>	Antéposition et inversion	In situ
Contexte formel	-0.40	-0.48	-0.14
Contexte informel	-0.30	-0.70	-0.21

TABLEAU 5.47 – Jugement moyen en fonction du type interrogatif et du contexte (MG3, ORIGINE)

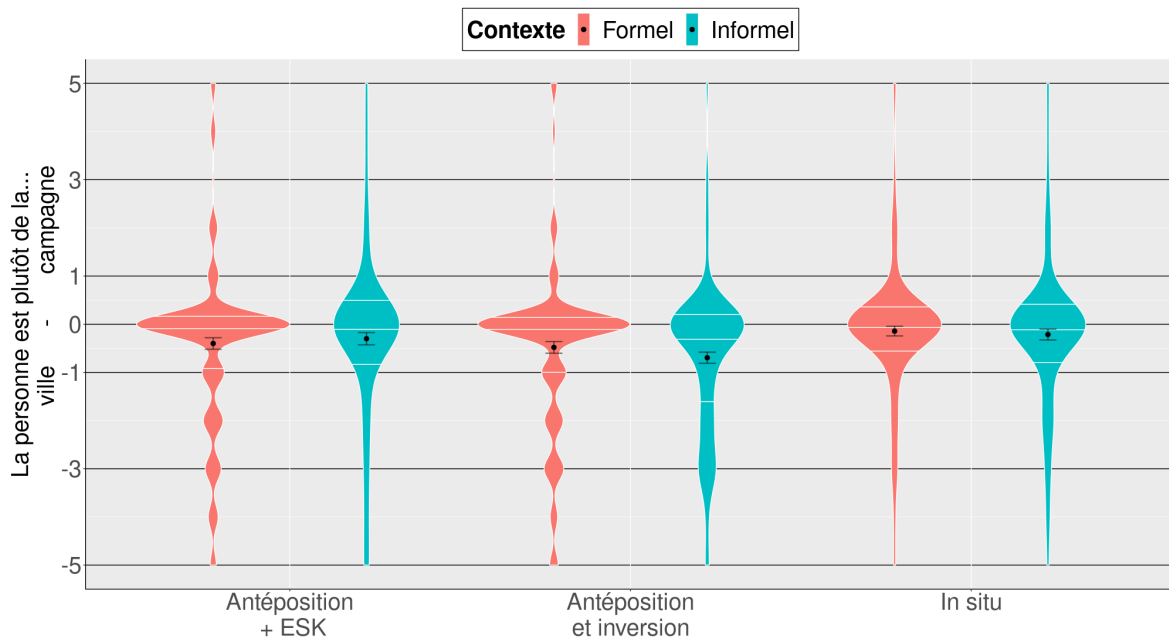


FIGURE 5.60 – Jugement moyen en fonction du type interrogatif et du contexte (MG3, ORIGINE)

Il apparaît ici que les interrogatives avec antéposition et inversion sont le plus associées à une origine urbaine dans des contextes informels. Comme dans ces contextes les antépositions en *est-ce que* sont pour leur part associées à des jugements au contraire plus hauts, l'écart entre les deux types se creuse dans cette condition.

5.3.3.4.12 Une interaction entre âge et contexte (échelle ORIGINE)

Le Tableau 5.48 et la Figure 5.61 approfondissent les résultats en synthétisant les moyennes des jugements donnés sur chaque type phrastique, en fonction du contexte mais aussi du groupe d'âge.

	-30 ans		+30 ans	
	Formel	Informel	Formel	Informel
Antéposition en <i>est-ce que</i>	-0.45	-0.50	-0.33	-0.01
Antéposition et inversion	-0.65	-0.65	-0.24	-0.76
In situ	-0.18	-0.12	-0.08	-0.35

TABLEAU 5.48 – Jugement moyen relatif aux types interrogatif, contexte et âge (MG3, ORIGINE)

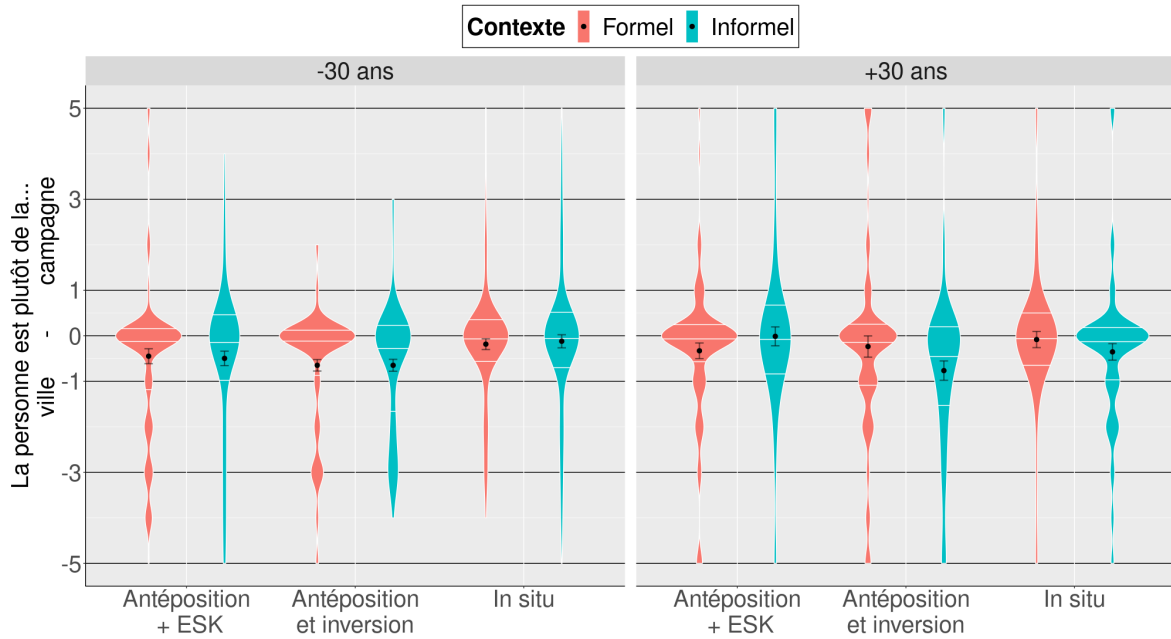


FIGURE 5.61 – Jugement moyen relatif aux types interrogatif, contexte et âge (MG3, ORIGINE)

Ces données illustrent que l'effet de contexte observé semble dû aux plus de 30 ans. Ce groupe montre le plus de différence dans ses jugements en fonction du contexte, pour les antépositions avec inversion. Il fait aussi une différence en fonction du contexte pour les autres types : l'in situ est associé à une origine plus urbaine en contexte informel, les antépositions en *est-ce que* le sont en contexte formel. Les moins de 30 ans modulent moins leurs jugements suivant le contexte.

5.3.3.4.13 Analyses bayésiennes (échelle ORIGINE)

Un modèle de régression bayésienne a été calculé sur la base des jugements fournis sur cette échelle de l'origine géographique, à nouveau strictement parallèle aux modèles calculés pour les échelles d'âge et des études (Annexe L.0.6). La Figure 5.62 en illustre les distributions postérieures.

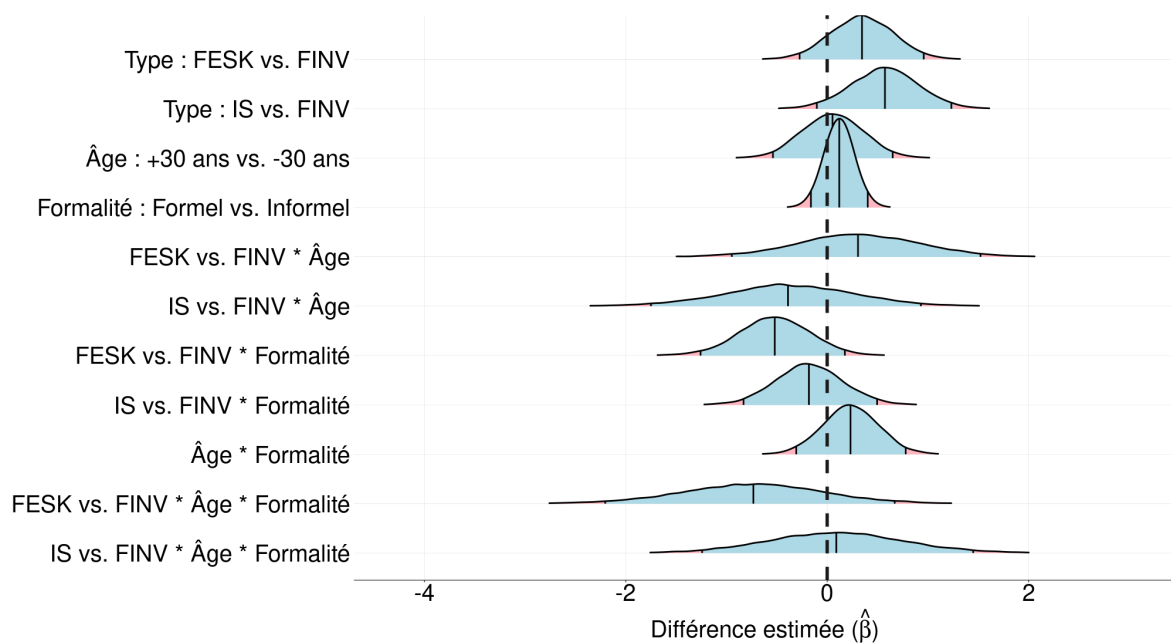


FIGURE 5.62 – Distributions postérieures pour le modèle brms (MG3, ORIGINE)

L'effet simple de TYPE syntaxique est confirmé. Les antépositions en *est-ce que* sont de manière générale associées à des jugements plus hauts que les antépositions avec inversion ($\hat{\beta} = 0.35$, 95%CrI=[-0.27,0.96], $P(\beta) > 0 = 0.86$). C'est aussi le cas, de façon plus robuste, pour les phrases in situ ($\hat{\beta} = 0.57$, 95%CrI=[-0.10,1.23], $P(\beta) > 0 = 0.95$).

Un effet de CONTEXTE est modérément corroboré par les données ($\hat{\beta} = 0.12$, 95%CrI=[-0.54,0.65], $P(\beta) > 0 = 0.80$), qui indique des jugements généralement plus hauts à travers les conditions syntaxiques dans les contextes formels. Ceci se prolonge dans des interactions TYPE*CONTEXTE. Les antépositions avec inversion sont liées à des jugements plus bas que les antépositions en *est-ce que* en contexte informel. En contexte formel, les antépositions avec inversion sont liées à des jugements plus hauts alors que ceux pour les antépositions en *est-ce que* ne varient pas beaucoup. La différence de jugements entre contextes est donc plus faible pour celles-ci ($\hat{\beta} = -0.52$, 95%CrI=[-1.26,0.18], $P(\beta) < 0 = 0.93$). Il semble que ce soit également le cas, mais sous forme de tendance à peine, pour les phrases in situ ($\hat{\beta} = -0.18$, 95%CrI=[-0.83,0.50], $P(\beta) < 0 = 0.70$).

Une légère interaction TYPE*ÂGE est également visible pour les phrases in situ ($\hat{\beta} = -0.39$, 95%CrI=[-1.75,0.93], $P(\beta) < 0 = 0.72$). Elle traduit le fait que les jugements pour les antépositions avec inversion sont proportionnellement plus élevés que ceux pour les phrases in situ chez les plus de 30 ans, par rapport aux moins de 30 ans. Ce phénomène n'est néanmoins pas corroboré pour les antépositions en *est-ce que*, sans doute parce que dans le même temps les jugements rendus par ce groupe sur les antépositions avec inversion sont aussi plus bas que ceux rendus par les plus de 30 ans (avec une présomption un peu plus forte d'origine urbaine).

Enfin, une interaction triple TYPE*ÂGE*CONTEXTE est confirmée ($\hat{\beta} = -0.74$, 95%CrI=[-2.20,0.67], $P(\beta) < 0 = 0.85$). Elle est le reflet de ce que le groupe des plus de 30 ans semble plus moduler ses jugements sur l'origine géographique, en fonction du contexte, que les moins de 30 ans. Les antépositions avec inversion sont liées à des jugements plus bas en contexte informel, comme il a été vu, tandis que c'est l'inverse pour les antépositions en *est-ce que*.

5.3.3.5 Bilan : L'antéposition en *est-ce que*, des stéréotypes *intermédiaires*

La Figure 5.63 propose une visualisation du « profil-type » des personnes utilisant les trois types phrastiques analysés, sur le modèle des graphiques de bilan proposés pour l'étude MG2.

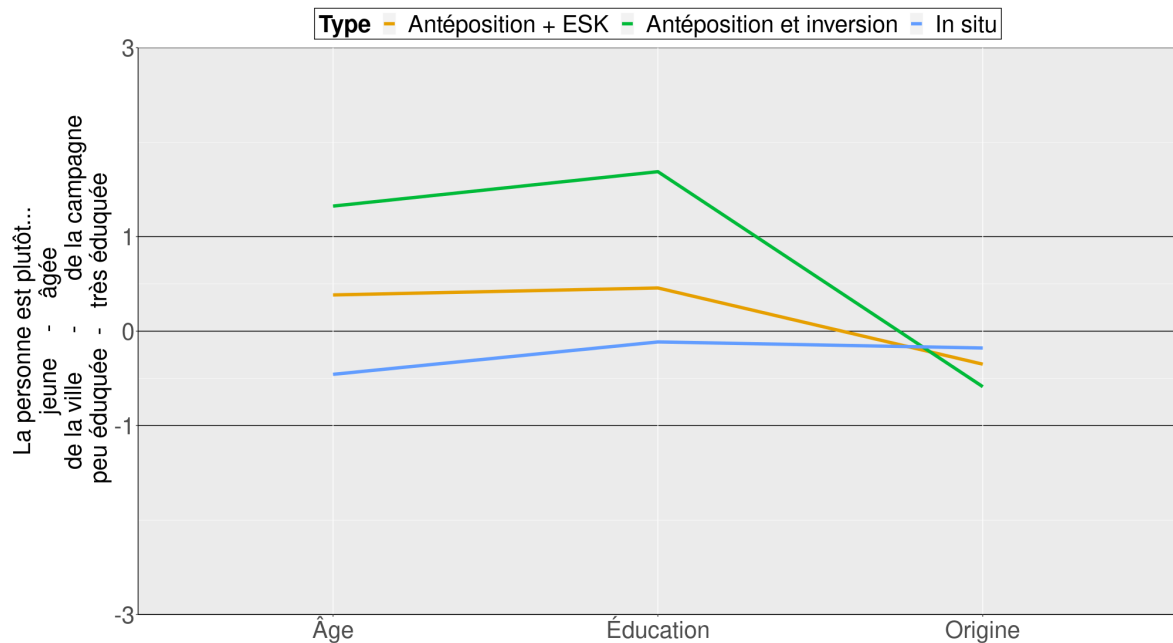


FIGURE 5.63 – Résultats généraux moyens pour toutes les échelles (MG3)

Même avec un nombre réduit d'échelles, cette étude a permis une évaluation des projections sociales liées à la variante interrogative avec antéposition en *est-ce que*, ce qu'illustre la Figure. Pour les trois stéréotypes (en abscisse), l'antéposition en *est-ce que* est associée à des jugements qui se situent entre ceux rendus pour les antépositions avec inversion et ceux rendus pour les phrases in situ. Ceci, même pour un stéréotype comme l'ORIGINE géographique, où les données sont plus nuancées. Ces constatations changent peu à travers les contextes, et peu à travers les groupes d'âge.

Ces résultats sont à mettre en relation avec les résultats de l'étude AJT4, dans lesquels ce type phrastique était aussi positionné *entre* l'antéposition avec inversion et les phrases in situ, pour ce qui est du « bon français ».

La valeur socialement « moyenne » de ce type phrastique fait ainsi ressortir l'opposition déjà observée, étude après étude, entre le type in situ et le type avec antéposition et inversion. L'association entre ce dernier et le stéréotype d'un âge plus avancé, d'une éducation plus longue, et dans une moindre mesure d'une origine plus urbaine, est en effet confirmée. En comparaison, les phrases in situ semblent se situer à un autre bout du réseau des représentations sociales.

5.3.4 Études de *matched-Guise* : Récapitulatif

La série des études MG1/MG2/MG3 décrite dans les sections précédentes apporte de nombreux éléments à plusieurs niveaux.

D'abord, du point de vue méthodologique, elles confirment qu'il est possible d'explorer les ramifications sociales du langage de manière systématique. Les résultats reproduits sur les trois études, par exemple le statut particulier des antépositions avec inversion, tendent à montrer aussi la valeur de reproductibilité de ce protocole, à l'oral comme à l'écrit.

Ensuite, ces études ouvrent une nouvelle perspective pour l'étude d'un phénomène d'alternance linguistique comme les questions partielles du français. Les jugements d'acceptabilité des Sections 5.2.1 à 5.2.5 ont permis de caractériser les préférences des francophones en matière d'interrogatives, mais elles ont fait appel pour cela à leur conscience explicite de locuteurs et locutrices du français. Ces tâches requièrent une méta-analyse sur le signal linguistique : « *est-ce que cette phrase est bien formée?* ». Les tâches de *matched-guise* permettent un accès plus profond, et en même temps plus direct, aux représentations mentales qui s'associent à l'interrogation linguistique : « *dans quelles catégories sociales puis-je placer cette personne que j'entends ou que je lis?* ».

Par ce biais, les études MG1/2/3 ont permis l'accès à un pan rarement envisagé de l'analyse des différentes formes interrogatives du français. Les diverses variantes disponibles pour les francophones ne sont pas interchangeables, ce qui a déjà été proposé par exemple par des études sur la structure informationnelle. Une partie des raisons pour lesquelles elles ne sont pas interchangeables réside dans leur usage *spécialisé*. Non seulement certains types phrastiques sont du « meilleur français » que d'autres, mais ils sont en plus associés à des stéréotypes sociaux très variés. Les choix faits par l'individu dans sa production linguistique, spontanée ou non, ne dépendent pas que de facteurs liés à l'énoncé en lui-même. Les enjeux sociaux de la situation d'interaction sont aussi un facteur influençant ces choix. Et, au premier rang de ces enjeux se trouve le maintien ou la modification de la *persona* [OCHS, 1992], ce masque social, au cours de l'échange. Si une variante linguistique, en l'occurrence une forme interrogative, est intégrée par les individus comme étant liée à des stéréotypes sociaux, que ceux-ci soient négatifs ou positifs, alors il devient possible de l'utiliser comme tel. Que ce soit à un niveau conscient ou non, maîtriser la variation et avoir intégré ses aspects devient ainsi un outil, qui module l'image des personnes participant à un échange.

5.4 Conclusion intermédiaire

Les études de corpus menées au Chapitre 4 avaient fait émerger l'hypothèse d'une utilisation spécialisée des interrogatives partielles du français, reposant sur des contraintes non pas internes à l'énoncé linguistique, mais externes. L'emploi des interrogatives varie selon le groupe social, selon la formalité du contexte, selon les objectifs et l'argumentation du discours, ou encore selon l'époque. Les études expérimentales menées dans ce chapitre ont permis d'explorer systématiquement certains de ces aspects.

D'abord, les études de jugements d'acceptabilité ont permis d'explorer les préférences des francophones et le regard, explicite et métalinguistique, porté sur les différentes variantes interrogatives à leur disposition. Il en ressort que la forme avec antéposition et inversion, notamment, relève d'un « meilleur » français que les autres types interrogatifs (AJT2, AJT3, AJT4). Il semble exister sur ce point un gradient. Les phrases *in situ* sont du moins bon français que les antépositions avec inversion (AJT1, AJT2, AJT3, AJT4). Les phrases avec antéposition simple sont du moins bon français que les phrases *in situ* (AJT1, AJT2, AJT3). Les phrases avec antéposition en *est-ce que* sont du moins bon français que les phrases avec antéposition et inversion, mais du meilleur français que les phrases *in situ* (AJT4). En combinant tous ces résultats, c'est le rapport des francophones à la langue standard qui se détaille. La structure interrogative standard du français est l'antéposition avec inversion.

Mais ce « standard » n'est pas toujours adapté en situation d'interaction. Utilisé à l'oral, dans

des situations quotidiennes, le type avec antéposition et inversion n'est pas le plus adapté (AJT3, AJT4). Il semble même parfois « mal adapté », dans des contextes informels ou pour certains groupes sociaux (définis par l'âge par exemple) (AJT3, AJT4). Ceci renvoie au faible usage de cette structure observée dans des corpus contemporains de français oral spontané. Si la question est posée explicitement, c'est-à-dire en affinant la définition de l'acceptabilité et en précisant s'il s'agit d'évaluer du « bon français » ou s'il s'agit de définir le français qui est « réellement utilisé », des réponses informatives peuvent être obtenues sur la discordance entre norme linguistique francophone et usages interrogatifs. Au quotidien, la structure la plus adaptée pour poser une question en français est peut-être la phrase in situ (AJT3, AJT4). En contexte informel, la structure avec antéposition en *est-ce que* est peut-être aussi adaptée (AJT4).

Ensuite, les études de *matched-guise* ont permis d'explorer les raisons de cette discordance. Le réseau des stéréotypes associés aux différentes formes interrogatives s'est ainsi ouvert. La structure avec antéposition et inversion, par exemple, est associée à la richesse, à l'éducation, à la ville par opposition à la campagne, à la lecture par opposition binaire aux sports, à l'âge. Ce, à un degré supérieur aux autres formes interrogatives, notamment le type in situ mais aussi l'antéposition simple (MG1, MG2). En revers, elle est associée à une moins grande sociabilité (MG2), ce qui vaut encore plus en contextes informels. Ces contextes informels réduisent par ailleurs, de manière générale, les oppositions entre types (MG2, MG3). Enfin, pour une vision complète, l'antéposition en *est-ce que* semble associée à des stéréotypes de degré « intermédiaire », à nouveau, entre l'antéposition avec inversion et le type in situ (MG3).

Ces études n'ont pas permis de départager les raisons de ces associations entre formes linguistiques et usages du français d'une part, et entre formes linguistiques et stéréotypes sociaux d'autre part. Il est probable, comme le montrent les études diachroniques (LARRIVÉE [2016, 2019], mais voir aussi le Chapitre 4), que ces associations se soient développées avec le temps. Quoi qu'il en soit, une fois mises au jour, leur existence doit être prise en compte.

Ces associations peuvent et doivent être affinées : en testant d'autres stéréotypes, en contrôlant la comparabilité des groupes de volontaires sur des critères précis et autres que l'âge, en contrôlant mieux les effets de la modalité écrit/oral, ou encore en étudiant d'autres variantes interrogatives.

L'impact de ces associations doit aussi être intégré à des études qui traditionnellement ne les ont pas intégrées. Le Chapitre 6 propose ainsi d'étudier le développement de ces associations entre formes interrogatives et stéréotypes sociaux chez l'enfant.

5.5 Apports majeurs

1. Les études de jugements d'acceptabilité ont établi un gradient de « bonne formation » des interrogatives en français. L'antéposition avec inversion est la forme du « meilleur » français. L'antéposition en *est-ce que* vient après, puis la forme in situ, puis l'antéposition simple.
2. Ces études ont aussi montré que les francophones ont conscience de la différence entre *norme linguistique* et *usage réel* des interrogatives. Toutes les formes ne peuvent pas s'utiliser indifféremment.
3. Notamment, l'antéposition avec inversion est moins adaptée en contexte informel. L'antéposition en *est-ce que* et le type in situ sont alors plus adaptées.
4. Les différentes formes interrogatives du français sont associées à des stéréotypes sociaux différents. Le profil type d'une personne utilisant l'antéposition avec inversion est ainsi : plus âgée, plus riche, plus éduquée, plus probablement lectrice, plus probablement originaire de la ville... qu'une personne utilisant les autres types. A contrario, elle est moins sociable, surtout en contexte informel. L'association des autres types avec ces stéréotypes semble se faire selon le gradient d'acceptabilité dégagé auparavant.
5. Le contexte et le groupe d'âge modifient l'attribution de ces stéréotypes sur les formes interrogatives.

Chapitre 6

Nouvelles données expérimentales chez l'enfant

Sommaire

6.1 Démarche et choix des protocoles retenus	177
6.2 Matched-Guise enfantin (MG-C) et stéréotypes sociolinguistiques	178
6.2.1 Hypothèses et prédictions	178
6.2.2 Protocole et items expérimentaux	178
6.2.3 Participantes et participants	179
6.2.4 Résultats	180
6.2.4.1 En surface : une différence entre types phrastiques	181
6.2.4.2 Des différences entre stéréotypes	182
6.2.4.3 Des interactions entre types et groupes d'âge	186
6.2.4.4 Bilan : Un profil stéréotypique évolutif des interrogatives partielles	192
6.3 Adaptation sociolinguistique en production : Activité de jeux de rôles (JDR-C)	193
6.3.1 Hypothèses et prédictions	193
6.3.2 Protocole - Constitution du corpus	194
6.3.3 Données générales sur le corpus	196
6.3.4 Productions recueillies et transcriptions	196
6.3.5 Annotations syntaxiques et résultats	198
6.3.5.1 Un effet de formalité	199
6.3.5.2 Un effet d'âge (classe) en interaction avec la formalité	201
6.3.5.3 Un effet de genre en interaction avec la formalité	206
6.3.5.4 Nuances sur la notion de formalité	207
6.4 Conclusion intermédiaire	212
6.5 Apports majeurs	213

Ce chapitre présente les résultats de deux expériences menées sur des enfants francophones pour explorer leur préférence en matière d'interrogative partielle, à mettre en relation avec les données adultes présentées dans le chapitre précédent. La problématique principale est ici de savoir si les facteurs sociolinguistiques intervenant dans la variation présente chez l'adulte ont une influence chez l'enfant et, si oui, à partir de quand et suivant quelles étapes. Après un exposé de la démarche générale et des protocoles expérimentaux retenus, les résultats d'une expérience de *matched-Guise* ainsi que d'une expérience de production spontanée sont présentés.

6.1 Démarche et choix des protocoles retenus

Les phénomènes de variation sociolinguistique ont été de manière générale peu explorés chez l'enfant, et encore moins de manière systématique avec des protocoles expérimentaux réitérables. Les raisons en sont multiples et dépendent en partie des phénomènes variationnels considérés.

Au sujet en particulier de la variation dans les interrogatives partielles, la tâche de *matched-Guise* s'est imposée assez rapidement pour tester l'intégration et la compréhension des stéréotypes sociolinguistiques par l'enfant à différents stades de son développement langagier. C'est un protocole déjà testé sur l'adulte avec succès (Chapitre 5), et conceptuellement simple : les enfants entendent des phrases, et doivent choisir qui les a prononcées entre des paires de personnages. Aux âges envisagés (à partir de 3-4 ans), la tâche est en réalité plus difficile à programmer techniquement qu'à mettre en place et à faire passer à un individu ou à un groupe d'individus. Que ce soit sur tablette ou sur ordinateur et pour peu qu'on prenne le temps de leur expliquer l'exercice (présenté comme jeu), la plupart des enfants ont facilement pris en main la tâche.

En termes de production, les contraintes étaient plus fortes. Étant dialogique par nature comme décrit dans le Chapitre 2, l'acte interrogatif est difficile à inciter d'une manière qui soit spontanée pour la personne testée et en même temps contrôlée pour une comparabilité des résultats entre groupes d'individus, ou pour qu'une analyse statistique soit possible par la suite. Une tâche ludique de jeu de rôle a été conçue, où l'enfant était dans la situation de recevoir des personnes dans son restaurant (personnages sous formes de marionnettes). L'idée était de plonger l'enfant dans une activité qui ressemblait le moins possible à un « exercice linguistique », ce qui a assez bien fonctionné, aussi bien au plus jeune âge (Grande Section, 4-6 ans) qu'au plus avancé (CM1, 9 ans).

Les deux études ont été menées dans la même école d'Argouges (Bayeux, Calvados). Il s'agit d'un établissement accueillant des enfants de maternelle et d'élémentaire. Quelques mois avant la mise en place, une convention a été élaborée entre la Direction Académique du Calvados et l'Université de Paris pour permettre l'entrée en école sur le temps scolaire et le Comité d'Éthique de la Recherche de l'Université Paris Descartes a validé en parallèle le projet scientifique (voir Annexe N). Les passations se sont faites entre septembre et novembre 2019, sur plusieurs sessions. La première étude a concerné la quasi intégralité des enfants de l'école de niveau Grande Section à CM2 (N=136), et était validée comme une activité pédagogique par les enseignantes concernées. La seconde étude a concerné une proportion moindre des enfants (N=69). Pour l'intégralité des enfants, la tâche de *matched-Guise* a été menée avant les vacances d'octobre, et la tâche de production après, afin de minimiser l'influence possible de la première sur la seconde. Un avantage à ce découpage a aussi été, lors de la seconde tâche, que les enfants étaient en confiance : le marionnettiste qui leur faisait face était inconnu, mais l'expérimentateur leur était déjà connu.

6.2 Matched-Guise enfantin (MG-C) et stéréotypes sociolinguistiques

136 enfants de l'école retenue ont contribué à la tâche de matched-guise qui leur était proposée. Dans celle-ci, présentée par leurs enseignantes comme un petit jeu linguistique, il leur était expliqué qu'il fallait aider des scientifiques à remettre de l'ordre dans un tas de fichiers mélangés par erreur. Ces données ont été récoltées à l'automne 2019, avant les vacances d'octobre.

6.2.1 Hypothèses et prédictions

Avec pour base les données récoltées chez l'adulte (étude MG2, Section 5.3.2), l'idée de ce protocole était de comparer la réaction des enfants, en réception, à des interrogatives partielles dans trois conditions syntaxiques différentes (les mêmes que celles observées chez l'adulte : *in situ*, antéposition simple, antéposition avec inversion). L'approche visait à explorer deux hypothèses complémentaires :

- les enfants n'ont par définition pas intégré toutes les contraintes sociolinguistiques liées à l'utilisation des interrogatives partielles en français. Les associations entre variante interrogative et stéréotype social ne devraient donc pas être les mêmes que chez l'adulte, voire être inexistantes ou illisibles sous un référentiel adulte.
- la compétence sociopragmatique se développe depuis le plus jeune âge jusqu'à l'adolescence (et même après), par un processus d'assimilation des règles sociales liées aux situations d'interaction dans lesquelles baigne l'individu linguistique (voir Section 2.4). En conséquence, les associations entre variante interrogative et stéréotype social devraient évoluer au cours du développement linguistique de l'enfant sur les âges considérés (ici, de 4 à 11 ans). Les résultats chez les enfants en fin de primaire devraient, au moins en partie, se rapprocher des résultats observés chez l'adulte.

6.2.2 Protocole et items expérimentaux

Présentée sous forme de jeu, la tâche consistait pour les enfants à écouter des sons très courts puis à choisir entre deux personnages dessinés celui qui avait parlé.

La complexité de la tâche adulte qui a servi de base (Section 5.3.2) ne permettait pas une transposition telle quelle pour des enfants à partir de 3 ans, incapables de manipuler en autonomie une souris d'ordinateur ou un dispositif de pointage quelconque avant 6 ans pour la plupart. Elle a été simplifiée. Une autre contrainte était le degré d'attention envisageable chez les enfants sur la durée et même sur un seul item. En conséquence, la tâche a été raccourcie (moins d'items), simplifiée (moins d'éléments visuels sur un écran), et limitée (moins de conditions dans le design). Si les items de base sont ceux de l'étude adulte MG2, la condition de CONTEXTE en a été retirée, pour ne se focaliser donc que sur la relation entre le TYPE syntaxique (*in situ*, *antéposition simple*, *antéposition avec inversion*) et trois stéréotypes sociaux (richesse, loisirs et éducation, comme étant les trois les plus susceptibles d'être compris par les enfants).

En tout, 18 items des expériences adultes ont été sélectionnés (notamment sur le critère d'être lexicalement accessible par des enfants, voir Annexe M.0.2). Seule la partie proprement interrogative de ces items a été conservée. Chacun de ces 18 items est passé trois fois, dans ses trois conditions syntaxiques. Pour écouter chaque item les enfants pouvaient cliquer - en autonomie ou avec l'aide de l'expérimentateur - sur un petit lecteur audio, en dessous duquel se trouvaient deux gros dessins

correspondant aux extrêmes des échelles stéréotypées utilisées chez l'adulte (Figure 6.1, Annexe M.0.3).



FIGURE 6.1 – Captures d'écrans (Étude MG-C, Stéréotype : RICHESSE)

Les trois variantes syntaxiques de chaque item étaient associées à trois stéréotypes sociaux différents au cours d'une même expérience, et les 18x3=54 phrases à faire écouter aux enfants étaient pseudo-randomisées. Ces 54 phrases étaient séparées en trois blocs successifs distincts, séparés par un écran de pause indiquant la progression du rangement des données et le degré de contentement des scientifiques (voir Annexe M.0.4). Ces écrans de pause avaient pour but à la fois de soulager les enfants et de leur permettre de relâcher leur attention un petit temps, mais aussi de vérifier si tout se passait bien. Au-delà des classes de CP les passations se déroulaient à trois ou quatre enfants en même temps, sur des postes informatiques différents et avec des casques audio les isolant les uns des autres. Chacun des trois blocs ne comprenait qu'une seule association TYPE-STÉRÉOTYPE de chaque item. Dans chaque liste, la moitié des items était lue par une femme et la moitié par un homme. Pour brasser les stimuli au maximum, autant en termes syntaxiques et stéréotypiques qu'au niveau de la voix, six listes ont été créées au total. Ces six listes ont été présentées de façon aussi équilibrée que possible au sein de chaque classe (Annexe M.0.1).

Il n'y avait pas de distracteurs, mais quatre phrases d'entraînement précédaient l'expérience à proprement parler. Le « jeu » était précédé d'un écran vierge avec quatre champs, remplis par l'expérimentateur en discutant avec l'enfant (âge, classe, genre, et un code aléatoire pour filtrer les données). Aucune autre donnée n'a été collectée.

6.2.3 Participantes et participants

Les participations de 136 enfants ont été enregistrées. Les seules informations disponibles à leur propos sont leur genre, leur âge et leur classe. 58 filles et 77 garçons ont participé, et 1 enfant qui n'a pas voulu préciser (label = « autre »). Les enfants avaient entre 3 et 11 ans ($\mu = 7.5$, $m = 7$), et appartenaient aux classes de l'école depuis la Grande Section jusqu'au CM2. Le détail de la répartition des enfants par classe, par niveau (la plupart des classes étant doubles), et par âge est donné dans les Tableaux 6.1, 6.2 et 6.3.

GS-1	GS-2	CP	CP-CE1	CE1-CE2	CE2-CM1	CM1-CM2-A	CM1-CM2-B
9	9	19	18	21	20	23	17

TABLEAU 6.1 – Répartition des enfants par classes (MG-C)

GS	CP	CE1	CE2	CM1	CM2
18	23	29	20	17	29

TABLEAU 6.2 – Répartition des enfants par niveaux (MG-C)

3 ans	4 ans	5 ans	6 ans	7 ans	8 ans	9 ans	10 ans	11 ans
1	3	14	23	28	20	24	22	1

TABLEAU 6.3 – Répartition des enfants par âges (MG-C)

Un mot aux parents avait été distribué par l'équipe pédagogique en amont des sessions expérimentales (voir Annexe N). La quasi-totalité des enfants des classes visées ont participé à cette première tâche, présentée aux enfants et intégrée comme une activité pédagogique par les enseignantes.

Avant les passations, qui duraient entre 10 et 20 minutes selon les enfants, et pendant celles-ci quand nécessaire, il était bien rappelé que le jeu n'était pas un jeu de rapidité et qu'il n'y avait pas de bonnes ou de mauvaises réponses, mais que seules leur réponse intuitive intéressait les scientifiques. Après les passations, un petit diplôme était délivré aux enfants pour les remercier de leur participation (voir Annexe M.0.5). Le diplôme était délivré indépendamment de tout autre paramètre, et y compris pour les enfants dont les résultats n'ont pas été enregistrés pour des raisons techniques et n'ont pas été comptabilisées (N=1, non-inclus dans les 136 mentionnés plus haut et par la suite).

6.2.4 Résultats

Les résultats de cette expérience sont à considérer en deux temps. En surface, il semble y avoir, chez les enfants pris en tant que groupe homogène, des associations entre forme linguistique et stéréotype social plus ou moins comparables à celles présentes chez les adultes. Il convient néanmoins de nuancer ceci, en examinant plus en détail les données, ce qui passe notamment par une analyse séparée de chacun des stéréotypes considérés, mais aussi par une séparation des enfants en groupes d'âge différents.

Les résultats présentés ci-après sont assortis d'analyses inférentielles bayésiennes, et portent sur une portion « nettoyée » des données. Hors phrases d'entraînement, chaque enfant devait effectuer $18 \times 3 = 54$ choix sur des items cibles. Pour 136 enfants, le nombre total de réponses recueillies était donc de 7344. Il est néanmoins apparu pendant les passations que les enfants qui voyaient deux fois la même image (le même stéréotype) de suite, avaient tendance à re-cliquer directement sur la même image, sans prendre le temps de vérifier que la phrase à écouter était différente (c'était pourtant toujours le cas, par design). À l'inverse, parfois cette succession poussait l'enfant à cliquer sur l'image opposée de façon automatique, comme s'il y avait une « erreur » à corriger. Par précaution, ce défaut du design impose de ne pas tenir compte des réponses où un même stéréotype a été présenté de façon redondante (N=1761).

Le temps de réaction pour chaque item étant par ailleurs enregistré, les choix effectués en moins de 2000 milli-secondes par les enfants ont également été retirés de l'analyse (N=308). De tels temps de réaction ne permettent pas d'avoir écouté les sons puis de les analyser puis de cliquer sur l'une ou l'autre image.

Au final, 5275 réponses ont été conservées pour analyse.

6.2.4.1 En surface : une différence entre types phrastiques

Pour rappel, au contraire des adultes qui avaient à leur disposition une échelle graduée de plusieurs niveaux pour situer la personne entendue sur chacun des stéréotypes, les enfants avaient des choix binaires à effectuer. Soit la personne entendue est pauvre, soit elle est riche ; soit elle fait souvent du sport, soit elle lit souvent ; soit elle est plus proche d'un caissier ou d'une caissière, soit elle est plus proche d'un ou d'une médecin. Des diagrammes en violon ne sont ici pas pertinents.

La Figure 6.2 permet une visualisation des choix effectués par les enfants après l'écoute des phrases selon leur TYPE syntaxique (*antéposition simple*, *antéposition avec inversion*, *in situ*). En ordonnée est donné, au-dessus de 0, le pourcentage de choix correspondant aux stéréotypes moins associés à la variante *in situ* chez l'adulte (*Riche*, *Lecture*, *Médecin*). En dessous de 0, il s'agit du pourcentage de choix correspondant aux stéréotypes plus associés à la variante *in situ* chez l'adulte (*Pauvre*, *Sports*, *Caissier* ou *Caissière*). Par exemple, sur 100% des phrases avec antéposition simple entendues par les enfants, il faut lire que dans 52.4% des cas leur choix s'est porté sur les images avec la personne pauvre, celle faisant du sport, ou celle qui était caissier ou caissière.

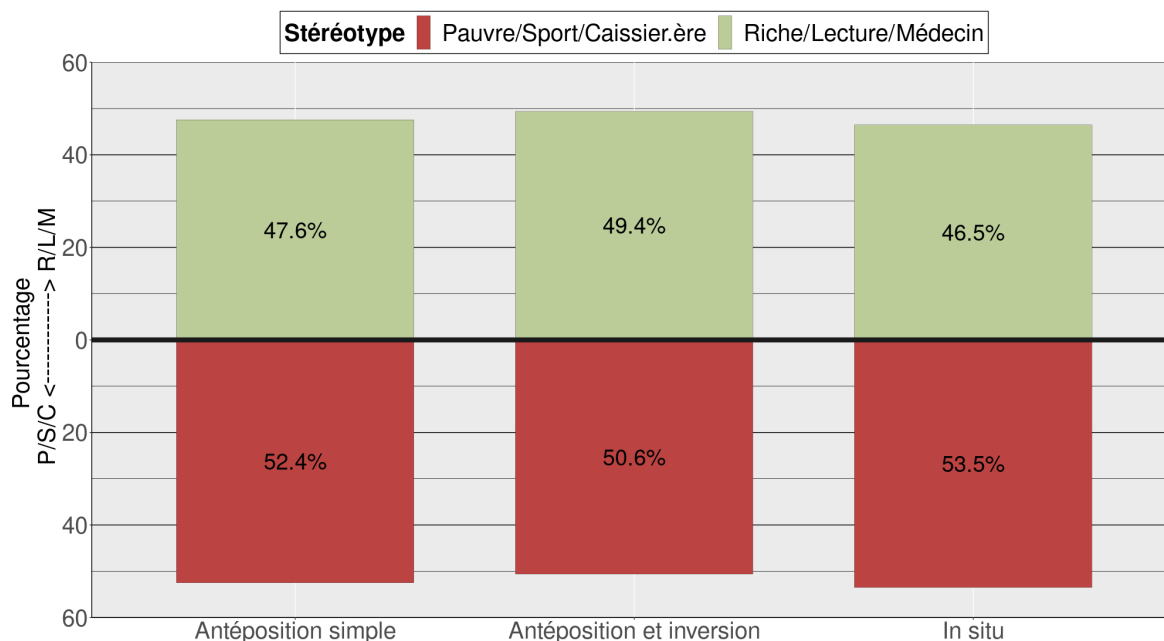


FIGURE 6.2 – Stéréotypes choisis en fonction du type phrastique (MG-C)

À ce niveau d'analyse, il semble y avoir pour le groupe des enfants une distinction entre les antépositions avec inversion d'une part et les antépositions simples et les phrases *in situ* d'autre part. C'est pour l'antéposition avec inversion que l'association avec des stéréotypes *Riche/Lecture/Médecin* est la plus fréquente, comparée aux deux autres types de phrase.

Un modèle de régression bayésienne a été calculé sur ces données (voir Annexe M.0.6). Le modèle avait pour variable dépendante la RÉPONSE choisie par l'enfant (les réponses *pauvre*,

sport, caissier ou caissière étant codées « 0 » et les autres « 1 »), et pour variable indépendante le TYPE syntaxique avec pour niveau de référence les phrases antéposées avec inversion. Des effets aléatoires ont été ajoutés pour tenir compte de la variabilité entre les enfants ("ID", avec pour pente le TYPE) et de la variabilité potentielle entre les items ("Item", avec pour pente le TYPE également). La variable dépendante étant binomiale, c'est la famille "bernoulli" qui a été sélectionnée pour faire tourner ce modèle, avec 4 chaînes et 3000 itérations par chaîne. La convergence des modèles a été atteinte et vérifiée en contrôlant que chaque Rhat soit à 1.00 pour tous les paramètres. Les distributions postérieures de ce modèle sont visibles en Figure 6.3.

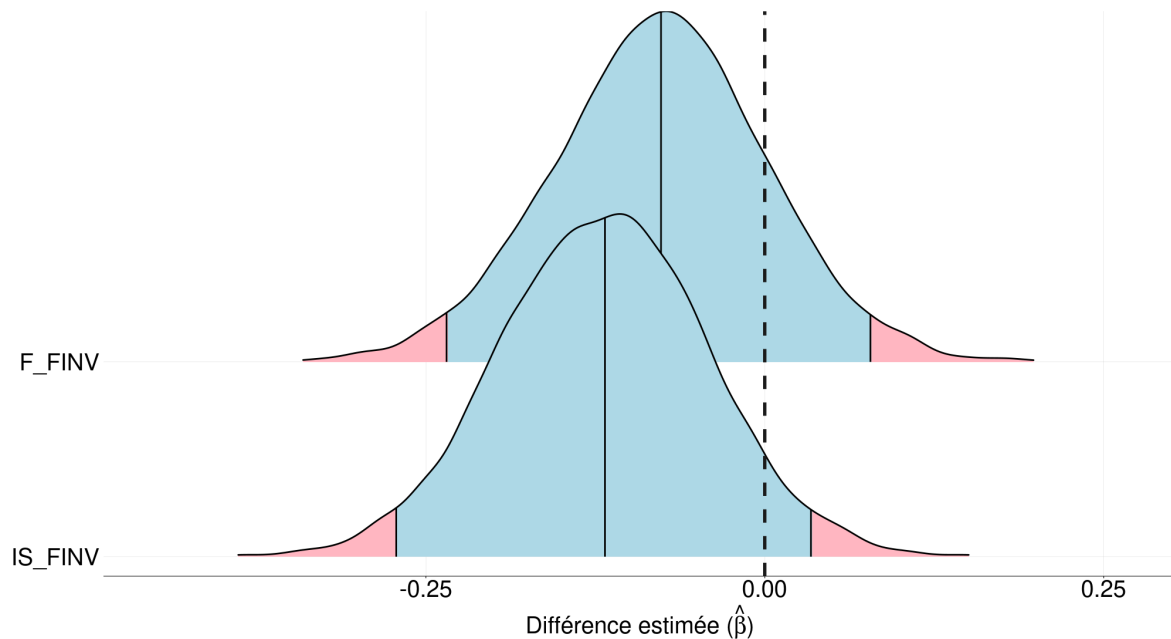


FIGURE 6.3 – Distributions postérieures pour le modèle brms (MG-C, effet TYPE)

Le modèle confirme de manière fiable que les interrogatives in situ sont moins associées aux stéréotypes tels *riche, lecture, médecin* que les antépositions avec inversion ($\hat{\beta} = -0.12$, 95%CrI=[-0.27,0.03], $P(\beta) < 0 = 0.94$). Ceci est également valable, mais de façon plus modérée, quand les antépositions simples sont comparées aux antépositions avec inversion ($\hat{\beta} = -0.08$, 95%CrI=[-0.23,0.08], $P(\beta) < 0 = 0.84$). Ceci rappelle à première vue les résultats obtenus sur la tâche de *matched-guise* avec stimuli audio (MG-2), et notamment le profil stéréotypique des utilisateurs de l'antéposition avec inversion qui s'en dégageait (Section 5.3.2, Figure 5.38, p. 140). En effet, sur ces trois stéréotypes, les adultes pris dans leur globalité associaient cette forme à un niveau de richesse plus élevé, une lecture plus fréquente, et des études plus longues (stéréotype explicitement associé à l'image du médecin dans la phase de préparation).

Ces résultats sont néanmoins à nuancer pour au moins deux raisons. D'une part, sous ce tableau général peuvent se dissimuler des différences entre stéréotypes. D'autre part, les enfants ayant participé à la tâche sont loin d'être un groupe homogène, ne serait-ce qu'à cause de la variation d'âge, et sans parler de leurs différences socioéconomiques ou socioculturelles.

6.2.4.2 Des différences entre stéréotypes

Regrouper les données en sous-groupes selon que les stéréotypes interrogés étaient la richesse, l'éducation ou les loisirs permet dans un premier temps d'affiner les associations effectuées par

les enfants entre formes linguistiques et stéréotypes. Il apparaît ainsi que ces trois stéréotypes ne fonctionnent pas de la même manière, même à un niveau global en considérant l'ensemble des enfants.

6.2.4.2.1 Stéréotype de la RICHESSE

La Figure 6.4 permet ainsi, sur le modèle de la Figure 6.2, de visualiser la proportion des choix effectués entre personne pauvre et personne riche, en fonction des formes interrogatives entendues (N=1786 réponses retenues pour analyse). Il semble que, si les deux types d'antépositions appellent des représentations sociales similaires, les phrases in situ s'en distinguent, par une association un peu plus fréquente avec les images de personnes riches (dans 52.3% des cas).

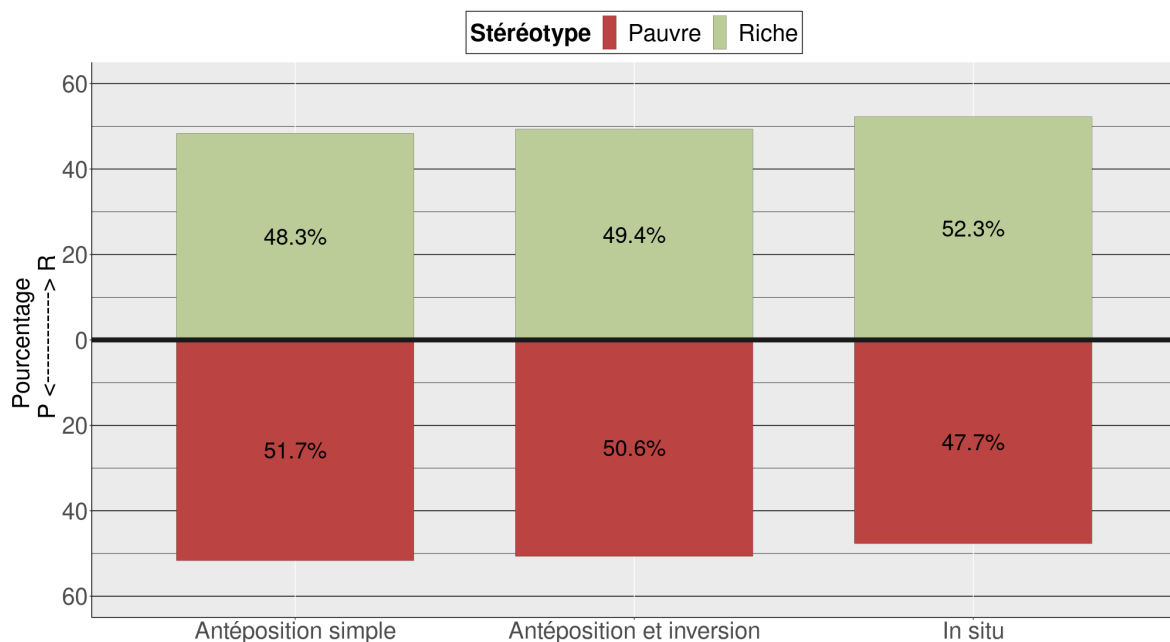


FIGURE 6.4 – Évaluation de la RICHESSE en fonction du type phrastique (MG-C)

Un modèle de régression bayésienne calculé sur ce sous-ensemble de données, avec les mêmes variables et les mêmes effets aléatoires que le précédent (voir Annexe M.0.7), supporte ces interprétations (de façon certes modérée). Comparées avec les antépositions avec inversion, les phrases in situ sont plus souvent associées au dessin de la personne riche ($\hat{\beta} = 0.12$, 95%CrI=[-0.14,0.39], $P(\beta) > 0 = 0.82$). À l'inverse, aucune différence n'est corroborée entre les antépositions avec inversion et les antépositions simples ($\hat{\beta} = -0.01$, 95%CrI=[-0.29,0.28], $P(\beta) > 0 = 0.47$).

Les résultats globaux masquaient ce résultat, qui entre en contradiction avec les données globales pour les adultes, dégagées dans la tâche MG-2 (Section 5.3.2). Une explication serait que, si et quand les enfants se construisent un stéréotype de la richesse, celui-ci n'est pas identique à celui des adultes. Suite aux résultats inattendus, à la fin de la tâche de production présentée en Section 6.3, il a été demandé aux enfants qui y participaient de lister quelques exemples de « personnes riches ». Parmi les réponses, il y avait beaucoup de personnalités du sport (football, course automobile, judo), de célébrités notamment présentes à la télévision (chanson, animation d'émission), ou encore de personnalités politiques très médiatisées comme le président français (en 2019, Emmanuel Macron) ou celui des États-Unis (en 2019, Donald Trump). De manière générale, la richesse semble donc pour les enfants associée à une réussite sociale mesurable en termes de

« visibilité », plus qu'à l'appartenance à une classe sociale privilégiée pour qui l'accès à la culture est facilité et le respect des normes linguistiques plus important, à l'inverse des adultes.

6.2.4.2.2 Stéréotype des LOISIRS

La Figure 6.5 permet de visualiser la proportion des choix effectués entre personne faisant souvent du sport et personne privilégiant la lecture, en fonction des formes interrogatives entendues (N=1746 réponses retenues pour analyse). Il apparaît que, si une différence existe entre formes interrogatives à ce stade, elle va plutôt dans le même sens que celle observée chez l'adulte, à savoir que les antépositions avec inversion sont plus souvent associées à une lecture fréquente qu'à du sport (dans 51.6% des cas contre 48.4%), à l'inverse des deux autres types phrastiques.

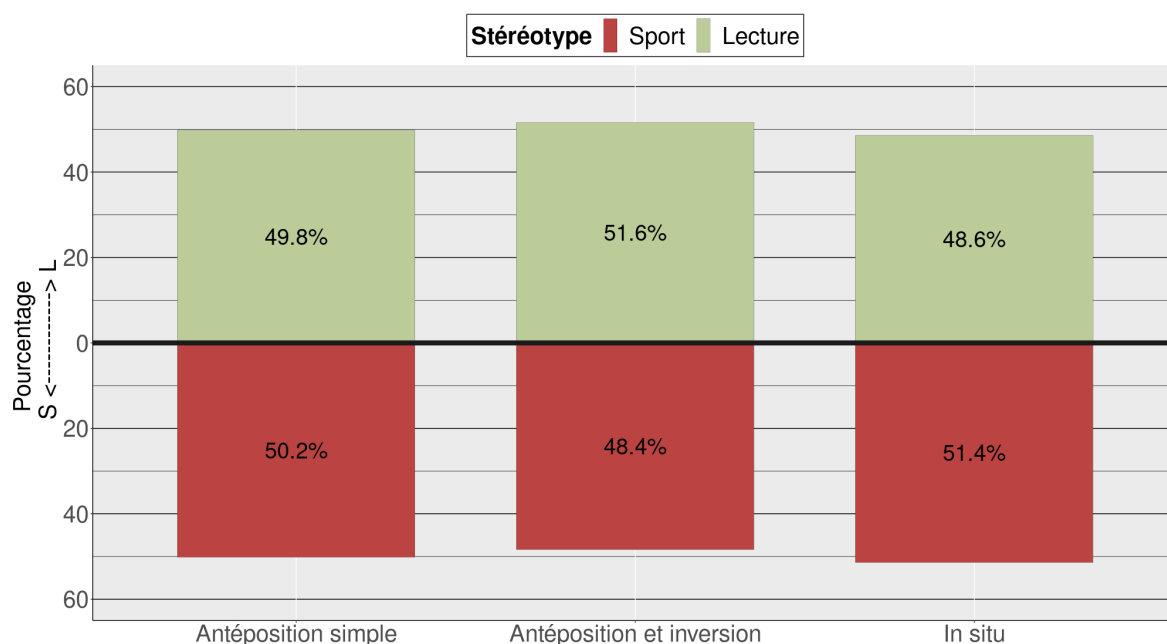


FIGURE 6.5 – Évaluation des LOISIRS en fonction du type phrastique (MG-C)

Un modèle de régression bayésienne calculé sur ce sous-ensemble de données, avec les mêmes variables et les mêmes effets aléatoires que les précédents (voir Annexe M.0.8), confirme ces tendances. C'est autant le cas quand les antépositions avec inversion sont comparées aux antépositions simples ($\hat{\beta} = -0.11$, 95%CrI=[-0.37,0.16], $P(\beta) < 0.79$), que quand elles sont comparées aux phrases in situ ($\hat{\beta} = -0.09$, 95%CrI=[-0.36,0.17], $P(\beta) < 0.75$). Ces valeurs ne permettent cependant pas plus que d'évoquer des tendances, dont la réalité serait à affiner.

6.2.4.2.3 Stéréotype de l'ÉDUCATION

La Figure 6.6 permet de visualiser la proportion des choix effectués entre personne peu éduquée et personne très éduquée, en fonction des formes interrogatives entendues (N=1743 réponses retenues pour analyse). Ces données peuvent se lire dans deux sens différents : soit une association plus forte des phrases antéposées avec inversion avec une personne médecin que pour les autres types (47.3% contre 44.5% et 38.4%), soit une association plus forte des phrases avec antéposition simple ou in situ avec une personne non-médecin. Quoi qu'il en soit, il semble y avoir une différence importante entre les antépositions avec inversion et les deux autres types phrastiques pour ce stéréotype. Il semble aussi que les phrases in situ soit particulièrement peu associées,

indépendamment des autres types phrastiques, avec une personne médecin (moins de 40% des réponses).

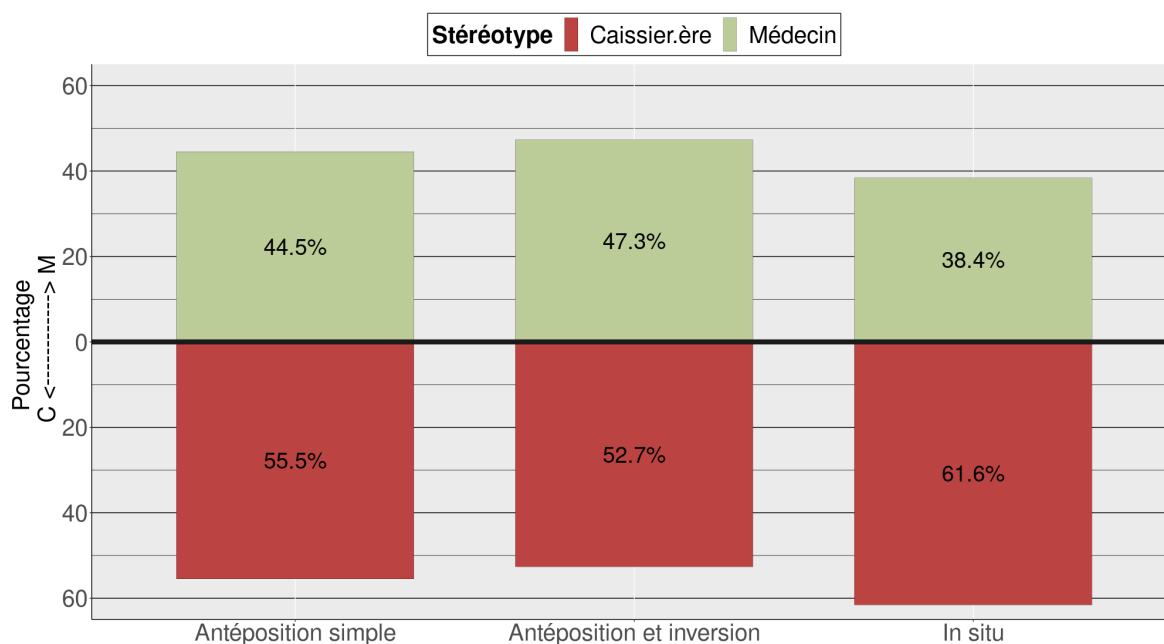


FIGURE 6.6 – Évaluation de l'ÉDUCATION en fonction du type phrastique (MG-C)

Un modèle de régression bayésienne calculé sur ce sous-ensemble de données, avec les mêmes variables et les mêmes effets aléatoires que précédemment (voir Annexe M.0.9), confirme ces tendances. En comparaison avec les antépositions avec inversion, les phrases in situ sont de manière fiable moins souvent associées aux images de médecin ($\hat{\beta} = -0.39$, 95%CrI=[-0.67,-0.12], $P(\beta) < 0=1$). Cette différence est similaire, quoique supportée plus modérément par les données, pour les antépositions simples ($\hat{\beta} = -0.11$, 95%CrI=[-0.38,0.15], $P(\beta) < 0=0.80$).

Il resterait cependant ici à affiner si c'est réellement le stéréotype des études et de l'éducation qui a été atteint à travers cette tâche. En effet, en complément d'exemples de personnes riches, il était demandé aux enfants participant à la tâche décrite en Section 6.3 le « métier qu'ils et elles voulaient faire plus tard ». Le cas échéant, en leur demandant ensuite si ils et elles savaient quelles études faire pour arriver à ce métier, ou à celui de médecin par exemple, il est souvent ressorti des discussions que, à part « bien travailler à l'école », le concept d'études courtes ou longues échappait à la plupart des enfants. Il faut donc s'interroger sur la valeur des résultats obtenus sur ce stéréotype, non pas pour les rejeter, mais plutôt en ce qu'ils sont sans doute le reflet d'une situation interactionnelle particulière, qui est celle d'aller chez le médecin. Dans cette situation, à laquelle les enfants ont déjà dû se confronter quel que soit leur âge, il semble déjà intériorisé que l'adulte parle un français « un peu différent de d'habitude », avec notamment moins de phrases in situ. Ceci, par opposition avec des situations où par exemple l'enfant fait les courses avec ses parents et les entend discuter avec le caissier ou la caissière du magasin.

6.2.4.2.4 Bilan intermédiaire

La première hypothèse de cette étude était que les associations entre variante interrogative et stéréotype social ne seraient pas les mêmes, par définition, chez les enfants que celles dégagées chez l'adulte, voire seraient inexistantes. En les considérant comme un groupe homogène, il ressort

de cette analyse séparée des trois stéréotypes qu'il existe bien des associations variante-stéréotype chez l'enfant. En parallèle, effectivement, ces associations ne sont pas toujours aussi fortes, et elles ne vont pas toujours dans le même sens que chez l'adulte. Observer les données en les reliant au niveau de développement des enfants peut permettre d'y voir plus clair, en faisant le pont avec la seconde hypothèse.

6.2.4.3 Des interactions entre types et groupes d'âge

Cette seconde hypothèse était que les associations entre variante interrogative et stéréotype social évoluent avec le temps, et que les plus jeunes enfants n'auront pas intégré les mêmes associations que celles et ceux des classes plus avancées en âge.

Le critère retenu ici pour distinguer parmi les enfants est en réalité proche de celui retenu par l'établissement scolaire pour son organisation interne : des classes mixtes, avec une répartition des enfants certes basée sur leur âge, mais également en fonction de leurs capacités (des élèves de CP qui seraient plus proches du niveau d'élèves de CE1 dans une classe de CP-CE1, par exemple, etc.). Plutôt que de regrouper les enfants par âge, ou par niveau absolu (Grande Section, CP, CE1, CE2, CM1, CM2), ce sont des groupes un peu intermédiaires qui semblent les plus pertinents après analyse.

Les 5275 réponses retenues pour analyse ont ainsi été réparties sur quatre « groupes-classe », dégagés depuis les classes du Tableau 6.1. Ces quatre groupes séparent les élèves de Grande Section (GS), les élèves de CP stricts (CP), les élèves de CM1-CM2 (CM1CM2) et les élèves des niveaux intermédiaires (CE1CM1). Il est à noter que ce dernier niveau, malgré son nom, comprend 4 élèves de CP. Le Tableau 6.4 présente précisément la répartition des réponses avec le filtre « groupe-classe ».

Groupe-classe	GS	CP	CE1CM1	CM1CM2
Nombre total d'enfants	18	19	59	40
Nombre d'enfants par niveau officiel	GS(18)	CP(19)	CP(4) CE1(29) CE2(20) CM1(6)	CM1(11) CM2(29)
Réponses retenues	677	741	2282	1575

TABLEAU 6.4 – Répartition des réponses retenues pour analyse par groupes-classe (MG-C)

Il y a une disparité manifeste entre les groupes, qui limite nécessairement la portée statistique des analyses inférentielles présentées ci-après, mais dont elles tiennent compte au maximum (analyses bayésiennes compatibles avec une forte versatilité des données et avec prise en compte d'effets aléatoires). Cette disparité est le reflet de la disparité des constitutions de classes au sein des établissements scolaires (avec ici, pour rappel, 29 CM2 et 29 CE1 contre à peine 17 CM1 et 20 CE2).

Ces groupes-classe permettent néanmoins de globalement séparer les enfants en fonction de leur rapport au monde extérieur. Les stéréotypes se développent en effet le plus dans l'interaction avec le monde incluant des adultes, hors famille (par des effets de saillance et des généralisations sur ces éléments saillants répétés plusieurs fois, voir Section 2.4.3.2 et BUSON [2009]; BUSON et al. [2014], inspirés de BYBEE [2013]; BYBEE et HOPPER [2001]). Or, de ce point de vue, les Grande Section sont des élèves dont la scolarisation commence à peine (avec des arrivées parfois en dernière année

de maternelle, et des modes de relation très disparates entre les enfants, lesquels entraînent des relations avec l'adulte, et notamment les enseignantes, encore très individualisées). Les CP en sont au premier stade d'une scolarisation à part entière, et leur mode de relation avec l'adulte devient plus homogène (même si la période de passation se situait en début d'année scolaire). Le groupe des CE1CM1 en est au milieu de la scolarisation à l'école primaire et constitue un groupe, certes encore hétérogène, mais surtout intermédiaire dans les processus d'apprentissage et de socialisation, encore très supervisés par les adultes. Le groupe de CM1CM2, enfin, est constitué d'enfants qui entrent dans la phase finale de leur scolarisation initiale (ce sont les « grands et les grandes » de l'école, par opposition aux plus jeunes), ce qui se traduit par plus d'indépendance vis-à-vis des adultes.

6.2.4.3.1 Stéréotype de la RICHESSE

La Figure 6.7 permet de visualiser l'évolution des associations entre les variantes interrogatives et les images de personnes pauvres ou personnes riches (N=1786), en fonction du groupe-classe des enfants. Par exemple, alors que les enfants de Grande Section ont associé les antépositions simples à l'image d'une personne riche dans 40.5% des cas, les enfants de CM1CM2 les y ont associées dans 48.3% des cas.

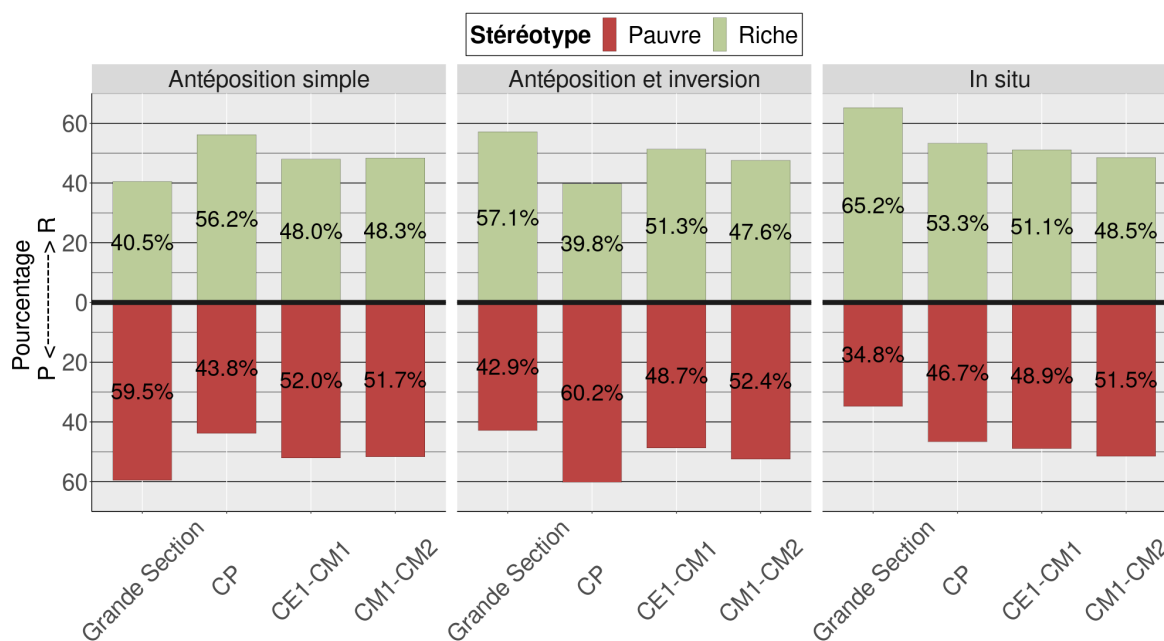


FIGURE 6.7 – Évaluation de la RICHESSE, par TYPE et CLASSE (MG-C)

Ces données indiquent que ces associations entre forme linguistique et stéréotype social ne sont pas figées dans le temps et qu'il existe une forte variabilité selon le groupe d'enfants considéré. Il est intéressant d'observer une grande instabilité entre les enfants des trois premiers groupes, avec les CP qui associent moins les antépositions simples aux images de personnes pauvres en comparaison des autres groupes. À l'inverse, ces mêmes CP associent moins les antépositions avec inversion aux images de personnes riches que les autres enfants.

Mais, plutôt qu'une comparaison groupe à groupe, ce sont les tendances globales qui sont à relever. Et, notamment, il apparaît ici que les interrogatives in situ sont de moins en moins associées aux images de personnes riches, à mesure que les enfants avancent dans les groupes-classe. C'est

même le cas entre le groupe CE1CM1 et CM1CM2, qui comprennent le plus d'enfants.

Un modèle de régression bayésienne a été calculé sur ces données (voir Annexe M.0.10). Comme pour les modèles précédents, celui-ci avait pour variable dépendante la RÉPONSE choisie par l'enfant (« 0 » pour *pauvre* et « 1 » pour *riche*). Cette fois la variable indépendante était l'interaction entre TYPE syntaxique (niveau de référence : antépositions avec inversion) et le groupe-classe (CLASSE) considéré comme une variable ordinale. Les mêmes effets aléatoires ont été intégrés : les enfants ("ID", avec pour pente le TYPE) et les items ("Item", avec pour pente l'interaction TYPE*CLASSE également). C'est toujours la famille "bernoulli" qui a été sélectionnée, avec de nouveau 4 chaînes et 3000 itérations par chaîne. La convergence des modèles a été atteinte et vérifiée selon des critères inchangés (Rhats = 1.00 pour tous les paramètres).

En plus de l'effet simple de TYPE, ce modèle confirme une influence (certes modérée) du groupe-CLASSE : en comparant leur évolution avec celle des phrases antéposées avec inversion, les phrases in situ sont de moins en moins associées au stéréotype *riche* quand les enfants grandissent ($\hat{\beta} = -0.15$, 95%CrI=[-0.41,0.11], $P(\beta) < 0 = 0.87$). En ce qui concerne les antépositions simples, ce n'est qu'une tendance mais l'évolution semble plutôt inverse, avec des associations comparativement plus fréquentes entre ces phrases et les images de personnes riches chez les enfants des classes les plus avancées ($\hat{\beta} = 0.08$, 95%CrI=[-0.17,0.34], $P(\beta) > 0 = 0.73$).

Cette perspective nuance l'effet principal présenté plus tôt où, à l'inverse des adultes, les enfants semblaient associer plus souvent les formes in situ à la richesse que les antépositions avec inversion. Il semble qu'il y ait une neutralisation progressive des stéréotypes initialement associés aux différentes formes interrogatives. Le groupe des CM1CM2 établit en effet, finalement, peu de différence entre les trois types phrastiques dans ses choix (48.3%, 47.6% et 48.5% d'associations avec des personnes riches).

Les enfants participant à la tâche de production (Section 6.3) semblaient concevoir la richesse différemment des adultes. Ces enfants n'appartenaient cependant qu'au groupe-classe des CE1-CM1 (deux élèves de CM1 seulement y ont participé, qui appartenaient à la classe de CE2-CM1, pour un âge maximal de 9 ans, contre 11 pour la présente tâche de *matched-Guise*). Il est concevable que les enfants des deux classes de CM1CM2 ont une vision plus nuancée de la richesse, peut-être moins influencée par les médias, et en tout cas plus proche, en ce sens, de celle des adultes.

6.2.4.3.2 Stéréotype des LOISIRS

La Figure 6.8 permet de visualiser l'évolution des associations entre les variantes interrogatives et les images de personnes pauvres ou personnes riches (N=1746), en fonction du groupe-classe des enfants. Contrairement au stéréotype de la richesse, il semble y avoir ici moins d'instabilité entre les groupes les plus jeunes, notamment en ce qui concerne les associations entre antépositions simples et activités sportives. En termes d'évolution globale, deux patrons ressortent.

D'abord, il semble y avoir une augmentation générale des choix *lecture* à mesure que les enfants grandissent, quel que soit le type phrastique (c'est même presque linéaire pour les phrases in situ). Ensuite, il faut observer que les antépositions avec inversion sont, pour les groupes CE1CM1 et CM1CM2, les phrases les plus souvent associées avec un choix *lecture* ((dans 51.9% et 60.6% des cas), notamment en contraste avec les phrases in situ et donc un peu à la manière des adultes.

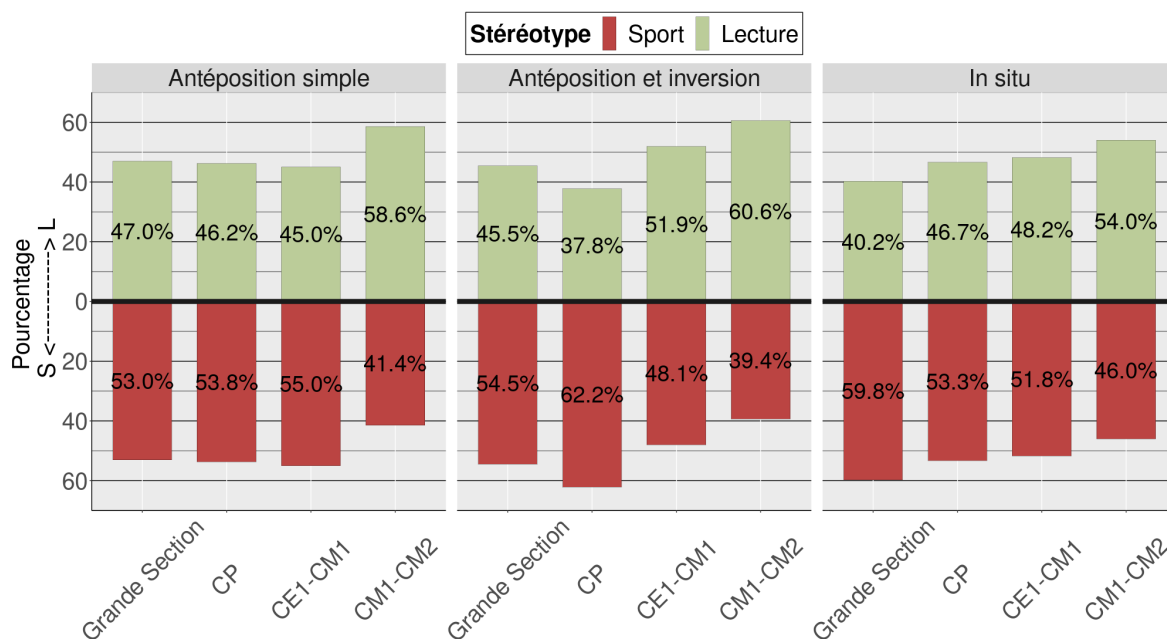


FIGURE 6.8 – Évaluation des LOISIRS, par TYPE et CLASSE (MG-C)

Un modèle de régression bayésienne calculé sur ce sous-ensemble des données (Annexe M.0.11) avec les mêmes variables, effets aléatoires, et critères de convergence que les précédents (mais 6000 itérations par chaîne pour atteindre la convergence cette fois), confirme cette dernière tendance.

Aucune interaction TYPE*GROUPE-CLASSE ne semble corroborée en comparant l'évolution des antépositions simples à celles des antépositions avec inversion ($\hat{\beta} = -0.05$, 95%CrI=[-0.32,0.21], $P(\beta) < 0=0.64$). Pour les phrases in situ, les données semblent bien indicatives d'une tendance selon laquelle ces phrases seraient, à mesure que les enfants grandissent, moins associées au stéréotype de la *lecture*, comparées aux phrases avec antéposition et inversion ($\hat{\beta} = -0.10$, 95%CrI=[-0.36,0.16], $P(\beta) < 0=0.77$).

Plusieurs idées sont liées à ces observations. D'abord, il est normal que les enfants augmentent leurs choix *lecture* de manière générale, indépendamment de tout autre facteur, à mesure que l'école les familiarise avec cette activité. Ensuite, l'augmentation de la prévalence des choix *lecture* pour l'antéposition avec inversion comparée au type in situ semble confirmer une évolution vers une association forme-stéréotype semblable à celle existant chez les adultes.

Le dernier point, un peu surprenant, est cette association assez forte entre antépositions simples et activité de lecture, qui semble apparaître dans le groupe des CM1CM2. Sans certitude, une hypothèse pourrait être que cette variante, relativement stigmatisée par la norme chez l'adulte, n'est pas une variante beaucoup employée au quotidien dans l'environnement local où ont eu lieu les passations. Par contraste, il se pourrait qu'elle acquière chez les enfants une valeur un peu littéraire, au sens de « sortie des livres » ou « c'est dans les livres que les personnages l'utilisent ». Un travail d'approfondissement serait ici nécessaire.

6.2.4.3.3 Stéréotype de l'ÉDUCATION

Enfin, la Figure 6.9 permet de visualiser l'évolution des associations entre les variantes interrogatives et les images de caissier-caissière ou de médecin (N=1743), en fonction du groupe-classe des enfants. Ici encore, il semble y avoir moins d'instabilité entre les groupes les plus jeunes dans leurs associations, sauf à nouveau sur l'antéposition avec inversion pour le groupe CP.

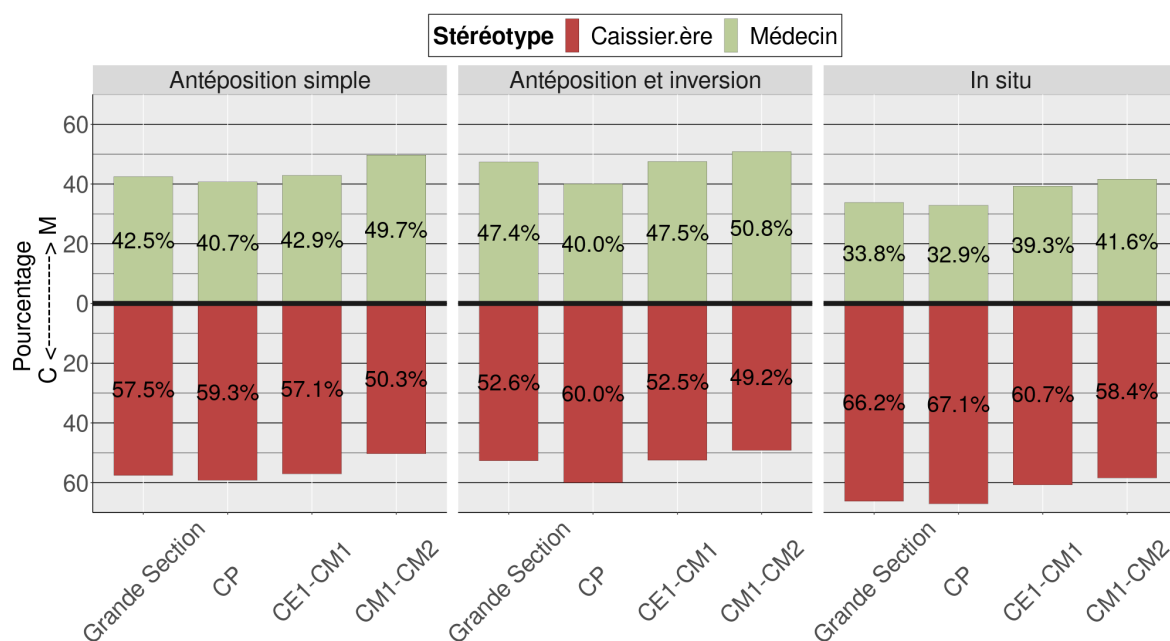


FIGURE 6.9 – Évaluation de l'ÉDUCATION, par TYPE et CLASSE (MG-C)

En termes d'évolution globale, deux résultats sont intéressants. Une observation générale est d'abord que les évolutions des associations forme-stéréotype sont plutôt stables à travers les groupes-classe, et les trois associations évoluent de manière très parallèle. Ensuite, et c'est lié, les phrases in situ restent de manière générale moins associées aux images de personnes médecin, quel que soit le groupe-classe, que les deux autres types phrastiques.

Un modèle de régression bayésienne, calculé avec les mêmes paramètres que précédemment sur ce sous-ensemble des données (avec 3000 itérations par chaîne suffisantes pour atteindre la convergence), confirme ces faits (Annexe M.0.12). Tout en confirmant de manière relativement fiable la différence générale entre les antépositions avec inversion et les phrases in situ ($\hat{\beta} = -0.49$, $95\%CrI = [-1.32, 0.34]$, $P(\beta) < 0 = 0.88$), le modèle ne corrobore aucune variation générale sur l'ensemble des groupes-classe. C'est autant le cas lorsque les antépositions simples sont comparées aux antépositions avec inversion ($\hat{\beta} = 0.03$, $95\%CrI = [-0.23, 0.29]$, $P(\beta) > 0 = 0.58$) que lorsque les phrases in situ leur sont comparées ($\hat{\beta} = 0.03$, $95\%CrI = [-0.24, 0.30]$, $P(\beta) > 0 = 0.59$).

6.2.4.3.4 Un modèle global

Confronter les données des trois stéréotypes en en faisant une analyse unique permet de mieux visualiser leur évolution différente. La Figure M.13 condense les informations des paragraphes précédents. Ces graphiques présentent le pourcentage des choix effectués par les enfants pour les images de personnes riches, de lecture, et de médecins, en fonction du groupe-classe (en abscisse) et du type phrastique (en couleur).

Ces graphiques font ressortir les différences entre stéréotype :

- pour la RICHESSE, une neutralisation des associations forme-stéréotype semble s'opérer avec le temps,
- pour les LOISIRS, la lecture est de plus en plus associée aux trois formes interrogatives, mais avec une association plus fréquente pour les antépositions avec inversion,
- pour l'ÉDUCATION, non seulement l'évolution des trois associations forme-stéréotype semble

très parallèle, mais il ressort aussi que les phrases in situ demeure très peu associées, comparativement, à une personne médecin, et ce dès le plus jeune âge (quel que soit le stéréotype qui est ici vraiment atteint : soit l'éducation, soit la position sociale, soit encore seulement la situation interactionnelle particulière consistant à « aller chez le médecin »).

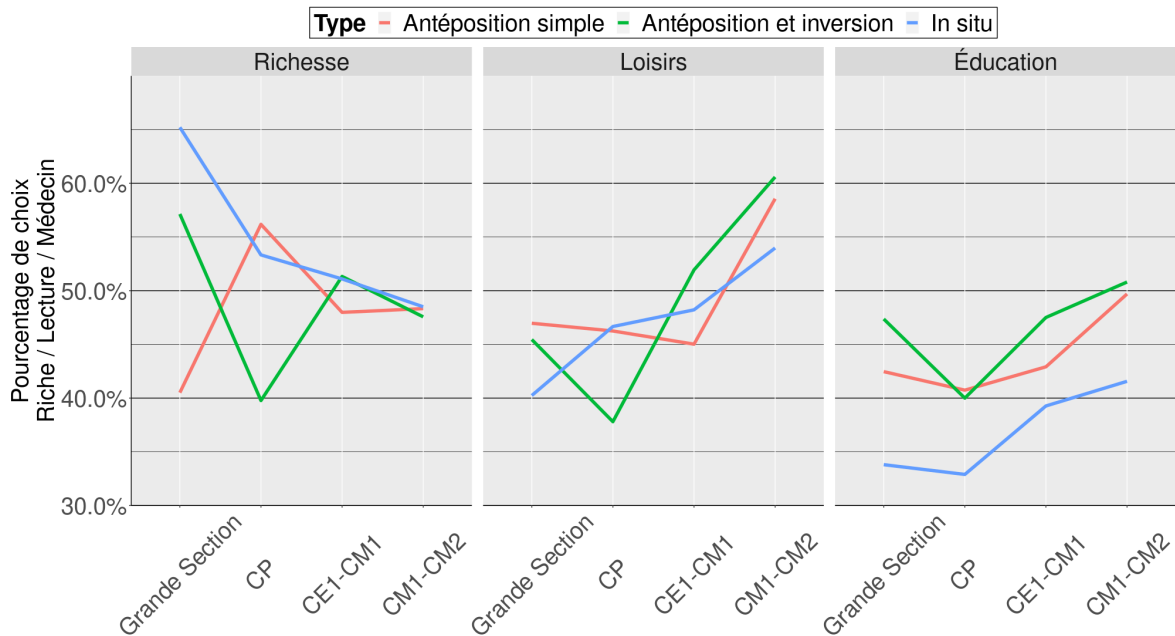


FIGURE 6.10 – Évolution des associations forme-stéréotype (MG-C)

Un modèle de régression bayésienne a été calculé sur l'ensemble des 5275 réponses en y incluant tous ces paramètres (Annexe M.0.13). La variable dépendante était toujours le choix de l'image (codé « 0 » pour les choix *pauvre, sport, caissier-caissière* et 1 pour les autres choix). La variable indépendante était l'interaction entre le TYPE phrastique, le GROUPE-CLASSE et le STÉRÉOTYPE (avec l'éducation pour niveau de référence, car c'est sur celui-ci qu'il semble y avoir le plus de stabilité au cours du temps). Des effets aléatoires "ID" (avec pour pente l'interaction TYPE*STÉRÉOTYPE) et "Item" (avec pour pente l'interaction TYPE*GROUPE-CLASSE*STÉRÉOTYPE) ont été ajoutés. Le modèle a tourné avec la famille "bernoulli" à nouveau, avec 4 chaînes et 3000 itérations par chaîne pour atteindre la convergence (Rhats = 1.00 pour tous les paramètres).

Si la complexité du modèle fait diminuer la valeur d'un certain nombre de paramètres, deux interactions triple TYPE*GROUPE-CLASSE*STÉRÉOTYPE restent notables, qui concernent la comparaison des phrases in situ avec les phrases avec antéposition et inversion. Lorsque ces deux types phrastiques sont comparés à l'échelle de toutes les données, et en comparant le stéréotype *richesse* à celui de l'*éducation*, il y a ainsi une influence modérée du groupe-classe sur le choix fait par les enfants ($\hat{\beta} = -0.16$, 95%CrI=[-0.54,0.21], $P(\beta) < 0.80$). Cette interaction confirme la neutralisation au fil du temps chez les enfants de l'association entre forme in situ et *richesse* observée précédemment, alors même que la distinction entre forme in situ et antéposition avec inversion reste constante.

En parallèle, en comparant ces deux types phrastiques et en contrastant le stéréotype *loisirs* à celui de l'*éducation*, une même tendance se dessine, bien que plus légère ($\hat{\beta} = -0.12$, 95%CrI=[-0.49,0.26], $P(\beta) < 0.73$). Cette interaction tend à confirmer que, alors même que les enfants semblent associer plus souvent les trois formes interrogatives avec la *lecture*, c'est plus fortement le cas pour l'antéposition avec inversion que pour la forme in situ.

6.2.4.4 Bilan : Un profil stéréotypique évolutif des interrogatives partielles

Une autre façon de regarder les données de la Figure 6.10 consiste à les condenser non pas par stéréotype mais par groupe-classe. C'est ce que fait la Figure 6.11, où sont visibles les pourcentages des choix effectués par les enfants pour les images de personnes *riches*, de *lecture*, et de *médecins*, en fonction cette fois du stéréotype (en abscisse). Cette visualisation peut directement être mise en parallèle du profil stéréotypique des interrogatives établi chez l'adulte (MG2, Section 5.3.2, Figure 5.38, p. 140). Par exemple, pour les enfants de Grande Section, une personne utilisant les phrases *in situ* est associée à plus de richesse que si elle utilisait une forme antéposée, mais elle est au contraire moins associée à la lecture et à une personne médecin.

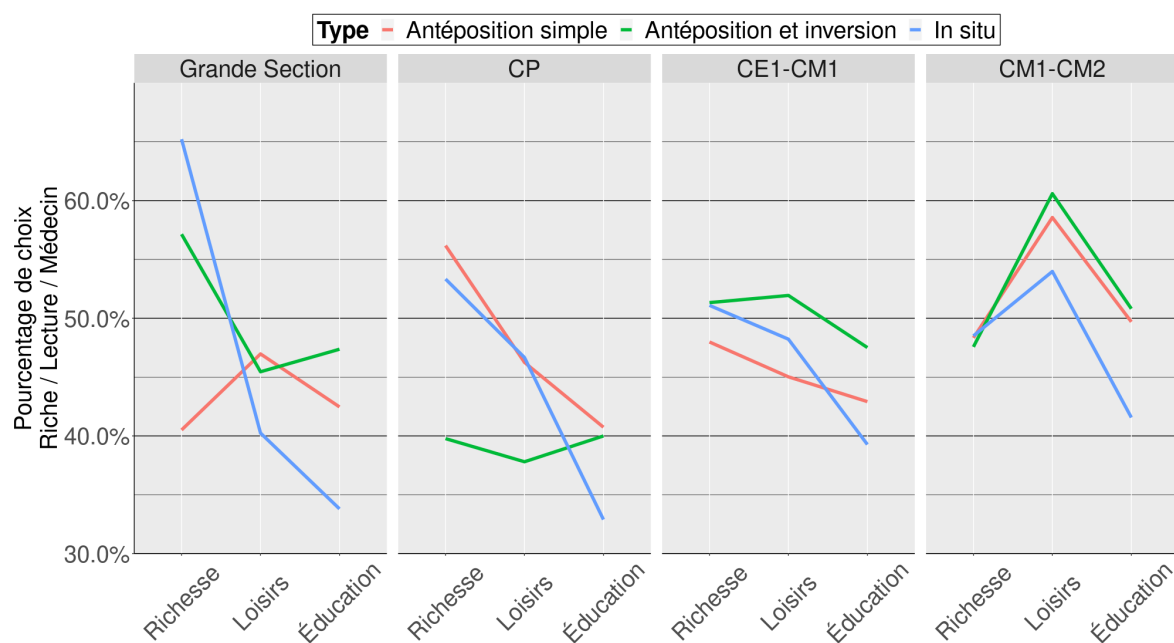


FIGURE 6.11 – Évolution du profil stéréotypique des interrogatives partielles (MG-C)

La comparaison de ces quatre graphiques permet de mieux caractériser le développement des « profils sociaux type » associés à chacune des variantes interrogatives. Ainsi

- la variante *in situ*, d'abord plus associée à la richesse que les autres types, perd peu à peu de sa spécificité. Cependant, elle ne devient jamais plus fréquemment associée à la lecture que l'antéposition avec inversion, et reste toujours la variante la moins associée à une personne médecin.
- la variante *antéposée avec inversion*, n'est de manière générale pas associée à plus de richesse que les deux autres types, mais devient la plus fréquemment associée à l'activité de lecture et à des personnes médecin.
- la variante *antéposée sans inversion* est la plus volatile des trois variantes testées, mais semble se positionner peu à peu comme voisine de l'antéposition avec inversion, notamment dans le même contraste avec le type *in situ* pour la lecture et l'association avec les médecins.

Ces données confirment l'hypothèse d'une évolution des associations forme-stéréotype chez l'enfant pour ce qui est des variantes d'interrogation partielle, au cours notamment de leur scolarité. À part pour la richesse, cette évolution conduit à une distinction semblable à celle observée chez les adultes entre les formes *in situ* et les formes avec antéposition et inversion. La proximité finale,

pour le groupe le plus avancé, entre antéposition simple et antéposition avec inversion, diffère cependant de ce qui est observé chez l'adulte. Une explication à ceci serait peut-être à aller chercher dans l'input linguistique des enfants, et à la fréquence à laquelle l'antéposition simple y est utilisée.

Un autre point important que ces données laissent également entrevoir est l'existence, même chez les plus jeunes enfants, d'une conscience d'un lien existant entre situation interactionnelle et production linguistique. Pour cette tâche de *matched-guise*, cette situation interactionnelle est définie par un « portrait social » minimal de l'interlocuteur ou de l'interlocutrice.

Les hypothèses de départ étaient que les enfants n'ont « par définition par intégré toutes les contraintes sociolinguistiques liées à l'utilisation des interrogatives partielles en français » mais que « les associations entre variante interrogative et stéréotype social devraient évoluer au cours du développement linguistique de l'enfant et (...) se rapprocher des résultats observés chez l'adulte ». Les données récoltées montrent que, certes, les associations entre stéréotype social et forme interrogative ne sont pas toujours semblables à celles de l'adulte, pour le groupe d'enfants et pour certains stéréotypes testés. Il y a cependant bien une évolution de ces associations au cours du temps. Par ailleurs, il semble également que, même aux premiers temps de leur scolarisation, les enfants ont intégré le principe d'une utilisation circonstanciée de la variation linguistique.

6.3 Adaptation sociolinguistique en production : Activité de jeux de rôles (JDR-C)

La tâche de *matched-guise* adaptée aux enfants décrite ci-dessus explore l'intégration et le développement de stéréotypes sociolinguistiques chez l'enfant. Des résultats se dégagent une conscience précoce des enfants quant au lien entre situation interactionnelle et production linguistique. Cette situation interactionnelle peut dépendre du statut des individus (« les gens parlent différemment selon leurs caractéristiques sociales propres ») mais aussi plus largement du contexte extra-linguistique pris au sens large (par exemple, « on parle différemment dans un contexte a priori informel et dans un contexte a priori formel »). Bien sûr, la formalité du contexte est en partie définie par les personnes y intervenant. C'est l'intégration par l'enfant de ces paramètres de formalité qui a été explorée dans la seconde tâche réalisée à l'école d'Argouges, en deuxième période de la phase de recueil de données, après les vacances d'octobre 2019.

6.3.1 Hypothèses et prédictions

Les hypothèses principales sous-tendant cet axe de recherche sont parallèles à celles sous-tendant la première tâche sur les stéréotypes :

- les enfants n'ont par définition pas intégré toutes les contraintes sociolinguistiques liées à l'utilisation des interrogatives partielles en français. Les associations entre variante interrogative et *formalité du contexte* ne devraient donc pas être les mêmes que chez l'adulte, voire être inexistantes ou illisibles sous un référentiel adulte.
- la compétence sociopragmatique se développe depuis le plus jeune âge jusqu'à l'adolescence (et même après), par un processus d'assimilation des règles sociales liées aux situations d'interaction dans lesquelles baigne l'individu linguistique (voir Section 2.4). En conséquence, les associations entre variante interrogative et *formalité du contexte* devraient évoluer au

cours du développement linguistique de l'enfant sur les âges considérés (ici, de 4 à 9 ans). Les résultats chez les enfants les plus âgés devraient évoquer une progression vers les utilisations adultes des variantes interrogatives (notamment : antépositions avec inversion, s'il y en a, dans des contextes plus formels).

6.3.2 Protocole - Constitution du corpus

L'enjeu était de déceler et d'observer les capacités adaptatives des enfants sur le plan socio-linguistique d'un point de vue des interrogatives partielles, et notamment en termes de formalité du contexte. Il fallait également conserver un degré de spontanéité dans les productions, tout en restant dans une perspective systématique et contrôlée à travers les tranches d'âges. L'option choisie a donc été une mise en situation de jeu de rôles, dans laquelle les enfants devaient gérer un restaurant face à des clients et clientes venant manger dans celui-ci. Un enregistrement des enfants était cette fois effectué, audio uniquement, et uniquement pendant la phase de jeu à proprement parler, avec un accord explicite recueilli auprès des parents par l'équipe pédagogique.

Le tout se déroulait dans une salle de classe disponible, proche de celles des enfants, autour d'une table dédiée représentant le périmètre du restaurant et avec de la dinette comprenant à la fois un service à couverts et une grande quantité d'aliments solides ou liquides. Tout l'environnement ainsi que le marionnettiste était présenté à chaque enfant avant le début des enregistrements par l'expérimentateur, déjà connu des enfants. Au moment où le jeu commençait, à l'initiative du marionnettiste la plupart du temps, l'expérimentateur était assis hors du champ de vision de l'enfant mais sans être caché (dans un coin un peu isolé de la pièce).

Pour distancier les enfants d'une relation à l'adulte risquant de biaiser leur comportement, des personnages marionnettes ont été utilisés, manipulés par un adulte non-familier avec la problématique de recherche, mais formé à l'interaction avec les enfants (BAFA), à l'improvisation même en suivant un script prédéfini, à l'incarnation de personnalités variées (formation de théâtre), et à la manipulation de marionnettes. Les personnages disponibles étaient au nombre de quatre, genrés pour permettre une identification plus facile et rapide avec les enfants.

Il y avait d'un côté un personnage de roi et un personnage de reine (Richard et Éléonore) associés à des contextes plus formels, et de l'autre un personnage de petit garçon et un personnage de petite fille (Jeannot et Marie) associés à des contextes plus familiers et proches des enfants. La Figure 6.12 présente ces quatre personnages.



FIGURE 6.12 – Marionnettes utilisées (JDR-C)

La Figure 6.13 donne une idée de l'environnement dans lequel les données ont été produites et récoltées.



FIGURE 6.13 – Environnement de jeu de rôle, marionnettes et matériel (JDR-C)

En concertation avec le marionnettiste, les personnages du roi et de la reine s'exprimaient autant que possible dans un registre soutenu, en vouvoyant l'enfant et en insérant des liaisons facultatives et des négations complètes (« ne... pas »), avec une prosodie bien détachée. Les personnages de Marie et de Jeannot pour leur part avaient pour traits typiques un grand relâchement prosodique, une certaine familiarité de comportement, des négations incomplètes, un tutoiement total de l'enfant, des centres d'intérêt plus « jeunes » et un vocabulaire plus familier.

Le script squelettique de chaque interaction était toujours le même; le personnage venait manger dans le restaurant de l'enfant, avec pour points chronologiquement ancrés :

1. l'attente d'une question sur ce que le personnage voulait à manger;
2. la présentation d'une journée « terrible », avec attente d'une question sur ce qu'il s'était passé;
3. divers points de discussion incitant l'enfant à la curiosité envers le personnage en face de lui (volonté d'explorer la région voisine de l'école, soins à un cheval pour les personnages du roi et de la reine, goût musicaux, passions diverses, sport, etc.)

Les points 1. et 2. ont plutôt bien fonctionné, le point 3. moins. Si la consigne du marionnettiste était au départ de ne pas produire lui-même d'interrogation pour influencer le moins possible sur les productions éventuelles des enfants, il est vite devenu difficile de procéder de la sorte, tant l'acte interrogatif est crucial dans la création de lien entre deux individus pour maintenir une conversation.

Les jeux commençaient toujours par un personnage d'enfant (Jeannot ou Marie), puis par un personnage entre le roi et la reine. Pour les enfants au-delà du CP, une troisième et parfois une quatrième marionnette intervenaient souvent ensuite, les interactions étant alors à la fois plus rapides, et plus productives avec les enfants.

La durée des interactions n'était pas limitée pour laisser le temps à l'enfant de s'approprier l'environnement du jeu et les personnages, et d'accepter ou non l'interaction avec l'adulte marionnettiste (l'accord était explicitement demandé pour chaque enfant). Dans un cas (CHI07), cette acceptation n'a pas eu lieu et l'enfant a bien sûr pu se retirer sans conséquence.

Après la phase de jeu, l'enregistrement était stoppé et il était demandé à l'enfant s'il était possible de lui poser quelques questions sur :

- ce qu'était quelqu'un de riche, avec des exemples si possible ;
- à quelle fréquence il lui arrive de lire, quoi et avec qui ;
- ses projets d'avenir en termes de métier et sur les moyens d'y parvenir si connus.

Ces questions visaient à approfondir la relation des enfants avec les stéréotypes manipulés dans la tâche de *matched-guise* enfantin décrite en Section 6.2.

Après tout ceci, il était remis à l'enfant un petit diplôme, sur le modèle du précédent (voir Annexe O.0.1). Le diplôme était donné indépendamment de tout autre paramètre, notamment de production, et y compris aux enfants qui n'ont rien dit pendant le petit jeu.

6.3.3 Données générales sur le corpus

68 sessions ont débouché sur un enregistrement, par la suite transcrit (section 6.3.4) puis annoté syntaxiquement. Ces enregistrements ont été obtenus auprès de 37 garçons et 31 filles, et correspondent au total à 15 heures et 56 minutes d'interaction entre les enfants et le marionnettiste (minimum = 05mn29s; maximum = 33mn05s; moyenne = 14mn04s). Les enfants avaient entre 4 et 9 ans (âge précis non conservé pour les classes de GS, moyenne pour les CP et au-delà = 7.08 ans), et appartenaient aux classes de Grande Section jusque celle de CE2-CM1. Le détail de la répartition des enfants par classe, par niveau scolaire et par âge est donnée dans les Tableaux 6.5-6.7.

GS-1	GS-2	CP	CP-CE1	CE1-CE2	CE2-CM1
7	11	14	12	12	12

TABLEAU 6.5 – Répartition des enfants par classes (JDR-C)

GS	CP	CE1	CE2	CM1
18	17	15	16	2

TABLEAU 6.6 – Répartition des enfants par niveaux (JDR-C)

4-6 ans (GS)	6 ans	7 ans	8 ans	9 ans
18	17	14	17	2

TABLEAU 6.7 – Répartition des enfants par âges (JDR-C)

Les enfants des niveaux CM1 et CM2 n'ont pas participé (sauf 2 de la classe double CE2-CM1 où beaucoup de CE2 ont participé). La principale raison à cela est que le jeu avec des marionnettes est moins engageant pour les enfants de ces âges avancés. La seconde raison, plus matérielle, est une raison de temps. Il n'était pas attendu que les enfants s'engagent autant dans l'activité et soient aussi prolifiques. Ce protocole s'est révélé beaucoup plus chronophage que prévu, à la fois pour récolter les données mais également pour les traiter par la suite.

6.3.4 Productions recueillies et transcriptions

Au total, 66517 mots constituent le corpus obtenu, pour 46446 mots produits par le marionnettiste et 20071 par les enfants. Les enfants ont été plus ou moins volubiles selon les enregistrements (par enregistrement, nombres de mots minimum = 0, maximum = 1516, moyenne = 295.16), de

même que l'adulte en face, qui s'adaptait (par enregistrement, nombres de mots minimum = 263, maximum = 1387, moyenne = 683.03).

L'objectif de ce corpus était la récolte d'interrogatives partielles produites spontanément, dans des situations d'interaction relativement naturelles pour les enfants (un jeu). Il n'était pas attendu d'obtenir une production linguistique aussi importante. La transcription s'est faite avec le souci principal de pouvoir retrouver rapidement les interrogatives produites. Il s'agissait également de pouvoir éventuellement retravailler sur d'autres éléments par la suite (par exemple étudier les phases de construction de l'interaction). Ce dernier point nécessitera une standardisation des annotations selon les protocoles habituels dans le domaine (alignement temporel des fichiers audio, utilisation d'un logiciel comme CLAN [MACWHINNEY, 2000, 2009], par exemple).

Les enregistrements ont été transcrits en distinguant chaque phase interactionnelle, à l'appréciation du transcripteur sur une base notamment prosodique. Plusieurs phases successives peuvent appartenir à la même personne, soit l'enfant soit le marionnettiste. Ces phases recouvrent aussi des phases de rire, ou des moments où l'enfant par exemple se déplace, ou encore fouille dans les conteneurs à dinette pour y trouver des aliments et des objets à utiliser. Un exemple de transcription est montré en Figure 6.14, correspondant au début de la séance avec l'enfant CHI33.

Personne	Phase	N° de Phase	Question
DEBUT		0	
MARIONNETTE	MARIE	1	
ADU	salut	2	
ADU	moi c'est marie	3	
CHI	(bouge)	4	
ADU	j'ai dix ans et euh	5	
ADU	j'ai super faim alors jviens manger dans ton restaurant	6	
CHI	qu'est-ce que tu veux	7	1
ADU	euh jveux bien du poulet	8	
CHI	(bouge)	9	
ADU	steuplait	10	
ADU	hmm merci	11	
CHI	(bouge)	12	
ADU	merci beaucoup	13	
ADU	hmm ça a l'air vachement bon	14	
ADU	(mange)	15	
ADU	hm c'est super bon	16	
CHI	(bouge)	17	
ADU	hmm mais c'est ça donne un ptit peu soif hein	18	
CHI	(bouge)	19	
ADU	ah merci	20	
ADU	(boit)	21	
ADU	oh ça fait du bien	22	
ADU	surtout après la matinée qu'j'ai passée	23	
ADU	c'était terrible	24	
ADU	y a des gens qui m'ont poursuivie dans la rue	25	
ADU	et euh après jsuis arrivée en rtard à l'école	26	
ADU	et la maitresse elle m'a grondée	27	
ADU	hn	28	
ADU	c'est injuste hein	29	1
CHI	oui	30	
CHI	mais xx	31	
CHI	mais pourquoi ils t'ont poursuie	32	1

FIGURE 6.14 – Exemple de transcription, JDR-C - CHI33, phases 0-32

La colonne « Question » indique la présence d'une phase interrogative, de quelque type que ce soit : interrogation partielle, interrogation totale, y compris les phrases averbales ou même

interjections du type « hmm? ». C'est sur la base de ce relevé initial que les interrogatives partielles ont été marquées comme telles puis annotées syntaxiquement pour une analyse plus poussée (voir Section 6.3.5).

Par la suite, les extractions de ce corpus seront citées sous la forme

(119) Mais **pourquoi** ils t'ont poursuie

(CHI33, Phase 32 - #5675)

Après la citation en elle-même se trouvent donc, entre parenthèses, le code attribué à l'enfant ou à l'adulte pour l'annotation, et le numéro de phase, valable au sein de chaque transcription. Le dernier nombre est la référence unique de cette ligne dans la compilation de tous les fichiers de transcription, et permet par exemple de retrouver l'extraction dans l'Annexe O.0.3, mais aussi dans le fichier disponible en ligne au répertoire OSF (osf.io/ud4ks).

Une attention particulière a été portée à transcrire les élisions de syllabes, mots (120) ou par exemple du schwa (121), et autres signes paraverbaux de changements de registre. La finalité en est une éventuelle exploitation future, qui analyserait plus largement la construction d'une interaction chez l'enfant de 4 à 9 ans.

(120) steuplait

(ADU, Phase 10 - #5653)

(121) surtout après la matinée qu'j'ai passée

(ADU, Phase 23 - #5666)

Dans le même ordre d'idée, les phases 4, 9, 12, 19 et 21 donnent un aperçu des phases interactionnelles non-linguistiques mentionnées plus haut. Elles correspondent principalement, à travers les transcriptions, aux moments où l'enfant bouge, fouille dans la pièce pour soutenir son activité de restauration, ou encore sert à manger aux marionnettes. Pour l'adulte il s'agit notamment des moments où la marionnette boit et mange ce qui lui est servi. Les phases de rire, notamment de l'enfant, sont aussi relevées de cette manière. Ces phases, entre parenthèses, n'ont pas été comptabilisées dans le comptage des mots, ni pour les enfants ni pour le marionnettiste.

La longueur et la densité linguistique très variables des interactions se traduisent aussi sur ce « nombre total de phases » par enregistrement : de 71 à 700 avec une moyenne de 274. La moyenne du nombre de mots par « phase », calculée pour chaque enregistrement, illustre aussi la densité très différente en production, d'un enfant à l'autre (minimum = 0, maximum = 4.28, moyenne = 0.86). Ces dernières données sont à comparer avec les mêmes chiffres pour l'adulte (minimum = 1.37, maximum = 5.75, moyenne = 2.85). Le détail de ces métadonnées est compilé en Annexe O.0.2.

6.3.5 Annotations syntaxiques et résultats

Les phrases annotées comme étant des questions durant la phase de transcription ont ensuite été annotées syntaxiquement, pour isoler les interrogatives partielles. 361 éléments interrogatifs partiels ont été repérés sur l'ensemble des productions enfantines, dont 70 dans des productions averbales ou incomplètes. Aucune interrogative sujet n'a été produite par les enfants, mais 3 interrogatives supplémentaires ont été exclues des analyses présentées ci-après, parce qu'elles ont été produites hors du champ de l'activité (avant l'apparition des marionnettes). 1 phrase supplémentaire a été exclue car son type mêlait in situ et antéposition en *est-ce que* (122).

(122) euh qu'est ce que y a quoi d'autre

(CHI58, Phase 306 - #10403)

17 interrogatives supplémentaires ont été exclues car elles présentaient une structure avec enchâssement, hors du champ de ce travail. Enfin, 1 interrogative de type idiomatique a été supprimée (« Comment dire? », phrase non-finie qui plus est). Au total, 269 phrases interrogatives partielles finies, non-sujet, et produites par les enfants, ont été extraites du corpus pour analyse. Elles sont reproduites en Annexe O.0.3. Ces 269 interrogatives ont été produites par 42 enfants sur les 68 ayant participé à la tâche. L'intégralité du reste du corpus et des fichiers d'annotation est ou sera mise en ligne et accessible par le lien osf.io/ud4ks.

Chacune de ces 269 interrogatives a été annotée selon son type syntaxique (antéposition simple, antéposition en *est-ce que*, antéposition avec inversion, ou in situ). Le type de mot interrogatif a également été noté. Le nom de la marionnette avec laquelle interagissait l'enfant au moment de la production de l'interrogative a été conservé (Marie, Jeannot, la reine Éléonore ou le roi Richard), ainsi que le degré de formalité du contexte associé : « 0 », informel, pour Marie et Jeannot ; « 1 » pour le roi et la reine. La présence d'une marque de tutoiement ou de vouvoiement dans l'interrogative a aussi été annotée (« tu », « vous », ou une autre marque comme un déterminant possessif ou un pronom).

La répartition générale des 269 phrases relevées selon leur type est résumée dans le Tableau 5.3. S'y observe déjà une forte présence des phrases in situ, mais les interrogatives en *est-ce que* sont également assez nombreuses. Les deux autres types d'antéposition, simple et avec inversion, ont été pour leur part peu utilisés par les enfants.

Antéposition simple	Antéposition + ESK	Antéposition et inversion	In situ
10	52	9	198
3.7 %	19.3%	3.3 %	73.6 %

TABLEAU 6.8 – Interrogatives partielles par type (JDR-C)

Ces 269 interrogatives constituent une partie certes très réduite du corpus global recueilli auprès des enfants. La visée première de cette tâche de jeu de rôle était de pouvoir mener une étude systématique de la production spontanée d'interrogative partielle dans un contexte relativement cadré, et c'est donc principalement de manière systématique que seront analysées les productions des enfants. Le but est ici de dépasser les spécificités interindividuelles pour essayer d'atteindre des tendances générales, mais il est bien sûr illusoire de penser que le groupe des enfants avait un comportement linguistique homogène, même au sein d'une même classe. Au-delà des effets généraux de formalité, d'âge ou de genre qui ressortent des données, la Section 6.3.5.4 ci-dessous viendra d'ailleurs rappeler qu'une étude plus qualitative des productions enfantines est très certainement nécessaire en parallèle pour compléter cette perspective systématique.

6.3.5.1 Un effet de formalité

La principale variable à laquelle il était attendu que les enfants s'adaptent, ou non, était la formalité du contexte associée aux différentes marionnettes manipulées. Pour rappel, l'hypothèse était ici que les enfants n'ont par définition pas intégré toutes les contraintes sociolinguistiques d'utilisation des interrogatives partielles, et que les productions des plus jeunes notamment de-

vraient le refléter. En parallèle, il était prédit que le développement sociopragmatique des enfants permette une adaptation plus fine au contexte avec le temps et l'âge.

Le Tableau 6.9 et la Figure 6.15 donnent une visualisation des productions en fonction de la formalité du contexte, informel (avec Jeannot et Marie, les marionnettes enfants) ou formel (avec le roi ou la reine). De ces données brutes, il ressort une diminution des productions *in situ* dans les contextes formels et, en contraste, une augmentation des productions des trois différents types antéposés.

	Antéposition simple	Antéposition + ESK	Antéposition et inversion	In situ	Total
Contexte informel	3	26	3	121	153
Contexte formel	7	26	6	77	117
Total	10	52	9	198	269

TABLEAU 6.9 – Interrogatives partielles produites selon le contexte (JDR-C)

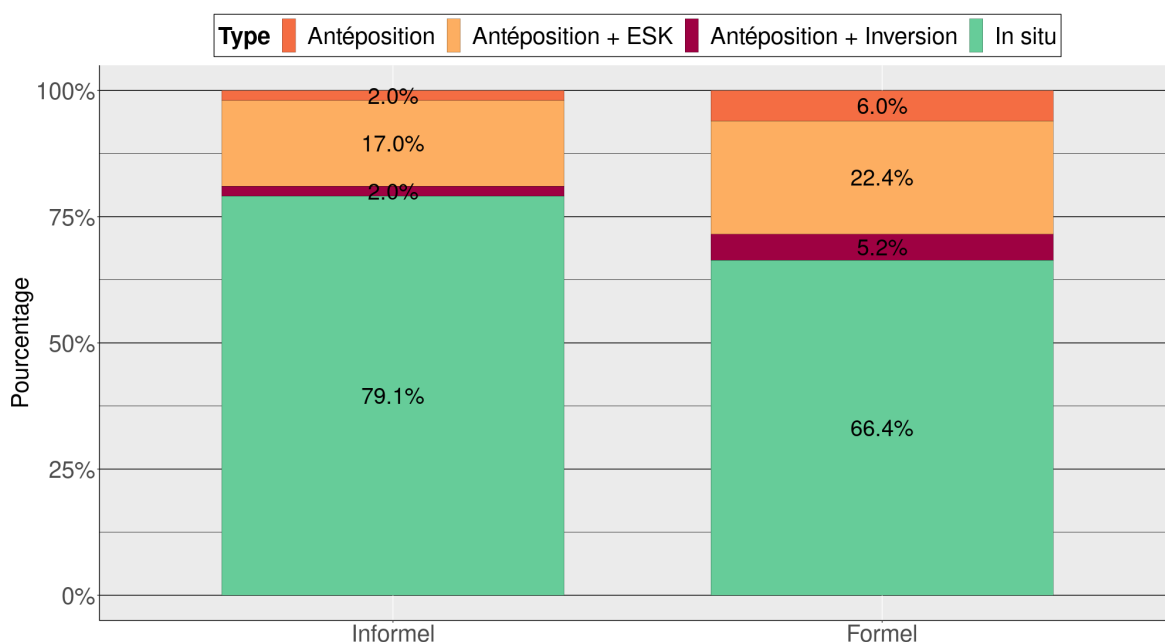


FIGURE 6.15 – Interrogatives partielles produites selon le contexte (JDR-C)

Un modèle de régression bayésienne a été calculé sur ces données avec pour variable dépendante le TYPE d'interrogative (variable catégorielle avec le type *in situ* pour niveau de référence), et pour variable indépendante le CONTEXTE (variable binaire, formel ou informel, avec *informel* pour niveau de référence). Un effet avec intercept aléatoire ("child") a été ajouté pour intégrer au mieux au modèle la variation interindividuelle entre les enfants. La convergence du modèle a été atteinte avec 4 chaînes et 6000 itérations par chaînes, et contrôlée selon les critères habituels (Rhats=1.00 pour tous les paramètres). Le modèle, dont le détail se trouve en Annexe P.0.1, confirme les observations descriptives.

En effet, quel que soit le type d'antéposition, en comparant l'évolution des productions de chaque type d'antéposition à l'évolution des productions *in situ* entre les contextes informels et les contextes formels, il y a une augmentation relative de leur proportion. Malgré les données peu nombreuses, cette augmentation est ainsi corroborée, avec modération, pour les antépositions

en *est-ce que* ($\hat{\beta} = 0.40$, 95%CrI=[-0.40,1.21], $P(\beta) > 0 = 0.84$). Cette augmentation est pleinement corroborée pour les antépositions avec inversion ($\hat{\beta} = 2.03$, 95%CrI=[0.11,4.28], $P(\beta) > 0 = 0.98$). Il ne s'agit cependant que d'une tendance pour les antépositions simples ($\hat{\beta} = 0.65$, 95%CrI=[-1.08,2.44], $P(\beta) > 0 = 0.77$).

Ces résultats semblent donc globalement indicatifs d'une capacité des enfants à modifier leur comportement linguistique en fonction d'une variation de contexte. On ne parle pas de la même façon à une reine ou à un roi qu'à un petit garçon ou une petite fille. Comme il a été mentionné cependant, et malgré la taille réduite du groupe des enfants pour cette tâche (N=68) par rapport à celui de la tâche précédente de *matched-guise* (N=136), ce groupe n'était pas homogène.

6.3.5.2 Un effet d'âge (classe) en interaction avec la formalité

Le premier filtre sous lequel considérer les productions enfantines peut être l'âge. Ceci permet d'apporter une réponse à la seconde hypothèse de cette étude, à savoir que, s'il y a chez les enfants une conscience sociolinguistique qui les pousse à changer leur comportement linguistique en fonction du contexte, cette conscience se développe avec le temps. Par le protocole expérimental établi et les contraintes sur le recrutement des enfants, deux manières d'approcher l'âge entrent néanmoins en compétition : l'âge réel et la classe d'appartenance. Les Tableaux 6.10 et 6.11 présentent la répartition des 269 interrogatives du corpus de jeu de rôle selon ces deux critères.

<6 ans	6 ans	7 ans	8 ans	9 ans
9	50	83	116	11

TABLEAU 6.10 – Répartition des interrogatives par âge (JDR-C)

Grande Section	CP	CP-CE1	CE1-CE2	CE2-CM1
9	45	65	68	82

TABLEAU 6.11 – Répartition des interrogatives par classe (JDR-C)

Hormis la correspondance entre Grande Section et le groupe d'enfants dont l'âge est inférieur à 6 ans, il est donc notable que les interrogatives produites sont mieux réparties en considérant les enfants par classe. Dans un souci de comparabilité statistique, ce sera donc ce filtre qui sera utilisé. Comme signalé en Section 6.2.4.3, ce filtre permet par ailleurs de tenir compte des spécificités de la situation scolaire de chaque niveau, avec par exemple un contraste entre les CP qui se trouvent dans une classe contenant aussi des CE1 et les autres. Le Tableau 6.12 et la Figure 6.16 donnent la distribution des formes interrogatives produites en fonction de la classe des enfants.

	Antéposition simple	Antéposition + ESK	Antéposition et inversion	In situ	Total
Grande Section	2	1		6	9
CP	5	8	4	28	45
CP-CE1	2	11	3	49	65
CE1-CE2		7		61	68
CE2-CM1	1	25	2	54	82
Total	10	52	9	198	269

TABLEAU 6.12 – Interrogatives partielles produites selon la classe (JDR-C)

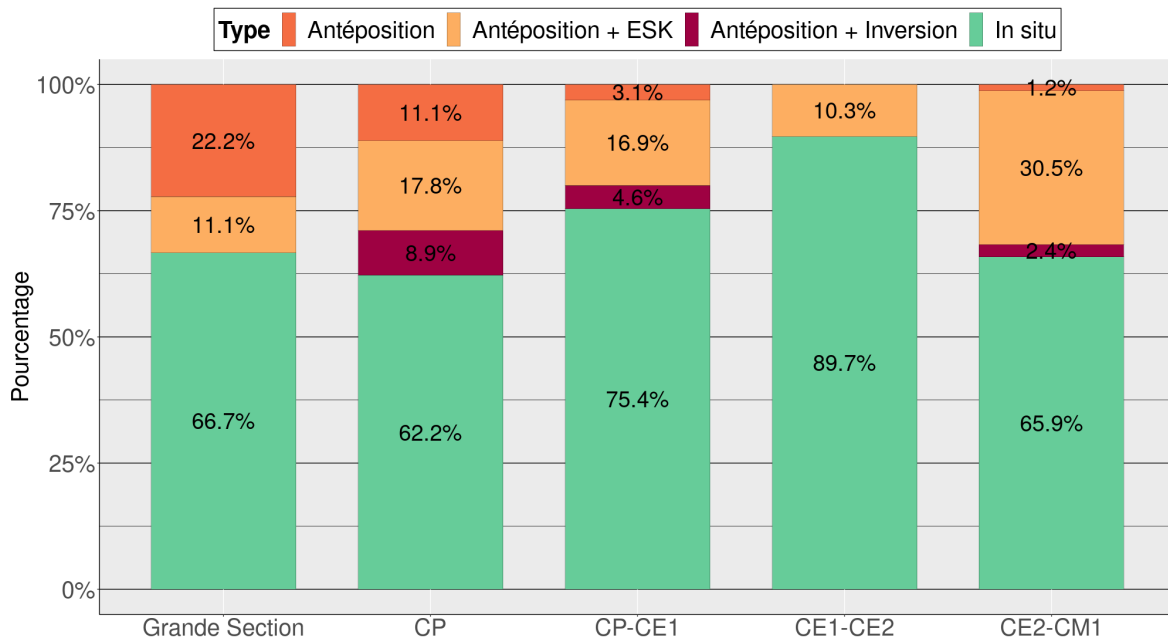


FIGURE 6.16 – Interrogatives partielles produites selon la classe (JDR-C)

Dans ces données il y a une fluctuation dans le recours aux phrases in situ, qui ne semble pas suivre de tendance générale particulière, depuis la Grande Section jusqu'à la classe de CE2-CM1, en passant par un pic en CE1-CE2. En parallèle, l'antéposition avec inversion est produite également aussi bien par des élèves de CE2-CM1 que, déjà, par des élèves de CP. L'antéposition en *est-ce que*, elle, semble plus utilisée par les enfants de CE2-CM1, à la fois en nombre absolu et en proportion, après un apparent délaissement progressif chez les enfants plus jeunes. Seule tendance générale durable, l'antéposition simple paraît cependant, de façon presque linéaire, de moins en moins utilisée à mesure que les enfants grandissent.

6.3.5.2.1 Un artefact du type d'élément interrogatif?

Avec la variabilité des situations de jeu de rôle, la production plus fréquente d'antépositions par les enfants en CP pourrait simplement résulter d'une diversité des types d'éléments interrogatifs utilisés. Pour reprendre les travaux de MYERS [2007] par exemple, si les éléments *pourquoi* et *comment* sont de manière générale plus souvent antéposés en français, et si les enfants les plus jeunes ont utilisé ces interrogatifs mais pas les enfants de CE1-CE2, ce phénomène pourrait au moins en partie s'expliquer.

Pour répondre à cette hypothèse, le Tableau 6.13 donne la liste complète des éléments interrogatifs utilisés et de leur nombre d'occurrences dans chacun des quatre types phrastiques.

	Antéposition simple	Antéposition + ESK	Antéposition et inversion	In situ	Total
à qui				1	1
à quoi				1	1
combien				3	3
comment	4		1		5
de quel SN				1	1
lequel				1	1
où	5			8	13
pourquoi	1				1
qu'/que/quoi		52	8	178	238
qui				5	5
Total	10	52	9	198	269

TABLEAU 6.13 – Répartition des éléments interrogatifs selon le type phrastique (JDR-C)

Les éléments interrogatifs qui ont été antéposés sont donc en effet, entre autres, *comment* et *pourquoi*. *Comment* a été produit 1 fois dès le CP (123), 2 fois en CP-CE1, et 2 fois en CE2-CM1 (dont, alors, une fois avec inversion (124)). L'unique occurrence de *pourquoi* a été produite en CP (125). Ces deux éléments n'ont jamais été produits in situ.

(123) et les tomates **comment** elles sont

(CHI33, Phase 593 - #6236)

(124) **comment** puis-je oublier ce coin de paradis

(CHI66, Phase 355 - #17865)

(125) mais **pourquoi** ils t'ont poursuivie

(CHI33, Phase 32 - #5675)

Un autre élément interrogatif, qui lui a été produit à la fois en position in situ et antéposé est l'élément *où*. Cet élément a été produit 2 fois dans une antéposition simple dès la Grande Section (126-127), 5 fois en CP (3 antépositions pour 2 in situ), 1 fois en CP-CE1 (in situ), 2 fois en CE1-CE2 (in situ) et 3 fois en CE2-CM1 (in situ). Ceci semblerait montrer, pour un élément comparable à travers les classes, que l'antéposition simple est de moins en moins privilégiée.

(126) **où** elle est ta licorne je la vois pas

(CHI12, Phase 188 - #1998)

(127) **où** il est ton rayo ton royaume

(CHI12, Phase 245 - #2055)

Enfin, le triptyque *qu'/quel/quoi*, dont la forme est conditionnée par sa position (respectivement antéposition avec *est-ce que*, antéposition avec inversion, et in situ), a été produit un total de 238 fois, selon une répartition résumée par le Tableau 6.14. Cette prépondérance de l'élément *qu'/quel/quoi* dans les interrogatives produites au cours de cette tâche favorise et facilite, au passage, son étude sous une perspective systématique.

	Antéposition + ESK	Antéposition + inversion	In situ	Total
Grande Section	1		6	7
CP	8	4	24	36
CP-CE1	11	3	46	62
CE1-CE2	7		57	64
CE2-CM1	25	1	45	69
Total	52	8	178	238

TABLEAU 6.14 – Productions de l'élément "qu'/que/quoi" selon la classe des enfants (JDR-C)

Cette répartition, à l'inverse de celle de *où*, est conforme au schéma visible sur la Figure 6.16, avec certes une diminution des antépositions en *est-ce que* en CE1-CE2, mais surtout une utilisation bien plus importante, en proportion autant qu'en chiffres bruts, par les élèves de CE2-CM1.

Ces différents éléments montrent que, quelle que soit leur classe, les enfants savent à la fois antéposer des éléments interrogatifs, mais aussi les placer in situ, y compris pour des éléments qui ne sont ni rares ni contraints dans leur placement syntaxique. De manière générale, il semble raisonnable de penser que, au-delà d'un possible artefact provoqué par les données récoltées et le caractère libre des interactions avec les enfants, ce n'est pas la seule explication possible.

6.3.5.2.2 Une interaction âge*formalité

Une autre explication peut se trouver en confrontant les données à la notion de formalité. L'hypothèse de départ n'était pas en effet qu'il y aurait une évolution des usages interrogatifs seulement en fonction de l'âge. L'hypothèse est qu'il devrait y avoir une évolution de la capacité à s'adapter sociolinguistiquement chez les enfants, et donc de leur capacité à mobiliser différemment des variantes linguistiques différentes dans des contextes différents. Le Tableau 6.15 et la Figure 6.17 donnent la distribution des interrogatives partielles en fonction de la classe des enfants mais aussi de la formalité du contexte.

En comparant les contextes formels aux contextes informels et en regardant comment évolue la différence entre les deux, il semble y avoir une tendance générale, à partir de la classe de CP-CE1, à recourir à moins de formes in situ et plus de formes d'un autre type dans les contextes formels. Ce, en parallèle d'un recours pourtant moins fréquent à l'antéposition simple ou à l'antéposition avec inversion avec le temps. C'est ainsi que, pour les élèves de CE2-CM1, il y a près de 50% de formes in situ en contexte formel, contre près de 75% en contexte informel. Pour ces enfants, c'est principalement la forme antéposée en *est-ce que* qui semble servir de relais pour adapter la production linguistique (la proportion en est quasiment doublée dans les contextes formels).

	Grande Section		CP		CP-CE1		CE1-CE2		CE2-CM1	
	Inf	For	Inf	For	Inf	For	Inf	For	Inf	For
Antéposition simple		2	3	2		2				1
Antéposition + ESK	1		5	3	7	4	3	4	10	15
Antéposition + inversion			1	3	1	2			1	1
In situ	5	1	12	16	32	17	37	24	35	19
TOTAL	6	3	21	24	40	25	40	28	46	36

 TABLEAU 6.15 – Interrogatives partielles par classe et contexte (*Informel/Formel*) (JDR-C)

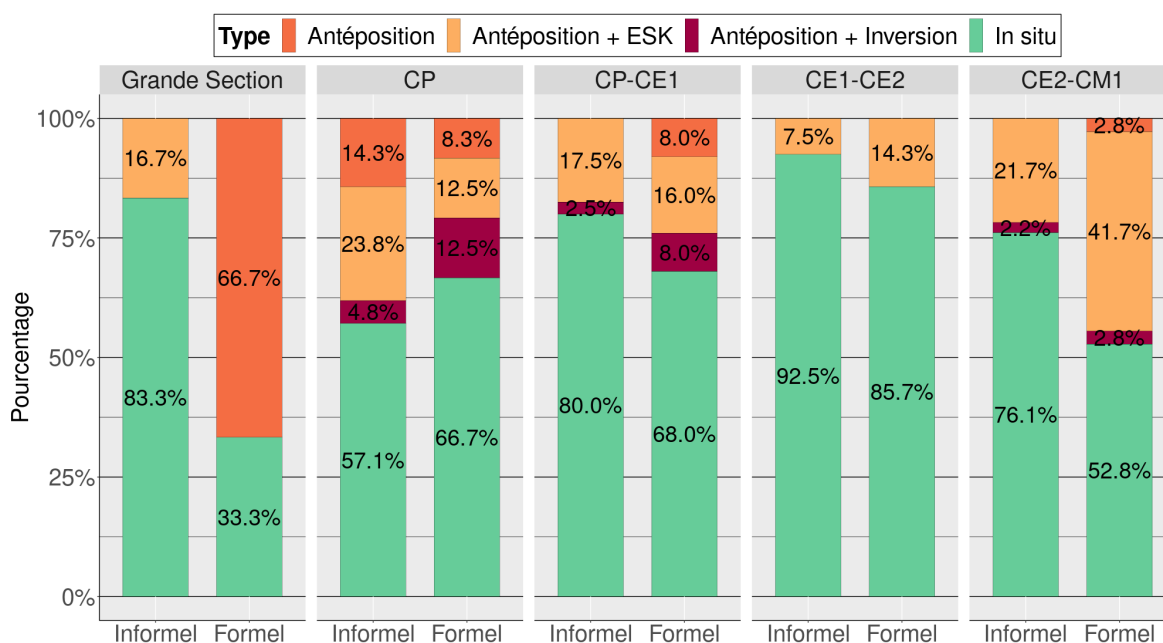


FIGURE 6.17 – Interrogatives partielles par classe et contexte (Informel/Formel) (JDR-C)

Un modèle de régression bayésienne a été calculé sur ces données (4 chaînes de 3000 itérations chacune, convergence atteinte et vérifiée en regardant que tous les Rhats=1.00). Semblable au précédent modèle présenté, seule la variable indépendante a été modifiée, puisqu'il s'agissait cette fois d'une interaction entre la formalité du CONTEXTE (*formel* ou *informel*, avec pour niveau de référence les contextes informels) et la CLASSE considérée comme variable ordinale (voir Annexe P0.2).

Ce modèle général confirme les éléments repérés plus tôt. En comparaison avec l'évolution des productions in situ à mesure que la classe des enfants augmente, il y a bien un délaissement des formes avec antéposition simple ($\hat{\beta} = -10.52$, 95%CrI=[-36.86,-0.50], $P(\beta) < 0=0.98$) et une tendance à l'augmentation du recours aux formes en *est-ce que* ($\hat{\beta} = 0.30$, 95%CrI=[-0.50,1.13], $P(\beta) > 0=0.77$). Aucune variation notable n'est corrélée par contre pour ce qui est des antépositions avec inversion ($\hat{\beta} = 0.78$, 95%CrI=[-3.80,6.84], $P(\beta) > 0=0.62$).

Pour ce qui est des interactions CONTEXTE*CLASSE, le modèle appuie l'idée d'une augmentation du recours aux divers types d'antépositions dans les contextes formels par les enfants à mesure que leur classe est avancée, en miroir de la diminution des productions in situ. Ainsi notamment, l'augmentation des productions d'antéposition en *est-ce que* par rapport aux productions in situ, quand les enfants passent d'un contexte informel à un contexte formel, est amplifiée à mesure que la classe est élevée ($\hat{\beta} = 0.65$, 95%CrI=[-0.08,1.40], $P(\beta) > 0=0.96$). C'est également le cas pour les antépositions simples ($\hat{\beta} = 13.35$, 95%CrI=[1.41,49.91], $P(\beta) > 0=1$) mais la faible quantité de données ne peut qu'inciter à la prudence.

De manière générale, ces données corroborent néanmoins une adaptation linguistique qui devient progressivement plus importante en contexte formel à mesure que les enfants poursuivent leur développement, notamment scolaire.

6.3.5.3 Un effet de genre en interaction avec la formalité

Cette adaptation progressive avec l'âge n'est pas cependant la seule donnée semblant influencer sur les choix faits en production par les enfants. La donnée du genre des enfants, non considérée lors de la conception de cette tâche, semble ainsi avoir une influence sur les productions linguistiques. Pour rappel, il y avait 31 filles et 37 garçons parmi les enfants qui ont participé à la tâche.

À l'abord, il ne semble pas y avoir de grande différence en production entre des groupes d'enfants définis par leur genre. Le Tableau 6.16 et la Figure 6.18 montrent ainsi que la proportion de phrases in situ produites ne semble pas varier grandement entre filles et garçons. Il semble pourtant y avoir moins de différences entre les productions avec antéposition chez les filles que chez les garçons. Dit autrement, les filles varient plus leurs stratégies interrogatives alors que les garçons se reposent surtout sur l'antéposition en *est-ce que*, en tant qu'alternative à l'in situ.

	Antéposition simple	Antéposition + ESK	Antéposition et inversion	In situ	Total
Filles	8	14	4	80	106
Garçons	2	38	5	118	163
Total	10	52	9	198	269

TABLEAU 6.16 – Interrogatives partielles produites selon le genre de l'enfant (JDR-C)

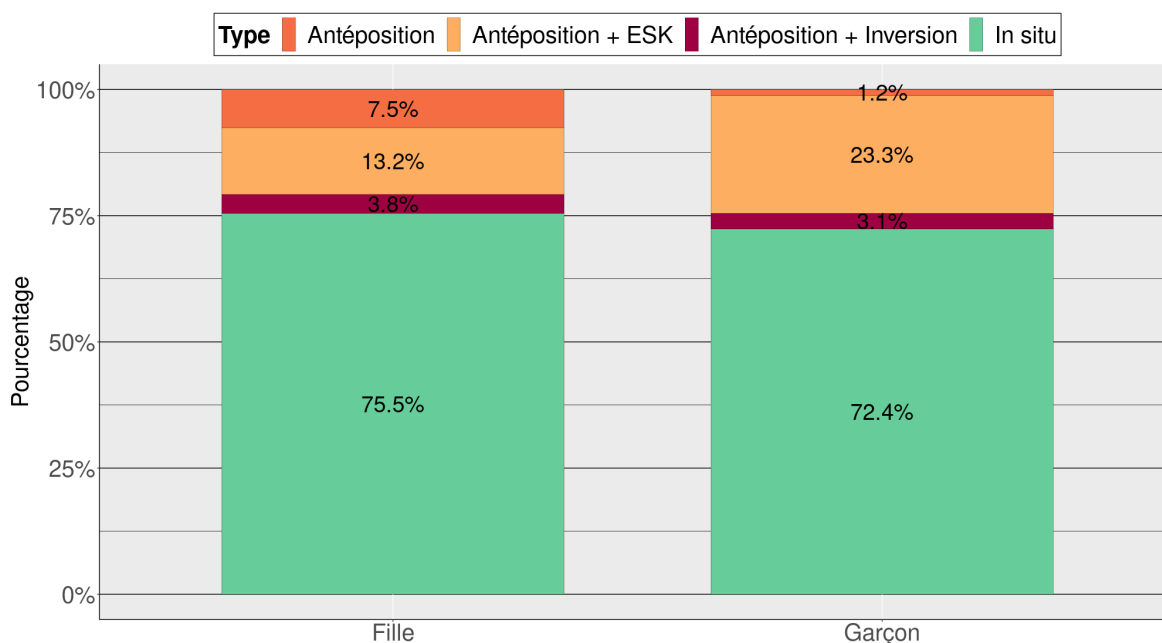


FIGURE 6.18 – Interrogatives partielles produites selon le genre de l'enfant (JDR-C)

C'est ce que confirme un modèle de régression bayésienne calculé avec pour variable indépendante le TYPE syntaxique des interrogatives, en variable indépendante le GENRE des enfants (masculin ou féminin, avec *féminin* pour référence), et un effet avec intercept aléatoire "enfant" (voir Annexe P.0.3). La convergence a été atteinte avec 4 chaînes et 6000 itérations par chaîne (Rhats = 1.00 pour tous les paramètres). Ce modèle confirme de façon fiable que, avec pour niveau de référence les productions in situ, les garçons produisent moins d'antépositions simples que les filles ($\hat{\beta} = -9.14$, 95%CrI=[-33.12,4.28], $P(\beta) < 0.91$), mais plus d'antépositions en *est-ce que* ($\hat{\beta} =$

1.59, 95%CrI=[-0.37,3.97], $P(\beta)>0=0.94$). En parallèle, il n'y a pas de différence visible entre leurs productions avec antéposition et inversion ($\hat{\beta} = -0.94$, 95%CrI=[-12.72,10.07], $P(\beta)<0=0.58$).

Un effet de genre n'était cependant pas attendu pour cette tâche. En outre, selon un rapide calcul, chaque fille a produit en moyenne 3.4 interrogatives (106/31) alors que chaque garçon en a produit en moyenne 4.4 (163/37). Cette différence pourrait indiquer que les garçons ont été généralement plus volubiles que les filles. Une possibilité serait que, l'équipe pédagogique de l'école étant quasi-uniquement composée d'enseignantes, un effet de contraste aurait restreint les filles dans leur production linguistique face à deux intervenants extérieurs hommes, l'expérimentateur et le marionnettiste. Un travail d'approfondissement, conçu avec cette perspective, serait nécessaire.

6.3.5.4 Nuances sur la notion de formalité

Un autre paramètre à prendre en compte est le caractère subjectif de la notion de *formalité*. La Section 6.3.5.1 a montré un effet général de formalité, mais celle-ci était définie *a priori* à la création de la tâche. Les marionnettes de Jeannot et de Marie devaient créer un contexte informel, avec pour l'enfant des destinataires proches et accessibles. De l'autre côté, les marionnettes du roi et de la reine incarnaient des personnages adultes, avec qui un certain degré de déférence serait de mise. Le Tableau 6.17 et la Figure 6.19 recensent les productions interrogatives des enfants en fonction de leur destinataire, et sont un indice que la problématique est plus complexe.

	Antéposition simple	Antéposition + ESK	Antéposition et inversion	In situ	Total
Jeannot	2	14		64	80
Marie	1	12	3	57	73
Reine	3	13	3	37	56
Roi	4	13	3	40	60
Total	10	52	9	198	269

TABLEAU 6.17 – Interrogatives partielles produites selon les destinataires (JDR-C)

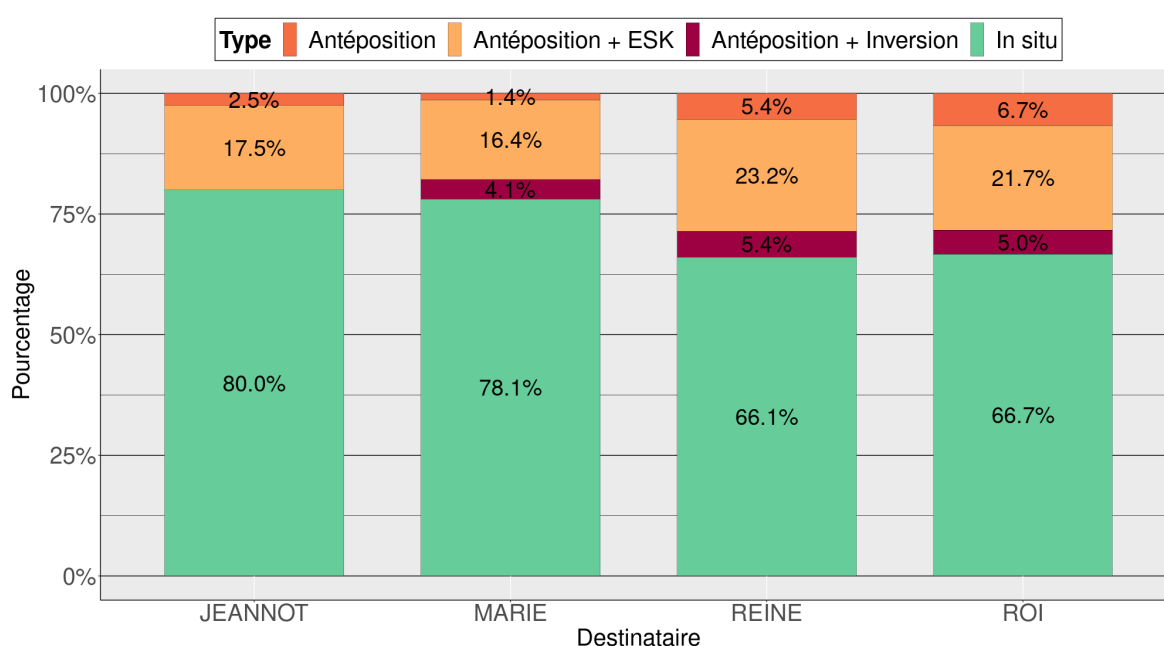


FIGURE 6.19 – Interrogatives partielles produites selon les destinataires (JDR-C)

Le comportement linguistique des enfants semble cohérent avec le roi et la reine, et globalement similaire entre les marionnettes de Marie et Jean. Même s'il s'agit de nombres très petits, l'absence totale de production d'antépositions avec inversion avec Jeannot est cependant notable. Il s'agit peut-être ici d'un écho de l'effet de genre précédemment évoqué, avec une proximité plus facile à percevoir pour les enfants avec un destinataire tel Jeannot, petit garçon qui joue au foot et parle de façon familière, qu'avec Marie. Cette différence de perception est peut-être explicable, à son tour, par une différence de performance du marionnettiste entre les deux personnages (ou alors celui-ci était simplement plus crédible en incarnant le petit garçon).

Quoi qu'il en soit, en allant plus loin encore dans les données et en mettant en parallèle la formalité estimée *a priori* des contextes et celle qui leur a été conférée par les enfants au cours de la tâche, une discordance apparaît. Le Tableau 6.18 et la Figure 6.20 comparent les productions d'interrogatives avec tutoiement ou vouvoiement pour chacune des marionnettes, pour les 233 interrogatives où un tel marqueur était présent (soit sur le verbe, soit dans un autre élément tel qu'un mot possessif).

	Jeannot	Marie	Reine	Roi	Total
Tutoiement	28	27	7	3	65
Vouvoiement	37	39	46	46	168
Total	65	66	53	49	233

TABLEAU 6.18 – Productions avec tutoiement/vouvoiement selon la marionnette (JDR-C)

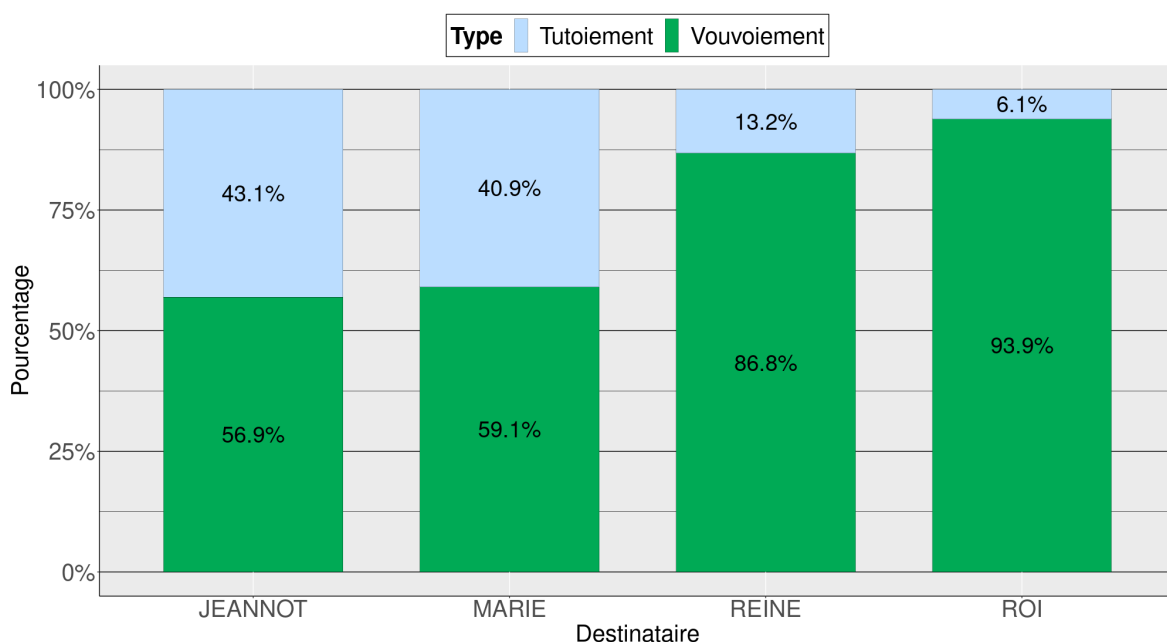


FIGURE 6.20 – Productions avec tutoiement/vouvoiement selon la marionnette (JDR-C)

Il n'est pas certain que l'utilisation par les enfants du tutoiement ou du vouvoiement correspondent à un marquage de leur part d'un niveau de formalité, et ce serait une hypothèse à explorer plus avant. En admettant cependant que ce soit le cas ici, il ressort ici que, si les marionnettes du roi et de la reine ont été correctement identifiées comme créant un contexte d'interaction « formel », celles de Jeannot et Marie ont par comparaison moins créé un contexte « informel ». S'il y a ainsi quelques productions discordantes avec les marionnettes royales, telles (128), il y en a en regard

plus avec les marionnettes enfantines, comme (129).

(128) qu'est-ce que tu voudrais comme légumes

(CHI58, Phase 354 (avec la reine) - #10451)

(129) pour vous faire digérer vous voulez quoi

(CHI40, Phase 531 (avec Marie) - #13512)

Une réanalyse des types phrastiques produits en fonction de cette dichotomie entre tutoiement et vouvoiement, sur les 233 interrogatives concernées, donne ainsi une vision un peu différente de celle présentée en Section 6.3.5.1. Le Tableau 6.19 et la Figure 6.21 permettent une visualisation des choix phrastiques effectués par les enfants dans leur production d'interrogative, en fonction du recours à une marque de tutoiement ou de vouvoiement.

	Antéposition simple	Antéposition + ESK	Antéposition et inversion	In situ	Total
Tutoiement	3	17		45	65
Vouvoiement		31	8	129	168
Total	3	48	8	174	233

TABLEAU 6.19 – Type interrogatif produit selon le tutoiement/vouvoiement (JDR-C)

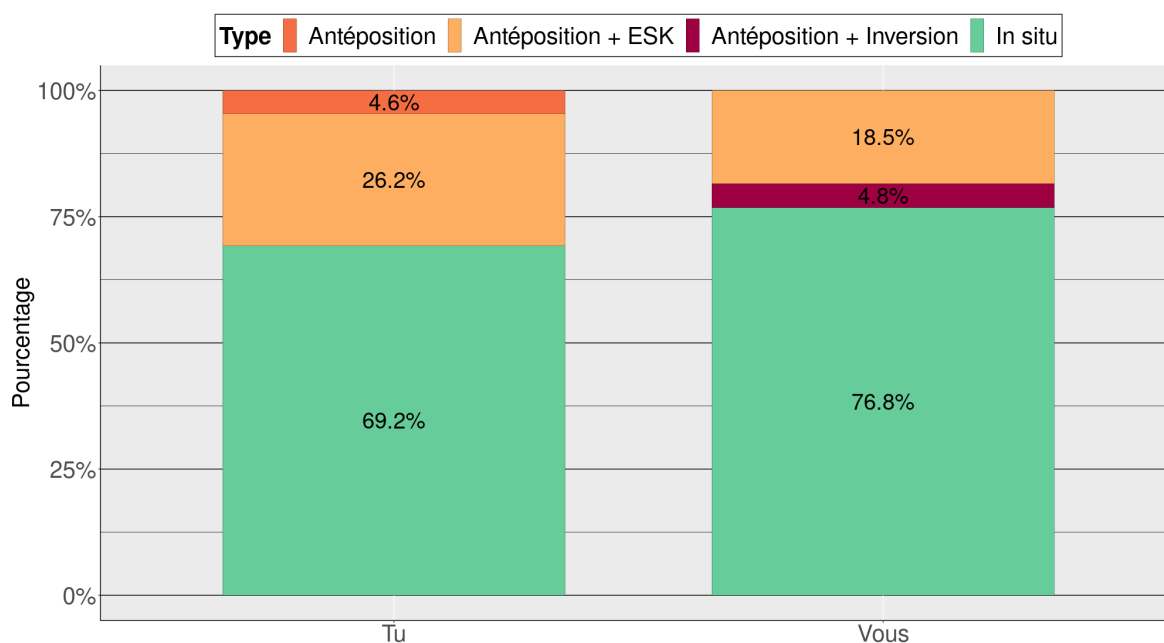


FIGURE 6.21 – Type interrogatif produit selon le tutoiement/vouvoiement (JDR-C)

Cette fois-ci la proportion de phrases in situ est plus importante lorsqu'il y a vouvoiement que lorsqu'il y a tutoiement. Pour rappel, c'était l'inverse lorsque les contextes *a priori* formels sont opposés aux contextes *a priori* informels. Outre ceci, il est notable que les formes plus marquées (en l'occurrence, moins nombreuses), l'antéposition simple et l'antéposition avec inversion sont plus clairement réparties : l'antéposition simple semble réservée à des usages avec tutoiement, alors que l'antéposition avec inversion semble réservée à des usages avec vouvoiement.

Cette observation, en tension avec les observations tirées d'une définition *a priori* de la formalité, mériterait un approfondissement dans une autre tâche expérimentale. Il serait notamment

important d'isoler les raisons qui poussent l'enfant, individuellement, à considérer un contexte comme informel.

Une dernière observation prolonge cette réflexion. Comme annoncé plus haut, même si cette tâche de jeu de rôle visait à isoler des tendances interpersonnelles à l'échelle d'un groupe d'enfants hétérogènes, il convient de rappeler que cette hétérogénéité a toute son importance. La variabilité des productions de chaque enfant ayant participé est visible dans la Figure 6.22, p. 211.

La diversité des stratégies enfantines y est manifeste. 4 enfants seulement sont ainsi responsables des 9 productions avec antéposition et inversion. 4 autres enfants également sont responsables des 10 productions avec antéposition simple. 23 enfants, sur les 42 ayant produit les 269 interrogatives retenues pour analyse, n'ont produit qu'un seul type interrogatif. Parmi les 8 enfants qui ont produit les types phrastiques marqués (antépositions simple et avec inversion), 1 n'a autrement produit que des phrases in situ (CHI69), et l'autre n'a produit que ce type marqué (CHI12). Au bilan :

- 23 enfants sur 42 n'ont produit qu'un seul type phrastique,
- 13 enfants sur 42 ont produit deux types phrastiques différents,
- 6 enfants sur 42 ont produit trois types phrastiques différents,
- aucun enfant n'a produit les quatre types phrastiques repérés.

Cette observation n'annule pas les tendances générales dégagées dans l'analyse. Chacun et chacune des enfants qui ont participé à ce jeu y ont réagi différemment, en fonction de leur humeur et de leurs pensées du moment. La variabilité de ces données est peut-être à relier aux propos de LAHIRE [2019]. L'auteur y présente une analyse des productions linguistiques d'enfants de milieux socio-économiques variés et scolarisés dans des écoles différentes. Ses conclusions sont que le comportement linguistique des enfants est le reflet de la condition socio-économique mais aussi socio-culturelle du foyer et du milieu dans lesquels ils et elles grandissent.

Cette raison pourrait expliquer notamment que les interrogatives avec antéposition et inversion aient été produites par des enfants aussi bien de CM1 (CHI68, 9 ans), de CE2 (CHI66, 8 ans), de CE1 (CHI63, 7 ans), que de CP (CHI30, 6 ans). Elle pourrait aussi expliquer que ces enfants ont produit ces formes marquées dans un contexte qui s'y prêtait (avec vouvoiement), et en majorité dans les contextes définis *a priori* comme formels lors de la création de la tâche. Le développement des associations entre forme linguistique et stéréotype social est sans doute plus précoce pour ces enfants.

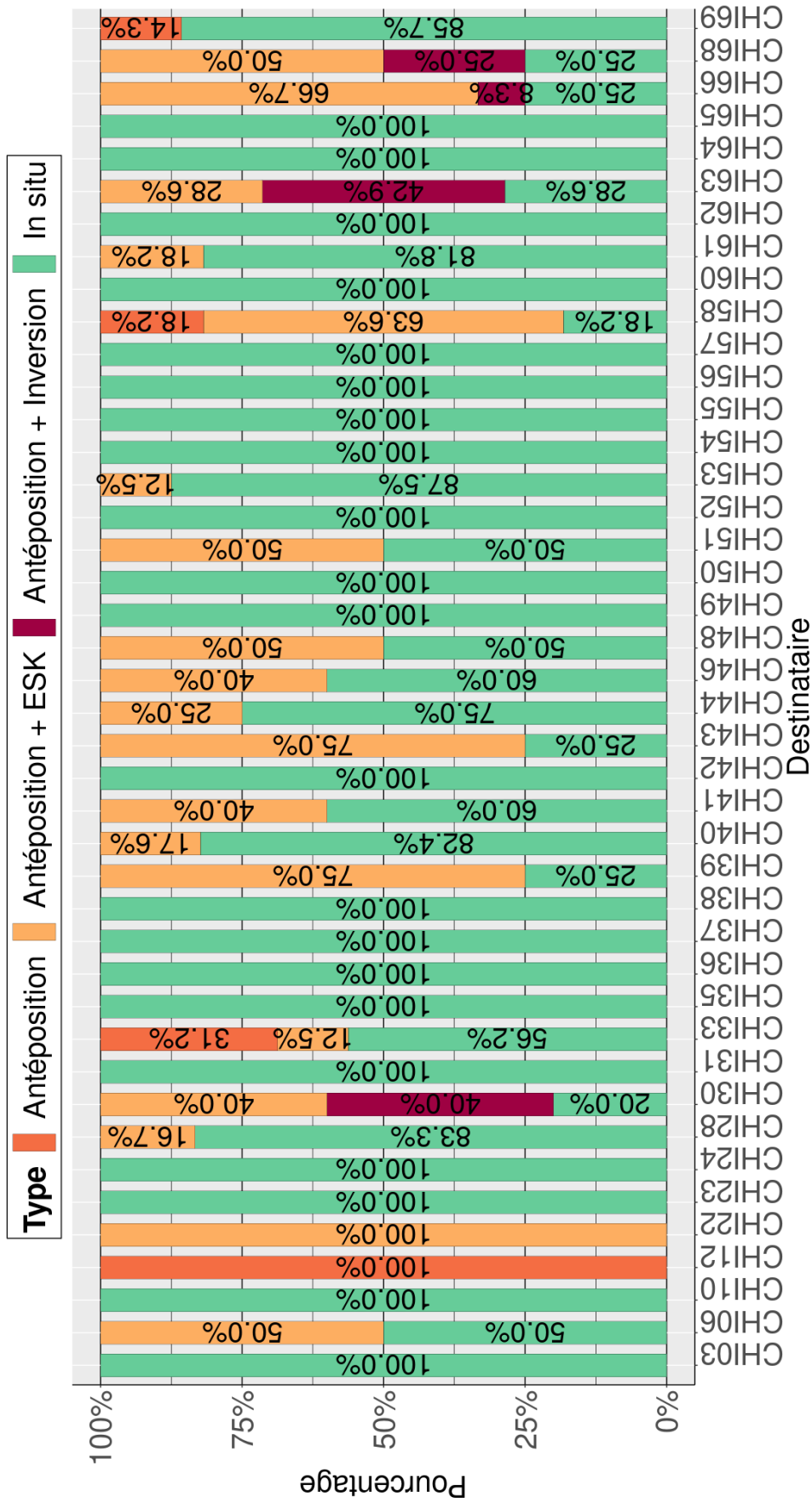


FIGURE 6.22 – Productions interrogatives par enfant (IDR-C)

6.4 Conclusion intermédiaire

La tâche de *matched-guise* enfantin (MG-C) décrite en Section 6.2 explore l'intégration et le développement de stéréotypes sociolinguistiques chez l'enfant. Les résultats corroborent une conscience précoce des enfants quant au lien entre production linguistique et situation interactionnelle. Celle-ci est en partie définie par le statut social, en l'occurrence des allocutaires : « les gens parlent différemment selon leurs caractéristiques sociales propres ».

Par ailleurs, cette étude a montré que les associations enfantines entre forme linguistique et stéréotype évoluent avec le temps. L'antéposition avec inversion devient par exemple la forme interrogative la plus associée à l'activité de lecture avec le temps, pour en arriver à une association conforme à celle observée chez l'adulte (MG2, Section 5.3.2). Sur l'association entre forme interrogative et richesse, on note aussi une évolution progressive, avec une neutralisation de la différence entre types phrastiques pour les enfants des classes avancées, alors que les enfants plus jeunes semblent associer les phrases *in situ* à des personnes riches. Il est envisageable que les représentations de ce stéréotype soit reconsidérées par les enfants à mesure que leur connaissance du monde s'étend, en partant d'une association « personne riche = personnalité médiatique » pour arriver à une vue plus nuancée. Cette nuance est peut-être d'ailleurs apportée en partie par la co-construction et la mise en relation avec d'autres stéréotypes (par exemple celui du statut social ou des études).

Certaines associations entre formes linguistiques et stéréotypes sociaux sont au contraire intégrées dès le plus jeune âge. Un exemple est l'association entre antéposition avec inversion et l'image d'une personne médecin, laquelle est déjà opérante chez les enfants de Grande Section. Cette association précoce entre forme linguistique et stéréotype social se voit par le contraste, maintenu dans le temps, avec le type interrogatif *in situ*. Comme chez l'adulte, celui-ci n'est pas associé par les enfants à l'image d'une personne médecin, et ce dès la maternelle jusque le CM2. Ici encore, la connaissance du monde est un paramètre important à considérer, puisque les enfants sont dès le plus jeune âge régulièrement en contact avec des personnes médecin. Dans la lignée des analyses proposées par BUSON [2009]; BUSON et al. [2014], il leur est donc possible très tôt, le cas échéant, d'établir des statistiques entre des éléments linguistiques *saillants* (par exemple la forme antéposée avec inversion) et une situation sociale *particulière* (par exemple celle d'aller chez le médecin).

L'activité de jeu de rôle (JDR-C) décrite en Section 6.3 explore l'influence du contexte extralinguistique pris au sens large sur les productions enfantines. Les résultats montrent une influence générale de la formalité telle que définie par le statut des personnages : les productions sont différentes selon que la marionnette avec laquelle interagit l'enfant est celle d'un enfant ou celle d'un adulte. En particulier, il semble y avoir plus de stratégies avec antéposition (quel que soit le type) et au contraire moins de productions *in situ* avec les marionnettes adultes.

Par ailleurs, un développement sociolinguistique progressif est mis en évidence puisque, à mesure que leur classe est avancée, les enfants semblent faire de plus en plus la distinction entre ces contextes formels et informels. Il y a notamment de moins en moins de productions *in situ* en contexte informel quand les enfants grandissent. La notion de formalité et de son appréciation par les enfants reste néanmoins à explorer et à affiner. Elle était en effet définie ici *a priori*, lors de la conception de l'étude, sans pouvoir vérifier avec certitude qu'elle était bien opérante lors des activités.

Enfin, une influence non-prévue du genre des enfants et une forte variabilité interindividuelle, détectées dans la tâche de jeu de rôles, appuient le besoin d'approfondir ces résultats. Une analyse qualitative des données, en parallèle de ces études plus systématiques, le permettrait sans doute.

Cependant, les données récoltées montrent néanmoins la pertinence méthodologique de tâches menées à grande échelle, de façon la plus systématique possible et ceci même en production, avec des populations enfantines. C'est ceci qui a permis un parallélisme, et donc la comparaison des résultats avec les données dégagées chez l'adulte, notamment pour la tâche de *matched-guise*. Ce protocole assez facile à mettre en oeuvre serait un outil pertinent, s'il est réutilisé, pour construire une vue plus globale sur le réseau des stéréotypes construits par les enfants. Ce type de tâche peut aussi permettre, par la confrontation de groupes d'âge différents, de visualiser le développement des stéréotypes sociaux associés à certains éléments linguistiques, et le moment où ceux-ci deviennent opérants chez l'enfant dans un sens semblable à celui de l'adulte.

6.5 Apports majeurs

1. **Les associations entre forme linguistique et stéréotypes sociaux apparaissent précocement chez l'enfant, au moins dès 3-4 ans.**
2. **Ces associations évoluent, en particulier car les stéréotypes intégrés par les enfants changent avec leur apprentissage du monde, mais aussi sans doute car leur confrontation à des situations sociales se diversifient.**
3. **Les enfants montrent rapidement une capacité à adapter leur production linguistique à la formalité du contexte.**
4. **Cette capacité et le degré auquel elle est mise en oeuvre se développe néanmoins avec le temps, et les enfants des classes avancées s'adaptent ainsi plus visiblement à la différence de contextes que les enfants plus jeunes.**
5. **L'influence d'autres facteurs (genre, conception de la formalité) sur le développement de cette capacité d'adaptation au contexte n'est pas à oublier, non plus que la forte variation interindividuelle entre les enfants.**

Chapitre 7

Conclusion Générale

Sommaire

7.1 Une réanalyse des usages interrogatifs en français	217
7.2 L'acquisition des interrogatives en français : un phénomène social	220
7.3 Réflexions sur la démarche et les méthodes utilisées	222
7.4 Perspectives	223

L'objectif général de ce travail était double, avec deux questions principales :

1. À quel point les francophones adultes projettent-ils des sens différents sur chacune des variantes interrogatives auxquelles ils ont accès?
2. Ces considérations peuvent-elles aussi influencer sur les productions enfantines précoces?

Ce chapitre résume les principales avancées empiriques décrites dans ce travail, et des éléments de réponse que les données issues de corpus et d'études expérimentales permettent d'avancer à ces questions. Une réflexion plus générale sur la démarche dans laquelle ces études ont été menées sera aussi fournie. Enfin, des pistes de prolongation des recherches menées pour cette thèse seront discutées.

7.1 Une réanalyse des usages interrogatifs en français

Le point de départ de ce travail est l'existence en français de formes linguistiques concurrentes pour réaliser un même acte de langage : l'interrogation partielle. Les francophones peuvent construire des phrases syntaxiquement assez différentes pour demander une même information manquante sur le monde. De multiples théories ont déjà été avancées pour expliquer ce phénomène. Des facteurs syntaxiques, et pragmatiques ont été avancés pour expliquer cette alternance des formes. Certains travaux ont élargi l'approche de ce fait linguistique en intégrant des paramètres liés aux individus produisant ces formes (groupe social) ou des effets diachroniques. L'objet du Chapitre 2 a été de montrer la nécessité d'élargir encore la perspective prise, et de prendre en compte toute la dimension utilitaire du langage.

La littérature sociolinguistique contemporaine s'affranchit en effet peu à peu d'une vision sociolectale de la variation et des phénomènes d'alternance linguistique. Une approche de « sociolinguistique enrichie », de type *troisième vague* [ECKERT, 2012], fait désormais de la variation un outil par lequel l'identité sociale de l'individu est modifiée tout au long d'un échange linguistique. Ceci passe par une spécialisation de certaines formes linguistiques, associées plus que d'autres à des représentations sociales stéréotypées. Par le recours à ces formes sociologiquement marquées, une personne peut certes revendiquer une identité de groupe (et donc ici la variation s'identifie en partie par des critères sociolectaux). Cependant, cette personne peut aussi modifier son image dans l'interaction (*persona*). C'est ainsi que le comportement linguistique du président américain Barack Obama change d'un contexte interactionnel à l'autre, suivant qu'il donne une interview ou participe à un barbecue public, dans les travaux de LABOV [2012].

De cette analyse de la variation découle la nécessité d'explorer, ou de réexplorer, l'usage fait par les francophones des différentes formes interrogatives à leur disposition. En premier lieu ce travail a donc consisté en une compilation de données empiriques, tirées de deux corpus de français oral spontané : EPAC [BAZILLON et al., 2011; ESTÈVE et al., 2010] et ESLO [ESHKOL-TARAVELLA et al., 2011; LLL, 2017]. Dans le Chapitre 4, le premier corpus a permis de dresser les contours d'une utilisation *stratégique* des différentes formes interrogatives du français, à contexte constant. Dans un contexte a priori formel d'émission radiophonique, l'antéposition avec inversion et fréquente et, entre autres, le type *in situ* permet d'amener de l'informalité dans l'interaction et un rapprochement momentané avec les allocutaires. Le second corpus a permis en complément de croiser les *contextes interactionnels* et d'y comparer les usages interrogatifs. Des contextes au public et aux enjeux différents sont liés à des proportions d'emploi différentes des formes interrogatives.

Un contexte informel comme un repas est lié à plus de réalisations in situ alors qu'un entretien avec une personne inconnue (membre d'une équipe de recherche) s'illustre par un recours plus important aux différentes formes d'antéposition.

Ces données empiriques de corpus ont montré un usage circonstancié des différents types interrogatifs (stratégique et dépendant du contexte). Cet usage, et notamment la proportion d'emploi importante des réalisations in situ, semble entrer en contradiction avec l'idée d'une forme antéposée avec inversion promue comme « standard » interrogatif de la langue française, par exemple dans des ouvrages de référence comme RIEGEL et al. [2014]. Les études expérimentales présentées au Chapitre 5 ont exploré la perception qu'ont les francophones de ces formes linguistiques, et les représentations sociales qui y sont associées.

Quatre études de jugements d'acceptabilité ont ainsi permis de sonder explicitement les francophones sur leurs préférences linguistiques. À l'abord (AJT1, AJT2) celles-ci ont semblé se porter sur le *standard*, avec une acceptabilité bien supérieure de l'antéposition avec inversion par rapport à l'antéposition simple et aux phrases in situ. L'étude AJT3 nuance néanmoins ceci et permet de voir que les francophones établissent une réelle distinction entre ce que doit être un « bon français » et leur usage réel de la langue. Sur la première question, un « gradient de bonne formation » semble apparaître (AJT3, AJT4 : antéposition avec inversion > antéposition en *est-ce que* > in situ > antéposition simple). Sur la seconde question en revanche, celle de l'adaptation à une situation d'interaction courante, les réponses des volontaires sont moins tranchées. L'antéposition avec inversion est moins adaptée que le type in situ dans certains contextes, notamment informels, et auprès de certains groupes de population (ici définis par l'âge, et en l'occurrence en-dessous d'un seuil de 30 ans). Ceci, en cohérence avec les usages observés sur corpus.

Trois études de *matched-guise* ont par ailleurs permis de dessiner un réseau de stéréotypes sociaux associés aux différentes formes interrogatives par les francophones. Ces associations entre forme linguistique et stéréotypes semblent modulées par des effets de contexte (formel vs. informel) et ici aussi par une notion de groupe d'âge. Globalement néanmoins, il semble que les interrogatives avec antéposition et inversion soit associées à des indices sociaux de différentes natures, et liées à des considérations générationnelles (ici, un âge plus élevé), mais aussi sociales (richesse et éducation plus importante, origine géographique plus urbaine, loisirs plutôt tournés vers la lecture que le sport), ou encore psychologique (ici, une sociabilité un peu moindre). Ce n'est pas dans l'absolu que ces associations se définissent, mais par opposition aux autres formes interrogatives. Le type in situ est ainsi associé à moins de richesse, au sport, à une éducation plus courte par comparaison, et à une sociabilité plus importante. On peut comme auparavant établir un gradient sur certains points, avec par exemple l'antéposition en *est-ce que* qui semble associée à des stéréotypes à un degré intermédiaire entre ces deux extrêmes. La Figure 7.1 rappelle le profil social *type* des personnes utilisant les variantes interrogatives, tel que dégagé dans les études MG2 et MG3.

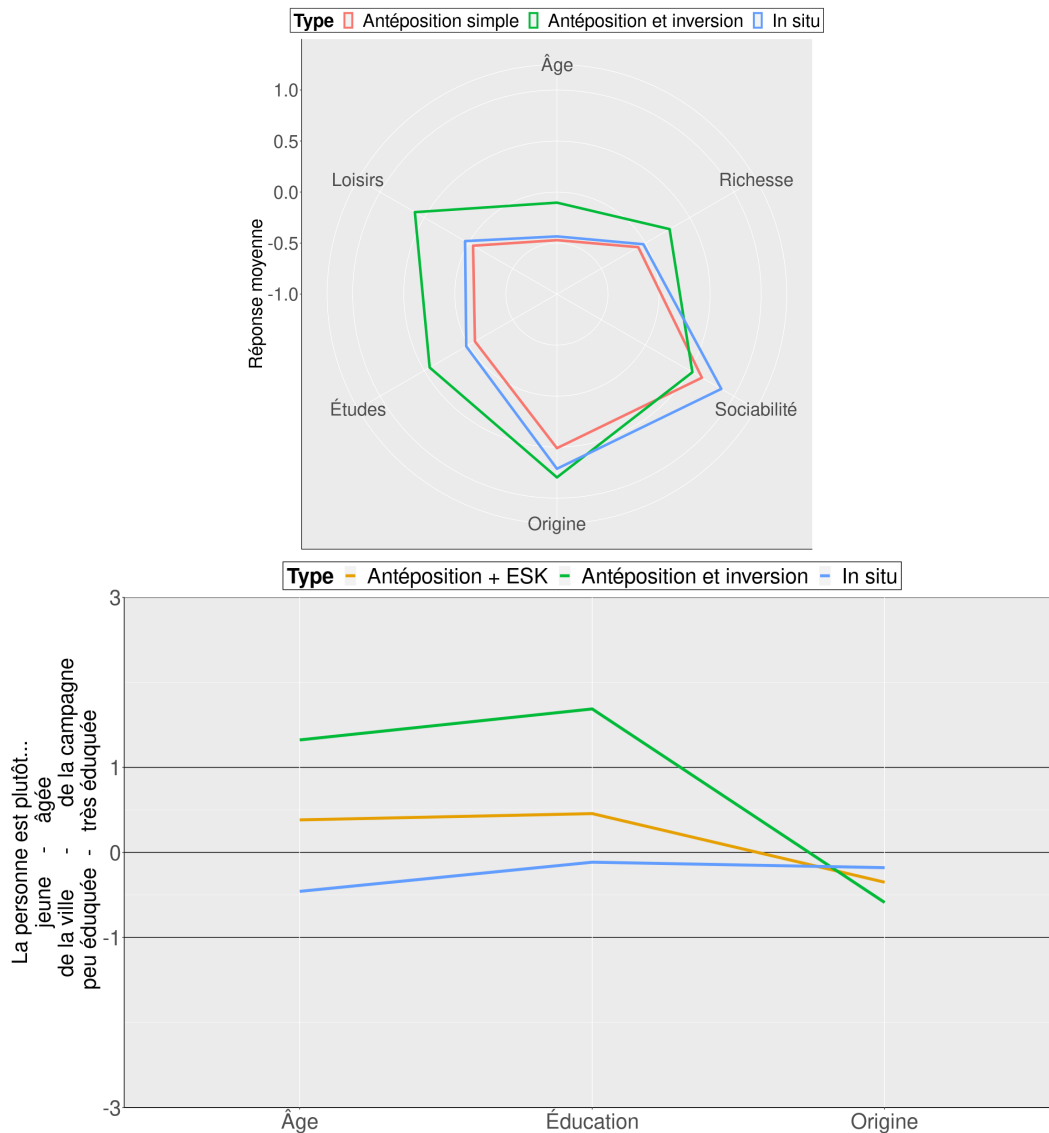


FIGURE 7.1 – Profils sociaux types liés à l'utilisation des variantes interrogatives (MG2-3)

Le lien entre ces préférences affichées par les francophones dans les jugements d'acceptabilité et les représentations mentales associées aux différents types phrastiques n'est pas trivial. Il est probable que la vision de l'antéposition avec inversion comme le *standard* français est entretenue par la vision stéréotypée des gens qui l'utilisent : des personnes riches, aux études longues, qui lisent souvent, etc. À l'inverse, la vision stéréotypée des personnes qui utilisent l'antéposition avec inversion auto-entretient sans doute elle aussi le statut de l'antéposition avec inversion comme forme de *référence* en matière de « bon français ». Indépendamment de ces considérations, les ramifications de cette interaction sont cependant importantes.

L'analyse des usages francophones de l'interrogation partielle doit inclure ce paramètre social élargi. La forme même des phrases interrogative porte un sens, social, qui doit être intégré dans leur étude ou leur formalisation. L'enjeu social qui pèse sur les francophones lors du recours à une interrogative partielle interfère certainement avec les autres facteurs déjà dégagés, que ceux-ci soient syntaxiques, phonotactiques ou pragmatiques. Par ailleurs, les analyses menées dans d'autres domaines doivent aussi intégrer ces facteurs sociaux étendus. C'est typiquement le cas des études menées en acquisition du langage.

7.2 L'acquisition des interrogatives en français : un phénomène social

À la seconde question posée en introduction, les données expérimentales présentées dans le Chapitre 6 apportent une réponse positive. En compréhension comme en production, les enjeux sociaux de la variation dans les interrogatives partielles sont déjà visibles pour les enfants.

Une première tâche de *matched-guise* a ainsi mis en évidence l'existence d'associations entre formes linguistiques et stéréotypes sociaux chez l'enfant. Cette idée n'est pas neuve en soi : les enfants connaissent nécessairement un développement sociolinguistique pour parvenir à une maîtrise adulte des enjeux sociaux du langage. Les travaux de BUSON [2009]; BUSON et al. [2014] sur l'émergence des stéréotypes sociolinguistiques par effets de saillance et un jeu de statistiques inscrit dans les théories de l'exemplaire de BYBEE [2013]; BYBEE et HOPPER [2001] le montrent déjà.

Cependant les théories sur l'acquisition des interrogatives en français doivent encore le prendre en compte. L'idée d'un développement progressif du répertoire interrogatif chez l'enfant reste encore largement inspiré par le postulat d'une complexité graduelle des structures interrogatives. Or la problématique reste réelle de savoir si la réalisation in situ est la plus simple pour l'enfant, ou si elle pourrait – par exemple – n'être que la conséquence d'un effet de fréquence dans l'input. Dans des théories non inspirée des cadres génératifs, l'antéposition n'est en effet pas plus complexe *a priori* que les type in situ. On pourrait même envisager qu'en termes de traitement de la phrase, le fait d'avoir l'élément interrogatif en début d'énoncé facilite la lecture de l'acte de langage comme une interrogation (sans rien présumer par ailleurs sur une éventuelle complexité lié à un ordre différent des constituants autres que l'élément interrogatif). L'effet de fréquence pourrait venir quant à lui du fait que les enfants évoluent dans un contexte généralement peu formel avec leurs proches.

L'objectif de ce travail n'est cependant pas de remettre en cause les théories sur la complexité des phrases interrogatives, et son influence possible sur les productions enfantines précoces. Les observations de travaux comme THIBERGE [2017], qui semblent corroborer une différence de comportement linguistique chez l'enfant comme chez l'adulte selon que l'élément interrogatif comporte ou non une préposition, sont difficiles à intégrer directement à une réflexion uniquement guidée par le sens social des types phrastiques. En outre, l'âge des enfants qui ont participé aux études de ce travail de thèse est déjà avancé (3-11 ans), et la production d'interrogatives commence avant cette période (pour rappel, 2 ans selon PRÉVOST [2009], malgré de fortes variations interindividuelles).

Il n'y a pas de doute sur le fait que les enfants dont l'âge est le plus avancé, parmi celles et ceux qui ont participé aux activités, maîtrisent l'antéposition de l'élément interrogatif. Malgré cela, ils produisent encore beaucoup de phrases in situ, de la même façon que les 15-25 ans du corpus ESLO en produisent aussi une proportion importante. Au vu des résultats notamment de la tâche de *matched-guise* enfantin, il paraît difficile de ne pas considérer une influence, sur les plus jeunes enfants, des mêmes facteurs qui ont une influence sur leurs aîné.e.s. Dès les plus jeunes âges, en effet, les enfants ont déjà intégré des associations entre forme linguistique et stéréotype social, comme le rappelle la Figure 7.2.

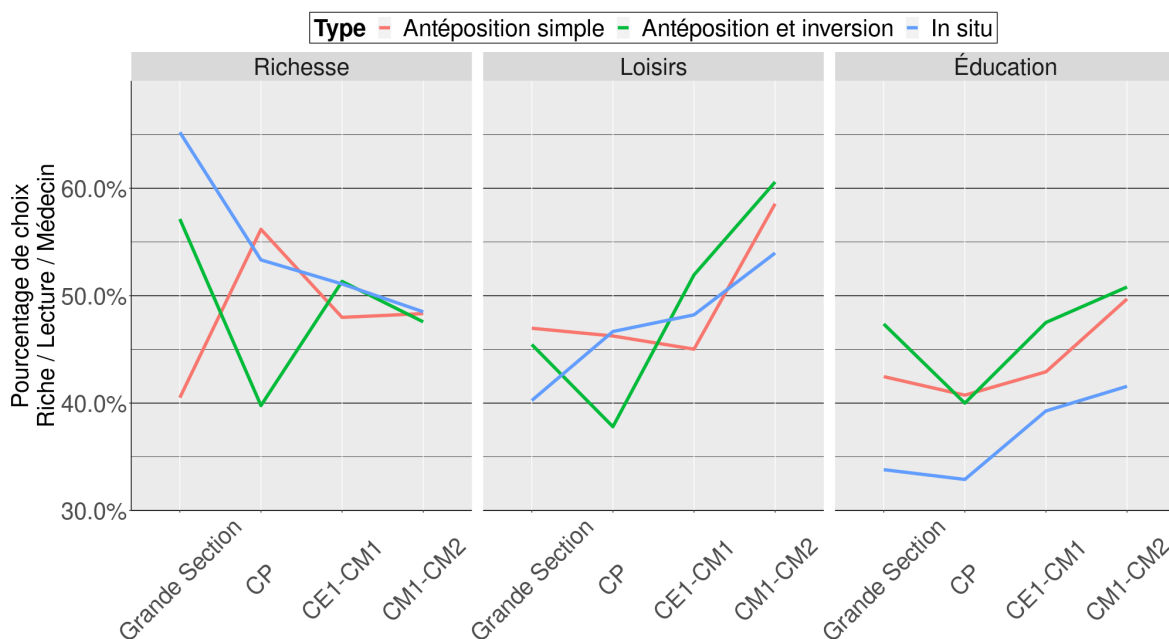


FIGURE 7.2 – Évolution du profil stéréotypique des interrogatives partielles (MG-C)

Dès la Grande Section, l'antéposition avec inversion est ainsi associée à des circonstances interactionnelles particulières comme une rencontre avec une personne médecin. Ou plutôt : le type in situ ne lui est pas associé. C'est une forme de phrase qui ne va pas avec cette situation. Une intégration aussi précoce des dimensions spécialisées des types phrastiques est à confronter aux productions linguistiques de l'enfant. Si l'enfant sait qu'une phrase ne correspond pas à une situation sociale c'est déjà le signe d'une conscience sociolinguistique associant une forme ou des formes syntaxiques particulières à un usage restreint.

Bien sûr, une nuance importante est ici que ces associations sont acquises petit à petit. De stéréotypes qui leur sont propres et correspondent à leur vision du monde, les enfants passent par étapes à une intégration des stéréotypes semblable à celle des adultes. C'est visible avec la conception de la richesse et à la « dé-spécialisation » progressive qui semble associée aux phrases in situ en relation avec ce stéréotype. Ceci est en lien avec une définition initiale de la richesse qui, d'après de premiers indices, pourrait être « être une personnalité publique de premier rang » (à la télévision par exemple, que ce soit pour dans le domaine sportif ou pas). À mesure que cette définition se modifie (s'élargit?) par la confrontation à d'autres exemples du monde qu'explore l'enfant, les formes interrogatives semblent se neutraliser sur ce point.

Cette évolution des stéréotypes en fonction de la confrontation de l'enfant au monde est importante. Elle rejoint en partie les thèses d'un développement linguistique influencé par l'usage [LANGACKER, 1987; TOMASELLO, 2009b]. Sans même prendre en compte l'influence des médias et des stéréotypes qu'ils véhiculent, les enfants reçoivent en permanence un input linguistique des adultes qui les entourent. Il peut s'agir de leurs parents, mais aussi des adultes qui les encadrent, à l'école, dans leurs activités de loisirs, etc. Un travail comparatif d'ampleur est encore nécessaire pour caractériser l'input parental en termes de types interrogatifs utilisés (voir ROWLAND et al. [2003] pour une description de ce type d'étude, sur l'influence de l'input maternel dans l'acquisition des éléments interrogatifs et des phrases interrogatives en anglais). Une étude semblable en français pourrait permettre de mieux comprendre la relation entre input et production enfantine, d'en

dégager les parallélismes, et éventuellement d’y déceler des effets de fréquence. Les extraits du corpus ESLO (130) et (131) donnent néanmoins déjà un aperçu de ce que peut entendre l’enfant au quotidien, dans un même contexte scolaire d’exercices.

- (130) le vendeur lui reprend son ancien appareil pour cent quatre euros et les accessoires qui le complètent pour trente-quatre euros **quelle somme** Monsieur Portrait devra-t-il ajouter? le vendeur accepte un paiement en trois versements (...)

(ESLO2ECOLE1278C (35-55) - #267)

- (131) si tu avais deux kilos de bonbons tu ferais **quoi** comme opération? réfléchis à cette question et ça te donnera la réponse

(ESLO2ECOLE1280C (35-55) - #297)

Ici, l’enfant en (130) entend une antéposition avec inversion, et en (131) une phrase in situ, durant le même type de tâche. Dans le premier exemple l’enseignant.e est en train de lire une consigne sans destinataire précis, alors que dans le second il ou elle est en train d’aider directement l’enfant destinataire. Si de tels exemples se répètent, la nature différente de ces interactions peut aider l’enfant à construire l’association entre informalisation et phrase in situ, ou au contraire entre énoncé à visée formelle (consigne) et antéposition avec inversion. De là, par saillance et stéréotypie, peut émerger une spécialisation des formes (similaire à la spécialisation de constructions décrite dans [GOLDBERG \[2006\]](#)).

La tâche de jeu de rôle présentée au Chapitre 6 (Section 6.3) a permis d’explorer le développement de la capacité infantine à s’adapter à des situations interactionnelles différentes. Par le biais du jeu avec des marionnettes au statut social différent (roi/reine vs. jeunes enfants), cette activité a permis de tester en production les associations forme-contexte intégrées en fonction de l’âge. Des indices ont émergé soutenant l’idée d’une installation progressive de la capacité, pour les enfants, à s’adapter à la formalité du contexte interactionnel et à modifier leur comportement linguistique en conséquence. Ici, néanmoins, la conception qu’ont les enfants de la notion de formalité demeure floue, et l’importance du facteur individuel est cruciale. Des études plus qualitatives, dans une démarche à la [LAHIRE \[2019\]](#), sont et seront un indispensable complément.

Toutes ces données invitent ainsi à une relecture des données disponibles sur l’acquisition des phrases interrogatives en français, et sur l’acquisition de phénomènes variationnels en général. Elles doivent par ailleurs être prolongées, et confrontées à d’autres données recueillies par des méthodologies différentes.

7.3 Réflexions sur la démarche et les méthodes utilisées

Ce travail a pris pour point de départ une base théorique déjà très diversifiée dans ses approches et ses moyens, pour décrire le phénomène variationnel de l’interrogation partielle chez l’adulte comme chez l’enfant en français. Même en réduisant le champ d’exploration aux facteurs sociaux influant sur les choix et les préférences des francophones, une diversité des méthodologies de recherche a été un précieux atout.

Les premières études expérimentales menées ici (décrites dans [THIBERGE \[2018\]](#)) ont permis de dégager des préférences globales chez les francophones. Les contraster aux études de corpus

préexistantes mais aussi menées ici (Chapitre 4, BADIN et al. [Accepté]; THIBERGE et al. [en révision]) a ainsi fait émerger les limites d'une conception simple de la notion d'acceptabilité. Une modulation des protocoles expérimentaux en alternant l'écrit et l'oral, en contrastant contextes formels et informels, et en complétant les jugements d'acceptabilité par des tâches de *matched-guise*, a permis de dresser un profil stéréotypique plus fin associé aux variantes interrogatives du français. Explorer les profils des volontaires aux études a par ailleurs permis de rendre compte de la variabilité des préférences selon le groupe social, en parallèle de phénomènes similaires observés en corpus.

Toutes ces données, assorties d'analyses statistiques permettant une appréciation fine et non-binaire des effets, ont donné une confirmation empirique supplémentaire des théories mentionnant un glissement des préférences linguistiques, et un changement dans la valeur données aux différentes variantes interrogatives, notamment dans la seconde moitié du 20^{ème} siècle. Cette théorie du changement linguistique elle-même peut expliquer les divergences et les difficultés à saisir dans un cadre formalisé synchronique, ou dans une modélisation qui serait statique temporellement, toutes les nuances régissant le phénomène variationnel.

Ce besoin d'un modèle évolutif qui puisse intégrer le glissement des valeurs sociales attachées à un phénomène variationnel, qu'il soit celui des interrogatives partielles en français ou un autre, apparaît plus important encore lorsque les données adultes sont confrontées à celles provenant d'enfants. L'individu jeune, en parallèle de son développement purement linguistique, doit intégrer des paramètres sociaux qui sont le reflet d'un monde contemporain toujours en changement. Un modèle qui ne prendrait pas en compte la différence entre norme et usage, entre habitudes de l'écrit et pratiques orales, ne peut pas rendre compte d'une langue intrinsèquement utilitaire. Ceci vaut également pour des modélisations plus abstraites, par exemple du langage en tant qu'aptitude humaine.

7.4 Perspectives

Il y a évidemment encore beaucoup à faire, ne serait-ce que sur la problématique des interrogatives partielles. En premier lieu, les protocoles utilisés pour cette thèse gagneraient à être modifiés sur certains points. Des échelles de stéréotypes plus largement inclusives seraient un plus, de même que multiplier les voix utilisées pour créer les stimuli. Tester d'autres dessins stéréotypés permettrait d'affiner la compréhension des réseaux d'associations formes-stéréotypes sociaux. Sa taille massive n'était pas prévue, et retravailler le corpus recueilli auprès des enfants est également nécessaire. En standardiser la transcription dans le cadre du projet CHILDES [MACWHINNEY, 2000, 2009], notamment, permettrait une meilleure étude de la construction de l'interaction chez l'enfant. Enfin, l'adaptation des protocoles de *matched-guise* adulte pour les enfants a montré qu'il était possible de sonder l'intégration implicite des relations forme-stéréotype à ces âges; il serait sans doute possible également de sonder les préférences explicites des enfants. Des tâches de jugements d'acceptabilité, modifiées et adaptées, seraient une piste de prolongation (voir AMBRIDGE [2010]; AMBRIDGE et ROWLAND [2013] pour des protocoles en ce sens utilisant des *smilies*). Elles permettraient de suivre l'évolution des connaissances métalinguistiques de l'enfant en parallèle de l'adulte, sur ce qui ressort du « bon » français, et sur le français qui peut s'utiliser dans différents contextes.

Plus largement, la simple comparaison de l'interrogation partielle à l'interrogation totale, avec des méthodologies similaires à celles utilisées ici, aurait doublé le travail. Une réflexion plus

générale sur le statut de l'inversion en français, en tant que marqueur d'un style ou d'une *persona* particulière, mériterait également de plus amples recherches.

Pour relier les résultats présentés ici à la thématique des contraintes cognitives pesant en acquisition du langage, une étude de la variation sociolinguistique par des méthodes plus directes (*online*) paraît une piste prometteuse. Les travaux menés par VAN BERKUM et al. [2008] ou VAN DEN BRINK et al. [2012] par exemple, qui analysent la violation de normes sociolinguistiques par des méthodes EEG, permettent d'envisager des protocoles qui pourraient s'appliquer au français. Grâce à ceux-ci, il paraît envisageable d'isoler le moment, à l'écoute ou à la lecture d'une interrogative partielle, où est attribué le stéréotype social qui lui est lié. Une éventuelle différence à ce niveau, entre les différentes variantes disponibles en français, relancerait la thématique d'un coût communicationnel ou procédural différent pour ces variantes. Un rapprochement avec les modélisations de MÜNSTER et KNOEFERLE [2018], plutôt bâti sur des recherches en suivi de mouvements oculaires, serait également intéressant.

La transposition chez l'enfant de ces protocoles expérimentaux (EEG et oculométrie), déjà complexes à mettre en oeuvre chez l'adulte, serait un nouveau défi à relever.

Malgré les nombreuses perspectives mentionnées ci-dessus, cette thèse a permis une mise en relief de l'ancrage fondamentalement social des productions linguistiques et des phénomènes variationnels. Cette base sociale se manifeste par des associations fortes entre des formes linguistiques concurrentes, ici les phrases interrogatives partielles du français, et des stéréotypes sociaux variés. Une théorie sur l'acquisition du langage ne peut en conséquence pas faire abstraction de ces enjeux sociaux dans son approche du fait linguistique, et doit au contraire les prendre compte comme un facteur déterminant.

Annexe A

Corpus EPAC : Extractions

Item	File	Speaker	Theme	Extraction
6	20040319	journalist	politics	comment fonctionne un conseil général?
8	20040319	journalist	politics	en quoi consistent les pouvoirs nouveaux qui leur sont ou vont leur être dévolus?
10	20040319	journalist	politics	je voudrais demander à nos invités ce que change la constitutionnalisation de la décentralisation d() constitutionnalisation qui a été () qu'est-ce que ça apporte de plus au chantier de la décentralisation?
16	20040322	expert	politics	() et c'est pas si simple que ça. qu'ont voulu dire les français là et maintenant? I(l) i(l) faut interpréter ce fameux message des français.
17	20040322	expert	politics	ça veut dire quoi? Ça veut dire qu'on est on souffre d'un certain nombre de situations qui sont mal acceptées...
18	20040322	journalist	politics	XD et FB n'ont pas arrêté pendant toute la campagne de dénoncer l'absence de projet de l'un et de l'autre : comment peuvent-ils se retrouver entre les deux tours et surtout autour de quel projet concret pour l'aquitaine?
25	20040325	journalist	health	// que peut-on dire à ceux par exemple dont le père a été victime de de ce type de cancer un père ou un oncle euh que peut-on leur dire? //
26	20040325	expert	health	... pourquoi lorsqu'on a eu un cancer on serait plus pénalisé que quand a eu un infarctus ou k() quand on a eu euh un accident de la route? c'est impossible et on doit résoudre ce problème
27	20040325	expert	health	Rassurez-vous c'est un peu le coeur de cette émission hein. Ça veut dire quoi se prémunir? c'est limiter les comportements à risque
28	20040325	journalist	health	à quoi concrètement va servir ce numéro de téléphone qui sera donc si j'ai bien compris disponible sur le cancer?

TABLEAU A.1 – Extractions du corpus EPAC retenues pour analyse

Item	File	Speaker	Theme	Extraction
30	20040325	expert	health	() d'abord i(l) faut qu'on parle du cancer dans sa globalité et quand même le cancer c'est quoi? c'est l'annonce quand même d'une terrible nouvelle ()
31	20040325	expert	health	le moyen () c'est () de faire le diagnostic de façon précoce. Alors comment on peut faire un diagnostic de façon plus précoce pas d'un mais des cancers c'est d'abord dédramatiser...
33	20040325	journalist	health	pascal il dit : « comment le cancer de la prostate est-il diagnostiqué? ()? »
34	20040326	expert	taxes	()? et si on est à temps partiel combien de temps on a travaillé dans l'année?
35	20040326	caller	taxes	euh je vous appelle parce que j'ai eu un petit arrêt maladie () et on me dit que () alors voilà ma question : pourquoi ressent-on () reçoit-on cet avis et doit-on le rajouter aux revenus imposables //
38	20040326	journalist	taxes	alors les raisons pour lesquelles vous recevez l'avis j'en sais rien cela dit () monsieur Iannucci que peut-on répondre? Alors
39	20040326	expert	taxes	() euh c'est vrai que cette affaire de plus-value immobilière était compliquée et qu'elle est () simplifiée à partir du (). en quoi est-ce modifié? Eh bien là où vous ()
42	20040326	journalist	taxes	j'héberge mon fils () dans quelle mesure puis-je faire figurer cette sorte de manque à gagner dans ma déclaration de revenus?
43	20040326	journalist	taxes	() est-ce impossible si oui dans l'affirmative où dois-je le mentionner?
44	20040326	journalist	taxes	j'aimerais qu'on m'explique pourquoi la retenue à la source semble si difficile à mettre en œuvre en france () pourquoi n'essaie-t-on pas au moins de le faire pour les salariés
45	20040326	journalist	taxes	c'est pas un revenu où peut-elle trouver les renseignements pour euh son cas personnel euh également auprès de
46	20040326	journalist	taxes	pendant combien de temps doit-on conserver le double des déclarations de revenus ()
47	20040326	journalist	taxes	() jusqu'à quand l'administration peut-elle procéder à un redressement
49	20040326	journalist	taxes	comment puis-je rentrer dans le rang?
51	20040326	caller	taxes	j'ai une question d'ordre pratique je dirais () en temps normal () on envoie certains justificatifs dois-je les envoyer quand même par courrier parallèlement ou ou qu'est-ce que j'en fais exactement?
54	20040329	journalist	politics	() demain soir la suite question : à quelles conditions les réformes sont-elles possibles dans ce pays?

TABLEAU A.1 – Extractions du corpus EPAC retenues pour analyse

Item	File	Speaker	Theme	Extraction
57	20040329	journalist	politics	pourquoi cette fois-ci le désaveu du pouvoir en place ne s'est-il pas traduit par l'abstention ou le vote aux extrêmes? » on avait parlé de vote protestataire le 21 avril ()
60	20040330	expert	politics	() les gouvernants ont externalisé la question de l'intérêt général sur des instance non démocratiques () qu'il y a un divorce qui est pas seulement de la question économique et sociale qui est une question de sens où est le contrat
61	20040330	expert	politics	() où est le politique les politiques pèsent-ils? ()
67	20040330	expert	politics	i(l) y a une personne qui est intervenue des EU en en en quoi le modèle américain ça peut être un modèle pour nous avec les réalité du modèle américain? i(l) ne faut pas uniquement voire les bénéfices des grosses sociétés ()
68	20040330	expert	politics	() i(l) faut pas mélanger tout euh de quoi parle-t-on? ()
71	20040330	expert	politics	() il me disait « schröder fait pareil »//. Schröder fait quoi? // Bah justement Schröder ne semble pas insérer ()
73	20040330	caller	politics	() ou que sais-je encore ()
75	20040402	journalist	internat	quand la slovaquie va-t-elle abandonner la couronne pour l'euro?
76	20040402	journalist	internat	() l'euro c'est pour quand? ()
77	20040402	expert	internat	() () le problème en fait i(l) faut peut être le poser différemment c'est à dire : si les tchèques ne s() et les slovaques ne s'étaient pas séparés dans quel type d'état auraient-ils vécu? Et la question () en fait la réponse à cette question elle est pas claire du tout
78	20040402	journalist	internat	à quelle échéance l'euro devrait-il être adopté en slovaquie?
81	20040402	journalist	internat	() mais comment là-bas les économies les structures sociales etc euh vont-elles s'adapter
83	20040402	journalist	internat	qu'en pense notre confrère ivan stefunko?
84	20040402	expert	internat	() () moi je soutiens ce qui a été dit par euh le journaliste slovaque : () la slovaquie demain sera l() plus grand producteur () mais est-ce que s() comment comment un pays peut-il se développer seulement avec une seule économie?
85	20040402	expert	internat	() () () une fois que les pays auront le droit de () que ses habitants () auront le droit de circuler () où iront ces roms? Resteront-ils en slovaquie ou non?
90	20040804	expert	health	() et qu'est-ce qu'on fait pour elles?
96	20040805	caller	health	euh comment dirais-je

TABLEAU A.1 – Extractions du corpus EPAC retenues pour analyse

Item	File	Speaker	Theme	Extraction
97	20040805	caller	health	alors je demande à que choisir : comment se fait-il qu'ils ont laissé passer une chose une aberration pareille
105	20040805	journalist	health	et la prévention ça repose sur quels principes essentiels qu'i(l) faudrait qu'on retienne ce soir euh professeur patrick tounian?
106	20040805	caller	health	donc je suis diététicienne. Moi je me tourne vers l'ex() les experts de l'AFSSA pour dire euh à à quand euh encourageront-ils le gouvernement à rédiger une loi pour encadrer la qualité des menus proposés en restauration scolaire?
107	20040805	caller	health	() Euh l'hygiène est strictement encadrée et contrôlée () alors que les enquêtes () montrent que la qualité () se dégrade () () donc à quand une loi pour que les menus soient enfin équilibrés?
109	20040805	journalist	health	au coeur de tout cela il y a ce que les spécialistes () appellent () l'obésité et notamment celle qui touche les jeunes comment en est-on arrivé là? ()
111	20040805	caller	health	comment on comment on définit l'obésité
113	20040806	caller	economy	combien un pays producteur disons dont on prélève les ressources touche par baril?
114	20040806	caller	economy	combien coûte le transport?
115	20040806	caller	economy	combien coûte le raffinage?
118	20040806	expert	economy	Il faut bien comprendre que les questions d'énergie euh se comment dirais-je ne se ne se ne se traitent pas...
119	20040806	journalist	economy	mais olivier rech quand seront épuisées les réserves de pétrole du monde?
122	20040806	journalist	economy	() () le brut à new york () peut atteindre () qu'en pensez vous?
124	20040806	journalist	economy	() // là j'ai rien compris hein j'ai rien compris à ce que vous avez dit là qu'est-ce que ça veut dire ça? // alors ça veut dire ()
125	20040806	journalist	economy	c'est vrai que tout récemment le président bush a déclaré qu'il était pas question de toucher () ces réserves c'est quoi exactement? Réponse de ()
127	20040809	journalist	spies	() la question posée ce soir () () comment agissent ces hommes de l'ombre peu habitués à révéler leurs secrets? Pour répondre je l'espère à vos questions... ()
128	20040809	journalist	spies	pour euh généraliser le débat qu'est-ce que vous entendez ()
129	20040809	journalist	spies	() d'ailleurs barbouze c'est assez vague qu'est-ce que vous entendez par barbouze exactement? Mercenaire?
130	20040809	journalist	spies	() () // claude moniquet comment vous vous expliquez vous ce désintérêt politique des des élus pour les chefs de service? //

TABLEAU A.1 – Extractions du corpus EPAC retenues pour analyse

Item	File	Speaker	Theme	Extraction
132	20040809	caller	spies	() () voilà ma question maintenant : euh qu'est-ce qu'il sait cet homme pour que dans sa situation les plus hautes autorités de l'état le nomment préfet hors cadre?
133	20040809	caller	spies	qu'est-ce qu'il sait?
135	20040809	journalist	spies	() // et à quoi servaient les caisses noires // alors les caisses noires
137	20040812	caller	religion	pourquoi la république française qui au nom de () pourquoi cette république française investit aujourd'hui énormément de moyens humains et donc financiers pour la venue du pape à lourdes alors que celui-ci vient non pas en tant que chef d'état mais en tant que pèlerin? Je dois dire que cette incohérence m'interpelle en tant que militant chrétien
141	20040812	journalist	religion	avant de prendre les premier appels () on va quand même détailler les les grandes lignes de ce sondage () que représente pour vous avant tout la ville de lourdes? Un lieu de pèlerinage à soixante-dix-sept pour cent
143	20040812	caller	religion	dans le sondage qui paraîtra demain dans la croix a-t-on demandé aux sondés s'ils ne souhaiteraient pas que le pape démissionne? () sinon pourquoi la croix ne pose-t-elle pas cette question?
144	20040812	journalist	religion	comment fait () comment fait-on dans le catholicisme pour justifier ce culte rendu aux saints que nous protestants nous refusons absolument? ()
147	20040812	expert	religion	() () () qu'est-ce que c'est que la sexualité? ()
148	20040812	expert	religion	() comment nous structure-t-elle? Est-ce que c'est ()
154	20040813	journalist	internat	() () où où en sommes nous précisément ce soir alors?
156	20040813	expert	internat	() () on parle beaucoup de moqtada sadr on a raison () mais ce que je constate on n'a presque pas parlé du programme de moqtada sadr qu'est-ce qu'il cherche moqtada sadr? ()
157	20040813	expert	internat	() qu'est-ce qu'il veut impliquer aux gens de la société irakien? ()
158	20040813	expert	internat	() pourquoi il prêche pour une rupture totale à l'intérieur de l'irak avec les autres et à l'extérieur? ()
161	20040813	caller	internat	() que pensent-ils et est-ce qu'ils font continuent à faire aveuglément confiance au gouvernement de georges bush?
165	20040813	journalist	internat	// que pouvez-vous répondre à cette question? //
166	20040813	caller	internat	(dans) quelle mesure on peut imposer la démocratie occidentale à des peuples qui sont gérés sur un système tribal depuis très longtemps?

TABLEAU A.1 – Extractions du corpus EPAC retenues pour analyse

Item	File	Speaker	Theme	Extraction
167	20040813	journalist	internat	pour terminer avec vous hosham daoud euh très très court () et pourtant (il) y a une tension qui dure euh elle est due à quoi alors si moqtada al sadr n'a pas beaucoup d'influence?
168	20040813	caller	internat	comment réagissent les américains désormais par rapport à ce qu'on vient de dire c'est à dire euh pensent-ils rester en Irak et ne pas être remplacés par les Nations Unies?
170	20040813	caller	internat	donc euh dans ce cas là qu'est-ce qu'il reste de la démocratie?
178	20040813	expert	internat	() à quel point les irakiens se décideront-ils qu'ils ont ras le bol de cette insurrection? S() pour moi c'est la question ()
181	20040813	caller	internat	() il y a quand même des analystes compétents des () et euh quelle part i(l)s ont dans le processus décisionnel? Est-ce que autour des états majors on en retrouve? ()
189	20040816	caller	economy	() pourquoi est-ce que les producteurs ne pourraient pas créer eux mêmes des filières de distribution où les produits seraient respectés?
190	20040816	caller	economy	() donc pourquoi est-ce qu'on m() n'en ferait pas autant pour les fruits et les légumes?
195	20040816	expert	economy	() comment la qualité détermine-t-elle le prix des fruits et légumes? ()
199	20040817	expert	politics	() parce que finalement où est l'ascenc() où est l'ascenseur social dans cette histoire?
200	20040817	expert	politics	où est l'égalité des chances?
201	20040817	expert	politics	donc il faut où est () où en est le mérite républicain?
203	20040817	expert	politics	() d'ailleurs la question qu'on peut se poser c'est pourquoi au fond si Nicolas Sarkozy est si populaire et efficace () pourquoi au fond le président de la République n'a pas souhaité le nommer premier ministre et lui confier de plus larges responsabilités?
204	20040817	expert	politics	pourquoi il souhaite même les renier?
205	20040817	expert	politics	à quel code secret au sein du gaullisme euh a-t-il manqué?
208	20040817	expert	politics	et la question que je voudrais poser à Brice s'il peut répondre ce serait important je crois c'est de savoir pour quelle raison compte tenu de la campagne intense qu'a menée () pourquoi est-ce qu'il n'a pas pris le risque d'annoncer qu'il était candidat ()
209	20040817	expert	politics	() si par hasard il n'est pas candidat () dans quel intérêt ceci serait-il fait?
210	20040817	expert	politics	pourquoi devrait-il bousculer son calendrier?

TABLEAU A.1 – Extractions du corpus EPAC retenues pour analyse

Item	File	Speaker	Theme	Extraction
214	20040817	expert	politics	la vraie question c'est euh aujourd'hui où va nous conduire cette situation de conflit interne à la majorité?
217	20040817	journalist	politics	pourquoi donc nicolas sarkozy s'est-il exprimé sur le départ de zinedine zidane? Très brèvement hélène jouan?
218	20040817	expert	politics	on lui a confié les con() les fonctions de numéro deux du gouvernement avec celles de ministre de l'économie () qu'est-ce que cela veut dire au-delà du sigle du ministère? Ça veut dire qu'on lui a confié ()
220	20040817	expert	politics	() alors naturellement ça nécessiterait d'assez longues réponses : à quoi doit servir une par() une formation politique dans notre société? ()
226	20040817	journalist	politics	vous nous appelez d'où bernard?
228	20040818	expert	economy	on va prendre un autre locataire parce que derrière (il) y en a dix qui sont prêts à donner toutes () pourquoi?
229	20040818	expert	economy	() ce qui majore les prix () et qui pose un problème de fond euh sur l'évolution de nos villes capitales et de nos grandes villes : à qui appartient la ville demain? Et c'est pourquoi euh la ville de Paris a comme souci majeur de ()
230	20040818	journalist	economy	() () la question qui revient souvent et c'est pour ça que je vous je vous la pose aussi : quel quel rôle quelle influence peuvent avoir les pouvoirs publics sur euh sur euh sur le prix de de l'immobilier? //
231	20040818	journalist	economy	nous entendons dire que les français acquièrent de plus en plus leur logement. Comment font les français? Puisque pour notre part nous n'arrivons pas ()
232	20040818	journalist	economy	() bien qu'ayant de bons revenus mon mari et moi. À qui prêtent les banques ou les organismes financiers?
233	20040818	journalist	economy	quel profil doit-on avoir pour se loger David Rodrigues?
234	20040818	caller	economy	() donc pourquoi les frais d'agence ne sont-ils pas réglés au-delà d'un certain montant et est-il envisagé de le faire?
235	20040818	journalist	economy	// mais alors comment se fait-il que euh tout ce que vous mentionnez là ne soit pas ne soit pas si évident sur le terrain qu'on n'en ressent pas les effets? //
236	20040819	expert	culture	c'est le danger de toute situation de monopole de toute façon dans quelle mesure i(ls) vont continuer à avoir le désir d'aller dans des démarches innovantes euh novatrices de cinéma? (en)fin c'est la c'est la vraie question
238	20040819	journalist	culture	mais pourquoi ce sentiment là prédomine? ()
240	20040819	expert	culture	à l'intérieur de ce chiffre là : comment se comporte le cinéma d'auteurs? Euh est-ce qu'il va vraiment bien ou est-ce que finalement ()()

TABLEAU A.1 – Extractions du corpus EPAC retenues pour analyse

Item	File	Speaker	Theme	Extraction
242	20040819	expert	culture	comment on peut faire
244	20040819	expert	culture	() si ces cinémas deviennent euh exsangues () où va-t-on euh si () où va-t-on diffuser des films d'auteurs?
245	20040819	expert	culture	() () () si demain ce cinéma disparaît demain à clermont-ferrand il n'y a plus de cinéma d'art et d'essai qu'est-ce qu'on fait? On prend la voiture on va à paris on va à lyon () ça paraît un peu inquiétant
254	20040820	expert	history	le général de gaulle va () pourquoi y va-t-il? Il y va pour faire acclamer ()
263	20040823	caller	politics	pour moi la question c'est euh que veut réellement le peuple corse? Est-ce que c'est seulement les nationalistes qu'on doit entendre?
264	20040823	expert	politics	() () () alors aujourd'hui que dit-on? On fait monter la pression ()
266	20040823	Journalist 2	politics	lorsqu'on voit ces graffitis () comment ne se sent-on pas concernés par ces euh atteintes à la dignité humaine?
273	20040823	journalist	politics	() () qu'en pense michel colaccioni?
279	20040824	journalist	economy	pourquoi ne prend-on pas en compte ce que nous dépensons régulièrement pour vivre ()
280	20040824	journalist	economy	() et pourquoi ne fait-on pas notre baromètre pour les achats inhabituels?
283	20040824	journalist	economy	pourquoi d'après l'étude de l'insee certaines catégories sociales ont vu une augmentation de leurs salaires? ()
285	20040824	expert	economy	pourquoi d'après l'étude de l'insee certaines catégories sociales ont vu une augmentation de leurs salaires? ()
289	20040824	caller	economy	() ()? () pourquoi l'état ne respecte pas au moins euh l() l'évolution du coût de la vie?
291	20040824	journalist	economy	comment vos interlocuteurs voient-ils l'évolution du pouvoir d'achat à court moyen et long terme ()
295	20040824	expert	economy	() les entreprises en fait ne pouvant pas ou () ont été hyper sélectives et qu'est-ce qu'elles ont fait? Elles ont augmenté ()
300	20040825	journalist	politics	pourquoi dans les débats sur la constitution on nous parle toujours d'économie et d'un volet social alors que dans une constitution logiquement il n'y a que des éléments juridiques
301	20040825	expert	politics	() ce que je veux te dire c'est que si nous ne donnons pas une réponse politique pourquoi ce syndicaliste ()
304	20040825	journalist	politics	() comment vous e() vous expliquez ce décalage?
306	20040825	expert	politics	le procédé actuel n'est pas démocratique mais surtout les gens i(ls) veulent l'europe pour quoi? Pour avoir la paix ()

TABLEAU A.1 – Extractions du corpus EPAC retenues pour analyse

Item	File	Speaker	Theme	Extraction
309	20040825	caller	politics	() () que peuvent me répondre vos intervenants à cela s'il vous plait?
317	20040826	journalist	culture	Pouvez-vous demander à anna gavalda : que lit-on quand on a fini son roman ensemble c'est tout?
319	20040826	journalist	culture	et surtout que fait-on quand on a fini d'écrire un tel livre?
320	20040826	expert	culture	que lit-on? On peut lire euh ()
321	20040826	expert	culture	() quand vous rencontrez un libraire i(l) vous parle de quoi? i(l) vous parle des livres de ()
322	20040826	journalist	culture	plus de six cents livres ça leur paraît énorme comment est-ce possible?
323	20040826	journalist	culture	comment travaillez vous?
325	20040826	journalist	culture	() comment anna gavalda a-t-elle pu avoir autant de succès alors que ()
327	20040826	journalist	culture	() comment vous gérez ça?
331	20040826	journalist	culture	des concours de k() // que // de de quels concours parlez-vous?
333	20040826	journalist	culture	question récurrente donc qui est revenue souvent : quelle réponse pouvez-vous approter?
337	20040826	expert	culture	() mais pourquoi écrit-on tous ces livres ()
338	20040826	expert	culture	() et pourquoi les lit-on? Il y a un élément totalement irrationnel là-dedans ()
340	20040826	journalist	culture	pourquoi euh l() les éditeurs à la base ne font pas eux une sélection plus euh plus sa() plus rigoureuse? ()
345	20040830	expert	society	(il) y a des plannings qui sont avec des trous où vont-ils les enfants? Au CDI c'est pas possible c'est pas ()
346	20040830	caller	society	et je trouve que c'est toujours très difficile pour un enfant de euh bah de revenir euh à la maison avec un zéro en dictée. Quelle image il a de l'éducation de lui-même de de toutes ces images que ça le () ça lui renvoie
348	20040830	journalist	society	peut être cet argent n'est pas bien réparti // mais comment est-il employé ventilé? Oui
349	20040830	caller	society	() la question que je pose c'est quand ()
351	20040830	caller	society	() où est la parole entre l'adulte est l'enfant? ()
352	20040830	caller	society	() quelle valeur on lui donne? ()
354	20040830	caller	society	() et ce discours que moi je tiens dans ma classe euh eh bien les enfants i(l)s ont l'air de dire « de quoi qu'elle cause? »
355	20040830	caller	society	() l'enseignement des langues vivantes en en primaire () comment ça se passe? Euh les enseignants sont-ils formés? ()
359	20040831	expert	internat	() () mais d'un autre côté il a donné une arme extraordinaire à l'administration parce que maintenant comment peut-il s'y opposer?

TABLEAU A.1 – Extractions du corpus EPAC retenues pour analyse

Item	File	Speaker	Theme	Extraction
362	20040831	journalist	internat	() à quelle catégorie par exemple économique ou sociale appartiennent-ils?
366	20040831	expert	internat	l'exemple le plus spectaculaire c'était richard nixon il y a quoi? En soixante-huit qui a dit ()
367	20040831	caller	internat	() alors c'est une question que je pose : comment ça se fait que euh notre ami bush fasse sa convention () sur un terrain qui paraît euh complètement démocrate même si le maire est complètement républicain? () est-ce que () () est-ce que?
369	20040902	caller	society	et puis euh qu'est-ce qu'on entend par dictée par exemple? La dictée euh je sais pas si () a mu() de multiples formes ()
370	20040902	caller	society	ma question euh euh ce sera sur les lycées professionnels euh qu'en est-il de leur avenir? () ()
371	20040902	caller	society	() () qu'en est-il de l'avenir des lycées professionnels avec une forte demande de des entreprises euh qui sont prêtes à recruter beaucoup de jeunes? ()
375	20040902	journalist	society	() comment le ministre et la hiérarchie () peuvent-ils tout simplement nous aider à retrouver et exercer à bon escient cette autorité sans laquelle ()
377	20040902	expert	society	() (il) y a des choses à faire pour aider les enseignants euh à à à redonner des repères au fond aux enfants () ces repères c'est quoi? c'est d'abord ()
379	20040906	journalist	internat	pourquoi la tchéchénie si pauvre semble-t-elle avoir quand même une telle importance politique pour les russes ()
380	20040906	journalist	internat	() et quoi ont-ils choisi de se comporter ainsi vis à vis des révoltés nationalistes?
381	20040906	expert	internat	() alors la real politique comment elle s'est traduite? Elle s'est traduite euh parallèlement à de très timides ()
383	20040906	expert	internat	() () () pour la première fois (il) y a quand même des questions posées : pourquoi le président n'a pas parlé de la tchéchénie dans son discours à la nation samedi? ()
385	20040906	expert	internat	() et puis cette question toute simple : pourquoi vit-on dans un pays où l'état n'assure pas la sécurité de ses citoyens?
390	20040906	expert	internat	() la question de l'auditrice était très intéressante pourquoi la tchéchénie qui est si pauvre représente tant d'importance pour la russie? En effet économiquement ()
391	20040906	caller	internat	pour quelle raison euh les chefs d'état européens restent assez silencieux face à ()()?
392	20040906	caller	internat	() pourquoi les chefs d'état européens ne réagissent pas à cette façon de de faire?

TABLEAU A.1 – Extractions du corpus EPAC retenues pour analyse

Item	File	Speaker	Theme	Extraction
393	20040906	journalist	internat	le le poids euh culturel les pesanteurs politiques euh le fait que () () en quoi maintenant ça joue dans la politique actuelle ou du moins ça pèse?
394	20040906	journalist	internat	disons la stratégie de poutine et du pouvoir russe : que veut-il?
395	20040906	expert	internat	c'est l'un des grands reproches () où étaient les ministres? ()
396	20040906	expert	internat	() où était le président de l'ossétie du nord euh que le commando avait appelé pour les négociations? ()
397	20040906	expert	internat	() où était le président de l'ingouchie que les le commando avait appelé pour les négociations? ()
398	20040906	expert	internat	() où étaient les hautes responsables russes? ()
399	20040906	expert	internat	() où était le ministre de l'intérieur? () personne n'est venu
404	20040907	caller	politics	j'aimerais savoir comment ça peut évoluer
405	20040907	expert	politics	() alors si c'est une compilation de traités on a envie de dire pourquoi vient-on nous ennuyer avec ça?
408	20040907	expert	politics	les partisans du oui nous disent mais non non mais ça ne change rien alors à quoi ça sert?
410	20040907	expert	politics	() et pourquoi il faut voter non? Bah parce que i(l) va falloir faire immédiatement quelque chose et rebondir immédiatement ()
411	20040907	expert	politics	() et comment peut on rebondir? Eh bien avec un nombre restreint de pays ()
412	20040907	caller	politics	comment peut-on aujourd'hui espérer que l'europe devienne une europe puissance aujourd'hui alors que la constitution ()
413	20040907	expert	politics	si c'est pas une constitution alors qu'est-ce qu'on fait? ()
415	20040907	caller	politics	si le danemark qui vote en premier vote non euh à quoi ça sert que les autres aillent voter ensuite?
416	20040907	caller	politics	() est-ce que cet équilibre existe dans la constitution européenne et sinon pourquoi est-ce qu'il n'existe pas?
423	20040908	journalist	culture	où vous voit-on ()
424	20040908	journalist	culture	quand vous a-t-on vu pour la dernière fois jouer à la télévision?
429	20040908	journalist	culture	Vous-même qui trouvez-vous dans ces textes?
431	20040908	journalist	culture	paris combo sous les étoiles exactement paris combo représenté cette nuit par belle du berry par mano et par françois il manque qui? // bah il manque ()
432	20040908	journalist	culture	de quoi parle Belle dans ses textes?
433	20040908	journalist	culture	les prochains concerts c'est où ()
436	20040908	journalist	culture	que vous réclame-t-on sinon on vous met en taule?
437	20040908	journalist	culture	comment sont-ils au courant?
438	20040908	journalist	culture	il vous remercie pour quoi? ()

Item	File	Speaker	Theme	Extraction
444	20040908	journalist	culture	maintenant « aquarium » de quoi est-il voilà
445	20040908	journalist	culture	où peut-on vous écrire?
454	20040908	journalist	culture	et je vous demande « paris combo c'est quel style? »
455	20040908	expert	culture	ta mémoire se situe à quel niveau?
459	20040909	caller	society	pourquoi ne publierait-on pas euh le taux de réussite des auto-é() euh des auto-écoles d'une manière libre sachant que ()
460	20040909	caller	society	pourquoi par exemple n'apprendrons () n'apprendrait-on pas à rattraper une situation d'urgence sur circuit? Est-ce que c'est un problème de coût?
461	20040909	journalist	society	() comment ça se passe? ()
472	20040910	caller	health	() () pourquoi est-ce que les autorités les pouvoirs publics ne euh demandent pas des dépistages beaucoup plus généralisés de cette maladie dont le traitement est simple? Par contre une personne non traitée ()
473	20040910	caller	health	() () qu'en est-il euh sur l'avenir de euh de cette maladie? Peut-on être optimistes?
474	20040910	journalist	health	comment vous médecin euh comment vous vous travaillez avec l'agence de la fondation de la recherche? ()
475	20040910	journalist	health	() en quoi est-ce que cet argent vous aide?
477	20040910	expert	health	() est-ce que ce sont des leucémies de bons pronostics? De quels euh traitements dispose-t-on actuellement? Et est-ce qu'on a une connaissance aussi sur ()
481	20040910	expert	health	() () pourquoi (il) y a un pic à quatre ans? Personne ne peut répondre à votre question
482	20040910	caller	health	mon fils a déclaré une thrombopénie autoimmune périphérique à l'âge de treize ans où en sont les recherches?
483	20040917	journalist	culture	bon patrimoine euh euh très bien mais pourquoi ces objets sont-ils pas tout le temps visibles?
489	20040917	Journalist 2	culture	() comment est-ce que l'on décide qu'un objet technique entre au patrimoine?
490	20040917	Journalist 2	culture	comment ça se passe dans ce cas concret des phares? Attendez-vous que ()?
495	20040917	journalist	culture	que nous reste-t-il des outils ou les instruments utilisés à cette époque par ces savants?

TABLEAU A.1 – Extractions du corpus EPAC retenues pour analyse

Annexe B

Corpus EPAC : Statistiques

B.0.1 Répartitions des interrogatives par thématique - Sous-corpus complet

	Antéposition simple	Antéposition + ESK	Antéposition et inversion	In situ
Histoire	-	-	1	-
Religion	1	1	4	-
Espionnage	1	4	2	-
Impôts	1	1	10	-
Société	6	1	8	1
Santé	6	3	10	3
Économie	4	4	16	2
Culture	7	2	22	5
International	9	3	24	3
Politique	10	5	27	4

TABLEAU B.1 – Interrogatives partielles dans EPAC par type et par thématique (tout)

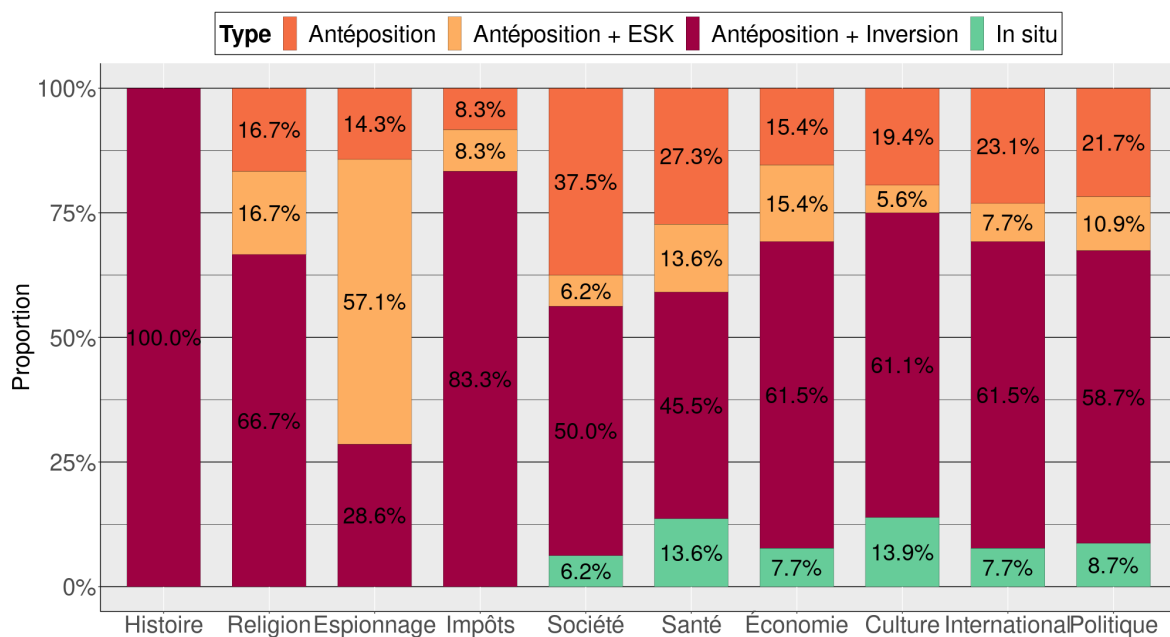


FIGURE B.1 – Interrogatives partielles dans EPAC par type et par thématique (tout)

B.0.2 Modèle *brms* et distributions postérieures pour un effet de THÉMATIQUE (4.1.3.1)**Family :** categorical**Links :** muFA = logit; muFESK = logit; muIS = logit**Formula :** construction ~ (int_pol + cult_pol + eco_pol + hlth_pol + soc_pol) + (1 | FILE)**Data :** eadat_thm (Number of observations : 185)**Samples :** 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-1.22	0.36	-2.04	-0.61	1.00	5350	4721
muFESK_Intercept	-3.16	1.06	-5.77	-1.62	1.00	2964	2851
muIS_Intercept	-2.12	0.39	-2.98	-1.46	1.00	6720	4656
muFA_int_pol	0.10	0.98	-1.82	2.12	1.00	4563	5496
muFA_cult_pol	0.21	1.00	-1.61	2.37	1.00	4603	4584
muFA_eco_pol	-0.42	1.09	-2.63	1.75	1.00	5524	5939
muFA_hlth_pol	0.71	1.07	-1.34	2.93	1.00	5011	5051
muFA_soc_pol	0.73	1.12	-1.47	2.99	1.00	4848	5474
muFESK_int_pol	-1.09	2.43	-6.60	3.41	1.00	4026	3672
muFESK_cult_pol	-0.21	2.22	-4.67	4.37	1.00	3690	4265
muFESK_eco_pol	1.36	2.19	-2.56	6.30	1.00	3874	3616
muFESK_hlth_pol	1.48	2.35	-2.85	6.72	1.00	3907	4004
muFESK_soc_pol	-0.78	2.81	-6.68	4.62	1.00	3762	3574
muIS_int_pol	-0.16	1.04	-2.24	1.92	1.00	6231	6229
muIS_cult_pol	0.31	0.98	-1.73	2.20	1.00	5649	4687
muIS_eco_pol	-0.36	1.19	-2.81	1.86	1.00	6177	6125
muIS_hlth_pol	0.63	1.11	-1.62	2.80	1.00	6454	5612
muIS_soc_pol	-0.59	1.56	-4.15	2.09	1.00	7824	5940

TABLEAU B.2 – Résumé du modèle *brms* (EPAC, effet THÉMATIQUE)

Effect	$\hat{\beta}$	$P(\beta > 0)$	$P(\beta < 0)$	min	max	l95	h95	l90	h90	l85	h85
F_International	0.10	0.53	0.47	-4.40	5.83	-1.82	2.12	-1.46	1.74	-1.24	1.48
F_Culture	0.21	0.57	0.43	-4.31	5.92	-1.61	2.37	-1.31	1.96	-1.14	1.69
F_Économie	-0.42	0.34	0.66	-5.51	4.84	-2.63	1.75	-2.21	1.34	-1.97	1.09
F_Santé	0.71	0.76	0.24	-4.61	6.34	-1.34	2.93	-0.97	2.52	-0.75	2.24
F_Société	0.73	0.76	0.24	-6.65	6.04	-1.47	2.99	-1.07	2.54	-0.82	2.30
FESK_International	-1.09	0.31	0.69	-14.06	9.41	-6.60	3.41	-5.20	2.62	-4.58	2.11
FESK_Culture	-0.21	0.45	0.55	-10.80	9.43	-4.67	4.37	-3.74	3.43	-3.26	2.87
FESK_Économie	1.36	0.75	0.25	-9.13	14.89	-2.56	6.30	-1.83	5.12	-1.42	4.46
FESK_Santé	1.48	0.75	0.25	-9.70	13.62	-2.85	6.72	-2.05	5.52	-1.51	4.89
FESK_Société	-0.78	0.39	0.61	-15.82	14.64	-6.68	4.62	-5.44	3.51	-4.69	2.88
IS_International	-0.16	0.43	0.57	-5.95	4.56	-2.24	1.92	-1.90	1.52	-1.66	1.26
IS_Culture	0.31	0.65	0.35	-4.77	4.41	-1.73	2.20	-1.32	1.85	-1.07	1.64
IS_Économie	-0.36	0.39	0.61	-7.98	4.81	-2.81	1.86	-2.34	1.48	-2.05	1.25
IS_Santé	0.63	0.73	0.27	-5.75	6.01	-1.62	2.80	-1.18	2.39	-0.92	2.15
IS_Société	-0.59	0.37	0.63	-12.53	6.20	-4.15	2.09	-3.35	1.65	-2.85	1.41

TABLEAU B.3 – Distributions postérieures pour le modèle brms (EPAC, effet THÉMATIQUE)

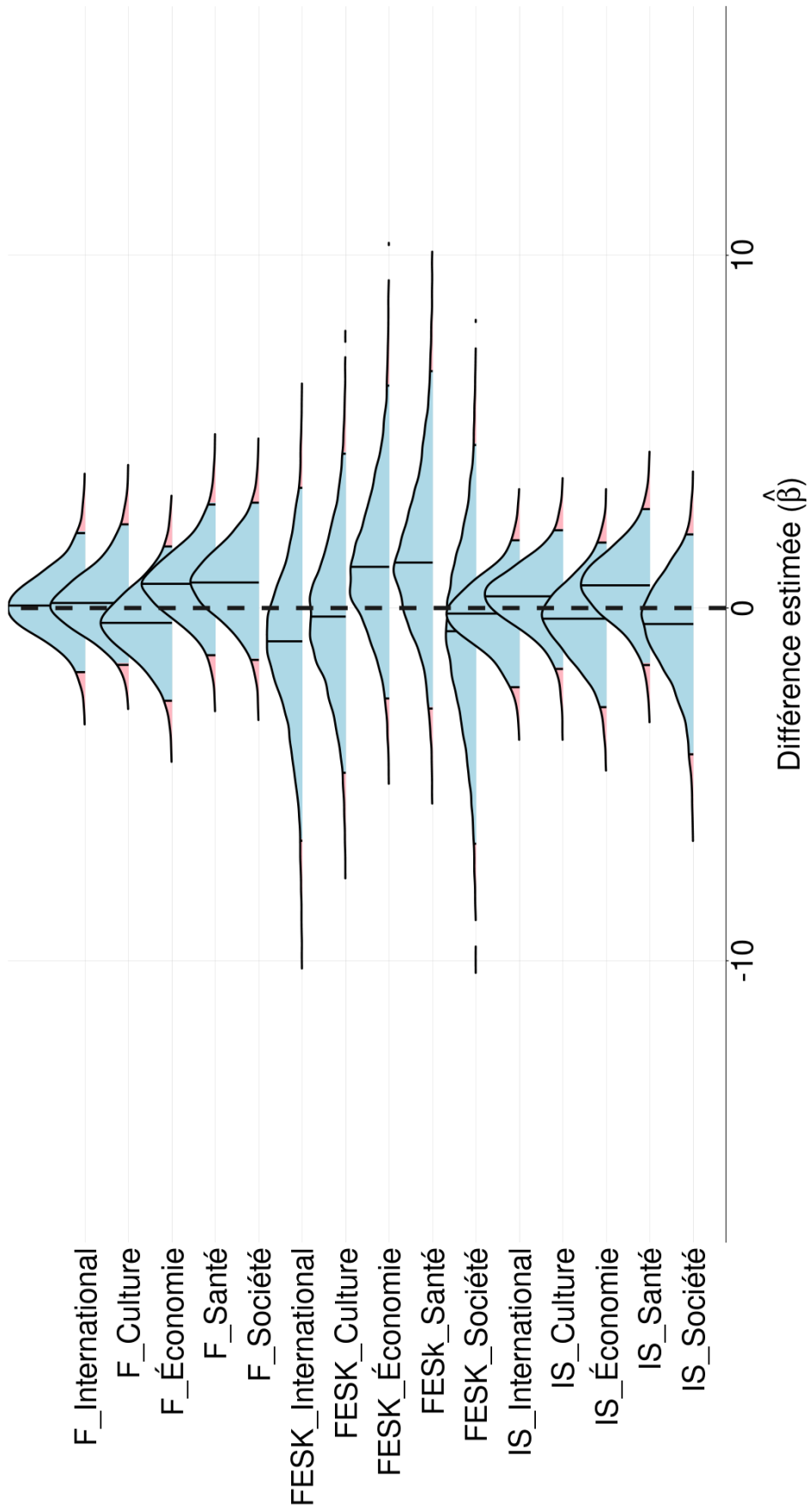


FIGURE B.2 – Distributions postérieures pour le modèle brms d'analyse de EPAC, effet de THÉMATIQUE

B.0.3 Modèle *brms* et distributions postérieures pour un effet de RÔLE (4.1.3.2)**Family :** categorical**Links :** muFA = logit; muFESK = logit; muIS = logit**Formula :** construction ~ (exp_jour + call_jour) + (1 | ID_file)**Data :** eadat_spk (Number of observations : 208)**Samples :** 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-1.19	0.24	-1.70	-0.76	1.00	5432	3725
muFESK_Intercept	-2.29	0.51	-3.46	-1.48	1.00	2178	2415
muIS_Intercept	-3.99	1.34	-7.18	-2.12	1.00	2807	2421
muFA_exp_jour	0.78	0.46	-0.11	1.69	1.00	5696	4892
muFA_call_jour	1.52	0.49	0.56	2.50	1.00	5703	4874
muFESK_exp_jour	1.35	0.73	-0.04	2.84	1.00	4802	4166
muFESK_call_jour	1.82	0.73	0.46	3.32	1.00	5166	4536
muIS_exp_jour	0.89	0.59	-0.23	2.09	1.00	6648	3846
muIS_call_jour	-8.56	5.70	-22.22	-0.97	1.00	2846	2648

TABLEAU B.4 – Résumé du modèle *brms* (EPAC, effet RÔLE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	195	h95	190	h90	l85	h85
F_Expert.e	0.78	0.96	0.04	-0.95	2.58	-0.11	1.69	0.02	1.55	0.11	1.45
F_Auditeur.trice	1.52	1.00	0.00	-0.55	3.37	0.56	2.50	0.72	2.32	0.82	2.22
FESK_Expert.e	1.35	0.97	0.03	-1.47	5.86	-0.04	2.84	0.19	2.61	0.33	2.43
FESK_Auditeur.trice	1.82	0.99	0.01	-0.76	5.70	0.46	3.32	0.69	3.06	0.82	2.89
IS_Expert.e	0.89	0.94	0.06	-1.51	3.34	-0.23	2.09	-0.06	1.89	0.06	1.76
IS_Auditeur.trice	-8.56	0.00	1.00	-40.52	1.37	-22.22	-0.97	-19.61	-1.51	-17.94	-1.94

TABLEAU B.5 – Distributions postérieures pour le modèle *brms* d'analyse pour EPAC, effet de RÔLE

B.0.4 Corrélations entre échelles - Détails (4.1.4)

	Informalisation	Proximité Destinataire	Reformulation	Auto-Répétition	Répétition autre	Fin argumentaire	Réponse auto-fournie	Opinion implicite	Prise de posture
Informalisation	1.00	0.24	0.01	0.13	0.14	-0.03	0.16	-0.14	0.15
Proximité Destinataire	0.24	1.00	-0.07	0.14	0.03	-0.16	0.10	-0.18	0.18
Reformulation	0.01	-0.07	1.00	0.25	0.11	-0.04	-0.02	0.17	-0.04
Auto-Répétition	0.13	0.14	0.25	1.00	-0.10	0.09	0.16	0.16	0.23
Répétition autre	0.14	0.03	0.11	-0.10	1.00	-0.05	-0.01	-0.07	-0.15
Fin argumentaire	-0.03	-0.16	-0.04	0.09	-0.05	1.00	0.15	0.13	0.17
Réponse auto-fournie	0.16	0.10	-0.02	0.16	-0.01	0.15	1.00	-0.17	0.44
Opinion implicite	-0.14	-0.18	0.17	0.16	-0.07	0.13	-0.17	1.00	0.07
Prise de posture	0.15	0.18	-0.04	0.23	-0.15	0.17	0.44	0.07	1.00

TABLEAU B.6 – Coefficients de corrélations entre critères interactionnels, méthode Spearman (EPAC)

	Informalisation	Proximité Destinataire	Reformulation	Auto-Répétition	Répétition autre	Fin argumentaire	Réponse auto-fournie	Opinion implicite	Prise de posture
Informalisation		0.00	0.83	0.06	0.05	0.62	0.02	0.05	0.03
Proximité Destinataire	0.00		0.32	0.04	0.68	0.02	0.16	0.01	0.01
Reformulation	0.83	0.32		0.00	0.11	0.56	0.72	0.02	0.58
Auto-Répétition	0.06	0.04	0.00		0.14	0.17	0.02	0.02	0.00
Répétition autre	0.05	0.68	0.11	0.14		0.43	0.92	0.33	0.03
Fin argumentaire	0.62	0.02	0.56	0.17	0.43		0.03	0.06	0.01
Réponse auto-fournie	0.02	0.16	0.72	0.02	0.92	0.03		0.01	0.00
Opinion implicite	0.05	0.01	0.02	0.02	0.33	0.06	0.01		0.31
Prise de posture	0.03	0.01	0.58	0.00	0.03	0.01	0.00	0.31	

TABLEAU B.7 – P-values pour les corrélations entre critères (EPAC)

B.0.5 Modèle *brms* et distributions postérieures pour des effets de CRITÈRES INTERACTIONNELS (4.1.4)

Family : categorical
Links : muFA = logit; muFESK = logit; muIS = logit
Formula : construction ~ inf2 + (1 | ID_file)
Data : eadat_lit (Number of observations : 211)
Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-1.12	0.29	-1.73	-0.59	1.00	3980	3255
muFESK_Intercept	-2.04	0.52	-3.22	-1.22	1.00	3350	3623
muIS_Intercept	-6.49	2.70	-13.22	-2.93	1.00	4049	3037
muFA_inf2	2.92	0.46	2.03	3.85	1.00	7485	4034
muFESK_inf2	2.97	0.59	1.87	4.18	1.00	6839	4162
muIS_inf2	10.49	4.28	4.79	21.15	1.00	4068	2851

TABLEAU B.8 – Résumé du modèle brms (EPAC, effet MARQUE D’INFORMALITÉ)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	195	h95	190	h90	185	h85
Antéposition	2.92	1.00	0.00	1.27	4.80	2.03	3.85	2.19	3.68	2.28	3.58
Antéposition & ESK	2.97	1.00	0.00	0.93	5.67	1.87	4.18	2.03	3.97	2.14	3.83
In situ	10.49	1.00	0.00	2.89	36.34	4.79	21.15	5.27	18.64	5.61	17.38

TABLEAU B.9 – Distributions postérieures pour le modèle brms (EPAC, effet MARQUE D’INFORMALITÉ)

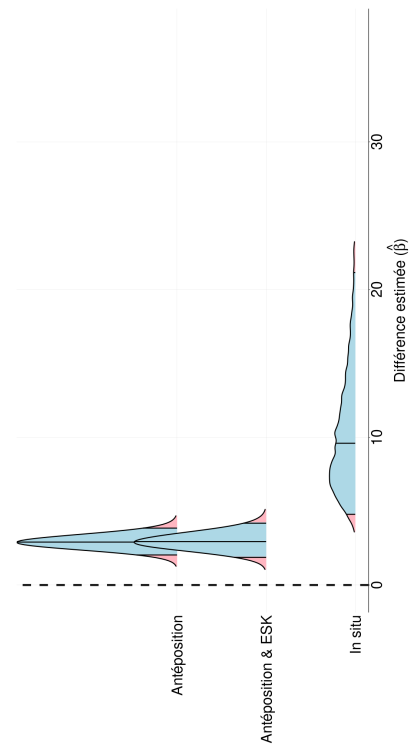


FIGURE B.3 – Distributions postérieures pour le modèle brms (EPAC, effet MARQUE D’INFORMALITÉ)

Family : categorical

Links : muFA = logit; muFESK = logit; muIS = logit

Formula : construction ~ prox2 + (1 | ID_file)

Data : eadat_lit (Number of observations : 211)

Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-1.10	0.25	-1.65	-0.64	1.00	3888	3147
muFESK_Intercept	-2.01	0.45	-3.05	-1.27	1.00	2102	2628
muIS_Intercept	-2.48	0.41	-3.42	-1.79	1.00	3585	2049
muFA_prox2	1.24	0.63	0.01	2.47	1.00	5944	4191
muFESK_prox2	0.91	0.86	-0.88	2.52	1.00	6657	4343
muIS_prox2	3.22	0.69	1.94	4.64	1.00	4448	3193

TABLEAU B.10 – Résumé du modèle brms (EPAC, effet PROXIMITÉ DESTINATAIRE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Antéposition	1.24	0.98	0.02	-0.93	3.83	0.01	2.47	0.22	2.27	0.35	2.14
Antéposition & ESK	0.91	0.86	0.14	-2.45	3.86	-0.88	2.52	-0.56	2.27	-0.34	2.12
In situ	3.22	1.00	0.00	0.79	5.86	1.94	4.64	2.13	4.39	2.27	4.23

TABLEAU B.11 – Distributions postérieures pour le modèle brms (EPAC, effet PROXIMITÉ DESTINATAIRE)

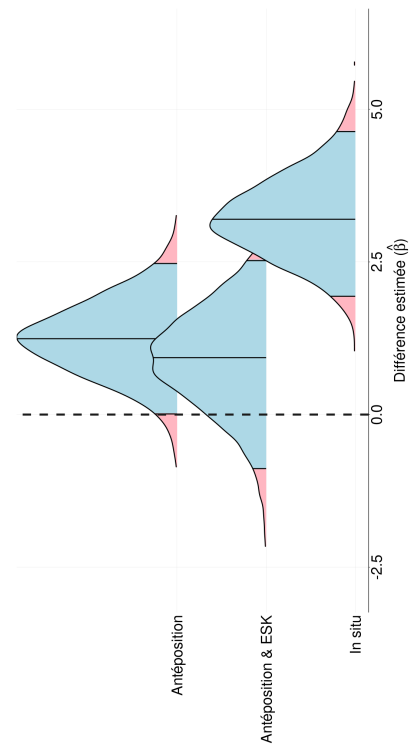


FIGURE B.4 – Distributions postérieures pour le modèle brms (EPAC, effet PROXIMITÉ DESTINATAIRE)

Family : categorical**Links :** muFA = logit; muFESK = logit; muIS = logit**Formula :** construction ~ stc2 + (1 | ID_file)**Data :** eadat, it (Number of observations : 211)**Samples :** 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-1.14	0.26	-1.69	-0.68	1.00	4365	3228
muFESK_Intercept	-2.08	0.46	-3.14	-1.33	1.00	2793	3010
muIS_Intercept	-2.36	0.43	-3.33	-1.67	1.00	3100	2268
muFA_stc2	-0.02	0.41	-0.84	0.76	1.00	6504	4776
muFESK_stc2	-0.39	0.63	-1.66	0.80	1.00	6954	4228
muIS_stc2	1.32	0.62	0.17	2.61	1.00	4845	3528

TABLEAU B.12 – Résumé du modèle brms (EPAC, effet POSTURE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	195	h95	190	h90	185	h85
Antéposition	-0.02	0.48	0.52	-1.67	1.27	-0.84	0.76	-0.70	0.64	-0.61	0.56
Antéposition & ESK	-0.39	0.27	0.73	-3.08	2.36	-1.66	0.80	-1.42	0.61	-1.29	0.49
In situ	1.32	0.99	0.01	-0.79	4.01	0.17	2.61	0.34	2.38	0.46	2.22

TABLEAU B.13 – Distributions postérieures pour le modèle brms (EPAC, effet POSTURE)

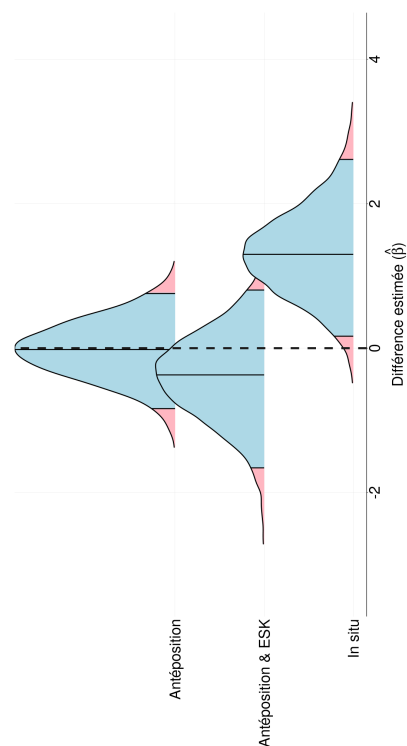


FIGURE B.5 – Distributions postérieures pour le modèle brms (EPAC, effet POSTURE)

Family : categorical

Links : muFA = logit; muFESK = logit; muIS = logit

Formula : construction ~ anm2 + (1 | ID_file)

Data : eadat_lit (Number of observations : 211)

Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-1.14	0.24	-1.66	-0.71	1.00	4401	3198
muFESK_Intercept	-2.16	0.50	-3.28	-1.34	1.00	2708	3403
muIS_Intercept	-2.13	0.33	-2.87	-1.55	1.00	5336	3829
muFA_anm2	-0.74	0.41	-1.56	0.03	1.00	8877	4620
muFESK_anm2	-1.15	0.66	-2.58	0.05	1.00	5794	3636
muIS_anm2	0.15	0.55	-0.98	1.22	1.00	8354	4814

TABLEAU B.14 – Résumé du modèle brms (EPAC, effet ANIMATION)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Antéposition	-0.74	0.03	0.97	-2.34	0.76	-1.56	0.03	-1.42	-0.08	-1.33	-0.16
Antéposition & ESK	-1.15	0.03	0.97	-5.24	0.71	-2.58	0.05	-2.29	-0.12	-2.12	-0.24
In situ	0.15	0.62	0.38	-2.37	2.52	-0.98	1.22	-0.79	1.04	-0.65	0.93

TABLEAU B.15 – Distributions postérieures pour le modèle brms (EPAC, effet ANIMATION)

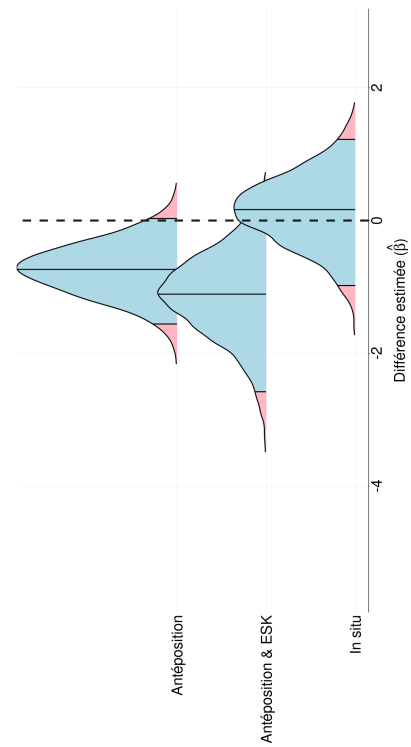


FIGURE B.6 – Distributions postérieures pour le modèle brms (EPAC, effet ANIMATION)

Family : categorical
Links : muFA = logit; muFESK = logit; muIS = logit
Formula : construction ~ refor2 + (1 | ID_file)
Data : eadat_lit (Number of observations : 211)
Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-1.13	0.24	-1.66	-0.69	1.00	4636	3190
muFESK_Intercept	-2.12	0.45	-3.14	-1.37	1.00	3780	3614
muIS_Intercept	-2.13	0.33	-2.85	-1.56	1.00	5100	3773
muFA_refor2	0.04	0.45	-0.88	0.91	1.00	8700	4403
muFESK_refor2	1.13	0.55	0.06	2.21	1.00	7233	4555
muIS_refor2	-0.32	0.74	-1.88	1.02	1.00	7823	3643

TABLEAU B.16 – Résumé du modèle brms (EPAC, effet REFORMULATION)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Antéposition	0.04	0.54	0.46	-1.61	1.73	-0.88	0.91	-0.73	0.76	-0.63	0.68
Antéposition & ESK	1.13	0.98	0.02	-1.04	3.42	0.06	2.21	0.23	2.04	0.35	1.92
In situ	-0.32	0.35	0.65	-3.98	2.05	-1.88	1.02	-1.61	0.82	-1.42	0.66

TABLEAU B.17 – Distributions postérieures pour le modèle brms (EPAC, effet REFORMULATION)

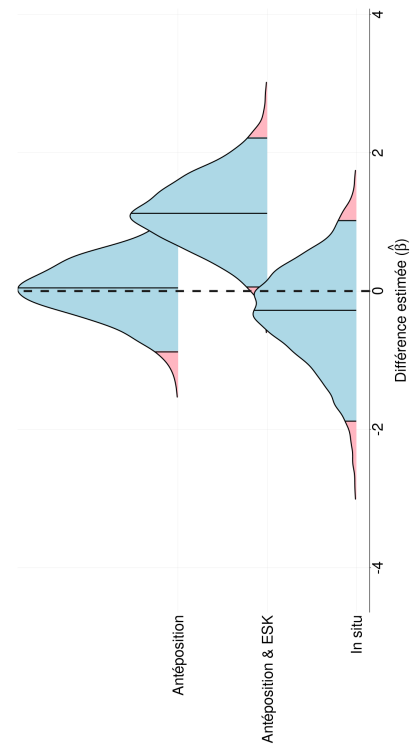


FIGURE B.7 – Distributions postérieures pour le modèle brms (EPAC, effet REFORMULATION)

Family : categorical

Links : muFA = logit; muFESK = logit; muIS = logit

Formula : construction ~ rself2 + (1 | ID_file)

Data : eadat_lit (Number of observations : 211)

Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-1.13	0.25	-1.70	-0.69	1.00	5308	3757
muFESK_Intercept	-2.31	0.53	-3.49	-1.47	1.00	3180	3132
muIS_Intercept	-2.31	0.43	-3.29	-1.61	1.00	4131	3135
muFA_rself2	-0.11	0.49	-1.08	0.83	1.00	8764	4436
muFESK_rself2	2.11	0.60	0.97	3.33	1.00	6213	4253
muIS_rself2	1.60	0.65	0.40	2.94	1.00	4964	4100

TABLEAU B.18 – Résumé du modèle brms (EPAC, effet AUTO-RÉPÉTITION)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Antéposition	-0.11	0.42	0.58	-2.58	1.52	-1.08	0.83	-0.91	0.68	-0.81	0.58
Antéposition & ESK	2.11	1.00	0.00	0.28	4.49	0.97	3.33	1.16	3.14	1.26	3.00
In situ	1.60	1.00	0.00	-0.94	4.32	0.40	2.94	0.57	2.72	0.70	2.55

TABLEAU B.19 – Distributions postérieures pour le modèle brms (EPAC, effet AUTO-RÉPÉTITION)

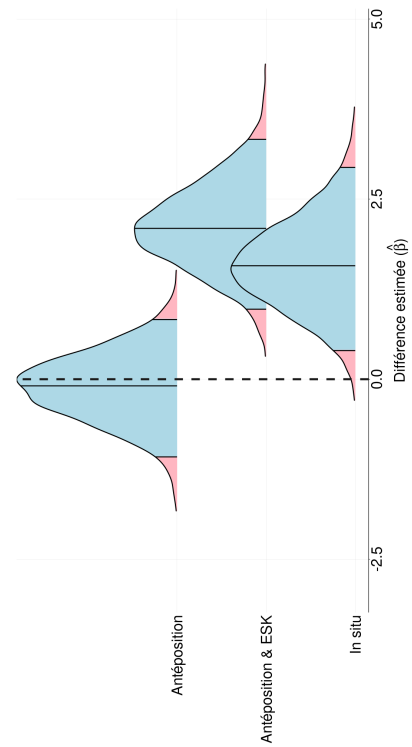


FIGURE B.8 – Distributions postérieures pour le modèle brms (EPAC, effet AUTO-RÉPÉTITION)

Family : categorical**Links :** muFA = logit; muFESK = logit; muIS = logit**Formula :** construction ~ ro2 + (1 | ID_file)**Data :** eadat_lit (Number of observations : 211)**Samples :** 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-1.17	0.26	-1.72	-0.71	1.00	4641	3575
muFESK_Intercept	-2.11	0.49	-3.21	-1.33	1.00	3606	3334
muIS_Intercept	-2.14	0.34	-2.89	-1.54	1.00	6157	3833
muFA_ro2	1.02	0.61	-0.15	2.21	1.00	9655	4547
muFESK_ro2	-0.71	1.16	-3.29	1.26	1.00	5829	3948
muIS_ro2	-0.75	1.30	-3.76	1.39	1.00	7905	4374

TABLEAU B.20 – Résumé du modèle brms (EPAC, effet RÉPÉTITION AUTRE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Antéposition	1.02	0.95	0.05	-1.48	3.84	-0.15	2.21	0.01	2.03	0.14	1.91
Antéposition & ESK	-0.71	0.28	0.72	-6.89	2.44	-3.29	1.26	-2.81	1.01	-2.49	0.78
In situ	-0.75	0.30	0.70	-9.39	2.57	-3.76	1.39	-3.07	1.12	-2.72	0.92

TABLEAU B.21 – Distributions postérieures pour le modèle brms (EPAC, effet RÉPÉTITION AUTRE)

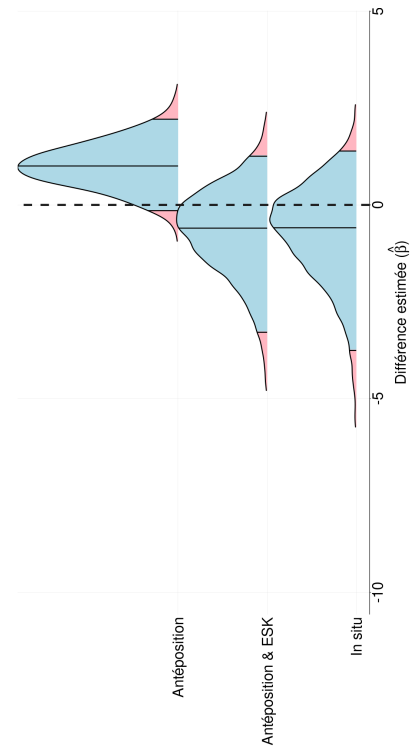


FIGURE B.9 – Distributions postérieures pour le modèle brms (EPAC, effet RÉPÉTITION AUTRE)

Family : categorical
Links : muFA = logit; muFESK = logit; muIS = logit
Formula : construction ~ thmc2 + (1 | ID_file)
Data : eadat_lit (Number of observations : 211)
Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-1.14	0.25	-1.69	-0.70	1.00	4184	2974
muFESK_Intercept	-2.09	0.45	-3.13	-1.37	1.00	3343	2333
muIS_Intercept	-2.18	0.34	-2.92	-1.59	1.00	5090	3107
muFA_thmc2	-0.06	0.39	-0.81	0.70	1.00	8224	4847
muFESK_thmc2	0.87	0.59	-0.22	2.10	1.00	6977	3879
muIS_thmc2	-0.87	0.55	-1.96	0.20	1.00	7104	3827

TABLEAU B.22 – Résumé du modèle brms (EPAC, effet FIN ARGUMENTAIRE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Antéposition	-0.06	0.44	0.56	-1.40	1.42	-0.81	0.70	-0.69	0.57	-0.60	0.49
Antéposition & ESK	0.87	0.94	0.06	-1.55	3.27	-0.22	2.10	-0.04	1.88	0.05	1.73
In situ	-0.87	0.06	0.94	-3.18	1.15	-1.96	0.20	-1.78	0.04	-1.65	-0.08

TABLEAU B.23 – Distributions postérieures pour le modèle brms (EPAC, effet FIN ARGUMENTAIRE)

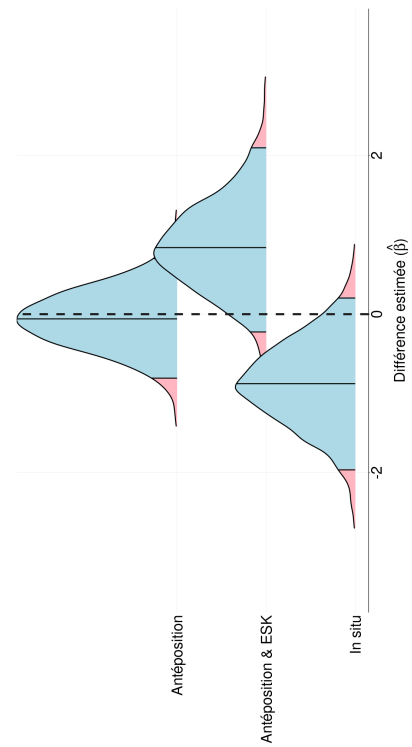


FIGURE B.10 – Distributions postérieures pour le modèle brms (EPAC, effet FIN ARGUMENTAIRE)

Family : categorical
Links : muFA = logit; muFESK = logit; muIS = logit
Formula : construction ~ aswg2 + (1 | ID_file)
Data : eadat_lit (Number of observations : 211)
Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-1.13	0.24	-1.65	-0.69	1.00	4599	4204
muFESK_Intercept	-2.13	0.50	-3.25	-1.34	1.00	3096	3209
muIS_Intercept	-2.58	0.49	-3.72	-1.78	1.00	2810	2503
muFA_aswg2	-0.01	0.57	-1.17	1.04	1.00	6499	4042
muFESK_aswg2	1.03	0.72	-0.39	2.46	1.00	5080	4133
muIS_aswg2	2.42	0.71	1.13	3.92	1.00	4034	3393

TABLEAU B.24 – Résumé du modèle brms (EPAC, effet RÉPONSE AUTO-FOURNIE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Antéposition	-0.01	0.51	0.49	-2.17	1.78	-1.17	1.04	-0.96	0.90	-0.85	0.79
Antéposition & ESK	1.03	0.93	0.07	-1.90	4.57	-0.39	2.46	-0.14	2.19	0.01	2.06
In situ	2.42	1.00	0.00	-0.04	5.99	1.13	3.92	1.31	3.64	1.44	3.46

TABLEAU B.25 – Distributions postérieures pour le modèle brms (EPAC, effet RÉPONSE AUTO-FOURNIE)

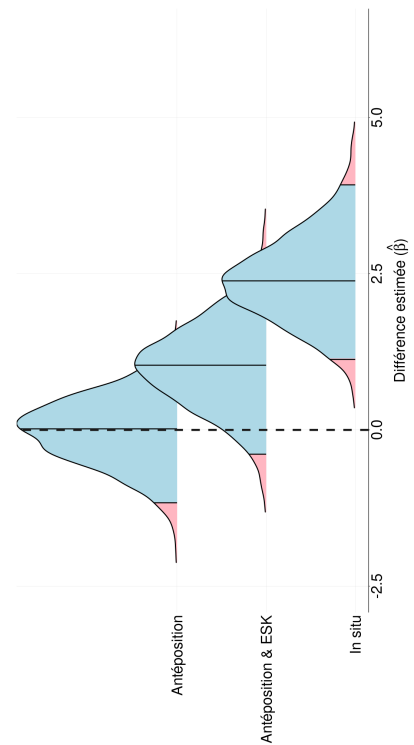


FIGURE B.11 – Distributions postérieures pour le modèle brms (EPAC, effet RÉPONSE AUTO-FOURNIE)

Family : categorical
Links : muFA = logit; muFESK = logit; muIS = logit
Formula : construction ~ opni2 + (1 | ID_file)
Data : eadat_lit (Number of observations : 211)
Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-1.14	0.25	-1.69	-0.69	1.00	5528	3229
muFESK_Intercept	-2.26	0.48	-3.37	-1.48	1.00	3347	2898
muIS_Intercept	-2.85	0.58	-4.21	-1.93	1.00	5410	2908
muFA_opni2	0.23	0.39	-0.55	0.97	1.00	8958	4567
muFESK_opni2	1.53	0.56	0.47	2.73	1.00	7386	4324
muIS_opni2	-2.95	1.29	-5.99	-0.97	1.00	5656	3190

TABLEAU B.26 – Résumé du modèle brms (EPAC, effet OPINION IMPLICITE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Antéposition	0.23	0.71	0.29	-1.18	1.57	-0.55	0.97	-0.41	0.86	-0.33	0.79
Antéposition & ESK	1.53	1.00	0.00	-0.47	3.90	0.47	2.73	0.64	2.50	0.74	2.37
In situ	-2.95	0.00	1.00	-10.05	-0.01	-5.99	-0.97	-5.37	-1.21	-4.94	-1.36

TABLEAU B.27 – Distributions postérieures pour le modèle brms (EPAC, effet OPINION IMPLICITE)

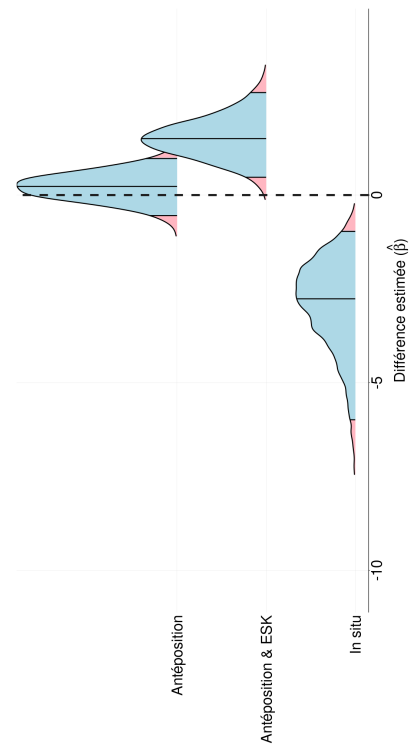


FIGURE B.12 – Distributions postérieures pour le modèle brms (EPAC, effet OPINION IMPLICITE)

Annexe C

Corpus ESLO : Extractions

Item	Reference	Age	Extraction
1	ESLO1ENT002C	35-55	oui les mots les mots comment dirais -je? malpolis malhon- nêtes oui
2	ESLO1ENT002C	35-55	aussi bien au point de vue euh comment dirais -je? moi loge- ment pur euh ambiance de vie et puis même au point de vue artistique ça ça habituera les gens à voir autre chose que des petites maisons comme dans ces rues -là où
3	ESLO1ENT009C	35-55	'oh oui voyez -vous avant si je vous dis avant d"avoir la voiture je faisais partie du d"un club je comment dirais -je? je faisais quelques voyages organisés des week-ends de deux jours alors ça c"était assez appréciable'
4	ESLO1ENT009C	35-55	'que le l"handicap moi je ne le sais pas j"ai pas suivi mes études mais ils savent bien par contre moi j"ai des nièces qui ont été je sais pas il y en a une qu"est qui a fait chimie qui a fait comment dirais -je? elle est oh des elle avait quel- commencé d"abord un départ de pharmacienne'
5	ESLO1ENT009C	35-55	'mais mais alors on fait peut-être euh y a peut-être pas assez des des des comment dirais -je? de je sais pas comment dirais -je? de d"étudiants de profs pas de professeurs mais enfin des vous savez pour les recherches ou autres on doit manquer un peu de cette société -là alors est -ce qu"on a été trop dur pour leur examen qu"ils se trouvent à lâcher ou alors est - ce que c"est le Français qui s"amuse à être sur les bancs de l"école jusqu"à temps que papa et maman les financent et qu"ils donnent la place parce qu"on a'
6	ESLO1ENT009C	35-55	'mais alors comment dirais -je euh? vous avez quand même une tenue très correcte et puis au fond votre enquête elle est sportive ça va avec votre tenue c"est c"est pas c"est pas quelque chose euh moi je sais pas'
7	ESLO1ENT009C	35-55	'et puis alors ce que je fais beaucoup à Orléans c"est c"est euh comment dirais -je? la conduite euh du patron la conduite de la maison'

Item	Reference	Age	Extraction
8	ESLO1ENT009C	35-55	'ah bé déformer les choses oh c"est in- incroyable attendez qu"est -ce que j"ai encore remarqué? euh la pa- la jeune fille du magasin qui a vingt ans elle dit à un jeune homme qu"elle connaissait euh quels sont les souliers de ta mère? oh les godasses de ma mère je sais pas ce qu"elle prend ce qu"elle a'
9	ESLO1ENT011C	35-55	'non que voulez -vous qu"on ait comme projet?'
10	ESLO1ENT011C	35-55	'mais un gosse euh qui arrive qui n"a qu"à se mettre les pieds sous table que voulez -vous qu"il fasse? il est pas habitué'
11	ESLO1ENT018C	35-55	je crois que les filles euh sont comment dirais -je? apprennent plus vite plus facilement que les garçons hein dans les jeunes âges
12	ESLO1ENT041C	35-55	'y a comment s"appelle-t -il? le Robert'
13	ESLO1ENT056C	35-55	'eh bien ça c"est une question qui est très vaste euh on fait assez s- mais qu"entendez -vous par là? euh au point d- à quel point de vue?'
14	ESLO1ENT056C	35-55	'sur euh le plan national surtout euh on on on s"est dit mais où va-t -on? et que va-t -il euh va-t -il'
15	ESLO1ENT056C	35-55	'sur euh le plan national surtout euh on on on s"est dit mais où va-t -on? et que va-t -il euh va-t -il'
16	ESLO1ENT056C	35-55	'que vouliez -vous qu"il fasse?'
17	ESLO1ENT064C	35-55	'on a une maison des jeunes comment marche-t -elle? euh j"oserais pas critiquer je n"en sais rien'
18	ESLO1ENT078C	35-55	'eh bien euh j"ai débuté euh dans dans dans ma carrière et dans une autre administration euh j"étais j"étais rentré dans l"enseignement j"étais instituteur et je n"y suis pas resté longtemps parce que euh la grammaire je ne m"y plaisais pas euh pour des raisons qui seraient d"ailleurs c"est délicat de vous expliquer ma foi euh j"ai bien réfléchi avant de partir parce que quand même on a beau être jeune toute la vie devant soi on se dit il ne faut pas faire de faux pas au départ et j"ai quitté l"enseignement pour aller dans cette branche la branche PTT ma foi qu"aurais -je pu faire d"autre? vous savez il faut être réaliste entre ce qu"on voudrait faire et ce qu"on peut faire il y a quelquefois une nuance'
19	ESLO1ENT078C	35-55	'euh il y a une affaire de maturité d"esprit à quel moment l"enfant réalise-t -il que le travail seul paie? ça c"est variable avec les enfants'
20	ESLO1ENT078C	35-55	'euh que demandent les habitants d"une ville? c"est assez assez terre-à-terre comme euh comme revendication ils demandent de payer le moins d"impôts possible et ils demandent par contre beaucoup de réalisations'
21	ESLO1ENT078C	35-55	'que voulez -vous c"est déplorable euh prenez euh un bon sujet de français ou un bon sujet euh une rédaction euh avec les meilleures idées les les plus belles idées bien exprimées s"il y a eu t- trop de fautes d"orthographe moi je ne suis pas hein?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
22	ESLO1ENT086C	35-55	'que voulez -vous? j'ai de du raisin de l'ananas du de l'abricot ou du pamplemousse et de l'orange'
23	ESLO1ENT094C	35-55	'et ça moi je fais des efforts avec les les manoeuvres de la maison pour me mettre pour les mettre à l'aise et je crois que c'est comme ça qu'on gagne leur confiance alors si vous me demandez est -ce que ça a de l'importance dans quel sens entendez -vous votre question?'
24	ESLO1ENT095C	35-55	'oui je sais bien elle est très compliquée mais moi je l'ai bien apprise les autres l'ont bien appri- c'est ce qui fait justement le charme moi je trouve d'être compliquée étant donné que maintenant l'école est obligatoire et que tous les enfants peuvent apprendre l'orthographe je sais bien qu'il y en a qui sont plus doués les uns que les autres je crois que pour l'orthographe il faut lire beaucoup mais non pas lire comme les élèves des fois nous disent ah je lis oui mais comment lisez -vous? il faut lire et dire tiens comment se fait -il tiens là y a E N T pourquoi ah oui parce que ça s'accorde avec telle chose oh ben alors madame on ne voit pas l'histoire ben j'ai dit si mais vous regardez tout de même comment les mots sont écrits et les enfants qui lisent beaucoup en général ont une bonne orthographe'
25	ESLO1ENT096C	35-55	'oh je regarde un peu les décès les comment dirais -je? moi je donne un coup d'oeil un peu sur tout'
26	ESLO1ENT105C	35-55	'alors ce qu'ils ont voulu faire ils veulent rattraper rattraper comment vous dirais -je? euh euh la production comme les autres pays comme en Allemagne où ils ont investi beaucoup en machines outils nous on n'a pas investi on a des machines out- y a des machines à l- ici à l'usine où je travaille qui ont qui sont de vingt ah disons qui marchent depuis dix-neuf vingt depuis vingt à peu près ils ont pas été changés'
27	ESLO1ENT105C	35-55	'mais non j'en ai été très contente très très contente d'abord c'était des soeurs mais assez évoluées voyez? euh pas pas trop non non non non franchement euh euh je vous dis je vous dis elles sont pas puis alors euh comme euh fréquentations euh c'était pas c'était pas comment je vous dirais -je? euh n'importe qui enfin c'est des filles de commerçants euh c'était des gens déjà enfin d'un certain milieu vous comprenez? alors euh franchement j'en ai été très satisfaite ah oui oui oui oui oui oui oui très très très d'ailleurs maintenant faut un examen euh elles étaient très gentilles très gentilles ça ça très j'en ai connues mais elles étaient très gentilles c'est des f- des filles de la de la Sagesse d'ailleurs'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
28	ESLO1ENT105C	35-55	'ça je vous jure que dans le quartier euh l- y a des gamins des gamines qui ont maintenant dix-sept dix-huit ans ils vous passent devant vous ils vous disent ni bonjour ni bonsoir et ça j"ai j"ai et ça j"ai horreur de ça ça je vous dis sincèrement c"est pas pour donner des gants à ma fille hein ça je je ah s- ça aurait été j"aurais eu elle aurait pas été comme il comme s- enfin comment je vous dirais -je? elle aurait pas dit bonjour ça je disais que mais alors ça je lui je lui disais toujours tu diras bonjour aux gens et dans le quartier tout le monde tout le monde me dit vous avez une fille qu"est -ce qu"elle est gentille et ça c"est vrai et ça c"est vrai et ça c"est vrai et elle dit bonjour à toutes les petites grands-mères que je rencontre ils sont tous à me demander des nouvelles'
29	ESLO1ENT105C	35-55	'oh non j"ai pas eu à m"en plaindre parce que elle a un parler assez correct ça franchement euh euh elle elle est pas elle est pas mal elle est comment vous dirais -je? elle dit pas de gros mots'
30	ESLO1ENT109C	35-55	'et il a cinquante-cinq ans maintenant que voulez -vous qu'on aille chercher ailleurs?'
31	ESLO1ENT111C	35-55	'en général très très satisfait dans cet atelier mais il faut je pourrais vous parler des autres où j"ai travaillé j"étais dix ans je vous assure que il c"était pas toujours bien euh voyons vous comment vous dirais -je euh? nous avions des conditions qui étaient assez au point de vue salaire premièrement et puis au point de vue propreté et tout ce qui s"ensuit y avait un peu de c"était délaissé si vous voulez un peu'
32	ESLO1ENT111C	35-55	'bien moi écoutez à l"heure actuelle euh moi quand on parle de la ville d"Orléans à l"heure actuelle vous savez que euh la ville d"Orléans euh comment vous dirais -je moi? nous avons des impôts'
33	ESLO1ENT114C	35-55	on se demande tiens euh cette ville -là où est -ce? euh ce personnage pour situer la date exacte
34	ESLO1ENT118C	35-55	oui euh où faudrait -il aller pour apprendre à parler un bon français?
35	ESLO1ENT118C	35-55	bien une une question de euh euh comment dirais -je? de de connaissances quoi hein de euh qui intervient
36	ESLO1ENT118C	35-55	'euh pourquoi? ben mon dieu parce que euh j"ai d"abord l"habitude hein? puisque c"est c"est pratique- c"est c"est mon métier et puis que euh bien souvent c"est une question de euh comment dirais -je? de de trésorerie de finance qui intervient voyez -vous?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
37	ESLO1ENT122C	35-55	'euh aide -toi et l"on t"aidera on fait toujours euh de trop pour les autres et jamais assez pour soi-même qu"est ce que voulez vous? que voulez -vous qu"on fasse? s"occuper d"une collectivité c"est attirer toutes les critiques alors il est certain que est -ce qu"on sait exactement ce qu"on veut? est -ce qu"on sait ce qui ce qu"on souhaite? et pour faire quelque chose euh mettons -nous à la place de celui qui doit décider est -ce qu"il a les moyens de pouvoir le faire? bon bien entendu s- la moindre chose faite pour la collectivité revient terriblement cher et et chacun va se plaindre de d"y participer financièrement alors euh on trouvera toujours quelque chose à à dire il manquera toujours quelque chose même quand on pensera tout avoir et et il est certain que y a des choses dominantes y a des besoins urgents hm y a toujours des besoins urgents si on nous dirait qu"il manque un piscine à Orléans bah il est certain qu"il qu"une piscine couverte à Orléans euh ça serait le souhait de tout le monde et puis celui qui souhaite cette piscine euh à Orléans n- n"y va jamais par exemple ou ou celui qui la souhaite pas s- irait peut-être assez souvent on fait on fait comme partout sur Orléans qu"est -ce que vous voulez? on fait comme partout euh'
38	ESLO1ENT127C	35-55	's"est répandue en France depuis quelques années et qui prend beaucoup d"expansion c"est tout le monde a des vêtements qui ont soit d"adultes soit d"enfants dans dans sa famille euh chez soi je veux dire euh qui qui ne sont pas abîmés qui sont en bon état et alors on les on les porte on les porte alors y a on les porte et là et euh y a la réception des vêtements ils sont achetés très peu cher remarquez c"est p- c"est pas cher mais enfin on arrive à faire pas mal d"argent avec quand même ça pa- parce que ça dépanne beaucoup de familles et c"est revendu ici à Orléans c"est à la salle des fêtes alors euh c"est fou ce qu"y a de vê- euh combien il y avait -il d"articles? un nombre incroyable d"articles et al- et ça se vend énormément les gens se jettent là-dessus parce que évidemment la vie est difficile la vie est dure'
39	ESLO1ENT127C	35-55	'ah à côté de ça je dis qu"y a des changements mais y en a qui ne se sont pas faits assez rapidement quand je pense qu"on apprend les langues avec tous les moyens dont on dispose maintenant exactement comme on les apprenait nous de notre temps alors là c"est c"est s- c"est vraiment déplorable enfin ils s"y mettent je sais que à au lycée comment s"appelle-t -il ce lycée?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
40	ESLO1ENT127C	35-55	'dans les élections nationales pourquoi y a-t-il deux tours? eh bien c'est ce sont les les deux qui ont le pourquoi je sais pas vous me demandez l'explication pourquoi s- les deux qui ont eu le plus de voix qui qui restent en course? alors y a faudra pas trop me pousser parce que y a une chose que j'ai jamais compris c'est la répartition des des sièges alors quand euh ah ça non je la trouve profondément injuste et je je comprends pas'
41	ESLO1ENT127C	35-55	où tombera-t-on?
42	ESLO1ENT127C	35-55	'c'est difficile à dire on a une bibliothèque combien y en a-t-il? oh là là'
43	ESLO1ENT133C	35-55	'j'ai mes petites habitudes et puis alors dans une euh comment dirais -je? c'est pas une casserole une poêle je mets du beurre je la fais au beurre moi pas à l'huile'
44	ESLO1ENT133C	35-55	ça y avait y avait pas mal de choses y avait pour les clients euh comment vous dirais -je? moi je sais pas y avait des comptes y
45	ESLO1ENT133C	35-55	comment dirais -je euh?
46	ESLO1ENT135C	35-55	comment voulez -vous?
47	ESLO1ENT137C	35-55	'j'y crois pas je vous le dis carrément à l'heure actuelle vous voyez on est sur la période du référendum oui non oui non euh non non oui bon bah euh j'ai mon idée ça d'accord euh on entend euh des discussions on on nous dit euh faut voter oui faut voter non euh m- monsieur De Gaulle nous a dit euh si vous votez pas oui euh je fais ma valise bon ben que voulez -vous y faire? on n'y peut rien'
48	ESLO2ECOLE1273C	35-55	'Barcelone alors c'est d'où? où est -ce?'
49	ESLO2ECOLE1273C	35-55	'tu m'inquiètes garçon hein cathédrale c'est alors quelle est la différence avec une église?'
50	ESLO2ECOLE1273C	35-55	bah y a une cathédrale à Orléans quelle est la différence entre une église et une cathédrale?
51	ESLO2ECOLE1273C	35-55	donc tu fer- en feras vite le tour y en a quelques-uns comme ça alors Arthur de quelle ville viens -tu?
52	ESLO2ECOLE1274C	35-55	'Tristan est deux fois plus petit que son père qui mesure un mètre quatre-vingt-quatre quelle est la taille de Tristan? un alpiniste achète soixante mètres de corde pour la somme de cent dix-huit euros quatre-vingts quel est le prix d'un mètre de cette corde?'
53	ESLO2ECOLE1274C	35-55	quel est le poids de la fillette? ça ces quatre là vous les faites sur le cahier de brouillon
54	ESLO2ECOLE1274C	35-55	après avoir été augmenté ça sera celui -là sur le cahier du jour de quarante-sept virgule soixante-trois euros par mois un journaliste gagne mensuellement mille neuf cent cinquante-deux virgule quatre-vingt-dix-huit euros combien gagnait -il avant son augmentation de salaire? allez et quand vous avez fini ça vous venez me voir

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
55	ESLO2ECOLE1274C	35-55	'chut je pense pas que tu aies besoin de causer euh eh le jour où tu seras sérieux on pour- on pourra travailler ensemble je t'en veux pas trois non le maître revient suis -moi toi hé je vois pas le rapport entre ça et ça voilà qui n'ai -je pas vu sur le premier?'
56	ESLO2ECOLE1279C	35-55	'Sherlock l'écoute en hochant la tête justement il est un peu plus de neuf heures du matin et voilà Billy Bouledogue qui quitte son poste il s'avance tranquillement dans le couloir d'un pas lent et assuré Lola s'approche pour lui parler Bonjour Billy comment s'est passée cette première nuit de travail? très bien sans aucun problème affirme Billy seulement une chose me tracasse cette nuit j'ai rêvé que l'alarme tombait en panne et il se trouve que mes rêves se réalisent souvent je pense que vous devriez faire vérifier les circuits électriques et il gonfle la poitrine l'air important Lola un peu étonnée lui répond qu'elle va tout de suite s'en occuper c'est bien dit Billy en s'éloignant et maintenant je rentre chez moi dormir un peu il faut que je sois en forme pour ce soir'
57	ESLO2ECOLE1279C	35-55	'et lui il a il a dit bonjour Billy comment s'est passée cette première nuit de travail? très bien sans aucun problème seulement une chose me tracasse cette nuit j'ai rêvé que l'alarme tombait en panne'
58	ESLO2ECOLE1279C	35-55	par où vas -tu commencer?
59	ESLO2ECOLE1280C	35-55	euh Arthur où sont les dialogues?
60	ESLO2ECOLE1284C	35-55	'c'est pour montrer la différence c'est un tout petit peu au-dessus mais où est le stylo bille?'
61	ESLO2ECOLE1284C	35-55	'où est celui -là que je t'ai donné en début d'année? vendredi matin rouge bleu vert noir crayon de papier gomme stylo-plume bien'
62	ESLO2ECOLE1284C	35-55	donc en France où sont débarqués principalement les poissons? en Bretagne est -ce que quelque part y a une logique?
63	ESLO2ECOLE1284C	35-55	'si tu lis la légende ça correspond à quoi cette zone avec ces petits traits horizontaux ces pointillés horizontaux? Manon c'est-à-dire que toute la zone de la France où il y a ces pointillés horizontaux les poissons qui seront livrés chez les poissonniers viendront plutôt de Bretagne c'est-à-dire toute cette zone'
64	ESLO2ECOLE1290C	35-55	'n'écoutes écoute bien dix par quels moyens développe-t-on le commerce? onze quelles industries sont développées? douze comment le peuple appréciait le roi Henry quatre? et treize comment meurt Henry quatre? dans cette dernière question je vous demande pas un article de journal me résumant comment il est mort'
65	ESLO2ECOLE1299C	35-55	'c'est pas une question ça combien coûte-t-il?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
66	ESLO2ECOLE1300C	35-55	'ceux qui sont moi si je le faisais je courrais je suis là un deux deux pieds là j'ai pas fait l'élan hein est -ce que vous comprenez? donc là les cerceaux sont pour vous donner un repère on est d'accord? et on mesure en triple saut à partir de la ligne bleue jusqu'au dernier endroit où vous avez ça marche? bien ça marche? bien donc vous ferez passerez plusieurs fois le tout est que vous acquériez le principe après on mesurera et puis je vais faire deux passages de haies les deux passages de haies ne sont pas chut seront dans des écarts différents donc vous passez les deux et ensuite vous courrez le plus vite possible sur l'un ou l'autre passage le but du jeu quand on passe les haies c'est d'aller le plus vite possible sans sauter le plus haut possible pourquoi ne saute-t-on pas au-dessus des haies? on retombe lentement non non si on saute haut on n'avance pas le but du jeu c'est d'avancer le plus vite possible donc si on avance le plus vite possible il faut être le plus près du sol pour pouvoir appuyer avec ses pieds pour pouvoir avancer plus vous sautez haut moins vous avancez alors faut pas se toucher les barrières parce que ça te ralentit ça te fait mal mais alors là vous inquiétez pas celui qui fait tomber il remet en place ça marche? et puis on va faire du saut en hauteur là après un atelier mesure de hauteur ça marche aller on s'échauffe là dans cette partie là pendant que j'installe aller vous échauffez toutes les articulations et vous laissez les tapis là où ils sont'
67	ESLO2ENT1034C	35-55	'parce que souvent ça ça colle pas avec euh avec euh nos nos comment dirais -je? si on souvent on est parti euh dans des choses comme ça c'est parce que euh on aime plutôt la'
68	ESLO2ENT1049C	35-55	'euh hm et donc euh hm quand euh parce que mon père voulait pas que j'aille à de l'autre côté euh quand euh j'étais en âge d'aller au collège le collège de Saint-Jean-le-Blanc qui se trouvait euh où se trouvait -il d'ailleurs celui -là? je ne sais plus euh si il se trouvait rue Demay euh il allait être déplacé rue Creuse c'est-à-dire de l'autre côté de la route de la grande route de Sandillon et mon père a pas souhaité que je me fade tous les jours le trajet là-bas surtout qu'il imaginait bien qu'il y aurait un vélo ou je sais pas quoi pour aller là-bas il avait pas envie de me voir courir ce risque -là et il a préféré à ce que j'aille au lycée qui était en fait derrière ma maison puisque la maison où j'habitais elle était du côté droit de la rue des Balles enfin du côté pair et donc adossé à la rue des Anguignis et le collège il prend en il était en partie au bout de mon jardin'
69	ESLO2ENT1057C	35-55	'euh j'ai des problèmes de euh comment dirais -je? euh j'entends bien la rythmique ça y a pas de'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
70	ESLO2REPAS1253C	35-55	'je dis mais où est Sereine sur la photo? tu sais j'ai j'ai une très mauvaise vue en s- en en ce moment mais en fait c'est pas Soren qui manque c'est Thibault donc euh Soren va croire que je l'ai pris pour Thibault et il va être vexé'
71	ESLO2REPAS1264C	35-55	'comment veux -tu t'en sortir?'
72	ESLO1ENT029C	15-25	quels étaient exactement les termes de votre question?
73	ESLO1ENT063C	15-25	'je sais pas si c'est fait qu'en France euh c'est pas fait en je sais pas dans tous les pays je crois les hôtels sont ils ont divers euh un nombre d'étoiles alors euh par exemple vous avez quatre étoiles euh y en a même pas à Orléans parce que c'est à Paris par exemple dans les grandes grandes villes euh quatre étoiles euh luxe alors ce sont les grands grands hôtels euh ensuite nous à Orléans nous avons euh les mieux ce sont trois étoiles ensuite deux étoiles une étoile je sais pas si le comment j- dirais -je? ceux qui s'occupent de l'hôtellerie qui classent euh par exemple un hôtel euh faut pour être classé dans les mieux dans les trois étoiles faut qu'il ait euh tant de chambres tant de salles de bain euh et caetera'
74	ESLO1ENT092C	15-25	'oui où est -ce que c'est la forêt d'Orléans?'
75	ESLO1ENT092C	15-25	et comment pouvez vous voir si ça marche?'
76	ESLO1ENT110C	15-25	'je sais pas comment c'est chez vous?'
77	ESLO1ENT121C	15-25	'de lettres bah c'est comment entendez -vous lettres?'
78	ESLO1ENT125C	15-25	'parce que il y a eu des discussions à à ce propose alors mais enfin je le dis quand même parce que je le pense hein c'est on laissait les gens on laissait les gens venir on les poussait pas si ça vous intéresse vous venez mais alors euh on présentait des des trucs euh c'était assez inaccessible quoi fallait même les étudiants on n'y mettait pas tellement les pieds quand on avait un quand on rentrait à cette maison de la culture on avait l'impression de rentrer dans une chapelle ou dans une église euh d'aller à d'aller à une cérémonie et puis quand vous arriviez euh moi je me souviens d'un cas très précis c'est une expérience personnelle quoi quand vous arriviez à neuf heures cinq on vous disait non monsieur c'est c'est commencé c'est pas la peine vous vous rentrez pas quoi c'est tout même si ça vous intéressait enfin vous vous sentez pas vous vous sentez pas chez vous c'est pas la maison où on peut se cultiver euh pas à bon marché mais enfin où on peut trouver quelque chose on se sent pas à l'aise alors euh comment voulez -vous que? enfin moi moi étant étudiant déjà bon certaines facilités pour euh m'introduire dans un certain milieu'
79	ESLO1REU294C	15-25	'c'est lesquels les autres garçons?'
80	ESLO224H1249C	15-25	'c'était où ça les barrettes?'
81	ESLO2CINE1180C	15-25	'd'accord merci que diriez -vous à quelqu'un pour le convaincre d'aller le voir?'
82	ESLO2ENTJEUN1235C	15-25	que vais -je prendre? voilà donc

Item	Reference	Age	Extraction
83	ESLO2ENTJEUN1235C	15-25	ce soir que vais -je vous faire? alors haricots en boîte
84	ESLO2ENT1003C	15-25	comment vous perçoivent les autres?
85	ESLO2REPAS1255C	15-25	Killian ça veut dire quoi?
86	ESLO1DIV292C	35-55	alors que à la au remembrement on vous donnait quoi?
87	ESLO1DIV292C	35-55	bon parce que où où vous êtes euh votre euh la ferme elle-même par rapport à ici en gros?
88	ESLO1DIV292C	35-55	'attribution de la surface combien vous avez dit tout à l'heure?'
89	ESLO1DIV292C	35-55	'bon alors lui monsieur NPERS lui qu'est -ce qu'il alors euh les onze pages ça se résume en quoi?'
90	ESLO1DIV292C	35-55	quand est -ce que vous avez eu cette décision -là?
91	ESLO1ENTCONT205C	35-55	'je ne sais pas laissez peut-être la porte ouverte hein parce que les clés sont dehors bon attendez pourriez -vous revenir? vous êtes là jusqu'à quand alors?'
92	ESLO1ENTCONT212C	35-55	mais alors à partir de euh à quelle heure que vous dites?
93	ESLO1ENTCONT212C	35-55	et vous êtes comment?
94	ESLO1ENT002C	35-55	voulez oui pourquoi pas?
95	ESLO1ENT005C	35-55	'ben je pense qu'on devrait leur apprendre euh enfin évidemment les enfants oui à quel âge vous pensez? euh la lecture et tout à ce point de vue là vous voulez dire non?'
96	ESLO1ENT007C	35-55	'où irait le français si le français n'avait pas d'orthographe?'
97	ESLO1ENT008C	35-55	'c'est quoi donc ça? c'est quoi donc ça?'
98	ESLO1ENT009C	35-55	'parce que il y en a d'autres qui ne savent pas qui savent et qui vraiment prennent le beau et qui savent pas le définir facilement elles disent ben pourquoi on ne la voit pas la différence?'
99	ESLO1ENT009C	35-55	'ah bah maintenant vous savez c'est qu'ils sont tous euh le plus jeune quel âge il a? hum il a dix-sept ans'
100	ESLO1ENT009C	35-55	'ah oui naturellement là c'est pareil si vous voulez moi je sais pas moi j'aime quelqu'un de bien habillé j'aime bon on est au travail on est au travail mais une fois quand même je sais pas c'est si agréable quand euh j'admets que vous les hommes c'est peut-être dur de supporter une chemise un veston surtout en pleine chaleur c'est un fait mais maintenant avec tous les tissus on arrive quand même à avoir quelque chose puis enfin nos parents y étaient bien pourquoi donc nous on peut pas avoir ce ce genre -là? voyez -vous moi je vous cache pas un représentant arrive au magasin vous n'en voyez pas un en polo d'ailleurs il y en aurait un qui viendrait en polo hein'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
101	ESLO1ENT009C	35-55	'oui parce que voyez -vous moi je connais un gros gros un homme d'affaires à Paris qui mais qu- qui est formidable il a des une écurie de course il a tout il a trois enfants eh bien il les laisse vraiment bien se il se dit moi je suis arrivé par mes propres moyens lui il était d'au moins huit ou neuf enfants alors il était boulanger de boulanger il est fait des ét- des cours du soir il est arrivé il a passé ses examens bon ben maintenant le voilà gros boursier le voilà je peux pas vous dire parce que vous voyez sa situation eh bien moi je sais pas ses enfants mais ils sont elles sont d'une ses enfants vous les voyez y en a un qu'a vingt-cinq ans l'autre qu'a vingt-six ans voilà en- ils sont encore dans les jupons de papa vous vous rendez compte ils n'ont pas encore de ce de métier alors ça a été jusqu'au bac ou je sais pas j'ai jamais osé demander si ça avait des examens ben alors qu'est -ce que ça fera ces jeunes -là quand les parents vont plus être là? ah naturellement alors le père il s'est mis marchand de fonds tout en gardant lui son il s'est mis marchand de fonds en à Paris pour occuper ses enfants mais vous voyez les deux gars sont pas mariés la fille après le le père me dit une fois ah il dit vous Yolande vous êtes toute seule vous vous débrouillez que c'est pas mes enfants qui prendront le taureau par les cornes parce qu'au fond moi puisque je n'avais jamais été chez les autres et partir à trente ans trente-deux trente et un an j'ai travaillé c'est un fait eh bien voyez il dit c'est pas mes enfants qui vont pas se lancer ben je dis pourtant il dit j'ai beau leur dire non il dit faut que je sois là derrière après il a pas voulu mettre sa fille euh chez les autres il a pris une petite librairie- enfin il a pris une librairie'
102	ESLO1ENT011C	35-55	'puis alors voilà combien d'années qu'on est partis? dix ans?'
103	ESLO1ENT011C	35-55	'ah oui à la commune oui oui et c'est quoi au juste? c'est une sorte de redistribution des de terres'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
104	ESLO1ENT012C	35-55	'c"est-à-dire d"une part euh du milieu familial d"autre part bien sûr euh des professeurs euh qui sont chargés de les enseigner je pense que tous les individus ont les mêmes aptitudes intellectuelles hein à priori encore faut -il savoir les développer et leur faire comprendre que euh cet effort intellectuel qu"on leur demande est un c"est un jeu hein il est aussi distrayant de faire travailler son esprit qu"il est distrayant de faire travailler ses muscles je m"excuse de faire toujours cette comparaison avec le sport mais ça me paraît fondamental et je crois que le drame des enfants qui ne réussissent pas c"est parce qu"on n"a pas su leur montrer le caractère distrayant au sens propre du terme de l"activité intellectuelle on en a fait pour eux une obligation et une préparation à l'exercice d'un métier pour gagner leur vie c"est ce qu"il ne faut pas dire et je pense que euh j"ai d"ailleurs quand même quelques exemples enfin de d"enfants qui ne réussissaient pas qui un beau jour réussissent pourquoi? ben c"est le jour où ils ont eu la chance de tomber sur quelqu'un qui leur a dit ça la plupart du temps les professeurs je pense que euh ce qui est dramatique c"est que de nombreux professeurs ne sont pas des pédagogues et eux-mêmes ne sont pas persuadés de cela ça ne peut pas marcher'
105	ESLO1ENT012C	35-55	'depuis quand nous avons Larousse en dix volumes? il y a pas très longtemps justement nous ne l"avons que depuis cette année'
106	ESLO1ENT012C	35-55	'de parler en anglais d"ailleurs en fait on parle en quoi on parle en langage binaire maintenant en fait on parle avec un ordinateur je pourrais vous montrer des plans ici qui ont été dessinés par des ordinateurs on parle quoi actuellement? on parle le Fortran on parle le Pal on parle'
107	ESLO1ENT012C	35-55	'quel type d"expression qui m"agace?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
108	ESLO1ENT018C	35-55	'ah alors là comment se fait -il qu"il y a des grands et des petits? je crois que c"est un peu ça il y a des enfants qui sont il y a d"abord des enfants qui sont plus intelligents que d"autres qui sont plus réceptifs euh qui ont des indiscutablement euh des enfants mis euh quarante enfants mis dans une classe euh il y en a qui réussiront qui auront assimilé très bien d"autres qui n"assimileront pas ça ne veut pas dire que ceux qui n"assimilent pas sont des imbéciles mais enfin euh faut bien faire une sélection quand même d"une façon ou d"une autre en général ceux qui réussissent bien à l"école on peut admettre qu"ils réussissent quand même assez bien dans la vie je crois quand même il y a certainement une relation il y a les exceptions euh le cancre de l"école qui devient un PDG extraordinaire bien sûr mais enfin ça sont quand même des exceptions il y a des les enfants réussissent bien à l"école euh d"autres réussissent mal je sais pas pourquoi on ne sait pas on n"en sait rien c"est une question d"intelligence je sais pas comment vous vous appelez ça hein pour moi c"est une question d"intelligence de milieu d"ambiance un tas de choses qui interviennent peut -être mais je crois aussi une question d"intelligence il y a des gens intelligents d"autres ne le sont pas il y a rien à faire d"ailleurs ça euh'
109	ESLO1ENT018C	35-55	'alors revendications personnelles non ben alors il y a il y a pas tellement de problèmes de ce côté -là alors il y a des améliorations qui sont importantes quelquefois euh on fait pas mal d'améliorations à l"usine à l"atelier euh améliorations de conditions de travail ça je crois que c"est important d"avoir d"être en relation directement avec les intéressés parce que très souvent ils m"ont fait des suggestions qui étaient parfaitement valables et je ne savais même pas que ça existait euh euh je ne sais pas quoi euh l"autre jour il y avait il y avait des pare-flammes par exemple qui n"existaient pas euh j"en savais rien euh je me dis mais si on faisait un pare-flammes ça serait quand même euh évidemment ce sont des suggestions intéressantes enfin c"est un je crois que c"est un dialogue qui est intéressant qui est valable mais il ne faut pas euh maintenant on veut avoir la participation alors je ne sais pas ce que ça va donner j"ai pas encore lu leur décret je pense que ça va être difficile d"application il ne faut pas que ça sombre dans l"anarchie je crois que dans une société comme dans toute société il y a actuellement le seul euh la seule méthode qu'on ait trouvée c"est de la hiérarchisation euh parce que l"homme est fait comme ça il en a besoin moi je veux bien faire une société anarchique je ne vois je vois pas d'inconvénient mais qu"est -ce que ça donnera?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
110	ESLO1ENT041C	35-55	'eh bien mais qu'est -ce que vous appelez? enfin bourgeois c'est c'est quoi vous euh?'
111	ESLO1ENT051C	35-55	'on me mettait tout seul sur un chantier on me l'expliquait je le faisais comme on ne peut pas aujourd'hui même j'ai vu au au bout de deux ans d'apprentissage il pense tout à fait à autre chose qu'à son travail dans ce temps -là on pensait davantage à son travail déjà premièrement alors ils comptent euh ça d'accord c'est bien de savoir écrire euh mais la main-d'œuvre compte beaucoup quand même c'est bien leurs centres d'apprentissage mais je crois que les jeunes s'intéressent beaucoup moins au travail euh manuel que que dans le temps je vois j'ai des apprentis là j'en ai un et puis alors c'est tout à fait normal il va pas chercher à comprendre c'est j'aimerais qu'il me demande ben vous vous faites ça comme ça pourquoi? ça me ferait plaisir qu'il me demande euh de le conseiller bah non euh je le fais euh lui il va pas chercher à le faire euh c'est tout à fait normal que je sache le faire mais et par contre à l'école euh il travaille pas non plus lui alors je me demande s- celui -là qu'est -ce qu'on va en sortir je sais pas ils cherchent pas à arriver à arriver à faire quelque chose ils croient que c'est la machine qui va faire tout et ça beaucoup je leur reproche'
112	ESLO1ENT051C	35-55	'Larousse là combien quelle quelle année qu'il a le dernier?'
113	ESLO1ENT053C	35-55	'oui je crois que ça c'est ça va être un projet qui doit certainement peut-être un jour euh ils vont peut-être avoir trente mille habitants mais enfin je crois que faut beaucoup d'années hein parce que les gens ne ça n'est pas intéressant de venir ici qu'est -ce que vous voulez? faudrait on ne fait partie d'abord de de la zone où les usines ne s'implantent plus'
114	ESLO1ENT053C	35-55	'et alors euh s- c'est vous savez y a plusieurs volumes y a combien? je crois douze volumes vous savez c'est vraiment c'est des livres où on peut trouver toutes sortes de choses dessus'
115	ESLO1ENT053C	35-55	'c'est euh comment ça s'appelle donc? quelquefois on va le chercher quand on a des lettres à faire remarquez c'est mais c'est très très vieux quand même comme euh comme genre de de français vous savez'
116	ESLO1ENT054C	35-55	mais vous avez choisi Orléans pourquoi parce que? vous aviez déjà des attaches sur Orléans ou bien
117	ESLO1ENT054C	35-55	'ben à l'origine bof vous savez c'est assez complexe tout ça comment ça s'est passé déjà à l'origine ça? c'est les étudiants bof euh il faut faut le juger l'analyser?'
118	ESLO1ENT054C	35-55	'ah oui bien ah oui c'est le même que j'ai moi et vous l'avez depuis quand?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
119	ESLO1ENT054C	35-55	'oh je me souviens plus comment ça s'appelle je l'ai là-haut remarquez bien mais c'était un genre de petite euh comment je pourrais appeler ça? un petit résumé de de savoir-faire on pourrait dire de de de savoir-faire en toutes choses'
120	ESLO1ENT054C	35-55	'oui si c'est euh si c'est à des amis euh comment vous avez dit?'
121	ESLO1ENT055C	35-55	pour alors euh continuer leurs études pour avoir de comment je dirais pour être au niveau des pour avoir les diplômes et tout enfin?
122	ESLO1ENT055C	35-55	comment ils font le choix?
123	ESLO1ENT055C	35-55	'comment j'expliquerais ça?'
124	ESLO1ENT055C	35-55	comment ça se traduit?
125	ESLO1ENT055C	35-55	comment je fais une omelette?
126	ESLO1ENT064C	35-55	qui coiffent mieux on tient compte de la personnalité de la cliente on tient quand même compte de de comment je je pourrais dire? de son physique
127	ESLO1ENT064C	35-55	'euh comment je dirais? des choses qui pourront vraiment leur être absolument nécessaires dans les contacts qu'ils auront par la suite'
128	ESLO1ENT065C	35-55	'mais toute façon je suis contre le communisme ça c'est c'est c'est c'est difficile à expliquer mais je suis je suis contre parce que je vais vous dire quand j'étais jeune euh en mille neuf cent trente-six vous savez qu'il y a eu les premières grèves pour obtenir la semaine de quarante heures euh et caetera et j'ai assisté à des défilés à Saint-Denis hein et j'ai remarqué que tous les communistes qui défilaient en chantant l'Internationale étaient à moitié ivres et beaucoup étaient des voyous et ça m'a toujours frappée je crois que c'est et s- ça m'a ça m'a toujours frappée ça euh je j'ai gardé cette image -là et dans je l'ai dans mon souvenir et je ne pourrai pas m'en je ne pourrai pas m'en séparer ça m'a toujours déplu et qu'est -ce qu'il y a ma biche?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
129	ESLO1ENT065C	35-55	'quand on est allé en Russie toutes les deux euh on est allé en voyage organisé bien entendu parce qu'on pouvait pas circuler euh seules on est parti en autocar donc on a visité les pays communistes l'Allemagne de l'Est euh la Pologne et enfin la Russie bien sûr et dans notre autocar y avait un jeune ménage qui était communiste enfin c'était mais c'était des gens qui avaient une grosse situation qui avaient leur voiture qui avaient leur appartement leur maison de campagne qui pouvaient tout de même se permettre de faire un voyage en Russie à deux hein et avant ils avaient fait un voyage dans toutes les capitales d'Europe centrale donc c'est des gens qui avaient des moyens et ils se disaient communistes alors moi ça m'a ça m'a choquée vous voyez c'est c'était inexplicable c'était j'ai peut-être tort mais en principe un communiste partage tout hein et quand on est communiste qu'on qu'on se permet de faire des revendications c'est qu'on en a besoin eux ils sont communistes pourquoi? et par contre aussi y avait une dame qui était là alors une dame seule qui était euh divorcée je ne sais plus une dame oui un petit peu plus âgée que moi et elle elle allait en Russie pour raconter à son bureau euh je suis allée en Russie parce que je suis communiste mais elle ne visitait rien elle restait assise dans l'autocar quand on allait visiter des musées'
130	ESLO1ENT077C	35-55	'la couronne la couronne roy- la couronne royale elle était là et tout hein comment qu'elle s'appelait là la couronne là enfin la reine?'
131	ESLO1ENT077C	35-55	'j'ai une nièce qui habite à la frontière suisse on est allé lui rendre visite on est resté cinq six jours en Suisse à combien de kilomètres de la Suisse? vingt kilomètres?'
132	ESLO1ENT077C	35-55	'euh où j'ai mis mon briquet là? merci monsieur oui ah non ça les prochaines vacances je sais pas on veut pas'
133	ESLO1ENT077C	35-55	'à le dresser là ça va faire combien de temps qu'il est dedans?'
134	ESLO1ENT077C	35-55	'il a réussi pourquoi parce que c'est la marine qui l'a hein?'
135	ESLO1ENT077C	35-55	'oh ben toutes les barricades qu'ils ont faits tout ça là c'est puis envahir les machins là les comment on appelle? théâtres tout ça'
136	ESLO1ENT077C	35-55	'on on y perd à chaque grève alors c'est pas compliqué l'autre fois là combien qu'on a de temps on a arrêté au mois de mai là un mois?'
137	ESLO1ENT077C	35-55	'moi j'ai mon petit neveu là quel âge qu'il a là? euh'
138	ESLO1ENT077C	35-55	'ah oui oui oui ah ça on fait attention oui oui oh oui alors là euh une orthographe ça c'est tout hein même si vous écrivez une lettre chez un patron ou quelque chose pour demander hein? si y a possibilité d'avoir de la rallonge ou quelque chose et puis si y a des fautes dedans il va leur dire euh qu'est-ce que c'est que ça? non l'orthographe ça ça fait beaucoup'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
139	ESLO1ENT077C	35-55	'c'est un copain qui m'avait fait ça comment que c'est son nom là? regarde voir c'est marqué dessus là'
140	ESLO1ENT078C	35-55	'tout à fait normal euh je n'y vois pas d'inconvénients je n'ai rien contre les enfants et puis vous me direz qu'est -ce qui peut arriver à un enfant dans un immeuble comme ça? y a déjà euh une affaire d'escalator escalier mécanique'
141	ESLO1ENT083C	35-55	'ben ça dépend ce que vous parlez de quoi d'élections présidentielles ou d'élections lé-gislatives? pour élire le président de la république? par exemple'
142	ESLO1ENT086C	35-55	'qu'est -ce qui a?'
143	ESLO1ENT086C	35-55	et y en a comme ça combien? dix
144	ESLO1ENT086C	35-55	mais vous descendez où comme ça euh quand vous venez en?
145	ESLO1ENT087C	35-55	'quand est -ce qu'on fait une omelette?'
146	ESLO1ENT087C	35-55	'je vous répondrais je vous dirais monsieur c'est pourquoi?'
147	ESLO1ENT090C	35-55	'c'est pour quel genre de choses?'
148	ESLO1ENT091C	35-55	'comment moi je fais une omelette? oh ben c'est tout simple y a rien de plus simple eh bien je casse mes oeufs je les je les bats avec mon fouet enfin pas à la main avec mon fouet électrique'
149	ESLO1ENT094C	35-55	'je dis d'abord entrez et suivant vous vous me donnez cet exemple mais je ne pourrais pas dire qu'il soit tellement bien choisi parce que quand on frappe je sais à peu près qui ça va être parce que si c'est quelqu'un de la maison par exemple il parlera d'abord avec les dames qui sont à côté et je reconnais très bien la voix si c'est quelqu'un de l'extérieur il viendra euh par ce j- il passe sous ma fenêtre donc je vois à peu près qui c'est mais donc pourquoi vous me posez cette question?'
150	ESLO1ENT094C	35-55	'pareil la grammaire c'est pour pour Nicolas je trouve ça très bien maintenant je ne sais pas comment à qu- qu- mais j'ai deux enfants qui vivent on vit tous ensemble enfin euh tous les quatre Marie-Chantal ne parle pas bien et fait même actuellement encore des fautes de français et elle est dans une école euh ou ou on parle bien dont la directrice parle très bien et mon fils qui qui est élevé comme Marie-Chantal parle très très bien à quoi ça tient?'
151	ESLO1ENT094C	35-55	'par exemple hier après-midi nous sommes allés chez nous à Olivet et hier c'était un jour de grande circulation et quand nous sommes arrivés à Olivet mon mari me dit maintenant on revient par quel pont le pont Royal ou ou le pont quel pont je prends? je réponds pas cinq minutes après il redit je sais pas quel pont je dois prendre parce que t- il réfléchissait depuis ah j'ai dit écoute prends n'importe lequel mais prends -en un ça moi je c'est et je lui ai donné un pont je lui ai dit n'importe le je lui ai donné un nom il a pris celui -là mais moi j'ai besoin qu'il y ait une solution tout de suite'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
152	ESLO1ENT094C	35-55	'une omelette? et pourquoi vous demandez ça? pour faire une omelette je casse des oeufs je fouette les oeufs peut-être que ce n'est pas le verbe que vous désirez que je dise mais enfin je fouette les oeufs parce que je prends un fouet donc je considère que je fouette les oeufs j'ajoute un peu de lait parce que ça fait une omelette plus légère je mets du beurre dans ma poêle quand ce beurre est fondu je verse mes oeufs fouettés dedans et évidemment le temps de cuisson et je retourne cette omelette dans un plat'
153	ESLO1ENT095C	35-55	'ça j'aime beaucoup écrire d'ailleurs mes amies le dir- elles seraient là le diraient elles diraient oh on a toujours des lettres ça c'est sûr quand elle écrit pas on dit qu'est-ce qu'il y a?'
154	ESLO1ENT096C	35-55	comment je fais?
155	ESLO1ENT096C	35-55	'qu'est-ce qui me vient?'
156	ESLO1ENT096C	35-55	'il y a le la gauche c'est quoi? c'est le parti qu'on a maintenant la gauche c'est ça? non?'
157	ESLO1ENT096C	35-55	'hm hm et puis maintenant avec toutes ces choses à Olivet c'est comment ça s'appelle déjà? La Source'
158	ESLO1ENT099C	35-55	'non je pense pas non enfin je crois pas j'arriverais ici qu'est-ce que je ferais euh pas grand-chose qu'est-ce que vous voulez faire quand on est rentré chez soi? rien si préparer la la la vaisselle pour la le repas le soir la soupe et puis c'est tout si c'était tout ce que je pourrais faire'
159	ESLO1ENT099C	35-55	'non je pense pas non enfin je crois pas j'arriverais ici qu'est-ce que je ferais euh pas grand-chose qu'est-ce que vous voulez faire quand on est rentré chez soi? rien si préparer la la la vaisselle pour la le repas le soir la soupe et puis c'est tout si c'était tout ce que je pourrais faire'
160	ESLO1ENT099C	35-55	'ça c'est difficile à c'est facile à c'est difficile à expliquer ça hein je sais moi qu'à Orléans euh c'est quand même assez c'est comment je vais vous expliquer moi? je sais pas trop moi comment vous dire maintenant vous savez les gens ils sont ils sont froids ils sont c'est c'est pas dans dans certaines villes c'est beaucoup plus beaucoup beau- beaucoup plus ami et à Orléans c'est pas du tout la même chose ma foi moi je vois chez nous en Normandie quand j'y vais c'est pas pareil moi je suis connu un peu partout là-bas alors euh j'ai beaucoup de d'amis mais ici à Orléans euh celui qui qu'on ne qu'on qu'on ne connaît pas on et puis c'est vrai c'est ce qui se passe souvent à Orléans mais point de vue différence entre euh moi je j'ai jamais eu l'occasion d'avoir affaire comme ça entre eux personnellement pour moi personnellement vous voyez je trouve pas qu'y a y a pas de différence moi j'ai jamais eu de difficulté de ce côté -là mais'
161	ESLO1ENT099C	35-55	'non mais je sais pas comment c'était la question déjà? euh oui'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
162	ESLO1ENT099C	35-55	'comment je vais vous dire moi? oh moi je pense que oui oui mais c'est'
163	ESLO1ENT099C	35-55	'on les garde y a ce que ce qui est les réclames les machins comme ça tout ça je fous ça en l'air moi ça sert à rien ça on va le mettre où? même les les publicités tout ça moi ça m'intéresse pas on les ouvre on les regarde'
164	ESLO1ENT104C	35-55	'd'autres raisons vous comprenez ça comment?'
165	ESLO1ENT104C	35-55	de quand il parle?
166	ESLO1ENT105C	35-55	'huit heures vous entendez? pas plus parce que on arrive le soir ça nous permet on sort à cinq heures moins le quart alors nous on n'est pas tellement loin ça me fait arriver à cinq heures ici hein mais vous savez on est claqué on est fatigué parce que dix heures dans presque entendre du bruit les les machines qui tournent les presses euh qu'est -ce que vous voulez? euh moi je dis que c'est trop la journée continue c'est très bien euh comme en Amérique ils veulent faire comme en Amérique mais c'est malheureux à dire les capitalistes Français ils trouveront rien vous comprenez? neuf heures neuf heures ben moi je dis que c'est de trop faire la journée continue faudrait sept heures huit heures au grand maximum vous comprenez?'
167	ESLO1ENT105C	35-55	'alors comment vous voulez qu'on puisse faire des reproductions? alors y a des y a c'est un service de métal un service de métal ben qu'est -ce qu'ils font? c'est si on est chronométré c'est c'est souvent ils nous donnent ils nous en retirent des centièmes mais c'est pas sous l'ouvrière a le travail manuel nous c'est du travail manuel c'est du travail c'est des machines perfectionnées qu'il faudrait ou que euh au lieu de de sortir mettons vingt mille pièces la machine elle n'en sorte quarante c'est pas c'est pas sur la personne humaine elle est la même'
168	ESLO1ENT105C	35-55	'mon mari il a eu une réunion il a été en à la haute direction avec d'autres camarades et là ils ont compris que fallait y a longtemps y a longtemps que les syndicats disent que fallait qu'ils investissent en France on n'a pas investi qu'est -ce que vous voulez? euh les f- le le le le patronat français c'est des vieux euh des vieux qui ont cinquante soixante ans soixante-dix ans ben ils sont encore à l'ancien régime'
169	ESLO1ENT105C	35-55	'ils ont obtenu les mêmes avantages mais ils sont combien? c'est une petite usine ils ont des des c'est une usine de machines outils ils font quoi? ils sont peut-être quinze vingt alors il gagne comme si il était à Paris vous en avez beaucoup à Orléans enfin à Orléans pas pas simplement à Orléans mais Orléans euh euh Vierzon qui montent travailler à Paris tous les jours'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
170	ESLO1ENT105C	35-55	'vous comprenez? non ah non en province on n"a pas on n"a même pas le t- je vous dis l"ouvrier l"ouvrier a un l"ouvrier il de province il est pas payé le même taux que ceux que ceux de de de la Paris et de la banlieue de parisienne à part peut-être des usines comme euh comme Peugeot euh à Sochaux ou comme ça alors ceux -là euh mais mais des petites usines non y en a encore en en France euh même même là le point de vue travail aussi dans des y a une usine Brive là qui travaille ben c"est la même chose euh ils sont payés quoi les ouvrières? y en a quelques-unes qui f- qui font des bons mois mais euh sans ça ils se font des petits mois'
171	ESLO1ENT105C	35-55	'ben oui ça pose des problèmes parce que moi je vois je travaille j"avais ma fille à et je travaillais et elle j- j"avais ben en arrivant le soir bon ben enfin elle était tout elle était toute seule encore ça va mais quand on en a trois quatre deux ou pas pas pas peut-être pas quatre mais mettons on en a au moins deux alors qu"est -ce que vous voulez? maman ceci maman cela alors euh c"est fatal et et faut dire de ranger ceci faut dire de ranger cela'
172	ESLO1ENT105C	35-55	'le midi ils arrivent ben maintenant mettons mettons que des ga- des gosses le midi ils mangent euh en cantine hein beaucoup ils mangent en cantine ceux qui ont des parents ils travaillent mais le soir ben et peut pas s"occuper des devoirs on peut pas s"occuper euh on n"est pas toujours à tu es rentré à l"heure? oui maman je suis rentré à l"heure mais il est bien rentré après l"heure qu"est -ce que vous voulez?'
173	ESLO1ENT105C	35-55	'moi j"ai vu on la petite belle-soeur à ma soeur là qui est en Angleterre Claire Marie-Claude sa maman elle travaille à la perception dans un bureau elle allait au collège c"est une enfant qui a toujours été première de sa classe sortant de l"école communale elle est entrée en sixième au collège eh bien elle a perdu un année pourquoi? parce que y avait une copine qui disait mais viens donc Marie-Claude on va aller se promener alors ils allaient faire les les les magasins ils allaient se promener et faisaient peut-être faisaient peut-être pas faisaient pas de mal mais elle arrivait ben les devoirs ils étaient pas faits ou elle arrivait juste un petit peu avant sa mère alors quand ils vont à une école ben ils ont beaucoup de devoirs à faire eh bien elle a perdu une année'
174	ESLO1ENT105C	35-55	'ou alors quand il y a la campagne ici ben moi je les vois y a des tuyaux en face et je te grimpe sur les tuyaux et qu"est -ce que vous voulez? ben c"est c"est c"est c"est c"est c"est fatal'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
175	ESLO1ENT105C	35-55	<p>'qu'est -ce que vous voulez? moi je vois mais j'ai des gens k- que la femme elle travaillait pas il y a des des pieds-noirs enfin des des gens des gens d'Algérie enfin les gens qui sont revenus depuis la guerre d'Algérie on appelle ça les les pieds-noirs enfin ce sont de nationalité soit italienne soit espagnole mais enfin ils sont pas ils sont français parce que c'était mais enfin si vous voulez ce qu'ils trouvaient drôle quand ils sont arrivés en France que la femme elle travaillait y a un camarade à mon mari il dit moi il dit je comprends pas pourquoi que ta femme elle travaille il dit elle peut pas rester chez toi alors mon mari mon mari il lui a dit non elle peut pas rester chez nous parce que moi j'ai une maison sur le dos et si y a pas la paye à ma femme ben comment tu veux que je paye les traites? c'est pas avec ma paye on peut pas alors mais lui ben maintenant sa femme elle travaille et il il a il a compris il a réalisé que maintenant que sa femme faut qu'elle travaille si il veut avoir quelque chose si on veut acheter si on veut avoir un peu de bien-être à moins à moins que vous avez quatre cinq gosses alors là y a les allocations familiales hein mais une un ménage qui a un ou deux enfants et qui veut avoir mettons le mari qui qui a une paye moyenne je vous parle d'une paye moyenne pas pas un cadre ni ni un agent technique ce qui fait déjà mieux qu'une paye moyenne eh bien qu'est -ce que vous voulez? cent cent dix mille mettons hein eh bien faut que la femme elle travaille si il veut entretenir la voiture euh si il veut s'acheter mettons la télévision ou si il veut changer un peu de meubles ou si il veut partir un peu un peu en vacances faut que la femme elle travaille ça serait -il qu'une demi-journée euh mais faut que la femme elle travaille les payes de l'homme n'a pas suivi euh euh à proportion le l'homme le enfin le mari est moins payé qu'avant la guerre de trente-neuf remarquez qu'on a de moins besoin on a moins besoin c'est peut-être pas la même vie vous comprenez maintenant on a on le ménage français on investit davantage en appareils ménagers c'est la télévision c'est l'aspirateur c'est le frigidaire c'est la machine à laver je vous dirai qu'on n'en change pas souvent en j- mais enfin quand même'</p>

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
176	ESLO1ENT105C	35-55	'maintenant plus hein c'est peut-être fini maintenant y a le tout-à-l'égout ils rarrangent tout ça ils ouvrent machin mais y en a beaucoup y en a y en a qui sont tombés comme ça y a pas tellement longtemps y a deux ans euh euh pas dans la rue de Bourgogne même mais sur dans dans des petites rues et c'est effondré et c'est rapport à ça parce que y avait certainement comprenez? les eaux usées ils allaient ils partaient je vais dire à la dérive dans les carrières et soi-disant que c'est c'est des ouvriers moi qui m'ont dit ça alors naturellement le le ça à force la terre ça la ravine et puis un beau jour ben qu'est -ce que vous voulez? ça diminue puis ça s'effondre oh la rue de Bourgogne c'est du r- vieil Orléans là hein'
177	ESLO1ENT105C	35-55	'c'est s- Orléans était une ville euh oh oh c'est c'est k- c'est resté comment je vous dirais -je? caractères euh euh royaux quoi c'est ça ça a'
178	ESLO1ENT105C	35-55	'ah ben il est je l'ai mis dans dans la chambre où y avait ma fille parce que ça fait un peu ça fait un peu comment je vous dirais? euh on met des livres'
179	ESLO1ENT105C	35-55	'c'est pas mal c'est pas mal ceux qui ont la facilité d'apprendre euh qu'est -ce que vous voulez euh?'
180	ESLO1ENT108C	35-55	'oui mais comment on fait le ch-? y a y a certainement un choix mais comment on le fait il y a les écoles privées qui sont souvent euh c'est un ensei- enfin ce sont des religieux et des religieuses qui f- enfin qui les dirigent qui les'
181	ESLO1ENT109C	35-55	'oui ça va être obligatoire jusqu'à dix-huit ans à partir de de quand?'
182	ESLO1ENT109C	35-55	oh y a combien? trois quatre ans?
183	ESLO1ENT109C	35-55	'et si on est malade alors qu'est -ce qui reste?'
184	ESLO1ENT111C	35-55	'qu'est -ce qui qu'est -ce compte le plus? euh comment est -ce'
185	ESLO1ENT111C	35-55	'alors les étudiants eux c'est plus pareil les étudiants à l'heure actuelle qu'est -ce qui se passe? c'était exactement pareil vous avez les meneurs chez eux alors meneurs sur quel parti je ne veux pas parler de partis politiques'
186	ESLO1ENT112C	35-55	'dieu cette dame -là comme elle est bien élevée mais où elle a pu être élevée? puis alors après quand j'ai su on peut pas questionner les gens comme puis ça se fait pas quand j'ai su qu'elle était fi- oh j'ai dit j'ai pas pu m'empêcher de lui dire une fois'
187	ESLO1ENT113C	35-55	on est combien?
188	ESLO1ENT113C	35-55	bon euh nous en étions où?
189	ESLO1ENT113C	35-55	écoutez une ville comme Orléans on est combien cent mille habitants à peu près?
190	ESLO1ENT113C	35-55	vous êtes à combien de Paris?
191	ESLO1ENT113C	35-55	ben où est -ce que tu les as mis?
192	ESLO1ENT113C	35-55	comment on fait une omelette?

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
193	ESLO1ENT114C	35-55	pourquoi je me plais à Orléans?
194	ESLO1ENT114C	35-55	'qu'est -ce que vous entendez par là ce qui compte le plus euh?'
195	ESLO1ENT114C	35-55	'bah euh oui parce que euh mais on n'est oh puis alors on est en f- en France on enfin nous on sommes en mater- on peut pas dire qu'on ait un programme mais enfin y a quand même un faut déjà leur a- quand même les initier un peu en lecture en tout ça en ça va peut-être un p- je sais pas peut-être un peu trop c'est peut-être un petit peu trop tôt pourquoi ils réussissent pas oh? puis parce qu'ils sont peut-être il faut qu'y a relation entre l'école et la famille pour qu'un enfant réussisse faut qu'il se qu'il sente une grande union entre les l'école et la famille à tous points de vue enfin là c'est vous me prenez au dépourvu je'
196	ESLO1ENT114C	35-55	à quoi on voit ça?
197	ESLO1ENT114C	35-55	'oh c'est-à-dire que j'en ai plusieurs j'en ai un dans ma classe j'en ai ici mais mon premier Larousse je l'ai eu j'avais quel âge? dix ans mon premier'
198	ESLO1ENT119C	35-55	'de comment que ça s'appelle? du Périgord du Périgord noir'
199	ESLO1ENT122C	35-55	'euh aide -toi et l'on t'aidera on fait toujours euh de trop pour les autres et jamais assez pour soi-même qu'est ce que voulez vous? que voulez -vous qu'on fasse? s'occuper d'une collectivité c'est attirer toutes les critiques alors il est certain que est -ce qu'on sait exactement ce qu'on veut? est -ce qu'on sait ce qui ce qu'on souhaite? et pour faire quelque chose euh mettons -nous à la place de celui qui doit décider est -ce qu'il a les moyens de pouvoir le faire? bon bien entendu s- la moindre chose faite pour la collectivité revient terriblement cher et et chacun va se plaindre de d'y participer financièrement alors euh on trouvera toujours quelque chose à à dire il manquera toujours quelque chose même quand on pensera tout avoir et et il est certain que y a des choses dominantes y a des besoins urgents hm y a toujours des besoins urgents si on nous dirait qu'il manque un piscine à Orléans bah il est certain qu'il qu'une piscine couverte à Orléans euh ça serait le souhait de tout le monde et puis celui qui souhaite cette piscine euh à Orléans n- n'y va jamais par exemple ou ou celui qui la souhaite pas s- irait peut-être assez souvent on fait on fait comme partout sur Orléans qu'est -ce que vous voulez? on fait comme partout euh'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
200	ESLO1ENT122C	35-55	'euh aide -toi et l'on t'aidera on fait toujours euh de trop pour les autres et jamais assez pour soi-même qu'est ce que voulez vous? que voulez -vous qu'on fasse? s'occuper d'une collectivité c'est attirer toutes les critiques alors il est certain que est -ce qu'on sait exactement ce qu'on veut? est -ce qu'on sait ce qui ce qu'on souhaite? et pour faire quelque chose euh mettons -nous à la place de celui qui doit décider est -ce qu'il a les moyens de pouvoir le faire? bon bien entendu s- la moindre chose faite pour la collectivité revient terriblement cher et et chacun va se plaindre de d'y participer financièrement alors euh on trouvera toujours quelque chose à à dire il manquera toujours quelque chose même quand on pensera tout avoir et et il est certain que y a des choses dominantes y a des besoins urgents hm y a toujours des besoins urgents si on nous dirait qu'il manque un piscine à Orléans bah il est certain qu'il qu'une piscine couverte à Orléans euh ça serait le souhait de tout le monde et puis celui qui souhaite cette piscine euh à Orléans n- n'y va jamais par exemple ou ou celui qui la souhaite pas s- irait peut-être assez souvent on fait on fait comme partout sur Orléans qu'est -ce que vous voulez? on fait comme partout euh'
201	ESLO1ENT122C	35-55	'c'est-à-dire euh ce qui s'est passé au point de vue national? comment je vois la le déroulement de la des événements?'
202	ESLO1ENT122C	35-55	'regardez tous ces mots composés de plus en plus il est pratiquement impossible de parler sans utiliser des mots composés et euh euh y en a combien? le pot de fleurs je prends celui -là le pot de fleurs euh y a un tas de mots où euh on a groupé deux trois mots ensemble pour en former un nouveau non on pourrait très bien avoir des termes plus courts euh en inventer éventuellement même dans la langue anglaise y a peut-être des mots courts simples faciles euh avec une permettrait de d'augmenter notre vocabulaire euh en utilisant un mot sans utiliser un terme euh très long et et d'une orthographe compliquée de l'orthographe à l'orthographe simple et'
203	ESLO1ENT127C	35-55	'comment est -ce qu'on décide son vote la seconde fois? ah ça dépend des gens y a des gens qui des- ils qui qui voteront contre quelqu'un contre euh oui contre quelqu'un par exemple je prends le dernier vote entre monsieur Couve de Murville et monsieur Rocart hein alors euh certaines personnes euh voulaient voter con- contre Couve contre Couve de Murville et ont vot- et euh auraient voté pour n'importe qui pourvu que ce soit contre Couve de Murville ça y a pas de doute hein et puis d'autres alors votent euh ben auraient voté pour lui par peur ont voté pour lui par p- par p-'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
204	ESLO1ENT127C	35-55	'mais je vais vous je vais vous dire que c'est ah c'est difficile voyez y a eu des élec- les des élections mon mari s'en est occupé puisque il lui il était co- conseiller municipal là-bas en Lorraine et c'est au moment des élections comment ça s'appelle? du conseiller général alors on avait proposé un un employé et qui était qui était quelqu'un de bien mais y a très très peu vraiment bien k- très très peu d'ingénieurs qui ont voté pour lui et c'est simplement pour la question qu'il était employé y a que ça qui a joué alors pour que voyez que c'est di-'
205	ESLO1ENT127C	35-55	'moi j'aime bien comment c'est l'après-midi? euh L'Invité du Dimanche alors quand il est bien quand c'est bien euh la dernière fois c'était bien'
206	ESLO1ENT127C	35-55	'ah ben moi je dirais plutôt non euh science-fiction non drame et épouvante non western euh non alors qu'est -ce qui reste? les autres par exemple L'Armée des Ombres dans quoi vous pouvez? ah un bon film gai'
207	ESLO1ENT127C	35-55	'ah ben moi je dirais plutôt non euh science-fiction non drame et épouvante non western euh non alors qu'est -ce qui reste? les autres par exemple L'Armée des Ombres dans quoi vous pouvez? ah un bon film gai'
208	ESLO1ENT127C	35-55	'dans quoi est -ce qu'on peut?'
209	ESLO1ENT127C	35-55	une bibliothèque comme ça sur de un deux trois quatre rayons on double euh devant et derrière combien ça peut faire? ah vous savez
210	ESLO1ENT127C	35-55	'ah bon ah ben un ordre ça va déjà mieux si il faut vous dire qu'une tranche ah par tranche j'aime mieux combien vous dites? alors un à mille mille c'est beaucoup mille hein il me semble que c'est beaucoup mille'
211	ESLO1ENT127C	35-55	combien ça fait mille livres? vous vous rendez compte vous? combien ça peut faire mille livres?
212	ESLO1ENT133C	35-55	'ah oui alors une omelette bien mousseuse vous l'aimez comment?'
213	ESLO1ENT133C	35-55	'mais alors faut qu'il écrive où?'
214	ESLO1ENT135C	35-55	'je sais pas pourquoi qu'est -ce qui se passe? est -ce qu'ils est -ce qu'ils ont des des professeurs trop jeunes? je sais pas ou alors les parents ça vient des parents mais enfin je ne sais pas ça alors c'est s- ça m- ça me choque ça'
215	ESLO1ENT135C	35-55	'dans quelque années j'ai un des statistiques prouvent que nous serons cin- il y aura combien? cinquante pour cent de Français qui auront moins de vingt-deux ans alors ces jeunes -là bah c'est eux les maîtres'
216	ESLO1ENT135C	35-55	'très peu ou pas du tout alors qu'est -ce ils apprenaient quoi? la langue de leurs parents'
217	ESLO1ENT135C	35-55	et comment ils et comment les parents parlaient? ça

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
218	ESLO1ENT137C	35-55	'moi je sais que s'ils f- prennent un métier en main bah parce qu'à l'heure actuelle on arrive à un point en France que à force d'étudier à force d'étudier bah les mômes ils vont sortir euh ils auront peut-être leur brevet ils auront peut-être enfin le leur bac ou n'importe quoi et puis qu'est -ce qu'ils font? par la suite si ils veulent si on peut pas euh continuer à'
219	ESLO1ENT137C	35-55	'ça euh les CRS à Paris euh c'est dommage j'ai pas mon livre là Le Peuple je sais pas où il est où c'est que tu l'as mis?'
220	ESLO1ENT139C	35-55	comment je faisais comment je les travaillais?
221	ESLO1ENT139C	35-55	'comment je vais vous expliquer m-? voyez nous on a quatorze ans treize quatorze ans m- on passe le euh certificat d'études notre premier examen après ça le brevet s- c'est ce qui vient après'
222	ESLO1ENT139C	35-55	'oh non ça c'est c'est révolu ça on garde voyez -vous? maintenant les les les costumes euh comme ça on les garde pour les fêtes quand il y a une fête et euh y a y a comment que ça s'appelle euh? y a des danses euh folkloriques alors à ce moment -là on ils mettent les les costumes mais autrement ça ne se fait plus du tout c'est complètement perdu ça'
223	ESLO1ENT139C	35-55	'euh comment que ça s'appelle? c'est ah oui ça ça doit être à l'Académie mais vous avez dit un nom mais c'est autre chose mais je me souviens plus du nom'
224	ESLO1ENT150C	35-55	'pourquoi j'en fais pas?'
225	ESLO1REPAS273C	35-55	ils partent pour combien de temps quinze jours?
226	ESLO1REPAS273C	35-55	vous la mettez où vous la mettez où votre cuve?
227	ESLO1REPAS273C	35-55	vous la mettez où vous la mettez où votre cuve?
228	ESLO1REPAS273C	35-55	'après le départ de Pierre il me dit oh dit donc il me dit ce que ça me plaît ce que Pierre nous a dit ce que ça me plaît mais il me dit tu sais il faudrait qu'il fasse un questionnaire très précis alors il me dit Pierre écoute Jean-Baptiste m'a dit qu'il faudrait que tu fasses quelque chose de très précis il me dit oh il me dit ah il me dit arrête va falloir que je vois ça que je mette ça sur pied et Jean-Baptiste était très emballé puis tout d'un coup le père Besançon me dit il a quel âge Jean-Baptiste? je dis écoute il a pas douze ans je dis tu l'as sans doute vu plus vieux que ça je dis ça va être un petit peu ennuyeux maintenant que tu lui as fait miroiter alors qu'il me dit ben c'est parce que sur les papiers j'ai mis pour les douze treize ans euh un duvet puis alors qu'il amène son duvet son matériel'
229	ESLO1REPAS273C	35-55	'la vigne vierge c'est quoi?'
230	ESLO1TEL307C	35-55	en principe euh et ça se passerait à quelle heure?
231	ESLO2ECOLE1273C	35-55	'et après c'est quoi étoile d'or?'
232	ESLO2ECOLE1273C	35-55	'ah c'était où ça le salon de l'agriculture?'
233	ESLO2ECOLE1273C	35-55	'bah euh c'est pas si évident que ça bon des animaux oui et alors quel est l'animal qui t'a le plus impressionné?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
234	ESLO2ECOLE1273C	35-55	'ouais ça fait plus d"une tonne et demie quoi une belle bête bien bien bien alors tu as vu quoi d"autre?'
235	ESLO2ECOLE1273C	35-55	'd"accord euh juste une question là euh pourquoi on dit Porte de Versailles?'
236	ESLO2ECOLE1273C	35-55	'c"était où?'
237	ESLO2ECOLE1273C	35-55	'tu as vu quoi d"autre?'
238	ESLO2ECOLE1273C	35-55	'qui a dessiné qui a conçu ce bâtiment il est et c"est quoi ce bâtiment?'
239	ESLO2ECOLE1273C	35-55	il sert à quoi ce bâtiment?
240	ESLO2ECOLE1273C	35-55	'ah ah ah une cathédrale c"est quand même un un rôle particulier quel est le but du jeu d"une cathédrale normalement? si tu enlèves les mains de ta bouche on va peut-être pouvoir entendre hein c"est embêtant hein'
241	ESLO2ECOLE1273C	35-55	Soissons est dans quel département?
242	ESLO2ECOLE1273C	35-55	'l"Aisne alors dans quel euh qui est -ce qui peut nous montrer euh l"Aisne allez Soissons il doit être sur la carte c"est sûr alors c"est dans quelle région de France?'
243	ESLO2ECOLE1273C	35-55	Nathan chut chut chut chut chut Loïs Maxence Loïs il habite où lui? Maxence ça a été signé où garçon? deux semaines ça a pas été suffisant et le reste a été signé? les devoirs les cahiers
244	ESLO2ECOLE1274C	35-55	'deux ah non ça je suis pas d"accord sur l"unité là c"est des quoi?'
245	ESLO2ECOLE1274C	35-55	'et comment je peux trouver des centimètres? sûrement pas et je peux pas avoir une autre je mets pas d"unités dans les o- dans les opérations en ligne je peux pas trouver des centimètres quand j"ai des mètres au départ ah bon? alors c"est des quoi?'
246	ESLO2ECOLE1275C	35-55	'trois trois trois trois trois trois trois trois y a trois tu as zéro tu pourrais le mettre ah très bien trois sauf que là c"est les mètres et là tu as des centimètres je sais pas comment tu peux arriver ce matin moi quand je partage des c- des concombres j"ai des concombres à la fin comment tu peux partager des mètres et trouver des centimètres? c"est pas possible quand je partage des mètres je trouve des mètres quand je partage des centimètres je trouve des centimètres y a un problème trois'
247	ESLO2ECOLE1275C	35-55	'non non tu peux pas avoir ça c"est des quoi? comment tu peux trouver des centimètres? y a un problème non ah oui tu as pas corrigé là c"est des mètres tu as des mètres au départ tu as des mètres à l"arrivée non bon'
248	ESLO2ECOLE1275C	35-55	'et tu as pas des des centimètres au départ pourquoi tu t"enquiquines? tu sais pas faire ça? celle -là on peut la garder en un combien de fois deux zéro trois quatre quatre chut euh c"était la langue ça aurait été mieux plutôt que le menton alors qui en est qui n"a pas fait le deux? tout le monde a fait le deux? tout le monde est au trois au moins?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
249	ESLO2ECOLE1275C	35-55	'tu l"as déjà ça a l"air quatre non un problème de virgule pour-quoi tu l"as mis? parce que tu mets ensemble toute façon tu peux me calculer le premier résultat parce que vingt euros vingt fois soixante ça fait quatre erreur de calcul le jour où tu cor- arriveras à compter ça va aller mieux hein c"est toujours un problème d"erreur là hein quatre quatre quatre tu travailles Arthur quatre quatre'
250	ESLO2ECOLE1275C	35-55	'opération qui est pas bonne j"ai pas de preuve de solution j"ai rien pas une mobilette rien tu la remets la virgule elle est là quatre cinq et euh quatre ah et puis ça coûte quoi? des dol-lars? quatre n"importe quoi cette addition ça pèse combien? alors j"ai une preuve de solution qui est mélangée avec une opération ça ressemble à quoi ça? rien à rien oui cahier du jour'
251	ESLO2ECOLE1275C	35-55	'cinq c"est mieux chut c"est quoi ça? ça un cinq? un neuf? un zéro? un? cahier du jour fais le deux bon vous me parlerez de vos skis'
252	ESLO2ECOLE1275C	35-55	'je m"en moque j"entends bien parler de ski toi tu continues tu colleras après si besoin est hein oui oui tiens tu m"en prends un suivant ah bon non mais attends tu m"écris n"importe quoi Julie c"est le numéro combien celui -là? combien? c"est du cas lourd oh tu as été plus rapide que moi chut chut chut chut cahier du jour ah c"est mieux mais alors quand y en a autant mets un S non cahier du jour ah c"est mieux suivant c"est mieux cahier du jour non il est là Arthur tu m"embêtes pas mon camarade qui a pas fini cahier du jour Maxence ça travaille? cahier du jour et enfin sept mars Lisa c"est pour de-main? trois jours pour me coller trois feuilles hé mais c"est pas trop tôt tu on commence à faire le problème en colonne c"est dur c"est dur allô bonjour oh je vous avais oublié aujourd'hui ah si si euh dix-huit je vérifie quand même trois quatre euh cent huit oui voilà au revoir'
253	ESLO2ECOLE1275C	35-55	'viens me voir déjà ah deux pour faire neuf dès qu"on forme bien les chiffres on comprend pas mieux que maintenant n"est -ce pas? ça c"est tu vois c"est des étourderies euh alors tu me l"as posée où cette opération? déjà si j"avais le bon texte ça irait mieux le texte il correspond à rien et combien il a de mètres?'
254	ESLO2ECOLE1275C	35-55	'je t"ai demandé combien il gagne mensuellement et on nous a posé la question mensuellement ça veut dire?'
255	ESLO2ECOLE1275C	35-55	'eh bah oui salaire c"est vrai que c"est écrit ici tu devrais être capable de le recopier ah non pourquoi y a rien d"écrit là?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
256	ESLO2ECOLE1275C	35-55	'non mais pourquoi vous êtes partis sur une division? qui a eu l'idée la prem- le premier? parce que vous êtes les seuls et c'est les seuls au mon- et c'est vous êtes tous au hasard à vouloir faire des divisions y aurait bien là une influence de quelques-uns ou de quelques-unes on parlait de quoi? hein ça va peut-être être mieux autre chose qu'une division? ah oui si tu soulignais ça irait mieux c'est disposé comme un goret'
257	ESLO2ECOLE1275C	35-55	'colle mais par contre si tu me faisais un trait pour séparer ça irait mieux fais -le avant bon alors c'est quoi qu'il faut me faire? il gagne combien aujourd'hui?'
258	ESLO2ECOLE1275C	35-55	combien? mille neuf cent vingt-deux bah oui euros avant il gagnait combien?
259	ESLO2ECOLE1276C	35-55	'c'est la faute de qui?'
260	ESLO2ECOLE1276C	35-55	tu as retrouvé quoi?
261	ESLO2ECOLE1276C	35-55	'donc c'est bien de ta faute au départ bien il reste qui?'
262	ESLO2ECOLE1276C	35-55	'juste une question euh alors aujourd'hui on va faire une technique préparatoire euh au au saut en hauteur parce que c'est l'atelier le plus euh je dirais à risque hein euh où il faut prendre quelques précautions et les autres euh euh fois on fera les quatre ateliers en même temps c'est-à-dire qu'on fera euh saut en longueur triple saut en saut en longueur et course euh course de haies on les fera en même temps et vous tournerez vous savez comme les ateliers qu'on faisait euh en gymnastique mais aujourd'hui on va tous passer sur la technique de la du saut en hauteur parce que c'est ça qui nécessite le plus de technique juste comme ça juste comme ça est -ce que y en a un qui oui? qui sait à peu près à quel niveau se situe la hauteur du record du monde? alors euh record du monde de saut en hauteur masculin parce que entre masculin et féminin c'est pas la même chose euh alors vous savez qu'il existe des records du monde en athlétisme en n- intérieur et en extérieur car on considère que c'est pas la même valeur pourquoi ça n'a pas la même valeur d'ailleurs? les records du monde sur pourtant la même distance ou la même discipline en intérieur ou en extérieur?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
263	ESLO2ECOLE1276C	35-55	'alors y a le ciseau ça c'est la plus ancienne effectivement on verra le ciseau pourquoi bah oui effectivement y a le saut ventral à enroulement autour du fil enfin ou la ou la perche et puis y a ce qu'on appelle le saut dans le dos vous voyez aujourd'hui tous ceux qui sautent alors vous ça on essaiera mais c'est très difficile c'est très difficile chut ça s'appelle le Fosbury-flop pourquoi? flop veut dire saut hein et Fosbury parce que c'est lui qui a fait ça le premier il s'appelait euh Fosbury monsieur Fosbury Fosbury il a d'ailleurs gagné avec ça les Jeux Olympiques de Mexico en mille neuf cent soixante-huit donc c'est avant mille neuf cent soixante-huit qu'on ne sautait jamais comme ça c'est lui qui a inventé ça et ça a permis de gagner quelques centimètres supplémentaires à la technique précédente'
264	ESLO2ECOLE1276C	35-55	tu es sûr? tu étais où?

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
265	ESLO2ECOLE1278C	35-55	<p>'non toutes les lettres ça serait pas mal Alexandre deux M Arthur quand tu auras fini de faire le clown tu corrigeras la totalité de l'exercice il est où? et tu le mets où? Nathan tu viens tu peux me montrer où tu as mis les verbes à l'infinif à la correction? aucun aucun la remise en route est difficile Lisa tu as été capable de corriger des choses qui existaient même pas moi là quand j'ai corrigé est -ce que j'ai mis les verbes conjugués? au tableau? non alors explique -moi pourquoi ils sont pas comme je les ai écrits au tableau ceux que tu as corrigés y a des mystères que je m'explique pas hein Théo tu soulignes de temps en temps puis tu tires un trait c'est pas compliqué il manque des lettres Arthur moi j'ai écrit ça comme ça? sûrement pas même dans la correction tu es pas bon hors de la ville pas souligné CCL arriver ça veut dire quoi? moi j'ai écrit au tableau complément circonstanciel de lieu CCL du verbe arret- arriver ou arrêter plutôt nulle part tu as marqué du verbe c'est quoi deux points? ça veut rien dire y a rien de souligné les verbes à l'infinif c'est en option Vincent NPERS tu es plus au ski hein? donc il va peut-être falloir arrêter de faire tout schuss c'est écrit je me demande comment tu as écrit avec les pieds? bon les accents j'en ai pas pas un les lignes c'est en option je suis pas sûr vraiment que l'ophtalmologiste que tu as été voir ne soit pas un charlatan parce qu'il dit que tu as plus besoin de lunettes mais tu écris na- nulle part nulle part sur les lignes je m'inquiète donne -nous l'adresse pour qu'on n'y aille pas Maxime la colle est en option? et tu souligneras Juliette le texte il est où? hé oui et tu as même corrigé des choses qui étaient justes non pour l'instant je veux pas te voir tu viendras après Grégoire pourquoi tu as pas rayé ce qui était une erreur? tu as quelque chose à rajouter? laisse-le tranquille Maxence allez recule souligne ra- tu tu tu rayes quand c'est faux à part ça je l'ai fait lou- euh Loïs alors toi tu as le pompon hein tu as écrit avec le pied gauche? tu as pas rayé ce qui était une erreur arbre H A B R E le R le H tu as des lettres en plus des lettres en moins tu coupes ça les mots au bout de la ligne n'importe comment tu sais pas que c'est par syllabe? bah alors? on sait pas euh c'est corrigé vraiment tu rayes pas les erreurs c'est vraiment n'importe quoi hein alors j'espère qu'aujourd'hui ça sera mieux euh les autres euh qui ont leur cahier euh là-bas tu peux les distribuer vous mettez la date mardi huit mars merci mademoiselle'</p>
266	ESLO2ECOLE1278C	35-55	<p>'toutes ces journées bah intéressantes certes parce qu'elles attirent l'attention sur différents problèmes mais normalement c'est toute l'année que ça devrait être euh euh qu'on s'en soucie normalement euh pourquoi on a in- on a inventé une journée des femmes comme ça? chut chut chut'</p>

Item	Reference	Age	Extraction
267	ESLO2ECOLE1278C	35-55	le vendeur lui reprend son ancien appareil pour cent quatre euros et les accessoires qui le complètent pour trente-quatre euros quelle somme Monsieur Portrait devra-t-il ajouter? le vendeur accepte un paiement en trois versements quel sera le montant de chaque versement? Vincent NPERS
268	ESLO2ECOLE1278C	35-55	quatre cent un on va simplifier on arrondit à combien?
269	ESLO2ECOLE1278C	35-55	tu vas avoir au moins combien?
270	ESLO2ECOLE1278C	35-55	'trois cents donc tu auras au moins cent même ça plus que ça hein? un petit peu plus donc quand tu me donnes le résultat en me disant combien tu m"as dit?'
271	ESLO2ECOLE1278C	35-55	'hein donc je pense qu"y a une erreur là y a quoi? tu es un train de faire quoi toi? tu as fait quoi? comment? alors tu penses qu"il y a quoi là comme erreur?'
272	ESLO2ECOLE1278C	35-55	tu as trouvé combien? cent trente-trois
273	ESLO2ECOLE1278C	35-55	tu as divisé quoi?
274	ESLO2ECOLE1278C	35-55	ah le prix normal mais tu as pas tenu compte du rachat hein de la euh de ce que lui donnait le vendeur contre son ancien appareil dans quelles circonstances on a aussi ce genre de calcul un peu identique? dans autres circonstances commerciales
275	ESLO2ECOLE1278C	35-55	'et comme ça tu l"aurais fait eh bah hier tu aurais posé aux k- aux copains la question oh dis donc euh j"ai essayé de le faire pendant les vacances j"ai pas trouvé tu peux me donner un renseignement? vous restez tous les trois en récréation les touristes ça va bien ta feuille elle est où toi?'
276	ESLO2ECOLE1278C	35-55	'oui Jean sortit quoi? la camionnette c"est bien l"objet de la sortie y a t -il des compléments circonstanciels?'
277	ESLO2ECOLE1278C	35-55	'temps d"accord du verbe sortir il l"a sortie quand? de bonne heure'
278	ESLO2ECOLE1278C	35-55	'complément circonstanciel de but du verbe sortir il l"a sortie pour quoi? dans quel but? pour aller au marché j"ai invité des camarades avec la permission de mes parents'
279	ESLO2ECOLE1278C	35-55	'des camarades je suis d"accord j"ai invité qui? des camarades y a-t -il un complément circonstanciel?'
280	ESLO2ECOLE1278C	35-55	une tarte aux fruits hein elle fait quoi une tarte aux fruits y a t -il un complément circonstanciel Léo?
281	ESLO2ECOLE1278C	35-55	'de temps hein quand est -ce qu"elle fait euh la tarte? eh bien chaque dimanche tout à fait bien merci vous avez révisé tout toutes vos règles pendant les vacances'
282	ESLO2ECOLE1278C	35-55	'non pourquoi on utilise l"un et pourquoi on utilise l"autre?'
283	ESLO2ECOLE1278C	35-55	'je vais t"expliquer pourquoi qu"est -ce que tu vois?'
284	ESLO2ECOLE1279C	35-55	Anouk lequel tu dis qui était trop bien?
285	ESLO2ECOLE1279C	35-55	'tu me le rapportes tu me l'amènes jeudi jeudi jeudi euh Marion alors attends tu as quoi là?'
286	ESLO2ECOLE1279C	35-55	pourquoi Billy Bouledogue est -il un mauvais veilleur de nuit?
287	ESLO2ECOLE1279C	35-55	pourquoi Billy Bouledogue est -il un v- mauvais veilleur de nuit?

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
288	ESLO2ECOLE1279C	35-55	'pourquoi en fait pourquoi il dit que c'est un mauvais veilleur de nuit?'
289	ESLO2ECOLE1279C	35-55	'Sherlock Heml"Os lui donne un coup de coude au guichet voisin Bruno Boxer et Larry Lévrier sont prêts à se battre Sherlock empoche les billets qu'il vient d'acheter et suivi de son fidèle ami s'empresse d'aller voir ce qu'il se passe en l'apercevant Charlie Caniche pousse un soupir de soulagement Sherlock Heml"Os dieu soit loué vous allez pouvoir régler ce problème de quoi s'agit -il? demande Sherlock Charlie lui explique que Bruno Boxer et Larry Lévrier ont exactement les mêmes valises'
290	ESLO2ECOLE1279C	35-55	'chut chut chut chut ont exactement les mêmes valises Larry accuse Bruno de vouloir lui voler sa valise et Bruno affirme que c'est la sienne et pour quelle raison Bruno voudrait -il voler ta valise Larry?'
291	ESLO2ECOLE1279C	35-55	'tu vas me laisser finir je vais t'interroger tout de suite hein pourquoi Sherlock Heml"Os a a-t -il accusé Fredo Frisé plutôt que Chico le ch- le chauve?'
292	ESLO2ECOLE1280C	35-55	'c'est gentil donc six dans quel là tu as deux L deux P pourquoi tu m'enlèves pas deux P? c'était la bonne question pas la bonne opération on peut calculer'
293	ESLO2ECOLE1280C	35-55	'en effet ceux qui étaient derrière avaient raison allô Madeleine? pourquoi j'ai soixante? faut que tu aies le même résultat là ou alors après tu vas mettre zéro virgule vingt et un mètre là effectivement deux c'est pas'
294	ESLO2ECOLE1280C	35-55	'ça pourquoi tu le mets pas là? c'est bien le même qui parle là c'est le même qui parle alors si tu me mets mais la dictée il va la recopier y a un problème ah d'accord la dictée oralement d'accord bah non j'ai pas mon pareil là euh il va à la photocopie dans la cours des ah ah ou- ah oui oui ah tu vas la photocopier d'accord non je veux'
295	ESLO2ECOLE1280C	35-55	et la récréation sert à quoi?
296	ESLO2ECOLE1280C	35-55	'tiens si déjà numéro six cahier de brouillon il me reste qui sur le contrôle là? euh c'est quand vous voulez les deux loulous là on va pas mettre le deuxième zéro saute une ligne sous le calcul oui sauf que là moi je le mettrais pas là alors que tu la poses comme ça c'est pas un problème attention écris'
297	ESLO2ECOLE1280C	35-55	si tu avais deux kilos de bonbons tu ferais quoi comme opération? réfléchis à cette question et ça te donnera la réponse
298	ESLO2ECOLE1280C	35-55	'tu veux acheter moins d'un kilogramme bah ça change rien Johanna rentre ta langue tu as pas sauté de ligne je corrige pas c'est quoi la phrase?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
299	ESLO2ECOLE1280C	35-55	'alors? tu pourrais me la? tu as la réponse bah tu f- tu la refais tu refais ton texte en sautant des lignes qu"est -ce que tu veux que je corrige? tu suis pas les consignes non et alors lequel? on va reprendre bon bah c"est bien c"est la bonne opération mais y a une erreur de calcul jeune homme par contre j"ai un problème de sens oui?'
300	ESLO2ECOLE1280C	35-55	tu en es où euh Robert?
301	ESLO2ECOLE1280C	35-55	'ah de il faut combien de temps pour sortir le cahier? c"est un peu long hein c"est que tu me fais? toi tu en es où sur mes bonbons?'
302	ESLO2ECOLE1280C	35-55	'euh Loïc tu veux quoi là? raconter quoi ta vie? j"ai corrigé'
303	ESLO2ECOLE1280C	35-55	'ah qu"est -ce que j"avais donné comme consigne? on ouvre comment? avec quoi?'
304	ESLO2ECOLE1280C	35-55	'oui bon c"était bon quoi? ça donne ça? bon c"est bon? non c"est bon alors c"est bon alors maison tu préfères laquelle?'
305	ESLO2ECOLE1284C	35-55	comment je sais moi après?
306	ESLO2ECOLE1284C	35-55	oui alors quel âge il avait Louis treize? en mille six cent dix-sept
307	ESLO2ECOLE1284C	35-55	'euh sous deux portraits alors c"est un petit peu particulier vous le voyez en pied hein euh c"est-à-dire qu"on le voit dans la totalité euh et vous le voyez à côté avec euh un portrait euh euh vu de face euh enfin de trois quarts de profil des deux côtés euh c"était euh une technique une habitude de l"époque et puis vous avez euh Louis treize au-dessus hein juste une petite chose oh vous voyez là là il est habillé en en seigneur de guerre puisque comment on peut savoir qu"il est en seigneur de guerre d"ailleurs?'
308	ESLO2ECOLE1284C	35-55	'presque la moitié parce que la moitié c"est combien de pourcents?'
309	ESLO2ECOLE1284C	35-55	'donc y a beaucoup moins de ports si l"on voulait comparer eh bien on pourrait comparer en pourcents c"est-à-dire que pour cent kilos de poissons qui arrivent en Belgique y en a combien qui arrivent dans ce port -là? vous voyez ça permettrait d"avoir des pourcentages de comparer oui?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
310	ESLO2ECOLE1284C	35-55	'd'accord? bien y y a une espèce de barrière c'est des espèces de frontière dans lesquelles vont intervenir euh la police et cetera et puis dépassée cette euh frontière qu'on ne voit pas bien sûr puisque c'est un nombre de kilomètres par rapport à la côte eh bien c'est ce qu'on appelle les eaux internationales donc bah par exemple ici peuvent pêcher aussi bien des Belges que des Espagnols des Portugais des Français des ce que tu veux hein alors évidemment euh quand tu arrives dans une zone ici les deux zones se touchent peut-être mais ici bah c'est international par exemple bon donc euh est -ce que vous comprenez pourcentage c'est pour facilement comparer les choses donc on va voir hein que alors après on a des comp on a décomposé un quarante-deux virgule trois y en a plus sur la côte sud que sur la côte nord puisque y en a trente-trois virgule euh trente-trois pour cent pardon sur la côte sud mais c'est normal parce que regardez combien de ports sur la Bretagne sud hein euh regardez sur la Bretagne sud vous avez le Douarnenez le Guilvinec euh Concarneau Lorient hein y en a quatre alors que la Bretagne du nord ben à part Saint-Malo euh c'est tout ce qu'y a hein donc euh et ça ne fait que neuf virgule trois mais c'est la Bretagne en général bon la deuxième place c'est qui?'
311	ESLO2ECOLE1284C	35-55	'qui traversent ces cubes qui forment les limites de ces cubes et lorsque l'on a deux cubes on a ça correspond à quelle ligne?'
312	ESLO2ECOLE1284C	35-55	on fait en quelle couleur?
313	ESLO2ECOLE1284C	35-55	'je parle bien des poissons frais hein pas les poissons congelés hein donc ici c'est logique il y a c'est le le plus proche ou c'est des moyens de déplacement regardez ici on a les autoroutes qu'est -ce qu'on a qui n'est pas obligatoirement desservi ici? c'est le'
314	ESLO2ECOLE1284C	35-55	'c'est un peu c'est théorique c'est pas parce que quand vous achetez du poisson à Saint-Hilaire que normalement le poisson il viendrait d'où?'
315	ESLO2ECOLE1284C	35-55	'oh tu as le cahier du jour? comment tu as pu le faire signer puisque je l'avais ramassé pour corriger?'
316	ESLO2ECOLE1284C	35-55	'eh je pense qu'il va venir vous le distribuer vous allez vous asseoir autrement je vais vite supprimer ça c'est quoi cette façon de faire? hein? Baptiste et Léa tu te retrouves là là-bas on a l'habitude d'en distribuer à certains et pas à d'autres?'
317	ESLO2ECOLE1292C	35-55	'ouais c'est en train voilà voilà donc vous voyez tout le schéma que fait l'eau vous voyez que l'eau elle parcourt de très nombreux méandres un labyrinthe important pour pouvoir au fur et à mesure s'évaporer l'eau s'évapore et elle devient de plus en plus salée pourquoi tu as cliqué? sur quoi?'
318	ESLO2ECOLE1292C	35-55	'vous avez suivi avec votre doigt le parcours de l'eau à quoi à quel animal ça vous fait penser?'

Item	Reference	Age	Extraction
319	ESLO2ECOLE1292C	35-55	si vous regardez bien ça ça rentre alors comment on appelle ça comme ça ?
320	ESLO2ECOLE1292C	35-55	'ouais c'est la taille est pas du tout la même voyez c'est les par exemple ici vous voyez hein tiens regarde là tiens regarde y a le tas là tu vois ? avec le le convoyeur bien chut bien donc ça c'est l'exploitation alors y a une image où oui ?
321	ESLO2ECOLE1292C	35-55	bon vous allez bien les filles ? eh vous allez voir que les huîtres dans les images elles sont cultivées comment ? enfin cultivées élevées dans des espèces de sacs effectivement et ces sacs sont placés où ? ces sacs
322	ESLO2ECOLE1292C	35-55	bon vous allez bien les filles ? eh vous allez voir que les huîtres dans les images elles sont cultivées comment ? enfin cultivées élevées dans des espèces de sacs effectivement et ces sacs sont placés où ? ces sacs
323	ESLO2ECOLE1292C	35-55	'dans l'eau mais où dans l'eau ? il manque quelque chose comme élément oui au-dessus de l'eau pas toujours parce qu'elles trempent dans l'eau à marée haute à marée basse elles sont en dehors de l'eau à marée haute alors j'attends la réponse non oui voilà chut mais non ils brûlent pas des huîtres ça a rien à voir alors nos sacs d'huîtres ils sont sur quoi ?
324	ESLO2ECOLE1292C	35-55	'voilà allô ? oui bonjour une demi-heure de retard lorsque j'ai prévu à quelle heure ? donc dix-huit heures trente à tout à l'heure'
325	ESLO2ECOLE1299C	35-55	il nous manque qui ?
326	ESLO2ECOLE1299C	35-55	'ça ça va pas être obligatoirement des difficultés mathématiques des des problèmes euh très très compliqués au niveau de l'opération c'est plutôt le raisonnement parce que souvent tous là vous avez des soucis de raisonnement pas de technique hein vous savez compter vous connaissez vos tables ou euh c'est plutôt vous organiser dans qu'est -ce qui est important qu'est -ce qui est pas important dans les informations qu'est -ce qu'il faut chercher hein donc on va travailler sur ce domaine -là et éviter de se perdre dans les vingt-deux résolutions hein Grégoire que tu des fois tu me proposes donc aujourd'hui et et jeudi on va travailler sur les seize premiers exercices qui est de trouver la question hein parce que souvent on sait pas trop ce qu'on cherche donc ce sont des solutions ce sont des phrases simples donc un dictionnaire totalise mille quatre-vingt huit pages la nouvelle édition compte quatre-vingts pages de plus question qu'est -ce qu'on peut trouver ? hein c'est c'est la question qu'est -ce qu'on pose et bien sûr vous le faites mais euh vous le résolvez mais c'est pas compliqué'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
327	ESLO2ECOLE1299C	35-55	'on va faire la phrase in- interrogative une question un point d"interrogation allez tu me mets un résultat et tu m"en mets un autre pourquoi l"opération atterrit là? pourquoi elle est pas en face? et maintenant'
328	ESLO2ECOLE1299C	35-55	'on va faire la phrase in- interrogative une question un point d"interrogation allez tu me mets un résultat et tu m"en mets un autre pourquoi l"opération atterrit là? pourquoi elle est pas en face? et maintenant'
329	ESLO2ECOLE1299C	35-55	et ensuite pourquoi tu rajoutes un T?
330	ESLO2ECOLE1299C	35-55	alors euh où est -ce que tu as trouvé ça? il est pas dans le texte de départ
331	ESLO2ECOLE1299C	35-55	'la question c"est combien les six pots font de kilogrammes? moi je veux bien'
332	ESLO2ECOLE1299C	35-55	'j"ai dit Tom j"ai dit quoi? ah je suis désolé ça ça peut pas être le résultat de ça le résultat de ça c"est ça'
333	ESLO2ECOLE1299C	35-55	comment tu passes de ça à ça?
334	ESLO2ECOLE1299C	35-55	'oui mais tu m"as écrit quoi là?'
335	ESLO2ECOLE1299C	35-55	'comment tu l"as écrit?'
336	ESLO2ECOLE1299C	35-55	où est -ce que tu me trouves ça?
337	ESLO2ECOLE1299C	35-55	'mais c"était pas ce que tu m"avais ça dépend il manque quelque chose oui panneua suivant oui mais tu aurais pu très bien le dire regarde et ça ça m"expliquait où est -ce que tu l"avais trouvé alors tu m"as solutionné le problème où là? ah tu m"as mis une virgule là? nenni dis -moi s- six fois ça ça fait bien ça mais ça me donne pas ça comment tu passes de ça à ça? d"un coup de baguette magique?'
338	ESLO2ECOLE1299C	35-55	'oh c"est quoi ça?'
339	ESLO2ECOLE1300C	35-55	'soixante là allez allez me chercher les plots qui sont là s"il vous plaît Charlie tu as quoi comme problème? c"est arrivé tu as pas été adroit alors fais voir mon dieu vas mettre un peu d"eau froide dessus'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
340	ESLO2ECOLE1300C	35-55	'j"ai fait exprès de mettre les plots pour vous o- obliger à tourner vous comprenez? hein pour longer le fil donc là vous commencez à votre coin à la- à l"autre coin vous avez droit à trois essais Christopher tu es où? voilà aller vas -y soixante centimètres donc là c"était l"échauffement soixante-dix hé touché deuxième essai hé tu es venu sauter ici donc tu pouvais pas tendre tes jambes faut que tu sautes là non non deuxième essai soixante-dix soixante-quinze très bien c"est bon ah touché deuxième essai c"est bon attends attends attends c"est bon c"est bon euh ceux qui me font du saut en longueur mor- mordu je veux qu'on arrive les deux pieds en avant quatre-vingts deuxième essai deuxième essai quatre-vingts ouais quatre-vingt-cinq hé hé touché deuxième essai touché deuxième essai touché deuxième essai c"est bon deuxième essai touché troisième essai c"est bon troisième essai troisième essai oui ah non donc quatre-vingts toi tu notes quatre-vingts tu restes là tu restes en face tu vas lever et on met un quatre-vingt-dix m- mets -toi en face sur le poteau et tu montes le fil jusqu"à quatre-vingt-dix la f- la ficelle jusqu"à quatre-vingt-dix voilà impeccable hein et puis tu restes là pour regarder allez ouais non deuxième regarde toujours s"il est à quatre-vingt-dix deuxième ouais non touché oui c"est bon ah refus de l"obstacle là c"est bon refus de l"obstacle troisième essai éliminé tu restes là au poteau hé viens là donc toi tu as réussi quatre-vingt-cinq on met à quatre-vingt-quinze tu vois là ici on lit quatre-vingt-quinze quatre-vingt-quinze non regardez si il est bien toujours à quatre-vingt-quinze refus de l"obstacle deuxième essai non non mais attends laisse passer les autres repose -toi ah ça commence à des hauteurs là ah non ah non y a un pied mais pas l"autre deuxième essai non non tu as les fesses qui retombent dessus faut te couper les fesses faut faire quelque chose ah non tu as les fesses qui retombent dessus bah regarde tu as vu comment il bouge ah non ah non tu as la deuxième jambe qui veut pas passer tu vois elle est prisonnière troisième et dernier essai ah non ah non touché hein non fini ah oui c"est bon oui deux quatr- alors on passe à quatre-vingt-quinze? non hein'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
341	ESLO2ECOLE1300C	35-55	'oui oui non touché troisième essai oui oui non éliminé quatre-vingt-dix tu viens tenir là oh oh euh score quatre-vingt réussite quatre-vingt tu viens tenir ici je te mets à quatre-vingt-dix là tu vois il est à quatre-vingt-dix tu vois et toi tu tiens là-bas quatre-vingt-dix ouais ouais non deuxième essai oui bah arrête arrête arrête non deuxième essai vérifiez bien que c'est à quatre-vingt-dix deuxième essai quatre-vingt-dix non deuxième essai quatre-vingt-dix non troisième et dernier essai Christopher euh tu descends de là non éliminé quatre-vingt-cinq la réussite non éliminé quatre-vingt-cinq la réussite vous mettez à quatre-vingt-quinze y en a plus que deux allez les deux à quatre-vingt-quinze un mètre ah non un mètre zéro cinq si besoin est non premier essai euh non touché deuxième essai Tom tout de suite non troisième Thomas non troisième et dernier non éliminé alors la meilleur score c'est quatre-vingt-dix vas -y non quatre-vingt-dix tu baisses le fil à au ras du sol euh vont aller se rhabiller le groupe des filles ceux qui étaient là-bas aux haies vous allez vous rhabiller euh baissez -moi cette haie les deux groupes qui sont passés au triple saut vous restez là parce que vous allez ceux qui sont passés au triple saut vous restez je vous attends je t'ai dit en-bas parce que tu as pas vu que c'est en train de tirer sur le poteau que ça va tomber tu es quand même hein hop vous me montrez tous un saut de triple saut puisque vous vous êtes entraînés vas -y Samuel montre -moi pour voir juste le geste je te demande pas une performance juste le geste non là pourquoi tu me fais recommence parce que ça m'a pas l'air très net ton affaire re- vas -y après bah pourquoi tu arrives à deux pieds? c'est trop loin c'est que vous êtes trop loin si vous partez de la ligne blanche c'est sûr allez vas -y ouais c'est bon vous allez pouvoir vous rhabiller tous les deux non tu as fait un tu as pas fait de cloche-pied au départ non tu as pas fait de cloche-pied oui alors t- euh elle est où la bé- là là vous prenez la ligne blanche ou la ligne bleue?'
342	ESLO2ENT1005C	35-55	'donc euh quand on peut on s'arrange et puis bah je fais tout ce qu'on font les mamans euh à la maison donc euh le rangement le linge le ménage les courses la cuisine qu'est -ce que je vous fais encore?'
343	ESLO2ENT1008C	35-55	'et puis c'était quoi l'autre truc que tu m'a demandé?'
344	ESLO2ENT1008C	35-55	'qu'est -ce que est -ce que tu aimes l'omelette?'
345	ESLO2ENT1008C	35-55	'ouais mais comment tu aimes l'omelette? nature euh cuite euh'
346	ESLO2ENT1012C	35-55	'ah ouais ben ouais j'ai même pas regardé la truc de clôture c'était quand cette nuit non?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
347	ESLO2ENT1012C	35-55	'ouais ah ben moi je voilà euh à droite à gauche moi vous avez quoi en ce moment? voilà puis euh en fonction de y a toujours des d'aut- y a toujours plein d'autres gens qui achètent des DVD qui téléchargent des films donc voilà moi je le fais jamais moi-même y en a toujours qui sont assez cons'
348	ESLO2ENT1012C	35-55	Kesk JE FAIS PENDANT MES quel type de tourisme je pratique?
349	ESLO2ENT1012C	35-55	Kesk JE FAIS PENDANT MES quel type de tourisme je pratique?
350	ESLO2ENT1012C	35-55	'enfin je veux dire c'est assez facile de trouver des thématiques et à part le Festival de Jazz c'était quoi l'autre truc?'
351	ESLO2ENT1012C	35-55	'en précisant que lui n'était pas prof d'université et que d'être maire pour lui c'était un travail à plein temps et qu'il faisait ce travail qu'à plein temps et il a rajouté comme argument et là pour c'est impossible à entendre qu'il était quand même père de cinq enfants donc là moi j'en ai rien à foutre qu'il en ait un ou dix ou cinquante pour moi ça justifie pas que on puisse multiplier par deux son salaire ça c'est une première chose et après donc il disait que l'autre sal- le le d'être Maire sur une ville comme Orléans c'était un un truc à temps plein et peu de temps après il s'est quand même pré- pré- présenté à la députation il a été élu député ah alors ça je trouve ça quand même formidable hein quand le le le quand on peut admettre qu'une ville aussi grande qu'Orléans nous prend autant de temps ce que je peux concevoir parce que c'est un vrai boulot enfin pour moi c'est un vrai boulot d'être maire d'une aussi grande ville c'est sûr que d'être euh maire d'une ville de cent vingt habitants c'est quand même pas la même chose que d'être maire d'une ville qu'en fait beaucoup plus qu'en fait combien de fois plus euh? combien il faut que je rajoute de zéro? trois quatre'
352	ESLO2ENT1016C	35-55	'vous avez le long de comment ça s'appelle à côté de Jargeau? Saint-Denis- de-l'Hôtel'
353	ESLO2ENT1020C	35-55	'euh comment ça s'appelle? euh ah si on a entendu des rumeurs sur notre euh sur notre maire'
354	ESLO2ENT1020C	35-55	'euh savoir que après si on met l'imparfait et non pas euh comment s'appelle? un p- un conditionnel présent hein'
355	ESLO2ENT1020C	35-55	pourquoi elle pleure là?
356	ESLO2ENT1020C	35-55	pourquoi il pleure?
357	ESLO2ENT1021C	35-55	ça v- ça signifie quoi Lutèce?

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
358	ESLO2ENT1027C	35-55	'c'est vrai que au niveau artisanat euh artisans y a pas non non non y a plus grand monde y a plus grand monde bah et moi je vois c'est c'est marrant quand j'étais à l'école et quand je m'étais fixé entre guillemets de de reprendre un petit peu le magasin mais j'étais manuel donc je voulais faire une filière manuelle donc on m'a tout de suite dirigé vers un un Bac mais moi je dis non je je veux pas faire un Bac moi je veux faire un BEP euh mécanicien monteur on m'a dit mais ça va pas qu'est -ce qui bah mon petit gars qu'est -ce que tu qu'est -ce qu'il se passe? tu es à la limite tu es malade euh v- bah je dis si moi je je veux faire du technique je veux bah euh non non euh mais s- si si et je me suis battu moi pour euh aller en BEP alors qu'on voulait me passer euh en en en première quoi et c'était incroyable'
359	ESLO2ENT1027C	35-55	'je dis tiens on mettrait bien une euh petite pergola ou un store ouais hop je tape s- store et puis euh voilà je après c'est euh on part où en vacances? oh bah je sais pas on pourrait essayer là bah là on est sur la Corse donc euh bah Corse alors je voyage en Corse en ce moment voilà je je préfère aller visiter des sites comme ça aller dans les forums voir euh les dialogues euh voilà ça c'est c'est quelque chose de sympa on apprend beaucoup de choses et puis euh une heure sur le le PC c'est c'est rien hein quand vous l'allumez vous allez voir vos mails vous répondez'
360	ESLO2ENT1029C	35-55	'il a une carte de handicap et tout hein vous pouvez appeler qui vous voulez la police tout ça personne ne fait rien tout le monde s'en fout je dis pas que la police s'en fout hein la police nous dit voilà monsieur brigade des mineurs sur Orléans nous sommes quatre neuf heures du matin dix-huit heures trente y a cent-cinquante gamins en moyenne à chercher par semaine sur toute l'agglomération comment on fait?'
361	ESLO2ENT1029C	35-55	'ça devient euh donc c'est ça hein quand je vois les là les pauvres parents euh je m-j'imagine ça tout à fait le nombre de fois où j'ai eu ce type de problèmes où j'ai appelé l'assistante sociale le conseil général le machin puis on vous dit bah qu'est -ce qu'on peut faire? alors bah si on vous dit bah nous on peut rien faire bon on va pas vous l'écrire hein'
362	ESLO2ENT1029C	35-55	'y a eu un très bon rapport y a dix ans pourquoi on l'a pas appliqué? pourquoi on est obligé d'en refaire un autre pour arriver au même résultat on va arriver au même résultat qui est côté jeune public marionnettes tout ça il n'y a pas de moyens pas de moyens pas de moyens pas de moyens et bon'
363	ESLO2ENT1031C	35-55	'c'est vrai qu'on était dans un quartier qui était pas euh comment je pourrais dire ça? qui qui qui se dégradait dans le temps'
364	ESLO2ENT1033C	35-55	'c'était quoi cette rumeur?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
365	ESLO2ENT1034C	35-55	'et donc en fait euh y a le hangar qui reste euh qu"à d- qu"à dû être une construction euh on va dire plus récente et y avait un y avait un un comment s"appelle? un pour faire de l'alcool un oh zut'
366	ESLO2ENT1034C	35-55	'euh les premiers boulots que j"ai trouvé sur Orléans c"était quoi? c"était quatre-vingt quatre-vingt onze quatre-vingt douze quatre-vingt treize de nouveau sur Orléans puis après quatre-vingt quatorze j"ai été embauchée en CD en contrat à durée déterminée'
367	ESLO2ENT1034C	35-55	'en Normandie l"autre à Pau euh une à Paris euh une dans le Cher qui est du coup pas très loin de chez mes parents une ici euh qui est -ce que j"ai oublié? et c"est tout c"est bon'
368	ESLO2ENT1034C	35-55	'euh mai- p- par contre ma mère est issue de enfin elle était euh issue elle est née en Normandie et ses parents euh sont partis son père a voulu s"installer pendant euh euh pendant la crise euh c"était laquelle? vingt-neuf où y a eu des euh des gros euh des gros cracks boursier et euh il s"est installé à cette époque -là ce qui fait qu"en fait au moment où il s"est installé il a acheté un truc et après il a fallu qu"il rembourse plus cher donc il il s"est très rapidement euh cassé la figure'
369	ESLO2ENT1034C	35-55	'donc euh toute une grande partie du de Saint Marceau qui doit être située entre l"avenue Dauphine en dessous ça doit être euh la rue de la Cigogne et puis euh euh ce qui remonte de la nationale vingt comment elle s"appelle? l"avenue Secrétain je crois ou Secrétain je sais plus comme ça s"écrit'
370	ESLO2ENT1034C	35-55	'mais pourquoi i- pourquoi traitent -ils? parce que bon c"est euh c- c"est quelque chose normalement euh c"est pas qu"ils ont suffisamment de monde c"est que normalement ils doivent toujours on doit pouvoir trouver des solutions pour qu"ils aient pas besoin de traiter et que ça leur prennent pas plus de temps faut simplement réfléchir différemment'
371	ESLO2ENT1034C	35-55	'je vois ce qu"il en est si ça ri- s- si y a un risque si y a pas de risque et j"envoie aux différentes personnes qui s"occupent des différentes cultures concernées qu"est -ce que je fais d"autre?'
372	ESLO2ENT1034C	35-55	'parce que le problème c"est que pour avoir une chance euh on en a laissé tombé combien?'
373	ESLO2ENT1034C	35-55	'quoi qu"est -ce tu dis?'
374	ESLO2ENT1034C	35-55	'ah oui au niveau de la région y a plein plein de choses euh y a tous les fromages euh s- si y a les fromages y a le cendré mais il est encore pas d"Orléans y a le cendré de Pithiviers et puis euh y a le comment il s"appelle? euh euh la camembert aux herbes de de Pithiviers'
375	ESLO2ENT1040C	35-55	'une chambre d"hôtel à Paris ça coû- une chambre à Paris ça coûte combien?'
376	ESLO2ENT1040C	35-55	pourquoi elle est saugrenue?

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
377	ESLO2ENT1043C	35-55	'je pense qu"il y a de tout mais oui oui oui enfin bien bien ça veut dire quoi un bon français c"est ça?'
378	ESLO2ENT1045C	35-55	'et en plus là dernièrement elle fait des étu- des comment ils appellent ça? des euh elle a quelques collégiens qu"elle les fait travailler'
379	ESLO2ENT1049C	35-55	vous vous vous baladez en v- vous êtes venue comment en transport en commun ou en voiture?
380	ESLO2ENT1049C	35-55	en voiture vous êtes passée par où?
381	ESLO2ENT1049C	35-55	'beaucoup de gens qui sont entrés à l"Insee euh par hasard moi j"y suis entré par choix euh j"ai décidé que ça ça m"intéressait euh et grâce notamment à ma professeure d"économie qui s"appelait madame NPERS à l"époque euh et qui m"a donné le goût de ces choses -là euh voilà euh j"ai fait une euh seconde euh euh comment ça s"appelait à l"époque? euh j"ai fait un Bac B mais la seconde qui y conduisait c"est une seconde AB'
382	ESLO2ENT1049C	35-55	'euh que ils nous ont lâché qu"il y aurait pas alors qu"il aurait fallu qu'on continue mais je pense que voilà euh et euh ils ont agi euh en gens responsables selon leur sens à eux et pas selon notre sens à nous euh et puis euh voilà donc j"ai un peu de mal avec les y compris les élites syndicales je suis pas sûr qu"elles nous représentent toujours bien en même temps quand Bernard Thibault est venu à Orléans dernièrement je sais pas si vous avez suivi il est venu inaugurer la nouvelle euh comment ça s"appelle? zut qui a été déplacée qui était quais des Auguch- quais de Cypierre qu- je sais plus bah qui était sur les quais avant quais des Augustins ou quais Cypierre je sais jamais euh la bourse du travail voilà qui a été déplacée'
383	ESLO2ENT1049C	35-55	'donc euh et euh de nos chers gouvernants quand même qui se foutent bien de notre gueule euh et euh hm et donc euh hm qu"est -ce que je voulais dire? euh'
384	ESLO2ENT1049C	35-55	'donc euh bon là ils devraient en rouvrir un d"ici euh vers le mois de septembre mais ça faisait presque deux ans ou deux ans je sais plus si c"est pas plus faudrait v- vérifier auprès des gens des commerçants de La Bolière parce que moi je vois pas le temps passer des fois que y en avait plus que Major enfin que euh comment ça s"appelait à la fin je ne sais même plus moi Stock? oui Stock avait fermé et que voilà y avait plus de supermarché donc euh ce qui est quand même un peu ennuyeux après euh les l- les autres c"est les hard discounts et eux ils z- ils sont fermés le dimanche matin je pense y a plus qu"un hard discount qui se trouve ça doit être le Rally euh qui se trouve à côté de la piscine là par là dans'
385	ESLO2ENT1049C	35-55	'c"est pas Rally comment il s"appelle? non'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
386	ESLO2ENT1049C	35-55	'exactement mais voilà bon euh voilà et ça a fait beaucoup de mal mais ceci dit euh bah hier soir j"ai regardé non c"était quand? qu"est -ce qu"il y a eu? y a eu un téléfilm à la télé là y a pas longtemps sur le service public il me semble mais je suis pas sûr c"était peut-être sur TF un parce que des fois y en a des biens sur TF un aussi euh qui parlait justement de la rumeur et de ses conséquences enfin des et c"était pas mal quoi on voit bien que ça reste toujours euh c"était toujours quelque chose euh et pas à l"époque euh dans laquelle on vit hein euh euh on a quand même euh moi j"ai le sentiment que les gens que globalement les gens vont pas bien et que euh ils ont besoin quand ça quand euh ça les démange comme disait Bedos en quatre vingt-trois euh j"aime beaucoup Bedos quand ça d- enfin lui il parlait du PC à l"époque euh donc des communistes il parlait pas du PC il parlait du communisme ça les démange alors faut qu'ils se grattent sur la peau des autres'
387	ESLO2ENT1049C	35-55	'j"aime pas trop faire de la pub mais j"a- j"a- j'avoue que j"aime beaucoup Nagui comme animateur de jeu euh on regarde aussi euh on regarde aussi euh comment ça s'appelle? quand j"ai le temps j"aime bien regarder Motus parce que je trouve Becca- enfin c'est un jeu'
388	ESLO2ENT1049C	35-55	'c- euh c'est sur ce genre de chose si y en a pas un qui pousse l'autre euh bah voilà c'est que ni l'un ni l'autre ne sont vraiment déterminés bon voilà euh non oui je l'embrassement des tours euh c'est pareil bon alors là en plus qu'est -ce qu"il nous a fait rou- Grouard cette année de quoi il a parlé alors au moment de l'embrassement des tours? non il a parlé d'un sujet sérieux je ne sais plus je sais que ça a été relevé dans les médias mais euh aux locaux au moins je ne sais plus ça vous dit rien'
389	ESLO2ENT1051C	35-55	'pour moi non non non moi je s- quand je suis arrivée là que j'arrivais de l'Est euh j'ai eu j'avais encore des expressions très typiques et qu'on me faisait remarquer parce que ça allait pas c'est quoi ça veut dire quoi?'
390	ESLO2ENT1053C	35-55	'et il c'est entre les deux alors on se dit on on se place où?'
391	ESLO2ENT1053C	35-55	'oui alors c'est quelles euh c'est quelles langues et et et où?'
392	ESLO2ENT1053C	35-55	comment on fait une omelette?
393	ESLO2ENT1053C	35-55	comment je fais une omelette?
394	ESLO2ENT1057C	35-55	'euh en premier c'est lequel qui a fermé? je pense que c'est le Royal qui a dû fermer en premier'
395	ESLO2ENT1057C	35-55	'euh tu viens d'où euh tu es d'où? bah quand ils me disent bah je viens de Saint-Denis'
396	ESLO2ENT1057C	35-55	'euh tu viens d'où euh tu es d'où? bah quand ils me disent bah je viens de Saint-Denis'
397	ESLO2ENT1057C	35-55	'et vous? vous la faites comment vous l'omelette?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
398	ESLO2ENT1060C	35-55	'qui parlaient elles disaient ah il est il est tout coti je dis qu"est -ce que vous racontez?'
399	ESLO2ENT1060C	35-55	'voilà qu"est -ce qui y avait d"autre euh? ah ça y est je m"en souviens plus j"essaye de m"en souvenir j"essaye de les garder'
400	ESLO2ENT1070C	35-55	'voilà et donc c"est donc l- petit homme qui rencontre Dieu et euh et tu imaginais Dieu comment? grand vieux avec une longue barbe blanche un air sévère une tunique bleu ciel mais les mains derrière le dos tout comme toi'
401	ESLO2ENT1076C	35-55	'comment vous êtes arrivée à c- dans ce quartier? euh c"est un c"est un hazard parce que avant ici c"était tout en friche ouais et euh'
402	ESLO2ENT1076C	35-55	'oui euh j"en ai un à la barrière Saint-Marc où est -ce qu"ils sont scolarisés?'
403	ESLO2ENT1076C	35-55	'ouais donc ça euh vous êtes allés? oui on y est allé euh c"était c"était quoi exactement? mon fils a- euh'
404	ESLO2ENT1076C	35-55	'et euh vous avez combien d"ordinateurs?'
405	ESLO2ENT1083C	35-55	'parce que la la cité les gens étaient bizarres ils avaient pas le même langage que tout le monde ils étaient chez eux ils disaient oui je suis chez moi alors qu"on qu"on est en France oui c"est c"est chez nous euh les policiers ils viennent pas i- y a quelque chose dans c"est pas normal et de là j"ai vu que c"était le c"était des descendants des hm comment dire euh? de la guerre des gens qui avaient eu euh des problèmes y avait une communauté qui était euh manouches des gens du voyage qui étaient mélangés c"était des sinistrés les mélanger avec des des gens du voyage j"ai vu qu"y avait plein plein de lacunes plein de choses que je connaissais pas du tout parce que je connaissais pas du tout ce monde -là j"étais enfermée dans mon ma bulle là euh j"ai toujours vécu là et euh là je tour- trouve des gens qui n"arrivent pas à finir le le mois qui ont je sais pas qui ont des soucis qui des choses que qui sont tout tout le temps en prison toute une toute une vie que je connaissais pas des maris qui battent leur femme euh matin et soir des hommes qui arrêtent pas de boire des femmes qui boivent des enfants qui sont laissés à eux-mêmes à l'ex- mais on dit mais je suis où? je voulais y aller mais je suis où? et et en allant et c"est moi qui ai fait euh qui ai déclenché la la chose pour qu"on fasse euh un historique euh et cetera'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
406	ESLO2ENT1083C	35-55	'parce que la la cité les gens étaient bizarres ils avaient pas le même langage que tout le monde ils étaient chez eux ils disaient oui je suis chez moi alors qu'on qu'on est en France oui c'est c'est chez nous euh les policiers ils viennent pas i- y a quelque chose dans c'est pas normal et de là j'ai vu que c'était le c'était des descendants des hm comment dire euh? de la guerre des gens qui avaient eu euh des problèmes y avait une communauté qui était euh manouches des gens du voyage qui étaient mélangés c'était des sinistrés les mélanger avec des des gens du voyage j'ai vu qu'y avait plein plein de lacunes plein de choses que je connaissais pas du tout parce que je connaissais pas du tout ce monde -là j'étais enfermée dans mon ma bulle là euh j'ai toujours vécu là et euh là je tour- trouve des gens qui n'arrivent pas à finir le le mois qui ont je sais pas qui ont des soucis qui des choses que qui sont tout tout le temps en prison toute une toute une vie que je connaissais pas des maris qui battent leur femme euh matin et soir des hommes qui arrêtent pas de boire des femmes qui boivent des enfants qui sont laissés à eux-mêmes à l'ex- mais on dit mais je suis où? je voulais y aller mais je suis où? et et en allant et c'est moi qui ai fait euh qui ai déclenché la la chose pour qu'on fasse euh un historique euh et cetera'
407	ESLO2ENT1083C	35-55	'et c'est là que je là euh où je suis intervenue à la Caf parce que quand ils y allaient ils leur parlaient mais eux ils me disaient mais elle me dit quoi là? comprenaient pas tout'
408	ESLO2ENT1083C	35-55	'cette euh cette question aussi bah ça fait euh ça fait deux jours qu'on t'a pas vue tu étais où? ça c'est bien parce que tu te dis bah j'ai j'ai j'ai une vie oui mais nous on sait pas où tu es non mais c'est c'est c- ça a l'air ça a l'air bête et euh s- les gens ils prennent soin de vous'
409	ESLO2ENT1083C	35-55	'voilà j'ai arrêté ça j'ai arrêté quoi euh? hm y a plein de choses que je fais plus parce que j'ai pas le temps c'est juste ça hein'
410	ESLO2ENT1083C	35-55	'ouais y a toujours un y a toujours un peu plus pour que bien viens manger avec nous ce midi tu manges quoi aujourd'hui? euh je sais pas euh bah écoute euh je fais ça bah tu viens voilà ou bien euh oh j'ai pas envie de faire à manger aujourd'hui oh bah viens euh viens viens euh viens euh ou bien voilà c'est comme ça'
411	ESLO2ENT1083C	35-55	'il s'est suicidé y a combien de temps? cinq mois peut-être ou six mois mais on s'en doutait'
412	ESLO2ENT1085C	35-55	'voilà mais euh moi quand je suis avec quelqu'un en va- en vacances avec mon frère ou ma soeur avant elle était célibataire on es- longtemps partis ensemble donc hop on se balade un petit petit coup à droite un petit coup à gauche un petit coup tout d'un coup on est arrivé à un stop bah d'après toi tu irais où là? hm'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
413	ESLO2ENT1085C	35-55	'des trucs comme ça tout ce qui peut être un peu un peu curieux un peu inattendu un peu bah oui voilà parce que parce que oui des fois oh tu vas faire ça toi? quand j"étais euh que j"allais rallumer la flamme euh au du soldat inconnu bah on m"a dit mais ça va pas mais tu es tarée qu"est -ce qui se passe? je fais attends ça je le referai jamais c"est ça doit être quelque chose'
414	ESLO2ENT1272C	35-55	'pourquoi y a trop de cela? parce que ça marche les si les gens ils avaient vu que bah pourtant euh si les gens ils avaient vu que faire plusieurs opticiens ça marchait même si ils peuvent pas forcément le faire parce que c"est quand même les commerces qui peuvent être euh accessibles c"est des commerces on connaissait lesquels mais même ne serait -ce que je sais pas une une agence de voyages une euh une agence d"assurances'
415	ESLO2ENT1272C	35-55	'qui sont vraiment euh qui marchent dans les quartiers ce qu"il marche c"est la sandiwcherie c"est euh la boucherie voilà ils aiment manger ils aiment ils ont que ça de toute façon qu"est ce que vous voulez faire ils les gens ils vont au boulot ils rentrent ils p- s"ils mangent pas bien si ils vont faire quoi? ils ont ils ont plus rien à faire en fait quoi bah donc ils profitent ils mangent ils et à chaque fois c"est le débat ici c"est de dire bah y a trop de boucheries y a trop de cela y a trop de voilà c"est quoi le débat en ce moment je sais pas si vous en avez entendu parler mais c"est tout le temps ça quoi euh des réunions à tire-larigot pour dire bah y a trop de sandiwcheries y a trop de y a pas assez de commerces blancs entre guillemets'
416	ESLO2ENT1329C	35-55	aux Groues euh oui bon voilà ils disaient bonjour aux jeunes et cetera ils se connaissaient ils avaient un différent mais ça ça empêchait pas de dire bonjour euh voilà euh voilà tu as fait quoi? au fait euh on est venu me dire euh tu étais là et cetera y avait un dialogue
417	ESLO2ENT1329C	35-55	mais vous étiez déjà combien au départ avec tes parents?
418	ESLO2ITI1092C	35-55	en quoi?
419	ESLO2ITI1107C	35-55	'oui euh c"est quoi les questions?'
420	ESLO2ITI1134C	35-55	vous êtes étudiantes en quoi?
421	ESLO2ITI1151C	35-55	mon Dieu vous y allez comment à pied?
422	ESLO2ITI1151C	35-55	vous me donnez quel âge?
423	ESLO2ITI1153C	35-55	vous êtes de quel euh?
424	ESLO2ITI1171C	35-55	'euh comment elle s"appelle cette rue là en face déjà?'
425	ESLO2ITI1174C	35-55	'un petit chemin pour a- pour accéder au tram là vous prenez le tram et vous descendez euh où est -ce que vous allez descendre? euh vous descendez à Orléans centre mais euh c"est comment l"arrêt euh République'
426	ESLO2ITI1174C	35-55	'euh ça s"appelle comment ici? euh'
427	ESLO2ITI1174C	35-55	'j"ai quel âge? quarante-huit ou quarante-neuf? quarante-neuf'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
428	ESLO2ITI1174C	35-55	'c"est quoi ça?'
429	ESLO2REPAS1247C	35-55	hm hm bah oui oui bah elle avait ça où donc? dans la petite chambre?
430	ESLO2REPAS1247C	35-55	elle est où la viande? ah là elle est là parce que je confonds la viande avec les patates moi
431	ESLO2REPAS1247C	35-55	'comment elle s"appelle le docteur?'
432	ESLO2REPAS1247C	35-55	'c"est quoi ça?'
433	ESLO2REPAS1247C	35-55	'c"est quoi ça?'
434	ESLO2REPAS1247C	35-55	'c"est quoi ça?'
435	ESLO2REPAS1247C	35-55	'les Grosses Têtes c"est quand même quel âge il a lui euh?'
436	ESLO2REPAS1247C	35-55	'il a quoi lui un micro c"est quoi son mic- le micro là?'
437	ESLO2REPAS1247C	35-55	'il a quoi lui un micro c"est quoi son mic- le micro là?'
438	ESLO2REPAS1247C	35-55	'hé mais au fait c"est pas mauvais ton c"est quoi c"était quoi ça?'
439	ESLO2REPAS1247C	35-55	'hé mais au fait c"est pas mauvais ton c"est quoi c"était quoi ça?'
440	ESLO2REPAS1253C	35-55	'et euh où est -ce qu"il va lui au collège? au lycée?'
441	ESLO2REPAS1253C	35-55	'c"est quoi ça?'
442	ESLO2REPAS1253C	35-55	'Soren c"est lui que j"ai fait réviser avec toi c"est quoi ce boulot?'
443	ESLO2REPAS1253C	35-55	'c"était sur quoi?'
444	ESLO2REPAS1253C	35-55	'c"est qui cette euh Mathilde c"est Mathilde NPERS?'
445	ESLO2REPAS1253C	35-55	'comment il s"appelle euh?'
446	ESLO2REPAS1253C	35-55	et pourquoi il veut pas vous les donner? il va
447	ESLO2REPAS1253C	35-55	'c"était parce que c"était avec quel prof?'
448	ESLO2REPAS1253C	35-55	tu ferais quoi?'
449	ESLO2REPAS1253C	35-55	'c"est qui ça Amédée?'
450	ESLO2REPAS1253C	35-55	'alors? alors? c"est quel groupe?'
451	ESLO2REPAS1253C	35-55	tu enlèves combien de temps le à midi?'
452	ESLO2REPAS1253C	35-55	'ça c"est quoi ça?'
453	ESLO2REPAS1254C	35-55	combien tu as payé ça?'
454	ESLO2REPAS1254C	35-55	'bah vous ferez ça où vous ferez quoi? il a tu sais il a dit qu"il demanderait une grande surface'
455	ESLO2REPAS1254C	35-55	'bah vous ferez ça où vous ferez quoi? il a tu sais il a dit qu"il demanderait une grande surface'
456	ESLO2REPAS1254C	35-55	'qu"est -ce qu"y a? mais ça si tu as pris de la verte tu as pris du comment ça s"appelle?'
457	ESLO2REPAS1254C	35-55	'qu"est -ce qu"y a? mais ça si tu as pris de la verte tu as pris du comment ça s"appelle?'
458	ESLO2REPAS1254C	35-55	'et c"était quoi comme texte de Martin Luther King?'
459	ESLO2REPAS1254C	35-55	mais il travaille où? enfin il est installé euh
460	ESLO2REPAS1254C	35-55	'avec qui j"ai mangé alors? ah non j"ai mangé avec une fille que j"avais'
461	ESLO2REPAS1254C	35-55	et alors comment tu dé- comment tu définis ce mot?'
462	ESLO2REPAS1254C	35-55	bah tu as eu quoi comme moyenne en histoire depuis le début?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
463	ESLO2REPAS1254C	35-55	'est -ce qu"y a un kiwi pour moi? s"il te plaît Victor tu fais quoi là? remets -toi à table j"ai pas mangé'
464	ESLO2REPAS1263C	35-55	'c"était quoi le principe des familles euh nombreuses? le principe de réduction que vous appliquez'
465	ESLO2REPAS1263C	35-55	'bon c"était quoi finalement euh la sanction enfin la peine euh appliquée?'
466	ESLO2REPAS1263C	35-55	'et tu travailles jusqu"à quelle heure?'
467	ESLO2REPAS1263C	35-55	je suppose tu fais quoi?
468	ESLO2REPAS1263C	35-55	ah oui? et alors vous faites quoi?
469	ESLO2REPAS1263C	35-55	'et le thème c"est quoi?'
470	ESLO2REPAS1264C	35-55	bah quand il la prend ils vont où?
471	ESLO2REPAS1264C	35-55	il fait quoi? il est appelé?
472	ESLO2REPAS1264C	35-55	bah comment tu le sais?
473	ESLO2REPAS1264C	35-55	bah pourquoi y a un trou là-dedans?
474	ESLO2REPAS1264C	35-55	'on se disait mais quand est -ce qu"il va nous nous servir son gâteau? qu"on s"en aille hein ah'
475	ESLO2REPAS1264C	35-55	je lui aurais dit non elle aurait fait quoi?
476	ESLO2REPAS1264C	35-55	'cinq six ans mais maintenant je peux plus hein mais enfin ça m"arrive encore mais faut que ce soit exceptionnel quoi qu"est -ce qu"on avait fait la veille au soir? ah bah oui on était on je me souviens plus il était pratiquement on est parti d"ici il était deux heures moins vingt hein'
477	ESLO2REPAS1264C	35-55	'la première nuit moi j"avais pas le moral je dis putain c"est pas vrai je vais pas m"y faire à ce matelas -là comment je vais faire?'
478	ESLO2REPAS1264C	35-55	mais là euh oh putain impressionnant le bruit oh bah je dis attends on est où? on est dans un blockhaus on est quoi ça?
479	ESLO2REPAS1264C	35-55	'qu"y a un frein un un feu rouge ou un stop pas loin hein je j"arrête d"accélérer ça sert à quoi de toute façon?'
480	ESLO2REPAS1264C	35-55	'c"est le mari le mari à Fabienne comment il s"appelle?'
481	ESLO2REPAS1264C	35-55	'on a mangé on a juste mangé avec eux puis on est parti à trois heures quoi on est repassé te voir et on est remonté après ouais mais euh non non moi le seul euh c"est Claudine qui nous a invitées Jocelyne et moi parce que elle elle nous aimait bien toutes les deux et quand ils ont vu l"ambiance qu"on mettait parce qu"on s"est déguisé dans en vamps là que Jocelyne avait oublié de d"amener une paire de chaussettes on avait fait tout on avait demandé à tous les hommes Pierre et tout ah non non j"ai pas confiance et puis qu"est -ce que vous allez me faire à vos à mes chaussettes? aucun homme ne voulait nous prêter ses chaussettes mais non tonton on va pas te les abîmer tes chaussettes t"inquiète pas on va juste les mettre ah bah dis donc on a été et donc quand ils ont vu comment qu"on mettait de l"ambiance'
482	ESLO2REPAS1264C	35-55	il reste qui? Jean-Claude
483	ESLO2REPAS1264C	35-55	bah oui tu veux quoi comme yaourt?

Item	Reference	Age	Extraction
484	ESLO1ENT015C	15-25	'qu"est -ce qui s"est passé exactement?'
485	ESLO1ENT015C	15-25	'combien de lettres? euh qu"est -ce que vous appelez lettres? lettres pourquoi euh?'
486	ESLO1ENT017C	15-25	alors euh c- à quoi vous pensez quand vous dites ces deux heures de loisirs?
487	ESLO1ENT017C	15-25	'ben tout s- tout après où est -ce qu"on va aller?'
488	ESLO1ENT017C	15-25	'moi je savais pas ce que c"était évidemment j"ai pas levé le doigt pour demander puis il y en avait plein dans la classe qui savaient pas bon ben un triptyque qu"est -ce que c"est? elle nous aurait dit euh une récapitulation avec trois feuilles ça aurait été'
489	ESLO1ENT017C	15-25	'combien vous en avez fait jusqu"à maintenant? on voulait en faire trois cents on va pas en faire tant normalement on en fait trois quatre par jour chacun'
490	ESLO1ENT063C	15-25	'alors qu"est -ce que j"ai fait? oui alors je par exemple si tous les matins c"est toujours pareil euh nous recevons le courrier alors euh j"enregistre le courrier ça consiste euh à enregistrer le courrier à tout noter euh toutes les lettres euh qu"on reçoit ensuite euh je les trie quoi y a certaines lettres que je réponds directement toute seule certaines euh je passe euh au chef de service euh ensuite euh je l"ai déjà ensuite alors quand j"ai fait ça enfin je fais d"autres travaux euh que j"avais déjà de la veille euh par exemple euh je fais des des rapports euh des comptes-rendus des choses comme ça qu"est -ce que j"ai fait encore? j"ai tapé quoi'
491	ESLO1ENT063C	15-25	euh disons quels élèves?
492	ESLO1ENT063C	15-25	'refait un un cadre disons euh du moment des des Gaulois enfin des choses comme ça ah ils vendaient d"a- d"ailleurs des casques romains des choses comme ça euh c"était le cadre de des Gaulois quoi y avait une partie qui était réservée justement à ça euh y avait des comment ça s"appelle déjà? des trucs des voyez des cases euh gauloises des choses comme ça des des c"était marrant enfin les Florales euh on a beaucoup parlé des Florales à Orléans mais en fait c"était ça a été plutôt en déficit hein oui'
493	ESLO1ENT063C	15-25	'où est -ce qu"on le garde?'
494	ESLO1ENT069C	15-25	'souvent des des in- euh des euh comment on appelle ça? des journalistes qui ne sont pas tellement euh qui n"ont pas un euh un français tellement châtié enfin'
495	ESLO1ENT092C	15-25	'pourquoi est ce-qu"il faut l"école publique? bah l"école publique ça vous fait voir tout le monde'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
496	ESLO1ENT092C	15-25	'qu"est -ce qu"on attend vous savez au point de vue politique alors là je vous je v- vous je peux vous dire tout de suite que on vote pour lui parce qu"on pense qu"il fera quelque chose mais ce qu"on attend euh euh qu"est ce qu"il peut faire d"abord? voyez on attend ce qu"il peut faire si il on il fera ce qu"il peut ou il fera rien euh vous savez alors là euh on n"attend pas grand chose disons'
497	ESLO1ENT092C	15-25	'pour la ville vous savez actuellement les gens n"ont pas de moyens et nos députés n"ont pas les moyens de faire quelque chose qu"est -ce que vous voulez qu"on attende d"eux?'
498	ESLO1ENT092C	15-25	'mais à condition de pas avoir des élèves de sixième qui vont dire euh actuellement vous avez des collègues comme ça où je reviens à la discipline de tout à l"heure que me demander dans l"enseignement qy"est -ce que bah on est des des collègues et des CEG c"est qu- que c"- c"est c"est comment on peut dire? la foire c"est les élèves de sixième de quatrième qui vont à des réunions de parents d"élèves et puis qui prennent la parole alors en arriver à ce point -là je crois que c"est assez léger maintenant que ce soit des élèves de faculté ou des élèves de terminale ou des élèves de de seconde de troisième même qui donnent leur avis à condition que ce soit un avis euh posé et calculé quoi c"est possible oui ces ces jeunes ont des idées certainement'
499	ESLO1ENT101C	15-25	combien faut compter?
500	ESLO1ENT101C	15-25	'jusque à quel âge euh tes parents ils ont été à l"école?'
501	ESLO1ENT110C	15-25	et puis y avait beaucoup trop de comment? de
502	ESLO1ENT110C	15-25	'enfin nous on connaît mal Orléans hein? on n"est pas touriste un jour y a un touriste dans la rue qui m"a dit ben qu"est -ce qu"y a à visiter comme mu- euh monuments musées à Orléans? et je lui ai répondu j"en sais rien'
503	ESLO1ENT110C	15-25	'oh oui j"ai entendu vaguement parler c"est comment déjà ça?'
504	ESLO1ENT110C	15-25	't"as fait quoi toi?'
505	ESLO1ENT110C	15-25	'en études t"as fait quoi toi?'
506	ESLO1ENT110C	15-25	oh y a y a euh écrit quoi? une lettre?
507	ESLO1ENT120C	15-25	'alors là y a pas de raison mais sans ça je suis pour pourquoi la femme ne travaillerait pas? après tout elle a bien le droit elle aussi de de de vivre avec la société d"être avec tout le monde pas rester enfermée y a qu"un seul ennui c"est les enfants'
508	ESLO1ENT120C	15-25	'mais sans ça euh euh de de toute manière y a y en a la moitié qui ont qui ont fait la grève sans savoir pourquoi alors euh vous savez si si vous enlevez là tous les gens qui étaient dans la ville de Paris ou qui ont fait la grève ou n"importe quoi hein moi moi moi je je parlais avec beaucoup de gens et la plupart pourquoi pourquoi tu fais pourquoi tu fais grève? oh ben tu comprends euh ben dis -moi ben euh'

Item	Reference	Age	Extraction
509	ESLO1ENT120C	15-25	'oh oui mais de toute manière y a c"est idiot le mot bourgeois y a bourgeois et bourgeois il il les les enfin je sais pas les socialistes ou les les gauchisants disons traitent de bourgeois d"une manière péjorative c"est-à-dire euh capitaliste euh euh euh comment ça s"appelle? enfin ramène tout à soi et gagner ses gros sous bon ça c"est ça c"est le bon c"est pas la définition du bourgeois la bourgeoisie bon ben c"est ce sont des vieilles familles qui qui ont amassé de l"argent et alors? mais pour pour la plupart des gens le mot bourgeois tout de suite c"est péjoratif'
510	ESLO1ENT120C	15-25	'ben c"est idiot c"est idiot parce que alors euh bourgeoise ça veut dire beaucoup de choses bourgeoise euh la la la la caste familiale aussi c"est c"est la caste c"est c"est les c"est le le clan c"est le c"est c"est c"est idiot aussi et et les ouvriers se soutiennent bien entre eux pourquoi pourquoi les autres se soutiendraient pas entre eux?'
511	ESLO1ENT128C	15-25	euh comment vous expliquer?'
512	ESLO1ENT128C	15-25	'des mots-croisés combien j"en fais par semaine?'
513	ESLO1ENT128C	15-25	'non il était de non il est devait être de attendez d"où c"est qu"il était? il était d"à côté de Paris qu"il nous avait dit euh'
514	ESLO1ENT130C	15-25	'des étiquettes à la peinture sinon euh pour faire des comment n- qu"on appelle ça? des inventaires des trucs comme ça quoi vraiment'
515	ESLO1ENT167C	15-25	euh je ne deviendrai pas quoi?'
516	ESLO1REU294C	15-25	bonjour alors comment va?'
517	ESLO1REU294C	15-25	'ah c"est toi bien alors y en a combien de Louveteaux là en tout?'
518	ESLO1REU294C	15-25	hier vous avez fait quoi?'
519	ESLO1REU294C	15-25	et comment ça marche dans la sixaine?'
520	ESLO1REU294C	15-25	votre feu il est fait comment? avec des barres en feu non?'
521	ESLO1REU294C	15-25	'alors ça c"est quoi?'
522	ESLO1REU294C	15-25	alors ça veut dire quoi rangement?'
523	ESLO1REU294C	15-25	'pourquoi les marrons qu"est -ce que ça veut dire?'
524	ESLO1REU294C	15-25	alors ça veut dire quoi la B A?'
525	ESLO1REU294C	15-25	alors en q- en quoi ça consiste?'
526	ESLO1REU294C	15-25	mais pourquoi pourquoi tu aimes bien? non non'
527	ESLO1REU294C	15-25	à quelle heure vous partez demain?'
528	ESLO1REU294C	15-25	'les jeux qu"est -ce qui te plaît le plus là-dedans?'
529	ESLO1REU294C	15-25	'et toi qu"est -ce qui te plaît le plus?'
530	ESLO1REU294C	15-25	'à quoi tu vois que c"est la Beauce?'
531	ESLO1REU294C	15-25	'très bien je vois que vous apprenez bien vos vos leçons hein ouais non mais même quand on parle de j- euh géographie tout le monde s"en va non y en a quatre qui sont partis et dans les habitants vous avez qu"est -ce que vous avez remarqué?'
532	ESLO1REU294C	15-25	'comment qu"il s"appelle?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
533	ESLO1REU294C	15-25	et le louveteau venait pas si il pleuvait il est pas venu? il est là il est où dont?
534	ESLO224H1249C	15-25	'OK ça marche ça marche ça marche tu vas faire quoi de beau toi aujourd'hui?'
535	ESLO224H1249C	15-25	'c'est clair vous aviez eu quoi avec elle l'année dernière?'
536	ESLO2CINE1187C	15-25	vous êtes en fa- vous êtes en fac de quoi?
537	ESLO2CINE1187C	15-25	'non mais c'est pourquoi en fait?'
538	ESLO2CINE1187C	15-25	'vous allez faire euh un dossier? c'est quoi?'
539	ESLO2CINE1187C	15-25	vous faites euh quoi comme euh fac de lettres vous faites euh?
540	ESLO2CINE1187C	15-25	ou je ne sais pas quoi?
541	ESLO2CINE1187C	15-25	pourquoi je le sentais?
542	ESLO2CINE1187C	15-25	vous êtes en quelle année?
543	ESLO2CINE1187C	15-25	y a quoi com ou euh?
544	ESLO2CINE1187C	15-25	'je ne sais pas c'est quoi votre UE?'
545	ESLO2CINE1187C	15-25	ça fait combien de temps? nous ça fait trois ans en fait?
546	ESLO2CINE1187C	15-25	'bon bah c'est bien et vous avez fait quoi comme bac sans indiscretion?'
547	ESLO2CINE1187C	15-25	vous étudiez quoi comme logiciels?
548	ESLO2CINE1187C	15-25	et en première année vous étiez combien?
549	ESLO2CINE1187C	15-25	qui se fait et puis généralement après en troisième année je ne sais plus on était combien en troisième année?
550	ESLO2CINE1187C	15-25	vous parlez combien de langues?
551	ESLO2ENTJEUN1228C	15-25	'euh ensuite euh bah les les les supermarchés euh hypermarchés euh euh qu'est-ce qui peut y avoir d'autre?'
552	ESLO2ENTJEUN1229C	15-25	'oui pour euh pour l'avenir et euh euh vous avez un ordre d'idée de dans combien de temps euh?'
553	ESLO2ENTJEUN1231C	15-25	comment on dit?
554	ESLO2ENTJEUN1233C	15-25	'qu'est-ce que je vais faire quand je me déplace sur Orléans?'
555	ESLO2ENTJEUN1233C	15-25	'comment ça j'en reçois tous les jours?'
556	ESLO2ENTJEUN1233C	15-25	'oh franchement c'est compliqué parce que quand tu quand je vais mettre une photo y a du monde qui va commenter alors bah tu tu vas regarder puis tu vas commenter aussi puis et puis bah comme l'autre soir on va discuter sur un truc tiens euh quelqu'un qui irait bien voir ce film bah alors je vais hop tiens bah c'est quand que tu veux aller le voir? oh puis tiens euh bah j'y vais aussi ah bah on pourrait aller au restaurant puis bon commentons commentons'
557	ESLO2ENTJEUN1234C	15-25	'euh c'était quoi qu'on avait fait l'acous- non c'était quoi? la phonétique?'
558	ESLO2ENTJEUN1234C	15-25	il a quel âge?
559	ESLO2ENTJEUN1234C	15-25	'c'est quoi cette question piège?'
560	ESLO2ENTJEUN1234C	15-25	'non mais euh ouais sinon où c'est que où est -ce que je vais manger? au Brin de Zinc euh les trucs comme ça'
561	ESLO2ENTJEUN1234C	15-25	comment ça se fait que tu connais Bartabas?
562	ESLO2ENTJEUN1234C	15-25	'hé ben le c'est quand c'est le week-end dernier? je le feu du radar qui est à côté du théâtre'

Item	Reference	Age	Extraction
563	ESLO2ENTJEUN1234C	15-25	'c"était quand?'
564	ESLO2ENTJEUN1234C	15-25	'c"est où que j"ai été l"autre jour?'
565	ESLO2ENTJEUN1234C	15-25	'genre euh c"est quoi? les chlops pour dire des chaussures'
566	ESLO2ENTJEUN1235C	15-25	'de classe euh ma prof me disait tu as tu as un bon tu as un très bon niveau et cetera j"ai dit ben j"avais quatre de moyenne l"année dernière elle m"a demandé tu étais avec quel prof?'
567	ESLO2ENTJEUN1235C	15-25	'au collègue on attend que l"acné passe au lycée ça passe à peu près donc on essaye de se chercher c"est quoi mon identité? qu"est -ce que je veux faire plus tard?'
568	ESLO2ENTJEUN1235C	15-25	'enfin le marché de Noël c"est dans quelle ville déjà? Strasbourg?'
569	ESLO2ENTJEUN1235C	15-25	'justement je me dis comment est -ce que je vais pouvoir utiliser mes capacités informatiques? dans lo- lorsque j"ai vu en master un qu"il y avait euh la création de logiciels et cetera avec la charte graphie j"ai fait'
570	ESLO2ENT1001C	15-25	'là c"est une boîte de finitions dans laquelle je suis en ce moment c"est tout ce qui est tout ce qui est des des murs euh euh pose de pose de toile pose de papier peint peinture et euh tout ce qui est des sols finitions donc euh parquet strat- carrelage lino- et euh y a quoi d"autre? faïence des bandes euh de temps en temps si derrière un chantier de placo y a de la bande à faire éventuellement on va sous-traiter un chantier de placo donc voilà'
571	ESLO2ENT1001C	15-25	'c"est quoi la rue où y a ma mutuelle là?'
572	ESLO2ENT1001C	15-25	sinon bah je sais pas y a quoi?'
573	ESLO2ENT1001C	15-25	'je sais pas Pris c"est où Ajaccio?'
574	ESLO2ENT1001C	15-25	pourquoi tu tu te marres tu sais le truc non?'
575	ESLO2ENT1001C	15-25	'et le but c"est quoi c"est c"est d"étudier le la façon de parler des gens?'
576	ESLO2ENT1003C	15-25	'j"ai enfin ouais ça m"a pas non ça m"a pas vraiment plu en fait et donc bah j"ai eu une période où je me suis dit qu"est -ce que je fais?'
577	ESLO2ENT1026C	15-25	comment je fais une omelette?'
578	ESLO2ENT1038C	15-25	'c"est quoi la question?'
579	ESLO2ENT1038C	15-25	'Hero Corp ça s"écrit comment?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
580	ESLO2ENT1038C	15-25	'j"ai bossé euh rue de la Rép- donc en fait tous les soirs pour euh rentrer chez moi fallait que je passe dans le marché de Noël ah bah je peux te dire que euh quand tu te tapes euh sept heures dans un magasin avec des gens euh pour les cadeaux de Noël qui te cassent les couilles qui connaissent pas les tailles qui euh qui savent pas euh qui le style de la personne ah bah je vais faire un cadeau oui et ben i- il fait quelle taille? oh bah je sais pas ah il aime quoi? ah bah je sais pas ah bah écoute tu as tu as cas lui acheter des gants euh enfin tu vois c"est des trucs les périodes de Noël c"est horrible quoi c"est le pire dans le vente et du coup ben euh le marché de Noël ça me rappelle un peu des mauvais souvenirs parce que'
581	ESLO2ENT1038C	15-25	'j"ai bossé euh rue de la Rép- donc en fait tous les soirs pour euh rentrer chez moi fallait que je passe dans le marché de Noël ah bah je peux te dire que euh quand tu te tapes euh sept heures dans un magasin avec des gens euh pour les cadeaux de Noël qui te cassent les couilles qui connaissent pas les tailles qui euh qui savent pas euh qui le style de la personne ah bah je vais faire un cadeau oui et ben i- il fait quelle taille? oh bah je sais pas ah il aime quoi? ah bah je sais pas ah bah écoute tu as tu as cas lui acheter des gants euh enfin tu vois c"est des trucs les périodes de Noël c"est horrible quoi c"est le pire dans le vente et du coup ben euh le marché de Noël ça me rappelle un peu des mauvais souvenirs parce que'
582	ESLO2ENT1047C	15-25	'et donc euh bah ça m"a ça ça m"a fini là puis souvent on me demande bah dans ce cas là pourquoi tu as pas fait pompier professionnel?
583	ESLO2ITI1125C	15-25	'c"est quoi ces conneries?'
584	ESLO2REPAS1253C	15-25	y a quoi ce soir à la télé?
585	ESLO2REPAS1253C	15-25	'merci je me disais bah pourtant i- il fait pas de de sport comment ça se fait? il fait bah bah j"étais scout et après j"ai repensé à ce qu"on avait dit en rentrant il fait'
586	ESLO2REPAS1253C	15-25	comment il lui dit?
587	ESLO2REPAS1253C	15-25	tu veux quoi?
588	ESLO2REPAS1253C	15-25	pourquoi on peut pas soigner ça?
589	ESLO2REPAS1253C	15-25	mais pourquoi pourquoi tu fais des commentaires? tu es pas à la télé
590	ESLO2REPAS1254C	15-25	'c"est à qui?'
591	ESLO2REPAS1254C	15-25	'c"est qui qui fait ça?'
592	ESLO2REPAS1254C	15-25	'c"est quoi?'
593	ESLO2REPAS1254C	15-25	'est -ce que je l"ai ramenée? est -ce que je l"ai ramenée? où est -ce que je l"ai mise ma copie d"histoire?'

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Item	Reference	Age	Extraction
594	ESLO2REPAS1254C	15-25	'c"est ma vie voilà comme comme tu l"as dit c"est comme la chanson euh Elle me Dit euh eh pourquoi tu gâches ma vie mais c"est ma vie fais ce que tu veux tant pis c"est comme dans la chanson pourquoi tu gâches ta vie? oh elle me dit fais comme si fais ce que tu veux tant pis'
595	ESLO2REPAS1254C	15-25	pourquoi on aura pas le temps jeudi?
596	ESLO2REPAS1255C	15-25	et Jacky il a pris quoi?
597	ESLO2REPAS1255C	15-25	elle est partie quand après mamie?
598	ESLO2REPAS1255C	15-25	je mets ça où?
599	ESLO2REPAS1255C	15-25	'ah faut aller dans le garage euh c"est laquelle de poubelle pour ça?'
600	ESLO2REPAS1263C	15-25	tu fais quoi cet après-midi?
601	ESLO2REPAS1263C	15-25	ouais ça fera plus près tu en veux combien?
602	ESLO2REPAS1263C	15-25	'd"accord et papa cet après-midi il bosse jusqu"à quelle heure?'
603	ESLO2REPAS1263C	15-25	et il redescend quand à Orléans?
604	ESLO2REPAS1263C	15-25	'bah attends je vais d"abord manger mon fromage tu vas faire quand les courses?'
605	ESLO2REPAS1263C	15-25	'hm elle est super bonne ta tarte aux pommes tu l"as fait comment? tu as mis quoi dessus?'
606	ESLO2REPAS1263C	15-25	'hm elle est super bonne ta tarte aux pommes tu l"as fait comment? tu as mis quoi dessus?'
607	ESLO2REPAS1269C	15-25	'bah i- ils venaient d"où les cordons bleus?'
608	ESLO2REPAS1269C	15-25	'ah elle reste pas j"étais en train de me dire c"est qui?'
609	ESLO2REPAS1269C	15-25	ça coûte combien une plaquette?
610	ESLO2REPAS1269C	15-25	'l"année dernière vous avez gagné quoi?'
611	ESLO2REPAS1269C	15-25	eh pourquoi y a des cheveux dans mon yaourt sérieux?
612	ESLO2REPAS1269C	15-25	tu vas te doucher quand?

TABLEAU C.1 – Extractions du corpus ESLO retenues pour analyse

Annexe D

Corpus ESLO : Modèles statistiques

D.0.1 Modèle *brms* et distributions postérieures pour un effet de CORPUS (4.2.4)

Family : categorical

Links : muFA = logit; muFESK = logit; muFINV = logit

Formula : construction ~ corp2 + (1 | ID_file)

Data : dat (Number of observations : 612)

Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-0.48	0.17	-0.83	-0.17	1.00	3380	3471
muFESK_Intercept	-1.41	0.21	-1.86	-1.05	1.00	3039	3433
muFINV_Intercept	-1.72	0.29	-2.31	-1.21	1.00	2585	3650
muFA_corp2	-1.41	0.33	-2.08	-0.78	1.00	3593	3794
muFESK_corp2	-1.43	0.35	-2.10	-0.75	1.00	4866	4138
muFINV_corp2	-2.09	0.46	-3.04	-1.23	1.00	3646	4065

TABLEAU D.1 – Résumé du modèle brms (ESLO, effet CORPUS)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Antéposition	-1.41	0.00	1.00	-2.76	-0.33	-2.08	-0.78	-1.96	-0.88	-1.89	-0.94
Antéposition & ESK	-1.43	0.00	1.00	-2.68	-0.19	-2.10	-0.75	-1.99	-0.87	-1.92	-0.94
Antéposition & Inversion	-2.09	0.00	1.00	-3.84	-0.39	-3.04	-1.23	-2.88	-1.37	-2.76	-1.44

TABLEAU D.2 – Distributions postérieures pour le modèle brms (ESLO, effet CORPUS)

D.0.2 Modèle *brms* et distributions postérieures pour un effet d'ÂGE (4.2.5)

Family : categorical

Links : muFA = logit; muFESK = logit; muFINV = logit

Formula : construction ~ age2 + (1 | ID_file)

Data : dat (Number of observations : 612)

Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-0.50	0.18	-0.87	-0.15	1.00	2360	2880
muFESK_Intercept	-1.53	0.23	-2.03	-1.12	1.00	2623	2951
muFINV_Intercept	-1.73	0.30	-2.38	-1.20	1.00	2202	2613
muFA_age2	0.00	0.37	-0.74	0.71	1.00	2704	3867
muFESK_age2	0.01	0.43	-0.82	0.87	1.00	3801	3965
muFINV_age2	0.27	0.55	-0.83	1.34	1.00	2763	3084

TABLEAU D.3 – Résumé du modèle *brms* (ESLO, effet ÂGE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	195	h95	190	h90	185	h85
Antéposition	0.00	0.51	0.49	-1.43	1.40	-0.74	0.71	-0.61	0.60	-0.53	0.54
Antéposition & ESK	0.01	0.50	0.50	-1.62	1.67	-0.82	0.87	-0.69	0.73	-0.60	0.63
Antéposition & Inversion	0.27	0.69	0.31	-2.30	2.63	-0.83	1.34	-0.65	1.17	-0.52	1.05

TABLEAU D.4 – Distributions postérieures pour le modèle *brms* (ESLO, ÂGE)

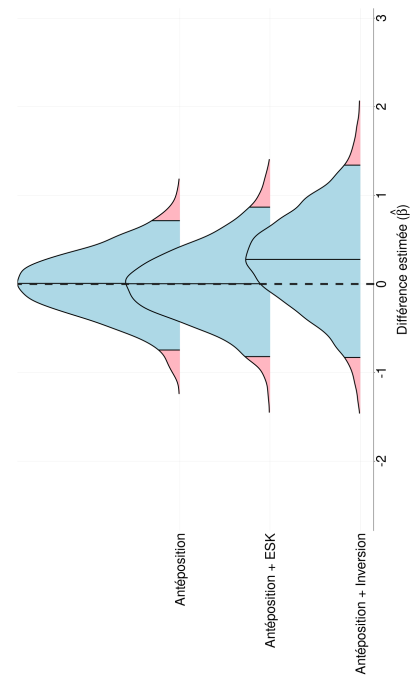


FIGURE D.1 – Distributions postérieures pour le modèle *brms* (ESLO, effet d'ÂGE)

D.0.3 Modèle *brms* et distributions postérieures pour un modèle avec CORPUS et ÂGE (4.2.6)

Family : categorical

Links : muFA = logit; muFESK = logit; muFINV = logit

Formula : construction ~ corp2*age2 + (1 | ID_file)

Data : dat (Number of observations : 612)

Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-0.47	0.16	-0.82	-0.17	1.00	4697	4597
muFESK_Intercept	-1.43	0.21	-1.89	-1.06	1.00	3918	3624
muFINV_Intercept	-1.75	0.29	-2.36	-1.22	1.00	3486	3703
muFA_corp2	-1.42	0.32	-2.07	-0.80	1.00	5037	4102
muFA_age2	0.02	0.37	-0.73	0.72	1.00	3789	3984
muFA_corp2:age2	0.78	0.73	-0.64	2.24	1.00	5800	4591
muFESK_corp2	-1.44	0.35	-2.12	-0.75	1.00	6020	4954
muFESK_age2	0.08	0.42	-0.77	0.91	1.00	6925	4449
muFESK_corp2:age2	1.13	0.81	-0.46	2.72	1.00	6075	4715
muFINV_corp2	-2.11	0.46	-3.04	-1.25	1.00	4231	4975
muFINV_age2	0.16	0.52	-0.87	1.19	1.00	5619	4647
muFINV_corp2:age2	0.08	1.06	-2.01	2.20	1.00	5396	4436

TABLEAU D.5 – Résumé du modèle *brms* (ESLO, effets CORPUS*ÂGE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
F_Corpus	-1.42	0.00	1.00	-2.65	-0.24	-2.07	-0.80	-1.95	-0.90	-1.89	-0.97
F_Âge	0.02	0.53	0.47	-1.44	1.26	-0.73	0.72	-0.61	0.61	-0.53	0.54
F_Corpus*Âge	0.78	0.86	0.14	-2.06	4.18	-0.64	2.24	-0.40	2.00	-0.26	1.84
FESK_Corpus	-1.44	0.00	1.00	-2.71	0.08	-2.12	-0.75	-2.02	-0.87	-1.95	-0.94
FESK_Âge	0.08	0.57	0.43	-1.41	1.79	-0.77	0.91	-0.62	0.78	-0.53	0.68
FESK_Corpus*Âge	1.13	0.92	0.08	-1.82	4.21	-0.46	2.72	-0.20	2.48	-0.03	2.29
FINV_Corpus	-2.11	0.00	1.00	-3.86	-0.45	-3.04	-1.25	-2.87	-1.39	-2.78	-1.47
FINV_Âge	0.16	0.62	0.38	-2.00	2.55	-0.87	1.19	-0.69	1.02	-0.59	0.90
FINV_Corpus*Âge	0.08	0.53	0.47	-3.56	4.46	-2.01	2.20	-1.66	1.80	-1.45	1.58

TABLEAU D.6 – Distributions postérieures pour le modèle *brms* (ESLO, CORPUS*ÂGE)

D.0.4 Modèle *brms* et distributions postérieures pour un effet de CONTEXTE (4.2.7)

Family : categorical

Links : muFA = logit; muFESK = logit; muFINV = logit

Formula : construction ~ (meal_school + interview_school) + (1 | ID_file)

Data : dat2 (Number of observations : 556)

Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-0.47	0.18	-0.85	-0.12	1.00	5228	4699
muFESK_Intercept	-1.48	0.23	-1.96	-1.09	1.00	3805	4424
muFINV_Intercept	-1.68	0.31	-2.34	-1.13	1.00	4863	4320
muFA_meal_school	-1.38	0.69	-2.80	-0.03	1.00	4157	4310
muFA_interview_school	0.59	0.46	-0.31	1.53	1.00	4699	4765
muFESK_meal_school	-0.22	0.75	-1.69	1.24	1.00	6176	4637
muFESK_interview_school	1.82	0.55	0.78	2.96	1.00	6435	4458
muFINV_meal_school	-2.37	1.07	-4.58	-0.38	1.00	4986	4464
muFINV_interview_school	0.35	0.65	-0.95	1.60	1.00	4004	4091

TABLEAU D.7 – Résumé du modèle *brms* (ESLO, effet CONTEXTE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
F_Repas	-1.38	0.02	0.98	-4.78	1.13	-2.80	-0.03	-2.52	-0.26	-2.38	-0.40
F_Entretiens	0.59	0.91	0.09	-1.27	2.54	-0.31	1.53	-0.15	1.37	-0.06	1.26
FESK_Repas	-0.22	0.39	0.61	-3.24	2.62	-1.69	1.24	-1.44	1.00	-1.28	0.84
FESK_Entretiens	1.82	1.00	0.00	-0.17	4.35	0.78	2.96	0.94	2.75	1.04	2.61
FINV_Repas	-2.37	0.01	0.99	-6.66	1.66	-4.58	-0.38	-4.18	-0.65	-3.96	-0.89
FINV_Entretiens	0.35	0.72	0.28	-2.19	3.13	-0.95	1.60	-0.75	1.38	-0.60	1.26

TABLEAU D.8 – Distributions postérieures pour le modèle *brms* (ESLO, CONTEXTE)

D.0.5 Modèle *brms* et distributions postérieures pour un modèle avec CORPUS et ÂGE (interviews, 4.2.8.1)

Family : categorical

Links : muFA = logit; muFESK = logit; muFINV = logit

Formula : construction ~ corp2*age2 + (1 | ID_file)

Data : inter (Number of observations : 334)

Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-0.24	0.26	-0.79	0.24	1.00	3617	3563
muFESK_Intercept	-0.79	0.23	-1.31	-0.37	1.00	4379	4412
muFINV_Intercept	-1.87	0.45	-2.83	-1.10	1.00	2533	3743
muFA_corp2	-1.28	0.48	-2.22	-0.35	1.00	3799	3789
muFA_age2	0.15	0.60	-1.05	1.32	1.00	4240	4397
muFA_corp2:age2	1.31	1.13	-0.85	3.55	1.00	3575	3232
muFESK_corp2	-0.86	0.42	-1.71	-0.03	1.00	5334	3768
muFESK_age2	0.04	0.49	-0.93	0.99	1.00	5033	4435
muFESK_corp2:age2	1.15	0.98	-0.75	3.07	1.00	6177	4422
muFINV_corp2	-2.96	0.70	-4.49	-1.71	1.00	4213	3877
muFINV_age2	-0.41	0.77	-1.97	1.11	1.00	4189	3940
muFINV_corp2:age2	-0.66	1.39	-3.33	2.06	1.00	4237	3831

TABLEAU D.9 – Résumé du modèle *brms* (ESLO, interviews, effets CORPUS*ÂGE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
F_Corpus	-1.28	0.00	1.00	-3.23	0.38	-2.22	-0.35	-2.08	-0.52	-1.98	-0.61
F_Âge	0.15	0.60	0.40	-2.54	2.29	-1.05	1.32	-0.85	1.12	-0.73	1.00
F_Corpus*Âge	1.31	0.89	0.11	-3.13	5.89	-0.85	3.55	-0.49	3.19	-0.28	2.96
FESK_Corpus	-0.86	0.02	0.98	-2.64	0.76	-1.71	-0.03	-1.55	-0.17	-1.45	-0.27
FESK_Âge	0.04	0.54	0.46	-1.81	1.85	-0.93	0.99	-0.76	0.84	-0.65	0.74
FESK_Corpus*Âge	1.15	0.88	0.12	-2.64	5.41	-0.75	3.07	-0.44	2.77	-0.24	2.55
FINV_Corpus	-2.96	0.00	1.00	-6.25	-1.05	-4.49	-1.71	-4.19	-1.89	-4.02	-2.00
FINV_Âge	-0.41	0.30	0.70	-3.25	2.56	-1.97	1.11	-1.70	0.82	-1.53	0.67
FINV_Corpus*Âge	-0.66	0.31	0.69	-7.60	5.98	-3.33	2.06	-2.87	1.63	-2.60	1.33

TABLEAU D.10 – Distributions postérieures pour le modèle *brms* (ESLO, interviews, CORPUS*ÂGE)

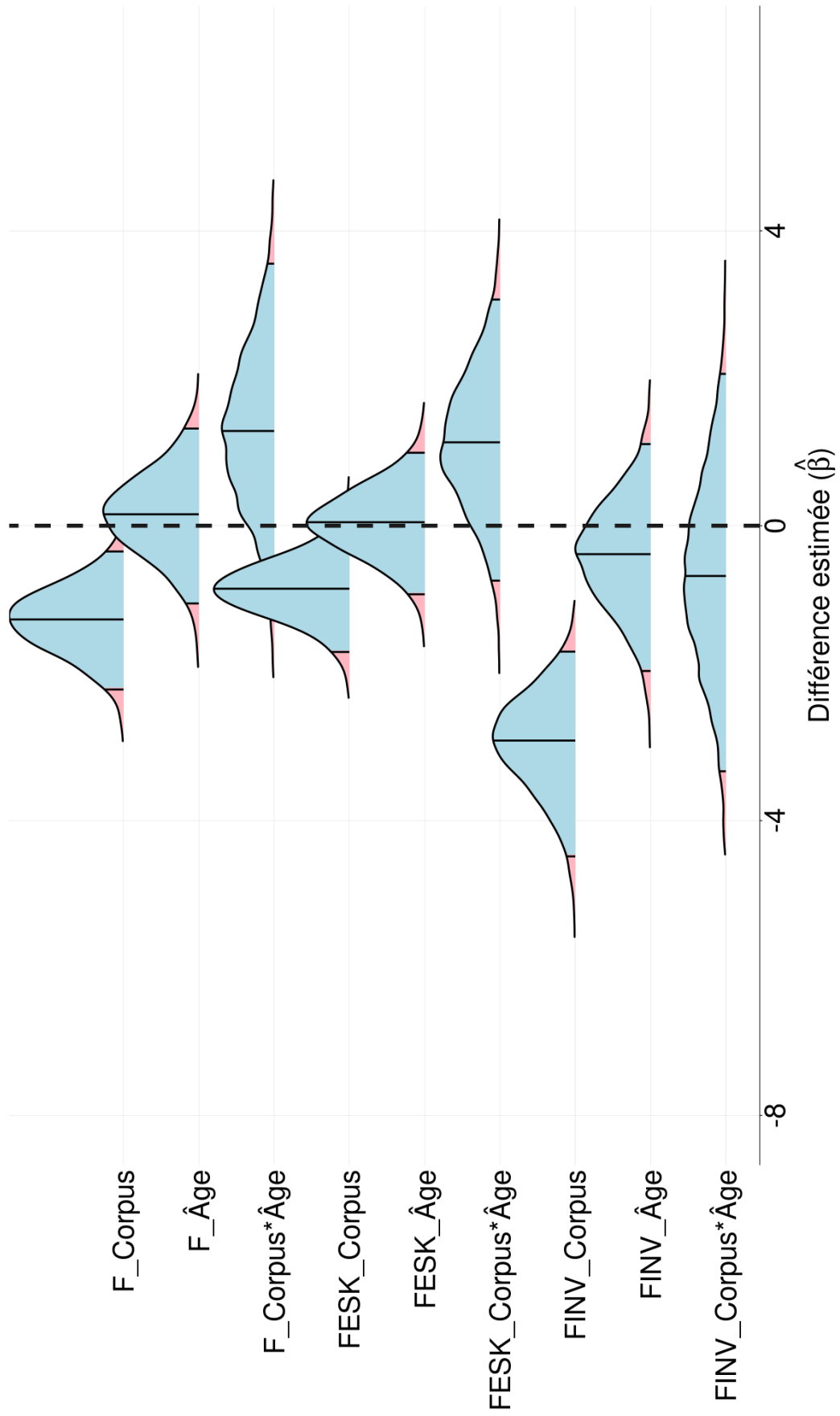


FIGURE D.2 – Distributions postérieures pour le modèle brms (ESLO, interviews, CORPUS*ÂGE)

D.0.6 Modèle *brms* et distributions postérieures pour un modèle avec CONTEXTE et ÂGE (ESLO 2, 4.2.8.2)

Family : categorical

Links : muFA = logit; muFESK = logit; muFINV = logit

Formula : construction ~ age2*meal_interview + (1 | ID_file)

Data : eslo2mealinter (Number of observations : 201)

Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
muFA_Intercept	-1.48	0.48	-2.59	-0.67	1.00	2950	3025
muFESK_Intercept	-2.01	0.48	-3.12	-1.26	1.00	3401	3559
muFINV_Intercept	-3.42	0.89	-5.71	-2.25	1.00	2422	1673
muFA_age2	-0.12	0.63	-1.41	1.04	1.00	4191	4012
muFA_meal_interview	-1.47	0.93	-3.53	0.24	1.00	3189	2531
muFA_age2:meal_interview	-1.63	1.17	-3.98	0.68	1.00	4967	4855
muFESK_age2	0.71	0.80	-0.81	2.38	1.00	5260	4292
muFESK_meal_interview	-1.65	0.84	-3.43	-0.08	1.00	4926	4215
muFESK_age2:meal_interview	0.62	1.66	-2.46	4.09	1.00	6421	4487
muFINV_age2	-0.21	1.05	-2.23	1.91	1.00	5252	3315
muFINV_meal_interview	-0.51	1.25	-2.77	2.23	1.00	3607	2208
muFINV_age2:meal_interview	1.20	2.11	-2.59	5.85	1.00	5078	3027

TABLEAU D.11 – Résumé du modèle *brms* (ESLO 2, effets ÂGE*CONTEXTE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	195	h95	190	h90	185	h85
F_Âge	-0.12	0.44	0.56	-3.49	2.63	-1.41	1.04	-1.18	0.87	-1.02	0.74
F_Contexte	-1.47	0.04	0.96	-6.05	2.56	-3.53	0.24	-3.09	-0.07	-2.82	-0.25
F_Âge*Contexte	-1.63	0.08	0.92	-6.59	3.75	-3.98	0.68	-3.56	0.28	-3.27	0.05
FESK_Âge	0.71	0.82	0.18	-2.18	4.21	-0.81	2.38	-0.57	2.06	-0.41	1.88
FESK_Contexte	-1.65	0.02	0.98	-5.32	2.69	-3.43	-0.08	-3.08	-0.38	-2.87	-0.54
FESK_Âge*Contexte	0.62	0.64	0.36	-5.13	7.41	-2.46	4.09	-1.96	3.46	-1.66	3.07
FINV_Âge	-0.21	0.41	0.59	-7.26	4.46	-2.23	1.91	-1.84	1.49	-1.62	1.26
FINV_Contexte	-0.51	0.30	0.70	-5.96	6.79	-2.77	2.23	-2.37	1.59	-2.14	1.22
FINV_Âge*Contexte	1.20	0.71	0.29	-8.65	10.65	-2.59	5.85	-2.01	4.94	-1.57	4.32

TABLEAU D.12 – Distributions postérieures pour le modèle *brms* (ESLO2, ÂGE*CONTEXTE)

Annexe E

AJT2 : Norming préalable à l'étude

Pour assurer la cohérence du design expérimental choisi et la fiabilité des résultats au regard des prédictions de départ, les deux niveaux de la variable CONTEXTE ont été vérifiés dans un norming.

E.0.1 Volontaires (AJT2p, 5.2.2)

Nombre de volontaires : 48 francophones

Genre des volontaires : 36 femmes + 10 hommes + 2 personnes

Âge des volontaires : 18-81 ans, $\mu=43$, $m=42$

Recrutement : RISC et réseaux sociaux

Compensation : Aucune

E.0.2 Protocole et design (AJT2p, 5.2.2)

Tâche : Jugements de la familiarité d'une phrase

Dates : Février 2018

Durée : 10 minutes

Modalité : Expérience en ligne, via IbexFarm [DRUMMOND, 2016]

Design : 1 variable (FORMALITÉ), 2 conditions (*formel* – *informel*)

NB : L'informalité des phrases était obtenue par trois procédés distincts : négation sans « ne », vocabulaire familier, et dislocation à gauche ou à droite d'un élément référencé dans la phrase racine par un pronom

Carré latin : Oui, chaque volontaire ne voyait qu'une seule version d'un item et voyait un nombre égal de chaque condition sur toute l'expérience

NB : Suite à une erreur de codage, l'item ajtp-2 n'est pas passé dans sa condition *formel*

Randomisation : Oui

Nombre d'items : 3 entraînements/contrôles + 30 items (les 30 items cibles de l'étude AJT2)
L'informalité des phrases était obtenue par trois procédés distincts : négation sans « ne », vocabulaire familier, et dislocation à gauche ou à droite d'un élément référencé dans la phrase par un pronom

Échelle : 1-7, de "Pas du tout formel" à "Très formel"

Consigne :

Dans cette expérience, il vous est demandé de juger la formalité des phrases qui vont apparaître à l'écran, sur une échelle de 1 (pas du tout formel) à 7 (très formel).

Une situation non formelle est une situation détendue, entre amis, en famille, ou entre gens qui se connaissent bien... Une situation formelle est au contraire une situation dont le contexte est plus sérieux, de type professionnel, officiel, administratif...

Utilisez s'il vous plaît l'intégralité de l'échelle.

L'ensemble devrait vous prendre une dizaine de minutes. Merci de votre participation !

L'ensemble des données collectées durant cette expérience le sont à des fins de recherche. Aucune donnée personnelle d'identification ne sera transmise à des tiers.

Votre âge :

Votre nationalité :

Votre genre :

- Homme
 Femme
 Autre
 Je ne veux pas répondre à cette question

Votre dernier diplôme obtenu :

- Avant le baccalauréat
 Baccalauréat ou équivalent
 Bac+2 ou équivalent
 Bac+3 ou équivalent
 Bac+5 ou équivalent
 Bac+8 ou équivalent

Votre domaine d'études ou d'activité professionnelle :

Votre région de naissance :

Votre langue maternelle :

Autres langues:

Niveau: de 1 (mauvais) à 10 (excellent)

Je suis majeur.e et j'accepte les termes et conditions de cette expérience.

→ [Cliquez ici pour continuer](#)

FIGURE E.1 – Consigne (AJT2p - Norming)

Capture d'écran :

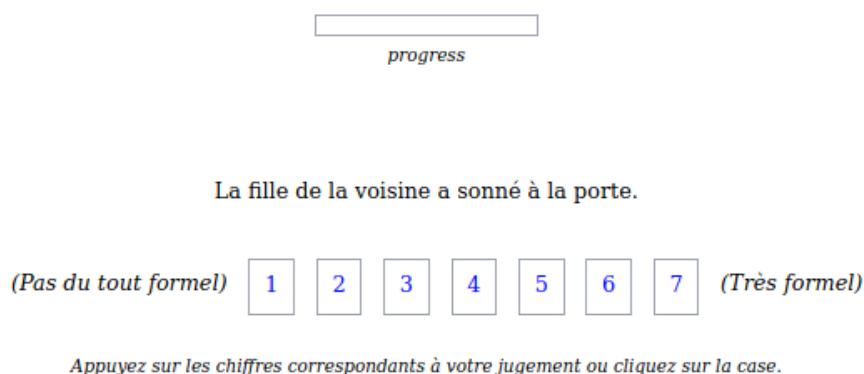


FIGURE E.2 – Capture d'écran (AJT2p - Norming)

E.0.3 Liste des items (AJT2p, 5.2.2)

ajtp-101	La fille de la voisine a sonné à la porte. <i>Formalité : neutre</i>
ajtp-102	Putain, tu fais chier! <i>Formalité : vulgaire, très informel</i>
ajtp-103	Indiscutablement, l'abnégation dont fait preuve le corps médical est remarquable. <i>Formalité : sophistiqué, formel</i>

TABLEAU E.1 – Items expérimentaux pour le norming préalable à l'étude AJT2 (contrôles 101-103)

ajtp-1 <i>formel</i>	Jean arrive demain.
ajtp-1 <i>informel</i>	Eh, Jean il arrive demain.
ajtp-2 <i>formel</i>	J'ai mis les clefs dans le tiroir.
ajtp-2 <i>informel</i>	Les clefs, je les ai foutues dans le tiroir.
ajtp-3 <i>formel</i>	J'ai cassé mon téléphone.
ajtp-3 <i>informel</i>	Pff, mon téléphone, je l'ai pétié.
ajtp-4 <i>formel</i>	Je pars demain.
ajtp-4 <i>informel</i>	Ouais, moi je me barre demain.
ajtp-5 <i>formel</i>	J'ai résolu l'équation du professeur!
ajtp-5 <i>informel</i>	Eh, l'équation du prof, je l'ai résolue!
ajtp-6 <i>formel</i>	Ils ont enfin choisi le nouveau chef.
ajtp-6 <i>informel</i>	T'as vu, ils l'ont enfin choisi, le nouveau chef.
ajtp-7 <i>formel</i>	Vous aurez les dates d'examen plus tard.
ajtp-7 <i>informel</i>	Bon, vous les aurez plus tard, les dates du partiel.
ajtp-8 <i>formel</i>	Cet homme est vraiment sympathique.
ajtp-8 <i>informel</i>	C'est vraiment un gars sympa, ce mec.
ajtp-9 <i>formel</i>	J'ai enfin trouvé mon sujet!
ajtp-9 <i>informel</i>	Yes, je l'ai enfin trouvé, mon sujet!
ajtp-10 <i>formel</i>	Il faut faire cela autrement.
ajtp-10 <i>informel</i>	Non mais il faut le faire autrement, ça.

TABLEAU E.2 – Items expérimentaux pour le norming préalable à l'étude AJT2 (cibles 1-10)

ajtp-11 <i>formel</i>	Non, ma sœur n'est pas là.
ajtp-11 <i>informel</i>	Non, elle est pas là ma sœur.
ajtp-12 <i>formel</i>	Ils ne répareront pas la digue tout de suite...
ajtp-12 <i>informel</i>	C'est pas tout de suite qu'ils la répareront, la digue...
ajtp-13 <i>formel</i>	Je ne sais pas ce que j'ai fait de ton agrafeuse...
ajtp-13 <i>informel</i>	Je sais pas ce que j'en ai fait, de ton agrafeuse...
ajtp-14 <i>formel</i>	Personne ne m'a rien donné!
ajtp-14 <i>informel</i>	Personne m'a rien donné, à moi!
ajtp-15 <i>formel</i>	Je n'ai pas réussi à suivre le débat en entier.
ajtp-15 <i>informel</i>	J'ai pas réussi à le suivre en entier, le débat.
ajtp-16 <i>formel</i>	Je n'ai jamais eu de problème avec cette machine.
ajtp-16 <i>informel</i>	Cette machine, j'ai jamais eu de problème avec!
ajtp-17 <i>formel</i>	Elle n'y sera pas toute seule.
ajtp-17 <i>informel</i>	Elle, elle y sera pas toute seule.
ajtp-18 <i>formel</i>	Je n'aime pas tellement les voitures rouges.
ajtp-18 <i>informel</i>	Moi j'aime pas tellement les voitures rouges.
ajtp-19 <i>formel</i>	Les premiers matches n'étaient pas simples, mais on a gagné!
ajtp-19 <i>informel</i>	Les premiers matches, ils étaient pas simples, mais on a gagné!
ajtp-20 <i>formel</i>	Il ne faudrait pas que tu oublies de venir demain.
ajtp-20 <i>informel</i>	Toi, il faudrait pas que tu oublies de venir demain.
ajtp-21 <i>formel</i>	Tu ne devineras jamais qui j'ai rencontré hier à l'université!
ajtp-21 <i>informel</i>	Tu devineras jamais qui j'ai croisé hier à la fac!
ajtp-22 <i>formel</i>	Je n'ai pas réussi à ne pas attraper de virus...
ajtp-22 <i>informel</i>	J'ai pas réussi à ne pas choper de virus...
ajtp-23 <i>formel</i>	Il n'y a pas trente-six solutions pour régler leur problème.
ajtp-23 <i>informel</i>	Y a pas trente-six soluces pour régler leur problème.
ajtp-24 <i>formel</i>	Je ne sais vraiment plus comment faire, je n'ai plus d'espoir.
ajtp-24 <i>informel</i>	Je sais vraiment plus comment faire, j'ai plus la foi.
ajtp-25 <i>formel</i>	Mon frère ne sait plus ce qu'il veut.
ajtp-25 <i>informel</i>	Mon frangin sait plus ce qu'il veut.
ajtp-26 <i>formel</i>	Il est vraiment impossible de raisonner cet homme.
ajtp-26 <i>informel</i>	C'est vraiment pas possible de raisonner ce mec.
ajtp-27 <i>formel</i>	Je ne veux pas aller dans cette région!
ajtp-27 <i>informel</i>	Je veux pas aller dans ce trou paumé!
ajtp-28 <i>formel</i>	Il n'y a aucune chance qu'elle vienne.
ajtp-28 <i>informel</i>	Y a aucune chance qu'elle se ramène.
ajtp-29 <i>formel</i>	Il ne sera plus du tout disponible pendant ses congés.
ajtp-29 <i>informel</i>	Il sera plus du tout dispo pendant ses vacances.
ajtp-30 <i>formel</i>	Il y a longtemps qu'on ne joue plus au poker avec le groupe.
ajtp-30 <i>informel</i>	Y a longtemps qu'on joue plus au poker avec le gang.

TABLEAU E.3 – Items expérimentaux pour le norming préalable à l'étude AJT2 (cibles 11-30)

E.0.4 Résultats (AJT2p, 5.2.2)

Plus de détails dans les fichiers bruts disponibles dans le répertoire osf.io/ud4ks.

Les notes reçues, entre 1 et 7, ont été codées entre -3 et +3 de façon à être centrées autour de 0. Une note négative signifie un item jugé *informel*, une note positive signifie un item jugé *formel*.

Condition	Item	Moyenne	Écart-type
Formel	1	0.67	1.52
Informel	1	-1.54	1.59
Formel	2		
Informel	2	-2.46	0.83
Formel	3	0.88	1.51
Informel	3	-2.38	1.28
Formel	4	1.04	1.30
Informel	4	-2.38	0.97
Formel	5	0.92	1.21
Informel	5	-1.67	1.76
Formel	6	0.88	1.51
Informel	6	-1.54	1.14
Formel	7	0.88	1.42
Informel	7	-1.38	1.56
Formel	8	1.92	1.10
Informel	8	-2.00	1.06
Formel	9	0.58	1.38
Informel	9	-1.50	1.89
Formel	10	1.25	1.36
Informel	10	-1.42	0.88
Formel	11	0.75	1.36
Informel	11	-1.71	1.37
Formel	12	0.92	1.38
Informel	12	-1.71	1.20
Formel	13	0.00	1.62
Informel	13	-1.67	1.61
Formel	14	0.17	1.76
Informel	14	-1.96	0.95
Formel	15	1.17	1.20
Informel	15	-1.42	1.47
Formel	16	1.42	1.02
Informel	16	-1.50	1.18
Formel	17	0.29	1.33
Informel	17	-1.88	1.33
Formel	18	1.08	1.10
Informel	18	-1.08	1.59
Formel	19	0.46	1.14
Informel	19	-1.46	1.38
Formel	20	1.12	1.48
Informel	20	-1.42	1.50

TABLEAU E.4 – Résultats du norming (AJT2p, items 1-20)

Condition	Item	Moyenne	Écart-type
Formel	21	0.33	1.55
Informel	21	-0.96	1.90
Formel	22	0.08	1.77
Informel	22	-2.12	1.03
Formel	23	0.50	1.35
Informel	23	-2.21	1.22
Formel	24	1.04	1.23
Informel	24	-1.46	1.59
Formel	25	0.83	1.17
Informel	25	-2.12	1.42
Formel	26	1.38	1.35
Informel	26	-1.75	1.11
Formel	27	0.25	1.45
Informel	27	-2.04	1.68
Formel	28	1.08	1.41
Informel	28	-2.25	1.07
Formel	29	1.33	1.09
Informel	29	-1.71	1.43
Formel	30	0.58	1.35
Informel	30	-1.79	1.38
Entraînement	101	0.02	1.56
Entraînement	102	-2.44	1.49
Entraînement	103	2.17	1.55

TABLEAU E.5 – Résultats du norming (AJT2p, items 21-30 - Entraînements)

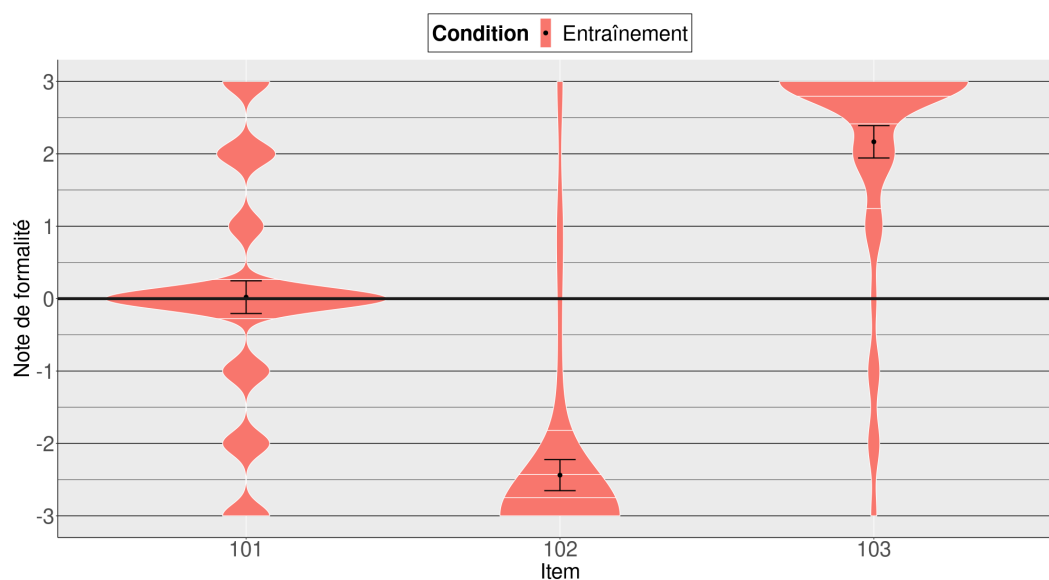


FIGURE E.3 – Résultats du norming (AJT2p, Entraînements)

E.0.5 -

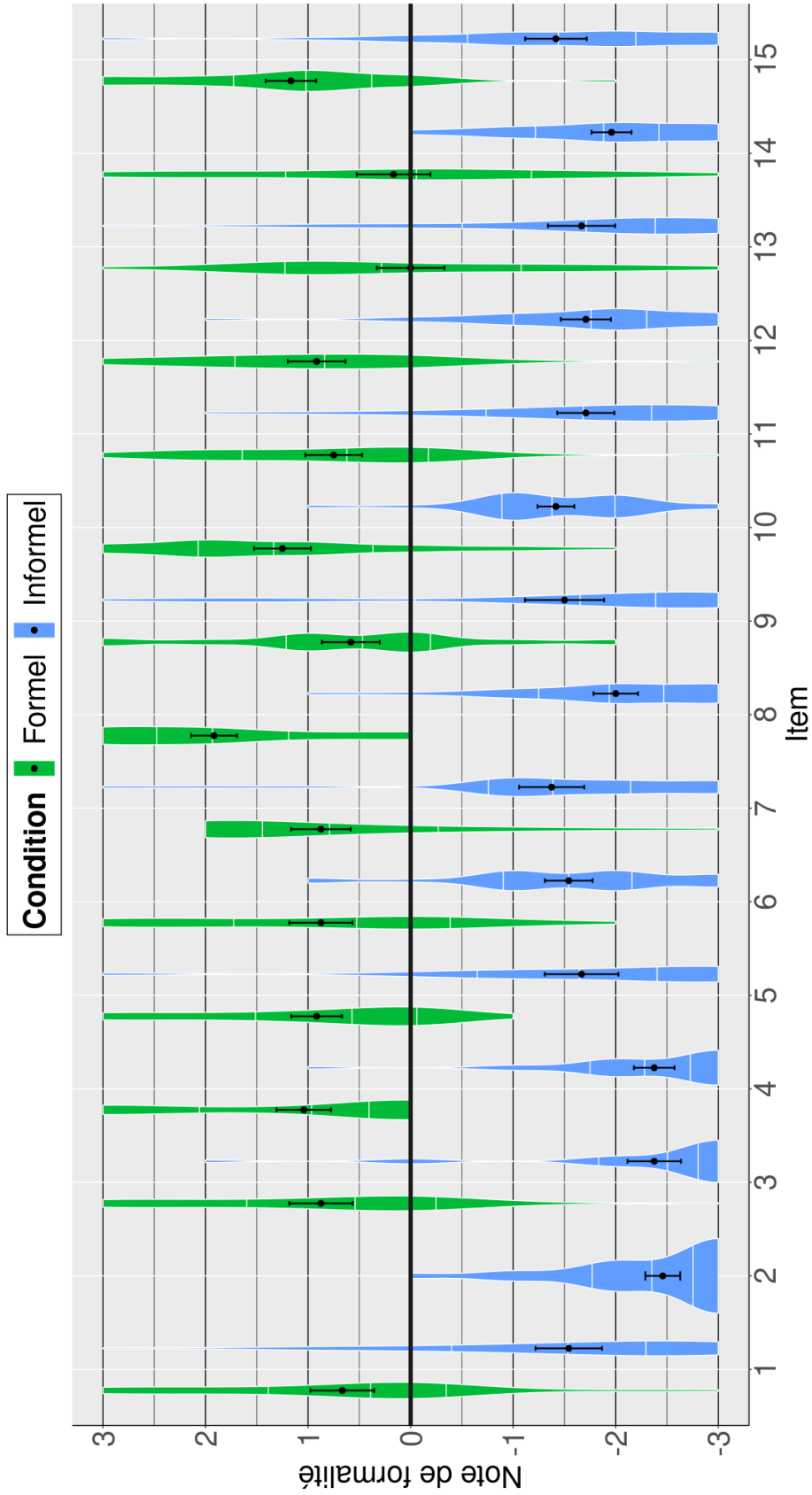


FIGURE E.4 – Résultats du norming (AJT2p, items 1-15)

E.0.6 -

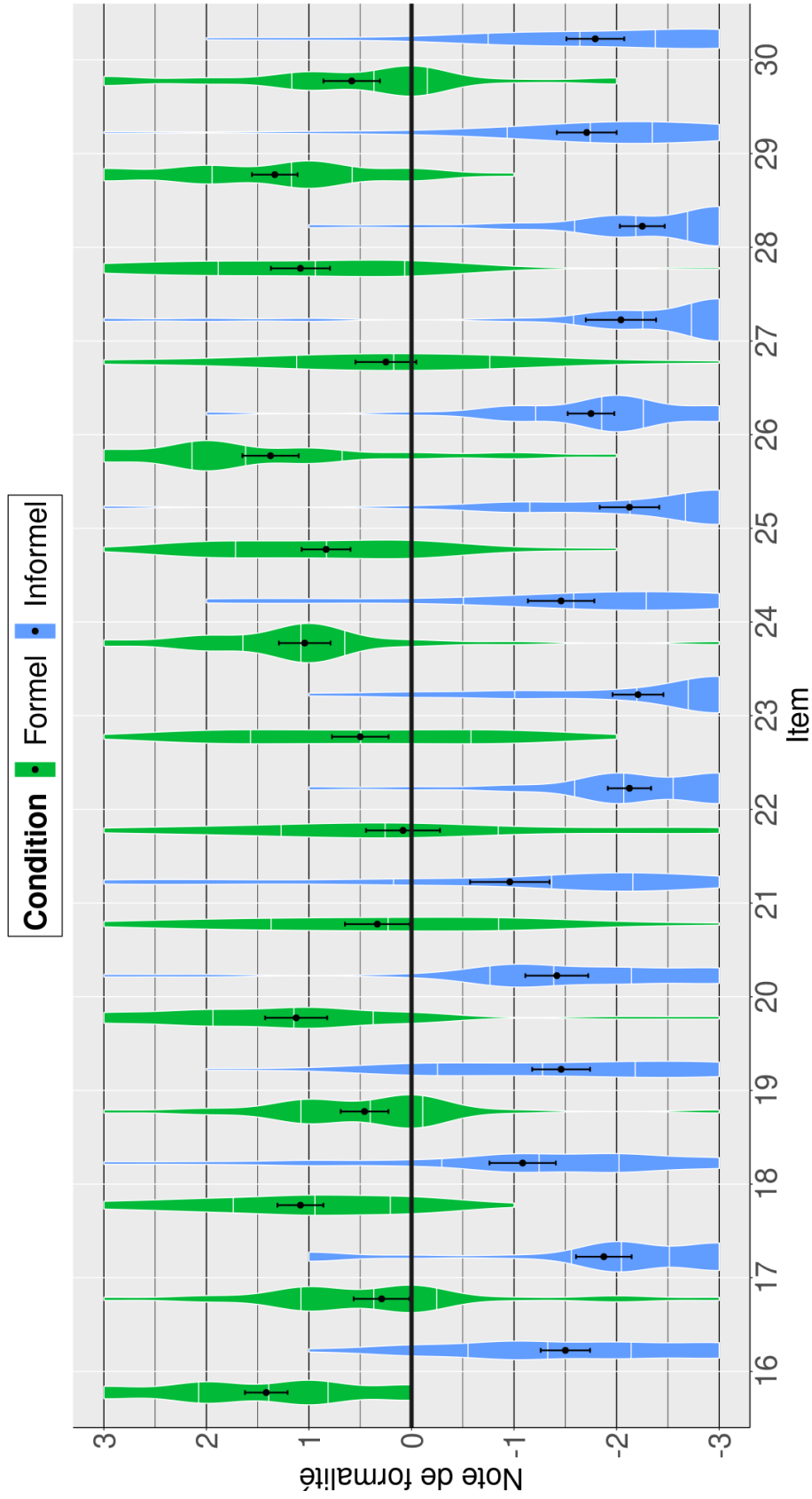


FIGURE E.5 – Résultats du norming (AJT2p, items 16-30)

Annexe F

AJT2 : Items et modèles statistiques

F.0.1 Items expérimentaux de l'étude AJT2 (5.2.2)

Type	Contexte	
	Informel	Formel
ajt2-001 F	A : Eh, Jean il arrive demain. B : À quelle heure il arrive?	A : Jean arrive demain. B : À quelle heure il arrive?
	A : Eh, Jean il arrive demain. B : À quelle heure arrive-t-il?	A : Jean arrive demain. B : À quelle heure arrive-t-il?
ajt2-001 IS	A : Eh, Jean il arrive demain. B : Il arrive à quelle heure?	A : Jean arrive demain. B : Il arrive à quelle heure?
	A : Les clefs, je les ai foutues dans le tiroir. B : Dans lequel tu les a mises?	A : J'ai mis les clefs dans le tiroir. B : Dans lequel tu les as mises?
ajt2-002 FINV	A : Les clefs, je les ai foutues dans le tiroir. B : Dans lequel les as-tu mises?	A : J'ai mis les clefs dans le tiroir. B : Dans lequel les as-tu mises?
	A : Les clefs, je les ai foutues dans le tiroir. B : Tu les as mises dans lequel?	A : J'ai mis les clefs dans le tiroir. B : Tu les as mises dans lequel?
ajt2-003 F	A : Pff, mon téléphone, je l'ai pétié. B : Comment tu l'as cassé?	A : J'ai cassé mon téléphone. B : Comment tu l'as cassé?
	A : Pff, mon téléphone, je l'ai pétié. B : Comment l'as-tu cassé?	A : J'ai cassé mon téléphone. B : Comment l'as-tu cassé?
ajt2-003 IS	A : Pff, mon téléphone, je l'ai pétié. B : Tu l'as cassé comment?	A : J'ai cassé mon téléphone. B : Tu l'as cassé comment?

TABLEAU F.1 – Items expérimentaux pour l'étude AJT2 (cibles 001 -003)

Type	Contexte	
	Informel	Formel
ajt2-004 F	A : Ouais, moi je me barre demain. B : Où tu vas?	A : Je pars demain. B : Où tu vas?
ajt2-004 FINV	A : Ouais, moi je me barre demain. B : Où vas-tu?	A : Je pars demain. B : Où vas-tu?
ajt2-004 IS	A : Ouais, moi je me barre demain. B : Tu vas où?	A : Je pars demain. B : Tu vas où?
ajt2-005 F	A : Eh, l'équation du prof, je l'ai résolue! B : Quel raisonnement tu as suivi?	A : J'ai résolu l'équation du professeur! B : Quel raisonnement tu as suivi?
ajt2-005 FINV	A : Eh, l'équation du prof, je l'ai résolue! B : Quel raisonnement as-tu suivi?	A : J'ai résolu l'équation du professeur! B : Quel raisonnement as-tu suivi?
ajt2-005 IS	A : Eh, l'équation du prof, je l'ai résolue! B : Tu as suivi quel raisonnement?	A : J'ai résolu l'équation du professeur! B : Tu as suivi quel raisonnement?
ajt2-006 F	A : T'as vu, ils l'ont enfin choisi, le nouveau boss! B : Qui c'est?	A : Ils ont enfin choisi le nouveau chef. B : Qui c'est?
ajt2-006 FINV	A : T'as vu, ils l'ont enfin choisi, le nouveau boss! B : Qui est-ce?	A : Ils ont enfin choisi le nouveau chef. B : Qui est-ce?
ajt2-006 IS	A : T'as vu, ils l'ont enfin choisi, le nouveau boss! B : C'est qui?	A : Ils ont enfin choisi le nouveau chef. B : C'est qui?

TABLEAU F.2 – Items expérimentaux pour l'étude AJT2 (cibles 003-006)

Type	Contexte	
	Informel	Formel
ajt2-007 F	A : Bon, vous les aurez plus tard, les dates du partiel. B : Quand on saura?	A : Vous aurez les dates d'examen plus tard. B : Quand on saura?
ajt2-007 FINV	A : Bon, vous les aurez plus tard, les dates du partiel. B : Quand saura-t-on?	A : Vous aurez les dates d'examen plus tard. B : Quand saura-t-on?
ajt2-007 IS	A : Bon, vous les aurez plus tard, les dates du partiel. B : On saura quand?	A : Vous aurez les dates d'examen plus tard. B : On saura quand?
ajt2-008 F	A : C'est vraiment un gars sympa, ce mec. B : De qui tu parles?	A : Cet homme est vraiment sympathique. B : De qui tu parles?
ajt2-008 FINV	A : C'est vraiment un gars sympa, ce mec. B : De qui parles-tu?	A : Cet homme est vraiment sympathique. B : De qui parles-tu?
ajt2-008 IS	A : C'est vraiment un gars sympa, ce mec. B : Tu parles de qui?	A : Cet homme est vraiment sympathique. B : Tu parles de qui?
ajt2-009 F	A : Yes, je l'ai enfin trouvé, mon sujet! B : Lequel tu as pris?	A : J'ai enfin trouvé mon sujet! B : Lequel tu as pris?
009 FINV	A : Yes, je l'ai enfin trouvé, mon sujet! B : Lequel as-tu pris?	A : J'ai enfin trouvé mon sujet! B : Lequel as-tu pris?
ajt2-009 IS	A : Yes, je l'ai enfin trouvé, mon sujet! B : Tu as pris lequel?	A : J'ai enfin trouvé mon sujet! B : Tu as pris lequel?

TABLEAU F.3 – Items expérimentaux pour l'étude AJT2 (cibles 007-009)

Type	Contexte
ajt2-010 F	A : Non mais il faut le faire autrement, ça! B : Avec quoi tu le ferais?
ajt2-010 FINV	A : Non mais il faut le faire autrement, ça! B : Avec quoi le ferais-tu?
ajt2-010 IS	A : Non mais il faut le faire autrement, ça! B : Tu le ferais avec quoi?
ajt2-011 F	A : Non, elle est pas là, ma sœur. B : Quand elle est partie?
ajt2-011 FINV	A : Non, elle est pas là, ma sœur. B : Quand est-elle partie?
ajt2-011 IS	A : Non, elle est pas là, ma sœur. B : Elle est partie quand?
ajt2-012 F	A : C'est pas tout de suite qu'ils la répareront, la digue... B : Pourquoi ils rouspètent?
ajt2-012 FINV	A : C'est pas tout de suite qu'ils la répareront, la digue... B : Pourquoi rouspètent-ils?
ajt2-012 IS	A : C'est pas tout de suite qu'ils la répareront, la digue... B : Ils rouspètent pourquoi?

TABLEAU F.4 – Items expérimentaux pour l'étude AJT2 (cibles 0010-012)

Type	Contexte	
	Informel	Formel
ajt2-013 F	A : Je sais pas ce que j'en ai fait, de ton agrafeuse... B : Où tu l'as utilisée?	A : Je ne sais pas ce que j'ai fait de ton agrafeuse... B : Où tu l'as utilisée?
ajt2-013 FINV	A : Je sais pas ce que j'en ai fait, de ton agrafeuse... B : Où l'as-tu utilisée?	A : Je ne sais pas ce que j'ai fait de ton agrafeuse... B : Où l'as-tu utilisée?
ajt2-013 IS	A : Je sais pas ce que j'en ai fait, de ton agrafeuse... B : Tu l'as utilisée où?	A : Je ne sais pas ce que j'ai fait de ton agrafeuse... B : Tu l'as utilisée où?
ajt2-014 F	A : Personne m'a rien donné, à moi! B : Alors à qui ils l'ont donné?	A : Personne ne m'a rien donné! B : Alors à qui ils l'ont donné?
ajt2-014 FINV	A : Personne m'a rien donné, à moi! B : Alors à qui l'ont-ils donné?	A : Personne ne m'a rien donné! B : Alors à qui l'ont-ils donné?
ajt2-014 IS	A : Personne m'a rien donné, à moi! B : Alors ils l'ont donné à qui?	A : Personne ne m'a rien donné! B : Alors ils l'ont donné à qui?
ajt2-015 F	A : J'ai pas réussi à le suivre en entier, le débat. B : À quel moment tu as arrêté?	A : Je n'ai pas réussi à suivre le débat en entier. B : À quel moment tu as arrêté?
ajt2-015 FINV	A : J'ai pas réussi à le suivre en entier, le débat. B : À quel moment as-tu arrêté?	A : Je n'ai pas réussi à suivre le débat en entier. B : À quel moment as-tu arrêté?
ajt2-015 IS	A : J'ai pas réussi à le suivre en entier, le débat. B : Tu as arrêté à quel moment?	A : Je n'ai pas réussi à suivre le débat en entier. B : Tu as arrêté à quel moment?

TABLEAU F.5 – Items expérimentaux pour l'étude AJT2 (cibles 013-015)

Type	Contexte	
ajt2-016 F	A : Cette machine, j'ai jamais eu de problème avec! B : Comment tu t'y prends?	A : Je n'ai jamais eu de problème avec cette machine. B : Comment tu t'y prends?
ajt2-016 FINV	A : Cette machine, j'ai jamais eu de problème avec! B : Comment t'y prends-tu?	A : Je n'ai jamais eu de problème avec cette machine. B : Comment t'y prends-tu?
ajt2-016 IS	A : Cette machine, j'ai jamais eu de problème avec! B : Tu t'y prends comment?	A : Je n'ai jamais eu de problème avec cette machine. B : Tu t'y prends comment?
ajt2-017 F	A : Elle, elle y sera pas tout seule. B : Qui elle connaît?	A : Elle n'y sera pas toute seule. B : Qui elle connaît?
ajt2-017 FINV	A : Elle, elle y sera pas tout seule. B : Qui connaît-t-elle?	A : Elle n'y sera pas toute seule. B : Qui connaît-t-elle?
ajt2-017 IS	A : Elle, elle y sera pas tout seule. B : Elle connaît qui?	A : Elle n'y sera pas toute seule. B : Elle connaît qui?
ajt2-018 F	A : Moi, j'aime pas tellement les voitures rouges. B : De quelle couleur tu les préfères?	A : Je n'aime pas tellement les voitures rouges. B : De quelle couleur tu les préfères?
ajt2-018 FINV	A : Moi, j'aime pas tellement les voitures rouges. B : De quelle couleur les préfères-tu?	A : Je n'aime pas tellement les voitures rouges. B : De quelle couleur les préfères-tu?
ajt2-018 IS	A : Moi, j'aime pas tellement les voitures rouges. B : Tu les préfères de quelle couleur?	A : Je n'aime pas tellement les voitures rouges. B : Tu les préfères de quelle couleur?

TABLEAU F.6 – Items expérimentaux pour l'étude AJT2 (cibles 016-018)

Type	Contexte	
	Informel	Formel
ajt2-019 F	A : Les premiers matches, ils étaient pas simples, mais on a gagné! B : Avec qui vous êtes battus?	A : Les premiers matches n'étaient pas simples, mais on a gagné! B : Avec qui vous êtes battus?
ajt2-019 FINV	A : Les premiers matches, ils étaient pas simples, mais on a gagné! B : Avec qui vous êtes-vous battus?	A : Les premiers matches n'étaient pas simples, mais on a gagné! B : Avec qui vous êtes-vous battus?
ajt2-019 IS	A : Les premiers matches, ils étaient pas simples, mais on a gagné! B : Vous vous êtes battus avec qui?	A : Les premiers matches n'étaient pas simples, mais on a gagné! B : Vous vous êtes battus avec qui?
ajt2-020 F	A : Toi, il faudrait pas que tu oublies de venir demain. B : Pourquoi tu fais dans le sous-entendu?	A : Il ne faudrait pas que tu oublies de venir demain. B : Pourquoi tu fais dans le sous-entendu?
ajt2-020 FINV	A : Toi, il faudrait pas que tu oublies de venir demain. B : Pourquoi fais-tu dans le sous-entendu?	A : Il ne faudrait pas que tu oublies de venir demain. B : Pourquoi fais-tu dans le sous-entendu?
ajt2-020 IS	A : Toi, il faudrait pas que tu oublies de venir demain. B : Tu fais dans le sous-entendu pourquoi?	A : Il ne faudrait pas que tu oublies de venir demain. B : Tu fais dans le sous-entendu pourquoi?
ajt2-021 F	A : Tu devineras jamais qui j'ai croisé hier à la fac! B : Qui tu as rencontré?	A : Tu ne devineras jamais qui j'ai croisé hier à l'université! B : Qui tu as rencontré?
ajt2-021 FINV	A : Tu devineras jamais qui j'ai croisé hier à la fac! B : Qui as-tu rencontré?	A : Tu ne devineras jamais qui j'ai croisé hier à l'université! B : Qui as-tu rencontré?
ajt2-021 IS	A : Tu devineras jamais qui j'ai croisé hier à la fac! B : Tu as rencontré qui?	A : Tu ne devineras jamais qui j'ai croisé hier à l'université! B : Tu as rencontré qui?

TABLEAU F.7 – Items expérimentaux pour l'étude AJT2 (cibles 019-021)

Type		Contexte
ajt2-022 F	A : J'ai pas réussi à ne pas choper de virus... B : Sur quels sites tu es allé?	A : Je n'ai pas réussi à ne pas attraper de virus... B : Sur quels sites tu es allé?
ajt2-022 FINV	A : J'ai pas réussi à ne pas choper de virus... B : Sur quels sites es-tu allé?	A : Je n'ai pas réussi à ne pas attraper de virus... B : Sur quels sites es-tu allé?
ajt2-022 IS	A : J'ai pas réussi à ne pas choper de virus... B : Tu es allé sur quels sites?	A : Je n'ai pas réussi à ne pas attraper de virus... B : Tu es allé sur quels sites?
ajt2-023 F	A : Y a pas trente-six soluces pour régler leur problème. B : Comment ton groupe ferait?	A : Il n'y a pas trente-six solutions pour régler leur problème. B : Comment ton groupe ferait?
ajt2-023 FINV	A : Y a pas trente-six soluces pour régler leur problème. B : Comment ferait ton groupe?	A : Il n'y a pas trente-six solutions pour régler leur problème. B : Comment ferait ton groupe?
ajt2-023 IS	A : Y a pas trente-six soluces pour régler leur problème. B : Ton groupe ferait comment?	A : Il n'y a pas trente-six solutions pour régler leur problème. B : Ton groupe ferait comment?
ajt2-024 F	A : Je sais vraiment plus comment faire, j'ai plus la foi. B : Avec quoi tu as des difficultés?	A : Je ne sais vraiment plus comment faire, je n'ai plus d'espoir. B : Avec quoi tu as des difficultés?
ajt2-024 FINV	A : Je sais vraiment plus comment faire, j'ai plus la foi. B : Avec quoi as-tu des difficultés?	A : Je ne sais vraiment plus comment faire, je n'ai plus d'espoir. B : Avec quoi as-tu des difficultés?
ajt2-024 IS	A : Je sais vraiment plus comment faire, j'ai plus la foi. B : Tu as des difficultés avec quoi?	A : Je ne sais vraiment plus comment faire, je n'ai plus d'espoir. B : Tu as des difficultés avec quoi?

TABLEAU F.8 – Items expérimentaux pour l'étude AJT2 (cibles 022-024)

Type	Contexte	
	Informel	Formel
ajt2-025 F	A : Mon frangin sait plus ce qu'il veut. B : Lequel il a pris?	A : Mon frère ne sait plus ce qu'il veut. B : Lequel il a pris?
ajt2-025 FINV	A : Mon frangin sait plus ce qu'il veut. B : Lequel a-t-il pris?	A : Mon frère ne sait plus ce qu'il veut. B : Lequel a-t-il pris?
ajt2-025 IS	A : Mon frangin sait plus ce qu'il veut. B : Il a pris lequel?	A : Mon frère ne sait plus ce qu'il veut. B : Il a pris lequel?
ajt2-026 F	A : C'est vraiment pas possible de raisonner ce mec. B : De qui tu parles?	A : Il est vraiment impossible de raisonner cet homme. B : De qui tu parles?
ajt2-026 FINV	A : C'est vraiment pas possible de raisonner ce mec. B : De qui parles-tu?	A : Il est vraiment impossible de raisonner cet homme. B : De qui parles-tu?
ajt2-026 IS	A : C'est vraiment pas possible de raisonner ce mec. B : Tu parles de qui?	A : Il est vraiment impossible de raisonner cet homme. B : Tu parles de qui?
ajt2-027 F	A : Je veux pas aller dans ce trou paumé! B : Où ils t'envoient?	A : Je ne veux pas aller dans cette région! B : Où ils t'envoient?
ajt2-027 FINV	A : Je veux pas aller dans ce trou paumé! B : Où t'envoient-ils?	A : Je ne veux pas aller dans cette région! B : Où t'envoient-ils?
ajt2-027 IS	A : Je veux pas aller dans ce trou paumé! B : Ils t'envoient où?	A : Je ne veux pas aller dans cette région! B : Ils t'envoient où?

TABLEAU F.9 – Items expérimentaux pour l'étude AJT2 (cibles 025-027)

Type		Contexte
ajt2-028 F	A : Y a aucune chance qu'elle se ramène. B : Pourquoi tu dis ça?	A : Il n'y a aucune chance qu'elle vienne. B : Pourquoi tu dis ça?
ajt2-028 FINV	A : Y a aucune chance qu'elle se ramène. B : Pourquoi dis-tu ça?	A : Il n'y a aucune chance qu'elle vienne. B : Pourquoi dis-tu ça?
ajt2-028 IS	A : Y a aucune chance qu'elle se ramène. B : Tu dis ça pourquoi?	A : Il n'y a aucune chance qu'elle vienne. B : Tu dis ça pourquoi?
ajt2-029 F	A : Il sera plus du tout dispo pendant ses vacances. B : Quand il revient?	A : Il ne sera plus du tout disponible pendant ses congés. B : Quand il revient?
ajt2-029 FINV	A : Il sera plus du tout dispo pendant ses vacances. B : Quand revient-il?	A : Il ne sera plus du tout disponible pendant ses congés. B : Quand revient-il?
ajt2-029 IS	A : Il sera plus du tout dispo pendant ses vacances. B : Il revient quand?	A : Il ne sera plus du tout disponible pendant ses congés. B : Il revient quand?
ajt2-030 F	A : Y a longtemps qu'on joue plus au poker avec le gang. B : À quoi vous jouez?	A : Il y a longtemps qu'on ne joue plus au poker avec le groupe. B : À quoi vous jouez?
ajt2-030 FINV	A : Y a longtemps qu'on joue plus au poker avec le gang. B : À quoi jouez-vous?	A : Il y a longtemps qu'on ne joue plus au poker avec le groupe. B : À quoi jouez-vous?
ajt2-030 IS	A : Y a longtemps qu'on joue plus au poker avec le gang. B : Vous jouez à quoi?	A : Il y a longtemps qu'on ne joue plus au poker avec le groupe. B : Vous jouez à quoi?

TABLEAU F.10 – Items expérimentaux pour l'étude AJT2 (cibles 028-030)

-	Item
ajt2-041	A : Il a bien embrassé le témoin pendant le mariage? B : Non, c'est la belle-mère qu'il a embrassée pendant le mariage.
ajt2-042	A : Avez-vous félicité les joueurs après le match? B : C'est le capitaine que j'ai félicité juste après.
ajt2-043	A : Tu vas bien acheter le champagne à une supérette? B : Non. La bière, je vais l'acheter à une supérette.
ajt2-044	A : Parle-t-il les langues comme un natif? B : Le vietnamien, il le parle depuis l'enfance.
ajt2-045	A : Elles ont bien trafiqué les lames sur places*? B : Non. Ce sont les pointes qu'elles ont trafiquées sur place.
ajt2-046	A : Devons-nous remplir des formulaires dans la semaine? B : C'est la déclaration d'emploi que vous devez remplir dans la semaine.
ajt2-047	A : Il a bien posté les colis la semaine dernière? B : Non. Les recommandés, ils les a postés la semaine dernière.
ajt2-048	A : As-tu installé les logiciels depuis les CDs? B : La suite bureautique, je l'ai installée depuis le CD n°1.
ajt2-049	A : Vous allez bien réparer l'alternateur en priorité? B : Non. C'est la batterie qu'on va réparer en priorité.
ajt2-050	A : Avez-vous couché les enfants avant vingt heures? B : C'est Léon que j'ai couché à dix-neuf heures.
ajt2-051	A : Il explorait bien l'Arctique pendant son dernier voyage? B : Non. L'Antarctique, il l'explorait pendant son dernier voyage.
ajt2-052	A : Passeront-elles les examens en fin d'année? B : Le brevet, elles le passeront début juin.
ajt2-053	A : Elle a bien battu la Belgique cette fois? B : Non. C'est le Luxembourg qu'elle a battu cette fois.
ajt2-054	A : A-t-elle viré des employés dernièrement? B : C'est la secrétaire de direction qu'elle a virée hier.
ajt2-055	A : Vous avez bien commencé "Twin Peaks" récemment? B : Non. "On the Air", on l'a commencé récemment.
ajt2-056	A : Doit-il faire le nettoyage ce week-end? B : Les fenêtres, il doit les faire samedi.
ajt2-057	A : Elle a bien fini la Comédie Humaine le mois dernier? B : Non. C'est les Rougeon*-Macquart qu'elle a finis le mois dernier.
ajt2-058	A : Dévorait-il les garçons dans son antre? B : Ce sont les mauvais garçons qu'il dévorait sur sa grande table.
ajt2-059	A : Je peux bien ranger les bermudas après la fermeture? B : Non. Les jeans, tu peux les ranger après la fermeture.
ajt2-060	A : Écrivait-il les films avec ses collaborateurs? B : Les scènes de combat, il les écrivait avec d'autres artistes martiaux.

Ces items (041-060) proviennent de [RIOU et HEMFORTH, 2015]

**Ces typos, présentes dans les items, ne sont imputables qu'à ma négligence.*

TABLEAU F11 – Items expérimentaux pour l'étude AJT2 (distracteurs 041-060)

	Item
ajt2-101	A : Oh, t'as vu, y a un oiseau dehors! B : Il est trop beau!
ajt2-102	A : Je suis tombé malade cette nuit. B : Moi je vais manger avec ma mère.
ajt2-103	A : Il paraît que le Brésil est en train de changer de politique étrangère. B : Je n'ai aucune idée d'où sont le Brésil sur une carte.
ajt2-031	A : J'ai vu mes parents hier... B : Ta mère pas va mieux?
ajt2-032	A : Y a un scooter qui m'a coupé la route tout à l'heure. B : Tu l'avais quand même pas renversé?
ajt2-033	A : Plus jamais je ne skierai avec elle. B : Elle va trop trop vite, hein?
ajt2-034	A : Tu voudrais pas venir m'aider? B : Non, clairement.
ajt2-035	A : À la fin, il est parti. B : Pourquoi donc pas?
ajt2-036	A : Le général de brigade nous a dit d'aller demander à son aide. B : Tu sais c'est qui son aide?
ajt2-037	A : Elles nous ont dit de venir demain soir. B : D'accord, on ira hier alors?
ajt2-038	A : Une éclipse, c'est quand le soleil cache la lune. B : Tu ne crois que c'est l'inverse?
ajt2-039	A : Les chats ne font pas des chiens... B : Ils font quoi?
ajt2-040	A : Les paris sont ouverts! B : Où ça?

TABLEAU F.12 – Items expérimentaux pour l'étude AJT2 (contrôles 031-040 et entraînements 101-103)

F.0.2 Consigne (AJT2) (5.2.2)

Dans cette expérience, il vous est demandé de noter l'acceptabilité des phrases soulignées qui vont apparaître à l'écran, sur une échelle de 0 (pas du tout acceptable) à 10 (très acceptable). Elles sont précédées d'un contexte, qui n'est pas à noter.

Une phrase est très acceptable si elle est bien formée et adaptée au contexte (= si vous pourriez vous-même la produire dans le contexte où elle apparaît). Une phrase n'est pas du tout acceptable si elle est mal formée et/ou non-adaptée, dans le contexte où elle apparaît.

Utilisez s'il vous plaît l'intégralité de l'échelle (de 0 à 10).

Merci de ne pas passer cette expérience si vous avez déjà répondu à un de mes appels précédents ces dernières semaines.

L'ensemble devrait vous prendre environ quinze minutes. Merci de votre participation !

L'ensemble des données collectées durant cette expérience le sont à des fins de recherche. Aucune donnée personnelle d'identification ne sera transmise à des tiers.

Votre âge :	<input type="text"/>
Votre nationalité :	<input type="text"/>
Votre genre :	<input type="radio"/> Homme <input type="radio"/> Femme <input type="radio"/> Autre <input type="radio"/> Je ne veux pas répondre à cette question
Votre dernier diplôme obtenu :	<input type="radio"/> Avant le baccalauréat <input type="radio"/> Baccalauréat ou équivalent <input type="radio"/> Bac+2 ou équivalent <input type="radio"/> Bac+3 ou équivalent <input type="radio"/> Bac+5 ou équivalent <input type="radio"/> Bac+8 ou équivalent
Votre domaine d'études ou d'activité professionnelle :	<input type="text"/>
Votre région de naissance :	<input type="text"/>
Votre langue maternelle :	<input type="text"/>
Autres langues:	Niveau: de 1 (mauvais) à 10 (excellent)
<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>
<input type="checkbox"/> Je suis majeur.e et j'accepte les termes et conditions de cette expérience.	

[→ Cliquez ici pour continuer](#)

FIGURE F1 – Consigne (AJT2)

F.0.3 Modèle *brms* et distributions postérieures (AJT2) (5.2.2)

Family : cumulative

Links : mu = logit; disc = identity

Formula : answer ~ (f_finv + is_finv) * old_young * formal_informal +
((f_finv + is_finv) * formal_informal | id) + ((f_finv + is_finv) * old_young * formal_informal |
Item)

Data : DON (Number of observations : 1320)

Samples : 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-5.33	0.40	-6.12	-4.54	1.00	1707	3600
Intercept[2]	-4.52	0.39	-5.28	-3.77	1.00	1615	3089
Intercept[3]	-3.41	0.37	-4.13	-2.70	1.00	1506	3054
Intercept[4]	-2.84	0.36	-3.56	-2.14	1.00	1442	2852
Intercept[5]	-2.42	0.36	-3.13	-1.72	1.00	1400	2788
Intercept[6]	-2.02	0.36	-2.72	-1.31	1.00	1370	2736
Intercept[7]	-1.16	0.35	-1.86	-0.47	1.00	1338	2644
Intercept[8]	-0.42	0.35	-1.12	0.27	1.00	1343	2666
Intercept[9]	0.44	0.35	-0.25	1.13	1.00	1315	2699
Intercept[10]	1.66	0.36	0.97	2.36	1.00	1335	2650
f_finv	-1.97	0.47	-2.90	-1.06	1.00	2598	4018
is_finv	-1.49	0.39	-2.28	-0.72	1.00	2824	4740
old_young	-1.31	0.56	-2.45	-0.24	1.00	1525	3173
formal_informal	0.26	0.18	-0.09	0.62	1.00	7616	8677
f_finv :old_young	-1.60	0.86	-3.34	0.11	1.00	2212	3715
is_finv :old_young	-1.70	0.76	-3.24	-0.21	1.00	2618	4388
f_finv :formal_informal	-0.39	0.36	-1.09	0.33	1.00	6580	7303
is_finv :formal_informal	-0.46	0.33	-1.08	0.19	1.00	7249	7969
old_young :formal_informal	0.08	0.34	-0.60	0.74	1.00	6960	8185
f_finv :old_young :formal_informal	-1.04	0.68	-2.43	0.29	1.00	7814	7825
is_finv :old_young :formal_informal	-0.01	0.63	-1.28	1.21	1.00	8885	8048

TABLEAU F.13 – Résumé du modèle *brms* d'analyse de AJT2

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Type : F vs. FINV	-1.97	0.00	1.00	-3.82	-0.04	-2.90	-1.06	-2.74	-1.20	-2.65	-1.31
Type : IS vs. FINV	-1.49	0.00	1.00	-3.29	-0.01	-2.28	-0.72	-2.14	-0.85	-2.05	-0.92
Âge : +30 ans vs. -30 ans	-1.31	0.01	0.99	-3.51	0.87	-2.45	-0.24	-2.25	-0.41	-2.13	-0.52
Formalité : Formel vs. Informel	0.26	0.93	0.07	-0.47	1.12	-0.09	0.62	-0.03	0.55	0.00	0.52
F vs. FINV * Âge	-1.60	0.03	0.97	-5.12	2.37	-3.34	0.11	-3.03	-0.19	-2.84	-0.39
IS vs. FINV * Âge	-1.70	0.01	0.99	-5.04	2.00	-3.24	-0.21	-2.97	-0.47	-2.77	-0.63
F vs. FINV * Formalité	-0.39	0.13	0.87	-1.89	1.18	-1.09	0.33	-0.96	0.20	-0.89	0.12
IS vs. FINV * Formalité	-0.46	0.08	0.92	-1.91	1.06	-1.08	0.19	-0.98	0.08	-0.92	0.01
Âge * Formalité	0.08	0.59	0.41	-1.57	1.56	-0.60	0.74	-0.48	0.64	-0.40	0.57
F vs. FINV * Âge * Formalité	-1.04	0.06	0.94	-3.92	2.31	-2.43	0.29	-2.17	0.08	-2.03	-0.08
IS vs. FINV * Âge * Formalité	-0.01	0.50	0.50	-3.13	2.36	-1.28	1.21	-1.05	1.02	-0.91	0.88

TABLEAU F.14 – Distributions postérieures pour le modèle brms maximal d'analyse de AJT2

Annexe G

AJT3 : Norming des voix préalables à l'étude

Pour préparer l'étude AJT3 (Section 5.2.3) sous une modalité orale, les items utilisés pour l'étude AJT2 (Annexe F.0.1) ont été enregistrés puis retravaillés avec le logiciel Audacity [AUDACITYTEAM, 1999-2020]. Pour assurer que le moins de paramètres extérieurs viennent fausser les résultats de l'expérience principale, un norming préalable a été effectué sur une partie des items oraux créés. Il s'agissait de vérifier qu'aucune voix ne soit plus *familière* par défaut que les autres. De plus, il s'agissait de regarder si les volontaires attribuaient un genre particulier aux quatre voix.

G.0.1 Volontaires (AJT3p, 5.2.3)

Nombre de volontaires : 35 francophones

Genre des volontaires : 26 femmes, 9 hommes

Âge des volontaires : 19-72 ans, $\mu=35.9$, $m=28$

Recrutement : RISC et réseaux sociaux

Compensation : Aucune

G.0.2 Protocole et design (AJT3p, 5.2.3)

Tâche : Jugements sur le genre et la familiarité d'une voix

Dates : Avril 2018

Durée : 10 minutes

Modalité : Expérience en ligne, via IbexFarm [DRUMMOND, 2016]

Design : 1 variable VOIX, 4 conditions (FEM1, FEM2, HOM1, HOM2)

- + pour échantillonner le mieux possible, variables inter-items : sur les 32 items,
 - 8 étaient de futurs distracteurs : *ajtp* 42, 44, 46, 48, 49, 52, 56, 58
 - 24 étaient de futures moitiés de dialogues cibles, dont :
 - 12 phrases de contexte (6 formels, 6 informels) : *ajtp* 02, 04, 05, 06, 09, 11, 16, 23, 26, 29, 30, 25
 - 12 phrases interrogatives (4 antéposition avec inversion, 4 in situ, 4 antéposition simple) : *ajtp* 01, 07, 10, 13, 14, 15, 17, 18, 19, 21, 22, 27

— les 4 entraînements étaient de futurs distracteurs : *ajtp* 31, 35, 38, 39

Carré latin : Oui, chaque volontaire n'écoutait qu'une seule version d'un item et écoutait un nombre égal de chaque condition de la variable intra-items sur toute l'expérience (8 items avec chacune des 4 VOIX).

Randomisation : Oui

Nombre d'items : 4 entraînements (1 par voix) + 32 items

Échelles :

— pourrait être un homme/une femme, 1-6 ("Pas du tout d'accord" – "Totalelement d'accord")

→ la moitié des items passaient avec *être un homme* / la moitié avec *être une femme*

— s'exprime de façon familière, 1-6 ("Pas du tout d'accord" – "Totalelement d'accord")

Consigne :

Dans cette expérience, il vous est demandé d'écouter de brefs extraits audio.

Vérifiez par ailleurs que le volume sonore est activé et bien réglé sur votre appareil avant de commencer l'expérience. Vous pourrez réécouter chaque extrait autant de fois que vous le voudrez.

Après votre écoute, il vous est demandé d'estimer :

- si la personne qui parle pourrait être un homme/une femme.
- si la personne qui parle s'exprime familièrement.

Utilisez s'il vous plaît l'intégralité des échelles fournies : de "Pas du tout d'accord" (gauche) à "Totalelement d'accord" (droite).

Quatre phrases d'entraînement vous seront proposées, avant de passer à l'expérience proprement dite.

L'ensemble devrait vous prendre moins de dix minutes.

Merci de votre participation !

Les données collectées durant cette expérience le sont à des fins de recherche. Aucune donnée personnelle d'identification ne sera transmise à des tiers.

Votre âge :

Votre genre :

- Homme Femme Autre
 Je ne veux pas répondre à cette question

Votre langue maternelle :

Je suis majeur.e et j'accepte les termes et conditions de cette expérience.

→ [Cliquez ici pour continuer.](#)

FIGURE G.1 – Consigne (AJT3p - Norming)

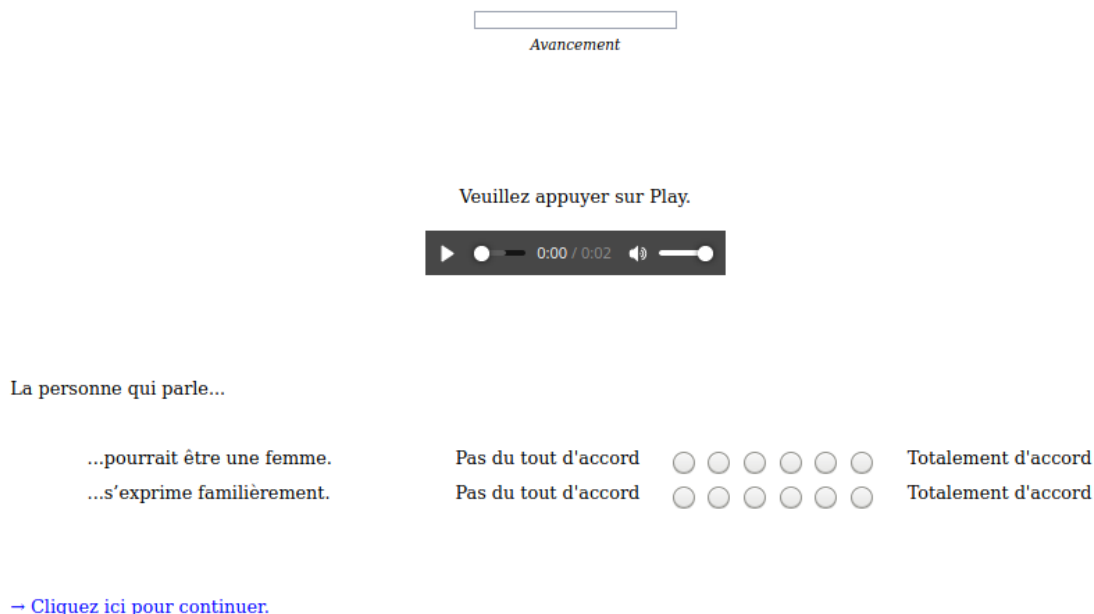
Capture d'écran :

FIGURE G.2 – Capture d'écran (AJT3p - Norming)

G.0.3 Liste des items (AJT3p, 5.2.3)

Les items sont disponibles en Annexe E0.1 sous les références : *ajtp* 01, 02, 04, 05, 06, 07, 09, 10, 11, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 25, 26, 27, 29, 30, 31, 35, 38, 39, 42, 44, 46, 48, 49, 52, 56, 58

G.0.4 Résultats - Genre des voix (AJT3p, 5.2.3)

Plus de détails dans les fichiers bruts disponibles dans le répertoire osf.io/ud4ks.

Les notes reçues, entre 1 et 6, ont été codées entre 0 et 5.

Voix	Échelle	0=Non, 5=Oui	Écart-type
FEM1	la personne pourrait être une femme	4.75	0.81
FEM1	la personne pourrait être un homme	0.36	1.05
FEM2	la personne pourrait être une femme	4.82	0.62
FEM2	la personne pourrait être un homme	0.27	0.89
HOM1	la personne pourrait être une femme	0.20	0.67
HOM1	la personne pourrait être un homme	4.71	0.91
HOM2	la personne pourrait être une femme	0.36	0.95
HOM2	la personne pourrait être un homme	4.78	0.61

TABLEAU G.1 – Résultats du norming - Genre des voix (AJT3p)

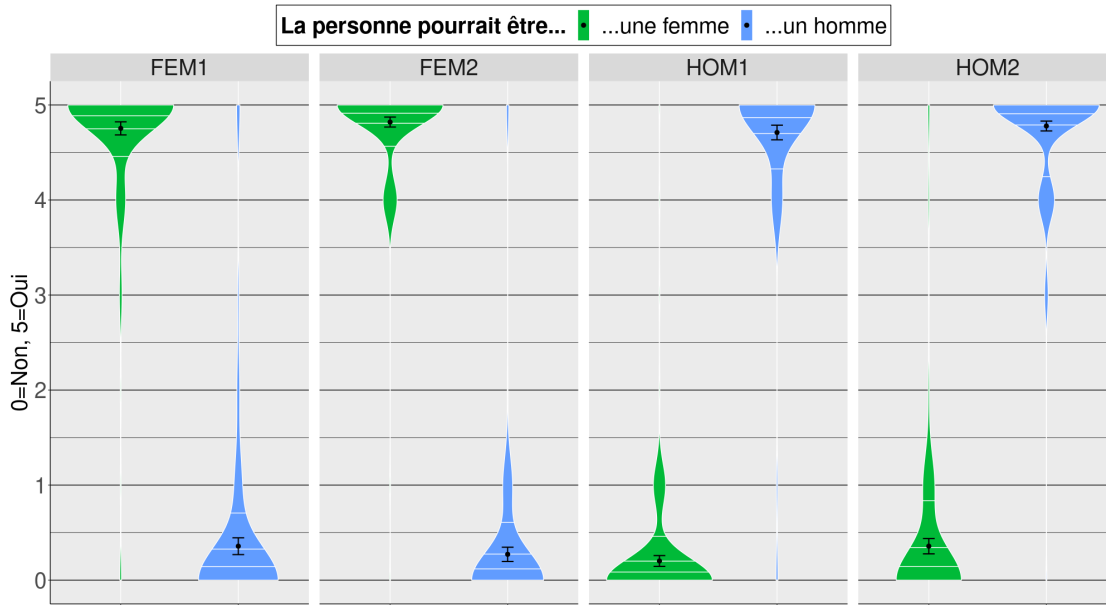


FIGURE G.3 – Résultats du norming - Genre des voix (AJT3p)

G.0.5 Résultats - Familiarité des voix (AJT3p, 5.2.3)

Les notes reçues, entre 1 et 6, ont été codées entre 0 et 5. Une note de 0 signifie une voix pas du tout familière; une note de 5 signifie une voix très familière.

Voix	Moyenne	Écart-type
FEM1	2.33	1.75
FEM2	2.49	1.71
HOM1	2.38	1.74
HOM2	2.41	1.74

TABLEAU G.2 – Résultats du norming - Familiarité des voix (AJT3p)

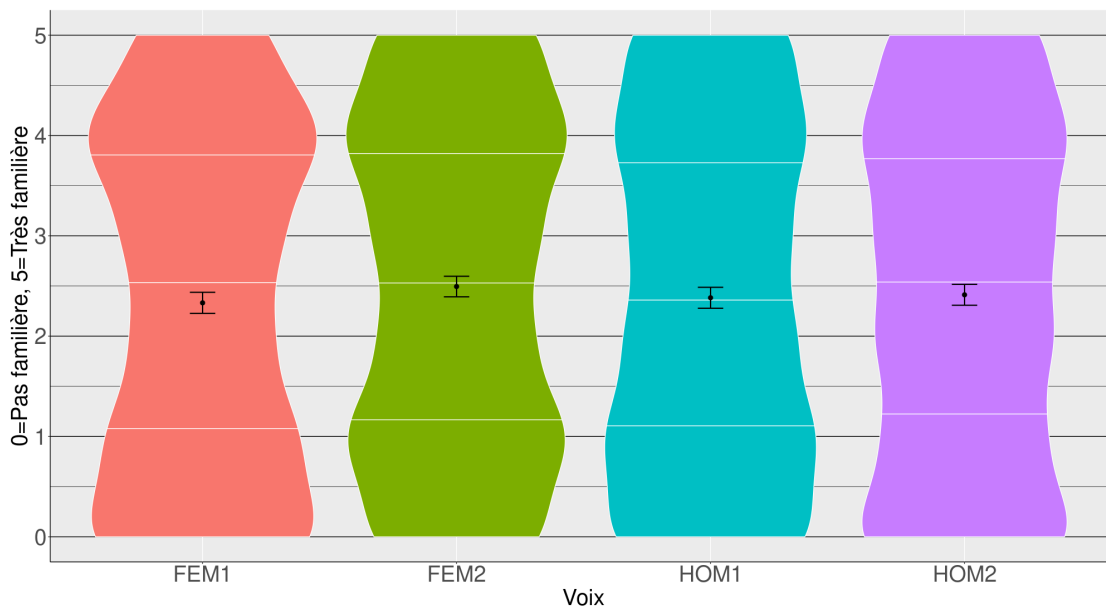


FIGURE G.4 – Résultats du norming - Familiarité des voix (AJT3p)

Annexe H

AJT3 : Matériels et modèles statistiques

H.0.1 Consigne (AJT3) (5.2.3)

Dans cette expérience, il vous est demandé d'écouter de brefs extraits audio.

Vérifiez que le volume sonore est activé et bien réglé sur votre appareil avant de commencer l'expérience. Vous pourrez réécouter chaque extrait autant de fois que vous le voudrez.

Après votre écoute, il vous est demandé de donner votre avis sur la seconde personne qui s'exprime :

- si elle parle un bon français (= grammaticalement correct).
- si elle parle un français adapté à la situation (= si vous pourriez vous exprimer ainsi dans le contexte).

Utilisez s'il vous plaît l'intégralité des échelles fournies : de "Pas du tout d'accord" (gauche) à "Totalelement d'accord" (droite).

Trois phrases d'entraînement vous seront proposées, avant de passer à l'expérience proprement dite.

L'ensemble devrait vous prendre une petite vingtaine de minutes.
Merci de votre participation !

Les données collectées durant cette expérience le sont à des fins de recherche. Aucune donnée personnelle d'identification ne sera transmise à des tiers.

Votre âge :	<input type="text"/>
Votre nationalité :	<input type="text"/>
Votre genre :	<input type="radio"/> Homme <input type="radio"/> Femme <input type="radio"/> Autre <input type="radio"/> Je ne veux pas répondre à cette question
Votre dernier diplôme obtenu :	<input type="radio"/> Avant le baccalauréat <input type="radio"/> Baccalauréat ou équivalent <input type="radio"/> Bac+2 ou équivalent <input type="radio"/> Bac+3 ou équivalent <input type="radio"/> Bac+5 ou équivalent <input type="radio"/> Bac+8 ou équivalent
Votre domaine d'études ou d'activité professionnelle :	<input type="text"/>
Votre région de naissance :	<input type="text"/>
Votre langue maternelle :	<input type="text"/>
Autres langues:	Niveau: de 1 (mauvais) à 10 (excellent)
<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>
<input type="checkbox"/> Je suis majeur.e et j'accepte les termes et conditions de cette expérience.	

→ [Cliquez ici pour continuer.](#)

FIGURE H.1 – Consigne (AJT3)

H.0.2 Modèle *brms* et distributions postérieures pour la question "La personne (...) qui s'exprime parle un *bon français*" (5.2.3)

Family : cumulative

Links : mu = logit; disc = identity

Formula : answer1 ~ (f_finv + is_finv) * old_young * formal_informal + ((f_finv + is_finv) * formal_informal | id) + ((f_finv + is_finv) * old_young * formal_informal | Item)

Data : DON2 (Number of observations : 1855)

Samples : 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-4.82	0.28	-5.38	-4.30	1.00	2098	4059
Intercept[2]	-3.88	0.26	-4.41	-3.39	1.00	1978	3854
Intercept[3]	-2.92	0.25	-3.41	-2.44	1.00	1927	3712
Intercept[4]	-1.55	0.24	-2.02	-1.09	1.00	1942	3795
Intercept[5]	0.32	0.23	-0.15	0.77	1.00	1979	3927
f_finv	-3.27	0.38	-4.04	-2.54	1.00	3001	4584
is_finv	-2.84	0.36	-3.55	-2.14	1.00	3137	4937
old_young	-0.01	0.39	-0.77	0.75	1.00	1596	3003
formal_informal	0.28	0.17	-0.06	0.60	1.00	5496	7636
f_finv:old_young	0.57	0.65	-0.69	1.83	1.00	2579	4671
is_finv:old_young	0.81	0.66	-0.51	2.12	1.00	2655	5193
f_finv:formal_informal	-0.46	0.37	-1.20	0.25	1.00	5721	6956
is_finv:formal_informal	-0.31	0.42	-1.14	0.50	1.00	5479	7647
old_young:formal_informal	0.16	0.33	-0.50	0.83	1.00	5400	7379
f_finv:old_young:formal_informal	0.49	0.69	-0.89	1.87	1.00	6688	7480
is_finv:old_young:formal_informal	0.52	0.82	-1.07	2.17	1.00	6144	7851

TABLEAU H.1 – Résumé du modèle *brms* d'analyse de AJT3 - Bon français

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Type : F vs. FINV	-3.27	0.00	1.00	-4.94	-1.85	-4.04	-2.54	-3.90	-2.66	-3.82	-2.73
Type : IS vs. FINV	-2.84	0.00	1.00	-4.38	-1.58	-3.55	-2.14	-3.44	-2.26	-3.36	-2.33
Âge : +30 ans vs. -30 ans	-0.01	0.49	0.51	-1.72	1.59	-0.77	0.75	-0.65	0.63	-0.57	0.55
Formalité : Formel vs. Informel	0.28	0.95	0.05	-0.39	0.96	-0.06	0.60	-0.00	0.55	0.03	0.52
F vs. FINV * Âge	0.57	0.81	0.19	-1.99	3.30	-0.69	1.83	-0.47	1.65	-0.35	1.50
IS vs. FINV * Âge	0.81	0.89	0.11	-1.68	3.79	-0.51	2.12	-0.29	1.90	-0.14	1.75
F vs. FINV * Formalité	-0.46	0.10	0.90	-2.30	0.92	-1.20	0.25	-1.07	0.13	-0.99	0.05
IS vs. FINV * Formalité	-0.31	0.23	0.77	-2.12	1.36	-1.14	0.50	-1.00	0.37	-0.91	0.28
Âge * Formalité	0.16	0.69	0.31	-1.22	1.69	-0.50	0.83	-0.40	0.71	-0.32	0.63
F vs. FINV * Âge * Formalité	0.49	0.76	0.24	-2.15	3.09	-0.89	1.87	-0.65	1.62	-0.51	1.47
IS vs. FINV * Âge * Formalité	0.52	0.74	0.26	-2.47	3.82	-1.07	2.17	-0.81	1.88	-0.65	1.71

TABLEAU H.2 – Distributions postérieures pour le modèle brms d'analyse de AJT3 - Bon français

H.0.3 Modèle *brms* et distributions postérieures pour la question "La personne (...) qui s'exprime parle un français adapté" (5.2.3)

Family : cumulative

Links : mu = logit; disc = identity

Formula : answer2 ~ (f_finv + is_finv) * old_young * formal_informal + ((f_finv + is_finv) * formal_informal | id) + ((f_finv + is_finv) * old_young * formal_informal | Item)

Data : DON2 (Number of observations : 1804)

Samples : 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-4.33	0.29	-4.91	-3.78	1.00	2004	3727
Intercept[2]	-3.42	0.27	-3.97	-2.90	1.00	1873	3475
Intercept[3]	-2.72	0.26	-3.25	-2.21	1.00	1827	3417
Intercept[4]	-1.62	0.26	-2.13	-1.12	1.00	1760	3422
Intercept[5]	-0.12	0.25	-0.62	0.37	1.00	1713	3203
f_finv	-0.01	0.22	-0.44	0.43	1.00	5709	7180
is_finv	0.41	0.23	-0.05	0.86	1.00	5957	7444
old_young	-0.29	0.34	-0.97	0.36	1.00	2118	3856
formal_informal	0.19	0.14	-0.09	0.45	1.00	8732	8095
f_finv:old_young	-0.70	0.38	-1.47	0.05	1.00	6208	7730
is_finv:old_young	-0.80	0.43	-1.65	0.05	1.00	6107	7158
f_finv:formal_informal	-0.59	0.33	-1.27	0.05	1.00	8896	8079
is_finv:formal_informal	-0.45	0.40	-1.24	0.34	1.00	7802	8120
old_young:formal_informal	0.26	0.26	-0.26	0.77	1.00	8352	8089
f_finv:old_young:formal_informal	-0.20	0.58	-1.35	0.94	1.00	10854	8962
is_finv:old_young:formal_informal	0.68	0.77	-0.83	2.21	1.00	6666	8603

TABLEAU H.3 – Résumé du modèle *brms* d'analyse de AJT3 - Français adapté

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Type : F vs. FINV	-0.01	0.48	0.52	-1.07	0.86	-0.44	0.43	-0.37	0.35	-0.32	0.30
Type : IS vs. FINV	0.41	0.96	0.04	-0.48	1.36	-0.05	0.86	0.03	0.78	0.08	0.73
Âge : +30 ans vs. -30 ans	-0.29	0.19	0.81	-1.65	1.02	-0.97	0.36	-0.85	0.25	-0.78	0.19
Formalité : Formel vs. Informel	0.19	0.92	0.08	-0.37	0.73	-0.09	0.45	-0.04	0.41	-0.00	0.39
F vs. FINV * Âge	-0.70	0.03	0.97	-2.35	0.93	-1.47	0.05	-1.33	-0.08	-1.24	-0.16
IS vs. FINV * Âge	-0.80	0.03	0.97	-2.47	0.98	-1.65	0.05	-1.51	-0.09	-1.42	-0.18
F vs. FINV * Formalité	-0.59	0.04	0.96	-1.96	1.16	-1.27	0.05	-1.14	-0.05	-1.07	-0.12
IS vs. FINV * Formalité	-0.45	0.12	0.88	-2.18	1.46	-1.24	0.34	-1.10	0.20	-1.02	0.12
Âge * Formalité	0.26	0.84	0.16	-0.76	1.32	-0.26	0.77	-0.18	0.69	-0.12	0.63
F vs. FINV * Âge * Formalité	-0.20	0.37	0.63	-2.57	2.17	-1.35	0.94	-1.15	0.76	-1.04	0.63
IS vs. FINV * Âge * Formalité	0.68	0.81	0.19	-2.76	3.85	-0.83	2.21	-0.58	1.96	-0.41	1.80

TABLEAU H.4 – Distributions postérieures pour le modèle brms d'analyse de AJT3 - Français adapté

Annexe I

AJT4 : Matériels et modèles statistiques

I.0.1 Écran de consentement (AJT4) (5.2.4)



Bonjour et bienvenue !

Titre de l'étude: Expérience de lecture et de jugement

Responsables de l'étude : Gabriel Thiberge et Heather Burnett

Institution de rattachement : Laboratoire de Linguistique Formelle, CNRS, Université de Paris

Chercheurs recueillant les données: Gabriel Thiberge et Simon Duverger

Objectif du projet

Nous vous invitons à participer à une étude analysant le jugement de phrases du français par des locutrices et des locuteurs dont la langue maternelle est le français. Les données recueillies nous permettront de mieux comprendre le traitement des phrases en français.

Pourquoi suis-je invité(e) à participer ?

Nous vous avons demandé de participer parce que nous souhaitons mesurer des jugements faits par des participant(e)s humain(e)s pour analyser quantitativement le type de réponses qui sont données.

Suis-je obligé(e) de participer ?

Cette étude n'est pas obligatoire. La participation est uniquement basée sur le volontariat. Vous pouvez vous retirer de cette étude à n'importe quel instant sans en donner la raison. Vous recevrez à la fin de l'expérience un code aléatoire qui vous permettra de vous retirer à tout moment après l'expérience (dans le mois qui suit). Vous aurez seulement à nous envoyer ce code par mail à expgs2020@gmail.com.

Que va-t-il se passer si je participe ?

Nous vous demanderons de juger des phrases selon deux échelles, puis de renseigner quelques informations sur votre profil. Au total, l'expérience dure environ 25 minutes.

Existe-t-il des risques pour cette étude ?

Il n'y a aucun risque connu pour ce genre d'expérience.

Quels sont les avantages pour moi ?

Vous contribuez à l'avancée de nos connaissances sur le langage, et vous avez toute notre gratitude pour cela. Il n'y a pas de compensation financière pour la passation de cette étude.

Comment sont traitées les données ?

Vos données resteront anonymes. Nous n'établissons aucun lien entre vos données et votre identité propre. Toutes les informations que vous nous transmettez (âge, sexe, langue maternelle, etc) resteront confidentielles ainsi que vos résultats. Les fichiers bruts de résultats dont nous disposerons seront stockés sur des supports informatiques cryptés.

Est-ce que cette recherche sera publiée ?

Les résultats de l'étude seront publiés sous une forme agrégée dans des revues et présentés dans des conférences. Une version condensée et anonyme du fichier de résultats globaux sera déposée sur un site public à destination de la communauté scientifique dans une démarche éthique d'ouverture de la science. Dans tous les cas, vous ne serez jamais identifié(e) ni identifiable.

Coordonnées du chercheur en charge du projet en cas de demande d'informations ou de plaintes.

Heather Burnett : heather.susan.burnett@gmail.com

Gabriel Thiberge : expgs2020@gmail.com

Vous pouvez envoyer un mail à l'adresse expgs2020@gmail.com pour avoir un résumé des résultats ou mieux comprendre les raisons de l'étude.

En participant à cette expérience, je consens aux points suivants :

1. Je suis d'accord pour participer à cette étude.
2. Je confirme que j'ai lu et compris comment mes données seront gardées et traitées.
3. J'ai plus de 18 ans.
4. Ma langue maternelle est le français.

Je comprends et j'accepte ce qui a été mentionné ci-dessus.

→ [Cliquez ici pour commencer la première partie](#)

FIGURE I.1 – Écran de consentement (AJ4)

I.0.2 Consignes (AJT4) (5.2.4)



Partie I: Expérience de lecture et de jugement

Merci de participer à cette expérience qui durera environ 25 minutes au total !
Dans cette première partie, vous allez lire attentivement de très courts dialogues et vous devrez juger si la dernière personne qui s'exprime (ce sera toujours la ligne commençant par "B") :

1. parle un bon français
(est-ce que pour vous la phrase est bien formée)
2. parle un français adapté à la situation
(est-ce que vous pourriez produire cette phrase dans la situation)

Vous devrez rendre vos jugements sur deux échelles séparées en faisant glisser un curseur entre "Pas du tout d'accord" et "Totalelement d'accord".

Ne passez pas trop de temps sur chaque phrase : il n'y a pas de mauvaise réponse et ce sont vos instincts sur la langue qui nous intéressent.

Cette première partie dure environ 20 minutes.

Il n'est pas possible de revenir en arrière au cours de l'expérience. Si vous cliquez sur la flèche "précédent" de votre navigateur, l'expérience ne fonctionnera plus.

Dans la dernière partie, nous vous demandons de renseigner des informations personnelles. Ces informations resteront strictement confidentielles, et ne seront en aucun cas transmises à un tiers. Cela dure environ 5 minutes.

→ [Cliquez ici pour commencer l'entraînement](#)

FIGURE I.2 – Consigne (AJT4)

I.0.3 Distracteurs (AJT4) (5.2.4)

	Item
-	
ajt4-031	A : J'ai vu mes parents hier... B : Ta mère pas va mieux?
ajt4-032	A : Y a un scooter qui m'a coupé la route tout à l'heure. B : Tu l'avais quand même pas renversé?
ajt4-033	A : Plus jamais je ne skierai avec elle. B : Elle va trop trop vite, hein?
ajt4-034	A : Tu voudrais pas venir m'aider? B : Non, clairement.
ajt4-035	A : À la fin, il est parti. B : Pourquoi donc pas?
ajt4-036	A : Le général de brigade nous a dit d'aller demander à son aide. B : Tu sais c'est qui son aide?
ajt4-037	A : Elles nous ont dit de venir demain soir. B : D'accord, on ira hier alors?
ajt4-038	A : Une éclipse, c'est quand le soleil cache la lune. B : Tu ne crois que c'est l'inverse?
ajt4-039	A : Les chats ne font pas des chiens... B : Ils font quoi?
ajt4-040	A : Les paris sont ouverts! B : Où ça?

TABLEAU I.1 – Distracteurs pour l'étude AJT4 (031-040)

	Item
-	
ajt4-051a	A : Le carton va tomber de tes bras. Je vais t'aider. B : Oui, prends-le-moi, ça serait gentil.
ajt4-051b	A : Le carton va tomber de tes bras. Je vais t'aider. B : Oui, prends-moi-le, ça serait gentil.
ajt4-052a	A : Il faut que je te donne le code pour entrer chez moi. B : Dis-le-moi, je vais arriver.
ajt4-052b	A : Il faut que je te donne le code pour entrer chez moi. B : Dis-moi-le, je vais arriver.
ajt4-053a	A : J'ai oublié de te donner le numéro de la nourrice. B : Donne-le-moi et je l'appelle.
ajt4-053b	A : J'ai oublié de te donner le numéro de la nourrice. B : Donne-moi-le et je l'appelle.
ajt4-054a	A : J'ai fini le dossier que vous m'avez demandé pour demain. B : Laisse-le-nous ce soir, avant de partir.
ajt4-054b	A : J'ai fini le dossier que vous m'avez demandé pour demain. B : Laisse-nous-le ce soir, avant de partir.
ajt4-055a	A : L'appel pour la visioconférence est en ligne. B : Passe-le-nous dans 5 minutes, s'il-te-plaît.
ajt4-055b	A : L'appel pour la visioconférence est en ligne. B : Passe-nous-le dans 5 minutes, s'il te plaît.
ajt4-056a	A : J'ai apporté mon diaporama sur une clef USB. B : Mets-le-nous sur l'écran, si ça ne te dérange pas.
ajt4-056b	A : J'ai apporté mon diaporama sur une clef USB. B : Mets-nous-le sur l'écran, si ça ne te dérange pas.
ajt4-057a	A : Tu as oublié de prendre le vin pour le pot de départ. B : Prends-le-moi, je te rembourserai.
ajt4-057b	A : Tu as oublié de prendre le vin pour le pot de départ. B : Prends-moi-le, je te rembourserai.
ajt4-058a	A : Ton ordinateur est resté allumé au bureau. B : Arrête-le-moi quand tu partiras.
ajt4-058b	A : Ton ordinateur est resté allumé au bureau. B : Arrête-moi-le quand tu partiras.
ajt4-059a	A : Je peux finir le plan pour mercredi ou jeudi. B : Fais-le-moi pour la fin de semaine.
ajt4-059b	A : Je peux finir le plan pour mercredi ou jeudi. B : Fais-moi-le pour la fin de semaine.
ajt4-060a	A : Un numéro inconnu cherche à vous joindre depuis trois jours. B : Appelle-le-nous, on va régler ça.
ajt4-060b	A : Un numéro inconnu cherche à vous joindre depuis trois jours. B : Appelle-nous-le, on va régler ça.
ajt4-061a	A : J'ai perdu le code d'activation, je vais le chercher. B : Trouve-le-nous avant qu'il ne soit trop tard.
ajt4-061b	A : J'ai perdu le code d'activation, je vais le chercher. B : Trouve-nous-le avant qu'il ne soit trop tard.
ajt4-062a	A : Si vous avez encore faim, j'ai fait un gâteau. B : Mets-le-nous de côté pour plus tard.
ajt4-062b	A : Si vous avez encore faim, j'ai fait un gâteau. B : Mets-nous-le de côté pour plus tard.

TABLEAU I.2 – Distracteurs pour l'étude AJT4 (051-062)

	Item
-	
ajt4-063a	A : Je n'arrive pas à te faire comprendre ce que je veux te dire. B : Mime-le-moi, ça sera plus simple.
ajt4-063b	A : Je n'arrive pas à te faire comprendre ce que je veux te dire. B : Mime-moi-le, ça sera plus simple.
ajt4-064a	A : Je suis sûr que toi aussi tu connais encore l'air de ce slogan. B : Siffle-le-moi pour que je m'en souviene.
ajt4-064b	A : Je suis sûr que toi aussi tu connais encore l'air de ce slogan. B : Siffle-moi-le pour que je m'en souviene.
ajt4-065a	A : Le montant est trop élevé pour que je te fasse un seul virement. B : Reverse-le-moi en trois fois, ça devrait aller.
ajt4-065b	A : Le montant est trop élevé pour que je te fasse un seul virement. B : Reverse-moi-le en trois fois, ça devrait aller.
ajt4-066a	A : J'ai du mal à apprendre ce passage, il faut que je m'entraîne. B : Déclame-le-nous en articulant bien.
ajt4-066b	A : J'ai du mal à apprendre ce passage, il faut que je m'entraîne. B : Déclame-nous-le en articulant bien.
ajt4-067a	A : J'ai vu que vous n'avez plus ce médicament en réserve. B : Achemine-le-nous ou nous ne pourrons plus soigner les malades.
ajt4-067b	A : J'ai vu que vous n'avez plus ce médicament en réserve. B : Achemine-nous-le ou nous ne pourrons plus soigner les malades.
ajt4-068a	A : Le colis que je voulais vous envoyer m'a été retourné. B : Réexpédie-le-nous à notre nouvelle adresse.
ajt4-068b	A : Le colis que je voulais vous envoyer m'a été retourné. B : Réexpédie-nous-le à notre nouvelle adresse.
ajt4-069a	A : Je vais réécrire en partie le discours, je le trouve un peu lisse. B : Pimente-le-moi un peu pour voir ce que ça donne.
ajt4-069b	A : Je vais réécrire en partie le discours, je le trouve un peu lisse. B : Pimente-moi-le un peu pour voir ce que ça donne.
ajt4-070a	A : Le dossier est trop gros pour que tu l'envoies par la Poste. B : Compresse-le-moi et je l'enverrai par e-mail.
ajt4-070b	A : Le dossier est trop gros pour que tu l'envoies par la Poste. B : Compresse-moi-le et je l'enverrai par e-mail.
ajt4-071a	A : Maman, mon bouillon est trop froid maintenant. B : Reverse-le-moi dans la casserole, je vais le réchauffer.
ajt4-071b	A : Maman, mon bouillon est trop froid maintenant. B : Reverse-moi-le dans la casserole, je vais le réchauffer.
ajt4-072a	A : Je ne vois pas comment améliorer le graphique. B : Modernise-le-nous pour la prochaine réunion.
ajt4-072b	A : Je ne vois pas comment améliorer le graphique. B : Modernise-nous-le pour la prochaine réunion.
ajt4-073a	A : Je ne peux pas conduire le train à la destination prévue. B : Achemine-le-nous vers la gare Montparnasse.
ajt4-073b	A : Je ne peux pas conduire le train à la destination prévue. B : Achemine-nous-le vers la gare Montparnasse.
ajt4-074a	A : Je ne peux pas faire grand-chose, ce chemin est trop dangereux pour être utilisé. B : Balise-le-nous pour qu'on puisse recommencer les visites.
ajt4-074b	A : Je ne peux pas faire grand-chose, ce chemin est trop dangereux pour être utilisé. B : Balise-nous-le pour qu'on puisse recommencer les visites.

TABLEAU I.3 – Distracteurs pour l'étude AJT4 (063-074)

I.0.4 Modèle *brms* et distributions postérieures pour la question "La personne (...) qui s'exprime parle un *bon français*" (5.2.4)

Family : cumulative
Links : mu = logit; disc = identity
Formula : answerA ~ (fesk_finv + is_finv) * old_young * formal_informal + ((fesk_finv + is_finv) * formal_informal | id) + ((fesk_finv + is_finv) * old_young * formal_informal | Item)
Data : DON (Number of observations : 2730)
Samples : 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-4.19	0.26	-4.71	-3.70	1.00	1094	2106
Intercept[2]	-3.66	0.25	-4.17	-3.18	1.00	1056	1906
Intercept[3]	-3.01	0.25	-3.51	-2.53	1.00	1024	1899
Intercept[4]	-2.49	0.24	-2.98	-2.01	1.00	1005	1868
Intercept[5]	-2.10	0.24	-2.59	-1.63	1.00	993	2017
Intercept[6]	-1.44	0.24	-1.92	-0.97	1.00	988	2023
Intercept[7]	-1.06	0.24	-1.54	-0.60	1.00	985	2094
Intercept[8]	-0.61	0.24	-1.08	-0.14	1.00	991	1945
Intercept[9]	0.03	0.24	-0.44	0.49	1.00	986	1993
Intercept[10]	0.65	0.24	0.18	1.12	1.00	979	2125
fesk_finv	-2.39	0.35	-3.08	-1.72	1.00	2472	5083
is_finv	-3.41	0.35	-4.11	-2.72	1.00	2601	4691
old_young	-0.87	0.41	-1.68	-0.09	1.00	902	1836
formal_informal	0.10	0.12	-0.14	0.33	1.00	8204	8022
fesk_finv:old_young	-1.69	0.53	-2.75	-0.65	1.00	2788	4700
is_finv:old_young	-1.38	0.60	-2.57	-0.24	1.00	2417	4543
fesk_finv:formal_informal	-0.23	0.26	-0.74	0.28	1.00	8212	7696
is_finv:formal_informal	-0.16	0.26	-0.67	0.36	1.00	7875	8236
old_young:formal_informal	0.08	0.21	-0.33	0.49	1.00	8826	8503
fesk_finv:old_young:formal_informal	0.42	0.48	-0.50	1.36	1.00	8585	8143
is_finv:old_young:formal_informal	0.41	0.46	-0.48	1.30	1.00	8731	8015

TABLEAU I.4 – Résumé du modèle *brms* (AJT4, *bon français*)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Type : FESK vs. FINV	-2.39	0.00	1.00	-3.93	-1.09	-3.08	-1.72	-2.97	-1.82	-2.89	-1.89
Type : IS vs. FINV	-3.41	0.00	1.00	-4.85	-2.07	-4.11	-2.72	-4.01	-2.83	-3.92	-2.90
Âge : +30 ans vs. -30 ans	-0.87	0.01	0.99	-2.32	0.88	-1.68	-0.09	-1.54	-0.22	-1.45	-0.30
Formalité : Formel vs. Informel	0.10	0.80	0.20	-0.37	0.58	-0.14	0.33	-0.10	0.29	-0.08	0.26
FESK vs. FINV * Âge	-1.69	0.00	1.00	-3.84	0.42	-2.75	-0.65	-2.58	-0.82	-2.47	-0.94
IS vs. FINV * Âge	-1.38	0.01	0.99	-4.11	1.02	-2.57	-0.24	-2.38	-0.41	-2.24	-0.52
FESK vs. FINV * Formalité	-0.23	0.19	0.81	-1.33	0.75	-0.74	0.28	-0.66	0.19	-0.60	0.14
IS vs. FINV * Formalité	-0.16	0.27	0.73	-1.18	0.84	-0.67	0.36	-0.59	0.28	-0.54	0.22
Âge * Formalité	0.08	0.65	0.35	-0.74	0.87	-0.33	0.49	-0.27	0.43	-0.22	0.38
FESK vs. FINV * Âge * Formalité	0.42	0.81	0.19	-1.46	2.58	-0.50	1.36	-0.35	1.21	-0.26	1.11
IS vs. FINV * Âge * Formalité	0.41	0.82	0.18	-1.32	2.27	-0.48	1.30	-0.34	1.16	-0.25	1.07

TABLEAU I.5 – Distributions postérieures pour le modèle brms (AJT4, bon français)

I.0.5 Modèle *brms* et distributions postérieures pour la question "La personne (...) qui s'exprime parle un français adapté" (5.2.4)

Family : cumulative
Links : mu = logit; disc = identity
Formula : answerB ~ (fesk_finv + is_finv) * old_young * formal_informal + ((fesk_finv + is_finv) * formal_informal | id) + ((fesk_finv + is_finv) * old_young * formal_informal | Item)
Data : DON (Number of observations : 2730)
Samples : 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-3.43	0.23	-3.87	-2.98	1.00	1329	2151
Intercept[2]	-3.13	0.23	-3.57	-2.69	1.00	1300	2234
Intercept[3]	-2.70	0.22	-3.13	-2.26	1.00	1269	2318
Intercept[4]	-2.32	0.22	-2.75	-1.89	1.00	1244	2202
Intercept[5]	-2.03	0.22	-2.46	-1.61	1.00	1234	2284
Intercept[6]	-1.34	0.22	-1.76	-0.92	1.00	1222	2263
Intercept[7]	-1.05	0.22	-1.47	-0.63	1.00	1219	2101
Intercept[8]	-0.60	0.21	-1.02	-0.19	1.00	1213	2022
Intercept[9]	-0.04	0.21	-0.46	0.38	1.00	1223	2026
Intercept[10]	0.49	0.21	0.08	0.91	1.00	1215	2125
fesk_finv	-0.49	0.19	-0.86	-0.13	1.00	4413	6811
is_finv	-0.05	0.20	-0.44	0.34	1.00	3809	6431
old_young	-0.11	0.30	-0.70	0.50	1.00	1119	1957
formal_informal	0.16	0.11	-0.06	0.37	1.00	9203	9055
fesk_finv:old_young	-0.59	0.32	-1.21	0.02	1.00	4297	6841
is_finv:old_young	-0.10	0.37	-0.82	0.61	1.00	4412	5751
fesk_finv:formal_informal	-0.52	0.25	-1.02	-0.05	1.00	8510	8823
is_finv:formal_informal	-0.52	0.27	-1.05	0.02	1.00	8270	7660
old_young:formal_informal	-0.01	0.19	-0.39	0.36	1.00	8281	8649
fesk_finv:old_young:formal_informal	-0.22	0.44	-1.08	0.66	1.00	9422	8503
is_finv:old_young:formal_informal	-0.05	0.47	-0.96	0.88	1.00	9220	8522

TABLEAU I.6 – Résumé du modèle *brms* (AJT4, français adapté)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Type : FESK vs. FINV	-0.49	0.01	0.99	-1.35	0.27	-0.86	-0.13	-0.80	-0.19	-0.76	-0.22
Type : IS vs. FINV	-0.05	0.40	0.60	-0.87	0.71	-0.44	0.34	-0.38	0.28	-0.34	0.23
Âge : +30 ans vs. -30 ans	-0.11	0.35	0.65	-1.24	1.13	-0.70	0.50	-0.60	0.39	-0.53	0.32
Formalité : Formel vs. Informel	0.16	0.93	0.07	-0.25	0.57	-0.06	0.37	-0.02	0.33	0.00	0.31
FESK vs. FINV * Âge	-0.59	0.03	0.97	-1.83	0.64	-1.21	0.02	-1.11	-0.07	-1.04	-0.13
IS vs. FINV * Âge	-0.10	0.39	0.61	-1.78	1.31	-0.82	0.61	-0.70	0.49	-0.63	0.42
FESK vs. FINV * Formalité	-0.52	0.02	0.98	-1.77	0.59	-1.02	-0.05	-0.94	-0.12	-0.88	-0.17
IS vs. FINV * Formalité	-0.52	0.03	0.97	-1.59	0.44	-1.05	0.02	-0.96	-0.08	-0.90	-0.14
Âge * Formalité	-0.01	0.48	0.52	-0.66	0.88	-0.39	0.36	-0.33	0.30	-0.29	0.26
FESK vs. FINV * Âge * Formalité	-0.22	0.31	0.69	-2.14	2.14	-1.08	0.66	-0.95	0.51	-0.85	0.42
IS vs. FINV * Âge * Formalité	-0.05	0.46	0.54	-1.72	1.99	-0.96	0.88	-0.82	0.72	-0.73	0.63

TABLEAU I.7 – Distributions postérieures pour le modèle brms (AJT4, français adapté)

Annexe J

MG2 : Création et norming des échelles utilisées dans l'étude

Pour assurer la cohérence du design expérimental choisi et la fiabilité des résultats au regard des prédictions de départ, les deux niveaux de la variable CONTEXTE ont été vérifiés dans un norming.

J.0.1 Volontaires (MG2p, 5.3.2)

Nombre de volontaires : 42 francophones

Genre des volontaires : 28 femmes + 14 hommes

Âge des volontaires : 19-77 ans, $\mu=41.1$, $m=33.5$

Recrutement : RISC et réseaux sociaux

Compensation : Aucune

J.0.2 Protocole et design (MG2p, 5.3.2)

Tâche : Jugements d'indices sociaux sur des images de personnages.

Dates : Février 2018

Durée : 5 minutes

Modalité : Expérience en ligne, via IbexFarm [DRUMMOND, 2016]

Design : Pas vraiment de variable/condition. Chaque volontaire voyait deux images et devait se prononcer sur quelle image remplissait le mieux les propositions des échelles. Les échelles, elles, étaient différentes selon deux listes, voir Captures d'écran) pour adresser les stéréotypes visés dans deux sens différents (être *pauvre* vs. être *riche*; il y a eu des différences).

Ce n'est pas vraiment une variable, mais chaque personne ne voyait qu'une version genrée de chaque item. Il y a peu d'influence sur les réponses reçues.

Ce n'est pas vraiment une variable, mais les personnages changeaient de côté d'une liste à l'autre. Il n'y a aucune incidence sur les réponses reçues.

Plus de détails dans les fichiers bruts disponibles dans le répertoire osf.io/ud4ks.

Carré latin : /

Randomisation : Oui

Nombre d'items : 9 items

Échelle : Neuf échelles (1-7), voir Captures d'écran plus bas.

Consigne :

Dans cette expérience, il vous est demandé d'observer plusieurs paires d'images, puis de donner votre avis sur un certain nombre d'affirmations les concernant.

Pour chaque affirmation, il vous est demandé de choisir si elle se rapporte plutôt à l'image de gauche ou plutôt à l'image de droite.

Vous pouvez nuancer votre avis sur une échelle comprenant sept niveaux, selon que vous trouvez que l'affirmation concerne une seule des deux images, plutôt les deux, ou aucune.

L'ensemble devrait vous prendre à peine cinq minutes.

Merci de votre participation !

Les données collectées durant cette expérience le sont à des fins de recherche. Aucune donnée personnelle d'identification ne sera transmise à des tiers.

Votre âge :

Votre nationalité :

Votre genre : Homme Femme Autre
 Je ne veux pas répondre à cette question

Votre dernier diplôme obtenu : Avant le baccalauréat
 Baccalauréat ou équivalent
 Bac+2 ou équivalent
 Bac+3 ou équivalent
 Bac+5 ou équivalent
 Bac+8 ou équivalent

Votre domaine d'études ou d'activité professionnelle :

Votre région de naissance :

Votre langue maternelle :

Je suis majeur.e et j'accepte les termes et conditions de cette expérience.

→ [Cliquez ici pour continuer.](#)

FIGURE J.1 – Consigne (MG2p - Norming)

Captures d'écran :

Regardez attentivement les deux images suivantes :

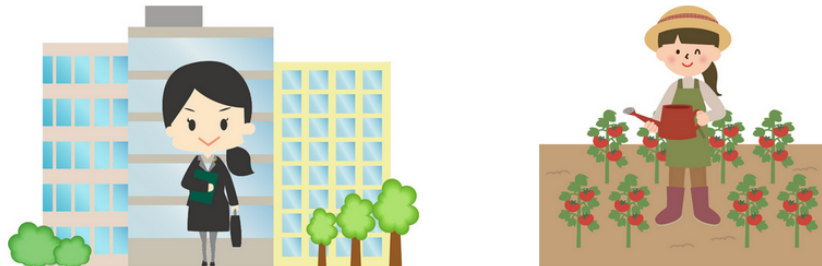


Maintenant, considérez les affirmations suivantes en indiquant à chaque fois à laquelle des deux images elle correspond le mieux :

La personne est plus jeune	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne est plus riche	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne est plus aimable	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne est plus solitaire	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne est plus citadine	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne est plus travailleuse	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne a fait plus d'études	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne est plus sportive	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne a plus de prestige	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.

– Cliquez ici pour continuer.

Regardez attentivement les deux images suivantes :



Maintenant, considérez les affirmations suivantes en indiquant à chaque fois à laquelle des deux images elle correspond le mieux :

La personne est plus âgée	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne est plus pauvre	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne est moins aimable	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne est plus sociable	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne est plus de la campagne	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne est plus paresseuse	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne a fait moins d'études	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne lit plus souvent	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.
La personne a moins de prestige	sur l'image de gauche	<input type="radio"/>	sur l'image de droite.

– Cliquez ici pour continuer.

FIGURE J.2 – Captures d'écran (MG2p - Norming)

J.0.3 Items et résultats (MG2p, 5.3.2)

Les réponses reçues sur des échelles à 7 crans ont été codées de -3 à 3 pour être centrées autour de 0.

Chaque couple d'images a été noté sur deux échelles différentes, par exemple "la personne est plus âgée" ou "la personne est plus jeune" selon la liste. Les stéréotypes ne sont pas toujours en opposition suivant le label. « A » réfère à des propositions issues de la première liste d'échelles, « B » à des propositions de la seconde liste (voir Captures d'écran).

Une note de -3 signifie que la proposition s'applique à la personne de gauche (homme ou femme), une note de +3 qu'elle s'applique à la personne de droite (homme ou femme).

Une note de +3 pour la proposition « A1_jeune+ » signifierait donc par exemple que la personne de droite est jugée sans doute plus jeune que la personne de gauche.

Un tableau des moyennes reçues par item sur chaque échelle précédera les graphiques des pages suivantes.

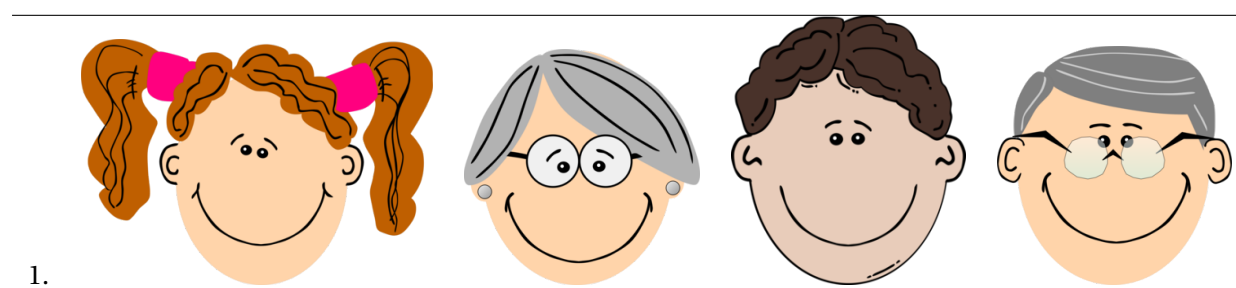


FIGURE J.3 – Item 1 (MG2p - Norming)

Proposition	Moyenne	Écart-type
A1_jeune+	-2.86	0.36
A2_riche+	1.33	1.46
A3_aimable+	0.05	0.59
A4_solitaire+	0.48	0.87
A5_citadin.e+	0.33	0.91
A6_travailleur.se+	0.19	1.08
A7_études+	1.24	1.04
A8_sports+	-0.90	1.18
A9_prestige+	0.67	1.02
B1_âgé.e+	2.81	0.51
B2_pauvre+	-0.14	0.79
B3_aimable-	-0.05	0.86
B4_sociable+	-0.24	1.04
B5_campagne+	0.05	0.80
B6_paresseux.se+	-0.24	0.70
B7_études-	-0.67	1.32
B8_lecture+	0.86	1.01
B9_prestige-	-0.24	0.94

TABEAU J.1 – Résultats du norming - Item 1 (MG2p)

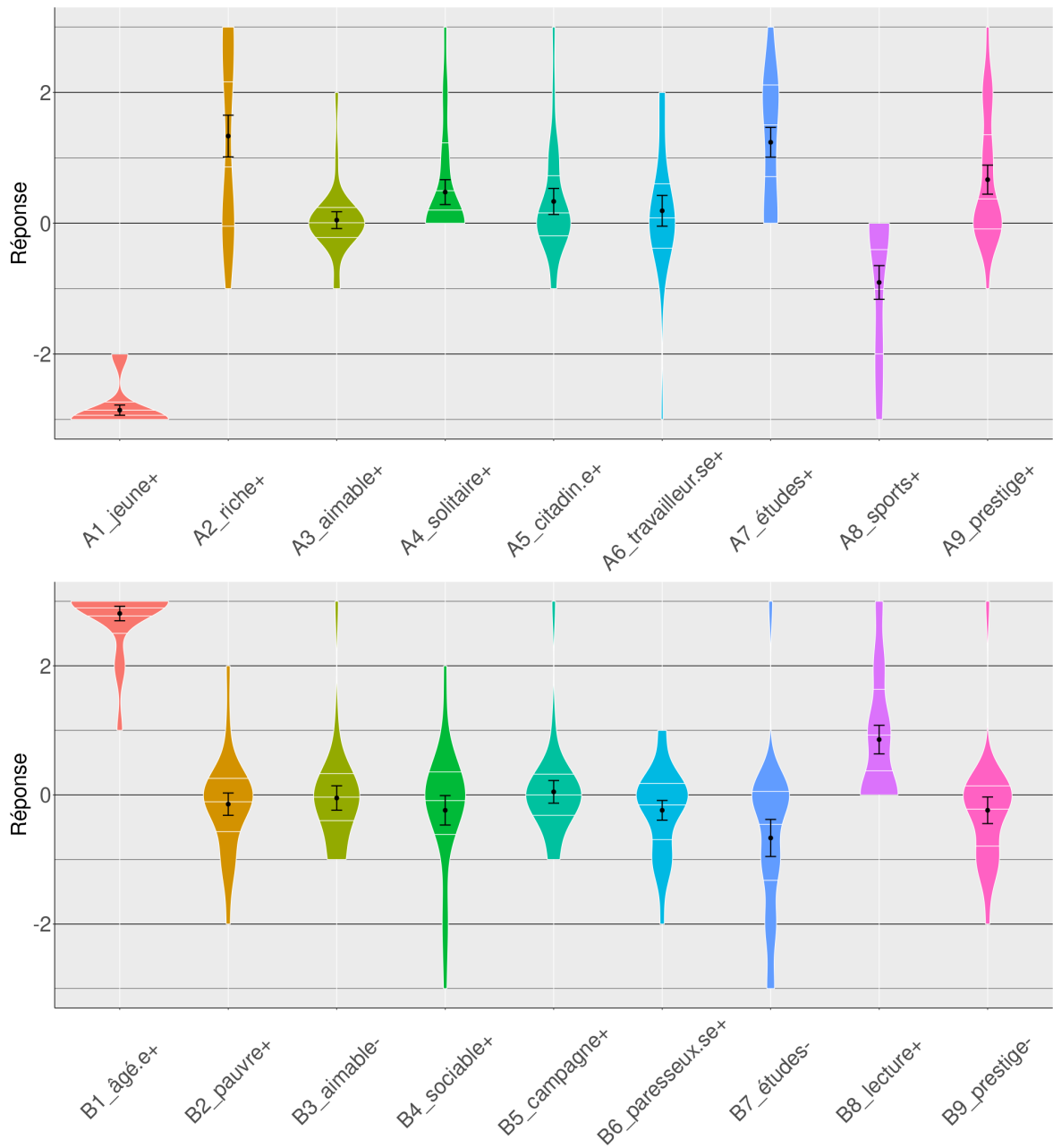


FIGURE J.4 – Résultats du norming - Item 1 (MG2p)

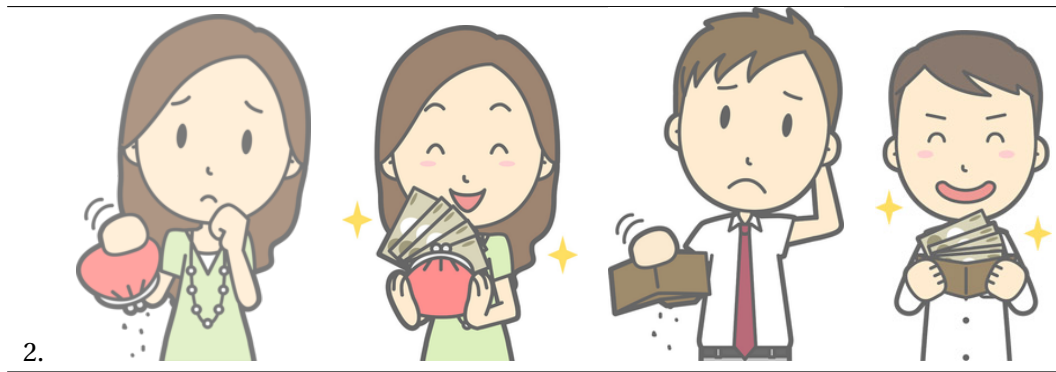


FIGURE J.5 – Item 2 (MG2p - Norming)

Proposition	Moyenne	Écart-type
A1_jeune+	-0.10	0.70
A2_riche+	2.14	1.65
A3_aimable+	0.29	0.90
A4_solitaire+	-0.14	0.36
A5_citadin.e+	0.00	0.55
A6_travailleur.se+	0.10	0.44
A7_études+	0.14	0.48
A8_sports+	0.10	0.30
A9_prestige+	0.24	0.83
B1_âgé.e+	0.10	0.77
B2_pauvre+	-1.62	1.53
B3_aimable-	0.24	1.14
B4_sociable+	0.19	0.87
B5_campagne+	-0.24	0.70
B6_paresseux.se+	-0.24	0.70
B7_études-	-0.29	0.85
B8_lecture+	0.05	0.80
B9_prestige-	-0.62	1.16

TABEAU J.2 – Résultats du norming - Item 2 (MG2p)

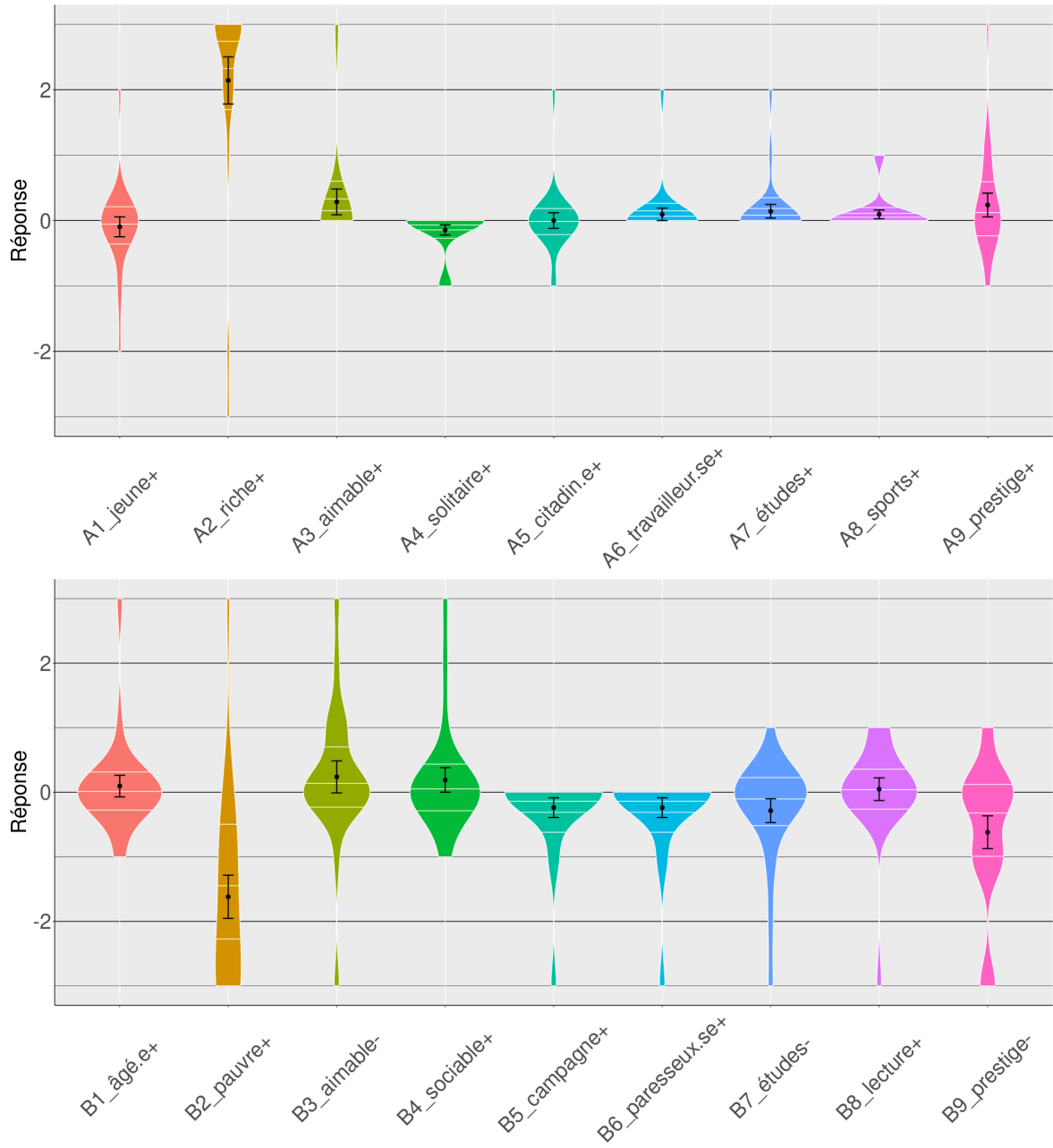


FIGURE J.6 – Résultats du norming - Item 2 (MG2p)

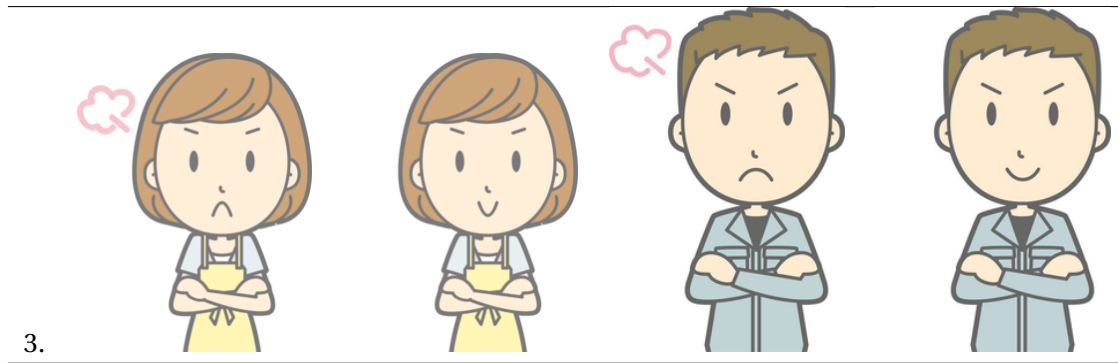


FIGURE J.7 – Item 3 (MG2p - Norming)

Proposition	Moyenne	Écart-type
A1_jeune+	0.05	0.38
A2_riche+	0.00	0.00
A3_aimable+	1.76	1.18
A4_solitaire+	-0.14	0.48
A5_citadin.e+	0.00	0.00
A6_travailleur.se+	0.24	0.54
A7_études+	0.00	0.00
A8_sports+	0.05	0.22
A9_prestige+	0.19	0.40
B1_âgé.e+	-0.10	0.30
B2_pauvre+	-0.14	0.48
B3_aimable-	-1.29	1.15
B4_sociable+	0.76	1.30
B5_campagne+	-0.14	0.48
B6_paresseux.se+	-0.05	0.38
B7_études-	-0.10	0.70
B8_lecture+	0.19	0.75
B9_prestige-	0.05	0.67

TABLEAU J.3 – Résultats du norming - Item 3 (MG2p)

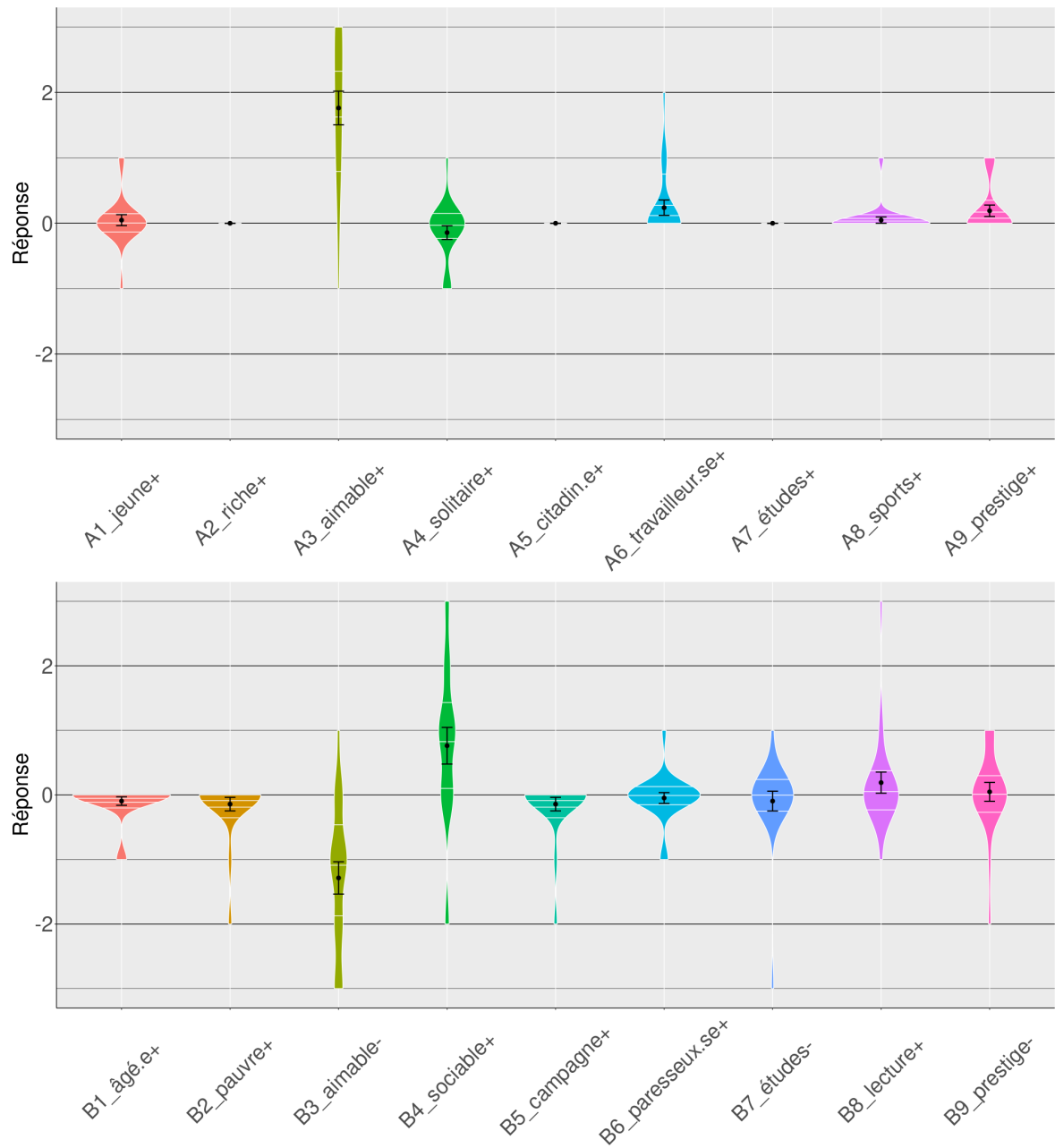


FIGURE J.8 – Résultats du norming - Item 3 (MG2p)

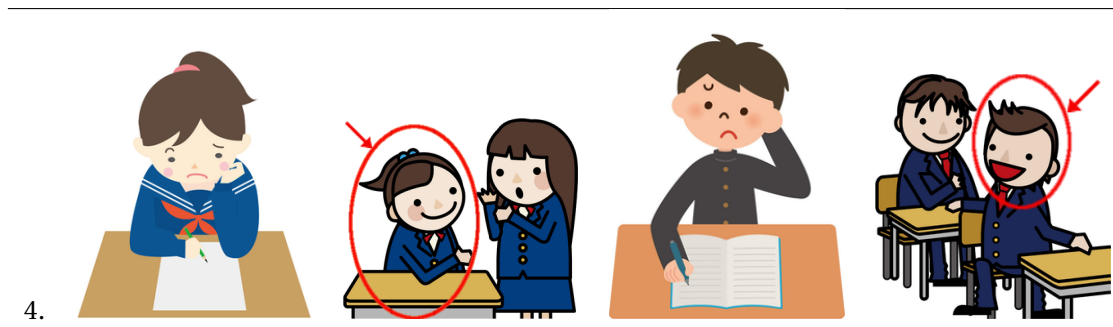


FIGURE J.9 – Item 4 (MG2p - Norming)

Proposition	Moyenne	Écart-type
A1_jeune+	0.24	1.18
A2_riche+	0.33	0.66
A3_aimable+	1.10	1.48
A4_solitaire+	-1.57	1.12
A5_citadin.e+	-0.10	0.30
A6_travailleur.se+	-1.24	1.45
A7_études+	-0.29	0.78
A8_sports+	-0.24	0.70
A9_prestige+	0.29	1.01
B1_âgé.e+	0.10	0.94
B2_pauvre+	-0.24	0.83
B3_aimable-	0.00	1.22
B4_sociable+	0.95	1.07
B5_campagne+	-0.33	0.58
B6_paresseux.se+	0.24	1.04
B7_études-	0.24	0.77
B8_lecture+	-0.48	0.81
B9_prestige-	-0.10	1.09

TABLEAU J.4 – Résultats du norming - Item 4 (MG2p)

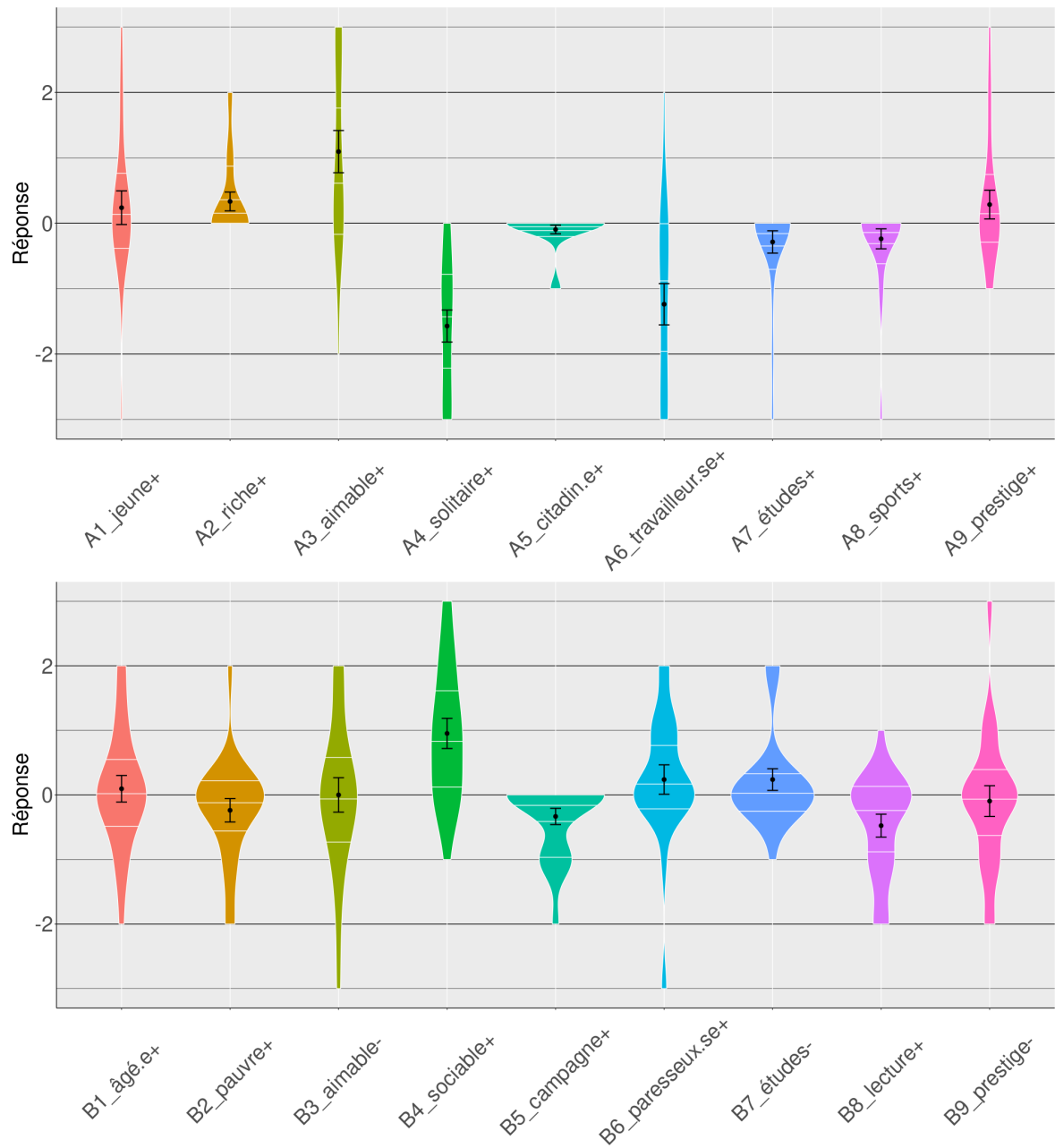


FIGURE J.10 – Résultats du norming - Item 4 (MG2p)

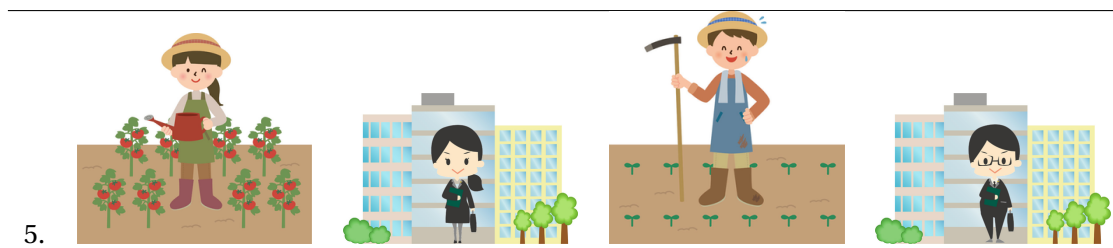


FIGURE J.11 – Item 5 (MG2p - Norming)

Proposition	Moyenne	Écart-type
A1_jeune+	-1.00	1.48
A2_riche+	1.71	1.19
A3_aimable+	-1.00	1.10
A4_solitaire+	0.29	1.19
A5_citadin.e+	2.76	0.54
A6_travailleur.se+	0.10	1.26
A7_études+	1.86	1.06
A8_sports+	-1.19	1.50
A9_prestige+	1.29	1.10
B1_âgé.e+	1.00	1.34
B2_pauvre+	-1.05	1.12
B3_aimable-	0.76	1.04
B4_sociable+	-0.71	1.15
B5_campagne+	-2.14	1.01
B6_paresseux.se+	0.33	0.66
B7_études-	-1.05	0.97
B8_lecture+	0.52	0.81
B9_prestige-	-0.67	0.80

TABLEAU J.5 – Résultats du norming - Item 5 (MG2p)

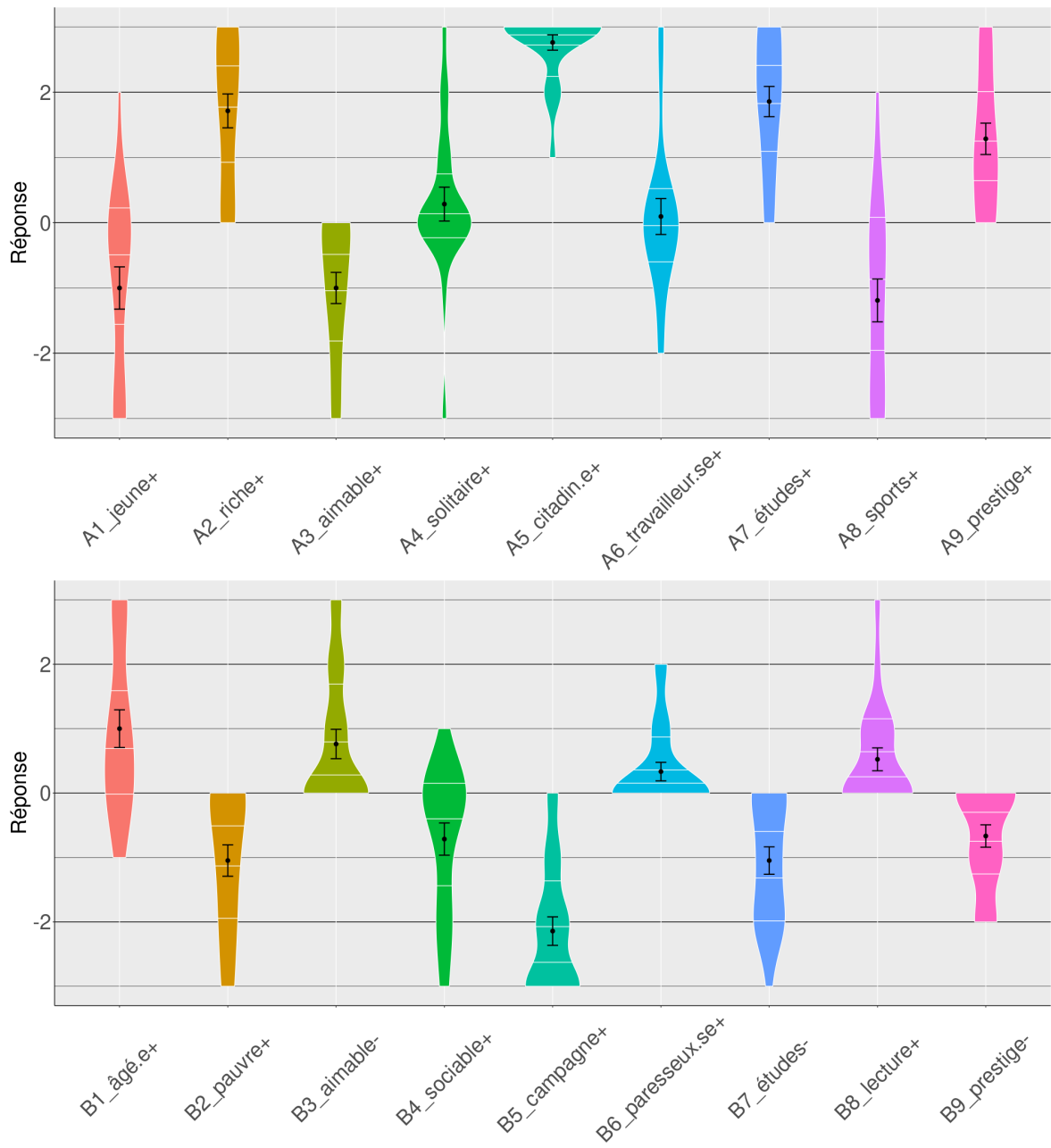


FIGURE J.12 – Résultats du norming - Item 5 (MG2p)

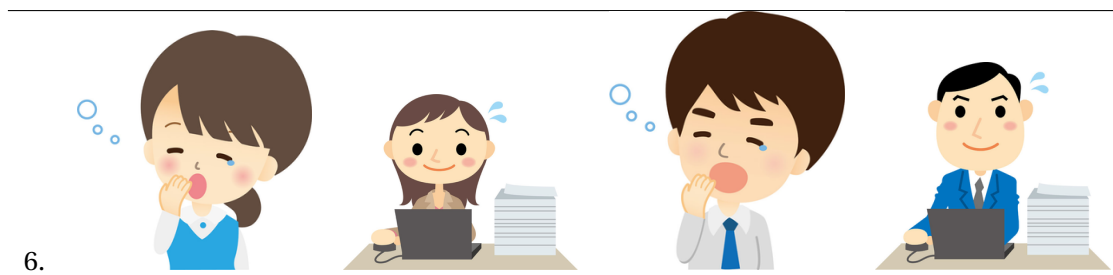


FIGURE J.13 – Item 6 (MG2p - Norming)

Proposition	Moyenne	Écart-type
A1_jeune+	-1.33	1.53
A2_riche+	0.86	1.31
A3_aimable+	0.38	0.92
A4_solitaire+	0.24	0.94
A5_citadin.e+	0.76	1.00
A6_travailleur.se+	1.33	1.11
A7_études+	0.33	1.20
A8_sports+	-0.05	1.07
A9_prestige+	0.71	1.01
B1_âgé.e+	1.00	1.30
B2_pauvre+	-0.29	0.78
B3_aimable-	0.19	1.03
B4_sociable+	0.19	0.60
B5_campagne+	0.00	0.45
B6_paresseux.se+	-0.43	1.03
B7_études-	-0.48	0.81
B8_lecture+	0.52	0.87
B9_prestige-	-0.48	0.81

TABLEAU J.6 – Résultats du norming - Item 6 (MG2p)

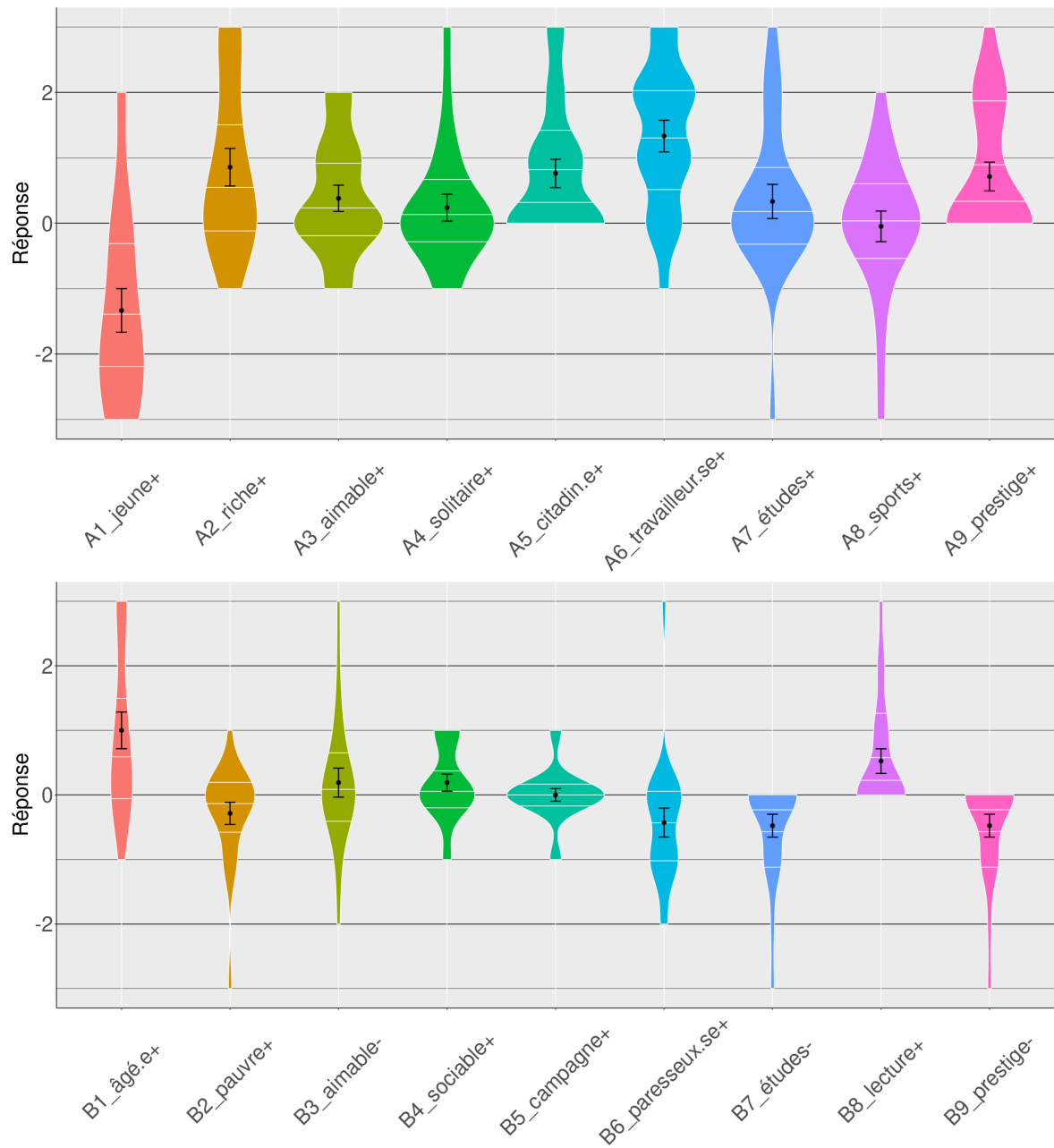


FIGURE J.14 – Résultats du norming - Item 6 (MG2p)

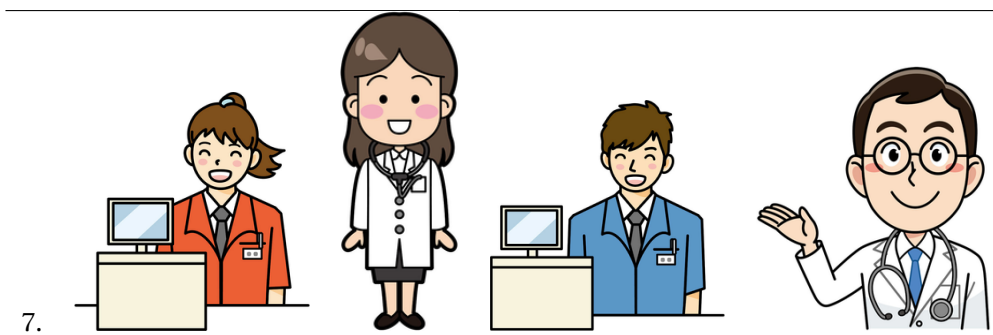


FIGURE J.15 – Item 7 (MG2p - Norming)

Proposition	Moyenne	Écart-type
A1_jeune+	-1.14	1.88
A2_riche+	2.14	1.15
A3_aimable+	0.10	0.94
A4_solitaire+	-0.10	0.83
A5_citadin.e+	0.05	0.50
A6_travailleur.se+	0.52	1.33
A7_études+	2.29	1.15
A8_sports+	-0.33	0.80
A9_prestige+	1.76	1.00
B1_âgé.e+	0.86	1.31
B2_pauvre+	-1.05	1.07
B3_aimable-	0.00	0.95
B4_sociable+	0.05	1.07
B5_campagne+	-0.29	0.85
B6_paresseux.se+	0.10	0.77
B7_études-	-1.00	1.10
B8_lecture+	0.05	1.07
B9_prestige-	-0.90	1.18

TABLEAU J.7 – Résultats du norming - Item 7 (MG2p)

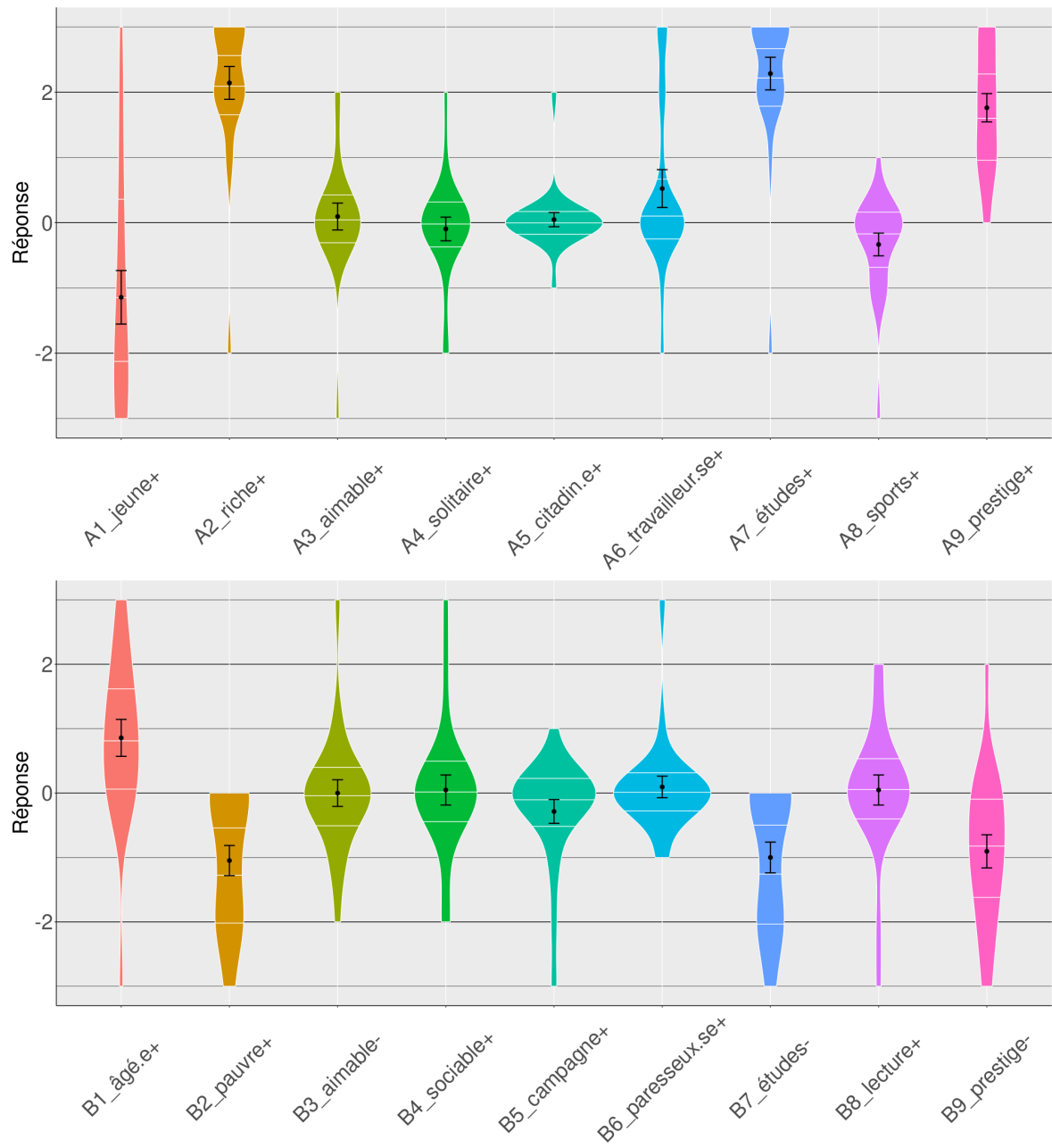


FIGURE J.16 – Résultats du norming - Item 7 (MG2p)



FIGURE J.17 – Item 8 (MG2p - Norming)

Proposition	Moyenne	Écart-type
A1_jeune+	-1.62	1.24
A2_riche+	0.10	1.04
A3_aimable+	0.14	0.96
A4_solitaire+	1.10	0.89
A5_citadin.e+	0.43	0.98
A6_travailleur.se+	0.95	1.02
A7_études+	1.10	1.22
A8_sports+	-2.00	1.00
A9_prestige+	0.24	1.04
B1_âgé.e+	0.86	1.28
B2_pauvre+	-0.05	0.97
B3_aimable-	0.24	0.70
B4_sociable+	-0.71	0.85
B5_campagne+	0.29	0.78
B6_paresseux.se+	0.19	0.68
B7_études-	-0.67	0.80
B8_lecture+	1.67	1.11
B9_prestige-	0.05	0.59

TABLEAU J.8 – Résultats du norming - Item 8 (MG2p)

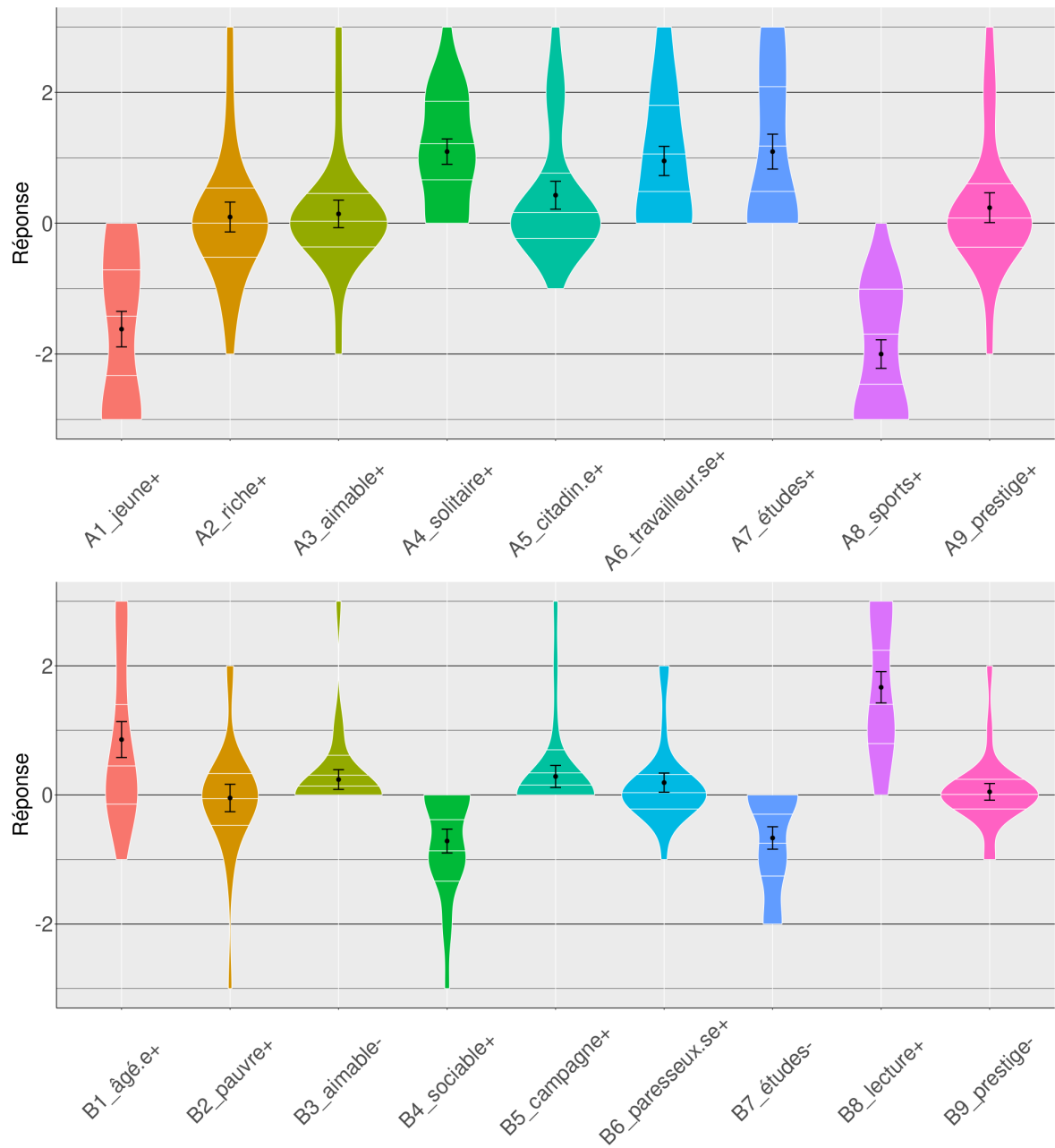


FIGURE J.18 – Résultats du norming - Item 8 (MG2p)

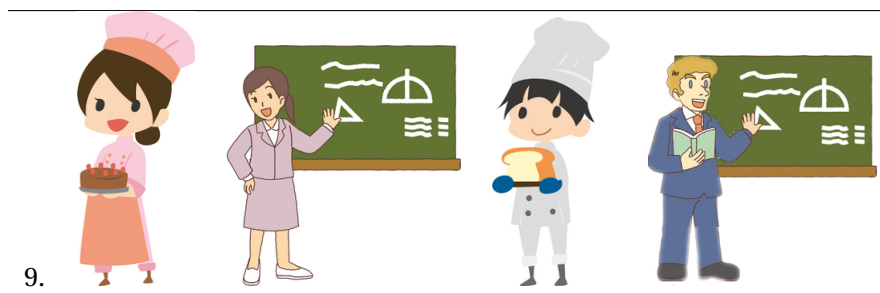


FIGURE J.19 – Item 9 (MG2p - Norming)

Proposition	Moyenne	Écart-type
A1_jeune+	-1.43	1.47
A2_riche+	0.67	1.24
A3_aimable+	-0.43	0.81
A4_solitaire+	0.33	0.80
A5_citadin.e+	0.57	0.87
A6_travailleur.se+	0.19	0.93
A7_études+	1.57	1.08
A8_sports+	0.14	0.96
A9_prestige+	0.57	1.03
B1_âgé.e+	1.67	1.35
B2_pauvre+	-0.24	1.14
B3_aimable-	0.33	0.73
B4_sociable+	-0.43	0.87
B5_campagne+	-0.24	0.94
B6_paresseux.se+	0.10	0.77
B7_études-	-0.86	1.01
B8_lecture+	0.62	1.16
B9_prestige-	-0.52	1.17

TABLEAU J.9 – Résultats du norming - Item 9 (MG2p)

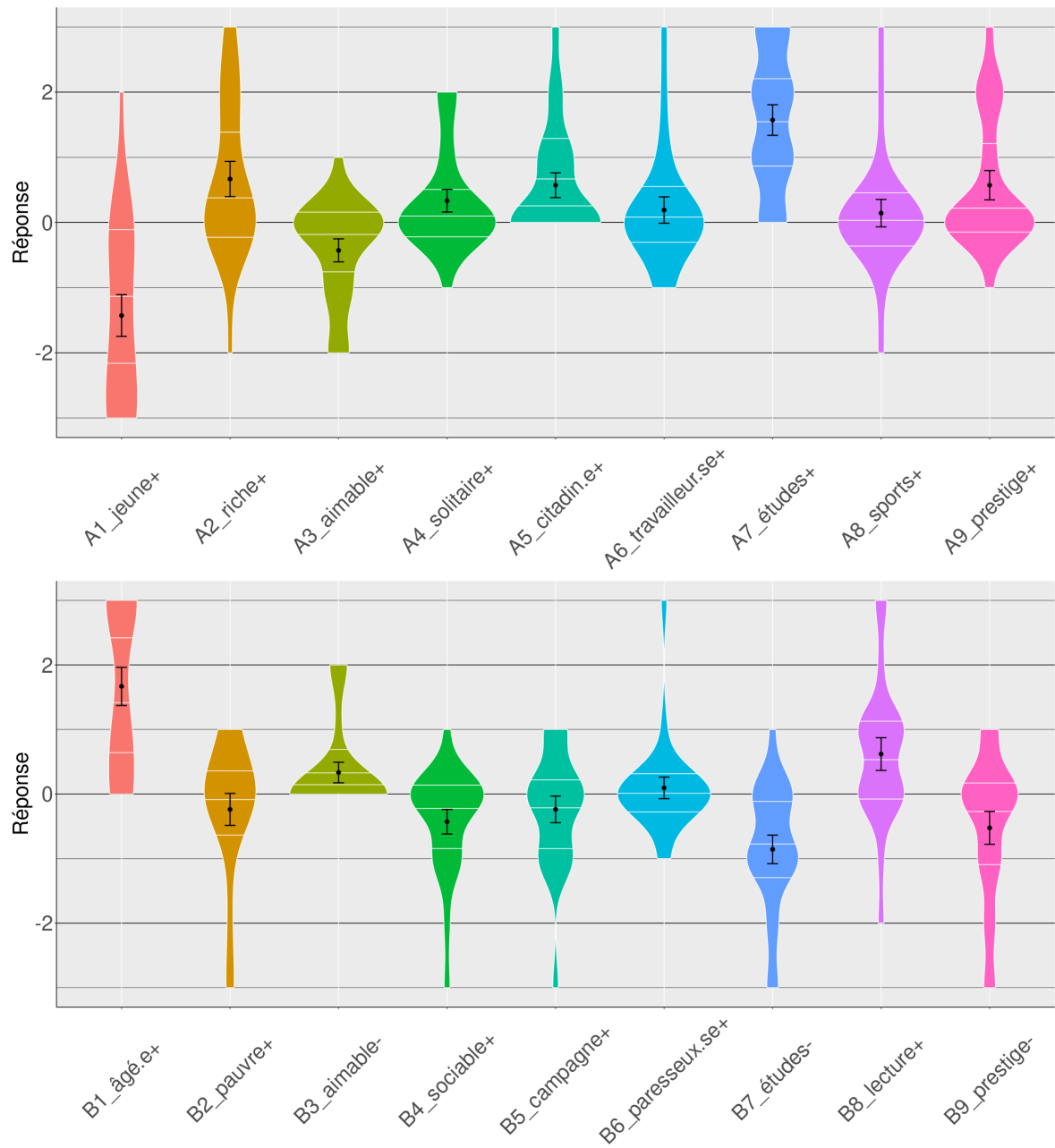


FIGURE J.20 – Résultats du norming - Item 9 (MG2p)

Annexe K

MG2 : Matériel et modèles statistiques

K.0.1 Consignes (MG2) (5.3.2)

Dans cette expérience, il vous est demandé d'écouter de brefs extraits audio.

Vérifiez que le volume sonore est activé et bien réglé sur votre appareil avant de commencer l'expérience. Vous pourrez réécouter chaque extrait autant de fois que vous le voudrez.

Après votre écoute, il vous est demandé d'associer un certain nombre de caractéristiques sociales à la seconde personne qui s'exprime. Ces caractéristiques sociales sont symbolisées par six échelles dont les extrémités sont à chaque fois deux images, par exemple :



Il n'y a pas de mauvaises réponses, en conséquence ne tentez pas de revenir en arrière. Prenez votre temps pour répondre.

Trois phrases d'entraînement vous seront proposées, avant de passer à l'expérience proprement dite.

L'ensemble devrait vous prendre une trentaine de minutes.
Merci de votre participation !

Les données collectées durant cette expérience le sont à des fins de recherche. Aucune donnée personnelle d'identification ne sera transmise à des tiers.

Votre âge :

Votre nationalité :

Votre genre : Homme Femme Autre
 Je ne veux pas répondre à cette question

Votre dernier diplôme obtenu : Avant le baccalauréat
 Baccalauréat ou équivalent
 Bac+2 ou équivalent
 Bac+3 ou équivalent
 Bac+5 ou équivalent
 Bac+8 ou équivalent

Votre domaine d'études ou d'activité professionnelle :

Votre région de naissance :

Votre langue maternelle :

Je suis majeur.e et j'accepte les termes et conditions de cette expérience.

[→ Cliquez ici pour continuer.](#)

FIGURE K.1 – Consigne (MG2)

K.0.2 Diaporama de la phase de familiarisation (5.3.2)

VOICI LÉO



VOICI LÉO



VOICI LOUIS



K.0.3 Diaporama de la phase de familiarisation (5.3.2)

VOICI LÉO

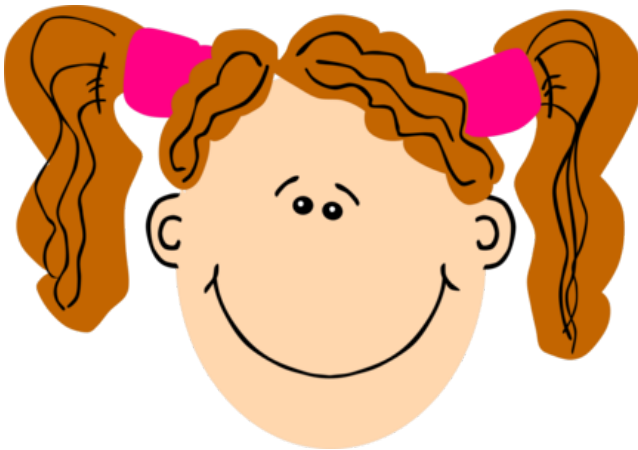


VOICI LOUIS







QUI EST LE PLUS JEUNE ?

VOICI NINA



K.0.4 Diaporama de la phase de familiarisation (5.3.2)

<p>VOICI <u>NINA</u></p> 	<p>VOICI <u>MARIE</u></p> 
<p>VOICI <u>NINA</u></p> 	<p>VOICI <u>MARIE</u></p> 

QUI EST LA PLUS JEUNE ?

K.0.5 Diaporama de la phase de familiarisation (5.3.2)

VOICI CARL



VOICI CARL



VOICI JEAN



K.0.6 Diaporama de la phase de familiarisation (5.3.2)

VOICI CARL



VOICI JEAN



QUI EST LE PLUS RICHE ?

VOICI JADE



K.0.7 Diaporama de la phase de familiarisation (5.3.2)

VOICI JADE



VOICI ELSA



VOICI JADE



VOICI ELSA



QUI EST LA PLUS RICHE ?

K.0.8 Diaporama de la phase de familiarisation (5.3.2)

VOICI DAVID



VOICI DAVID



VOICI THOMAS



K.0.9 Diaporama de la phase de familiarisation (5.3.2)

VOICI DAVID



VOICI THOMAS



QUI EST LE PLUS SOCIABLE ?

VOICI NADIA

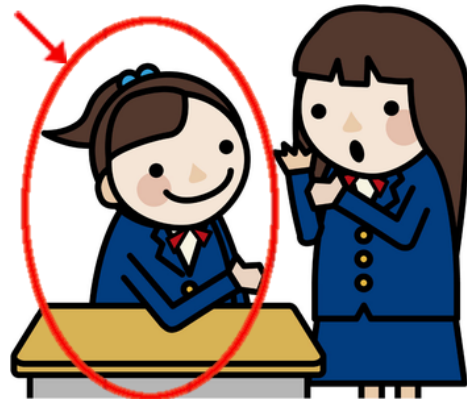


K.0.10 Diaporama de la phase de familiarisation (5.3.2)

VOICI NADIA



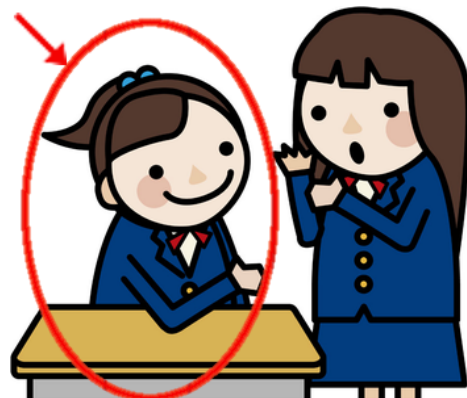
VOICI CHLOÉ



VOICI NADIA



VOICI CHLOÉ



QUI EST LA PLUS SOCIABLE ?

K.0.11 Diaporama de la phase de familiarisation (5.3.2)

VOICI JEAN



VOICI JEAN



VOICI SAMUEL



K.0.12 Diaporama de la phase de familiarisation (5.3.2)

VOICI JEAN



VOICI SAMUEL

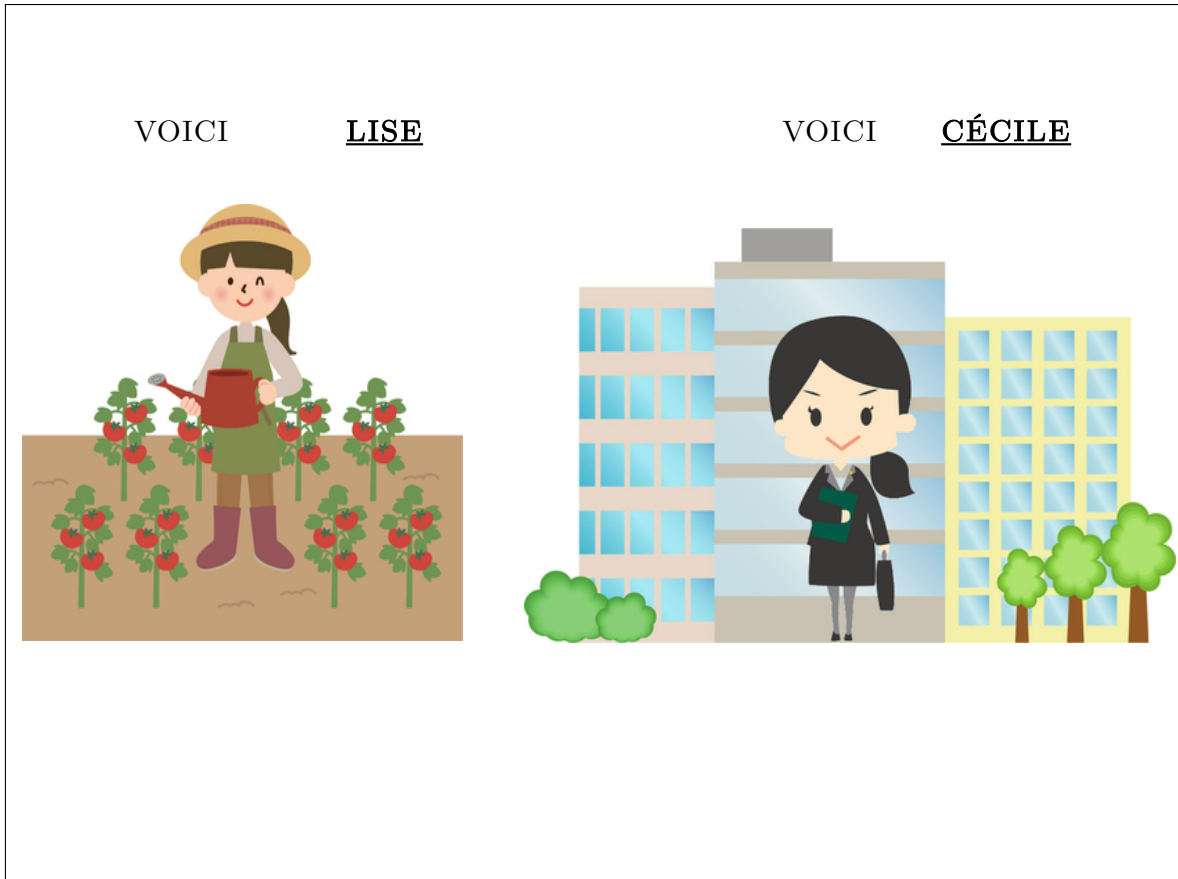


QUI VIENT DE LA CAMPAGNE ?

VOICI LISE



K.0.13 Diaporama de la phase de familiarisation (5.3.2)



K.0.14 Diaporama de la phase de familiarisation (5.3.2)

VOICI PAUL



VOICI PAUL



VOICI JACQUES



K.0.15 Diaporama de la phase de familiarisation (5.3.2)

VOICI PAUL



VOICI JACQUES



QUI A FAIT LES PLUS LONGUES ÉTUDES ?

VOICI SARAH



K.0.16 Diaporama de la phase de familiarisation (5.3.2)

VOICI SARAH



VOICI ANNE



VOICI SARAH



VOICI ANNE



QUI A FAIT LES PLUS LONGUES ÉTUDES ?

K.0.17 Diaporama de la phase de familiarisation (5.3.2)

VOICI LUCAS



VOICI LUCAS



VOICI MARC



K.0.18 Diaporama de la phase de familiarisation (5.3.2)

VOICI

LUCAS



VOICI

MARC



QUI PASSE LE PLUS DE TEMPS À LIRE ?

VOICI

CLAIRE



K.0.19 Diaporama de la phase de familiarisation (5.3.2)

VOICI

CLAIRE



VOICI

ANGÈLE



VOICI

CLAIRE



VOICI

ANGÈLE



QUI PASSE LE PLUS DE TEMPS À LIRE ?

K.0.20 Corrélations entre échelles - Détails

	Âge	Richesse	Sociabilité	Origine	Études	Loisirs
Âge	1.00	0.42	-0.01	0.11	0.32	0.29
Richesse	0.42	1.00	0.05	0.26	0.46	0.40
Sociabilité	-0.01	0.05	1.00	0.16	0.03	-0.05
Origine	0.11	0.26	0.16	1.00	0.41	0.34
Études	0.32	0.46	0.03	0.41	1.00	0.57
Loisirs	0.29	0.40	-0.05	0.34	0.57	1.00

TABLEAU K.1 – Coefficients de corrélations entre échelles, méthode Spearman (MG2)

	Âge	Richesse	Sociabilité	Origine	Études	Loisirs
Âge		0.00	0.83	0.00	0.00	0.00
Richesse	0.00		0.05	0.00	0.00	0.00
Sociabilité	0.83	0.05		0.00	0.17	0.06
Origine	0.00	0.00	0.00		0.00	0.00
Études	0.00	0.00	0.17	0.00		0.00
Loisirs	0.00	0.00	0.06	0.00	0.00	

TABLEAU K.2 – P-values pour les corrélations entre échelles, méthode Spearman (MG2)

K.0.21 Résultats moyens combinés entre toutes les échelles, par participant

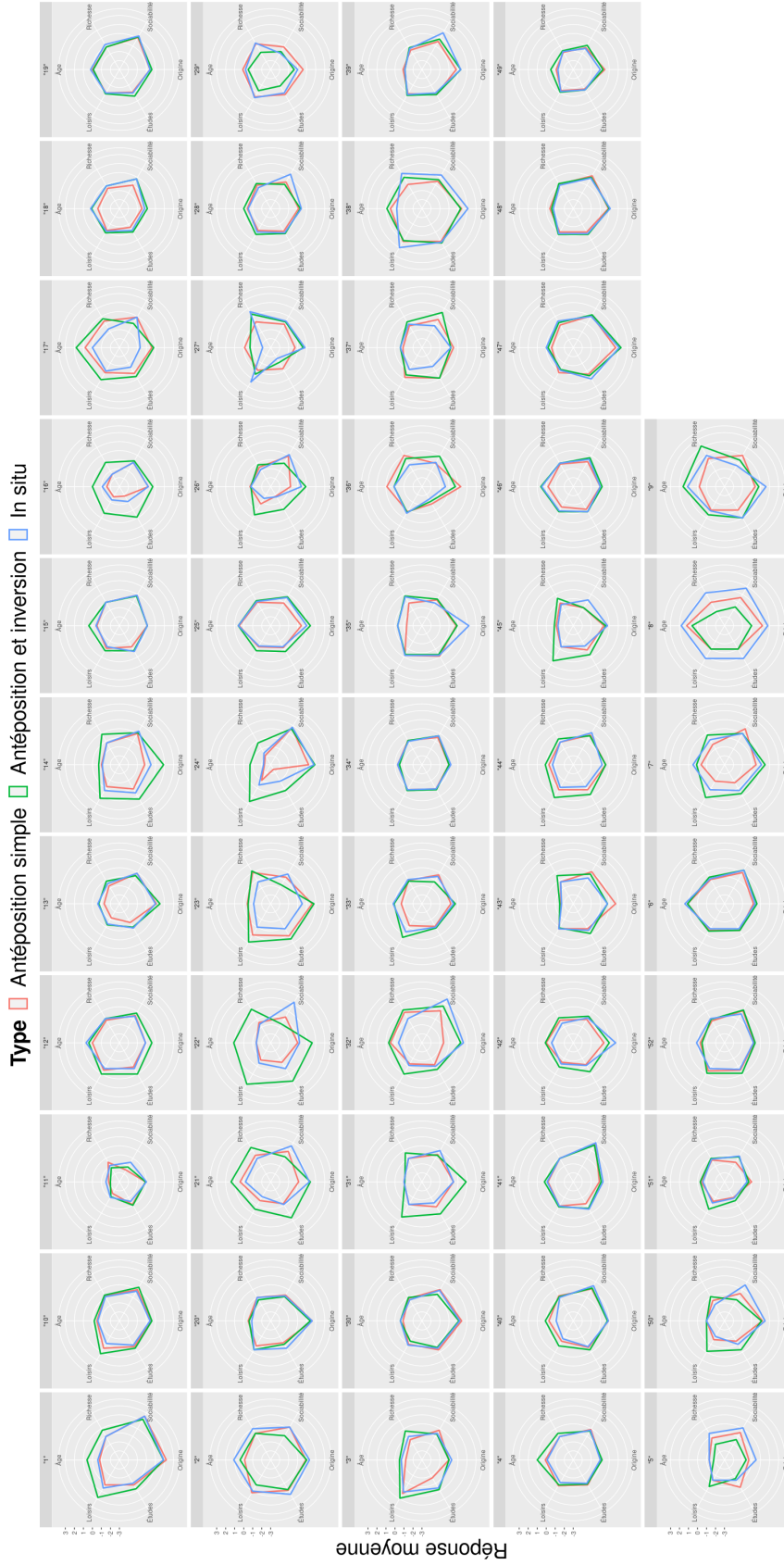


FIGURE K.2 – Résultats moyens combinés entre toutes les échelles, par participant (MG2)

K.0.22 Résultats détaillés pour l'influence de la variable GROUPE d'ÂGE (+/- 30 ans)

NB : La différence de taille entre les barres d'erreur sur les graphiques s'explique par la différence de nombre dans les groupes d'âge autour de 30 ans (volontaires de +30 ans, N=10).

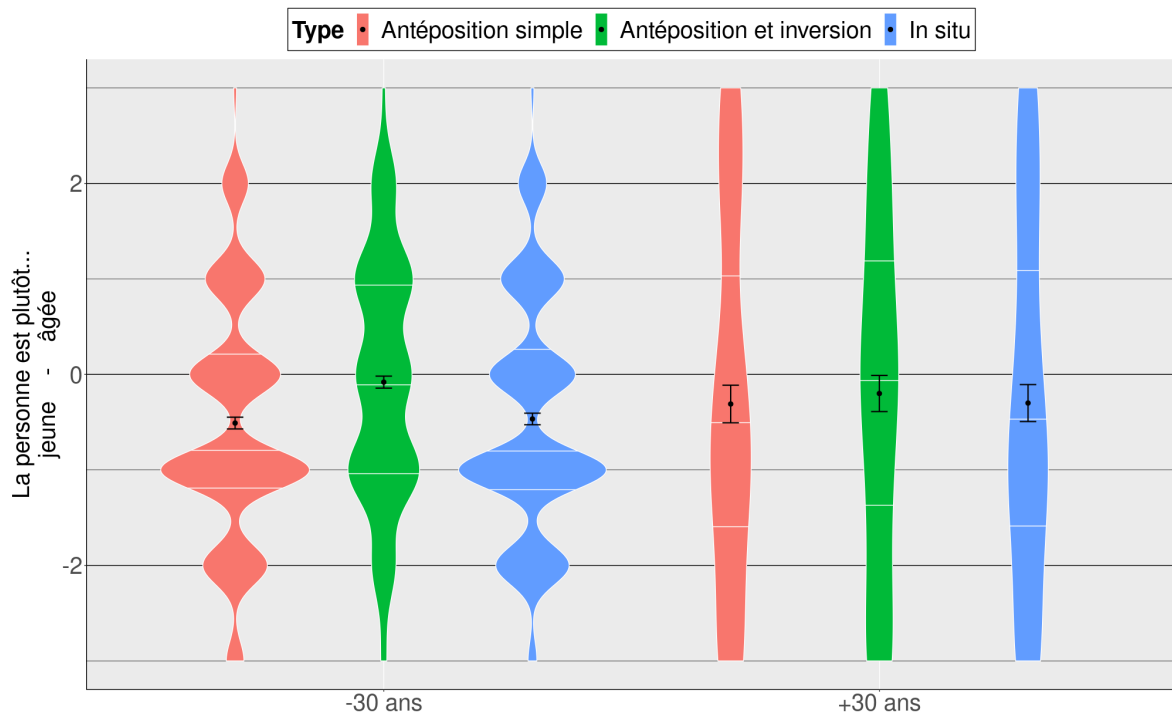


FIGURE K.3 – Note moyenne en fonction du type interrogatif et de l'âge (MG2, ÂGE)

	Antéposition simple	Antéposition et inversion	In situ
-30 ans	-0.51	-0.08	-0.47
+30 ans	-0.31	-0.20	-0.30

TABEAU K.3 – Note moyenne en fonction du type interrogatif et de l'âge (MG2, ÂGE)

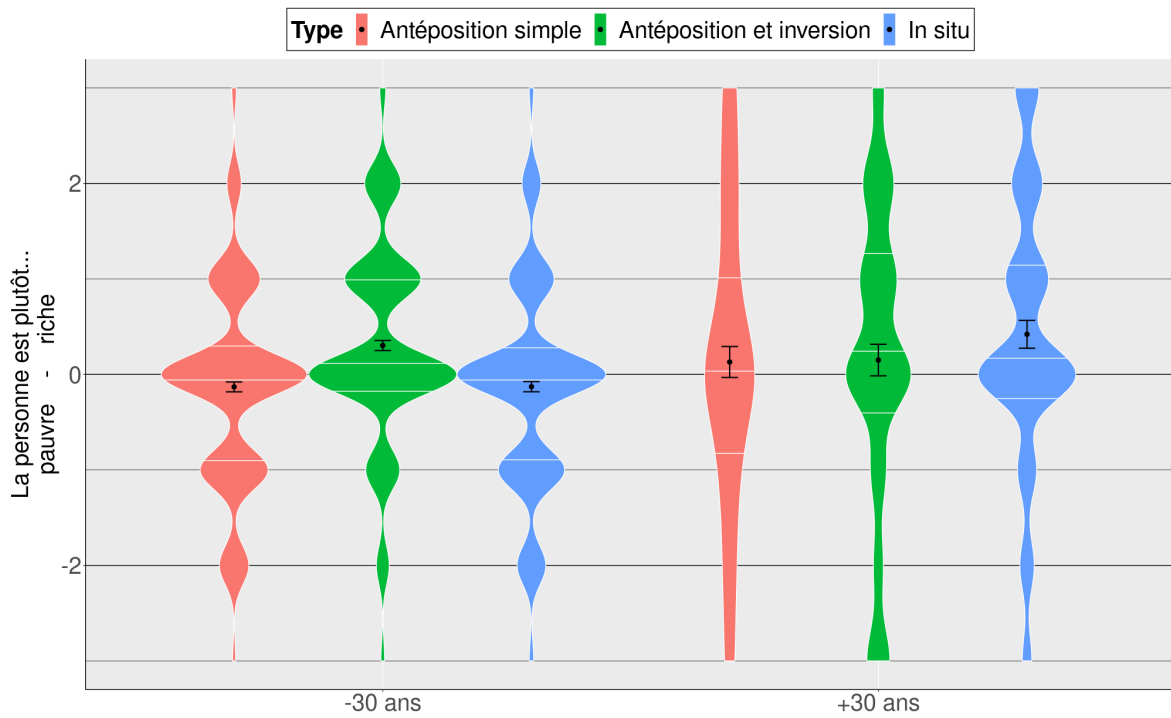


FIGURE K.4 – Note moyenne en fonction du type interrogatif et de l'âge (MG2, RICHESSE)

	Antéposition simple	Antéposition et inversion	In situ
-30 ans	-0.13	0.30	-0.13
+30 ans	0.13	0.15	0.42

TABEAU K.4 – Note moyenne en fonction du type interrogatif et de l'âge (MG2, RICHESSE)

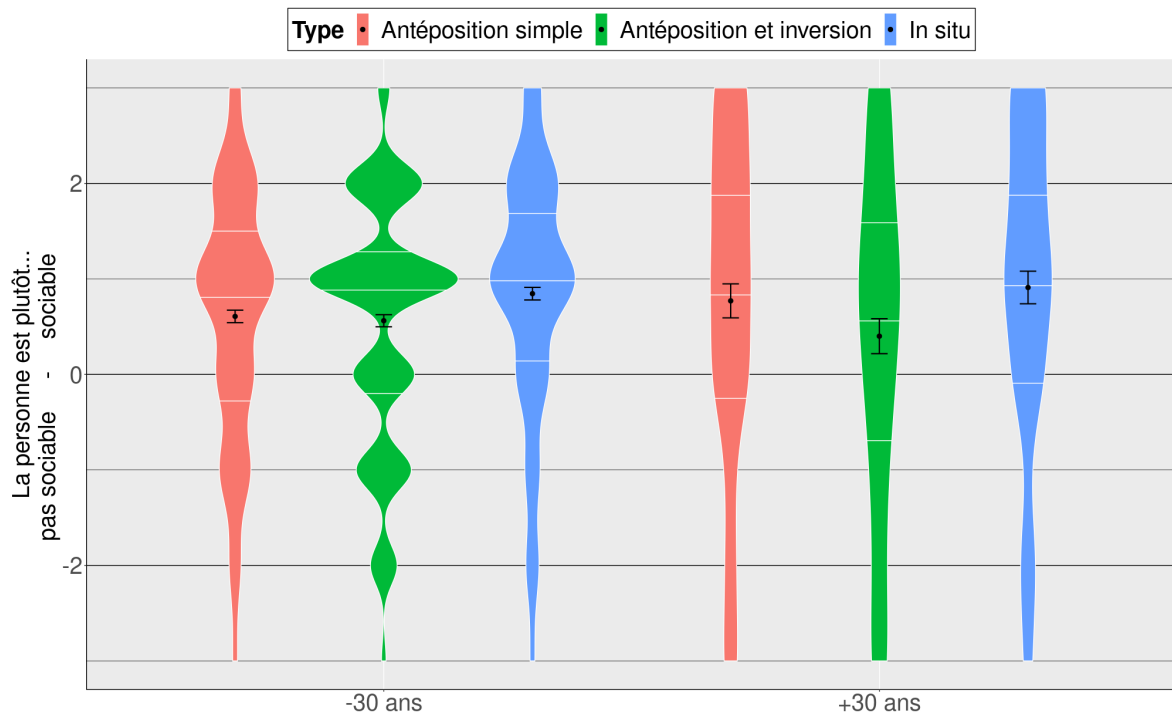


FIGURE K.5 – Note moyenne en fonction du type interrogatif et de l'âge (MG2, SOCIABILITÉ)

	Antéposition simple	Antéposition et inversion	In situ
-30 ans	0.61	0.56	0.85
+30 ans	0.77	0.40	0.91

TABLEAU K.5 – Note moyenne en fonction du type interrogatif et de l'âge (MG2, SOCIABILITÉ)

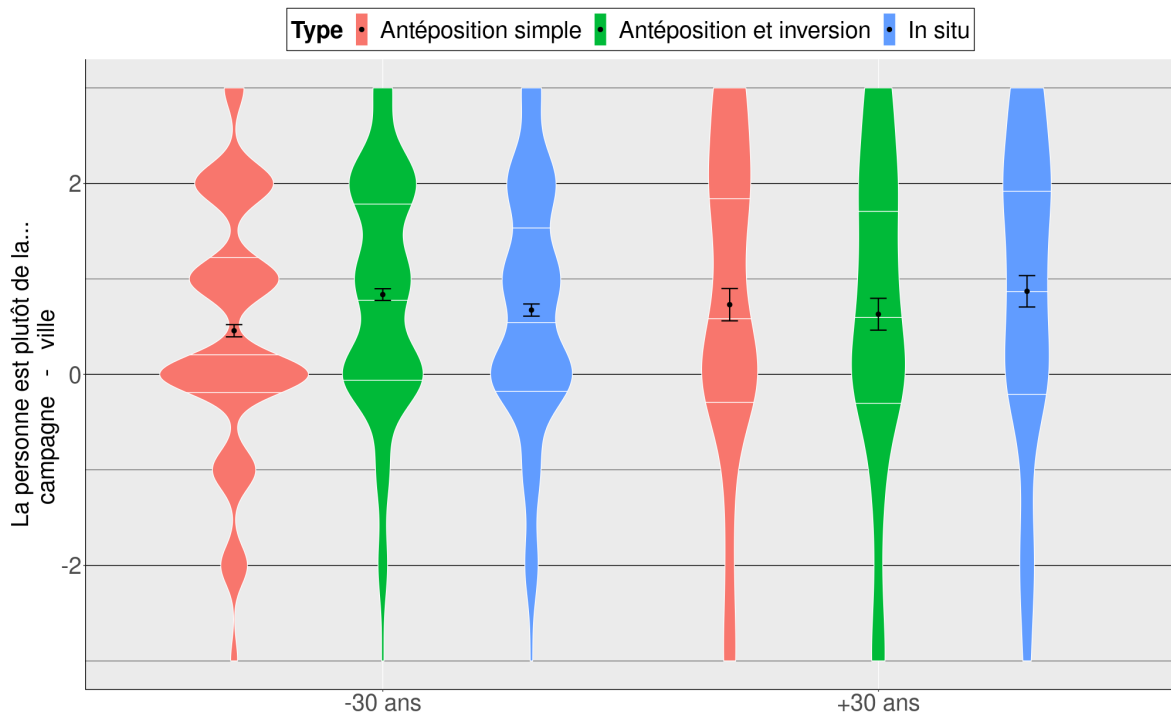


FIGURE K.6 – Note moyenne en fonction du type interrogatif et de l'âge (MG2, ORIGINE)

	Antéposition simple	Antéposition et inversion	In situ
-30 ans	0.46	0.84	0.67
+30 ans	0.73	0.63	0.87

TABLEAU K.6 – Note moyenne en fonction du type interrogatif et de l'âge (MG2, ORIGINE)

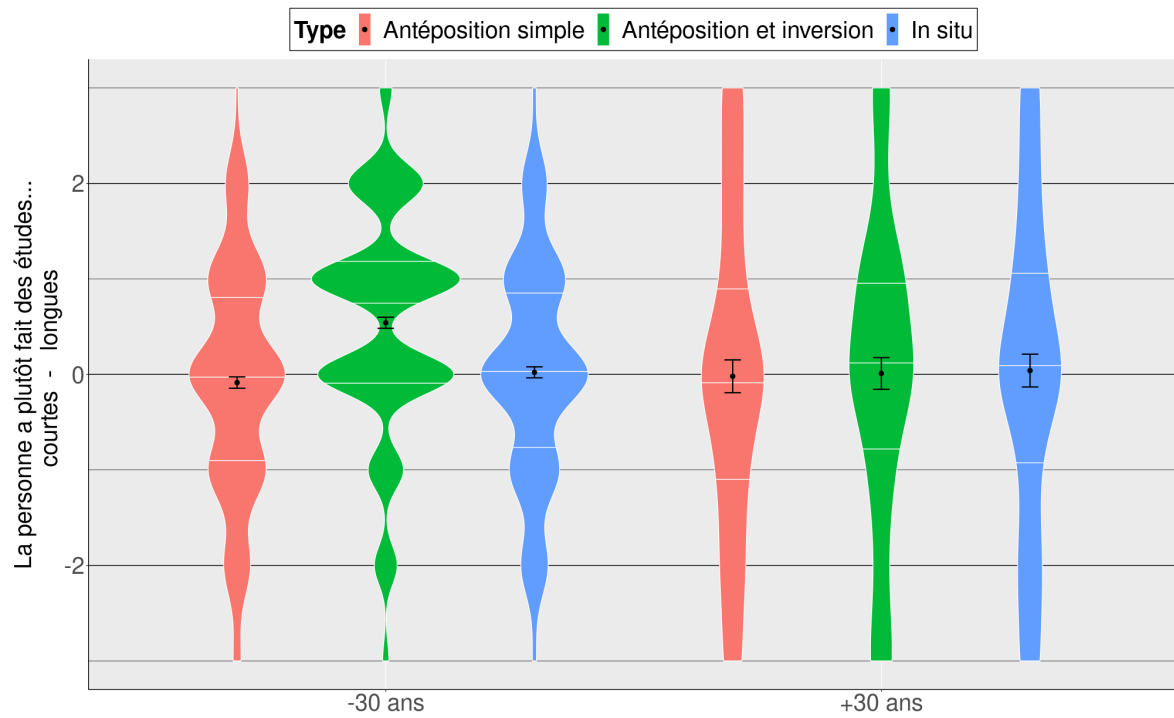


FIGURE K.7 – Note moyenne en fonction du type interrogatif et de l'âge (MG2, ÉTUDES)

	Antéposition simple	Antéposition et inversion	In situ
-30 ans	-0.09	0.54	0.02
+30 ans	-0.02	0.01	0.04

TABLEAU K.7 – Note moyenne en fonction du type interrogatif et de l'âge (MG2, ÉTUDES)

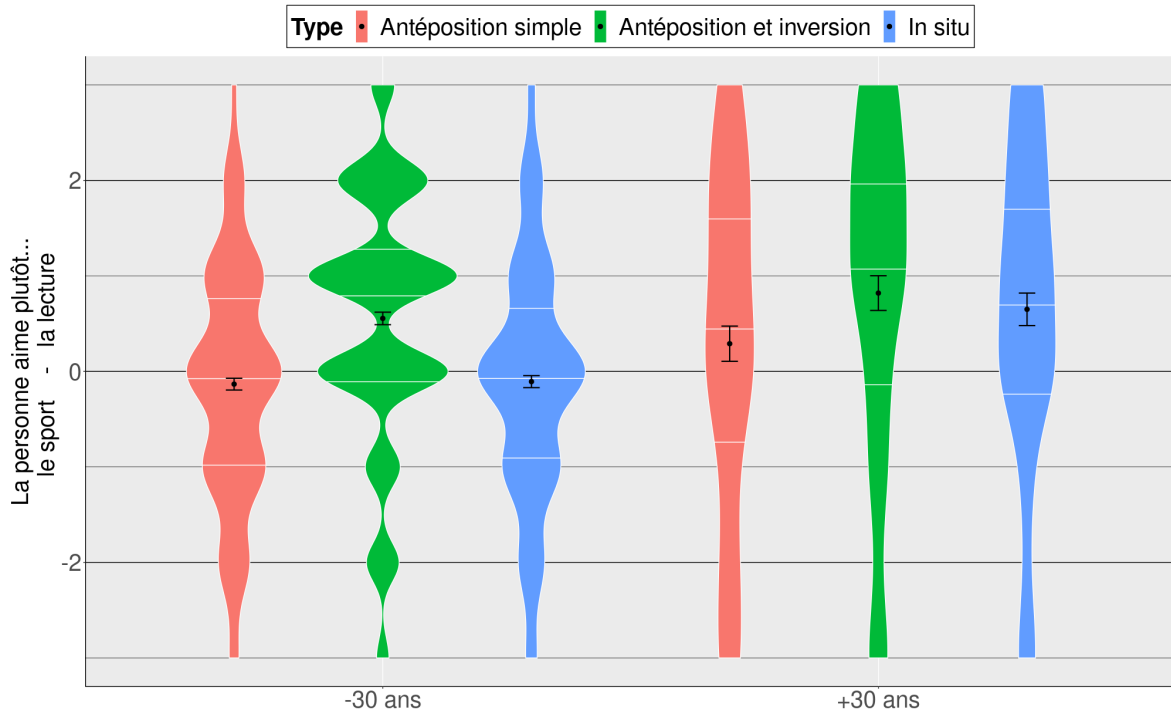


FIGURE K.8 – Note moyenne en fonction du type interrogatif et de l'âge (MG2, LOISIRS)

	Antéposition simple	Antéposition et inversion	In situ
-30 ans	-0.13	0.55	-0.11
+30 ans	0.29	0.82	0.65

TABLEAU K.8 – Note moyenne en fonction du type interrogatif et de l'âge (MG2, LOISIRS)

K.0.23 Résultats détaillés pour l'influence de la variable CONTEXTE (Formel/Informel)

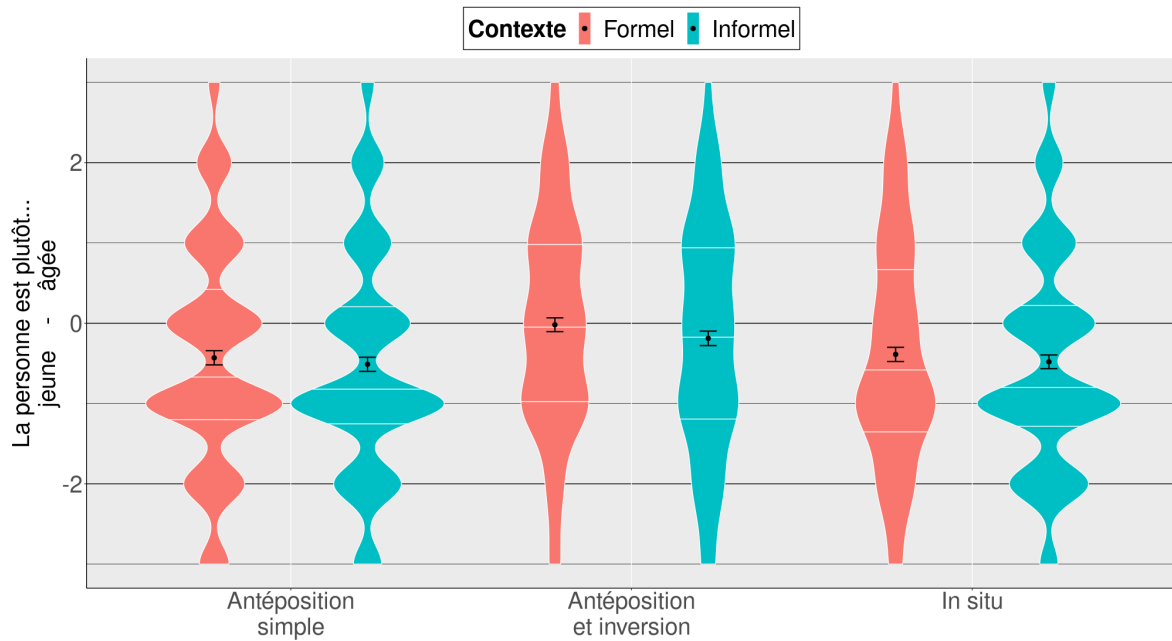


FIGURE K.9 – Note moyenne en fonction du type interrogatif et du contexte (MG2, ÂGE)

	Antéposition simple	Antéposition et inversion	In situ
Formel	-0.43	-0.02	-0.39
Informel	-0.51	-0.19	-0.48

TABLEAU K.9 – Note moyenne en fonction du type interrogatif et du contexte (MG2, ÂGE)

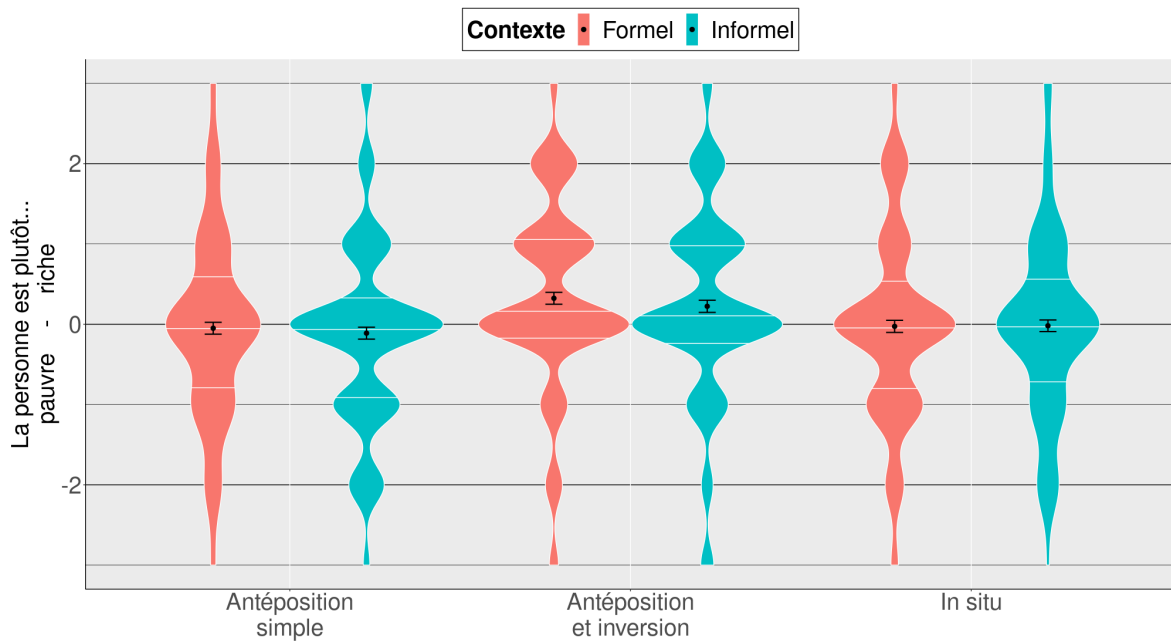


FIGURE K.10 – Note moyenne en fonction du type interrogatif et du contexte (MG2, RICHESSE)

	Antéposition simple	Antéposition et inversion	In situ
Formel	-0.05	0.32	-0.03
Informel	-0.11	0.22	-0.02

TABLEAU K.10 – Note moyenne en fonction du type interrogatif et du contexte (MG2, RICHESSE)

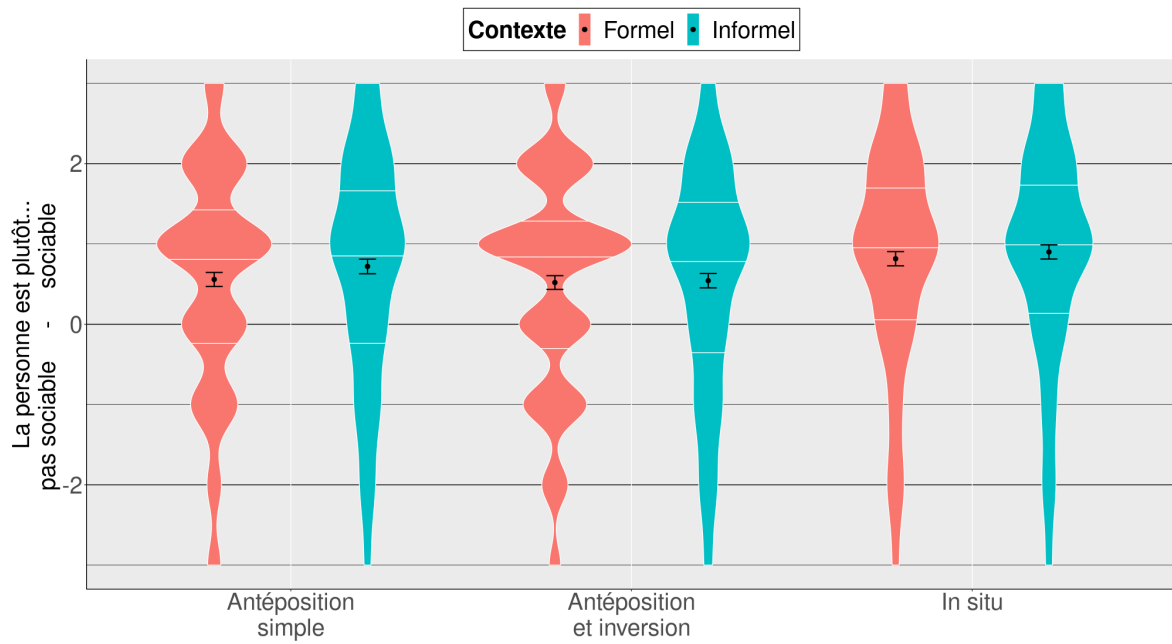


FIGURE K.11 – Note moyenne en fonction du type interrogatif et du contexte (MG2, SOCIABILITÉ)

	Antéposition simple	Antéposition et inversion	In situ
Formel	0.56	0.52	0.82
Informel	0.72	0.54	0.90

TABLEAU K.11 – Note moyenne en fonction du type interrogatif et du contexte (MG2, SOCIABILITÉ)

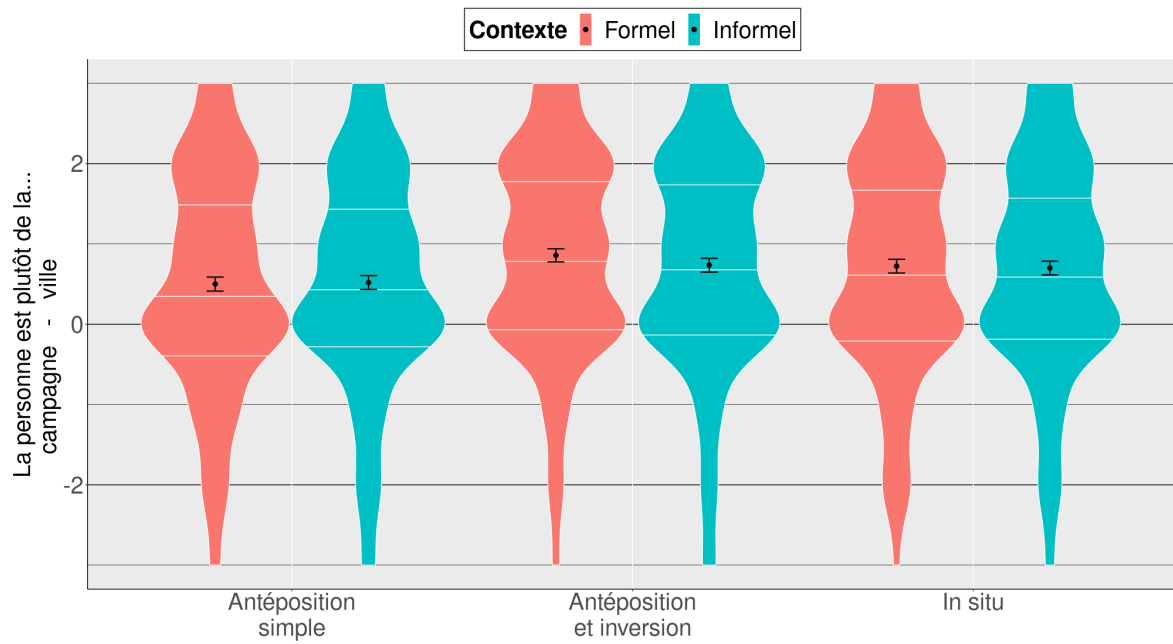


FIGURE K.12 – Note moyenne en fonction du type interrogatif et du contexte (MG2, ORIGINE)

	Antéposition simple	Antéposition et inversion	In situ
Formel	0.50	0.86	0.72
Informel	0.52	0.73	0.70

TABLEAU K.12 – Note moyenne en fonction du type interrogatif et du contexte (MG2, ORIGINE)

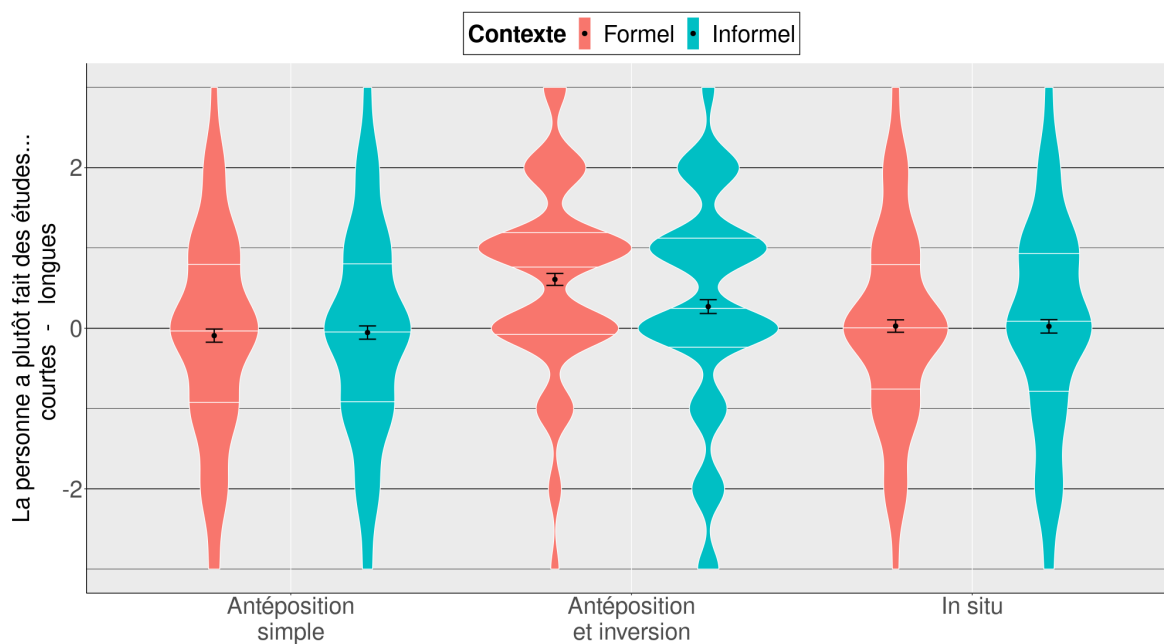


FIGURE K.13 – Note moyenne en fonction du type interrogatif et du contexte (MG2, ÉTUDES)

	Antéposition simple	Antéposition et inversion	In situ
Formel	-0.09	0.61	0.03
Informel	-0.05	0.27	0.02

TABLEAU K.13 – Note moyenne en fonction du type interrogatif et du contexte (MG2, ÉTUDES)

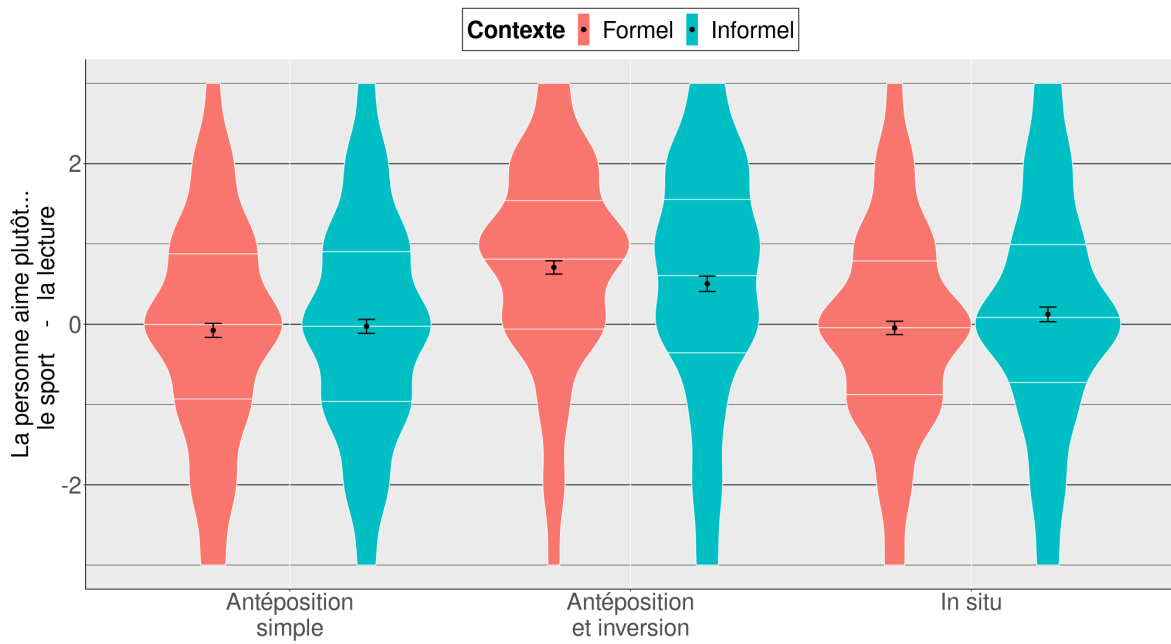


FIGURE K.14 – Note moyenne en fonction du type interrogatif et du contexte (MG2, LOISIRS)

	Antéposition simple	Antéposition et inversion	In situ
Formel	-0.08	0.71	-0.05
Informel	-0.03	0.50	0.12

TABEAU K.14 – Note moyenne en fonction du type interrogatif et du contexte (MG2, LOISIRS)

K.0.24 Résultats détaillés pour l'influence simultanée des variables GROUPE d'ÂGE (+/- 30 ans) et CONTEXTE (Formel / Informel)

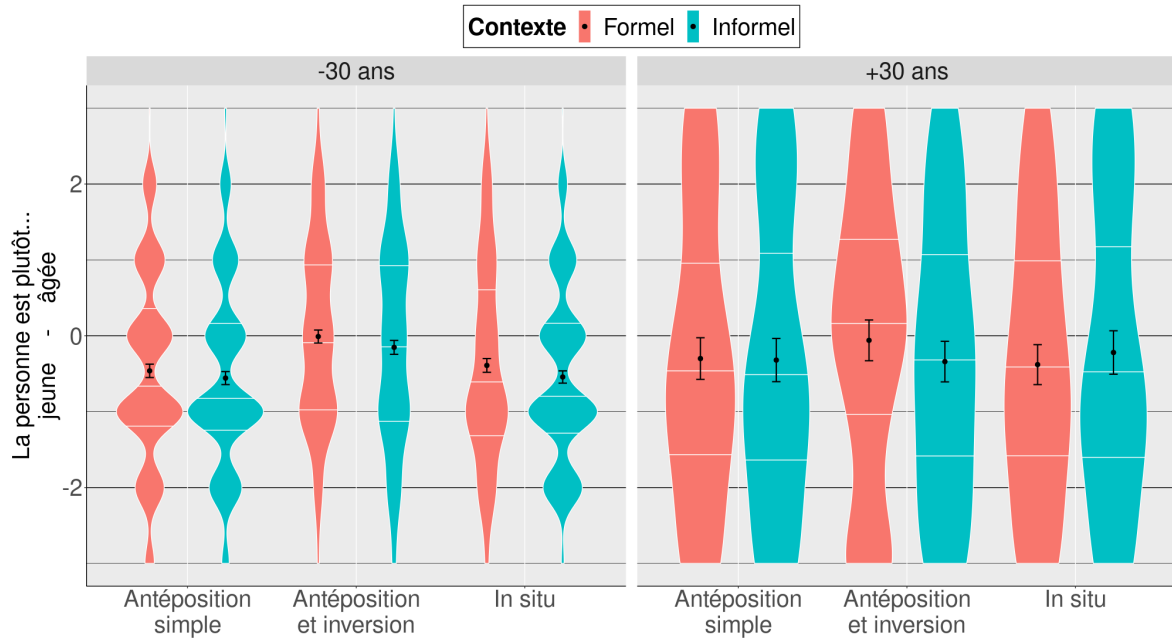


FIGURE K.15 – Note moyenne en fonction des type interrogatif, âge et contexte (MG2, ÂGE)

	-30 ans		+30 ans	
	Formel	Informel	Formel	Informel
Antéposition simple	-0.46	-0.56	-0.30	-0.32
Antéposition et inversion	-0.01	-0.15	-0.06	-0.34
In situ	-0.39	-0.54	-0.38	-0.22

TABLEAU K.15 – Note moyenne en fonction des type interrogatif, âge et contexte (MG2, ÂGE)

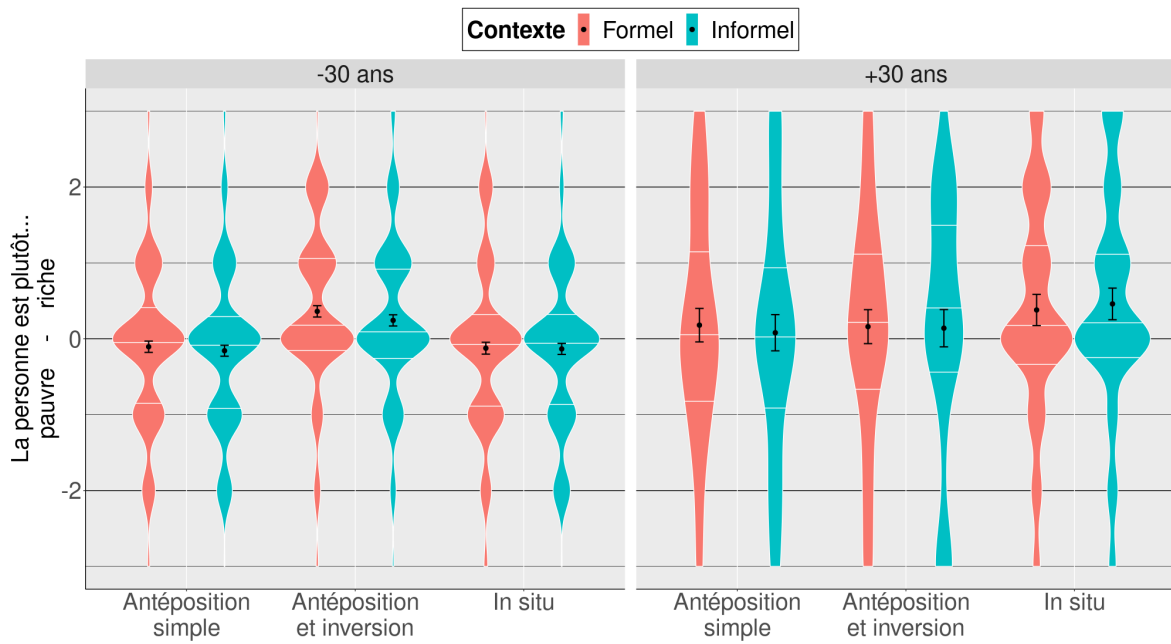


FIGURE K.16 – Note moyenne en fonction des type interrogatif, âge et contexte (MG2, RICHESSE)

	-30 ans		+30 ans	
	Formel	Informel	Formel	Informel
Antéposition simple	-0.10	-0.16	0.18	0.08
Antéposition et inversion	0.36	0.24	0.16	0.14
In situ	-0.12	-0.13	0.38	0.46

TABLEAU K.16 – Note moyenne en fonction des type interrogatif, âge et contexte (MG2, RICHESSE)

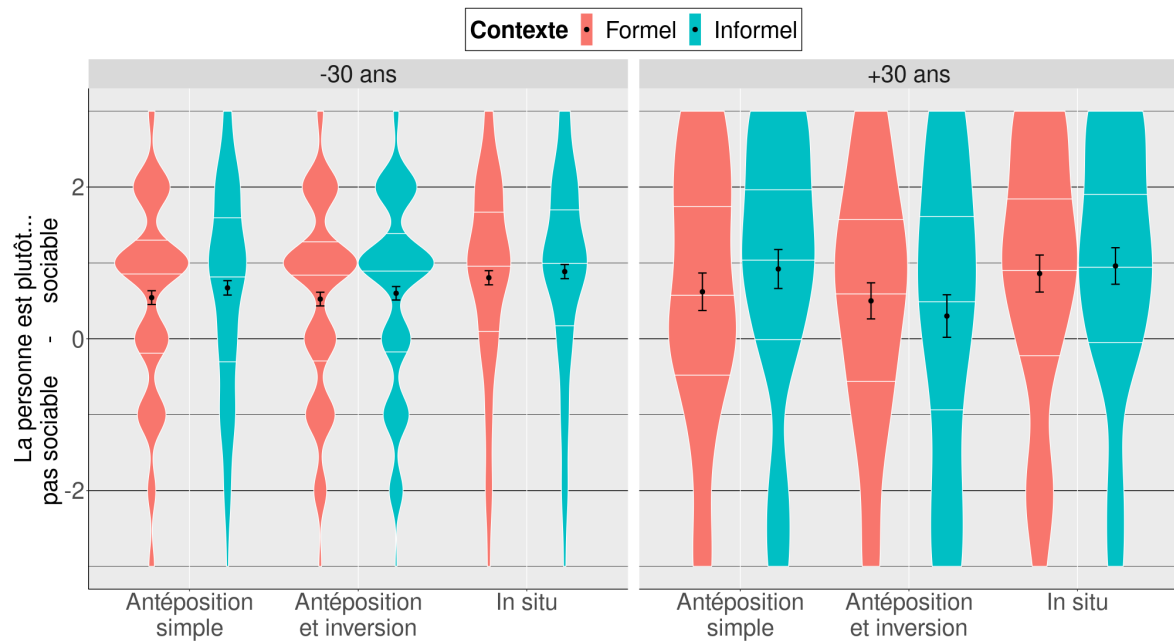


FIGURE K.17 – Note moyenne en fonction des type interrogatif, âge et contexte (MG2, SOCIABILITÉ)

	-30 ans		+30 ans	
	Formel	Informel	Formel	Informel
Antéposition simple	0.54	0.67	0.62	0.92
Antéposition et inversion	0.52	0.60	0.50	0.30
In situ	0.80	0.89	0.86	0.96

TABEAU K.17 – Note moyenne en fonction des type interrogatif, âge et contexte (MG2, SOCIABILITÉ)

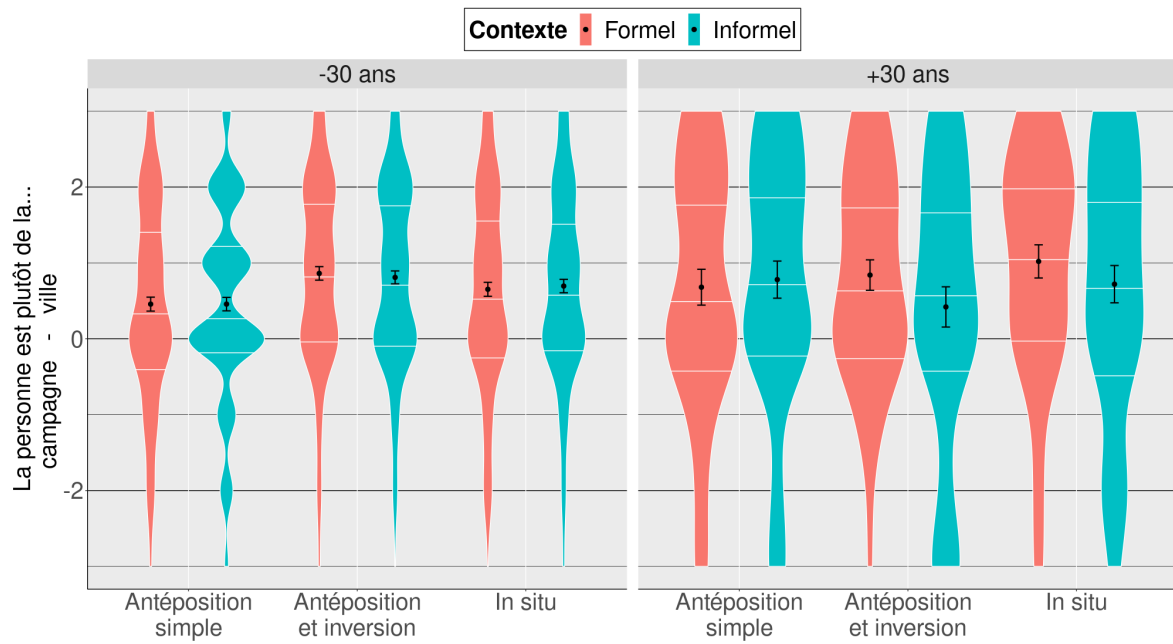


FIGURE K.18 – Note moyenne en fonction des type interrogatif, âge et contexte (MG2, ORIGINE)

	-30 ans		+30 ans	
	Formel	Informel	Formel	Informel
Antéposition simple	0.46	0.46	0.68	0.78
Antéposition et inversion	0.86	0.81	0.84	0.42
In situ	0.65	0.70	1.02	0.72

TABLEAU K.18 – Note moyenne en fonction des type interrogatif, âge et contexte (MG2, ORIGINE)

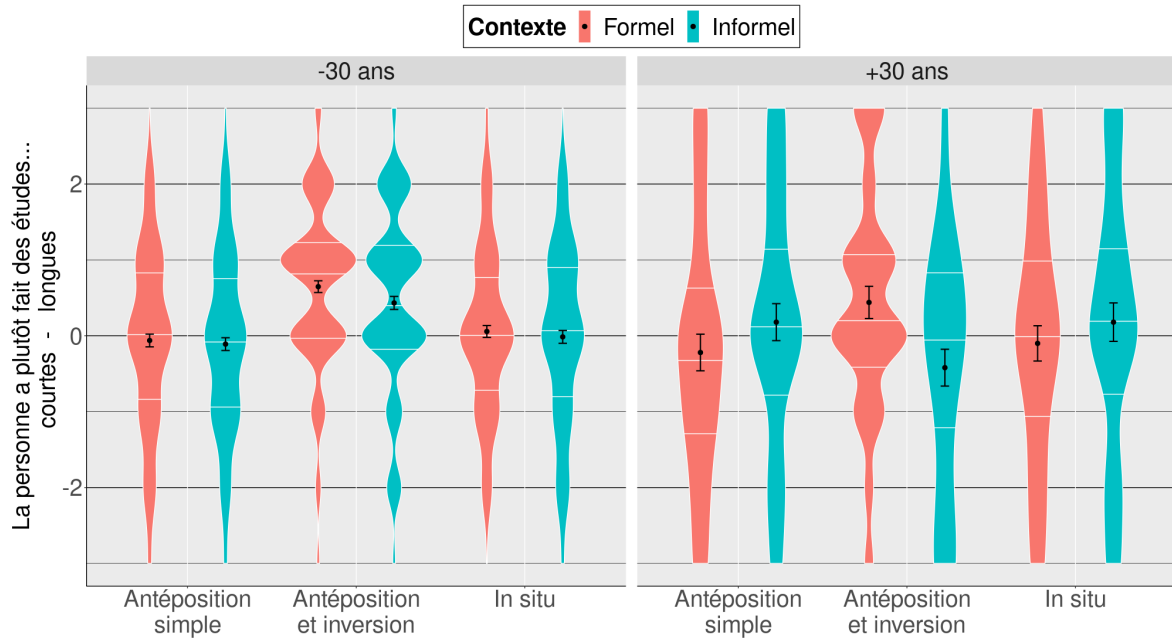


FIGURE K.19 – Note moyenne en fonction des type interrogatif, âge et contexte (MG2, ÉTUDES)

	-30 ans		+30 ans	
	Formel	Informel	Formel	Informel
Antéposition simple	-0.06	-0.11	-0.22	0.18
Antéposition et inversion	0.65	0.43	0.44	-0.42
In situ	0.06	-0.01	-0.10	0.18

TABLEAU K.19 – Note moyenne en fonction des type interrogatif, âge et contexte (MG2, ÉTUDES)

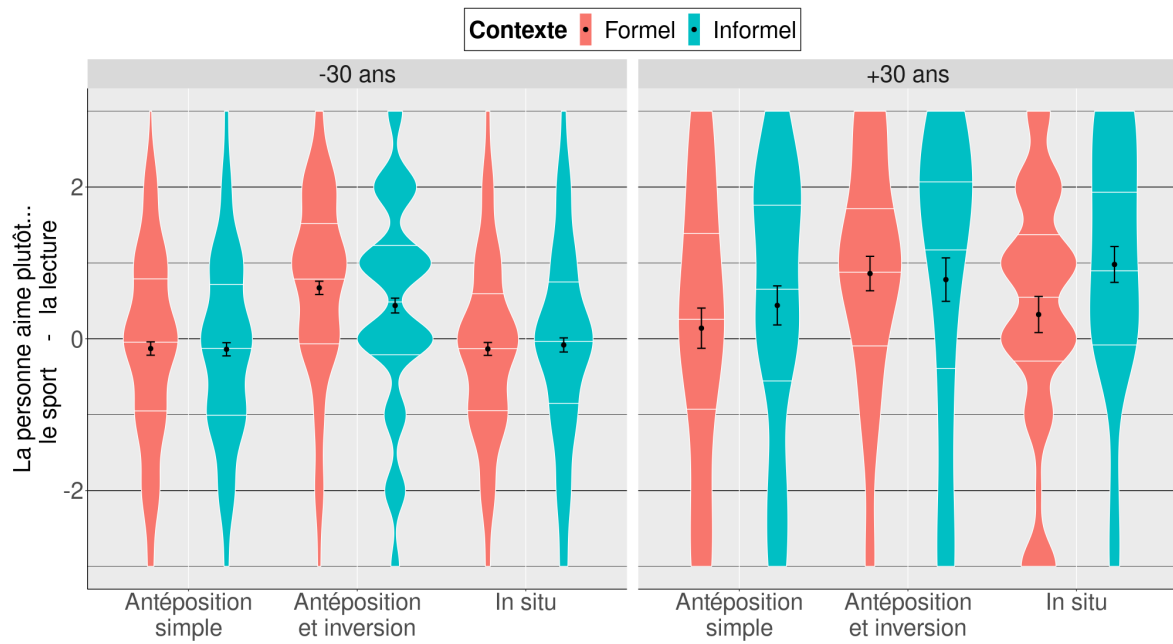


FIGURE K.20 – Note moyenne en fonction des type interrogatif, âge et contexte (MG2, LOISIRS)

	-30 ans		+30 ans	
	Formel	Informel	Formel	Informel
Antéposition simple	-0.13	-0.13	0.14	0.44
Antéposition et inversion	0.67	0.44	0.86	0.78
In situ	-0.13	-0.08	0.32	0.98

TABLEAU K.20 – Note moyenne en fonction des type interrogatif, âge et contexte (MG2, LOISIRS)

K.0.25 Modèle *brms* et distributions postérieures pour l'échelle ÂGE (5.3.2)**Family :** cumulative**Links :** mu = logit; disc = identity**Formula :** age ~ (f_finv + is_finv) * old_young * formal_informal + ((f_finv + is_finv) * formal_informal | id) + ((f_finv + is_finv) * old_young * formal_informal | Item)**Data :** DONmg (Number of observations : 1560)**Samples :** 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-3.80	0.25	-4.30	-3.33	1.00	2179	3765
Intercept[2]	-1.96	0.22	-2.38	-1.53	1.00	1877	3466
Intercept[3]	0.06	0.21	-0.35	0.48	1.00	1794	3418
Intercept[4]	1.32	0.21	0.91	1.74	1.00	1789	3778
Intercept[5]	2.74	0.23	2.30	3.19	1.00	1959	4450
Intercept[6]	4.73	0.29	4.17	5.29	1.00	2723	6212
f_finv	-0.63	0.16	-0.94	-0.33	1.00	7734	7652
is_finv	-0.54	0.17	-0.88	-0.20	1.00	7301	7899
old_young	0.08	0.53	-0.96	1.12	1.00	2317	4241
formal_informal	0.22	0.12	-0.01	0.44	1.00	9638	8351
f_finv :old_young	0.38	0.56	-0.69	1.49	1.00	5998	6568
is_finv :old_young	0.60	0.60	-0.57	1.78	1.00	6925	7645
f_finv :formal_informal	-0.06	0.28	-0.61	0.50	1.00	8848	7953
is_finv :formal_informal	-0.16	0.30	-0.74	0.45	1.00	8638	8918
old_young :formal_informal	-0.06	0.39	-0.81	0.73	1.00	7931	8812
f_finv :old_young :formal_informal	-0.08	0.85	-1.73	1.60	1.00	8492	6695
is_finv :old_young :formal_informal	-0.77	1.04	-2.83	1.30	1.00	7625	8487

TABLEAU K.21 – Résumé du modèle *brms* (MG2, ÂGE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Type : F vs. FINV	-0.63	0.00	1.00	-1.24	-0.02	-0.94	-0.33	-0.89	-0.38	-0.85	-0.41
Type : IS vs. FINV	-0.54	0.00	1.00	-1.22	0.20	-0.88	-0.20	-0.83	-0.26	-0.79	-0.30
Âge : +30 ans vs. -30 ans	0.08	0.56	0.44	-1.94	2.20	-0.96	1.12	-0.79	0.95	-0.68	0.84
Formalité : Formel vs. Informel	0.22	0.97	0.03	-0.24	0.72	-0.01	0.44	0.03	0.41	0.05	0.38
F vs. FINV * Âge	0.38	0.75	0.25	-2.23	3.59	-0.69	1.49	-0.53	1.30	-0.41	1.17
IS vs. FINV * Âge	0.60	0.85	0.15	-1.80	2.94	-0.57	1.78	-0.38	1.60	-0.26	1.46
F vs. FINV * Formalité	-0.06	0.41	0.59	-1.16	1.05	-0.61	0.50	-0.53	0.40	-0.46	0.34
IS vs. FINV * Formalité	-0.16	0.29	0.71	-1.39	1.34	-0.74	0.45	-0.64	0.33	-0.59	0.26
Âge * Formalité	-0.06	0.44	0.56	-1.61	1.48	-0.81	0.73	-0.68	0.59	-0.61	0.51
F vs. FINV * Âge * Formalité	-0.08	0.45	0.55	-3.70	3.91	-1.73	1.60	-1.47	1.32	-1.28	1.13
IS vs. FINV * Âge * Formalité	-0.77	0.22	0.78	-5.95	3.49	-2.83	1.30	-2.49	0.94	-2.28	0.70

TABLEAU K.22 – Distributions postérieures pour le modèle brms (MG2, ÂGE)

K.0.26 Modèle *brms* et distributions postérieures pour l'échelle *RICHESS* (5.3.2)**Family :** cumulative**Links :** mu = logit; disc = identity**Formula :** richness ~ (f_finv + is_finv) * old_young * formal_informal + ((f_finv + is_finv) * formal_informal | id) + ((f_finv + is_finv) * old_young * formal_informal | Item)**Data :** DONmg (Number of observations : 1560)**Samples :** 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-4.50	0.24	-4.98	-4.04	1.00	3125	5700
Intercept[2]	-2.78	0.17	-3.13	-2.45	1.00	2896	5449
Intercept[3]	-1.35	0.15	-1.66	-1.05	1.00	2655	5357
Intercept[4]	1.11	0.15	0.81	1.41	1.00	2780	5182
Intercept[5]	2.61	0.17	2.28	2.95	1.00	3293	6319
Intercept[6]	4.36	0.23	3.92	4.82	1.00	4577	7297
f_finv	-0.70	0.17	-1.03	-0.37	1.00	6193	7513
is_finv	-0.59	0.20	-0.98	-0.21	1.00	5208	7241
old_young	0.42	0.40	-0.36	1.21	1.00	2800	5208
formal_informal	0.15	0.12	-0.10	0.39	1.00	9348	8530
f_finv:old_young	0.96	0.55	-0.11	2.06	1.00	5374	6314
is_finv:old_young	1.62	0.61	0.45	2.86	1.00	4787	6618
f_finv:formal_informal	-0.11	0.28	-0.66	0.46	1.00	9223	9007
is_finv:formal_informal	-0.21	0.27	-0.74	0.33	1.00	9470	8720
old_young:formal_informal	0.16	0.48	-0.81	1.12	1.00	6966	8281
f_finv:old_young:formal_informal	0.20	1.03	-1.80	2.22	1.00	6862	7933
is_finv:old_young:formal_informal	0.12	0.88	-1.61	1.89	1.00	6539	7087

TABLEAU K.23 – Résumé du modèle *brms* (MG2, RICHESSE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Type : F vs. FINV	-0.70	0.00	1.00	-1.32	0.01	-1.03	-0.37	-0.97	-0.43	-0.94	-0.46
Type : IS vs. FINV	-0.59	0.00	1.00	-1.51	0.29	-0.98	-0.21	-0.92	-0.27	-0.87	-0.31
Âge : +30 ans vs. -30 ans	0.42	0.86	0.14	-1.03	2.12	-0.36	1.21	-0.23	1.08	-0.15	1.00
Formalité : Formel vs. Informel	0.15	0.88	0.12	-0.28	0.63	-0.10	0.39	-0.05	0.35	-0.03	0.32
F vs. FINV * Âge	0.96	0.96	0.04	-1.07	3.18	-0.11	2.06	0.06	1.88	0.17	1.76
IS vs. FINV * Âge	1.62	1.00	0.00	-0.76	4.24	0.45	2.86	0.64	2.63	0.76	2.49
F vs. FINV * Formalité	-0.11	0.35	0.65	-1.10	0.87	-0.66	0.46	-0.57	0.36	-0.50	0.30
IS vs. FINV * Formalité	-0.21	0.21	0.79	-1.22	0.78	-0.74	0.33	-0.65	0.24	-0.59	0.18
Âge * Formalité	0.16	0.63	0.37	-1.98	2.54	-0.81	1.12	-0.63	0.95	-0.54	0.84
F vs. FINV * Âge * Formalité	0.20	0.58	0.42	-5.02	4.57	-1.80	2.22	-1.47	1.88	-1.25	1.65
IS vs. FINV * Âge * Formalité	0.12	0.55	0.45	-3.43	3.91	-1.61	1.89	-1.31	1.58	-1.12	1.38

TABLEAU K.24 – Distributions postérieures pour le modèle brms (MG2, RICHESSE)

K.0.27 Modèle *brms* et distributions postérieures pour l'échelle *SOCIABILITÉ* (5.3.2)**Family :** cumulative**Links :** mu = logit; disc = identity**Formula :** sociable ~ (f_finv + is_finv) * old_young * formal_informal + ((f_finv + is_finv) * formal_informal | id) + ((f_finv + is_finv) * old_young * formal_informal | Item)**Data :** DONmg (Number of observations : 1560)**Samples :** 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-4.13	0.22	-4.59	-3.71	1.00	2963	4946
Intercept[2]	-2.81	0.19	-3.18	-2.46	1.00	2250	4453
Intercept[3]	-1.79	0.17	-2.14	-1.46	1.00	2010	4714
Intercept[4]	-0.78	0.17	-1.11	-0.45	1.00	1965	4151
Intercept[5]	1.14	0.17	0.82	1.47	1.00	2090	4536
Intercept[6]	3.12	0.19	2.74	3.50	1.00	2698	5477
f_finv	0.18	0.14	-0.09	0.44	1.00	9638	8938
is_finv	0.57	0.19	0.19	0.95	1.00	6088	6720
old_young	0.26	0.40	-0.52	1.04	1.00	2325	4073
formal_informal	-0.14	0.11	-0.37	0.09	1.00	10762	8912
f_finv:old_young	0.62	0.40	-0.17	1.41	1.00	7901	8470
is_finv:old_young	0.56	0.54	-0.51	1.61	1.00	5984	7765
f_finv:formal_informal	-0.23	0.31	-0.83	0.39	1.00	7441	8320
is_finv:formal_informal	-0.08	0.26	-0.58	0.43	1.00	10912	9212
old_young:formal_informal	-0.03	0.34	-0.72	0.64	1.00	8590	8160
f_finv:old_young:formal_informal	-0.80	0.88	-2.56	0.91	1.00	7733	8027
is_finv:old_young:formal_informal	-0.44	0.87	-2.15	1.31	1.00	8051	8332

TABLEAU K.25 – Résumé du modèle *brms* (MG2, SOCIABILITÉ)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Type : F vs. FINV	0.18	0.90	0.10	-0.35	0.75	-0.09	0.44	-0.04	0.40	-0.02	0.37
Type : IS vs. FINV	0.57	1.00	0.00	-0.30	1.37	0.19	0.95	0.26	0.89	0.30	0.85
Âge : +30 ans vs. -30 ans	0.26	0.74	0.26	-1.62	1.84	-0.52	1.04	-0.39	0.92	-0.31	0.82
Formalité : Formel vs. Informel	-0.14	0.11	0.89	-0.60	0.30	-0.37	0.09	-0.33	0.05	-0.31	0.02
F vs. FINV * Âge	0.62	0.94	0.06	-0.89	2.26	-0.17	1.41	-0.03	1.27	0.06	1.19
IS vs. FINV * Âge	0.56	0.85	0.15	-1.91	2.90	-0.51	1.61	-0.32	1.44	-0.21	1.33
F vs. FINV * Formalité	-0.23	0.23	0.77	-1.52	0.97	-0.83	0.39	-0.74	0.28	-0.67	0.22
IS vs. FINV * Formalité	-0.08	0.38	0.62	-1.04	0.97	-0.58	0.43	-0.49	0.35	-0.44	0.30
Âge * Formalité	-0.03	0.46	0.54	-1.46	1.46	-0.72	0.64	-0.60	0.53	-0.53	0.45
F vs. FINV * Âge * Formalité	-0.80	0.18	0.82	-4.20	2.54	-2.56	0.91	-2.25	0.61	-2.06	0.46
IS vs. FINV * Âge * Formalité	-0.44	0.30	0.70	-4.07	3.27	-2.15	1.31	-1.85	1.00	-1.66	0.79

TABLEAU K.26 – Distributions postérieures pour le modèle brms (MG2, SOCIABILITÉ)

K.0.28 Modèle *brms* et distributions postérieures pour l'échelle *ORIGINE* (5.3.2)**Family :** cumulative**Links :** mu = logit; disc = identity**Formula :** origin ~ (f_finv + is_finv) * old_young * formal_informal + ((f_finv + is_finv) * formal_informal | id) + ((f_finv + is_finv) * old_young * formal_informal | Item)**Data :** DONmg (Number of observations : 1560)**Samples :** 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-4.46	0.25	-4.97	-3.97	1.00	2874	5184
Intercept[2]	-3.10	0.20	-3.52	-2.72	1.00	2420	4214
Intercept[3]	-2.27	0.19	-2.65	-1.91	1.00	2254	4460
Intercept[4]	-0.19	0.18	-0.55	0.15	1.00	2164	3986
Intercept[5]	1.03	0.18	0.68	1.39	1.00	2223	4152
Intercept[6]	2.99	0.20	2.60	3.39	1.00	2698	5214
f_finv	-0.43	0.16	-0.76	-0.10	1.00	7448	7018
is_finv	-0.15	0.14	-0.42	0.13	1.00	8961	8829
old_young	0.24	0.40	-0.52	1.03	1.00	2641	4711
formal_informal	0.11	0.11	-0.11	0.34	1.00	8687	8555
f_finv:old_young	0.81	0.42	-0.02	1.65	1.00	6396	7105
is_finv:old_young	0.69	0.41	-0.12	1.50	1.00	6879	7696
f_finv:formal_informal	-0.31	0.25	-0.81	0.19	1.00	9535	8120
is_finv:formal_informal	-0.18	0.29	-0.75	0.39	1.00	7925	7503
old_young:formal_informal	0.45	0.36	-0.26	1.16	1.00	7342	7773
f_finv:old_young:formal_informal	-0.77	0.71	-2.15	0.65	1.00	8612	8715
is_finv:old_young:formal_informal	-0.08	0.81	-1.69	1.51	1.00	7400	8101

TABLEAU K.27 – Résumé du modèle *brms* (MG2, ORIGINE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Type : F vs. FINV	-0.43	0.01	0.99	-1.25	0.31	-0.76	-0.10	-0.70	-0.16	-0.66	-0.19
Type : IS vs. FINV	-0.15	0.14	0.86	-0.74	0.40	-0.42	0.13	-0.37	0.08	-0.34	0.05
Âge : +30 ans vs. -30 ans	0.24	0.73	0.27	-1.38	1.81	-0.52	1.03	-0.41	0.90	-0.33	0.81
Formalité : Formel vs. Informel	0.11	0.83	0.17	-0.31	0.63	-0.11	0.34	-0.08	0.30	-0.05	0.27
F vs. FINV * Âge	0.81	0.97	0.03	-0.91	2.68	-0.02	1.65	0.12	1.51	0.20	1.42
IS vs. FINV * Âge	0.69	0.95	0.05	-0.91	2.45	-0.12	1.50	0.01	1.35	0.11	1.27
F vs. FINV * Formalité	-0.31	0.11	0.89	-1.23	0.65	-0.81	0.19	-0.74	0.11	-0.68	0.05
IS vs. FINV * Formalité	-0.18	0.27	0.73	-1.47	1.01	-0.75	0.39	-0.65	0.30	-0.59	0.24
Âge * Formalité	0.45	0.90	0.10	-0.96	2.09	-0.26	1.16	-0.15	1.04	-0.07	0.96
F vs. FINV * Âge * Formalité	-0.77	0.14	0.86	-3.75	2.34	-2.15	0.65	-1.93	0.41	-1.78	0.24
IS vs. FINV * Âge * Formalité	-0.08	0.46	0.54	-2.88	3.64	-1.69	1.51	-1.40	1.23	-1.22	1.07

TABLEAU K.28 – Distributions postérieures pour le modèle brms (MG2, ORIGINE)

K.0.29 Modèle *brms* et distributions postérieures pour l'échelle *ÉTUDES* (5.3.2)**Family :** cumulative**Links :** mu = logit; disc = identity**Formula :** education ~ (f_finv + is_finv) * old_young * formal_informal + ((f_finv + is_finv) * formal_informal | id) + ((f_finv + is_finv) * old_young * formal_informal | Item)**Data :** DONmg (Number of observations : 1560)**Samples :** 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-4.15	0.22	-4.59	-3.72	1.00	3047	4492
Intercept[2]	-2.54	0.18	-2.91	-2.20	1.00	2781	4500
Intercept[3]	-1.28	0.16	-1.61	-0.96	1.00	2726	4260
Intercept[4]	0.64	0.16	0.32	0.95	1.00	2695	4237
Intercept[5]	2.31	0.17	1.98	2.65	1.00	2913	5151
Intercept[6]	4.15	0.22	3.73	4.59	1.00	3668	7672
f_finv	-0.92	0.19	-1.30	-0.55	1.00	4935	7017
is_finv	-0.74	0.17	-1.09	-0.40	1.00	5843	7803
old_young	-0.29	0.38	-1.04	0.47	1.00	2880	4999
formal_informal	0.19	0.11	-0.03	0.41	1.00	9160	8663
f_finv:old_young	1.07	0.55	0.01	2.15	1.00	4613	7055
is_finv:old_young	1.14	0.49	0.19	2.11	1.00	6112	7404
f_finv:formal_informal	-0.66	0.26	-1.17	-0.15	1.00	9470	8270
is_finv:formal_informal	-0.60	0.29	-1.18	-0.05	1.00	7884	7049
old_young:formal_informal	0.01	0.41	-0.79	0.82	1.00	7522	8023
f_finv:old_young:formal_informal	-2.15	0.78	-3.71	-0.64	1.00	7498	7743
is_finv:old_young:formal_informal	-1.67	0.86	-3.39	0.00	1.00	6783	7116

TABLEAU K.29 – Résumé du modèle *brms* (MG2, *ÉTUDES*)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Type : F vs. FINV	-0.92	0.00	1.00	-1.70	-0.17	-1.30	-0.55	-1.23	-0.61	-1.19	-0.65
Type : IS vs. FINV	-0.74	0.00	1.00	-1.36	-0.07	-1.09	-0.40	-1.03	-0.46	-0.99	-0.50
Âge : +30 ans vs. -30 ans	-0.29	0.23	0.77	-2.17	1.21	-1.04	0.47	-0.91	0.34	-0.83	0.26
Formalité : Formel vs. Informel	0.19	0.96	0.04	-0.25	0.66	-0.03	0.41	0.01	0.37	0.03	0.35
F vs. FINV * Âge	1.07	0.98	0.02	-1.33	3.29	0.01	2.15	0.18	1.98	0.29	1.87
IS vs. FINV * Âge	1.14	0.99	0.01	-1.02	3.32	0.19	2.11	0.34	1.94	0.44	1.84
F vs. FINV * Formalité	-0.66	0.01	0.99	-1.68	0.37	-1.17	-0.15	-1.09	-0.23	-1.03	-0.28
IS vs. FINV * Formalité	-0.60	0.02	0.98	-1.96	0.52	-1.18	-0.05	-1.07	-0.14	-1.02	-0.20
Âge * Formalité	0.01	0.51	0.49	-1.94	2.39	-0.79	0.82	-0.66	0.68	-0.57	0.59
F vs. FINV * Âge * Formalité	-2.15	0.00	1.00	-5.98	0.86	-3.71	-0.64	-3.45	-0.88	-3.27	-1.04
IS vs. FINV * Âge * Formalité	-1.67	0.03	0.97	-5.05	1.88	-3.39	0.00	-3.10	-0.27	-2.90	-0.45

TABLEAU K.30 – Distributions postérieures pour le modèle brms (MG2, ÉTUDES)

K.0.30 Modèle *brms* et distributions postérieures pour l'échelle *LOISIRS* (5.3.2)**Family :** cumulative**Links :** mu = logit; disc = identity**Formula :** hobby ~ (f_finv + is_finv) * old_young * formal_informal + ((f_finv + is_finv) * formal_informal | id) + ((f_finv + is_finv) * old_young * formal_informal | Item)**Data :** DONmg (Number of observations : 1560)**Samples :** 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-3.80	0.21	-4.22	-3.39	1.00	3806	5580
Intercept[2]	-2.51	0.18	-2.88	-2.16	1.00	3304	5601
Intercept[3]	-1.33	0.17	-1.67	-0.99	1.00	3178	5425
Intercept[4]	0.43	0.17	0.10	0.76	1.00	3149	5339
Intercept[5]	1.97	0.18	1.62	2.33	1.00	3487	5720
Intercept[6]	3.79	0.21	3.38	4.22	1.00	4366	7164
f_finv	-1.12	0.20	-1.52	-0.72	1.00	4769	7283
is_finv	-1.02	0.21	-1.43	-0.60	1.00	5140	7377
old_young	0.87	0.34	0.19	1.55	1.00	4282	6242
formal_informal	0.02	0.11	-0.19	0.23	1.00	12684	8930
f_finv:old_young	0.15	0.54	-0.90	1.21	1.00	5453	8002
is_finv:old_young	0.76	0.55	-0.32	1.85	1.00	5566	7878
f_finv:formal_informal	-0.40	0.26	-0.93	0.11	1.00	10944	8690
is_finv:formal_informal	-0.63	0.31	-1.24	-0.03	1.00	8628	8517
old_young:formal_informal	-0.36	0.43	-1.21	0.47	1.00	9073	8600
f_finv:old_young:formal_informal	-0.48	0.79	-2.03	1.05	1.00	10003	8899
is_finv:old_young:formal_informal	-1.17	1.01	-3.18	0.81	1.00	7904	7982

TABLEAU K.31 – Résumé du modèle *brms* (MG2, LOISIRS)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Type : F vs. FINV	-1.12	0.00	1.00	-1.88	-0.31	-1.52	-0.72	-1.45	-0.79	-1.41	-0.83
Type : IS vs. FINV	-1.02	0.00	1.00	-1.91	-0.10	-1.43	-0.60	-1.36	-0.67	-1.32	-0.72
Âge : +30 ans vs. -30 ans	0.87	0.99	0.01	-0.55	2.24	0.19	1.55	0.30	1.43	0.38	1.36
Formalité : Formel vs. Informel	0.02	0.56	0.44	-0.51	0.48	-0.19	0.23	-0.16	0.20	-0.14	0.17
F vs. FINV * Âge	0.15	0.62	0.38	-2.21	2.18	-0.90	1.21	-0.73	1.05	-0.62	0.94
IS vs. FINV * Âge	0.76	0.92	0.08	-1.63	3.12	-0.32	1.85	-0.16	1.67	-0.03	1.55
F vs. FINV * Formalité	-0.40	0.06	0.94	-1.55	0.72	-0.93	0.11	-0.83	0.03	-0.78	-0.03
IS vs. FINV * Formalité	-0.63	0.02	0.98	-1.97	0.65	-1.24	-0.03	-1.14	-0.14	-1.08	-0.20
Âge * Formalité	-0.36	0.19	0.81	-2.31	1.43	-1.21	0.47	-1.07	0.34	-0.98	0.24
F vs. FINV * Âge * Formalité	-0.48	0.27	0.73	-3.78	2.72	-2.03	1.05	-1.76	0.80	-1.61	0.63
IS vs. FINV * Âge * Formalité	-1.17	0.12	0.88	-5.37	2.89	-3.18	0.81	-2.83	0.49	-2.62	0.27

TABLEAU K.32 – Distributions postérieures pour le modèle brms (MG2, LOISIRS)

Annexe L

MG3 : Matériels et modèles statistiques

L.0.1 Écran de consentement (MG3) (5.3.3)



Bonjour et bienvenue !

Titre de l'étude: Expérience de lecture et de jugement

Responsables de l'étude : Gabriel Thiberge et Heather Burnett

Institution de rattachement : Laboratoire de Linguistique Formelle, CNRS, Université de Paris

Chercheurs recueillant les données: Gabriel Thiberge et Simon Duverger

Objectif du projet

Nous vous invitons à participer à une étude analysant le jugement de phrases du français par des locutrices et des locuteurs dont la langue maternelle est le français. Les données recueillies nous permettront de mieux comprendre le traitement des phrases en français.

Pourquoi suis-je invité(e) à participer ?

Nous vous avons demandé de participer parce que nous souhaitons mesurer des jugements faits par des participant(e)s humain(e)s pour analyser quantitativement le type de réponses qui sont données.

Suis-je obligé(e) de participer ?

Cette étude n'est pas obligatoire. La participation est uniquement basée sur le volontariat. Vous pouvez vous retirer de cette étude à n'importe quel instant sans en donner la raison. Vous recevrez à la fin de l'expérience un code aléatoire qui vous permettra de vous retirer à tout moment après l'expérience (dans le mois qui suit). Vous aurez seulement à nous envoyer ce code par mail à expgs2020@gmail.com.

Que va-t-il se passer si je participe ?

Nous vous demanderons de juger des phrases selon deux échelles, puis de renseigner quelques informations sur votre profil. Au total, l'expérience dure environ 25 minutes.

Existe-t-il des risques pour cette étude ?

Il n'y a aucun risque connu pour ce genre d'expérience.

Quels sont les avantages pour moi ?

Vous contribuez à l'avancée de nos connaissances sur le langage, et vous avez toute notre gratitude pour cela. Il n'y a pas de compensation financière pour la passation de cette étude.

Comment sont traitées les données ?

Vos données resteront anonymes. Nous n'établissons aucun lien entre vos données et votre identité propre. Toutes les informations que vous nous transmettez (âge, sexe, langue maternelle, etc) resteront confidentielles ainsi que vos résultats. Les fichiers bruts de résultats dont nous disposerons seront stockés sur des supports informatiques cryptés.

Est-ce que cette recherche sera publiée ?

Les résultats de l'étude seront publiés sous une forme agrégée dans des revues et présentés dans des conférences. Une version condensée et anonyme du fichier de résultats globaux sera déposée sur un site public à destination de la communauté scientifique dans une démarche éthique d'ouverture de la science. Dans tous les cas, vous ne serez jamais identifié(e) ni identifiable.

Coordonnées du chercheur en charge du projet en cas de demande d'informations ou de plaintes.

Heather Burnett : heather.susan.burnett@gmail.com

Gabriel Thiberge : expgs2020@gmail.com

Vous pouvez envoyer un mail à l'adresse expgs2020@gmail.com pour avoir un résumé des résultats ou mieux comprendre les raisons de l'étude.

En participant à cette expérience, je consens aux points suivants :

1. Je suis d'accord pour participer à cette étude.
2. Je confirme que j'ai lu et compris comment mes données seront gardées et traitées.
3. J'ai plus de 18 ans.
4. Ma langue maternelle est le français.

Je comprends et j'accepte ce qui a été mentionné ci-dessus.

– Cliquez ici pour commencer la première partie

FIGURE L.1 – Écran de consentement (MG3)

L.0.2 Consignes (MG3) (5.3.3)


Partie I: Expérience de lecture et de jugement

Merci de participer à cette expérience qui durera environ 25 minutes au total !

Dans cette première partie, vous allez lire attentivement de très courts dialogues et vous devrez juger si la dernière personne qui s'exprime (ce sera toujours la ligne commençant par "B") est plutôt :

1. jeune ou âgée
2. peu éduquée ou très éduquée
3. de la ville ou de la campagne

Vous devrez rendre vos jugements sur trois échelles séparées en faisant glisser un curseur entre les deux possibilités.

Ne passez pas trop de temps sur chaque phrase : il n'y a pas de mauvaise réponse et ce sont vos réactions instinctives qui nous intéressent.

Cette première partie dure environ 20 minutes.

Il n'est pas possible de revenir en arrière au cours de l'expérience. Si vous cliquez sur la flèche "précédent" de votre navigateur, l'expérience ne fonctionnera plus.

Dans la dernière partie, nous vous demandons de renseigner des informations personnelles. Ces informations resteront strictement confidentielles, et ne seront en aucun cas transmises à un tiers. Cela dure environ 5 minutes.

→ [Cliquez ici pour commencer l'entraînement](#)

FIGURE L.2 – Consigne (MG3)

L.0.3 Corrélations entre échelles - Détails (MG3)

	Âge	Études	Origine
Âge	1.00	0.49	-0.01
Études	0.49	1.00	-0.28
Origine	-0.01	-0.28	1.00

TABEAU L.1 – Coefficients de corrélations entre échelles, méthode Spearman (MG3)

	Âge	Études	Origine
Âge	NA	0	0.61
Études	0.00	NA	0.00
Origine	0.61	0	NA

TABEAU L.2 – P-values pour les corrélations entre échelles, méthode Spearman (MG3)

L.0.4 Modèle *brms* et distributions postérieures pour l'échelle ÂGE (5.3.3)

Family : cumulative
Links : mu = logit; disc = identity
Formula : answerA ~ (fesk_finv + is_finv) * old_young * formal_informal + ((fesk_finv + is_finv) * formal_informal | id) + ((fesk_finv + is_finv) * old_young * formal_informal | Item)
Data : DON (Number of observations : 1260)
Samples : 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-5.28	0.32	-5.93	-4.66	1.00	2751	4961
Intercept[2]	-4.54	0.29	-5.09	-3.99	1.00	2504	5170
Intercept[3]	-3.42	0.26	-3.92	-2.93	1.00	2168	4716
Intercept[4]	-2.30	0.24	-2.77	-1.84	1.00	1993	4127
Intercept[5]	-1.74	0.24	-2.21	-1.28	1.00	1946	3927
Intercept[6]	0.40	0.23	-0.06	0.85	1.00	1953	3896
Intercept[7]	1.24	0.23	0.78	1.70	1.00	1989	3957
Intercept[8]	2.47	0.25	1.98	2.96	1.00	2082	4296
Intercept[9]	3.69	0.27	3.17	4.22	1.00	2280	4721
Intercept[10]	4.58	0.30	4.00	5.17	1.00	2702	5174
fesk_finv	-1.20	0.28	-1.76	-0.67	1.00	3023	5479
is_finv	-2.26	0.33	-2.92	-1.63	1.00	2918	4951
old_young	0.75	0.39	-0.02	1.54	1.00	1365	2625
formal_informal	0.06	0.14	-0.23	0.33	1.00	6206	7046
fesk_finv:old_young	-1.11	0.56	-2.21	-0.01	1.00	3008	5172
is_finv:old_young	-0.72	0.66	-2.01	0.58	1.00	3035	4889
fesk_finv:formal_informal	0.07	0.31	-0.55	0.68	1.00	8880	7313
is_finv:formal_informal	0.24	0.29	-0.34	0.82	1.00	9529	9001
old_young:formal_informal	-0.04	0.28	-0.60	0.52	1.00	5772	6396
fesk_finv:old_young:formal_informal	-0.02	0.64	-1.29	1.25	1.00	8485	7304
is_finv:old_young:formal_informal	0.31	0.63	-0.93	1.55	1.00	8575	8957

TABLEAU L.3 – Résumé du modèle *brms* (MG3, ÂGE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Type : FESK vs. FINV	-1.20	0.00	1.00	-2.33	-0.05	-1.76	-0.67	-1.66	-0.75	-1.60	-0.80
Type : IS vs. FINV	-2.26	0.00	1.00	-3.73	-1.00	-2.92	-1.63	-2.81	-1.74	-2.74	-1.79
Âge : +30 ans vs. -30 ans	0.75	0.97	0.03	-0.70	2.55	-0.02	1.54	0.10	1.40	0.19	1.31
Formalité : Formel vs. Informel	0.06	0.68	0.33	-0.70	0.62	-0.23	0.33	-0.18	0.29	-0.14	0.26
FESK vs. FINV * Âge	-1.11	0.02	0.98	-3.55	1.18	-2.21	-0.01	-2.02	-0.21	-1.91	-0.32
IS vs. FINV * Âge	-0.72	0.13	0.87	-3.76	1.64	-2.01	0.58	-1.79	0.35	-1.65	0.22
FESK vs. FINV * Formalité	0.07	0.60	0.40	-1.21	1.26	-0.55	0.68	-0.45	0.58	-0.38	0.51
IS vs. FINV * Formalité	0.24	0.80	0.20	-1.01	1.44	-0.34	0.82	-0.25	0.72	-0.18	0.66
Âge * Formalité	-0.04	0.45	0.55	-1.18	1.06	-0.60	0.52	-0.50	0.42	-0.45	0.37
FESK vs. FINV * Âge * Formalité	-0.02	0.49	0.51	-2.45	2.64	-1.29	1.25	-1.10	1.04	-0.96	0.90
IS vs. FINV * Âge * Formalité	0.31	0.69	0.31	-2.47	3.40	-0.93	1.55	-0.72	1.35	-0.58	1.21

TABLEAU L.4 – Distributions postérieures pour le modèle brms (MG3, ÂGE)

L.0.5 Modèle *brms* et distributions postérieures pour l'échelle ÉTUDES (5.3.3)

Family : cumulative

Links : mu = logit; disc = identity

Formula : answerB ~ (fesk_finv + is_finv) * old_young * formal_informal + ((fesk_finv + is_finv) * formal_informal | id) + ((fesk_finv + is_finv) * old_young * formal_informal | Item)

Data : DON (Number of observations : 1260)

Samples : 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-6.13	0.37	-6.87	-5.42	1.00	3376	5447
Intercept[2]	-5.50	0.32	-6.14	-4.89	1.00	3212	5276
Intercept[3]	-4.57	0.27	-5.11	-4.05	1.00	2773	4987
Intercept[4]	-3.48	0.24	-3.95	-3.01	1.00	2399	4009
Intercept[5]	-2.43	0.22	-2.87	-1.99	1.00	2261	4184
Intercept[6]	0.31	0.21	-0.10	0.73	1.00	2124	4342
Intercept[7]	1.31	0.21	0.90	1.73	1.00	2161	4398
Intercept[8]	2.60	0.23	2.16	3.04	1.00	2373	4565
Intercept[9]	3.78	0.25	3.29	4.28	1.00	2798	4335
Intercept[10]	5.04	0.30	4.48	5.64	1.00	3357	5703
fesk_finv	-1.82	0.34	-2.49	-1.16	1.00	2150	4197
is_finv	-2.77	0.40	-3.57	-1.99	1.00	2415	4445
old_young	0.23	0.36	-0.48	0.93	1.00	1594	2873
formal_informal	-0.01	0.15	-0.31	0.29	1.00	5955	6873
fesk_finv:old_young	-1.24	0.66	-2.54	0.05	1.00	2262	3332
is_finv:old_young	-1.31	0.78	-2.85	0.25	1.00	2328	4275
fesk_finv:formal_informal	0.00	0.30	-0.60	0.60	1.00	7911	8164
is_finv:formal_informal	0.13	0.32	-0.49	0.75	1.00	7083	8062
old_young:formal_informal	-0.02	0.30	-0.63	0.55	1.00	6160	6917
fesk_finv:old_young:formal_informal	-0.74	0.60	-1.94	0.41	1.00	7568	8609
is_finv:old_young:formal_informal	-0.29	0.65	-1.56	0.97	1.00	7626	7575

TABLEAU L.5 – Résumé du modèle *brms* (MG3, ÉTUDES)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Type : FESK vs. FINV	-1.82	0.00	1.00	-3.46	-0.58	-2.49	-1.16	-2.38	-1.26	-2.31	-1.33
Type : IS vs. FINV	-2.77	0.00	1.00	-4.31	-0.83	-3.57	-1.99	-3.44	-2.12	-3.34	-2.21
Âge : +30 ans vs. -30 ans	0.23	0.74	0.26	-1.01	1.76	-0.48	0.93	-0.36	0.83	-0.28	0.75
Formalité : Formel vs. Informel	-0.01	0.48	0.52	-0.70	0.65	-0.31	0.29	-0.26	0.24	-0.23	0.21
FESK vs. FINV * Âge	-1.24	0.03	0.97	-4.30	1.17	-2.54	0.05	-2.33	-0.15	-2.20	-0.30
IS vs. FINV * Âge	-1.31	0.05	0.95	-4.52	1.75	-2.85	0.25	-2.59	-0.02	-2.41	-0.19
FESK vs. FINV * Formalité	0.00	0.51	0.49	-1.33	1.66	-0.60	0.60	-0.50	0.50	-0.43	0.44
IS vs. FINV * Formalité	0.13	0.66	0.34	-1.30	1.34	-0.49	0.75	-0.39	0.64	-0.32	0.58
Âge * Formalité	-0.02	0.47	0.53	-1.12	1.16	-0.63	0.55	-0.52	0.46	-0.46	0.40
FESK vs. FINV * Âge * Formalité	-0.74	0.11	0.89	-3.40	1.71	-1.94	0.41	-1.73	0.23	-1.62	0.11
IS vs. FINV * Âge * Formalité	-0.29	0.32	0.68	-3.29	2.18	-1.56	0.97	-1.36	0.76	-1.23	0.62

TABLEAU L.6 – Distributions postérieures pour le modèle brms (MG3, ÉTUDES)

L.0.6 Modèle *brms* et distributions postérieures pour l'échelle ORIGINE (5.3.3)

Family : cumulative
Links : mu = logit; disc = identity
Formula : answerC ~ (fesk_finv + is_finv) * old_young * formal_informal + ((fesk_finv + is_finv) * formal_informal | id) + ((fesk_finv + is_finv) * old_young * formal_informal | Item)
Data : DON (Number of observations : 1260)
Samples : 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-4.75	0.27	-5.30	-4.23	1.00	4857	6678
Intercept[2]	-3.84	0.23	-4.31	-3.39	1.00	4644	6713
Intercept[3]	-2.66	0.20	-3.05	-2.28	1.00	4236	6824
Intercept[4]	-1.88	0.19	-2.25	-1.53	1.00	4009	6105
Intercept[5]	-1.23	0.18	-1.59	-0.89	1.00	3907	6531
Intercept[6]	2.66	0.19	2.28	3.05	1.00	4355	6932
Intercept[7]	3.12	0.21	2.73	3.53	1.00	4637	7391
Intercept[8]	3.86	0.23	3.43	4.33	1.00	5416	8134
Intercept[9]	4.32	0.25	3.84	4.83	1.00	6203	8389
Intercept[10]	4.90	0.29	4.34	5.48	1.00	7329	8021
fesk_finv	0.35	0.31	-0.27	0.96	1.00	3195	5980
is_finv	0.57	0.34	-0.10	1.23	1.00	3071	5640
old_young	0.05	0.30	-0.54	0.65	1.00	3076	5637
formal_informal	0.12	0.14	-0.16	0.40	1.00	11039	9295
fesk_finv:old_young	0.31	0.62	-0.95	1.52	1.00	3301	5785
is_finv:old_young	-0.39	0.67	-1.75	0.93	1.00	3399	5233
fesk_finv:formal_informal	-0.52	0.36	-1.26	0.18	1.00	9601	8888
is_finv:formal_informal	-0.18	0.34	-0.83	0.50	1.00	10883	8850
old_young:formal_informal	0.23	0.28	-0.31	0.78	1.00	12132	8524
fesk_finv:old_young:formal_informal	-0.74	0.72	-2.20	0.67	1.00	9388	8488
is_finv:old_young:formal_informal	0.09	0.68	-1.24	1.45	1.00	10058	8814

TABLEAU L.7 – Résumé du modèle *brms* (MG3, ORIGINE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Type : FESK vs. FINV	0.35	0.86	0.14	-1.08	1.65	-0.27	0.96	-0.17	0.86	-0.10	0.80
Type : IS vs. FINV	0.57	0.95	0.05	-1.06	2.00	-0.10	1.23	0.01	1.13	0.08	1.05
Âge : +30 ans vs. -30 ans	0.05	0.57	0.43	-1.34	1.12	-0.54	0.65	-0.44	0.55	-0.38	0.49
Formalité : Formel vs. Informel	0.12	0.80	0.20	-0.41	0.80	-0.16	0.40	-0.11	0.36	-0.08	0.32
FESK vs. FINV * Âge	0.31	0.69	0.31	-2.62	2.66	-0.95	1.52	-0.72	1.33	-0.58	1.20
IS vs. FINV * Âge	-0.39	0.28	0.72	-3.17	2.37	-1.75	0.93	-1.51	0.70	-1.36	0.57
FESK vs. FINV * Formalité	-0.52	0.07	0.93	-1.97	0.87	-1.26	0.18	-1.12	0.06	-1.04	-0.01
IS vs. FINV * Formalité	-0.18	0.30	0.70	-1.50	1.29	-0.83	0.50	-0.73	0.39	-0.66	0.32
Âge * Formalité	0.23	0.80	0.20	-0.96	1.43	-0.31	0.78	-0.22	0.68	-0.16	0.63
FESK vs. FINV * Âge * Formalité	-0.74	0.15	0.85	-4.15	2.25	-2.20	0.67	-1.95	0.43	-1.79	0.28
IS vs. FINV * Âge * Formalité	0.09	0.55	0.45	-2.60	2.91	-1.24	1.45	-1.04	1.23	-0.89	1.08

TABLEAU L.8 – Distributions postérieures pour le modèle brms (MG3, ORIGINE)

Annexe M

MG enfants : Matériel et modèles statistiques

M.0.1 Répartition des passations de la tâche par liste (MG-C, 6.2)

Classe	Liste 1	Liste 2	Liste 3	Liste 4	Liste 5	Liste 6
GS-1	2	2	1	1	1	2
GS-2	1	2	2	1	2	1
CP	3	3	2	3	4	4
CPCE1	2	2	4	4	3	3
CE1CE2	5	4	3	3	3	3
CE2CM1	3	4	5	3	2	3
CM1CM2-A	3	4	3	5	4	4
CM1CM2-B	3	2	3	3	3	3

TABLEAU M.1 – Nombre de passations par liste par classe (MG-C)

M.0.2 Items expérimentaux par liste (MG-C, 6.2)

Item	Voix	Stéréotype	Phrase
101	FEM1	Richesse	Il est trop beau!
102	FEM2	Études	Moi je vais manger avec ma mère.
050	HOM1	Loisirs	C'est Léon que j'ai couché à dix-neuf heures.
056	HOM2	Richesse	Les fenêtres, ils doit les faire samedi.

TABLEAU M.2 – Item d'entraînement pour toutes les listes (MG-C)

Item	Voix	Bloc	Condition	Stéréotype	Phrase
01	FEM1	b1	f	Richesse	À quelle heure il arrive?
02	FEM1	b1	is	Loisirs	Tu les as mises dans lequel?
03	FEM1	b1	finv	Éducation	Tu l'as cassé comment?
04	FEM1	b1	f	Loisirs	Où tu vas?
06	FEM1	b1	is	Éducation	C'est qui?
07	FEM2	b1	finv	Richesse	Quand saura-t-on?
09	FEM2	b1	f	Éducation	Lequel tu as pris?
10	FEM2	b1	is	Richesse	Tu le ferais avec quoi?
11	FEM2	b1	finv	Loisirs	Quand est-elle partie?
13	HOM1	b1	f	Richesse	Où tu l'as utilisée?
17	HOM1	b1	is	Loisirs	Qui connaîtra-t-elle?
19	HOM1	b1	finv	Éducation	Avec qui vous êtes-vous battus?
21	HOM1	b1	f	Loisirs	Qui tu as rencontré?
25	HOM1	b1	is	Éducation	Il a pris lequel?
26	HOM2	b1	finv	Richesse	De qui parles-tu?
28	HOM2	b1	f	Éducation	Pourquoi tu dis ça?
29	HOM2	b1	is	Richesse	Il revient quand?
30	HOM2	b1	finv	Loisirs	À quoi jouez-vous?
01	FEM1	b2	finv	Éducation	À quelle heure arrive-t-il?
02	FEM1	b2	f	Richesse	Dans lequel tu les as mises?
03	FEM1	b2	is	Loisirs	Comment l'as tu cassé?
04	FEM1	b2	finv	Richesse	Où vas-tu?
06	FEM1	b2	f	Loisirs	Qui c'est?
07	FEM2	b2	is	Éducation	On saura quand?
09	FEM2	b2	finv	Loisirs	Lequel as-tu pris?
10	FEM2	b2	f	Éducation	Avec quoi tu le ferais?
11	FEM2	b2	is	Richesse	Elle est partie quand?
13	HOM1	b2	finv	Éducation	Où l'as-tu utilisée?
17	HOM1	b2	f	Richesse	Qui elle connaîtra?
19	HOM1	b2	is	Loisirs	Vous vous êtes battus avec qui?
21	HOM1	b2	finv	Richesse	Qui as-tu rencontré?
25	HOM1	b2	f	Loisirs	Lequel il a pris?
26	HOM2	b2	is	Éducation	Tu parles de qui?
28	HOM2	b2	finv	Loisirs	Pourquoi dis-tu ça?
29	HOM2	b2	f	Éducation	Quand il revient?
30	HOM2	b2	is	Richesse	Vous jouez à quoi?
01	FEM1	b3	is	Loisirs	Il arrive à quelle heure?
02	FEM1	b3	finv	Éducation	Dans lequel les as-tu mises?
03	FEM1	b3	f	Richesse	Comment tu l'as cassé?
04	FEM1	b3	is	Éducation	Tu vas où?
06	FEM1	b3	finv	Richesse	Qui est-ce?
07	FEM2	b3	f	Loisirs	Quand on saura?
09	FEM2	b3	is	Richesse	Tu as pris lequel?
10	FEM2	b3	finv	Loisirs	Avec quoi le ferais-tu?
11	FEM2	b3	f	Éducation	Quand elle est partie?
13	HOM1	b3	is	Loisirs	Tu l'as utilisée où?
17	HOM1	b3	finv	Éducation	Elle connaîtra qui?
19	HOM1	b3	f	Richesse	Avec qui vous vous êtes battus?
21	HOM1	b3	is	Éducation	Tu as rencontré qui?
25	HOM1	b3	finv	Richesse	Lequel a-t-il pris?
26	HOM2	b3	f	Loisirs	De qui tu parles?
28	HOM2	b3	is	Richesse	Tu dis ça pourquoi?
29	HOM2	b3	finv	Loisirs	Quand revient-il?
30	HOM2	b3	f	Éducation	À quoi vous jouez?

TABLEAU M.3 – Items expérimentaux - Liste 1 (MG-C)

Item	Voix	Bloc	Condition	Stéréotype	Phrase
01	FEM1	b1	finv	Richesse	À quelle heure arrive-t-il?
02	FEM1	b1	f	Loisirs	Dans lequel tu les as mises?
03	FEM1	b1	is	Éducation	Tu l'as cassé comment?
04	FEM1	b1	finv	Loisirs	Où vas-tu?
06	FEM1	b1	f	Éducation	Qui c'est?
07	FEM2	b1	is	Richesse	On saura quand?
09	FEM2	b1	finv	Éducation	Lequel as-tu pris?
10	FEM2	b1	f	Richesse	Avec quoi tu le ferais?
11	FEM2	b1	is	Loisirs	Elle est partie quand?
13	HOM1	b1	finv	Richesse	Où l'as-tu utilisée?
17	HOM1	b1	f	Loisirs	Qui elle connaîtra?
19	HOM1	b1	is	Éducation	Vous vous êtes battus avec qui?
21	HOM1	b1	finv	Loisirs	Qui as-tu rencontré?
25	HOM1	b1	f	Éducation	Lequel il a pris?
26	HOM2	b1	is	Richesse	Tu parles de qui?
28	HOM2	b1	finv	Éducation	Pourquoi dis-tu ça?
29	HOM2	b1	f	Richesse	Quand il revient?
30	HOM2	b1	is	Loisirs	Vous jouez à quoi?
01	FEM1	b2	is	Éducation	Il arrive à quelle heure?
02	FEM1	b2	finv	Richesse	Dans lequel les as-tu mises?
03	FEM1	b2	f	Loisirs	Comment tu l'as cassé?
04	FEM1	b2	is	Richesse	Tu vas où?
06	FEM1	b2	finv	Loisirs	Qui est-ce?
07	FEM2	b2	f	Éducation	Quand on saura?
09	FEM2	b2	is	Loisirs	Tu as pris lequel?
10	FEM2	b2	finv	Éducation	Avec quoi le ferais-tu?
11	FEM2	b2	f	Richesse	Quand elle est partie?
13	HOM1	b2	is	Éducation	Tu l'as utilisée où?
17	HOM1	b2	finv	Richesse	Qui connaîtra-t-elle?
19	HOM1	b2	f	Loisirs	Avec qui vous vous êtes battus?
21	HOM1	b2	is	Richesse	Tu as rencontré qui?
25	HOM1	b2	finv	Loisirs	Lequel a-t-il pris?
26	HOM2	b2	f	Éducation	De qui tu parles?
28	HOM2	b2	is	Loisirs	Tu dis ça pourquoi?
29	HOM2	b2	finv	Éducation	Quand revient-il?
30	HOM2	b2	f	Richesse	À quoi vous jouez?
01	FEM1	b3	f	Loisirs	À quelle heure il arrive?
02	FEM1	b3	is	Éducation	Tu les as mises dans lequel?
03	FEM1	b3	finv	Richesse	Comment l'as tu cassé?
04	FEM1	b3	f	Éducation	Où tu vas?
06	FEM1	b3	is	Richesse	C'est qui?
07	FEM2	b3	finv	Loisirs	Quand saura-t-on?
09	FEM2	b3	f	Richesse	Lequel tu as pris?
10	FEM2	b3	is	Loisirs	Tu le ferais avec quoi?
11	FEM2	b3	finv	Éducation	Quand est-elle partie?
13	HOM1	b3	f	Loisirs	Où tu l'as utilisée?
17	HOM1	b3	is	Éducation	Elle connaîtra qui?
19	HOM1	b3	finv	Richesse	Avec qui vous êtes-vous battus?
21	HOM1	b3	f	Éducation	Qui tu as rencontré?
25	HOM1	b3	is	Richesse	Il a pris lequel?
26	HOM2	b3	finv	Loisirs	De qui parles-tu?
28	HOM2	b3	f	Richesse	Pourquoi tu dis ça?
29	HOM2	b3	is	Loisirs	Il revient quand?
30	HOM2	b3	finv	Éducation	À quoi jouez-vous?

TABLEAU M.4 – Items expérimentaux - Liste 2 (MG-C)

Item	Voix	Bloc	Condition	Stéréotype	Phrase
01	FEM1	b1	is	Richesse	Il arrive à quelle heure?
02	FEM1	b1	finv	Loisirs	Dans lequel les as-tu mises?
03	FEM1	b1	f	Éducation	Comment tu l'as cassé?
04	FEM1	b1	is	Loisirs	Tu vas où?
06	FEM1	b1	finv	Éducation	Qui est-ce?
07	FEM2	b1	f	Richesse	Quand on saura?
09	FEM2	b1	is	Éducation	Tu as pris lequel?
10	FEM2	b1	finv	Richesse	Avec quoi le ferais-tu?
11	FEM2	b1	f	Loisirs	Quand elle est partie?
13	HOM1	b1	is	Richesse	Tu l'as utilisée où?
17	HOM1	b1	finv	Loisirs	Qui connaîtra-t-elle?
19	HOM1	b1	f	Éducation	Avec qui vous vous êtes battus?
21	HOM1	b1	is	Loisirs	Tu as rencontré qui?
25	HOM1	b1	finv	Éducation	Lequel a-t-il pris?
26	HOM2	b1	f	Richesse	De qui tu parles?
28	HOM2	b1	is	Éducation	Tu dis ça pourquoi?
29	HOM2	b1	finv	Richesse	Quand revient-il?
30	HOM2	b1	f	Loisirs	À quoi vous jouez?
01	FEM1	b2	f	Éducation	À quelle heure il arrive?
02	FEM1	b2	is	Richesse	Tu les as mises dans lequel?
03	FEM1	b2	finv	Loisirs	Comment l'as tu cassé?
04	FEM1	b2	f	Richesse	Où tu vas?
06	FEM1	b2	is	Loisirs	C'est qui?
07	FEM2	b2	finv	Éducation	Quand saura-t-on?
09	FEM2	b2	f	Loisirs	Lequel tu as pris?
10	FEM2	b2	is	Éducation	Tu le ferais avec quoi?
11	FEM2	b2	finv	Richesse	Quand est-elle partie?
13	HOM1	b2	f	Éducation	Où tu l'as utilisée?
17	HOM1	b2	is	Richesse	Elle connaîtra qui?
19	HOM1	b2	finv	Loisirs	Avec qui vous êtes-vous battus?
21	HOM1	b2	f	Richesse	Qui tu as rencontré?
25	HOM1	b2	is	Loisirs	Il a pris lequel?
26	HOM2	b2	finv	Éducation	De qui parles-tu?
28	HOM2	b2	f	Loisirs	Pourquoi tu dis ça?
29	HOM2	b2	is	Éducation	Il revient quand?
30	HOM2	b2	finv	Richesse	À quoi jouez-vous?
01	FEM1	b3	finv	Loisirs	À quelle heure arrive-t-il?
02	FEM1	b3	f	Éducation	Dans lequel tu les as mises?
03	FEM1	b3	is	Richesse	Tu l'as cassé comment?
04	FEM1	b3	finv	Éducation	Où vas-tu?
06	FEM1	b3	f	Richesse	Qui c'est?
07	FEM2	b3	is	Loisirs	On saura quand?
09	FEM2	b3	finv	Richesse	Lequel as-tu pris?
10	FEM2	b3	f	Loisirs	Avec quoi tu le ferais?
11	FEM2	b3	is	Éducation	Elle est partie quand?
13	HOM1	b3	finv	Loisirs	Où l'as-tu utilisée?
17	HOM1	b3	f	Éducation	Qui elle connaîtra?
19	HOM1	b3	is	Richesse	Vous vous êtes battus avec qui?
21	HOM1	b3	finv	Éducation	Qui as-tu rencontré?
25	HOM1	b3	f	Richesse	Lequel il a pris?
26	HOM2	b3	is	Loisirs	Tu parles de qui?
28	HOM2	b3	finv	Richesse	Pourquoi dis-tu ça?
29	HOM2	b3	f	Loisirs	Quand il revient?
30	HOM2	b3	is	Éducation	Vous jouez à quoi?

TABLEAU M.5 – Items expérimentaux - Liste 3 (MG-C)

Item	Voix	Bloc	Condition	Stéréotype	Phrase
01	HOM1	b1	f	Richesse	À quelle heure il arrive?
02	HOM1	b1	is	Loisirs	Tu les as mises dans lequel?
03	HOM1	b1	finv	Éducation	Tu l'as cassé comment?
04	HOM1	b1	f	Loisirs	Où tu vas?
06	HOM1	b1	is	Éducation	C'est qui?
07	HOM2	b1	finv	Richesse	Quand saura-t-on?
09	HOM2	b1	f	Éducation	Lequel tu as pris?
10	HOM2	b1	is	Richesse	Tu le ferais avec quoi?
11	HOM2	b1	finv	Loisirs	Quand est-elle partie?
13	FEM1	b1	f	Richesse	Où tu l'as utilisée?
17	FEM1	b1	is	Loisirs	Qui connaîtra-t-elle?
19	FEM1	b1	finv	Éducation	Avec qui vous êtes-vous battus?
21	FEM1	b1	f	Loisirs	Qui tu as rencontré?
25	FEM1	b1	is	Éducation	Il a pris lequel?
26	FEM2	b1	finv	Richesse	De qui parles-tu?
28	FEM2	b1	f	Éducation	Pourquoi tu dis ça?
29	FEM2	b1	is	Richesse	Il revient quand?
30	FEM2	b1	finv	Loisirs	À quoi jouez-vous?
01	HOM1	b2	finv	Éducation	À quelle heure arrive-t-il?
02	HOM1	b2	f	Richesse	Dans lequel tu les as mises?
03	HOM1	b2	is	Loisirs	Comment l'as tu cassé?
04	HOM1	b2	finv	Richesse	Où vas-tu?
06	HOM1	b2	f	Loisirs	Qui c'est?
07	HOM2	b2	is	Éducation	On saura quand?
09	HOM2	b2	finv	Loisirs	Lequel as-tu pris?
10	HOM2	b2	f	Éducation	Avec quoi tu le ferais?
11	HOM2	b2	is	Richesse	Elle est partie quand?
13	FEM1	b2	finv	Éducation	Où l'as-tu utilisée?
17	FEM1	b2	f	Richesse	Qui elle connaîtra?
19	FEM1	b2	is	Loisirs	Vous vous êtes battus avec qui?
21	FEM1	b2	finv	Richesse	Qui as-tu rencontré?
25	FEM1	b2	f	Loisirs	Lequel il a pris?
26	FEM2	b2	is	Éducation	Tu parles de qui?
28	FEM2	b2	finv	Loisirs	Pourquoi dis-tu ça?
29	FEM2	b2	f	Éducation	Quand il revient?
30	FEM2	b2	is	Richesse	Vous jouez à quoi?
01	HOM1	b3	is	Loisirs	Il arrive à quelle heure?
02	HOM1	b3	finv	Éducation	Dans lequel les as-tu mises?
03	HOM1	b3	f	Richesse	Comment tu l'as cassé?
04	HOM1	b3	is	Éducation	Tu vas où?
06	HOM1	b3	finv	Richesse	Qui est-ce?
07	HOM2	b3	f	Loisirs	Quand on saura?
09	HOM2	b3	is	Richesse	Tu as pris lequel?
10	HOM2	b3	finv	Loisirs	Avec quoi le ferais-tu?
11	HOM2	b3	f	Éducation	Quand elle est partie?
13	FEM1	b3	is	Loisirs	Tu l'as utilisée où?
17	FEM1	b3	finv	Éducation	Elle connaîtra qui?
19	FEM1	b3	f	Richesse	Avec qui vous vous êtes battus?
21	FEM1	b3	is	Éducation	Tu as rencontré qui?
25	FEM1	b3	finv	Richesse	Lequel a-t-il pris?
26	FEM2	b3	f	Loisirs	De qui tu parles?
28	FEM2	b3	is	Richesse	Tu dis ça pourquoi?
29	FEM2	b3	finv	Loisirs	Quand revient-il?
30	FEM2	b3	f	Éducation	À quoi vous jouez?

TABLEAU M.6 – Items expérimentaux - Liste 4 (MG-C)

Item	Voix	Bloc	Condition	Stéréotype	Phrase
01	HOM1	b1	finv	Richesse	À quelle heure arrive-t-il?
02	HOM1	b1	f	Loisirs	Dans lequel tu les as mises?
03	HOM1	b1	is	Éducation	Tu l'as cassé comment?
04	HOM1	b1	finv	Loisirs	Où vas-tu?
06	HOM1	b1	f	Éducation	Qui c'est?
07	HOM2	b1	is	Richesse	On saura quand?
09	HOM2	b1	finv	Éducation	Lequel as-tu pris?
10	HOM2	b1	f	Richesse	Avec quoi tu le ferais?
11	HOM2	b1	is	Loisirs	Elle est partie quand?
13	FEM1	b1	finv	Richesse	Où l'as-tu utilisée?
17	FEM1	b1	f	Loisirs	Qui elle connaîtra?
19	FEM1	b1	is	Éducation	Vous vous êtes battus avec qui?
21	FEM1	b1	finv	Loisirs	Qui as-tu rencontré?
25	FEM1	b1	f	Éducation	Lequel il a pris?
26	FEM2	b1	is	Richesse	Tu parles de qui?
28	FEM2	b1	finv	Éducation	Pourquoi dis-tu ça?
29	FEM2	b1	f	Richesse	Quand il revient?
30	FEM2	b1	is	Loisirs	Vous jouez à quoi?
01	HOM1	b2	is	Éducation	Il arrive à quelle heure?
02	HOM1	b2	finv	Richesse	Dans lequel les as-tu mises?
03	HOM1	b2	f	Loisirs	Comment tu l'as cassé?
04	HOM1	b2	is	Richesse	Tu vas où?
06	HOM1	b2	finv	Loisirs	Qui est-ce?
07	HOM2	b2	f	Éducation	Quand on saura?
09	HOM2	b2	is	Loisirs	Tu as pris lequel?
10	HOM2	b2	finv	Éducation	Avec quoi le ferais-tu?
11	HOM2	b2	f	Richesse	Quand elle est partie?
13	FEM1	b2	is	Éducation	Tu l'as utilisée où?
17	FEM1	b2	finv	Richesse	Qui connaîtra-t-elle?
19	FEM1	b2	f	Loisirs	Avec qui vous vous êtes battus?
21	FEM1	b2	is	Richesse	Tu as rencontré qui?
25	FEM1	b2	finv	Loisirs	Lequel a-t-il pris?
26	FEM2	b2	f	Éducation	De qui tu parles?
28	FEM2	b2	is	Loisirs	Tu dis ça pourquoi?
29	FEM2	b2	finv	Éducation	Quand revient-il?
30	FEM2	b2	f	Richesse	À quoi vous jouez?
01	HOM1	b3	f	Loisirs	À quelle heure il arrive?
02	HOM1	b3	is	Éducation	Tu les as mises dans lequel?
03	HOM1	b3	finv	Richesse	Comment l'as tu cassé?
04	HOM1	b3	f	Éducation	Où tu vas?
06	HOM1	b3	is	Richesse	C'est qui?
07	HOM2	b3	finv	Loisirs	Quand saura-t-on?
09	HOM2	b3	f	Richesse	Lequel tu as pris?
10	HOM2	b3	is	Loisirs	Tu le ferais avec quoi?
11	HOM2	b3	finv	Éducation	Quand est-elle partie?
13	FEM1	b3	f	Loisirs	Où tu l'as utilisée?
17	FEM1	b3	is	Éducation	Elle connaîtra qui?
19	FEM1	b3	finv	Richesse	Avec qui vous êtes-vous battus?
21	FEM1	b3	f	Éducation	Qui tu as rencontré?
25	FEM1	b3	is	Richesse	Il a pris lequel?
26	FEM2	b3	finv	Loisirs	De qui parles-tu?
28	FEM2	b3	f	Richesse	Pourquoi tu dis ça?
29	FEM2	b3	is	Loisirs	Il revient quand?
30	FEM2	b3	finv	Éducation	À quoi jouez-vous?

TABLEAU M.7 – Items expérimentaux - Liste 5 (MG-C)

Item	Voix	Bloc	Condition	Stéréotype	Phrase
01	HOM1	b1	is	Richesse	Il arrive à quelle heure?
02	HOM1	b1	finv	Loisirs	Dans lequel les as-tu mises?
03	HOM1	b1	f	Éducation	Comment tu l'as cassé?
04	HOM1	b1	is	Loisirs	Tu vas où?
06	HOM1	b1	finv	Éducation	Qui est-ce?
07	HOM2	b1	f	Richesse	Quand on saura?
09	HOM2	b1	is	Éducation	Tu as pris lequel?
10	HOM2	b1	finv	Richesse	Avec quoi le ferais-tu?
11	HOM2	b1	f	Loisirs	Quand elle est partie?
13	FEM1	b1	is	Richesse	Tu l'as utilisée où?
17	FEM1	b1	finv	Loisirs	Qui connaîtra-t-elle?
19	FEM1	b1	f	Éducation	Avec qui vous vous êtes battus?
21	FEM1	b1	is	Loisirs	Tu as rencontré qui?
25	FEM1	b1	finv	Éducation	Lequel a-t-il pris?
26	FEM2	b1	f	Richesse	De qui tu parles?
28	FEM2	b1	is	Éducation	Tu dis ça pourquoi?
29	FEM2	b1	finv	Richesse	Quand revient-il?
30	FEM2	b1	f	Loisirs	À quoi vous jouez?
01	HOM1	b2	f	Éducation	À quelle heure il arrive?
02	HOM1	b2	is	Richesse	Tu les as mises dans lequel?
03	HOM1	b2	finv	Loisirs	Comment l'as tu cassé?
04	HOM1	b2	f	Richesse	Où tu vas?
06	HOM1	b2	is	Loisirs	C'est qui?
07	HOM2	b2	finv	Éducation	Quand saura-t-on?
09	HOM2	b2	f	Loisirs	Lequel tu as pris?
10	HOM2	b2	is	Éducation	Tu le ferais avec quoi?
11	HOM2	b2	finv	Richesse	Quand est-elle partie?
13	FEM1	b2	f	Éducation	Où tu l'as utilisée?
17	FEM1	b2	is	Richesse	Elle connaîtra qui?
19	FEM1	b2	finv	Loisirs	Avec qui vous êtes-vous battus?
21	FEM1	b2	f	Richesse	Qui tu as rencontré?
25	FEM1	b2	is	Loisirs	Il a pris lequel?
26	FEM2	b2	finv	Éducation	De qui parles-tu?
28	FEM2	b2	f	Loisirs	Pourquoi tu dis ça?
29	FEM2	b2	is	Éducation	Il revient quand?
30	FEM2	b2	finv	Richesse	À quoi jouez-vous?
01	HOM1	b3	finv	Loisirs	À quelle heure arrive-t-il?
02	HOM1	b3	f	Éducation	Dans lequel tu les as mises?
03	HOM1	b3	is	Richesse	Tu l'as cassé comment?
04	HOM1	b3	finv	Éducation	Où vas-tu?
06	HOM1	b3	f	Richesse	Qui c'est?
07	HOM2	b3	is	Loisirs	On saura quand?
09	HOM2	b3	finv	Richesse	Lequel as-tu pris?
10	HOM2	b3	f	Loisirs	Avec quoi tu le ferais?
11	HOM2	b3	is	Éducation	Elle est partie quand?
13	FEM1	b3	finv	Loisirs	Où l'as-tu utilisée?
17	FEM1	b3	f	Éducation	Qui elle connaîtra?
19	FEM1	b3	is	Richesse	Vous vous êtes battus avec qui?
21	FEM1	b3	finv	Éducation	Qui as-tu rencontré?
25	FEM1	b3	f	Richesse	Lequel il a pris?
26	FEM2	b3	is	Loisirs	Tu parles de qui?
28	FEM2	b3	finv	Richesse	Pourquoi dis-tu ça?
29	FEM2	b3	f	Loisirs	Quand il revient?
30	FEM2	b3	is	Éducation	Vous jouez à quoi?

TABLEAU M.8 – Items expérimentaux - Liste 6 (MG-C)

M.0.3 Choix entre stéréotypes présentés aux enfants (MG-C, 6.2)



FIGURE M.1 – Échelle de l'étude MG-C. Stéréotype : RICHESSE

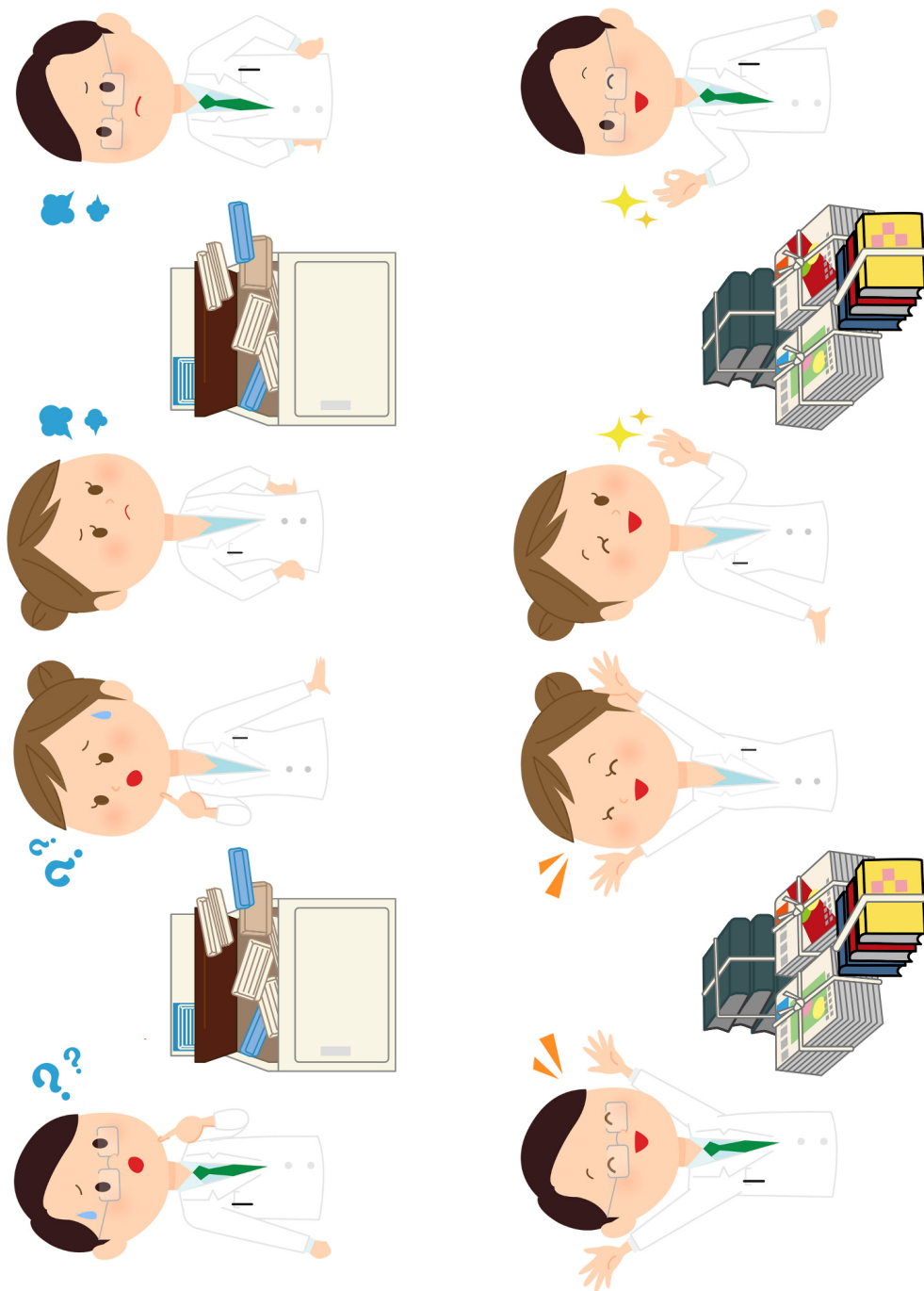


FIGURE M.2 – Échelle de l'étude MG-C. Stéréotype : ÉDUCATION



FIGURE M.3 – Échelle de l'étude MG-C. Stéréotype : LOISIRS

M.0.4 Écrans de pause et final pour le jeu présenté aux enfants (MG-C, 6.2)



NB : De gauche à droite et de bas en haut : stades 1, 2, 3 puis 4.

FIGURE M.4 – Écrans de progression (Étude MG-C.)

M.0.5 Diplôme donné aux enfants après l'expérience (MG-C, 6.2)



FIGURE M.5 – Diplôme de participation à l'étude MG-C

M.0.6 Modèle *brms* et distributions postérieures pour un effet de TYPE (MG-C, 6.2.4)

Family : bernoulli
Links : mu = logit
Formula : answernum ~ (f_finv + is_finv) + (f_finv + is_finv | id) + (f_finv + is_finv | item)
Data : DON (Number of observations : 5275)
Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	-0.10	0.06	-0.21	0.02	1.00	5531	4416
f_finv	-0.08	0.08	-0.23	0.08	1.00	9475	4759
is_finv	-0.12	0.08	-0.27	0.03	1.00	10699	4906

TABLEAU M.9 – Résumé du modèle *brms* (MG-C, effet TYPE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l85	h85
Antéposition simple	-0.08	0.16	0.84	-0.41	0.23	-0.23	0.08	-0.19	0.03
In situ	-0.12	0.06	0.94	-0.42	0.17	-0.27	0.03	-0.23	-0.01

TABLEAU M.10 – Distributions postérieures pour le modèle *brms* (MG-C, effet TYPE)

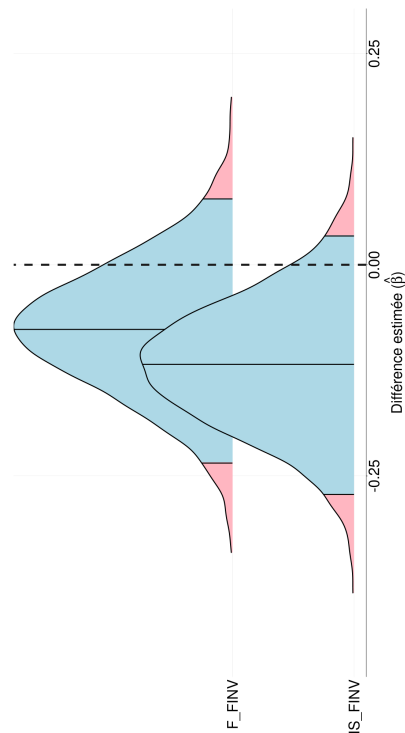


FIGURE M.6 – Distributions postérieures pour le modèle *brms* (MG-C, effet TYPE)

M.0.7 Modèle *brms* et distributions postérieures pour un effet de TYPE sur le stéréotype RICHESSE (MG-C, 6.2.4)

Family : bernoulli

Links : mu = logit

Formula : answernum ~ (f_finv + is_finv) + (f_finv + is_finv | id) + (f_finv + is_finv | item)

Data : subset(DON, stere=="r") (Number of observations : 1786)

Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	0.00	0.09	-0.18	0.18	1.00	4854	4535
f_finv	-0.01	0.14	-0.29	0.28	1.00	7723	4596
is_finv	0.12	0.14	-0.14	0.39	1.00	8077	4683

TABLEAU M.11 – Résumé du modèle *brms* (MG-C, stéréotype RICHESSE, effet TYPE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Antéposition simple	-0.01	0.47	0.53	-0.56	0.62	-0.29	0.28	-0.25	0.23	-0.21	0.19
In situ	0.12	0.82	0.18	-0.47	0.77	-0.14	0.39	-0.09	0.34	-0.07	0.32

TABLEAU M.12 – Distributions postérieures pour le modèle *brms* (MG-C, stéréotype RICHESSE, effet TYPE)

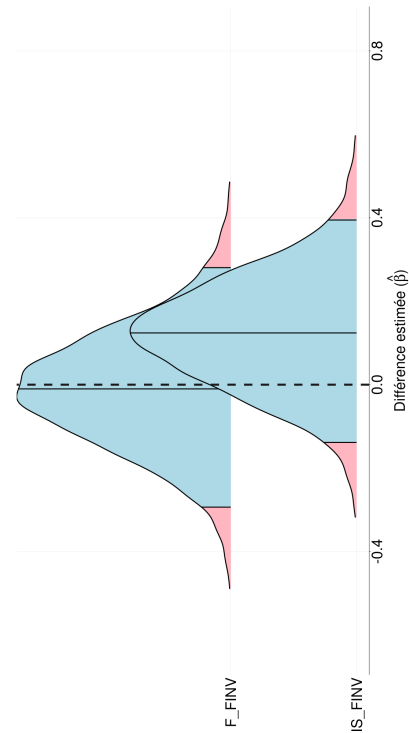


FIGURE M.7 – Distributions postérieures pour le modèle *brms* (MG-C, stéréotype RICHESSE, effet TYPE)

M.0.8 Modèle *brms* et distributions postérieures pour un effet de TYPE sur le stéréotype LOISIRS (MG-C, 6.2.4)

Family : bernoulli

Links : mu = logit

Formula : answernum ~ (f_finv + is_finv) + (f_finv + is_finv | id) + (f_finv + is_finv | item)

Data : subset(DON, stere=="h") (Number of observations : 1746)

Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	-0.02	0.12	-0.26	0.22	1.00	4513	4170
f_finv	-0.11	0.14	-0.37	0.16	1.00	9418	4527
is_finv	-0.09	0.14	-0.36	0.17	1.00	8661	4533

TABEAU M.13 – Résumé du modèle *brms* (MG-C, stéréotype LOISIRS, effet TYPE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
Antéposition simple	-0.11	0.21	0.79	-0.57	0.50	-0.37	0.16	-0.33	0.12	-0.30	0.09
In situ	-0.09	0.25	0.75	-0.65	0.48	-0.36	0.17	-0.32	0.13	-0.29	0.10

TABEAU M.14 – Distributions postérieures pour le modèle *brms* (MG-C, stéréotype LOISIRS, effet TYPE)

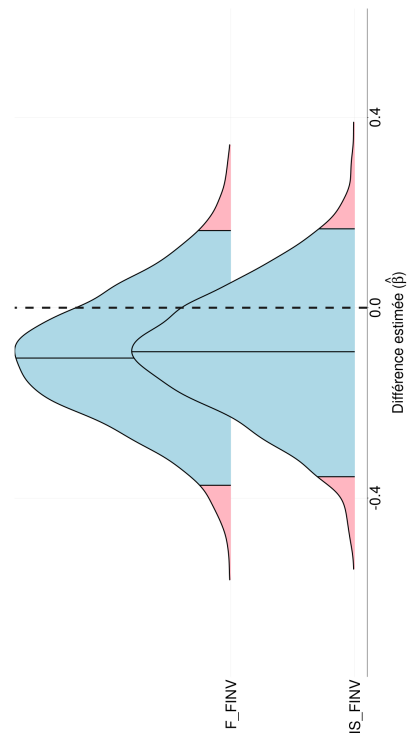


FIGURE M.8 – Distributions postérieures pour le modèle *brms* (MG-C, stéréotype LOISIRS, effet TYPE)

M.0.9 Modèle *brms* et distributions postérieures pour un effet de TYPE sur le stéréotype ÉDUCATION (MG-C, 6.2.4)

Family : bernoulli
Links : mu = logit
Formula : answernum ~ (f_finv + is_finv) + (f_finv + is_finv | id) + (f_finv + is_finv | item)
Data : subset(DON, stere=="e") (Number of observations : 1743)
Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	-0.28	0.08	-0.44	-0.12	1.00	5567	4409
f_finv	-0.11	0.13	-0.38	0.15	1.00	7810	4778
is_finv	-0.39	0.14	-0.67	-0.12	1.00	7353	4759

TABLEAU M.15 – Résumé du modèle brms (MG-C, stéréotype ÉDUCATION, effet TYPE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l85	h85
Antéposition simple	-0.11	0.20	0.80	-0.62	0.41	-0.38	0.15	-0.30	0.08
In situ	-0.39	0.00	1.00	-1.00	0.13	-0.67	-0.12	-0.60	-0.19

TABLEAU M.16 – Distributions postérieures pour le modèle brms (MG-C, stéréotype ÉDUCATION, effet TYPE)

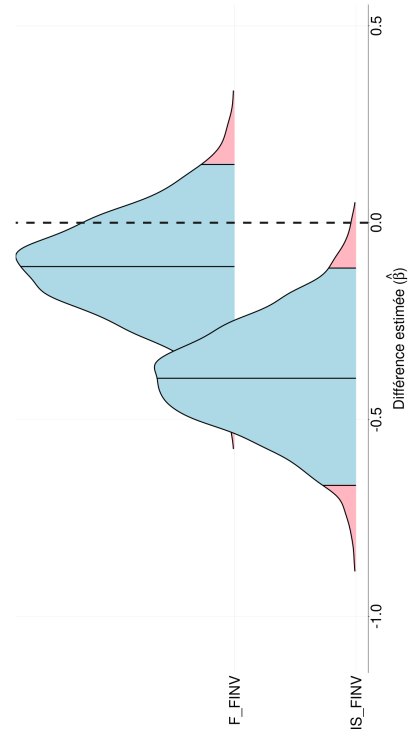


FIGURE M.9 – Distributions postérieures pour le modèle brms (MG-C, stéréotype ÉDUCATION, effet TYPE)

M.0.10 Modèle *brms* et distributions postérieures pour une interaction TYPE*GROUPE-CLASSE sur le stéréotype RICHESSE (MG-C, 6.2.4.3)

Family : bernoulli

Links : mu = logit

Formula : $\text{answernum} \sim (\text{f_finv} + \text{is_finv}) * \text{class5num} + ((\text{f_finv} + \text{is_finv}) | \text{id}) + (\text{class5num} * (\text{f_finv} + \text{is_finv}) | \text{item})$

Data : subset(DON, stere == "r") (Number of observations : 1786)

Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	0.18	0.23	-0.26	0.63	1.00	4172	5006
f_finv	-0.25	0.40	-1.03	0.52	1.00	3930	4101
is_finv	0.58	0.41	-0.25	1.38	1.00	3865	4499
class5num	-0.06	0.08	-0.21	0.09	1.00	3967	4940
f_finv :class5num	0.08	0.13	-0.17	0.34	1.00	3828	3834
is_finv :class5num	-0.15	0.13	-0.41	0.11	1.00	3862	4393

TABLEAU M.17 – Résumé du modèle *brms* (MG-C, stéréotype RICHESSE, effet TYPE*CLASSE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	195	h95	190	h90	185	h85
F_FINV	-0.25	0.27	0.73	-1.75	1.31	-1.03	0.52	-0.93	0.39	-0.84	0.31
IS_FINV	0.58	0.92	0.08	-1.17	2.22	-0.25	1.38	-0.11	1.26	-0.01	1.17
F_FINV*GROUPECLASSE	0.08	0.73	0.27	-0.44	0.57	-0.17	0.34	-0.13	0.30	-0.10	0.28
IS_FINV*GROUPECLASSE	-0.15	0.13	0.87	-0.69	0.38	-0.41	0.11	-0.37	0.07	-0.34	0.04

TABLEAU M.18 – Distributions postérieures du modèle *brms* (MG-C, stéréotype RICHESSE, effet TYPE*CLASSE)

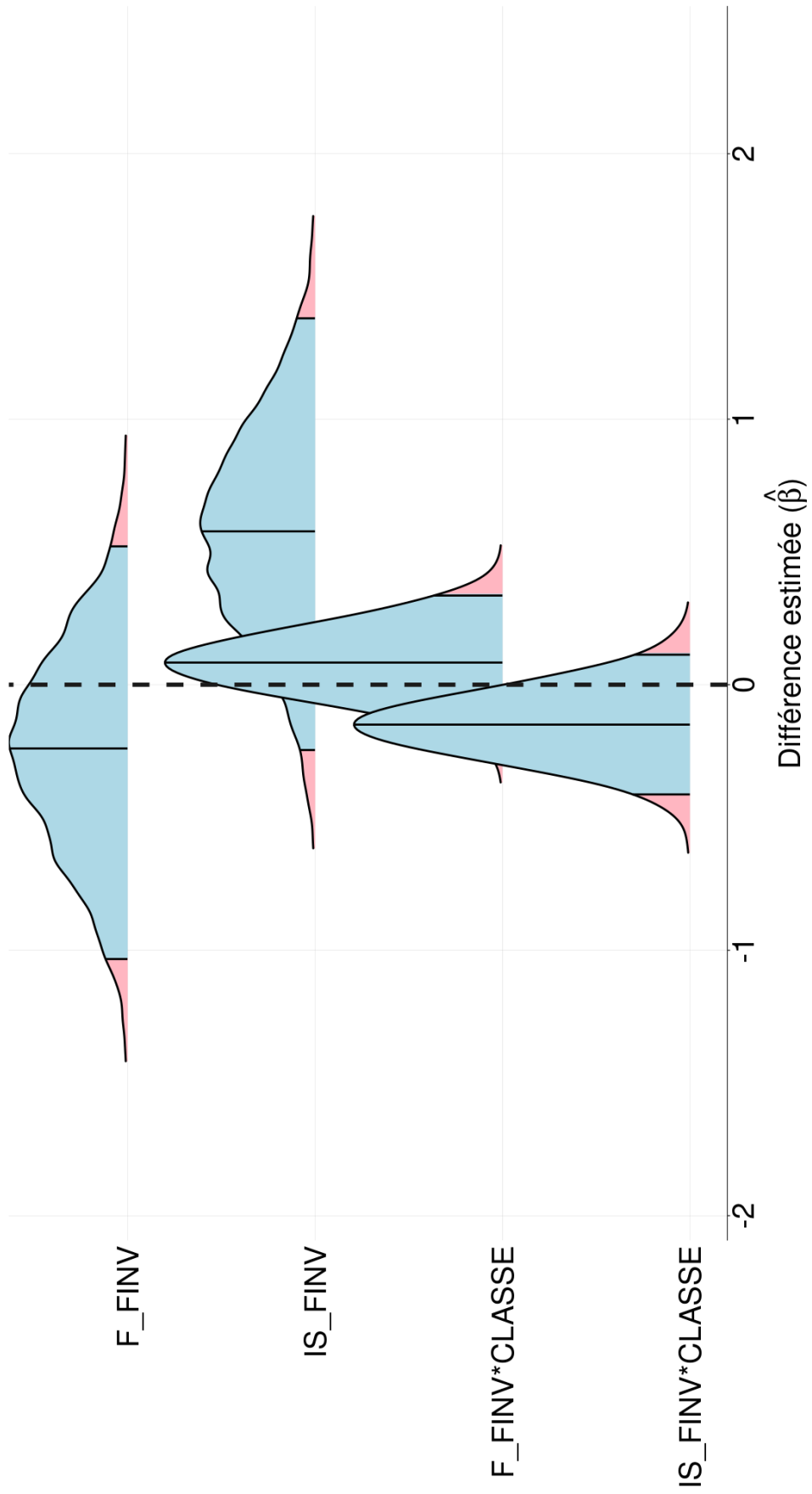


FIGURE M.10 – Distributions postérieures pour le modèle brms (MG-C, stéréotype RICHESSE, effet TYPE*CLASSE)

M.0.11 Modèle *brms* et distributions postérieures pour une interaction TYPE*GROUPE-CLASSE sur le stéréotype LOISIRS (MG-C, 6.2.4.3)

Family : bernoulli

Links : mu = logit

Formula : answernum ~ (f_finv + is_finv) * class5num + ((f_finv + is_finv) | id) + (class5num * (f_finv + is_finv) | item)

Data : subset(DON, stere == "h") (Number of observations : 1746)

Samples : 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	-0.76	0.26	-1.28	-0.24	1.00	6624	8743
f_finv	0.02	0.42	-0.79	0.86	1.00	6245	7414
is_finv	0.19	0.41	-0.60	1.00	1.00	6461	7289
class5num	0.26	0.08	0.09	0.42	1.00	7255	8067
f_finv :class5num	-0.05	0.14	-0.32	0.21	1.00	6464	7267
is_finv :class5num	-0.10	0.13	-0.36	0.16	1.00	6393	7916

TABLEAU M.19 – Résumé du modèle *brms* (MG-C, stéréotype LOISIRS, effet TYPE*CLASSE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l85	h85
F_FINV	0.02	0.52	0.48	-1.67	1.66	-0.79	0.86	-0.67	0.72
IS_FINV	0.19	0.69	0.31	-1.48	1.64	-0.60	1.00	-0.47	0.86
F_FINV*GROUPECLASSE	-0.05	0.36	0.64	-0.62	0.47	-0.32	0.21	-0.27	0.17
IS_FINV*GROUPECLASSE	-0.10	0.23	0.77	-0.59	0.42	-0.36	0.16	-0.32	0.12

TABLEAU M.20 – Distributions postérieures du modèle *brms* (MG-C, stéréotype LOISIRS, effet TYPE*CLASSE)

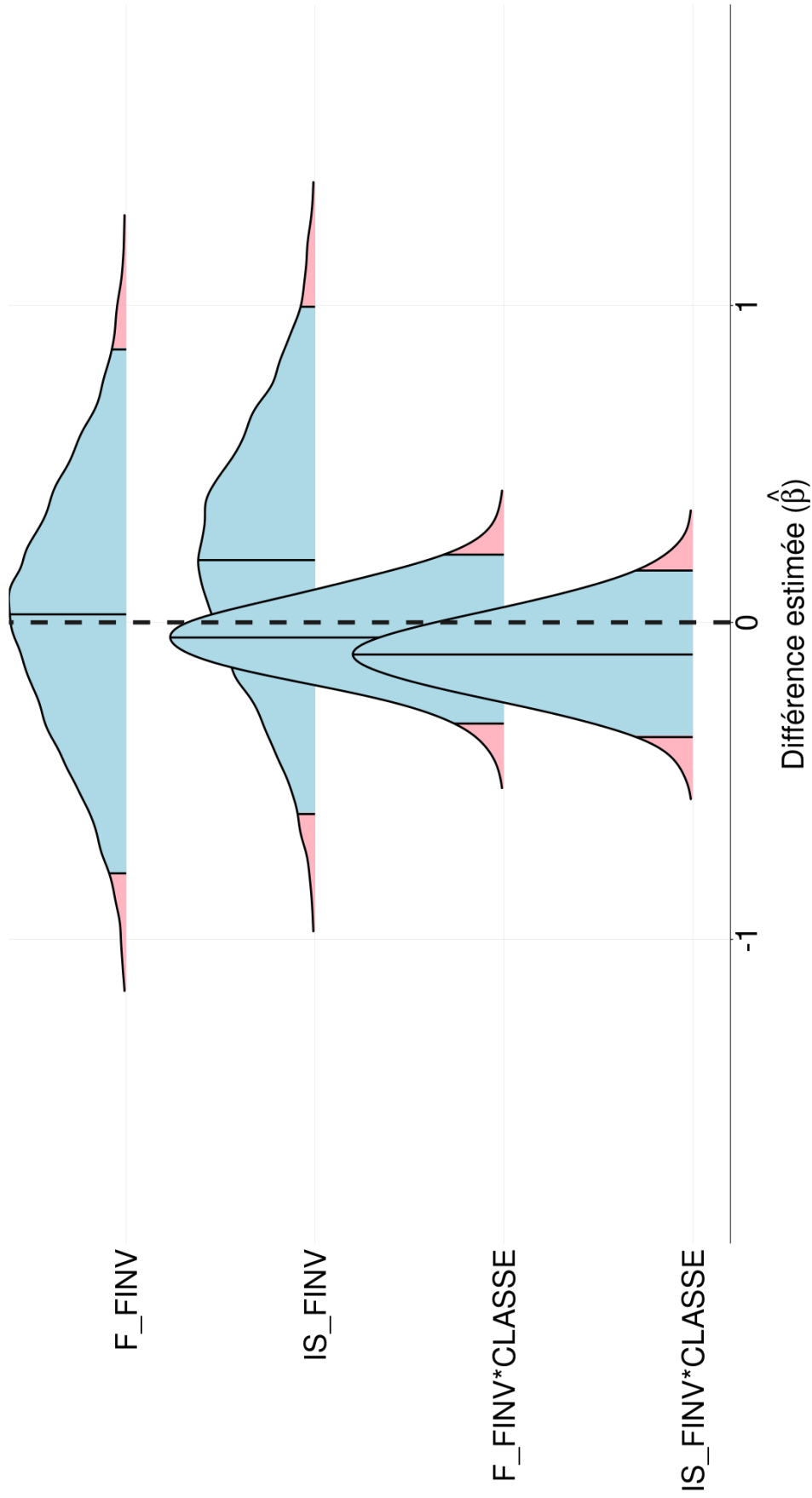


FIGURE M.1.1 – Distributions postérieures pour le modèle brms (MG-C, stéréotype LOISIRS, effet TYPE*CLASSE)

M.0.12 Modèle *brms* et distributions postérieures pour une interaction TYPE*GROUPE-CLASSE sur le stéréotype ÉDUCATION (MG-C, 6.2.4.3)

Family : bernoulli

Links : mu = logit

Formula : answernum ~ (f_finv + is_finv) * class5num + ((f_finv + is_finv) | id) + (class5num * (f_finv + is_finv) | item)

Data : subset(DON, stere == "e") (Number of observations : 1743)

Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	-0.63	0.20	-1.02	-0.25	1.00	9013	4833
f_finv	-0.19	0.41	-1.00	0.61	1.00	6767	4912
is_finv	-0.49	0.42	-1.32	0.34	1.00	7022	5158
class5num	0.12	0.06	-0.01	0.25	1.00	8563	4527
f_finv :class5num	0.03	0.13	-0.23	0.29	1.00	6771	4736
is_finv :class5num	0.03	0.14	-0.24	0.30	1.00	6838	5126

TABLEAU M.21 – Résumé du modèle *brms* (MG-C, stéréotype ÉDUCATION, effet TYPE*CLASSE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
F_FINV	-0.19	0.32	0.68	-1.79	1.18	-1.00	0.61	-0.88	0.47	-0.79	0.38
IS_FINV	-0.49	0.12	0.88	-1.98	1.24	-1.32	0.34	-1.18	0.20	-1.08	0.11
F_FINV*GROUPECLASSE	0.03	0.58	0.42	-0.47	0.52	-0.23	0.29	-0.19	0.25	-0.16	0.22
IS_FINV*GROUPECLASSE	0.03	0.59	0.41	-0.56	0.52	-0.24	0.30	-0.19	0.25	-0.16	0.22

TABLEAU M.22 – Distributions postérieures du modèle *brms* (MG-C, stéréotype ÉDUCATION, effet TYPE*CLASSE)

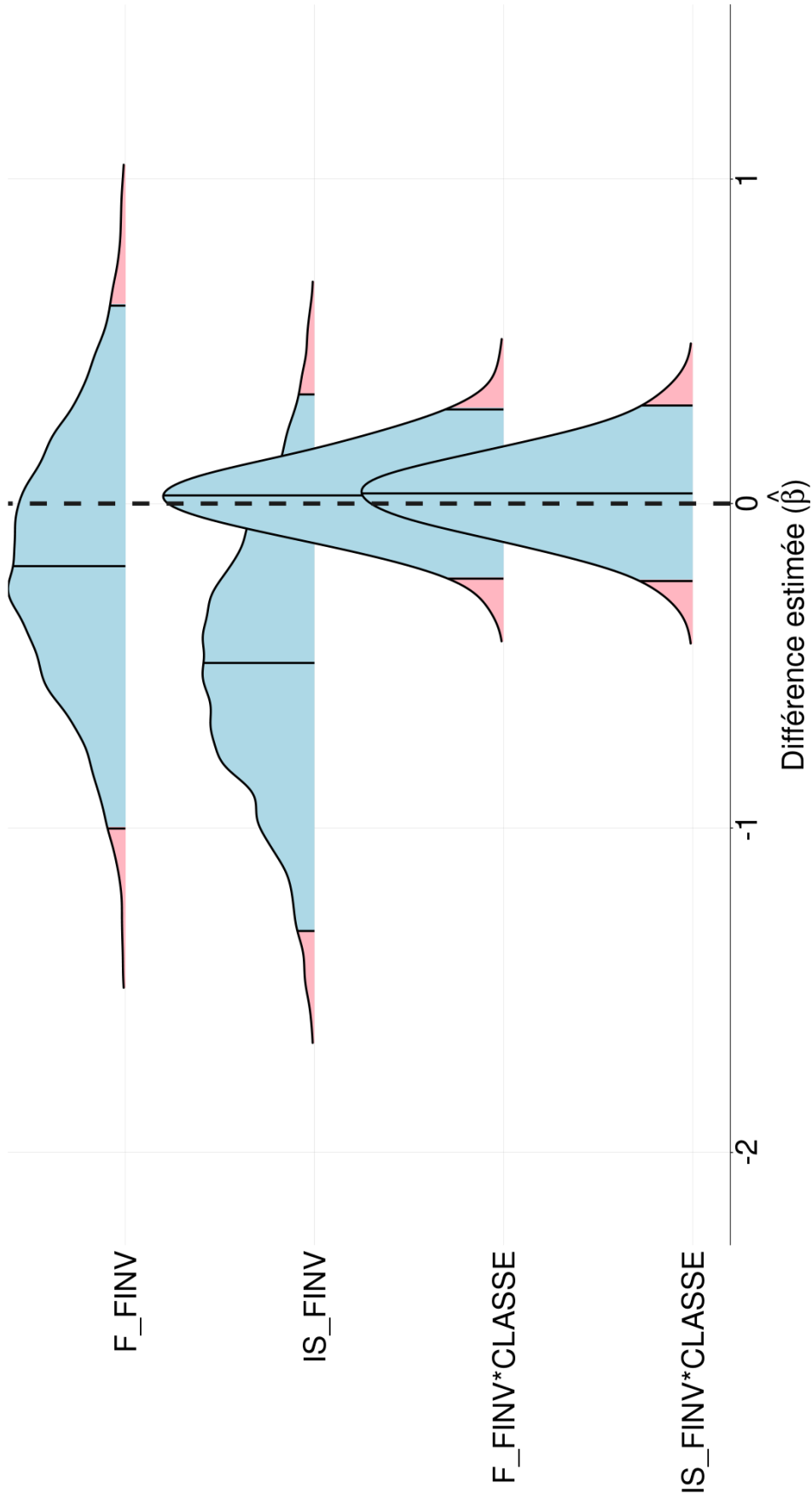


FIGURE M.12 – Distributions postérieures pour le modèle brms (MG-C, stéréotype ÉDUCATION, effet TYPE*CLASSE)

M.0.13 Modèle *brms* et distributions postérieures pour un modèle global (MG-C, 6.2.4.3.4)**Family :** bernoulli**Links :** mu = logit**Formula :** $\text{answernum} \sim (\text{f_finv} + \text{is_finv}) * (\text{r_e} + \text{h_e}) * \text{class5num} + ((\text{f_finv} + \text{is_finv}) * (\text{r_e} + \text{h_e}) | \text{id}) + ((\text{f_finv} + \text{is_finv}) * (\text{r_e} + \text{h_e}) * \text{class5num} | \text{item})$ **Data :** DON (Number of observations : 5275)**Samples :** 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	-0.40	0.14	-0.68	-0.11	1.00	6222	5029
f_finv	-0.14	0.24	-0.60	0.32	1.00	7772	4843
is_finv	0.09	0.23	-0.36	0.54	1.00	7691	5144
r_e	0.84	0.29	0.27	1.41	1.00	6589	4786
h_e	-0.10	0.30	-0.67	0.49	1.00	7097	4851
class5num	0.10	0.05	0.01	0.19	1.00	6396	5382
f_finv:r_e	-0.09	0.57	-1.19	1.04	1.00	5617	4559
f_finv:h_e	0.24	0.57	-0.87	1.35	1.00	5769	4840
is_finv:r_e	1.01	0.59	-0.14	2.15	1.00	5845	4386
is_finv:h_e	0.63	0.58	-0.51	1.79	1.00	5984	4573
f_finv:class5num	0.02	0.08	-0.13	0.18	1.00	7798	4976
is_finv:class5num	-0.07	0.08	-0.22	0.08	1.00	7667	4969
r_e:class5num	-0.19	0.10	-0.38	-0.00	1.00	6242	4934
h_e:class5num	0.12	0.09	-0.06	0.31	1.00	6921	4447
f_finv:r_e:class5num	0.06	0.19	-0.31	0.43	1.00	5530	4612
f_finv:h_e:class5num	-0.08	0.18	-0.44	0.28	1.00	5843	4677
is_finv:r_e:class5num	-0.16	0.19	-0.54	0.21	1.00	5926	4810
is_finv:h_e:class5num	-0.12	0.19	-0.49	0.26	1.00	5881	4322

TABLEAU M.23 – Résumé du modèle *brms* global (MG-C)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
F_FINV	-0.14	0.28	0.72	-0.98	0.67	-0.60	0.32	-0.53	0.25	-0.49	0.20
IS_FINV	0.09	0.65	0.35	-0.75	0.96	-0.36	0.54	-0.29	0.47	-0.24	0.42
Rich_Edu	0.84	1.00	0.00	-0.33	1.86	0.27	1.41	0.36	1.31	0.43	1.25
Loisirs_Edu	-0.10	0.37	0.63	-1.12	1.09	-0.67	0.49	-0.58	0.39	-0.52	0.33
F_FINV*Rich_Edu	-0.09	0.44	0.56	-2.01	2.07	-1.19	1.04	-1.02	0.84	-0.90	0.74
F_FINV*Lois_Edu	0.24	0.66	0.34	-1.82	2.49	-0.87	1.35	-0.69	1.18	-0.58	1.07
IS_FINV*Rich_Edu	1.01	0.96	0.04	-0.98	3.10	-0.14	2.15	0.03	1.96	0.15	1.85
IS_FINV*Lois_Edu	0.63	0.86	0.14	-1.64	2.80	-0.51	1.79	-0.32	1.59	-0.21	1.47
F_FINV*CLASSE	0.02	0.61	0.39	-0.26	0.29	-0.13	0.18	-0.11	0.15	-0.09	0.13
IS_FINV*CLASSE	-0.07	0.18	0.82	-0.33	0.22	-0.22	0.08	-0.20	0.05	-0.18	0.04
Rich_Edu*CLASSE	-0.19	0.02	0.98	-0.53	0.23	-0.38	-0.00	-0.35	-0.03	-0.33	-0.06
Lois_Edu*CLASSE	0.12	0.91	0.09	-0.22	0.49	-0.06	0.31	-0.03	0.28	-0.01	0.26
F_FINV*Rich_Edu*CLASSE	0.06	0.63	0.37	-0.65	0.80	-0.31	0.43	-0.24	0.37	-0.20	0.33
F_FINV*Lois_Edu*CLASSE	-0.08	0.33	0.67	-0.84	0.59	-0.44	0.28	-0.38	0.22	-0.35	0.18
IS_FINV*Rich_Edu*CLASSE	-0.16	0.20	0.80	-0.83	0.54	-0.54	0.21	-0.47	0.15	-0.44	0.11
IS_FINV*Lois_Edu*CLASSE	-0.12	0.27	0.73	-0.77	0.69	-0.49	0.26	-0.43	0.20	-0.39	0.16

TABLEAU M.24 – Distributions postérieures pour un modèle brms global (MG-C)

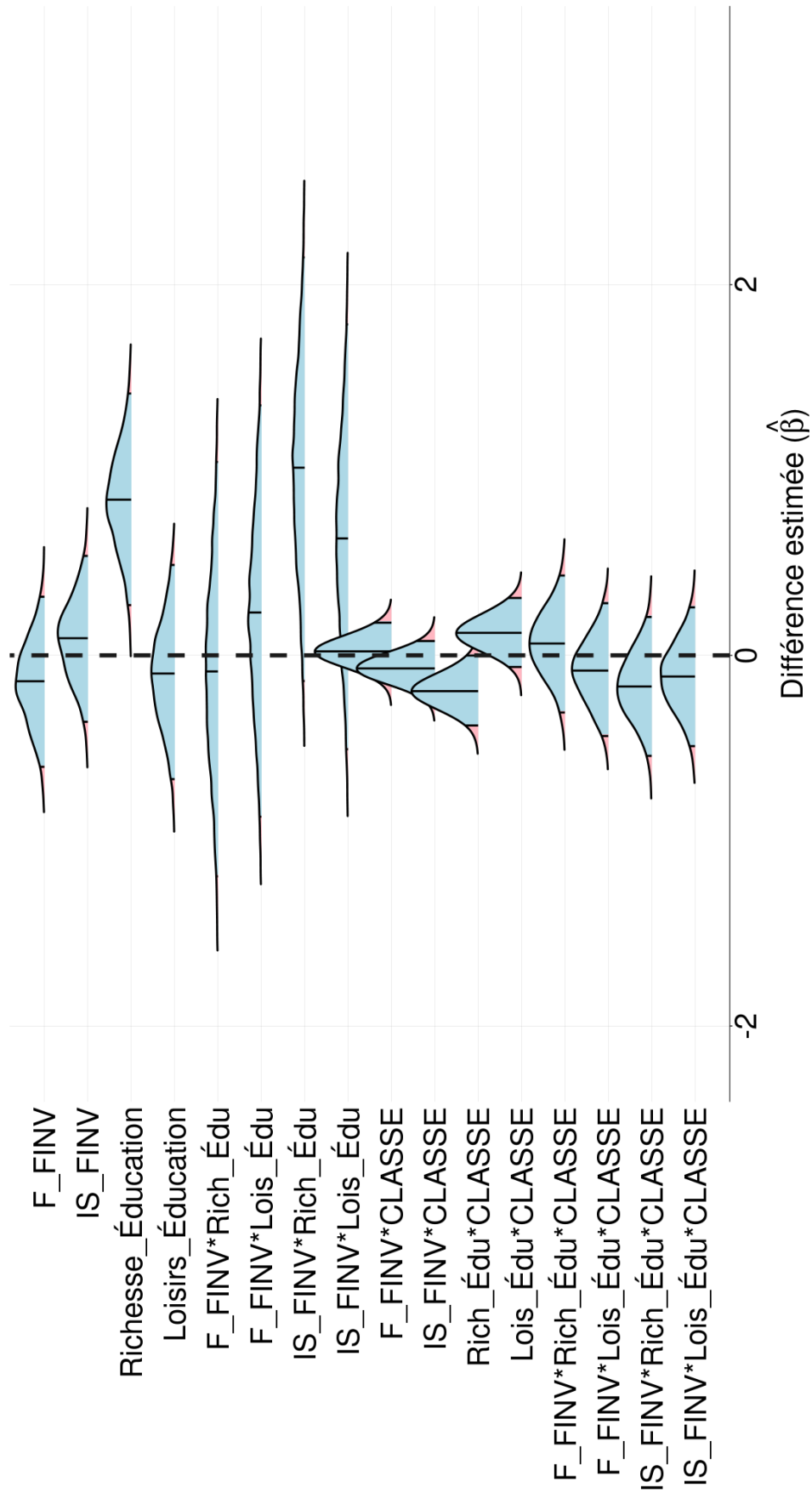


FIGURE M.13 – Distributions postérieures pour un modèle brms global (MG-C)

Annexe N

Production Enfantine : Dossier Éthique

N.0.1 Convention signée entre la Direction Académique de Caen et l'Université de Paris pour une intervention en classe)

REFUDD7: C005125/0

Convention d'accueil
Pour un doctorant
« Monsieur Gabriel Thiberge »

Circonscription
08 JUL. 2019
de Bayeux

Entre les soussignés

L'Université Paris Diderot
Établissement Public à caractère Scientifique Culturel et Professionnel (EPSCP),
dont le siège est situé 6 rue Thomas Mann, 75205 Paris cedex 13
N° SIREN 197 517 238, code APE 8542Z,
Représenté par Monsieur Reiner VEITA en sa qualité d'Administrateur Provisoire

Ci -après désigné par « **Établissement employeur** ».

Agissant au nom et pour le compte du Laboratoire de linguistique formelle LLF (UMR 7110) dirigé par M.
BONAMI olivier

Ci -après désigné par « **Le laboratoire** »

Et

La Direction des services départementaux de l'éducation nationale (DSDEN) du Calvados, représentée par le
Directeur Académique des Services de l'Éducation Nationale (DASEN) du Calvados : Mathias Bouvier, 2 place
de l'Europe à Hérouville Saint Clair.

L'Université Paris Diderot, la DSDEN

Ci -après désignés individuellement « la Partie » et conjointement par « les Parties ».

Et

Monsieur Gabriel Thiberge : chercheur responsable du projet, doctorant rattaché à l'École Doctorale 132
« Sciences du Langage » et au Laboratoire de Linguistique Formelle (UMR7110, CNRS).

Ci -après désigné par « **Le doctorant** »

N.0.2 Convention signée entre la Direction Académique de Caen et l'Université de Paris pour une intervention en classe)

Il est préalablement exposé que :

Monsieur Gabriel Thiberge est inscrit à l'Université Paris Diderot (Etablissement employeur) en 3ème année de doctorat, en qualité de doctorant contractuel. Le doctorant est rattaché au Laboratoire de linguistique formelle LLF (UMR 7110) où il réalise ses travaux de recherche sous la direction de Mme Barbara HEMFORTH, Directrice de Recherche au CNRS.

Le sujet de thèse du doctorant s'intitule : « "Structure syntaxique des interrogatives : entre optionnalité linguistique et contraintes cognitives" »

Dans le cadre de sa thèse, le doctorant est amené à entreprendre des observations de classe, des questionnaires, des entretiens au sein de l'École primaire et maternelle d'Argouges. C'est dans cette perspective, que l'Établissement employeur pour le compte du Laboratoire LLF, propose la mise en place de la présente convention d'accueil pour que le doctorant puisse intervenir au sein de l'école précitée.

Il est convenu ce qui suit

Article 1 : Objet

La présente convention vise à définir les conditions et modalités d'accueil du doctorant au sein de l'École primaire et maternelle d'Argouges pour la réalisation de son projet de recherche.

1-1 Objectifs du projet

Ce projet d'intervention en école vise à mieux comprendre le développement de la compétence sociolinguistique chez les enfants francophones de naissance, de 4 à 11 ans. Il consiste en une récolte de données langagières en école, permettant une analyse des productions de phrases interrogatives par les enfants et du développement des stéréotypes sociaux intériorisés à mesure que leur âge progresse.

1-2 Activités confiées au chercheur en fonction de l'objectif du projet :

Récolter des données linguistiques auprès d'enfants francophones à deux niveaux :

1°/ leurs préférences et les stéréotypes développés au fil du temps en fonction de la structure syntaxique des interrogatives, par le biais d'un jeu d'appariement entre des stimuli sonores (voix de francophones adultes préenregistrées) et des visages dessinés ;

2°/ leurs productions spontanées dans des situations de jeux de rôle où les enfants revêtent le rôle d'un serveur ou d'une serveuse accueillant les clients d'un restaurant.

Article 2 : Lieu et durée du projet

2-1 Lieu précis de déroulement du projet : École primaire et maternelle d'Argouges, 14400 BAYEUX

2-2 Dates de déroulement du projet dans l'établissement : entre le 20 septembre 2019 et le 19 novembre 2019, en concertation avec les équipes pédagogiques et la direction de l'établissement pour minimaliser les perturbations dans les activités pédagogiques. Les tâches à réaliser doivent l'être en deux sessions distinctes, espacées de quelques semaines.

Article 3 : Suivi du doctorant

Au sein de l'établissement employeur : Mme Barbara Hemforth, Directrice de Recherches CNRS, chargée du suivi du projet (direction de thèse).

Au sein de l'École primaire et maternelle d'Argouges: Mme Élisabeth Houel, directrice de l'établissement. A ce titre, le doctorant doit respecter les consignes données par cette dernière.

Le doctorant ne peut remplacer un personnel de l'école absent (notamment dans le cadre de la surveillance et de l'encadrement d'élèves qu'il ne peut assurer). De principe, il ne peut donc être seul avec un élève.

La direction de l'École primaire et maternelle d'Argouges ne doit pas mettre le doctorant dans une situation de risque ou de danger et s'assurer qu'à aucun moment la présence du doctorant n'est source de perturbation ou de risque pour les élèves.

Article 4 : Protection sociale et Responsabilité

4.1 Protection sociale

Pendant la période stipulée à l'article 2, l'Université Paris Diderot reste l'employeur du doctorant En qualité d'établissement employeur, il s'acquittera des obligations sociales et fiscales relevant de l'exécution de son statut

N.0.3 Convention signée entre la Direction Académique de Caen et l'Université de Paris pour une intervention en classe)

ou du contrat de travail, notamment en ce qui concerne les accidents de travail, de trajet et les maladies professionnelles.

En cas d'accident de travail du doctorant sur le site de l'École primaire et maternelle d'Argouges, la Direction des Ressources Humaines de l'établissement employeur devra être prévenue sous 48 heures afin que celle-ci procède à la déclaration d'accident du travail lui incombant en sa qualité d'employeur.

4.2 Responsabilité.

L'Établissement employeur conserve à l'égard du doctorant sa responsabilité d'employeur. À ce titre, il est seul détenteur du pouvoir disciplinaire à l'égard du doctorant, et à ce titre il peut être saisi par les parties. En cas de comportement fautif du doctorant, l'Établissement employeur sera immédiatement informé et il pourra être mis fin à l'accueil du doctorant. Dans cette hypothèse, la résiliation interviendra de plein droit, 15 jours après l'envoi d'une lettre recommandée avec accusé de réception.

Le doctorant est placé sous l'autorité fonctionnelle de l'École en qualité d'établissement d'accueil. Il est soumis dans l'exercice de ses activités, aux règles de fonctionnement de l'École, à son règlement intérieur et à ses règles en matière d'hygiène, de sécurité, d'éthique, d'utilisation et de ses systèmes d'information. Tout document réglementaire (Règlement intérieur, charte) devra être transmis au doctorant. Il doit aussi respecter, dans ses interventions, les obligations imposées par les services académiques.

Le doctorant doit également soumettre, préalablement, à la directrice d'école toute demande de collecte de données personnelles, certaines pouvant nécessiter en amont un accord exprès des parents.

En cas d'absence, le doctorant doit informer l'établissement employeur et la direction de l'école dans les 24h ouvrables précédant sa venue dans l'école quel qu'en soit le motif (maladie, congés exceptionnels, congés payés, etc.).

Article 5 : Assurances

Chaque Partie est responsable, des dommages que ses agents ou les personnes relevant de son autorité pourraient causer aux tiers, y compris les dommages résultant de l'utilisation d'équipements ou de matériels appartenant à l'autre Partie.

Toutefois, la responsabilité des parties ne saurait être engagée du fait de la présence du doctorant dans les locaux de l'École à l'occasion des dommages éventuels qu'il pourrait occasionner résultant d'une faute lourde ou intentionnelle, tant matériels qu'immatériels, sur sa personne, sur un usager ou un personnel de l'École. C'est la responsabilité propre du doctorant qui sera actionnée. La responsabilité des parties ne pouvant être engagée, le doctorant devra donc souscrire une assurance civile personnelle.

Article 6 : Prestations

Le doctorant ne perçoit aucun salaire ni gratification de la part de l'éducation nationale.

Article 7 : Confidentialité - Publication

7.1 Chacune des Parties ainsi que le doctorant s'engagent à ne pas publier ni divulguer de quelque façon que ce soit des informations à caractère confidentiel appartenant à l'autre Partie ou ses mandants ou dont elle pourrait avoir connaissance à l'occasion de la présente convention et ce, tant que ces informations ne seront pas du domaine public. Les dispositions ci-dessus ne sont pas applicables aux informations qui font déjà partie du domaine public à ce jour, ou qui viendront à y tomber ultérieurement sans faute de la part de la Partie à qui elles ont été communiquées, ou qui étaient déjà en sa possession auparavant sans obligation de secret à leur égard. De même, le doctorant s'engage à ne pas divulguer de données personnelles au sens de l'article 2 de la loi n°78-17 du 6 janvier 1978 et en application du règlement général sur la protection des données entré en application le 25 mai 2018.

7.2 Le doctorant après concertation avec sa directrice de recherche et avec Mme Elisabeth Houel, directrice de l'école pourra réaliser des publications de recherche prenant pour objet les activités du projet (tout en respectant l'anonymat des personnes et des lieux). Ces publications et communications devront mentionner le concours apporté par chacune des Parties à la réalisation dudit travail. Pour l'établissement employeur, le **Doctorant** devra respecter la Charte de signature des publications scientifiques de l'Université Paris Diderot en vigueur.

N.0.4 Convention signée entre la Direction Académique de Caen et l'Université de Paris pour une intervention en classe)

Les données recueillies ou retranscrites par le biais de mesures ou questionnaires resteront anonymes et aucune information ne sera donnée sur l'identité des élèves participant au projet.

Article 8 : Bilan

Le doctorant adressera à l'IA-DASEN ayant validé le projet un exemplaire du bilan du projet. Un double sera également transmis à la direction de l'Ecole primaire et maternelle d'Argouges, à sa demande.

Article 9 : Durée et résiliation de la convention d'accueil

La présente convention est en vigueur pour la période mentionnée à l'article 2 soit du 20 septembre 2019 au 19 novembre 2019.

Si l'une des trois parties décide d'interrompre définitivement le projet, elle doit en informer les autres parties le plus rapidement possible par un courrier motivé.

Toutefois, en cas de manquement à la discipline et/ou de faute grave du doctorant ou en cas de perturbation du fonctionnement de l'école, l'IA-DASEN, en lien avec la direction de l'école, se réserve le droit de mettre fin au projet sans préavis et en informe immédiatement l'établissement employeur.

Article 10 « Création de l'Université de Paris »:

« En vertu des dispositions du décret n° 2019-209 du 20 mars 2019 portant création de l'université de Paris et approbation de ses statuts, les biens, droits et obligations de l'université Paris-VII (Paris Diderot) seront automatiquement transférés à l'Université de Paris au 1er janvier 2020. »

Fait à Hérouville Saint Clair, le 03/07/2019

Pour l'Université Paris Diderot

En qualité d'Administrateur Provisoire
Reiner VEITIA

UNIVERSITÉ
PARIS
DIDEROT
PARIS 7

Administrateur provisoire

Reiner VEITIA

Le directeur académique des services de
l'éducation nationale du Calvados,

Mathias BOUVIER

Le doctorant responsable du projet
Gabriel Thiberge

Visa :

Visa :

N.0.5 Convention signée entre la Direction Académique de Caen et l'Université de Paris pour une intervention en classe)


DESCRIPTIF DU PROJET

Le projet global s'inscrit dans la thèse de doctorat de Gabriel Thiberge. Il consiste en une étude du développement sociolinguistique de l'enfant francophone, c'est-à-dire de sa capacité à adapter ses productions langagières au contexte social d'une interaction. Un phénomène linguistique permettant d'accéder à cette compétence sont les phrases interrogatives du français, dont les différentes variantes ne sont pas associées aux mêmes stéréotypes sociaux chez les adultes (comparer « Quand viens-tu? » à « Tu viens quand? »). Ce projet est financé par l'École Doctorale 132 « Sciences du Langage » et réalisé au sein de l'Université Paris Diderot, du Laboratoire de Linguistique Formelle (CNRS, UMR 7110), et du LabEx « EFL » Empirical Foundations of Linguistics.

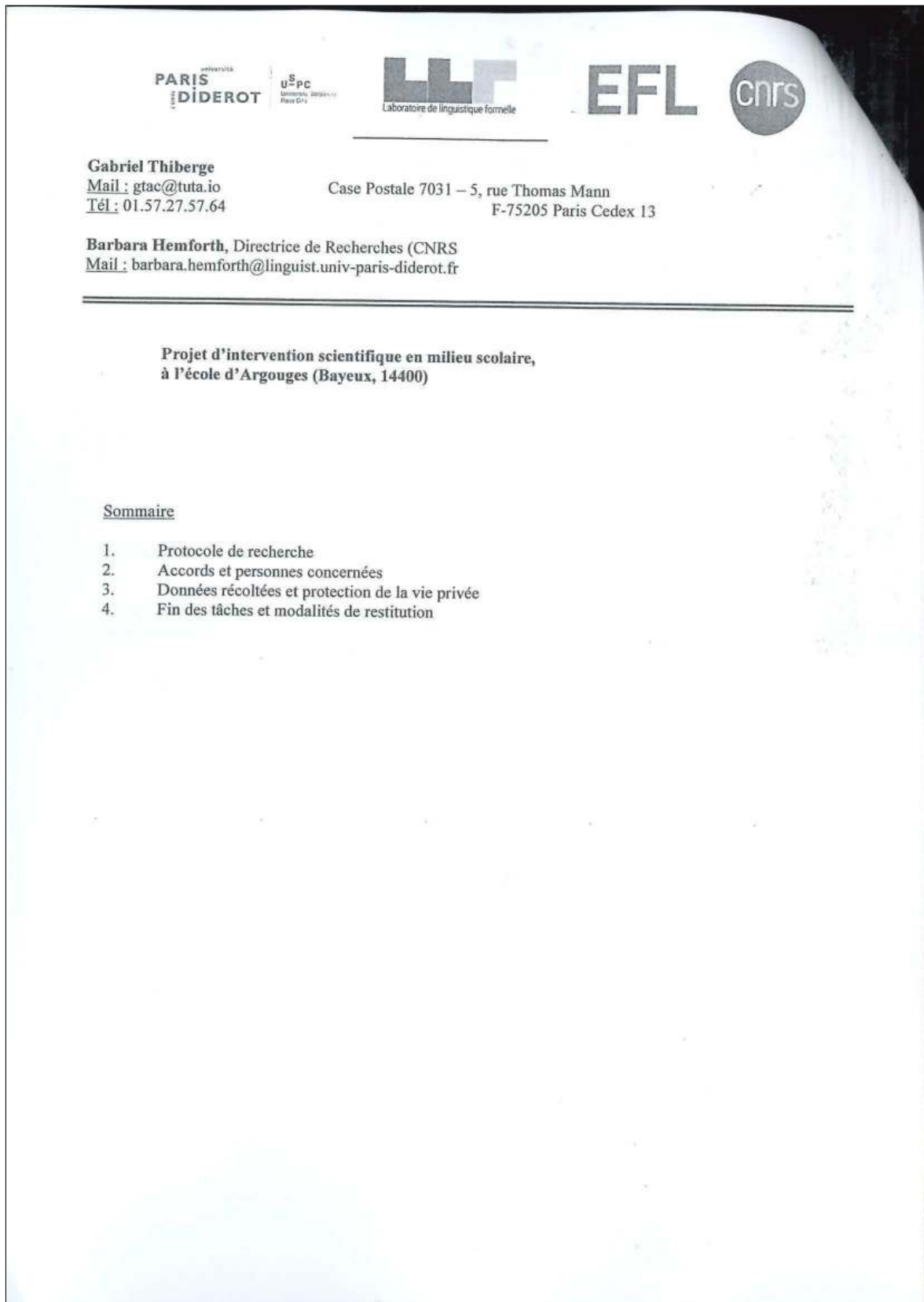
Le projet précis consiste ici en une récolte de données langagières en école, permettant une analyse des productions de phrases interrogatives par les enfants et du développement des stéréotypes sociaux intériorisés à mesure que leur âge progresse.

Un descriptif plus détaillé du projet et des tâches est annexé à la présente convention.

Date : le 09 avril 2019 Signature : Gabriel Thiberge



N.0.6 Convention signée entre la Direction Académique de Caen et l'Université de Paris pour une intervention en classe)



N.0.7 Convention signée entre la Direction Académique de Caen et l'Université de Paris pour une intervention en classe)

1. Protocole de Recherche

Deux tâches distinctes sont prévues, sans aucune obligation de participer aux deux (cf. Point 2 sur les accords requis). Chaque tâche commencerait par une période de familiarisation entre l'intervenante et les enfants participant, pendant laquelle les consignes seront expliquées, sous forme de petit jeu.

Session 1. Tâche du « Qui a dit quoi ? »

→ Écoute de sons et choix parmi deux images

La première tâche consisterait pour les enfants, sous forme de jeu, à écouter de courts enregistrements d'adultes francophones, puis à identifier parmi des paires de personnages (dessins) lequel ou laquelle a parlé. Il n'y aurait pas de bonne ou mauvaise réponse. Le but est de connaître les associations privilégiées par les enfants entre des voix et des images. À l'enfant, les choses seraient présentées sous une forme du genre : *« Les scientifiques ont mélangé tous leurs fichiers ! Est-ce que tu veux les aider à remettre de l'ordre dans leurs travaux ? Il faut entendre des personnes parler et dire qui a dit quoi. »*

Les entretiens seraient organisés de façon individuelle entre les enfants et un chercheur, sous surveillance distante d'un tiers familier de l'enfant éventuellement, d'une durée approximative de 15 minutes. Tout serait fait pour perturber au minimum les activités pédagogiques des classes concernées, et un endroit un peu calme et isolé avec une table et deux chaises serait le seul matériel nécessaire. Cette tâche pourrait aussi être menée, si le personnel pédagogique le juge pertinent, dans la salle informatique dont dispose l'école, par groupes de plusieurs élèves.

Les élèves que cette tâche concernerait auraient, dans l'idéal, entre 4 et 11 ans, et appartiendraient à des classes allant de la Grande Section de maternelle au CM2. Une dizaine d'enfants par classe serait idéal (soit 10x6=60 élèves), pour une validité optimale des calculs statistiques menés par la suite. Bien sûr, si plus de volontaires se présentent, nous les accepterons et les incorporerons à notre étude.

Chaque enfant entendrait, dans un ordre aléatoire, au maximum 45 phrases parmi les 90 suivantes :

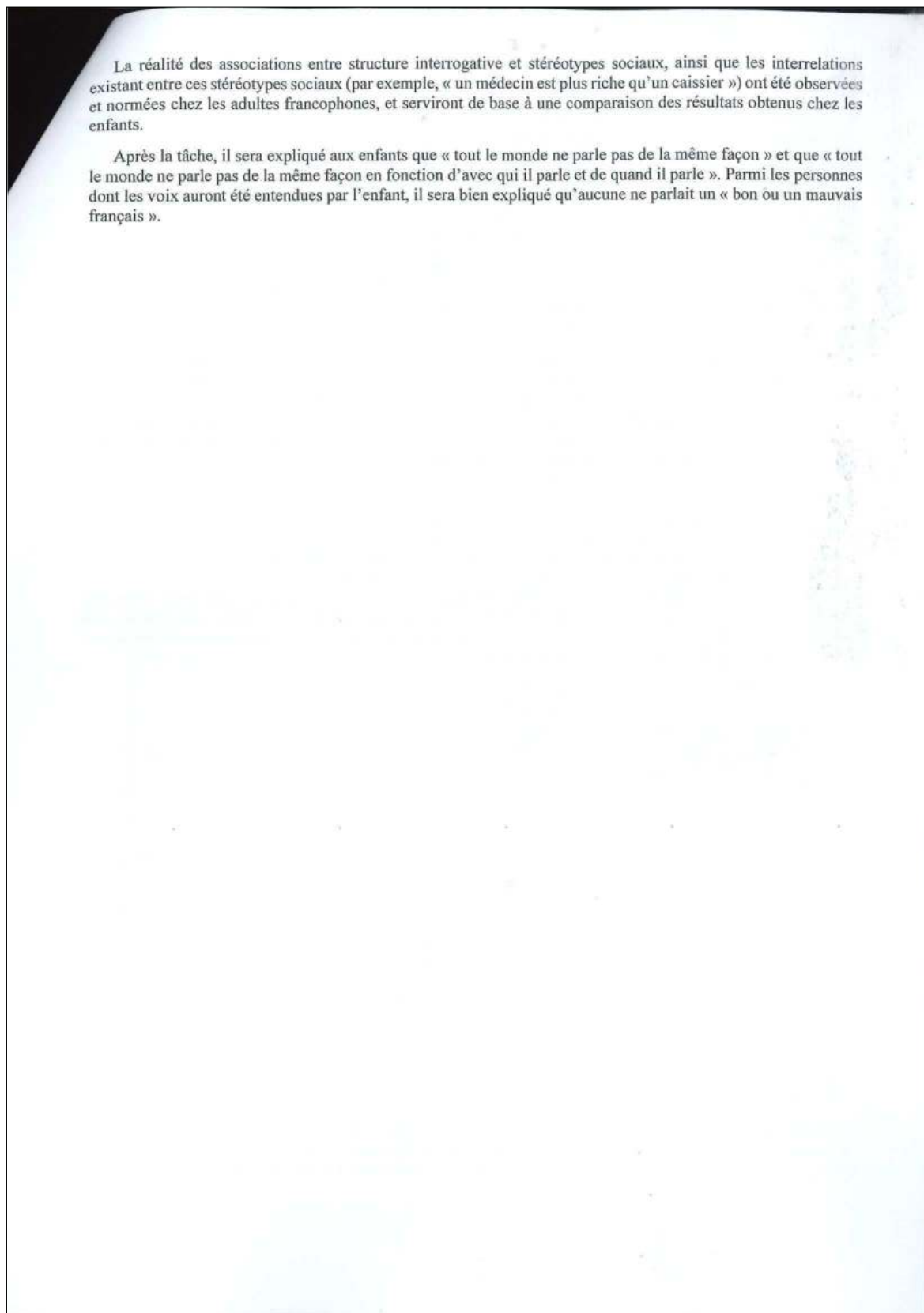
Condition 1	Condition 2	Condition 3
arrive à quelle heure ?	quelle heure il arrive ?	quelle heure arrive-t-il ?
les as mises dans lequel ?	ans lequel tu les as mises ?	ans lequel les as-tu mises ?
l'as cassé comment ?	omment tu l'as cassé ?	omment l'as-tu cassé ?
vas où ?	tu vas ?	vas-tu ?
as suivi quel raisonnement ?	quel raisonnement tu as suivi ?	quel raisonnement as-tu suivi ?
est qui ?	si c'est ?	si est-ce ?
saura quand ?	and on saura ?	and saura-t-on ?
parles de qui ?	qui tu parles ?	qui parles-tu ?
as pris lequel ?	quel tu as pris ?	quel as-tu pris ?
le ferais avec quoi ?	ec quoi tu le ferais ?	ec quoi le ferais-tu ?
est partie quand ?	and elle est partie ?	and est-elle partie ?
rechignent pourquoi ?	urquoi ils rechignent ?	urquoi rechignent-ils ?
l'as utilisée où ?	tu l'as utilisée ?	l'as-tu utilisée ?
ors ils l'ont donné à qui ?	ors à qui ils l'ont donné ?	ors à qui l'ont-ils donné ?
as arrêté à quel moment ?	quel moment tu as arrêté ?	quel moment as-tu arrêté ?
t'y prends comment ?	omment tu t'y prends ?	omment t'y prends-tu ?
elle connaîtra qui ?	si elle connaîtra ?	si connaîtra-t-elle ?
les préfères de quelle couleur ?	quelle couleur tu les préfères ?	quelle couleur les préfères-tu ?
us vous êtes battus avec qui ?	ec qui vous vous êtes battus ?	ec qui vous êtes-vous battus ?
fais dans le sous-entendu pourquoi ?	urquoi tu fais dans le sous-entendu ?	urquoi fais-tu dans le sous-entendu ?
as rencontré qui ?	si tu as rencontré ?	si as-tu rencontré ?
es allé sur quels sites ?	r quels sites tu es allé ?	r quels sites es-tu allé ?
n groupe ferait comment ?	omment ton groupe ferait ?	omment ferait ton groupe ?
as des difficultés avec quoi ?	ec quoi tu as des difficultés ?	ec quoi as-tu des difficultés ?
a pris lequel ?	quel tu as pris ?	quel as-tu pris ?
parles de qui ?	qui tu parles ?	qui parles-tu ?
t'envoient où ?	ils t'envoient ?	t'envoient-ils ?
dis ça pourquoi ?	urquoi tu dis ça ?	urquoi dis-tu ça ?
evient quand ?	and il revient ?	and revient-ils ?
us jouez à quoi ?	quoi vous jouez ?	quoi jouez-vous ?

N.0.8 Convention signée entre la Direction Académique de Caen et l'Université de Paris pour une intervention en classe)

Après avoir entendu chaque phrase, l'enfant devra choisir entre deux personnes susceptibles de l'avoir prononcée, et choisir celle qui lui paraît la plus probable, parmi les paires suivantes, qui auront été présentées avant :



N.0.9 Convention signée entre la Direction Académique de Caen et l'Université de Paris pour une intervention en classe)



N.0.10 Convention signée entre la Direction Académique de Caen et l'Université de Paris pour une intervention en classe)

Session 2. « Bonjour, c'est pour manger. »
→ Jeu de rôles et enregistrement sonore

La deuxième tâche, menée au cours d'une deuxième session, quelques temps après la première (environ un mois), consisterait à impliquer les enfants dans une situation de jeu de rôle. Les enfants joueraient le rôle d'un ou d'une employé.e de restaurant, accueillant la clientèle et prenant sa commande. Ces client.e.s seraient un ou une même adulte appartenant à l'équipe de recherche et ayant une expérience dans l'interaction avec des enfants, sans pour autant connaître celles et ceux de l'école. Cette personne revêtirait successivement deux rôles différents : celui d'un ou d'une jeune adulte s'exprimant dans un français quotidien, et celui d'une personne plus âgée s'exprimant dans un français plus soutenu, avec une marionnette. Les jeux dureraient au maximum 10 minutes.

Le but est ici de mettre l'enfant dans une situation d'interaction habituelle, où il ou elle aurait besoin de poser des questions à l'adulte, pour préciser sa commande et répondre à ses demandes de consommateur ou consommatrice. Pour un traitement des données ultérieur, les enfants seraient cette fois enregistrés pendant leur interaction avec l'adulte, mais seulement par audio. Aucun dispositif vidéo ne sera présent dans la pièce.

Les enfants que cette tâche impliquerait auraient, dans l'idéal, entre 4 à 8 ans maximum, et appartiendraient à des classes allant de la Grande Section de maternelle au CE2. Le nombre d'enfants que nous espérons pouvoir observer a moins d'influence sur l'analyse de cette tâche de production, mais nous avons le même objectif initial, une dizaine d'élèves par classe par établissement, soit $10 \times 2 \times 4 = 80$ élèves.

Les discussions seront minimalement scriptées, pour laisser place à une certaine improvisation de la part de la personne jouant le rôle des malades, en fonction de ce que dira l'enfant. La personne incarnant le rôle des malades dirait des choses comme :

«Bonjour, nous voudrions...»	ou	«Bonsoir, on voudrait...»
«Nous n'aimons pas trop...»	ou	«Moi j'aime pas trop...»
«Il serait peut-être préférable de...»	ou	«Y faudrait ptet...»
«Je m'excuse de vous déranger...»	ou	«Chuis désolé.e d'vous déranger...»
«Nous allons prendre des pommes de terre...»	ou	«On va prendre des patates...»

et autres variations du même genre, destinées autant à favoriser la production de questions pour préciser les demandes des client.e.s, qu'à introduire pour l'enfant un contexte d'interaction plutôt relâché, très quotidien, ou bien au contraire assez formel. Les restaurants seront également d'un standing différent pour chaque client, et présentés comme tels aux enfants.

N.0.11 Convention signée entre la Direction Académique de Caen et l'Université de Paris pour une intervention en classe)

2. Accords et personnes concernées

La réalisation de ces tâches est soumise à la contrainte d'un accord préalable cumulé : des services de l'éducation nationale et de la direction académique de la DSDEN, de l'établissement scolaire considéré, des enseignant.e.s de la classe d'appartenance de l'enfant, des représentant.e.s de l'autorité légale de l'enfant, et enfin de l'enfant. Si l'ensemble des autres intervenant.e.s a donné son accord écrit et éclairé, mais si l'enfant manifeste par un quelconque signe sa volonté de l'arrêter, le doctorant impliqué dans le recueil de données cessera immédiatement l'expérience, sans pénalité pour qui que ce soit.

Les parents recevront, en amont de ces tâches, un mot du type :

Pour information :

M. Gabriel Thiberge, doctorant en sciences du langage à l'université Paris Diderot, viendra prochainement à l'école pour faire participer individuellement ou en petit groupe les élèves dont les parents sont d'accord à des activités ludiques, sur deux sessions différentes.

Ces activités ludiques auront pour but d'analyser et de comprendre l'acquisition et le développement du langage chez les enfants de 4 à 11 ans. Pour la première session, il s'agira d'écouter des phrases et de les attribuer à un personnage dessiné parmi plusieurs. Pour la seconde session, il s'agira de jouer le rôle d'un serveur ou d'une serveuse de restaurant face à des clients. Les phrases produites par les enfants qui participeront à la seconde session seront enregistrées, mais resteront anonymes.

Bien sûr, les enfants pourront ne pas participer s'ils ne le souhaitent pas. Les sessions ludiques pourront être interrompues à tout moment si l'enfant le souhaite.

En donnant votre autorisation, vous conservez un droit d'accès et de rétractation des données recueillies sur simple demande. Si vous souhaitez plus d'informations ou désirez avoir connaissance des résultats globaux, vous pouvez contacter M. Thiberge par mail à l'adresse gtac@tuta.io.

Restant à votre disposition pour toutes informations complémentaires, nous vous demandons de bien vouloir remplir l'autorisation ci-dessous.

Bien cordialement,




université
PARIS
DIDEROT



U-S-PC
Université Sorbonne
Paris Cité



cnrs



Laboratoire de linguistique formelle

L'équipe pédagogique

Je, soussigné.e autorise mon fils / ma fille à participer aux activités ludiques organisées dans le cadre de l'étude sur le langage, les ... / ... et ... / ... 2019.

→ à la première session : oui non (entourer le choix)

→ à la seconde session : oui non (entourer le choix)

NOM et Prénom de l'enfant : _____ (obligatoire)
 DATE DE NAISSANCE : _____ (obligatoire)

Signature des parents

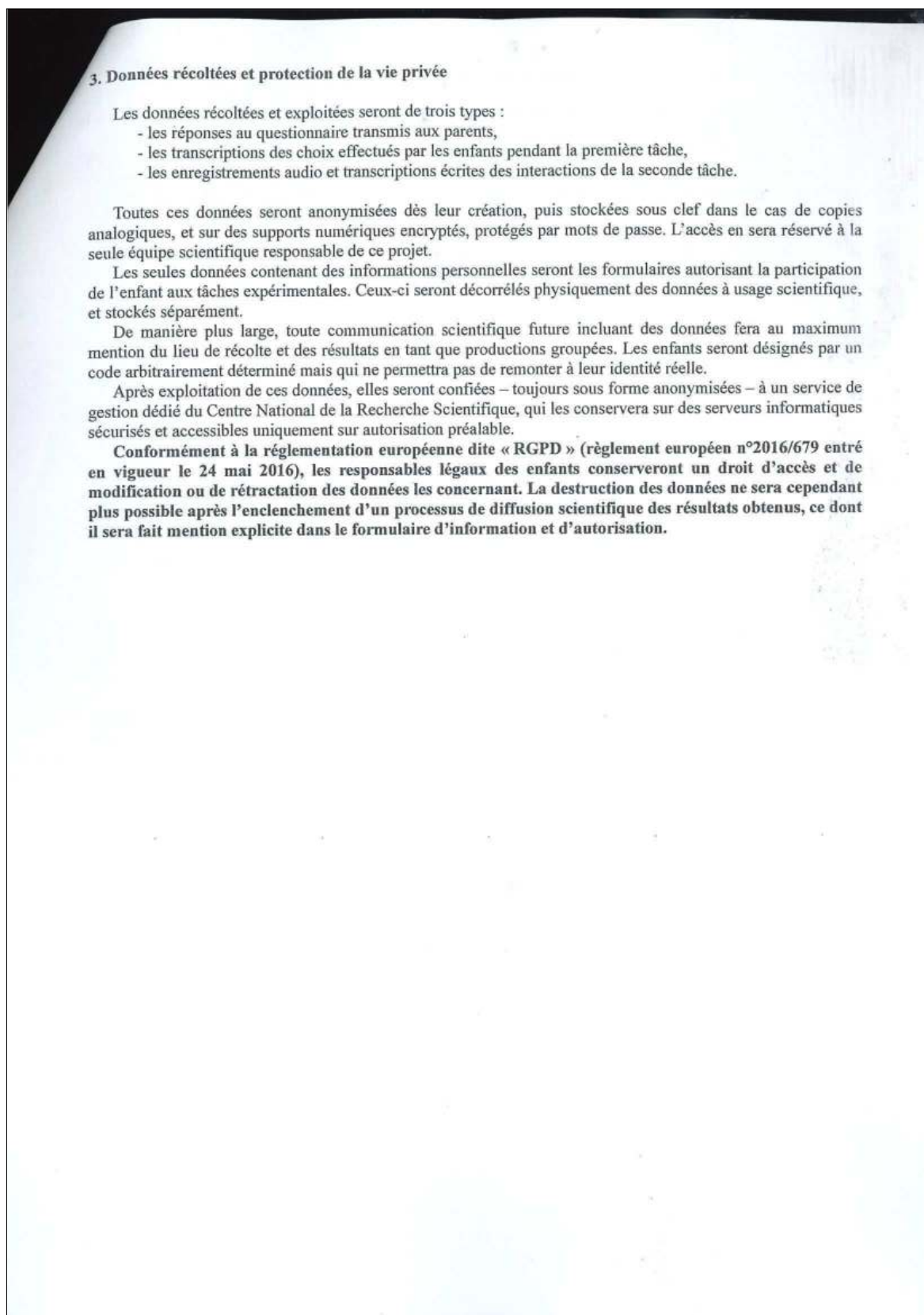
Du côté scientifique, la réalisation de ce projet est en cours de validation par un comité d'éthique (CER de l'Université Paris Descartes) dont la mission est de statuer sur l'acceptabilité des contraintes qu'il fait peser sur toutes les personnes concernées, ici avant tout les enfants. Le projet ne démarrera pas sans cet accord supplémentaire.

N.0.12 Convention signée entre la Direction Académique de Caen et l'Université de Paris pour une intervention en classe)

Critères d'inclusion, d'exclusion et de non-inclusion

Malgré les objectifs numériques indiqués pour les deux tâches présentées (justifiés par un besoin de validité statistique des analyses), l'intégralité des élèves dont la participation a été autorisée participeront aux études. Il n'y a pas de critères d'exclusion autre que la classe d'appartenance, de la Grande Section au CM2). Un traitement a posteriori sera effectué pour tenir compte des troubles apparents (de l'attention, du comportement, langagiers) repérés pendant ou après les passations, et les données liées aux participations non-retenues seront détruites.

N.0.13 Convention signée entre la Direction Académique de Caen et l'Université de Paris pour une intervention en classe)



N.0.14 Convention signée entre la Direction Académique de Caen et l'Université de Paris pour une intervention en classe)

4. Fin des tâches et modalités de restitution des résultats

Les enfants participant aux expériences recevront un petit diplôme attestant de leur contribution à la science, en remerciement de leur implication.

Les parents et les équipes pédagogiques pourront assister à une réunion grand public organisée à leur attention, pour présenter les enjeux, résultats et potentielles implications dans leur vie quotidienne, des recherches permises par leur implication. Sur simple demande lors de cette réunion ou par la suite, une version des productions scientifiques (actes de conférences, article de revue...) mentionnant les résultats obtenus par ces tâches leur sera communiquée.

La direction-académique DSDEN recevra également une version de ces productions scientifiques. D'autres modalités de restitution peuvent bien sûr être envisagées et discutées.

N.0.15 Avis favorable du CER de l'Université Paris Descartes)



CER-Paris Descartes
(Comité d'Éthique de la Recherche)

Présidente: Jacqueline Fagard

**N° 2019- 19-
THIBERGE-
HEMFORTH**

PROTOCOLE : Mesure de l'acquisition de la variation sociolinguistique par les enfants francophones

Noms du/ des chercheur(s) : Gabriel THIBERGE, Barbara HEMFORTH

Email pour la correspondance : gtac@tuta.io; barbara.hemforth@linguist.univ-paris-diderot.fr

Labo/ Service: Laboratoire de Linguistique Formelle; université Paris-Diderot

Évalué à la séance du 28/03/2019

AVIS : Favorable

Les investigateurs faisant appel aux services du CER-Paris Descartes s'engagent à lui signaler tout événement non anticipé survenant en cours d'étude. Ces éléments seront utilisés aux fins d'amélioration des futurs services et conseils que le CER-Paris Descartes pourrait donner.

L'avis du CER-Paris Descartes n'exonère pas des formalités réglementaires. A cet égard, il vous appartient notamment, si vous traitez des données se rapportant à un individu directement ou indirectement identifiable, de vous conformer au règlement européen sur la protection des données (RGPD) en vigueur depuis 2018. Pour cela, vous pouvez solliciter les conseils du Correspondant informatique et libertés (CIL) ou du service juridique de votre université ou de votre organisme de recherche.

N°IRB : 00012019 - 19

Jacqueline Fagard

email : cer_parisdescartes@services.cnrs.fr ; jacqueline.fagard@parisdescartes.fr ; david.brami@parisdescartes.fr

Annexe O

Production Enfantine : Corpus

O.0.1 Diplômes donnés aux enfants après le jeu de rôles, 6.3



FIGURE O.1 – Diplôme de participation au jeu de rôle (F)



FIGURE O.2 – Diplôme de participation au jeu de rôle (M)

O.0.2 Données générales sur les enregistrements des jeux de rôles - détail, 6.3

Enfant	Durée	Niveau	Classe	Âge	Genre	Nombre de mots		Phases	Mots par phase	
						CHI	ADU		CHI	ADU
CHI01	00:12:35	GS	GS-1	4-6	masc	34	502	187	0.18	2.68
CHI02	00:13:16	GS	GS-1	4-6	masc	38	823	204	0.19	4.03
CHI03	00:18:36	GS	GS-1	4-6	fém	186	832	252	0.74	3.3
CHI04	00:12:22	GS	GS-1	4-6	fém	0	644	112	0	5.75
CHI05	00:11:35	GS	GS-1	4-6	masc	46	758	167	0.28	4.54
CHI06	00:16:01	GS	GS-1	4-6	masc	432	595	249	1.73	2.39
CHI08	00:12:00	GS	GS-1	4-6	masc	173	697	193	0.9	3.61
CHI09	00:12:49	GS	GS-1	4-6	masc	73	706	191	0.38	3.7
CHI10	00:08:49	GS	GS-1	4-6	fém	238	502	143	1.66	3.51
CHI11	00:07:34	GS	GS-1	4-6	fém	12	460	101	0.12	4.55
CHI12	00:13:34	GS	GS-1	4-6	fém	1198	609	280	4.28	2.18
CHI13	00:06:00	GS	GS-2	4-6	fém	0	418	84	0	4.98
CHI14	00:08:12	GS	GS-2	4-6	fém	65	531	138	0.47	3.85
CHI15	00:06:55	GS	GS-2	4-6	masc	18	362	119	0.15	3.04
CHI16	00:07:30	GS	GS-2	4-6	masc	55	433	136	0.4	3.18
CHI17	00:05:29	GS	GS-2	4-6	masc	4	401	73	0.05	5.49
CHI18	00:07:17	GS	GS-2	4-6	fém	0	618	134	0	4.61
CHI19	00:10:06	GS	GS-2	4-6	masc	191	696	216	0.88	3.22
CHI20	00:08:27	CP	CP	6	fém	39	497	134	0.29	3.71
CHI21	00:09:05	CP	CP	6	masc	39	446	164	0.24	2.72
CHI22	00:10:35	CP	CP	6	fém	115	467	154	0.75	3.03
CHI23	00:12:00	CP	CP	6	fém	67	441	136	0.49	3.24
CHI24	00:10:59	CP	CP	6	fém	175	314	171	1.02	1.84
CHI25	00:05:57	CP	CP	6	fém	3	263	71	0.04	3.7
CHI26	00:06:10	CP	CP	6	masc	12	327	91	0.13	3.59
CHI27	00:11:49	CP	CP	6	masc	67	531	196	0.34	2.71
CHI28	00:15:28	CP	CP	6	masc	390	768	336	1.16	2.29
CHI29	00:16:48	CP	CP	6	masc	258	668	230	1.12	2.9
CHI30	00:21:53	CP	CP	6	masc	661	1025	453	1.46	2.26
CHI31	00:17:22	CP	CP	6	fém	337	797	317	1.06	2.51
CHI32	00:11:49	CP	CP	6	fém	76	657	179	0.42	3.67
CHI33	00:33:05	CP	CP	6	fém	1516	1195	636	2.38	1.88
CHI34	00:15:26	CP	CP-CE1	6	fém	295	928	279	1.06	3.33
CHI35	00:31:37	CP	CP-CE1	6	masc	398	914	472	0.84	1.94
CHI36	00:11:04	CP	CP-CE1	6	fém	94	574	189	0.5	3.04
CHI37	00:15:48	CE1	CP-CE1	7	fém	340	667	313	1.09	2.13
CHI38	00:16:31	CE2	CE2-CM1	8	masc	464	1072	429	1.08	2.5
CHI39	00:16:24	CE2	CE2-CM1	8	masc	174	786	305	0.57	2.58
CHI40	00:24:00	CE2	CE2-CM1	8	masc	932	1387	631	1.48	2.2

TABLEAU O.1 – Métadonnées du corpus de Jeux de rôles - CHI 01-40

ANNEXE O. PRODUCTION ENFANTINE : CORPUS

Enfant	Durée	Niveau	Classe	Âge	Genre	Nombre de mots		Phases	Mots par phase	
						CHI	ADU		CHI	ADU
CHI41	00:16:30	CE2	CE2-CM1	8	masc	395	629	312	1.27	2.02
CHI42	00:25:57	CE2	CE2-CM1	8	fém	491	732	425	1.16	1.72
CHI43	00:16:38	CE2	CE2-CM1	8	masc	432	725	333	1.3	2.18
CHI44	00:21:07	CE2	CE2-CM1	8	fém	396	926	404	0.98	2.29
CHI45	00:20:55	CE1	CE1-CE2	7	fém	51	582	165	0.31	3.53
CHI46	00:25:45	CE1	CE1-CE2	7	masc	1319	959	700	1.88	1.37
CHI47	00:11:51	CE1	CE1-CE2	7	fém	115	668	210	0.55	3.18
CHI48	00:11:40	CE1	CE1-CE2	7	masc	179	639	246	0.73	2.6
CHI49	00:18:26	CE1	CE1-CE2	7	masc	626	788	463	1.35	1.7
CHI50	00:12:31	CE1	CE1-CE2	7	masc	327	682	318	1.03	2.14
CHI51	00:17:46	CE2	CE1-CE2	8	masc	323	770	329	0.98	2.34
CHI52	00:18:47	CE2	CE1-CE2	8	masc	797	934	527	1.51	1.77
CHI53	00:13:11	CE2	CE1-CE2	8	masc	225	765	287	0.78	2.67
CHI54	00:14:51	CE2	CE1-CE2	8	fém	438	877	372	1.18	2.36
CHI55	00:16:13	CE2	CE1-CE2	8	masc	738	884	463	1.59	1.91
CHI56	00:08:40	CE2	CE1-CE2	8	masc	76	480	148	0.51	3.24
CHI57	00:20:02	CE1	CP-CE1	7	masc	517	941	451	1.15	2.09
CHI58	00:17:38	CE1	CP-CE1	7	masc	540	771	389	1.39	1.98
CHI59	00:10:40	CE1	CP-CE1	7	fém	49	709	207	0.24	3.43
CHI60	00:08:54	CE1	CP-CE1	7	masc	122	599	204	0.6	2.94
CHI61	00:10:29	CE1	CP-CE1	7	masc	264	664	290	0.91	2.29
CHI62	00:10:10	CE1	CP-CE1	8	fém	313	655	349	0.9	1.88
CHI63	00:12:11	CE1	CP-CE1	7	fém	181	690	269	0.67	2.57
CHI64	00:19:31	CE1	CP-CE1	7	fém	417	783	433	0.96	1.81
CHI65	00:12:25	CE2	CE2-CM1	8	fém	268	786	286	0.94	2.75
CHI66	00:14:19	CE2	CE2-CM1	8	fém	428	787	365	1.17	2.16
CHI67	00:07:42	CE2	CE2-CM1	8	masc	114	464	193	0.59	2.4
CHI68	00:19:19	CM1	CE2-CM1	9	masc	434	675	370	1.17	1.82
CHI69	00:09:44	CM1	CE2-CM1	9	fém	281	571	267	1.05	2.14

TABLEAU O.2 – Métadonnées du corpus de Jeux de rôles - CHI 41-69

	Durée	Niveau	Classe	Âge	Genre	Nombre de mots		Phases	Mots par phase	
						CHI	ADU		CHI	ADU
Total	15:56:49					20071	46446	18640		
Moyenne	00:14:04					295.16	683.03	274.12	0.86	2.85
Minimum	00:05:29					0	263	71	0	1.37
Maximum	00:33:05					1516	1387	700	4.28	5.75

TABLEAU O.3 – Métadonnées du corpus de Jeux de rôles - Global

O.0.3 Interrogatives retenues pour analyse, 6.3

Ref	Child	Class	Age	Gender	Type	Puppet	Extraction
422	CHI03	gs1	na	fem	is	MARIE	c'est quoi
457	CHI03	gs1	na	fem	is	MARIE	tu veux quoi maintenant
466	CHI03	gs1	na	fem	is	MARIE	tu veux quoi en dessert
538	CHI03	gs1	na	fem	is	REINE	tu veux quoi à manger
1006	CHI06	gs1	na	male	is	JEANNOT	tu veux quoi
1017	CHI06	gs1	na	male	fesk	JEANNOT	ah qu'est-ce tu fais
1589	CHI10	gs1	na	fem	is	MARIE	tu veux quoi
1998	CHI12	gs1	na	fem	fs	REINE	où elle est ta licorne je la vois pas
2055	CHI12	gs1	na	fem	fs	REINE	où il est ton rayo ton royaume
3398	CHI22	cpcp	6	fem	fesk	REINE	qu'est-ce que vous voulez monsieur
3536	CHI23	cpcp	6	fem	is	REINE	vous voulez manger quoi
3596	CHI24	cpcp	6	fem	is	MARIE	vous voulez quoi
3603	CHI24	cpcp	6	fem	is	MARIE	voulez quoi
3618	CHI24	cpcp	6	fem	is	MARIE	voulez quoi
3680	CHI24	cpcp	6	fem	is	REINE	vous voulez quoi
4141	CHI28	cpcp	6	male	fesk	JEANNOT	qu'est-ce qu'y a
4177	CHI28	cpcp	6	male	is	JEANNOT	euh ça c'est quoi
4194	CHI28	cpcp	6	male	is	JEANNOT	c'est quoi ça
4346	CHI28	cpcp	6	male	is	JEANNOT	et ça j'le mets où au fait
4385	CHI28	cpcp	6	male	is	ROI	vous voulez quoi
4449	CHI28	cpcp	6	male	is	ROI	c'est quoi ça
4780	CHI30	cpcp	6	male	finv	ROI	que désirez-vous avoir
4894	CHI30	cpcp	6	male	finv	REINE	que voulez-vous
4897	CHI30	cpcp	6	male	finv	REINE	que voulez-vous comme euh repas
4911	CHI30	cpcp	6	male	is	REINE	vous vouliez quoi avec les pommes de terre
4986	CHI30	cpcp	6	male	fesk	REINE	qu'est-ce qu'on a d'autre euh
5011	CHI30	cpcp	6	male	fesk	MARIE	qu'est-ce que que vous désiriez vous aller avoir
5018	CHI30	cpcp	6	male	finv	MARIE	que voulez-vous avec
5037	CHI30	cpcp	6	male	is	MARIE	vous voulez quoi avec les frites
5067	CHI30	cpcp	6	male	fesk	MARIE	qu'est-ce que vous voulez comme dessert par contre
5110	CHI30	cpcp	6	male	fesk	MARIE	bien sûr qu'est-ce que vous voulez
5223	CHI31	cpcp	6	fem	is	REINE	tu veux quoi
5233	CHI31	cpcp	6	fem	is	REINE	tu veux quoi
5324	CHI31	cpcp	6	fem	is	ROI	vous voulez quoi monsieur
5388	CHI31	cpcp	6	fem	is	ROI	il est où le sel
5415	CHI31	cpcp	6	fem	is	JEANNOT	tu veux quoi je
5416	CHI31	cpcp	6	fem	is	JEANNOT	tu veux quoi
5428	CHI31	cpcp	6	fem	is	JEANNOT	t'en veux combien de frites
5650	CHI33	cpcp	6	fem	fesk	MARIE	qu'est-ce que tu veux
5675	CHI33	cpcp	6	fem	fs	MARIE	mais pourquoi ils t'ont poursuie
5683	CHI33	cpcp	6	fem	is	MARIE	tu veux quoi en dessert

TABLEAU O.4 – Interrogatives du Jeu de Rôle enfantin retenues pour analyse

Ref	Child	Class	Age	Gender	Type	Puppet	Extraction
5758	CHI33	cpcp	6	fem	fesk	REINE	qu'est-ce que vous voulez
5781	CHI33	cpcp	6	fem	is	REINE	vous voulez quoi
5834	CHI33	cpcp	6	fem	is	REINE	vous voulez quoi
5904	CHI33	cpcp	6	fem	is	JEANNOT	vous voulez quoi
5916	CHI33	cpcp	6	fem	fs	JEANNOT	bah où elle est ma salade
5918	CHI33	cpcp	6	fem	fs	JEANNOT	mais où y sont
6102	CHI33	cpcp	6	fem	is	ROI	vous voulez quoi j'ai tout ça
6113	CHI33	cpcp	6	fem	fs	ROI	alors où elles sont mes pommes de terre
6163	CHI33	cpcp	6	fem	is	ROI	comme accompagnement vous voulez quoi
6210	CHI33	cpcp	6	fem	is	ROI	vous voulez quoi en à boire
6236	CHI33	cpcp	6	fem	fs	ROI	et les tomates comment elles sont
6243	CHI33	cpcp	6	fem	is	ROI	vous voulez quoi en dessert
6246	CHI33	cpcp	6	fem	is	ROI	vous voulez de quelle couleur
6573	CHI35	cpce1	6	male	is	JEANNOT	tu veux quoi
6583	CHI35	cpce1	6	male	is	JEANNOT	et en dessert il fallait quoi
6828	CHI35	cpce1	6	male	is	REINE	euh vous voulez quoi
6892	CHI35	cpce1	6	male	is	REINE	c'est quoi ça
7077	CHI36	cpce1	6	fem	is	MARIE	on peut y acheter quoi comme vêtements
7254	CHI37	cpce1	7	fem	is	MARIE	alors elles sont où les tasses
7311	CHI37	cpce1	7	fem	is	REINE	vous voulez quoi
7388	CHI37	cpce1	7	fem	is	ROI	vous voulez quoi à manger
7489	CHI37	cpce1	7	fem	is	JEANNOT	vous voulez quoi d'autre
7498	CHI37	cpce1	7	fem	is	JEANNOT	vous voulez quoi en dessert
7714	CHI46	ce1ce2	7	male	fesk	JEANNOT	qu'est-ce que tu veux
7744	CHI46	ce1ce2	7	male	is	JEANNOT	vous voulez quoi
7886	CHI46	ce1ce2	7	male	fesk	ROI	qu'est-ce que vous voulez
7889	CHI46	ce1ce2	7	male	fesk	ROI	qu'est-ce que vous voulez
7986	CHI46	ce1ce2	7	male	is	ROI	vous voulez un quoi ensuite
8093	CHI46	ce1ce2	7	male	fesk	REINE	qu'est-ce vous xxx proposer
8285	CHI46	ce1ce2	7	male	is	JEANNOT	vous c'est quoi
8292	CHI46	ce1ce2	7	male	is	JEANNOT	c'est quoi
8318	CHI46	ce1ce2	7	male	is	JEANNOT	c'est quoi
8325	CHI46	ce1ce2	7	male	is	JEANNOT	c'est qui
8714	CHI48	ce1ce2	7	male	fesk	ROI	qu'est-ce que vous voulez
8814	CHI48	ce1ce2	7	male	is	REINE	vous voulez quoi
8874	CHI49	ce1ce2	7	male	is	JEANNOT	vous désirez quoi
8973	CHI49	ce1ce2	7	male	is	ROI	vous désirez quoi
8980	CHI49	ce1ce2	7	male	is	ROI	et vous désirez quoi d'autre
8988	CHI49	ce1ce2	7	male	is	ROI	et à manger vous désirez quoi
8999	CHI49	ce1ce2	7	male	is	ROI	vous voulez quoi comme dessert
9095	CHI49	ce1ce2	7	male	is	REINE	vous désirez quoi
9127	CHI49	ce1ce2	7	male	is	REINE	euh vous désirez quoi
9149	CHI49	ce1ce2	7	male	is	REINE	vous voulez quoi d'autre en dessert
9185	CHI49	ce1ce2	7	male	is	MARIE	bon euh vous désirez quoi
9209	CHI49	ce1ce2	7	male	is	MARIE	vous voulez quoi en dessert
9231	CHI49	ce1ce2	7	male	is	MARIE	vous désirez quoi d'autre

TABLEAU O.4 – Interrogatives du Jeu de Rôle enfantin retenues pour analyse

Ref	Child	Class	Age	Gender	Type	Puppet	Extraction
9298	CHI49	ce1ce2	7	male	is	MARIE	vous voulez quoi
9360	CHI50	ce1ce2	7	male	is	JEANNOT	vous voulez quoi
9425	CHI50	ce1ce2	7	male	is	ROI	tenez vous voulez quoi
9469	CHI50	ce1ce2	7	male	is	ROI	vous voulez quoi en dessert
9505	CHI50	ce1ce2	7	male	is	MARIE	vous voulez quoi
9530	CHI50	ce1ce2	7	male	is	MARIE	vous voulez quoi en dessert
9564	CHI50	ce1ce2	7	male	is	MARIE	toi t'es où
9593	CHI50	ce1ce2	7	male	is	REINE	vous voulez quoi d'autre avec le hotdog
9617	CHI50	ce1ce2	7	male	is	REINE	vous voulez quoi en dessert
9656	CHI57	cpce1	7	male	is	JEANNOT	du coup vous voulez quoi
9745	CHI57	cpce1	7	male	is	ROI	bah vous voulez quoi
9835	CHI57	cpce1	7	male	is	MARIE	vous voulez quoi
9841	CHI57	cpce1	7	male	is	MARIE	vous en voulez combien
9960	CHI57	cpce1	7	male	is	REINE	donc vous voulez quoi
10074	CHI57	cpce1	7	male	is	REINE	vous voulez quoi d'autre
10112	CHI58	cpce1	7	male	fesk	JEANNOT	qu'est-ce que vous voudrez
10184	CHI58	cpce1	7	male	is	JEANNOT	tu voudrais quoi comme dessert
10216	CHI58	cpce1	7	male	is	ROI	vous voulez quoi
10233	CHI58	cpce1	7	male	fs	ROI	comment on appelle déjà
10242	CHI58	cpce1	7	male	fs	ROI	comment il a fait pour se faire très mal
10249	CHI58	cpce1	7	male	fesk	ROI	après qu'est-ce que vous voulez d'autre
10313	CHI58	cpce1	7	male	fesk	MARIE	et qu'est-ce que tu voudrais
10336	CHI58	cpce1	7	male	fesk	MARIE	en dessert qu'est-ce que tu voudrais
10357	CHI58	cpce1	7	male	fesk	MARIE	après qu'est-ce que j'pourrais
10451	CHI58	cpce1	7	male	fesk	REINE	qu'est-ce que tu voudrais comme légumes
10469	CHI58	cpce1	7	male	fesk	REINE	et qu'est-ce que tu voudrais d'autre après
10705	CHI60	cpce1	7	male	is	JEANNOT	tu veux quoi
10780	CHI60	cpce1	7	male	is	ROI	vous voulez quoi
10852	CHI60	cpce1	7	male	is	MARIE	tu veux quoi
10910	CHI61	cpce1	7	male	is	JEANNOT	tu veux quoi
10920	CHI61	cpce1	7	male	is	JEANNOT	tu veux quoi comme sauce
10933	CHI61	cpce1	7	male	is	JEANNOT	tu veux quoi comme dessert
10959	CHI61	cpce1	7	male	is	JEANNOT	tu veux quoi comme boisson
11015	CHI61	cpce1	7	male	fesk	ROI	qu'est-ce que vous voulez monsieur
11030	CHI61	cpce1	7	male	is	ROI	vous voulez quoi comme dessert
11055	CHI61	cpce1	7	male	is	ROI	vous voulez quoi
11102	CHI61	cpce1	7	male	fesk	MARIE	qu'est-ce que vous voulez
11123	CHI61	cpce1	7	male	is	MARIE	ah vous voulez combien d'saucisses
11160	CHI61	cpce1	7	male	is	MARIE	euh tu veux quoi en dessert
11174	CHI61	cpce1	7	male	is	MARIE	tu veux quoi en boisson
11202	CHI62	cpce1	8	fem	is	MARIE	tu veux quoi
11243	CHI62	cpce1	8	fem	is	MARIE	tu veux quoi
11282	CHI62	cpce1	8	fem	is	REINE	vous voulez quoi
11287	CHI62	cpce1	8	fem	is	REINE	et vous voulez quoi
11376	CHI62	cpce1	8	fem	is	JEANNOT	tu veux quoi
11378	CHI62	cpce1	8	fem	is	JEANNOT	euh c'est quoi

TABLEAU O.4 – Interrogatives du Jeu de Rôle enfantin retenues pour analyse

Ref	Child	Class	Age	Gender	Type	Puppet	Extraction
11383	CHI62	cpce1	8	fem	is	JEANNOT	tu veux quoi d'autre
11391	CHI62	cpce1	8	fem	is	JEANNOT	c'est quoi ça
11452	CHI62	cpce1	8	fem	is	ROI	vous voulez quoi
11550	CHI63	cpce1	7	fem	finv	MARIE	que voulez-vous
11564	CHI63	cpce1	7	fem	fesk	MARIE	qu'est-ce qui t'est arrivé
11616	CHI63	cpce1	7	fem	finv	REINE	que voulez-vous
11684	CHI63	cpce1	7	fem	fesk	JEANNOT	qu'est-ce que vous faites
11686	CHI63	cpce1	7	fem	is	JEANNOT	et vous voulez manger quoi
11731	CHI63	cpce1	7	fem	is	JEANNOT	tu veux quoi comme dessert
11760	CHI63	cpce1	7	fem	finv	ROI	que voulez-vous
11827	CHI64	cpce1	7	fem	is	MARIE	tu veux quoi à manger
11870	CHI64	cpce1	7	fem	is	MARIE	tu veux quoi en dessert
11934	CHI64	cpce1	7	fem	is	REINE	vous voulez quoi
11954	CHI64	cpce1	7	fem	is	REINE	vous voulez quoi comme légumes
12043	CHI64	cpce1	7	fem	is	JEANNOT	vous voulez quoi
12057	CHI64	cpce1	7	fem	is	JEANNOT	vous voulez quoi d'autre sinon
12113	CHI64	cpce1	7	fem	is	JEANNOT	vous voulez quoi comme dessert
12175	CHI64	cpce1	7	fem	is	ROI	vous voulez quoi comme manger
12254	CHI38	ce2cm1	8	male	is	JEANNOT	vous voulez manger quoi
12265	CHI38	ce2cm1	8	male	is	JEANNOT	il t'est arrivé quoi
12272	CHI38	ce2cm1	8	male	is	JEANNOT	et tu veux quoi
12284	CHI38	ce2cm1	8	male	is	JEANNOT	tu veux quoi
12333	CHI38	ce2cm1	8	male	is	ROI	vous voulez quoi
12447	CHI38	ce2cm1	8	male	is	REINE	vous voulez quoi
12569	CHI38	ce2cm1	8	male	is	MARIE	vous voulez quoi
12606	CHI38	ce2cm1	8	male	is	MARIE	vous préférez qui
12646	CHI38	ce2cm1	8	male	is	MARIE	et vous jouez où
12655	CHI38	ce2cm1	8	male	is	MARIE	ah et elle habite où
12684	CHI39	ce2cm1	8	male	fesk	JEANNOT	ke qu'est-ce que vous voulez
12732	CHI39	ce2cm1	8	male	is	JEANNOT	tu veux quoi
12755	CHI39	ce2cm1	8	male	fesk	ROI	qu'est-ce que vous voulez
12915	CHI39	ce2cm1	8	male	fesk	MARIE	qu'est-ce que tu veux
12991	CHI40	ce2cm1	8	male	is	JEANNOT	et jvous sers quoi
13009	CHI40	ce2cm1	8	male	fesk	JEANNOT	qu'est-ce qui vous esT arrivé c'matin
13027	CHI40	ce2cm1	8	male	is	JEANNOT	vous voulez quoi en dessert
13090	CHI40	ce2cm1	8	male	fesk	JEANNOT	qu'est-ce que tu pourrais aimer
13115	CHI40	ce2cm1	8	male	is	JEANNOT	à votre avis c'est quoi mon club
13164	CHI40	ce2cm1	8	male	is	ROI	ah bah alors vous voulez quoi
13173	CHI40	ce2cm1	8	male	fesk	ROI	qu'est-ce qu'on a d'autre
13183	CHI40	ce2cm1	8	male	is	ROI	je vous sers quoi comme boisson
13193	CHI40	ce2cm1	8	male	is	ROI	alors vous voulez quoi
13282	CHI40	ce2cm1	8	male	is	REINE	je vous sers quoi du coup
13351	CHI40	ce2cm1	8	male	is	REINE	et c'est quoi votre joueur préféré
13363	CHI40	ce2cm1	8	male	is	REINE	je vous sers quoi
13421	CHI40	ce2cm1	8	male	is	MARIE	voulez quoi
13470	CHI40	ce2cm1	8	male	is	MARIE	vous voulez quoi d'autre en dessert

TABLEAU O.4 – Interrogatives du Jeu de Rôle enfantin retenues pour analyse

Ref	Child	Class	Age	Gender	Type	Puppet	Extraction
13512	CHI40	ce2cm1	8	male	is	MARIE	pour vous faire digérer vous voulez quoi
13533	CHI40	ce2cm1	8	male	is	MARIE	c'est qui
13583	CHI40	ce2cm1	8	male	is	MARIE	et toi c'est co lequel ton joueur préféré préfé- féfé
13620	CHI41	ce2cm1	8	male	is	JEANNOT	tu veux quoi
13680	CHI41	ce2cm1	8	male	is	JEANNOT	tu veux quoi
13702	CHI41	ce2cm1	8	male	fesk	ROI	qu'est-ce que vous voulez
13704	CHI41	ce2cm1	8	male	fesk	ROI	et qu'est-ce que vous voulez
13834	CHI41	ce2cm1	8	male	is	REINE	vous voulez faire quoi
14350	CHI42	ce2cm1	8	fem	is	JEANNOT	c'est à qui
14391	CHI43	ce2cm1	8	male	is	MARIE	vous voulez quoi en dessert
14462	CHI43	ce2cm1	8	male	fesk	REINE	qu'est-ce que jvous sers
14519	CHI43	ce2cm1	8	male	fesk	REINE	qu'est-ce que vous voulez
14569	CHI43	ce2cm1	8	male	fesk	JEANNOT	alors qu'est-ce que jvous sers
14581	CHI43	ce2cm1	8	male	fesk	JEANNOT	qu'est-ce que vous voulez
14623	CHI43	ce2cm1	8	male	fesk	REINE	alors qu'est-ce que jvous sers
14627	CHI43	ce2cm1	8	male	is	REINE	vous voulez quoi comme légumes
14664	CHI43	ce2cm1	8	male	fesk	REINE	qu'est-ce que vous voulez
14730	CHI44	ce2cm1	8	fem	is	MARIE	vous voulez quoi d'autre
14800	CHI44	ce2cm1	8	fem	fesk	REINE	qu'est-ce que vous voulez
14996	CHI44	ce2cm1	8	fem	is	JEANNOT	vous voulez quoi
15054	CHI44	ce2cm1	8	fem	is	JEANNOT	vous voulez quoi d'autre
15111	CHI51	ce1ce2	8	male	fesk	JEANNOT	euhm qu'est-ce qui t'est arrivé c'matin
15267	CHI51	ce1ce2	8	male	is	MARIE	tu voudrais quoi
15434	CHI52	ce1ce2	8	male	is	JEANNOT	bah vous voulez quoi
15447	CHI52	ce1ce2	8	male	is	JEANNOT	eh et vous voulez quoi à boire
15458	CHI52	ce1ce2	8	male	is	JEANNOT	eh mais vous avez fait quoi
15475	CHI52	ce1ce2	8	male	is	JEANNOT	eh mais vous voulez quoi en dessert
15500	CHI52	ce1ce2	8	male	is	JEANNOT	eh mais vous avez envie d'aller où
15558	CHI52	ce1ce2	8	male	is	JEANNOT	c'est quoi
15642	CHI52	ce1ce2	8	male	is	ROI	mais vous voulez quoi
15712	CHI52	ce1ce2	8	male	is	ROI	et vous voulez quoi à boire
15726	CHI52	ce1ce2	8	male	is	ROI	eh mais vous voulez quoi en dessert
15832	CHI52	ce1ce2	8	male	is	MARIE	eh mais vous voulez quoi
15845	CHI52	ce1ce2	8	male	is	MARIE	et vous voulez quoi en plus
15910	CHI52	ce1ce2	8	male	is	MARIE	mais vous voulez quoi en dessert
15960	CHI53	ce1ce2	8	male	is	JEANNOT	tu veux quoi
15974	CHI53	ce1ce2	8	male	fesk	JEANNOT	qu'est-ce qui t'est arrivé
16036	CHI53	ce1ce2	8	male	is	JEANNOT	tu veux quoi
16066	CHI53	ce1ce2	8	male	is	ROI	vous voulez quoi
16083	CHI53	ce1ce2	8	male	is	ROI	vous voulez quoi
16114	CHI53	ce1ce2	8	male	is	ROI	vous voulez quoi
16165	CHI53	ce1ce2	8	male	is	MARIE	tu veux quoi
16215	CHI53	ce1ce2	8	male	is	MARIE	tu veux quoi
16248	CHI54	ce1ce2	8	fem	is	MARIE	tu veux quoi
16267	CHI54	ce1ce2	8	fem	is	MARIE	vous voulez quoi en dessert

TABLEAU O.4 – Interrogatives du Jeu de Rôle enfantin retenues pour analyse

Ref	Child	Class	Age	Gender	Type	Puppet	Extraction
16317	CHI54	ce1ce2	8	fem	is	REINE	vous voulez quoi
16425	CHI54	ce1ce2	8	fem	is	JEANNOT	vous voulez quoi
16500	CHI54	ce1ce2	8	fem	is	JEANNOT	vous voulez quoi en dessert
16537	CHI54	ce1ce2	8	fem	is	ROI	vous voulez quoi
16624	CHI55	ce1ce2	8	male	is	JEANNOT	vous voulez quoi
16669	CHI55	ce1ce2	8	male	is	JEANNOT	ça rssemble à quoi d'la grenadine
16744	CHI55	ce1ce2	8	male	is	ROI	vous voulez quoi
16868	CHI55	ce1ce2	8	male	is	MARIE	vous voulez quoi monsieur
16952	CHI55	ce1ce2	8	male	is	MARIE	vous voulez quoi d'autre
16964	CHI55	ce1ce2	8	male	is	MARIE	ça c'est quoi
17140	CHI56	ce1ce2	8	male	is	ROI	c'est quoi
17166	CHI56	ce1ce2	8	male	is	ROI	ça c'est quoi
17239	CHI65	ce2cm1	8	fem	is	MARIE	vous voulez quoi
17274	CHI65	ce2cm1	8	fem	is	MARIE	vous voulez quoi après
17316	CHI65	ce2cm1	8	fem	is	MARIE	vous voulez quoi
17340	CHI65	ce2cm1	8	fem	is	REINE	alors vous voulez quoi
17366	CHI65	ce2cm1	8	fem	is	REINE	vous voulez quoi après
17381	CHI65	ce2cm1	8	fem	is	REINE	vous préférez qui euh chez vous
17384	CHI65	ce2cm1	8	fem	is	REINE	dans votre famille c'est qui vot préféré
17394	CHI65	ce2cm1	8	fem	is	REINE	vous voulez quoi après à boire
17428	CHI65	ce2cm1	8	fem	is	JEANNOT	vous voulez quoi
17438	CHI65	ce2cm1	8	fem	is	JEANNOT	vous voulez quoi comme sauce
17521	CHI66	ce2cm1	8	fem	is	MARIE	tu veux quoi
17532	CHI66	ce2cm1	8	fem	fesk	MARIE	qu'est-ce qui t'est arrivé
17546	CHI66	ce2cm1	8	fem	fesk	MARIE	qu'est-ce que tu veux à boire
17557	CHI66	ce2cm1	8	fem	is	MARIE	tu veux quoi comme dessert
17600	CHI66	ce2cm1	8	fem	fesk	REINE	qu'est-ce que vous voulez
17651	CHI66	ce2cm1	8	fem	fesk	REINE	et qu'est-ce que vous voulez comme dessert
17733	CHI66	ce2cm1	8	fem	fesk	JEANNOT	et qu'est-ce que tu veux en dessert
17794	CHI66	ce2cm1	8	fem	fesk	ROI	qu'est-ce que vous voulez
17803	CHI66	ce2cm1	8	fem	fesk	ROI	et qu'est-ce que tu veux à boire
17824	CHI66	ce2cm1	8	fem	is	ROI	tu veux quoi comme dessert
17833	CHI66	ce2cm1	8	fem	fesk	ROI	qu'est-ce que tu veux
17865	CHI66	ce2cm1	8	fem	finv	ROI	comment puis-je oublier ce coin de paradis
18078	CHI68	ce2cm1	9	male	fesk	JEANNOT	qu'est-ce que vous voulez
18156	CHI68	ce2cm1	9	male	fesk	ROI	euh bah qu'est-ce que voulez pour euh manger
18229	CHI68	ce2cm1	9	male	is	ROI	elles sont où les tasses
18258	CHI68	ce2cm1	9	male	finv	MARIE	que voulez-vous euh marie
18451	CHI69	ce2cm1	9	fem	is	MARIE	vous voulez manger quoi
18495	CHI69	ce2cm1	9	fem	is	MARIE	vous voulez quoi
18546	CHI69	ce2cm1	9	fem	is	REINE	vous voulez quoi à manger
18570	CHI69	ce2cm1	9	fem	fs	REINE	comment ça s'appelle
18591	CHI69	ce2cm1	9	fem	is	REINE	et vous voulez quoi comme euh dessert
18639	CHI69	ce2cm1	9	fem	is	JEANNOT	vous voulez quoi comme euh à manger
18685	CHI69	ce2cm1	9	fem	is	JEANNOT	vous voulez quoi comme dessert

TABLEAU O.4 – Interrogatives du Jeu de Rôle enfantin retenues pour analyse

Annexe P

Production Enfantine : Statistiques

P0.1 Modèle *brms* et distributions postérieures pour un effet de CONTEXTE sur le type syntaxique (JDR-C, 6.3.5.1)

Family : categorical

Links : mufs = logit; mufesk = logit; mufinv = logit

Formula : ques ~ formal_informal + (1 | child)

Data : DON (Number of observations : 269)

Samples : 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
mufs_Intercept	-13.86	7.95	-35.24	-5.06	1.00	3830	4353
mufesk_Intercept	-2.31	0.56	-3.56	-1.38	1.00	4338	6293
mufinv_Intercept	-12.30	6.58	-29.14	-5.20	1.00	3630	3769
mufs_formal_informal	0.65	0.90	-1.08	2.44	1.00	14924	8410
mufesk_formal_informal	0.40	0.41	-0.40	1.21	1.00	13119	9012
mufinv_formal_informal	2.03	1.06	0.11	4.28	1.00	12098	8034

TABLEAU P.1 – Résumé du modèle *brms* (JDR-C, effet de CONTEXTE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	195	h95	190	h85	185	h85
FS_IS*Formel_Informel	0.65	0.77	0.23	-3.01	5.18	-1.08	2.44	-0.80	2.16	-0.63	1.96
FESK_IS*Formel_Informel	0.40	0.84	0.16	-1.41	1.98	-0.40	1.21	-0.28	1.08	-0.19	0.99
FINV_IS*Formel_Informel	2.03	0.98	0.02	-1.89	7.24	0.11	4.28	0.40	3.86	0.59	3.59

TABLEAU P.2 – Distributions postérieures pour le modèle *brms* (JDR-C, effet de CONTEXTE)

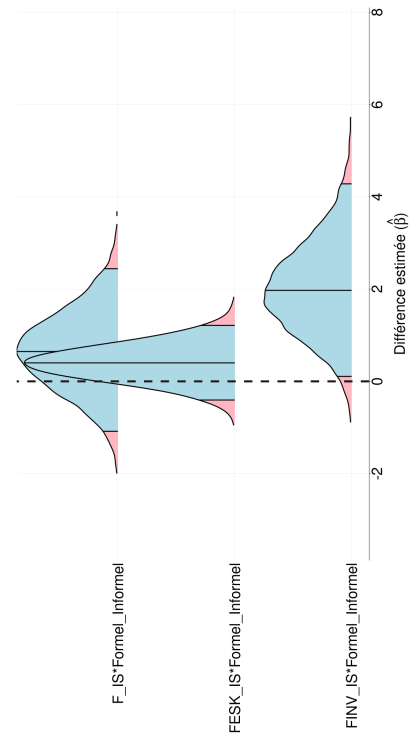


FIGURE P.1 – Distributions postérieures pour le modèle *brms* (JDR-C, effet de CONTEXTE)

P.0.2 Modèle *brms* et distributions postérieures pour une interaction CONTEXTE*CLASSE sur le type syntaxique (JDR-C, 6.3.5.2)

Family : categorical

Links : mufs = logit; mufesk = logit; mufinv = logit

Formula : ques ~ classnum*formal_informal + (1 | child)

Data : DON (Number of observations : 269)

Samples : 4 chains, each with iter = 3000; warmup = 1500; thin = 1; total post-warmup samples = 6000

Population-Level Effects :	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
mufs_Intercept	-6.86	19.28	-51.58	25.24	1.00	2734	2192
mufesk_Intercept	-3.48	1.63	-6.99	-0.56	1.00	3101	3073
mufinv_Intercept	-18.99	14.41	-57.37	-2.64	1.00	2516	1883
mufs_formal_informal	-28.05	26.19	-102.14	-2.98	1.00	1544	988
mufesk_classnum	-10.52	9.58	-36.86	-0.50	1.00	1594	1063
mufinv_formal_informal :classnum	13.35	12.99	1.41	49.91	1.00	1542	996
mufesk_formal_informal	-2.02	1.52	-5.03	0.96	1.00	5194	3899
mufesk_classnum	0.30	0.41	-0.50	1.13	1.00	3061	3189
mufesk_formal_informal :classnum	0.65	0.38	-0.08	1.40	1.00	5268	3723
mufinv_formal_informal	3.88	3.29	-2.07	10.86	1.00	6223	3852
mufinv_classnum	0.78	2.61	-3.80	6.84	1.00	2594	2461
mufinv_formal_informal :classnum	-0.56	0.93	-2.41	1.25	1.00	6541	4060

TABLEAU P.3 – Résumé du modèle *brms* (JDR-C, effet de CLASSE*CONTEXTE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	l95	h95	l90	h90	l85	h85
F_IS*Formel_Informel	-28.05	0.00	1.00	-223.10	3.02	-102.14	-2.98	-78.15	-4.72	-67.17	-5.78
F_IS*Classe	-10.52	0.02	0.98	-86.69	11.21	-36.86	-0.50	-29.17	-1.32	-24.47	-1.85
F_IS*Formel_Informel*Classe	13.35	1.00	0.00	-0.62	110.34	1.41	49.91	2.11	38.43	2.54	32.90
FESK_IS*Formel_Informel	-2.02	0.09	0.91	-8.41	3.58	-5.03	0.96	-4.58	0.43	-4.26	0.11
FESK_IS*Classe	0.30	0.77	0.23	-1.81	2.43	-0.50	1.13	-0.37	0.98	-0.28	0.88
FESK_IS*Formel_Informel*Classe	0.65	0.96	0.04	-0.61	2.24	-0.08	1.40	0.04	1.27	0.11	1.20
FINV_IS*Formel_Informel	3.88	0.88	0.12	-6.16	19.90	-2.07	10.86	-1.21	9.69	-0.70	8.79
FINV_IS*Classe	0.78	0.62	0.38	-16.74	23.17	-3.80	6.84	-2.82	5.11	-2.25	4.28
FINV_IS*Formel_Informel*Classe	-0.56	0.27	0.73	-4.62	2.86	-2.41	1.25	-2.11	0.95	-1.91	0.76

TABLEAU P4 – Distributions postérieures pour le modèle brms (JDR-C, effet de CONTEXTE*CLASSE)

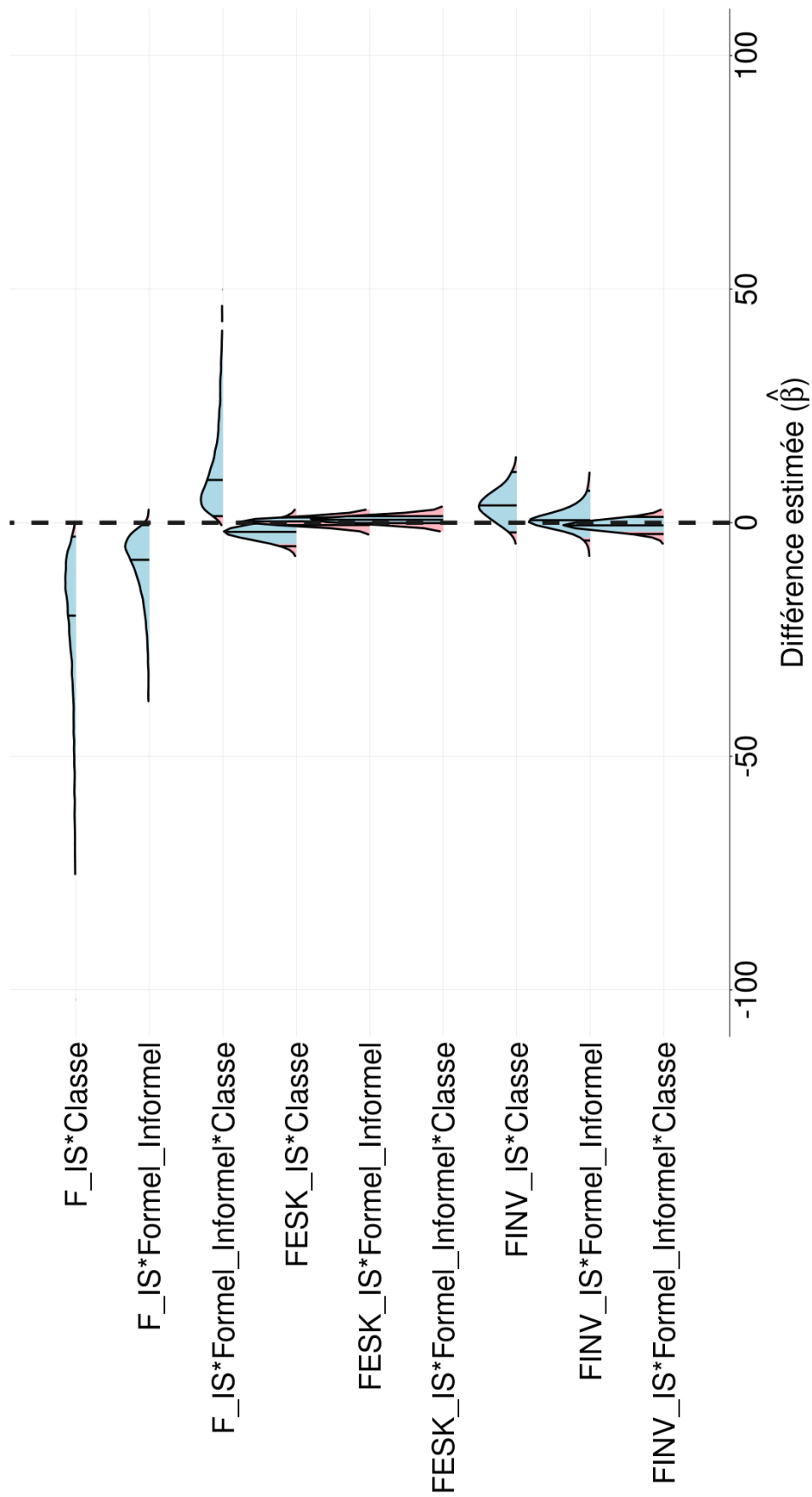


FIGURE P.2 – Distributions postérieures pour le modèle brms (JDR-C, effet de CONTEXTE*CLASSE)

P.0.3 Modèle brms et distributions postérieures pour un effet de GENRE sur le type syntaxique (JDR-C, 6.3.5.3)

Family : categorical
Links : mufs = logit; mufesk = logit; mufinv = logit
Formula : ques ~ mal_fem + (1 | child)
Data : DON (Number of observations : 269)
Samples : 4 chains, each with iter = 6000; warmup = 3000; thin = 1; total post-warmup samples = 12000

Population-Level Effects :

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
mufs_Intercept	-20.95	12.48	-55.24	-6.25	1.00	1377	524
mufesk_Intercept	-2.38	0.61	-3.78	-1.39	1.00	2382	1893
mufinv_Intercept	-14.20	7.67	-35.03	-5.36	1.00	3997	3440
mufs_mal_fem	-9.14	9.43	-33.12	4.28	1.00	2443	2085
mufesk_mal_fem	1.59	1.10	-0.37	3.97	1.00	2961	1947
mufinv_mal_fem	-0.94	5.54	-12.72	10.07	1.00	3812	4799

TABLEAU P.5 – Résumé du modèle brms (JDR-C, effet de GENRE)

Effect	$\hat{\beta}$	P($\beta > 0$)	P($\beta < 0$)	min	max	195	h95	190	h90	185	h85
F_IS*M_F	-9.14	0.09	0.91	-77.29	33.02	-33.12	4.28	-27.92	1.94	-24.20	0.59
FESK_IS*M_F	1.59	0.94	0.06	-2.97	7.21	-0.37	3.97	-0.06	3.48	0.13	3.18
FINV_IS*M_F	-0.94	0.42	0.58	-44.15	31.88	-12.72	10.07	-9.70	7.30	-8.17	5.92

TABLEAU P.6 – Distributions postérieures pour le modèle brms (JDR-C, effet de GENRE)

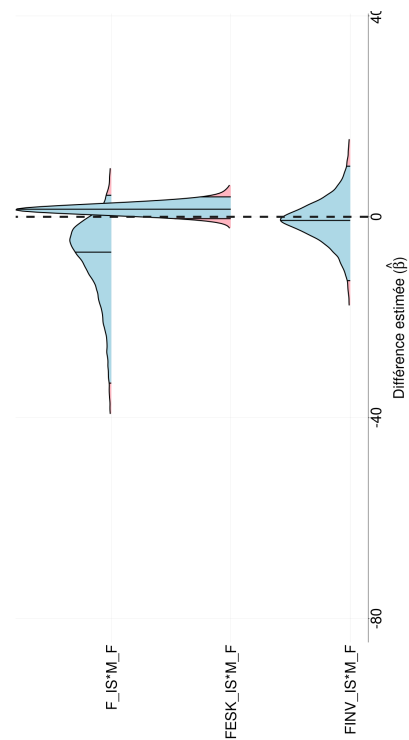


FIGURE P.3 – Distributions postérieures pour le modèle brms (JDR-C, effet de GENRE)

Annexe Q

Liste des acronymes

BAFA Brevet d’Aptitude aux Fonctions d’Animateur. [194](#)

CQL Corpus Query Language. [69](#)

F Antéposition simple, ou *Fronting*. [14](#)

FESK Antéposition en *est-ce que*, ou *Fronting with ESK*. [14](#)

FINV Antéposition avec inversion, ou *Fronting with INVersion*. [14](#)

IS in situ. [14](#)

Annexe R

Liste des symboles

α valeur alpha, seuil au-delà duquel un calcul statistique est jugé significatif. 47

m médiane. 88, 91, 101, 116, 128, 132, 155, 179, XCIII, CXVII, CXXXVII

μ moyenne. 88, 91, 101, 116, 128, 132, 155, 179, XCIII, CXVII, CXXXVII

N nombre d'éléments dans un groupe (participants, items, extractions). 15, 16, 57, 59, 68, 76, 78–80, 88, 140, 143, 177, 180, 181, 183, 184, 187–189, 201

r coefficient de corrélation. 47, 61, 102, 116, 134, 139, 156

Références

- ABEILLÉ, A. 2007, *Les grammaires d'unification*, Hermès, 344 pp.. URL <https://halshs.archives-ouvertes.fr/halshs-00748332>, 978-2-7462-1251-0. 18
- ABEILLÉ, A., B. CRABBÉ, D. GODARD et J.-M. MARANDIN. 2012, "French questioning declaratives : a corpus study", dans *Proceedings of SemDial 2012 (SeineDial) : The 16th Workshop on the Semantics and Pragmatics of Dialogue*, p. 70. URL <http://www.llf.cnrs.fr/Gens/Abeille/AA-BC-DG-JM.pdf>. 3, 14
- ABEILLÉ, A. et D. GODARD. 2012, "La grande grammaire du français et la variété des données", *Langue française*, vol. 4, pp. 47–68. URL <https://www.cairn.info/revue-langue-francaise-2012-4-page-47.htm>. 13
- ABEILLÉ, A. et E. WINCKEL. 2019, "Extracting out of the subject in French : experimental evidence", dans *Proceedings of the First Workshop on Quantitative Syntax (Quasy, SyntaxFest 2019)*, Association for Computational Linguistics, Paris, France, pp. 68–74, doi :10.18653/v1/W19-7908. URL <https://www.aclweb.org/anthology/W19-7908>. 19
- ADLI, A. 2006, "French wh-in-situ questions and syntactic optionality : Evidence from three data types", *Zeitschrift für Sprachwissenschaft*, vol. 25, n° 2, pp. 163–203. URL <https://www.degruyter.com/downloadpdf/journals/zfs/25/2/article-p163.pdf>. 25, 33
- ADLI, A. 2013, "Syntactic variation in french wh-questions : A quantitative study from the angle of bourdieu's sociocultural theory", *Linguistics*, vol. 51, n° 3, doi :<https://doi.org/10.1515/ling-2013-0018>, pp. 473–515. URL https://romanistik.phil-fak.uni-koeln.de/sites/romanistik/Mitarbeiter/Adli/Adli_2013.pdf. 25
- ADLI, A. 2015, "What you like is not what you do : Acceptability and frequency in syntactic variation", *Variation in language : Usage-based vs. system-based approaches*, pp. 173–200. URL https://romanistik.phil-fak.uni-koeln.de/sites/romanistik/Mitarbeiter/Adli/Adli_2015.pdf. 16, 17, 23, 24, 57, 71, 98, 108, 113
- ALMOR, A. 1999, "Noun-phrase anaphora and focus : The informational load hypothesis.", *Psychological review*, vol. 106, n° 4, p. 748. URL <https://www.mccauslandcenter.sc.edu/aLab/sites/sc.edu.alab/files/attachments/almor-1999.pdf>. 22
- AMBRIDGE, B. 2010, "Children's judgments of regular and irregular novel past-tense forms : New data on the english past-tense debate.", *Developmental Psychology*, vol. 46, n° 6, p. 1497. URL <https://pdfs.semanticscholar.org/85d4/935b523ac75cb835536df505ce9f1bc3e10f.pdf>. 223

- AMBRIDGE, B. 2019, “Against stored abstractions : A radical exemplar model of language acquisition”, *First Language*, p. 0142723719869 731. URL <https://journals.sagepub.com/doi/pdf/10.1177/0142723719869731>. 35
- AMBRIDGE, B. et C. F. ROWLAND. 2013, “Experimental methods in studying child language acquisition”, *Wiley Interdisciplinary Reviews : Cognitive Science*, vol. 4, n° 2, pp. 149–168. URL <http://pcwww.liv.ac.uk/~ambridge/Papers/FINALAmbridgeRowland.pdf>. 223
- ANALYTICS, R. et S. WESTON. 2019, *doMC : Foreach Parallel Adaptor for 'parallel'*. URL <https://CRAN.R-project.org/package=doMC>, r package version 1.3.6. ix
- ASHBY, W. 1977, “Interrogative forms in parisian french”, *Semasia*, vol. 4, pp. 35–52. 3, 16, 25
- ATILF, INIST et LIPN. 2012, “Perceo : un projet d’etiqueteur robuste pour l’écrit et pour l’oral”, URL <https://hdl.handle.net/11403/perceo/v1>, ORTOLANG (Open Resources and TOols for LANGuage) –www.ortolang.fr. 69
- ATILF-CNRS. 1998-2020, “Base textuelle frantext”, URL <https://www.frantext.fr>. 23
- AUDACITYTEAM. 1999-2020, “Audacity(r) : Free audio editor and recorder [computer application]”, URL <https://audacityteam.org/>, version 2.4.2, retrieved Sep 20th 2020. 99, CXVII
- BADIN, F., G. THIBERGE, L. LIÉGEOIS et C. PARISSÉ. Accepté, “Prise en main d’un corpus de français oral pour l’analyse des interrogatives partielles en micro-diachronie”, *Corpus*. 68, 69, 223
- BAIÃO, V. et M. F. H. S. L. GONÇALVES. 2014, “Aquisição de interrogativas preposicionadas no português europeu”, *Textos Seleccionados do XXIX Encontro Nacional da APL*, pp. 57–70. URL https://run.unl.pt/bitstream/10362/29170/1/Aquisi_o_de_interrogativas_preposicionadas_em_PE.pdf. 30
- BAIÃO, V. L. M. D. R. 2012, *Aquisição de interrogativas preposicionadas em português europeu*, thèse de doctorat, Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa. URL <https://run.unl.pt/bitstream/10362/11286/1/Tese%20Final%20-%20Vera%20baiao.pdf>. 30
- BARNES, B. 1984, “Minnesota corpus”, *University of Minnesota*. 21
- BARR, D. J., R. LEVY, C. SCHEEPERS et H. J. TILY. 2013, “Random effects structure for confirmatory hypothesis testing : Keep it maximal”, *Journal of memory and language*, vol. 68, n° 3, pp. 255–278. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3881361/>. 49
- BARRAS, C., E. GEOFFROIS, Z. WU et M. LIBERMAN. 2001, “Transcriber : development and use of a tool for assisting speech corpora production”, *Speech Communication*, vol. 33, n° 1-2, pp. 5–22. URL <https://hal.archives-ouvertes.fr/hal-01690349/file/transcriber.pdf>. 69
- BAZILLON, T., B. MAZA, M. ROUVIER, F. BECHET et A. NASR. 2011, “Qui êtes-vous? catégoriser les questions pour déterminer le rôle des locuteurs dans des conversations orales”, dans *Actes des conférences TALN 2011 et Recital 2011*, p. 147. URL <https://hal.archives-ouvertes.fr/hal-01196021>. 41, 53, 54, 88, 113, 217
- BEHNSTEDT, P. 1973, *Viens-tu? Est-ce que tu viens? Tu viens?*, vol. 41, *Tübinger Beiträge zur Linguistik*. 3, 15, 24, 57

- BENZITOUN, C., K. FORT et B. SAGOT. 2012, “Tcof-pos : un corpus libre de français parlé annoté en morphosyntaxe”, dans *JEP-TALN 2012 - Journées d’Études sur la Parole et conférence annuelle du Traitement Automatique des Langues Naturelles*, Grenoble, France, pp. 99–112. URL <https://hal.archives-ouvertes.fr/hal-00709187>. 69
- BEYSSADE, C. 2006, “La structure de l’information dans les questions : quelques remarques sur la diversité des formes interrogatives en français”, *Linx. Revue des linguistes de l’université Paris X Nanterre*, vol. 55, pp. 173–193. URL https://jeannicod.ccsd.cnrs.fr/ijn_00356329/document. 3, 19, 20
- BEYSSADE, C. et J.-M. MARANDIN. 2006, “The speech act assignment problem revisited : Disentangling speaker’s commitment from speaker’s call on addressee”, *Empirical issues in syntax and semantics*, vol. 6, n° 37-68. URL www.cssp.cnrs.fr/eiss6/beyssade-marandin-eiss6.pdf. 3, 11
- BOECKX, C. 1999, “Decomposing french questions”, *University of Pennsylvania Working Papers in Linguistics*, vol. 6, n° 1, p. 6. URL <https://repository.upenn.edu/cgi/viewcontent.cgi?article=1414&context=pwpl>. 3, 19
- BONFANTE, G., B. GUILLAUME et G. PERRIER. 2018, *Application of Graph Rewriting to Natural Language Processing, Logic, Linguistics and Computer Science Set*, vol. 1, ISTE Wiley, 272 pp.. URL <https://hal.inria.fr/hal-01814386>. 41
- BOŠKOVIĆ, Ž. 2002, “On multiple wh-fronting”, *Linguistic inquiry*, vol. 33, n° 3, pp. 351–383. URL <http://semantics.uchicago.edu/kennedy/classes/w07/readings/boskovic02.pdf>. 19
- BOUCHER, P. 2010, “L’interrogation partielle en français : l’interface syntaxe/sémantique”, *Syntaxe et sémantique*, vol. 1, pp. 55–82. URL <https://www.cairn.info/revue-syntaxe-et-semantique-2010-1-page-55.htm>. 20
- BOURDIEU, P. et L. BOLTANSKI. 1975, “Le fétichisme de la langue”, *Actes de la recherche en sciences sociales*, vol. 1, n° 4, pp. 2–32. URL https://www.persee.fr/doc/arss_0335-5322_1975_num_1_4_3417. 26
- BRANCA-ROSOFF, S., S. FLEURY, F. LEFEUVRE et M. PIRES. 2000, “Discours sur la ville”, *Corpus de français parlé parisien des années*, vol. 2009. URL <https://core.ac.uk/download/pdf/48362576.pdf>. 41
- BREHM, L. et P. M. ALDAY. 2020, “A decade of mixed models : It’s past time to set your contrasts”, dans *the 26th Architectures and Mechanisms for Language Processing Conference (AMLap 2020)*, pp. –. URL <https://amlap2020.github.io/a/131.pdf>. 48
- VAN DEN BRINK, D., J. J. VAN BERKUM, M. C. BASTIAANSEN, C. M. TESINK, M. KOS, J. K. BUITELAAR et P. HAGOORT. 2012, “Empathy matters : Erp evidence for inter-individual differences in social language processing”, *Social cognitive and affective neuroscience*, vol. 7, n° 2, pp. 173–183. URL <https://academic.oup.com/scan/article/7/2/173/1623141>. 224

- BUCHOLTZ, M. et K. HALL. 2005, “Identity and interaction : A sociocultural linguistic approach”, *Discourse studies*, vol. 7, n° 4-5, pp. 585–614. URL <https://escholarship.org/content/qt5tk670n8/qt5tk670n8.pdf>. 27
- BÜRKNER, P.-C. 2018, “Advanced bayesian multilevel modeling with the r package brms. the r journal, 10 (1), 395–411. doi : 10.32614”, *cahier de recherche*, RJ-2018-017. URL <https://arxiv.org/pdf/1705.11123.pdf>. ix, 47
- BÜRKNER, P.-C. et al.. 2017, “brms : An r package for bayesian multilevel models using stan”, *Journal of Statistical Software*, vol. 80, n° 1, pp. 1–28. ix, 47
- BURNETT, H. 2017, “Sociolinguistic interaction and identity construction : The view from game-theoretic pragmatics”, *Journal of Sociolinguistics*, vol. 21, n° 2, pp. 238–271. URL http://www.heatherburnett.net/uploads/9/6/6/0/96608942/jofs_burnett.pdf. 5, 27
- BUSON, L. 2009, *Stylistic variation and socialisation networks at school : uses, representations, acquisition and educational application with children aged 5 to 11*, Theses, Université Stendhal - Grenoble III. URL <https://tel.archives-ouvertes.fr/tel-00440455>. 35, 36, 186, 212, 220
- BUSON, L., J.-P. CHEVROT, A. NARDY et M. ABOUZAÏD. 2014, “Catégorisations et représentations sociolinguistiques : les variétés stylistiques existent-elles?”, dans *SHS Web of Conferences*, vol. 8, EDP Sciences, pp. 1407–1418. URL https://www.shs-conferences.org/articles/shsconf/pdf/2014/05/shsconf_cm1f14_01149.pdf. 35, 186, 212, 220
- BUTTELMANN, D., M. CARPENTER et M. TOMASELLO. 2009, “Eighteen-month-old infants show false belief understanding in an active helping paradigm”, *Cognition*, vol. 112, n° 2, pp. 337–342. URL https://www.eva.mpg.de/documents/Elsevier/Buttelmann_Eighteen-month-old_Cognition_2009_1553014.pdf. 34
- BYBEE, J. L. 2013, “Usage-based theory and exemplar representations of constructions”, dans *The Oxford handbook of construction grammar*, –, pp. –. URL <https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780195396683.001.0001/oxfordhb-9780195396683-e-4>. 35, 186, 220
- BYBEE, J. L. et P. J. HOPPER. 2001, *Frequency and the emergence of linguistic structure*, vol. 45, John Benjamins Publishing. URL <https://benjamins.com/catalog/ts1.45>. 35, 186, 220
- CAMPBELL-KIBLER, K. 2007, “Accent,(ing), and the social logic of listener perceptions”, *American speech*, vol. 82, n° 1, pp. 32–64. URL <https://pdfs.semanticscholar.org/25b3/00b036484d51bb3c9b50ba2580316f9c4a20.pdf>. 28, 42, 87
- CARPENTER, B., A. GELMAN, M. D. HOFFMAN, D. LEE, B. GOODRICH, M. BETANCOURT, M. BRUBAKER, J. GUO, P. LI et A. RIDDELL. 2017, “Stan : A probabilistic programming language”, *Journal of statistical software*, vol. 76, n° 1. URL <https://www.osti.gov/servlets/purl/1430202>. ix, 47
- CHANG, L. 1997, *Wh-in-situ phenomena in French*, thèse de doctorat, University of British Columbia. URL <https://open.library.ubc.ca/cIRcle/collections/ubctheses/831/items/1.0087736>. 19

- CHESHIRE, J. 1982, “Variation in an english dialect : A sociolinguistic study”, *Cambridge Studies in Linguistics London*, vol. 37. 27
- CHEVROT, J.-P., A. NARDY et S. BARBU. 2011, “Developmental dynamics of ses-related differences in children’s production of obligatory and variable phonological alternations”, *Language Sciences*, vol. 33, n° 1, pp. 180–191. URL https://hal.archives-ouvertes.fr/file/index/docid/718025/filename/Chevrot_Nardy_et_Barbu_2011_.pdf. 35
- CHOMSKY, N. 1965, “Aspects of the theory of syntax”, *Cambridge, MA : MIT Press*. 30
- CHOMSKY, N. 1977, “On wh-movement. formal syntax, ed. by peter w. culicover, thomas wasow, and adrian akmajian, 71-132”, URL https://babel.ucsc.edu/~hank/On_WH-Movement.pdf. 30
- CHOMSKY, N. 1981, “Lectures on government and binding (dordrecht : Foris)”, *Studies in generative grammar*, vol. 9. 18
- CLESTHIA. 2018, “Cfpp2000”, URL <https://hdl.handle.net/11403/cfpp2000/v1>, ORTO-LANG (Open Resources and TOols for LANGuage) –www.ortolang.fr. 41
- COVENEY, A. 2011, “L’interrogation directe”, *Travaux de linguistique*, vol. 2, pp. 112–145. URL <https://www.cairn.info/revue-travaux-de-linguistique-2011-2-page-112.htm>. 3, 13, 14, 15, 17, 23, 24, 57, 113
- CRISMA, P. 1992, “On the acquisition of wh-questions in french”, *GenGenP*, pp. 115–122. 32
- DAHL, D. B., D. SCOTT, C. ROOSEN, A. MAGNUSSON et J. SWINTON. 2019, *xtable : Export Tables to LaTeX or HTML*. URL <https://CRAN.R-project.org/package=xtable>, r package version 1.8-4. ix
- DELAVEAU, A., A. DAGNAC et P. CAPPEAU. 2021, “Les phrases interrogatives”, dans *La grande grammaire du français*, édité par A. Abeillé et D. Godard, Actes Sud, pp. –. 3, 12, 13, 14, 24
- DONATI, C. 2006, “On wh-head movement”, *Current Studies in Linguistics Series*, vol. 42, pp. 21–46. URL https://www.researchgate.net/profile/Caterina_Donati/publication/261932598_On_Wh_Head_Movement/links/02e7e535f810234c01000000.pdf. 18
- DOSTIE, G. 2016, “Le corpus de français parlé au québec (cfpq) et la langue des conversations familiales : Exemple de mise à profit des données à partir d’un examen lexico-sémantique de la séquence je sais pas”, *Corpus*, vol. 15. URL <https://journals.openedition.org/corpus/2945>. 24
- DRUMMOND, A. 2016, “Ibexfarm (version 0.3. 9)”. . 87, 101, 115, 132, 155, XCIII, CXVII, CXXXVII
- DÉPREZ, V., K. SYRETT et S. KAWAHARA. 2013, “The interaction of syntax, prosody, and discourse in licensing french wh-in-situ questions”, *Lingua*, vol. 124, doi :<https://doi.org/10.1016/j.lingua.2012.03.002>, pp. 4–19, ISSN 0024-3841. URL <http://www.sciencedirect.com/science/article/pii/S0024384112000617>, sI : Prosody at the Grammatical Crossroads — Japanese and Beyond. 3
- ECKERT, P. 1989, *Jocks and burnouts : Social categories and identity in the high school*, Teachers college press. 27

- ECKERT, P. 2000, *Language variation as social practice : The linguistic construction of identity in Belten High*, Wiley-Blackwell. 27
- ECKERT, P. 2008, “Variation and the indexical field 1”, *Journal of sociolinguistics*, vol. 12, n° 4, pp. 453–476. URL <https://www.semanticscholar.org/paper/Variation-and-the-indexical-field-Eckert/aa41220ba5129ea9e1d5ee412de13a84c3a07cfe?p2df>. 5
- ECKERT, P. 2012, “Three waves of variation study : The emergence of meaning in the study of sociolinguistic variation”, *Annual review of Anthropology*, vol. 41, pp. 87–100. URL <https://pdfs.semanticscholar.org/89d3/de3d9d756fd39cd0b2b1270c309feb4a49a7.pdf>. 26, 28, 30, 129, 217
- ELSIG, M. 2009, *Grammatical variation across space and time : The French interrogative system*, vol. 3, John Benjamins Publishing. URL <https://benjamins.com/catalog/silv.3>. 17
- ENGBAHL, E. 2006, “Information packaging in questions”, *Empirical issues in syntax and semantics*, vol. 6, n° 1, pp. 93–111. URL <http://llf.linguist.jussieu.fr/cssp/eiss6/eiss6.pdf#page=93>. 20
- ENGELMANN, F., S. GRANLUND, J. KOLAK, M. SZREDER, B. AMBRIDGE, J. PINE, A. THEAKSTON et E. LIEVEN. 2019, “How the input shapes the acquisition of verb morphology : Elicited production and computational modelling in two highly inflected languages”, *Cognitive psychology*, vol. 110, pp. 30–69. URL <https://www.sciencedirect.com/science/article/pii/S0010028518301105>. 49
- ESHKOL-TARAVELLA, I., O. BAUDE, D. MAUREL, L. HRIBA, C. DUGUA et I. TELLIER. 2011, “Un grand corpus oral “ disponible ” : le corpus d’Orléans 1 1968-2012”, *Traitement Automatique des Langues*, vol. 53, n° 2, pp. 17–46. URL <https://halshs.archives-ouvertes.fr/halshs-01163053>. 22, 41, 53, 68, 88, 113, 217
- ESTÈVE, Y., T. BAZILLON, J.-Y. ANTOINE, F. BÉCHET et J. FARINAS. 2010, “The EPAC corpus : manual and automatic annotations of conversational speech in French broadcast news”, dans *LREC 2010*, pp. 1686–1689. URL <https://hal.archives-ouvertes.fr/hal-01433895>. 41, 53, 113, 217
- FALK, Y. 2012, “Superiority effects”, Ms. *Hebrew University of Jerusalem*. URL <http://pluto.huji.ac.il/~msyfalk/Superiority.pdf>. 19
- FARMER, K. 2013, ““de quoi tu parles?” : A diachronic study of sociopragmatic interrogative variation in french films”, *University of Pennsylvania Working Papers in Linguistics*, vol. 19, n° 2, p. 8. URL <https://repository.upenn.edu/cgi/viewcontent.cgi?article=1304&context=pwpl>. 25
- FAURE, R. et K. PALASIS. 2020, “Exclusivity! wh-fronting is not optional wh-movement in colloquial french”, *Natural Language & Linguistic Theory*, doi:10.1007/s11049-020-09476-w. URL <https://link.springer.com/article/10.1007/s11049-020-09476-w>. 32
- FOSTER, E. D. et A. DEARDORFF. 2017, “Open science framework (osf)”, *Journal of the Medical Library Association : JMLA*, vol. 105, n° 2, p. 203. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5370619/>. vii, 42

- GABRY, J. 2018, *shinystan : Interactive Visual and Numerical Diagnostics and Posterior Analysis for Bayesian Models*. URL <https://CRAN.R-project.org/package=shinystan>, r package version 2.5.0. ix
- GINZBURG, J. et I. SAG. 2000, *Interrogative investigations*, Stanford : CSLI publications. URL <https://www.bookdepository.com/Interrogative-Investigations-Jonathan-Ginzburg/9781575862781>. 3, 11
- GOLDBERG, A. E. 2006, *Constructions at work : The nature of generalization in language*, Oxford University Press on Demand. URL <https://global.oup.com/academic/product/constructions-at-work-9780199268528>. 222
- GOODMAN, N. D. et D. LASSITER. 2015, “21 probabilistic semantics and pragmatics : Uncertainty in language and thought”, *The handbook of contemporary semantic theory*, pp. 655–686. URL <http://cocolab.stanford.edu/papers/GoodmanLassiter2015-Chapter.pdf>. 5, 27
- GURYEV, A. et F. DELAFONTAINE. 2015, “La variabilité formelle des questions dans les écrits sms”, *Revue Tranel (Travaux neuchâtelois de linguistique)*, vol. 63, pp. 129–152. URL https://doc.rero.ch/record/258199/files/Guryev_Alexander_-_La_variabilite_formelle_des_questions_dans_les_crits_SMS_20160122.pdf. 23
- HAMANN, C. 2000, “The acquisition of constituent questions and the requirements of interpretation”, *The acquisition of syntax : Studies in comparative developmental linguistics*, pp. 170–201. 3, 32
- HAMANN, C. 2006, “Speculations about early syntax : The production of wh-questions by normally developing french children and french children with sli”, *Catalan Journal of Linguistics*, vol. 5, n° 1, pp. 143–189. URL <https://revistes.uab.cat/catJL/article/viewFile/v5-hamann/80>. 30
- HAMLAOUI, F. 2009a, *La focalisation à l'interface de la syntaxe et de la phonologie : le cas du français dans une perspective typologique*, thèse de doctorat, Paris 3. 16, 17, 20, 23, 24, 32, 57, 71, 98, 108, 113
- HAMLAOUI, F. 2009b, “A prosodic study of wh-questions in french natural discourse”, *Proceedings of the LangUE*, pp. 27–38. 3, 21, 22
- HAMLAOUI, F. 2011, “On the role of phonology and discourse in francilian french wh-questions”, *Journal of linguistics*, vol. 47, n° 1, pp. 129–162. URL <https://semanticsarchive.net/Archive/DZjZmI1Y/Hamlaoui2008FrWh.pdf>. 3, 20
- HARRELL JR, F. E., WITH CONTRIBUTIONS FROM CHARLES DUPONT et MANY OTHERS. 2020, *Hmisc : Harrell Miscellaneous*. URL <https://CRAN.R-project.org/package=Hmisc>, r package version 4.4-0. ix
- HARRIS, Z. 1968, “Mathematical structures of language”, *Interscience tracts in pure and applied mathematics*. 42

- HASPELMATH, M. 1999, “Why is grammaticalization irreversible?”, *Linguistics*, vol. 37, n° 6, pp. 1043–1068. URL https://www.researchgate.net/profile/Martin_Haspelmath/publication/40852086_Why_is_grammaticalization_irreversible/links/5405c6f20cf2bba34c1d94f2/Why-is-grammaticalization-irreversible.pdf. 24, 82
- HEIDEN, S., J.-P. MAGUÉ et B. PINCEMIN. 2010, “Txm : Une plateforme logicielle open-source pour la textométrie-conception et développement”, dans *10th International Conference on the Statistical Analysis of Textual Data-JADT 2010*, vol. 2, Edizioni Universitarie di Lettere Economia Diritto, pp. 1021–1032. URL https://halshs.archives-ouvertes.fr/halshs-00549779/file/Heiden_al_jadt2010.pdf. 68
- HEIDER, F. 1946, “Attitudes and cognitive organization”, *The Journal of psychology*, vol. 21, n° 1, pp. 107–112. 30
- HEMFORTH, B., S. COLONNA, C. PETRONE et M. D’IMPERIO. 2013, “Length matters : Informational load in ambiguity resolution”, *Discours*, vol. 12. URL <https://journals.openedition.org/discours/8780>. 22
- HENRY, L., H. WICKHAM et W. CHANG. 2019, *ggstance : Horizontal 'ggplot2' Components*. URL <https://CRAN.R-project.org/package=ggstance>, r package version 0.3.3. ix
- HULK, A. 1996, “The syntax of wh-questions in child french”, *Amsterdam Series in Child Language Development*, pp. 129–172. 3, 30, 32
- IRVINE, J. 2001, “Style’as distinctiveness : The culture and ideologies of linguistic differentiation”, *Penelope Eckert y Jane Rickford (eds.)*, pp. 21–41. 27
- JAKUBOWICZ, C. 2005, “The language faculty :(ab) normal development and interface constraints”, *Comunicação apresentada no GALA*, pp. 8–10. 31
- JAKUBOWICZ, C. 2011, “Measuring derivational complexity : New evidence from typically developing and sli learners of l1 french”, *Lingua*, vol. 121, n° 3, pp. 339–351. URL <https://www.sciencedirect.com/science/article/pii/S0024384110002433>. 4, 31, 32, 43
- JAKUBOWICZ, C. et N. STRIK. 2008, “Scope-marking strategies in the acquisition of long distance wh-questions in french and dutch”, *Language and Speech*, vol. 51, n° 1-2, pp. 101–132. URL <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.861.8880&rep=rep1&type=pdf>. 4, 30, 31
- JUDGE, J., A. CAHILL et J. VAN GENABITH. 2006, “QuestionBank : Creating a corpus of parse-annotated questions”, dans *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Sydney, Australia, pp. 497–504, doi :10.3115/1220175.1220238. URL <https://www.aclweb.org/anthology/P06-1063>. 41
- KAISER, G. et S. QUAGLIA. 2015, “In search of wh in-situ in romance : an investigation in detective stories”, URL http://kops.uni-konstanz.de/bitstream/handle/123456789/34161/Kaiser_0-338855.pdf?sequence=3&isAllowed=y. 17, 108

- KAYNE, R. S. 1981, “Ecp extensions”, *Linguistic inquiry*, vol. 12, n° 1, pp. 93–133. URL <https://www.jstor.org/stable/4178206?seq=1>. 18
- KELLER, R. 1994, *On language change : The invisible hand in language*, Psychology Press. 24, 82
- LABOV, W. 1964, “Stages in the acquisition of standard english”, *Social dialects and language learning*, pp. 77–104. 34
- LABOV, W. 1966, “The social stratification of english in new york city. washington, dc : Center for applied linguistics.. 1972. sociolinguistic patterns”, *Philadelphia : University of Pennsylvania Press.(1972b). Language in the inner city : Studies in the Black English Vernacular. Philadelphia : University of Pennsylvania Press.(1973). The linguistic consequences of being a lame. Language in Society*, vol. 2, pp. 81–1. 26
- LABOV, W. 1972, “Some principles of linguistic methodology”, *Language in society*, pp. 97–120. URL http://danielezrajohnson.com/labov_1972.pdf. 26
- LABOV, W. 2001, “Principles of linguistic change volume 2 : Social factors”, *LANGUAGE IN SOCIETY- OXFORD-*, vol. 29. 26
- LABOV, W. 2012, *Dialect diversity in America : The politics of language change*, University of Virginia Press. 28, 29, 217
- LAHIRE, B. 2019, *Enfances de classe. De l'inégalité parmi les enfants*, Seuil. 35, 210, 222
- LAI-SHEN CHENG, L. et J. ROORYCK. 2000, “Licensing wh-in-situ”, *Syntax*, vol. 3, n° 1, pp. 1–19. 19
- LAMBERT, W. E., R. C. HODGSON, R. C. GARDNER et S. FILLENBAUM. 1960, “Evaluational reactions to spoken languages.”, *The Journal of Abnormal and Social Psychology*, vol. 60, n° 1, p. 44. 28, 42, 87, 128
- LANGACKER, R. W. 1987, *Foundations of cognitive grammar : Theoretical prerequisites*, vol. 1, Stanford university press. 33, 221
- LARRIVÉE, P. 2016, “Les interrogatives in-situ sont-elles pragmatiquement marquées en français vernaculaire? investigation synchronique et historique”, dans *4e colloque Repenser l'histoire du français – 7–8 avril 2016*, Ludwig Maximilians Universität München, pp. –. URL https://www.researchgate.net/publication/303863679_Les_interrogatives_in-situ_sont-elles_pragmatiquement_marquees_en_francais_vernaculaire_Investigation_synchronique_et_historique. 3, 22, 72, 82, 172
- LARRIVÉE, P. 2019, “Historical pragmatics, explicit activation and wh in situ in french”, *Romance Languages and Linguistic Theory 15 : Selected papers from'Going Romance'30, Frankfurt*, vol. 15, p. 113. URL https://www.researchgate.net/profile/Pierre_Larrivee/publication/336369194_Historical_pragmatics_explicit_activation_and_wh_in_situ_in_French/links/5d9e0aed458515df0ae88e21/Historical-pragmatics-explicit-activation-and-wh-in-situ-in-French.pdf. 3, 22, 72, 82, 172
- LEWIS, D. 1969, *Convention. A philosophical study*, Harvard University Press. 5, 27

- LIEGEOIS, L. 2014, *Phonological variables usage in a corpus of parents-child interaction : cognitive devices of learning and impact of language exposure*, Theses, Université Blaise Pascal - Clermont-Ferrand II. URL <https://tel.archives-ouvertes.fr/tel-01274860>. 41, 43
- LIÉGEOIS, L., T. CHANIER et D. CHABANAL. 2014, “Corpus globaux alipe : Interactions parents-enfant annotées pour l’étude de la liaison”, *Nancy, Ortolang*. URL <http://hdl.handle.net/11041/alipe-000853>. 41
- LIÉGEOIS, L., C. ETIENNE, C. PARISSÉ, C. BENZITOUN et C. CHANARD. 2015, “Using the tei as a pivot format for oral and multimodal language corpora”, dans *Text Encoding Initiative Conference and Member’s meeting 2015*, Lyon, France, pp. –. URL <https://halshs.archives-ouvertes.fr/halshs-01345777>. 69
- LLL. 2017, “Eslo”, URL <https://hdl.handle.net/11403/eslo/v1>, ORTOLANG (Open Resources and TOols for LANGuage) –www.ortolang.fr. 22, 41, 53, 68, 88, 113, 217
- LRL. 2017, “Alipe (acquisition de la liaison et interactions parents enfants)”, URL <https://hdl.handle.net/11403/alipe-000853/v1.1>, ORTOLANG (Open Resources and TOols for LANGuage) –www.ortolang.fr. 41
- LÜDECKE, D. 2019, *sjPlot : Data Visualization for Statistics in Social Science*, doi :10.5281/zenodo.1308157. URL <https://CRAN.R-project.org/package=sjPlot>, r package version 2.8.1. ix
- MACWHINNEY, B. 2000, *The CHILDES Project : Tools for analyzing talk. transcription format and programs*, vol. 1, Psychology Press. URL https://www.researchgate.net/publication/243532847_The_CHILDES_project_tools_for_analyzing_talk. 197, 223
- MACWHINNEY, B. 2009, “The childes project”, *Tools for analyzing talk–electronic edition*, vol. 2. URL <https://wizzion.com/childes/CHILDES/manuals/CLAN.pdf>. 32, 43, 197, 223
- MARANDIN, J.-M. 2018, “Les types de phrases en français contemporain”, URL <https://hal.archives-ouvertes.fr/hal-01882110>. 3, 11, 14
- MARTIN, N., J.-P. CHEVROT et S. BARBU. 2010, “Research note stylistic variations in the social network of a 10-year-old child : Pragmatic adjustments or automatic alignment? 1”, *Journal of sociolinguistics*, vol. 14, n° 5, pp. 678–692. URL https://hal.archives-ouvertes.fr/file/index/docid/721685/filename/Martin_Chevrot_Barbu_2010_.pdf. 35
- MATHIEU, E. 2004, “The mapping of form and interpretation : The case of optional wh-movement in french”, *Lingua*, vol. 114, n° 9-10, pp. 1090–1132. URL <http://people.umass.edu/roeper/711-05/Mtheiu%20wh-opt02.pdf>. 19
- MATHIEU, É. 2009, “Les questions en français : Micro et macro-variation”, *Le Français d’Ici : Études Linguistiques et Sociolinguistiques de la Variation*, pp. 37–66. 24
- MILROY, L. 1980, “Language and social networks”, . 27
- MODYCO. 2016, “teicorpo”, URL <https://hdl.handle.net/11403/teicorpo/v1>, ORTOLANG (Open Resources and TOols for LANGuage) –www.ortolang.fr. 69

- MOREY, R. D., R. HOEKSTRA, J. N. ROUDER, M. D. LEE et E.-J. WAGENMAKERS. 2016, “The fallacy of placing confidence in confidence intervals”, *Psychonomic bulletin & review*, vol. 23, n° 1, pp. 103–123. URL <https://link.springer.com/article/10.3758/s13423-015-0947-8>. 48
- MORGENSTERN, A. 2006, *Un Je en construction : genèse de l'auto-désignation chez le jeune enfant*, Editions OPHRYS. URL <https://hal.archives-ouvertes.fr/halshs-00355481/>. 41, 43
- MORGENSTERN, A. et C. PARISSÉ. 2007, “Codage et interprétation du langage spontané d’enfants de 1 à 3 ans”, *Corpus*, vol. 6, pp. 55–78. 4, 32, 41, 43
- MORGENSTERN, A. et M. SEKALI. 2009, “What can child language tell us about prepositions?”, dans *Studies in Language and Cognition*, édité par J. Zlatev, M. J. Falck, C. Lundmark et M. André, Cambridge Scholars Publishing, pp. 261–275. URL <https://halshs.archives-ouvertes.fr/halshs-00376186>. 32
- MÜNSTER, K. et P. KNOEFERLE. 2018, “Extending situated language comprehension (accounts) with speaker and comprehender characteristics : Toward socially situated interpretation”, *Frontiers in psychology*, vol. 8, p. 2267. URL <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.02267/full>. 224
- MYERS, L. L. 2007, *WH-interrogatives in spoken French : A corpus-based analysis of their form and function*, thèse de doctorat, University of Texas. URL <https://repositories.lib.utexas.edu/bitstream/handle/2152/3235/myers119237.pdf?sequence=2&isAllowed=y>. 21, 202
- NARDY, A. 2008, *Acquisition des variables sociolinguistiques entre 2 et 6 ans : facteurs sociologiques et influences des interactions au sein du réseau social*, thèse de doctorat, Université Stendhal - Grenoble III. URL https://tel.archives-ouvertes.fr/file/index/docid/466276/filename/THESE_Nardy_2008.pdf. 35
- NARDY, A., J.-P. CHEVROT et S. BARBU. 2013, “The acquisition of sociolinguistic variation : looking back and thinking ahead”, *Linguistics*, vol. 51, n° 2, pp. 255–284. URL https://hal.archives-ouvertes.fr/file/index/docid/706704/filename/Nardy_Chevrot_Barbu_to_appear_.pdf. 35
- NARDY, A., J.-P. CHEVROT et S. BARBU. 2014, “Sociolinguistic convergence and social interactions within a group of preschoolers : A longitudinal study”, *Language Variation and Change*, vol. 26, n° 3, doi :10.1017/S0954394514000131, p. 273–301. URL <https://hal.archives-ouvertes.fr/hal-01241715/file/Nardy%2C%20Chevrot%20%26%20Barbu%20%282014%29.pdf>. 35
- NEWCOMB, T. M. 1961, “The acquaintance process as a prototype of human interaction.”, –. 30
- OCHS, E. 1992, “14 indexing gender”, *Rethinking context : Language as an interactive phenomenon*, vol. 11, p. 335. URL http://www.sscnet.ucla.edu/anthro/faculty/ochs/articles/Ochs_1993_Indexing_Gender.pdf. 5, 27, 171
- PALASIS, K. 2013, “The case for diglossia : Describing the emergence of two grammars in the early acquisition of metropolitan french”, *Journal of French Language Studies*, vol. 23, doi :10.1017/S0959269512000348. URL <https://doi.org/10.1017/S0959269512000348>. 33

- PALASIS, K., R. FAURE et F. LAVIGNE. 2019, “Explaining variation in wh-position in child french : A statistical analysis of new seminaturalistic data”, *Language Acquisition*, vol. 26, n° 2, pp. 210–234. URL <https://childes.talkbank.org/access/French/0docs/Palasis2019.pdf>. 33
- PODESVA, R. J., J. REYNOLDS, P. CALLIER et J. BAPTISTE. 2015, “Constraints on the social meaning of released/t : A production and perception study of us politicians”, *Language Variation and Change*, vol. 27, n° 1, pp. 59–87. URL <https://doi.org/10.1017/S0954394514000192>. 28, 42, 87, 128
- POHL, J. et G. STRAKA. 1965, “Observations sur les formes d’interrogation dans la langue parlée et dans la langue écrite non littéraire”, dans *Actes du Xe congrès international de linguistique et philologie romanes*, Klincksieck, pp. 501–513. 3, 25
- PRÉVOST, P. 2009, *The acquisition of French : The development of inflectional morphology and syntax in L1 acquisition, bilingualism, and L2 acquisition*, vol. 51, John Benjamins Publishing. URL <https://www.jbe-platform.com/content/books/9789027288851>. 30, 220
- PRÉVOST, P., N. STRIK et L. TULLER. 2014, “Wh-questions in child l2 french : Derivational complexity and its interactions with l1 properties, length of exposure, age of exposure, and the input”, *Second Language Research*, vol. 30, n° 2, pp. 225–250. URL <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.824.262&rep=rep1&type=pdf>. 31
- PRÉVOST, P., L. TULLER, M. A. BARTHEZ, J. MALVY et F. BONNET-BRILHAULT. 2017, “Production and comprehension of french wh-questions by children with autism spectrum disorder : A comparative study with specific language impairment”, *Applied psycholinguistics*, vol. 38, n° 5, pp. 1095–1131. URL <http://laca.humanities.uva.nl/wp/wp-content/uploads/2018/05/Pr%C3%A9vost-et-al.-2017-ApplPsycho.pdf>. 31
- QUILLARD, V. 2001, “La diversité des formes interrogatives : comment l’interpréter?”, *Langage et société*, vol. 1, pp. 57–72. URL <https://www.cairn.info/revue-langage-et-societe-2001-1-page-57.htm>. 3, 25, 57, 71
- REINHARDT, J. 2019, *Regularity and Variation in French Direct Interrogatives : The Morphosyntax and Intonation of Question Forms in Reality TV Shows, Audio Books and Teaching Materials*, thèse de doctorat, Universität Konstanz, Konstanz. URL <http://kops.uni-konstanz.de/handle/123456789/45941>. 32
- RIEGEL, M., J.-C. RELLAT et R. RIOUL. 2014, *Grammaire méthodique du français.*, Presses Universitaires de France. 14, 15, 17, 98, 99, 118, 139, 218
- RIOU, É. et B. HEMFORTH. 2015, “Dislocation clitique de l’objet à gauche en français écrit”, *Discours. Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics*, vol. 16. URL <https://journals.openedition.org/discours/9037>. 90, CXI
- ROBERTS, J. L. 1994, *ACQUISITION OF VARIABLE RULES :(-t, d) DELETION AND*, thèse de doctorat, University of Pennsylvania. URL https://repository.upenn.edu/cgi/viewcontent.cgi?article=1095&context=ircs_reports. 35

- ROSS, J. R. 1967, *CONSTRAINTS ON VARIABLES IN SYNTAX*, thèse de doctorat, MASSACHUSETTS INSTITUTE OF TECHNOLOGY. URL <https://files.eric.ed.gov/fulltext/ED016965.pdf>. 18
- ROWLAND, C. F., J. M. PINE, E. V. LIEVEN et A. L. THEAKSTON. 2003, “Determinants of acquisition order in wh-questions : Re-evaluating the role of caregiver speech”, *Journal of Child Language*, vol. 30, n° 3, pp. 609–635. URL https://pure.mpg.de/rest/items/item_2344479/component/file_2344478/content. 221
- RSTUDIO TEAM. 2015, *RStudio : Integrated Development Environment for R*, RStudio, Inc., Boston, MA. URL <http://www.rstudio.com/>. ix
- RUDIN, C. 1988, “On multiple questions and multiple wh fronting”, *Natural Language & Linguistic Theory*, vol. 6, n° 4, pp. 445–501. URL <http://semantics.uchicago.edu/kennedy/classes/w07/readings/rudin88.pdf>. 19
- SAG, I. A. 2010, “English filler-gap constructions”, *Language*, pp. 486–545. URL <https://www.jstor.org/stable/40961690?seq=1>. 18
- SCHEIDNES, M. et L. TULLER. 2010, “Syntactic movement in the production of frenchwh-questions : The role of computational complexity versus l1 transfer inadultl2 acquisition”, *Movement and clitics : Adult and child grammar*, pp. 185–213. 4, 31
- SCHMID, H. 1994, “Part-of-speech tagging with neural networks”, dans *Proceedings of the 15th conference on Computational linguistics-Volume 1*, Association for Computational Linguistics, pp. 172–176. URL <https://arxiv.org/pdf/cmp-lg/9410018.pdf>. 69
- SCHMID, H. 1995, “Improvements in part of speech tagging with an application to german. in proceedings of the 7th conference of the european chapter of the acl (eacl-95)”, dans *Workshop SIGDAT, Dublin, Ireland*, pp. –. URL <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.52.2255&rep=rep1&type=pdf>. 69
- SEDDAH, D. et M. CANDITO. 2016, “Hard Time Parsing Questions : Building a QuestionBank for French”, dans *Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, Proceedings of the 10th edition of the Language Resources and Evaluation Conference (LREC 2016), Portorož, Slovenia, pp. –. URL <https://hal.archives-ouvertes.fr/hal-01457184>. 41
- SOARES, C. 2003, “The c-domain and the acquisition of european portuguese : The case of wh-questions”, *Probus*, vol. 15, n° 1, pp. 147–176. URL <https://www.degruyter.com/view/journals/prbs/15/1/article-p147.xml>. 30
- SOARES, C. 2004, “Computational complexity and the acquisition of the cp field in european portuguese”, dans *Proceedings of Console*, vol. 12, pp. 125–140. URL https://www.researchgate.net/profile/Carla_Soares-Jesel/publication/265308338_Computational_complexity_and_the_acquisition_of_the_CP_field_in_European_Portuguese/links/568cf4b908ae153299b880a3.pdf. 30

- SOARES, C. 2006, *La syntaxe de la périphérie gauche en portugais européen et son acquisition*, thèse de doctorat, Paris 8. URL https://www.researchgate.net/profile/Carla_Soares-Jesel/publication/284726044_La_Syntaxe_de_la_Peripherie_Gauche_en_Portugais_Europeen_et_son_Acquisition/links/56581fc008ae1ef9297ca47c.pdf. 30
- SÖLL, L. 1983, “L’interrogation directe dans un corpus de langage enfantin”, *Études de grammaire française descriptive*, pp. 45–54. 3, 15, 24
- SORENSEN, T., S. HOHENSTEIN et S. VASISHTH. 2016, “Bayesian linear mixed models using stan : A tutorial for psychologists, linguists, and cognitive scientists”, *The Quantitative Methods for Psychology*, vol. 12, n° 3, doi :10.20982/tqmp.12.3.p175, pp. 175–200. URL <http://www.tqmp.org/RegularArticles/vol12-3/p175/p175.pdf>. 47
- STOCKWELL, R. P. 1977, *Foundations of syntactic theory*, Prentice Hall. 18
- STRIK, N. 2007, “L’acquisition des phrases interrogatives chez les enfants francophones”, *Psychologie française*, vol. 52, n° 1, pp. 27–39. URL <https://www.sciencedirect.com/science/article/abs/pii/S0033298406000707>. 30
- STRIK, N. 2012, “Wh-questions in child bilingual acquisition of french : Derivational complexity and cross-linguistic influence”, *Canadian Journal of Linguistics/Revue canadienne de linguistique*, vol. 57, n° 1, pp. 133–151. URL https://www.researchgate.net/publication/317136069_Wh-questions_in_child_bilingual_acquisition_of_French_Derivational_complexity_and_cross-linguistic_influence. 31
- TERRY, R. M. 1970, *Contemporary French interrogative structures*, vol. 5, Editions Cosmos. 3, 15, 24, 57
- THIBERGE, G. 2017, *Acquisition en français L1 des syntagmes interrogatifs prépositionnels*, mémoire de master, Université Paris Diderot. URL http://www.llf.cnrs.fr/sites/llf.cnrs.fr/files/biblio//1706_gt_M2.pdf. 4, 32, 220
- THIBERGE, G. 2018, “Position du syntagme wh- en français : réelle optionnalité ou biais sociolinguistique?”, *ELIS - Échanges Linguistiques en Sorbonne, Université Paris Sorbonne*, vol. 5, p. 64–91. URL <https://halshs.archives-ouvertes.fr/halshs-01802590>. 4, 88, 90, 92, 97, 128, 130, 222
- THIBERGE, G., F. BADIN et L. LIÉGEOIS. en révision, “French partial interrogatives : a microdiachronic study of variation and new perspectives in a refined pragmatics framework”, *TBA*. 68, 71, 89, 223
- TOMASELLO, M. 2000, “Do young children have adult syntactic competence?”, *Cognition*, vol. 74, n° 3, pp. 209–253. URL <https://www.sciencedirect.com/science/article/abs/pii/S0010027799000694>. 33
- TOMASELLO, M. 2009a, *Constructing a language*, Harvard university press. URL <https://www.hup.harvard.edu/catalog.php?isbn=9780674017641>. 33
- TOMASELLO, M. 2009b, “The usage-based theory of language acquisition”, dans *The Cambridge handbook of child language*, Cambridge Univ. Press, pp. 69–87. 33, 221

- TOMASELLO, M. 2018, “How children come to understand false beliefs : A shared intentionality account”, *Proceedings of the National Academy of Sciences*, vol. 115, n° 34, pp. 8491–8498. URL <https://www.pnas.org/content/pnas/115/34/8491.full.pdf>. 34
- VAN BERKUM, J. J., D. VAN DEN BRINK, C. M. TESINK, M. KOS et P. HAGOORT. 2008, “The neural integration of speaker and message”, *Journal of cognitive neuroscience*, vol. 20, n° 4, pp. 580–591. URL https://www.researchgate.net/publication/5794892_The_Neural_Integration_of_Speaker_and_Message. 224
- VASISHTH, S., Z. CHEN, Q. LI et G. GUO. 2013, “Processing chinese relative clauses : Evidence for the subject-relative advantage”, *PloS one*, vol. 8, n° 10, p. e77 006. URL <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0077006>. 49
- VENEZIANO, E. 2009, “Language and internal states : A long developmental history at different levels of functioning.”, *Rivista di Psicolinguistica applicata /Journal of Applied Psycholinguistic*, vol. IX, n° 3, pp. pp 9–27. URL <https://halshs.archives-ouvertes.fr/halshs-00413314>. 34
- VENEZIANO, E. 2015, “Compétences pragmatiques, théorie de l’esprit et non linéarité discursive chez l’enfant.”, *Travaux linguistiques du CerLiCO*, vol. 28, n° 1, pp. 181–194. URL <https://halshs.archives-ouvertes.fr/halshs-01273744>. 34
- WAGNER, S. E. 2012, “Age grading in sociolinguistic theory”, *Language and Linguistics Compass*, vol. 6, n° 6, pp. 371–382. URL <https://scinapse.io/papers/1664970667>. 98
- WEI, T. et V. SIMKO. 2017, *R package "corrplot" : Visualization of a Correlation Matrix*. URL <https://github.com/taiyun/corrplot>, (Version 0.84). ix
- WICKHAM, H. 2007, “Reshaping data with the reshape package”, *Journal of Statistical Software*, vol. 21, n° 12, pp. 1–20. URL <http://www.jstatsoft.org/v21/i12/>. ix
- WICKHAM, H. 2011, “The split-apply-combine strategy for data analysis”, *Journal of Statistical Software*, vol. 40, n° 1, pp. 1–29. URL <http://www.jstatsoft.org/v40/i01/>. ix
- WICKHAM, H. 2016, *ggplot2 : Elegant Graphics for Data Analysis*, Springer-Verlag New York, ISBN 978-3-319-24277-4. URL <https://ggplot2.tidyverse.org>. ix
- WICKHAM, H. 2018, *scales : Scale Functions for Visualization*. URL <https://CRAN.R-project.org/package=scales>, r package version 1.0.0. ix
- WICKHAM, H., M. AVERICK, J. BRYAN, W. CHANG, L. D. MCGOWAN, R. FRANÇOIS, G. GROLEMUND, A. HAYES, L. HENRY, J. HESTER, M. KUHN, T. L. PEDERSEN, E. MILLER, S. M. BACHE, K. MÜLLER, J. OOMS, D. ROBINSON, D. P. SEIDEL, V. SPINU, K. TAKAHASHI, D. VAUGHAN, C. WILKE, K. WOO et H. YUTANI. 2019, “Welcome to the tidyverse”, *Journal of Open Source Software*, vol. 4, n° 43, doi :10.21105/joss.01686, p. 1686. URL <https://joss.theoj.org/papers/10.21105/joss.01686.pdf>. ix
- WICKHAM, H., R. FRANÇOIS, L. HENRY et K. MÜLLER. 2020, *dplyr : A Grammar of Data Manipulation*. URL <https://CRAN.R-project.org/package=dplyr>, r package version 0.8.5. ix

- WICKHAM, H., J. HESTER et R. FRANCOIS. 2018, *readr : Read Rectangular Text Data*. URL <https://CRAN.R-project.org/package=readr>, r package version 1.3.1. ix
- WILKE, C. O. 2020, *ggridges : Ridgeline Plots in 'ggplot2'*. URL <https://CRAN.R-project.org/package=ggridges>, r package version 0.5.2. ix
- WIMMER, H. et J. PERNER. 1983, “Beliefs about beliefs : Representation and constraining function of wrong beliefs in young children’s understanding of deception”, *Cognition*, vol. 13, n° 1, pp. 103–128. URL <https://www.sciencedirect.com/science/article/abs/pii/0010027783900045>. 33
- WINCKEL, E. et A. ABEILLÉ. accepté, “French subject island constraint, a discourse-based hpsg account”, *Linguisticæ Investigationes*. 19
- ZLATKA, G. et J.-P. DESCLÉS. 1991, “Test et acceptabilité”, *Histoire Épistémologie Langage*, tome 13, fascicule 2, 1991. *Théories et données.*, doi :10.3406/hel.1991.2330, ISSN 0750-8069. URL https://www.persee.fr/doc/hel_0750-8069_1991_num_13_2_2330. 42
- ZUCKERMAN, S. et A. HULK. 2001, “Acquiring optionality in french wh-questions : An experimental study”, *Revue québécoise de linguistique*, vol. 30, n° 2, pp. 71–97. URL <https://www.erudit.org/en/journals/rql/1900-v1-n1-rql412/000520ar.pdf>. 4, 30