

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LES PARTICULARITÉS DU FRANÇAIS SOURD : LE CAS DES VERBES PRONOMINAUX

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE
DE LA MAITRISE EN LINGUISTIQUE

PAR

LAURIE THIVIERGE CÔTÉ

JANVIER 2024

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.04-2020). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Un grand merci du fond du cœur à ma directrice, Reine Pinsonneault. Je n'aurais jamais pu mener ce projet à terme sans ta patience et ta flexibilité. Merci pour nos discussions enrichissantes et pour ton approche pédagogique; merci de m'avoir fait sentir écoutée. Je te suis énormément reconnaissante d'avoir cru en moi jusqu'au bout.

Merci à Anne-Marie Parisot, qui m'a fait découvrir la richesse du français sourd et qui a accepté de me donner accès aux ressources du groupe de recherche sur la LSQ et le bilinguisme sourd. L'idée derrière ce mémoire est née dans ton cours d'analyse de corpus; merci de m'avoir aidée à en construire la fondation et de m'avoir guidée dans le terrain inconnu de la recherche sur la LSQ.

Merci à mes collègues du groupe LSQ (tant les membres que j'ai côtoyé·e·s dans les bureaux du groupe que les chargé·e·s de cours qui m'ont acceptée dans leur classe) de m'avoir chaleureusement accueillie et de m'avoir permis d'approfondir mes connaissances sur la LSQ. Un merci tout particulier à Amélie Voghel et Daz Saunders, qui ont pris le temps de répondre à mes listes de questions décousues.

Merci à Jill Vandermeerschen, à Denis Larocque et à Denis Foucambert de m'avoir accordé des heures de leur temps afin de répondre à mes questions de débutante en statistiques; votre aide m'a été précieuse. Merci aussi à Philippe Ducharme, qui m'a aidée à apprendre comment rédiger efficacement; ton support a été un pilier pendant mes moments d'incertitude.

Merci à la gang du CAF du CVM. C'est là que mon amour de la grammaire a pris naissance; il a grandi au fil de conversations avec Catherine Brunet, Jocelyne Cyr et Laurence Olivier (et avec les *bums*, évidemment!). Merci à Charles Guilbert d'avoir vu le potentiel en moi.

Merci à mes ami·e·s, qui ont ensoleillé mes journées, qui m'ont supportée tout au long de mon (long) parcours et à qui j'ai demandé des jugements de grammaticalité à toute heure du jour. Merci à ma très chère Chloé Larouche; aux fantastiques Laurence Gagnon, Mylène Vézina-Boucher, Ariane Sarrazin et Isabelle Boyer; et à mon premier allié, Olivier Brosseau-Côté. Merci à mes collègues de monitorat et aux membres de l'asso qui ont participé aux séances de *coworking* et aux

retraites de rédaction. Merci à mes *homies*, Sarah-Florence Benjamin, Aglaé Duhamel-Dubé et Estelle Bourbeau, de partager leur quotidien avec moi, de me faire rire très fort et de constituer un solide réseau d'entraide. Merci à Laurent Lasnier et à Fauve Paris-Vallières, qui m'endurent depuis plus de deux décennies, sans se tanner. Merci à Patricio Calcagno, à Alexandra Kack et à Xavier Bouffard-Chevalier pour leur amour et leur support infailible; je pouvais toujours compter sur vous pour une période de rédaction efficace. Merci à Mariette Cliche-Galarza pour les meilleurs *one on one*. Merci à la belle gang Horap, qui m'a permis de découvrir un nouveau monde – surtout à Jessée Riach-Tremblay, à Vincent Alarie et à Tasneem Hoque, pour leurs encouragements constants.

Merci à François Côté, à Alex Thivierge Côté, à Honorine Youmbissi, à Makéda Ékoué et à Rokiya Ékoué pour leur amour inconditionnel; j'ai de la chance de vous avoir.

Merci, bien sûr, à mon plus grand coéquipier, Alexandre Vallée. Merci énormément d'avoir été là pour moi quand j'ai priorisé mon mémoire au détriment du reste. Merci pour ton soutien inébranlable, pour tous les efforts que tu as faits pour moi, pour ta patience et pour ton amour.

Cette recherche a été financée par le Fonds québécois de recherche sur la société et la culture.

DÉDICACE

*À ma maman,
qui n'aura jamais su que j'ai hérité de son amour pour la grammaire.*



TABLE DES MATIÈRES

Remerciements	ii
Dédicace	iv
Liste des figures	viii
Liste des tableaux	x
Liste des abréviations	xii
Résumé	xvi
CHAPITRE 1 : INTRODUCTION	1
CHAPITRE 2 : PROBLÉMATIQUE	4
2.1 Définition de la référence	4
2.2 Structures du marquage de la référence en français et en LSQ	6
2.2.1 Grammaire du français : les pronoms personnels	6
2.2.2 Grammaire de la LSQ : l'association et l'activation de locus	8
2.3 Le pronom SE	15
2.4 État de la question	16
2.5 Objectif général de recherche	21
CHAPITRE 3 : CADRE THÉORIQUE	22
3.1 L'écriture sourde dans une approche bilingue et biculturelle	22
3.1.1 Transferts linguistiques dans l'acquisition d'une langue seconde	23
3.1.2 Le français sourd – théorie	24
3.1.3 Le français sourd – description	26
3.1.4 La distance entre le français et la LSQ dans la référence syntaxique	29
3.1.5 L'emploi des pronoms personnels en français sourd	32
3.1.6 Les formes de la LSQ correspondant aux éléments réalisés par SE en français	34
3.2 Contextes morphosyntaxiques et sémantiques	35
3.2.1 Verbes pronominaux : définition et classements	36
3.2.1.1 Verbes essentiellement pronominaux	37
3.2.1.2 Verbes pronominaux autonomes	38
3.2.1.3 Verbes pronominaux réfléchis et réciproques	38

3.2.1.4	Verbes pronominaux passifs	39
3.2.1.5	Verbes pronominaux neutres.....	41
3.2.2	Temps de conjugaison.....	42
3.2.3	Classement aspectuel.....	43
3.2.4	Classement sémantique	44
3.2.5	Structure transitive	48
3.2.6	Forme du sujet.....	49
3.2.7	Traits sémantiques du sujet	49
3.3	Questions spécifiques de recherche.....	51
CHAPITRE 4 : MÉTHODE.....		53
4.1	Présentation du corpus Groupe LSQ 98 et de l'ensemble de données.....	54
4.1.1	Participant.e.s.....	54
4.1.2	Procédures de création du corpus.....	57
4.1.3	Données du corpus	57
4.2	Identification des occurrences.....	62
4.2.1	Catégories d'occurrences	62
4.2.2	Motivation des critères d'identification des occurrences.....	63
4.2.3	Protocole d'analyse grammaticale	64
4.2.4	Protocole d'identification des occurrences	65
4.2.5	Protocole de jugement de la norme	68
4.3	Présentation de la grille d'analyse du corpus.....	70
4.3.1	Variables d'identification.....	70
4.3.2	Variables d'observation de l'emploi du pronom SE.....	71
4.3.3	Variables contextuelles	73
4.4	Analyses statistiques	75
4.4.1	Test du chi-carré.....	75
4.4.2	Analyse factorielle.....	76
4.4.3	Analyse des correspondances multiples.....	77
4.4.4	Analyse factorielle multiple	79
4.4.5	Description des modalités	81
CHAPITRE 5 : RÉSULTATS.....		84
5.1	Effet des variables d'identification	84

5.2	Analyse générale (ACM)	89
5.3	Analyse orientée sur les variables d'observation (AFM).....	93
5.4	Analyse fine des variables d'observation (<i>catdes</i>).....	97
5.4.1	Résultats <i>catdes</i> de [PresSe].....	98
5.4.2	Résultats <i>catdes</i> de [RspNorme]	100
5.4.3	Résultats <i>catdes</i> de [CatOcc].....	102
5.4.4	Résultats <i>catdes</i> de [NormePres].....	105
5.4.5	Interaction de [NormePres] avec les résultats	107
CHAPITRE 6 : DISCUSSION.....		111
6.1	Représentativité des analyses	111
6.2	Tendances générales.....	113
6.3	Effet des variables contextuelles	114
6.3.1	Effet du type de verbe pronominal	114
6.3.2	Effet du temps de conjugaison	115
6.3.3	Effet de la classe aspectuelle	116
6.3.4	Effet du classement lexico-sémantique	117
6.3.5	Effet de la structure transitive	118
6.3.6	Effet de la forme du sujet	121
6.3.7	Effet des traits sémantiques du sujet	122
6.4	Observations liées à la question de recherche 1 (respect de la norme)	124
6.4.1	Description des occurrences respectant la norme d'emploi des verbes pronominaux.....	125
6.4.2	Description des occurrences présentant un écart à la norme d'emploi des verbes pronominaux.....	128
6.4.3	Description des occurrences d'ajout non normé du SE.....	131
6.5	Observations liées à la question de recherche 2 (présence du SE).....	134
6.6	Hypothèses diverses	136
CHAPITRE 7 : CONCLUSION		138
RÉFÉRENCES.....		142

LISTE DES FIGURES

Figure 2.1 – Signe POINTÉ dirigé vers un locus	9
Figure 2.2 – Exemples de classificateurs de forme	10
Figure 2.3 – Exemples de classificateurs sémantiques	11
Figure 2.4 – Structure de narration.....	11
Figure 2.5 – Structure de représentation corporelle	11
Figure 2.6 – Exemples de flexion verbale en LSQ	13
Figure 2.7 – Signe APPARTENIR.....	14
Figure 2.8 – Signe MOI-MÊME	14
Figure 3.1 – Les classes sémantico-syntaxiques, les sous-classes et les sous-types syntaxiques de la classe générique X	47
Figure 4.1 – Exemple de <i>scree plot</i>	77
Figure 4.2 – Exemple de <i>biplot</i> d’ACM.....	78
Figure 4.3 – Exemple de carré des liaisons.....	79
Figure 4.4 – Exemple de contribution d’un groupe de variable.....	80
Figure 4.5 – Exemple de graphique des axes partiels	81
Figure 5.1 – <i>Scree plot</i> de l’ACM.....	89
Figure 5.2 – Carré des liaisons de l’ACM.....	89
Figure 5.3 – Graphique principal de l’ACM.....	90
Figure 5.4 – Qualité de représentation des variables de l’ACM	91
Figure 5.5 – Regroupement des individus de l’ACM	92
Figure 5.6 – Qualité de représentation des individus de l’ACM.....	92
Figure 5.7 – <i>Scree plot</i> de l’AFM	93
Figure 5.8 – Carré des liaisons de l’AFM	93
Figure 5.9 – Graphique principal de l’AFM	94

Figure 5.10 – Qualité de représentation des individus de l’AFM	95
Figure 5.11 – Regroupement des individus de l’AFM.....	95
Figure 5.12 – Axes partiels de l’AFM.....	96
Figure 6.1 – Graphique des axes partiels de l’AFM (5 dimensions).....	135

LISTE DES TABLEAUX

Tableau 2.1 – Les pronoms personnels de première et deuxième personne	7
Tableau 2.2 – Les pronoms personnels de troisième personne	7
Tableau 2.3 – Comparaison des erreurs de verbes pronominaux (FS, FL1 et FL2)	19
Tableau 3.1 – Les classes génériques de LVF	46
Tableau 3.2 – Les classes sémantico-syntaxiques de LVF	47
Tableau 4.1 – Description des participant·e·s	55
Tableau 4.2 – Nombre de mots par participant·e et par texte	58
Tableau 4.3 – Nombre de verbes par participant·e et par texte.....	59
Tableau 4.4 – Nombre d’occurrences par participant·e et par texte	60
Tableau 4.5 – Catégories d’occurrences	62
Tableau 4.6 – Combinaisons possibles des variables d’observation.....	71
Tableau 4.7 – Combinaisons représentées par la variable [RspNorme]	72
Tableau 4.8 – Combinaisons représentées par la variable [CatOcc].....	72
Tableau 4.9 – Exemple de <i>catdes</i> d’une modalité.....	82
Tableau 5.1 – Chi-carré de [Texte] + [PresSe].....	85
Tableau 5.2 – Chi-carré de [Texte] + [RspNorme]	85
Tableau 5.3 – Chi-carré de [Particip] + [RspNorme].....	85
Tableau 5.4 – Chi-carré de [Particip] + [PresSe]	86
Tableau 5.5 – Chi-carré de [Particip] + [NormePres]	86
Tableau 5.6 – Tableau croisé de [Particip] + [PresSe].....	87
Tableau 5.7 – Tableau croisé de [Particip] + [NormePres].....	88
Tableau 5.8 – Contribution (%) des groupes de variables aux dimensions de l’AFM	96
Tableau 5.9 – Qualité de représentation des modalités des variables d’observation (AFM).....	97

Tableau 5.10 – Résultat <i>catdes</i> pour [PresSe] (variables).....	98
Tableau 5.11 – Résultat <i>catdes</i> pour [PresSe] (modalités).....	99
Tableau 5.12 – Résultat <i>catdes</i> pour [RspNorme] (variables)	101
Tableau 5.13 – Résultat <i>catdes</i> pour [RspNorme] (modalités)	101
Tableau 5.14 – Résultat <i>catdes</i> pour [CatOcc] (variables).....	102
Tableau 5.15 – Résultat <i>catdes</i> pour [CatOcc] (modalités)	103
Tableau 5.16 – Résultat <i>catdes</i> pour [NormePres] (variables).....	105
Tableau 5.17 – Résultat <i>catdes</i> pour [NormePres] (modalités)	106
Tableau 5.18 – Résultat <i>catdes</i> pour [PresSe] (variables, considérant [NormePres]).....	108
Tableau 5.19 – Résultat <i>catdes</i> pour [PresSe] (modalités, considérant [NormePres])	108
Tableau 5.20 – Résultat <i>catdes</i> pour [RspNorme] (modalités, considérant [NormePres])	109
Tableau 5.21– Résultat <i>catdes</i> pour [CatOcc] (variables, considérant [NormePres])	109
Tableau 5.22 – Résultat <i>catdes</i> pour [CatOcc] (modalités, considérant [NormePres])	110

LISTE DES ABRÉVIATIONS

TERMES GÉNÉRAUX

CD	Complément direct du verbe
CI	Complément indirect du verbe
FTS	Forme(s) typique(s) sourde(s)
FL1	Français langue maternelle
FL2	Français langue seconde
FS	Français sourd
L1	Langue maternelle
L2	Langue seconde
LSQ	Langue des signes québécoise
LVF	Les verbes français de Dubois et Dubois-Charlier (1997)

ANALYSES STATISTIQUES

ACM	Analyse des correspondances multiples
AFM	Analyse factorielle multiple
<i>catdes</i>	Test de description des modalités

VARIABLES

[]	Variable
< >	Modalité

[Anime] Trait d'animéité du sujet

< Ani > Animé

< Inani > Inanimé

[Aspect] Classe aspectuelle de l'occurrence

< aAchv > Achèvement

< aAcmp > Accomplissement

< aAct > Activité

< aEtat > État

< AspNA > Non applicable (semi-auxiliaire ou forme typique sourde ne pouvant être analysée)

- [CatOcc] Catégorie d'occurrence
- < Ajout > Ajout du SE (écart à la norme)
 - < AN > Absence normée du SE
 - < PNF > Présence normée facultative du SE
 - < PN > Présence normée du SE
 - < Omis > Omission du SE (écart à la norme)
- [LVF] Classement selon Les verbes du français (Dubois et Dubois-Charlier, 1997)
- < LVF_A > La lettre A représente la catégorie sémantico-syntaxique (voir la liste des catégories au tableau 3.1, section 3.2.4)
- [Nb] Trait de nombre du référent du sujet
- < sg > Singulier
 - < pl > Pluriel
- [NormePres] Norme quant à la présence du SE : selon la norme, le SE est-il facultatif, obligatoire ou interdit dans cette occurrence?
- < PN_O > Présence selon la norme oui : selon la norme, la présence du SE est obligatoire
 - < PN_N > Présence selon la norme non : selon la norme, la présence du SE est interdite
 - < PN_Fac > Présence selon la norme facultative : selon la norme, la présence du SE est facultative
- [Particip] Participant·e
- < P00 > Code alphanumérique du ou de la participant·e (de P01 à P15)
- [Pers] Traits de personne du référent du sujet
- < 0 > Le numéro associé au trait de personne (de 1 à 3)
- [PresSe] Présence du SE : est-ce que le SE est présent dans l'occurrence?
- < PresO > Présence oui : l'occurrence contient un SE
 - < PresN > Présence non : l'occurrence ne contient pas un SE
- [RoleT] Rôle thématique du sujet
- < agent > Agent
 - < experi > Expérimenteur
 - < instru > Instrument

- < pati > Patient
- < sieg > Siège
- < them > Thème
- < RolND > Rôle thématique non disponible (sujet impersonnel)

[RspNorme] Respect norme : l'occurrence respecte-t-elle la norme quant à la présence du SE?

- < RN_o > Respect norme oui : respect de la norme (présence ou absence normée du SE)
- < RN_n > Respect norme non : écart à la norme (ajout ou omission du SE)

[StEmp] Structure transitive employée dans l'occurrence

- < eAtt > Emploi attributif
- < eInt > Emploi intransitif
- < eTD > Emploi transitif direct
- < eTI > Emploi transitif indirect
- < eBiTr > Emploi bitransitif

[StNorme] Norme de structure transitive : quelle serait la structure transitive normée?

- < nAtt > Selon la norme, l'occurrence devrait avoir une structure attributive
- < nInt > Selon la norme, l'occurrence devrait avoir une structure intransitive
- < nTD > Selon la norme, l'occurrence devrait avoir une structure transitive directe
- < nTI > Selon la norme, l'occurrence devrait avoir une structure transitive indirecte
- < nBiTr > Selon la norme, l'occurrence devrait avoir une structure à deux compléments
- < nND > Il n'y a pas de structure transitive normée pour cette occurrence

[StRsp] Respect structure transitive : la structure transitive employée est-elle normée?

- < RSTo > Respect structure transitive oui (la structure employée respecte la norme)
- < RSTn > Respect structure transitive non (la structure employée présente un écart à la norme)
- < RSTchx > Mauvais choix de la préposition (la structure employée présente un écart à la norme dans le choix de la préposition qui introduit un complément)

[Sujet] Forme du sujet de l'occurrence

- < pronom > Le sujet a la forme d'un pronom personnel
- < ProRel > Le sujet a la forme d'un pronom relatif
- < SN > Le sujet a la forme d'un syntagme nominal dont le noyau n'est pas un pronom

<AbsCoord> Absence coordination : absence du sujet à cause d'une ellipse dans une coordination

< SsEnt > Sous-entendu (l'occurrence est à l'infinitif, à l'impératif ou un participe présent)

< AbsNn > Absence non normée du syntagme sujet

[Temps] Temps de conjugaison du verbe

< TC > Temps composé

< TS > Temps simple

[Texte] Numéro du texte, correspondant aux 6 amorces données aux participant·e·s

< T0 > Numéro du texte, correspondant aux 6 amorces données aux participant·e·s

[vPro] Verbe pronominal : à quel type de verbe pronominal correspond l'occurrence?

< AutoEP > Verbe pronominal autonome ou essentiellement pronominal

< Neutre > Verbe pronominal neutre

< Passif > Verbe pronominal passif

< PasPro > Verbe jamais pronominal ou ne pouvant pas l'être dans le contexte de l'occurrence

< RefCD > Verbe pronominal ou réfléchi dont le SE a la fonction de complément direct

<RefCI > Verbe pronominal ou réfléchi dont le SE a la fonction de complément indirect

[VvInf] Présence d'une construction *verbe + verbe infinitif*

< VinfO > Présence d'un semi-auxiliaire ou d'une phrase subordonnée infinitive

< VinfN > Absence de semi-auxiliaire et de phrase subordonnée infinitive

RÉSUMÉ

Ce mémoire a pour but la description de l'emploi des verbes pronominaux dans un corpus de textes préalablement identifié comme étant porteur des caractéristiques du français sourd. De nature exploratoire, cette recherche vise à déterminer si l'emploi du pronom préverbal « se » est régi par une norme sourde et, si c'est le cas, quels sont les facteurs qui influencent sa production.

Les occurrences de verbes retenues pour l'étude ont été analysées en fonction de variables morphologiques, syntaxiques et sémantiques (type de verbe pronominal, temps de conjugaison, classement aspectuel, classification sémantico-syntaxique, structure transitive, forme du sujet et traits sémantiques du sujet). Quatre types d'occurrences ont été retenus : la présence normée du « se », l'ajout non normé du « se », l'omission du « se », ainsi que l'absence normée du « se » dans les phrases où sa présence aurait été possible. La description des données est réalisée à l'aide d'une analyse des correspondances multiples et d'une analyse factorielle multiple précisée par des descriptions de modalités.

De manière générale, les résultats sont compatibles avec la littérature sur le sujet : lorsqu'il est obligatoire, le « se » est produit 63% du temps, mais il l'est seulement dans 10% des occurrences où sa présence est facultative. Lorsqu'il est produit, sa présence est normée dans 71% des cas, et les écarts à la norme comprennent davantage d'omissions que d'ajouts.

Les résultats démontrent que les variables observées indiquent si une présence ou une absence du « se » respectera la norme, mais pas si un verbe sera pronominal ou non dans ce genre de contexte. En d'autres mots, il semblerait que les facteurs sur lesquels les scripteurs et scriptrices sourd·e·s se basent pour décider si un verbe est pronominal ou non ne proviennent pas du contexte de la phrase. Cela soulève l'hypothèse que cette décision est le produit d'un transfert linguistique dans l'utilisation de la langue des signes québécoise comme langue de référence, comme le propose la littérature sur le français sourd.

Mots clés : Verbe pronominal, français écrit, français sourd.

CHAPITRE 1 : INTRODUCTION

Les langues des signes n'ont pas de système d'écriture, contrairement à certaines langues vocales (c'est-à-dire les langues transmises par la voix, par exemple le français). Ainsi, afin de produire un texte, une personne ayant une langue des signes en tant que langue maternelle ou en tant que langue de référence¹ devra employer le code écrit d'une langue seconde. Des études portant sur des textes de Sourd·e·s² rédigés en diverses langues indiquent que le profil de cette population ressemble à celui de personnes entendant·e·s écrivant dans leur langue seconde. En effet, les textes écrits en langue seconde (L2) présentent des écarts à la norme qui ne se retrouvent pas dans les écrits en langue maternelle (L1). Bref, les textes en L2 se distinguent de ceux en L1 par certaines formes typiques des scripteurs et scriptrices L2, tant entendant·e·s que Sourd·e·s. Toutefois, au sein même des textes L2, la population sourde se distingue. En effet, les textes de Sourd·e·s comportent des formes caractéristiques, que l'on ne retrouve pas dans les textes en L2 d'apprenant·e·s entendant·e·s.

Les particularités des textes de Sourd·e·s proviennent du fait qu'ils sont produits selon un système linguistique propre à cette population, donnant lieu à l'écriture sourde. Les règles qui composent ce système mènent à la production de formes typiques sourdes (dorénavant FTS). Ces formes typiques sourdes, présentées et définies à la section 3.1, regroupent toutes les formes produites de manière régulière dans les écrits sourds, peu importe si elles sont uniques à l'écriture sourde ou si elles sont aussi produites par d'autres populations bilingues. Ce mémoire cherche à décrire l'une de ces formes typiques sourdes dans le but d'approfondir notre compréhension du système derrière l'écriture sourde.

¹ Les notions de *langue de référence* et de *marquage de la référence* sont récurrentes dans ce mémoire. Toutefois, il s'agit de deux notions distinctes et indépendantes : la *langue de référence* désigne une langue déjà acquise sur laquelle on peut se baser dans une situation de bilinguisme (voir 3.1.1), tandis que le *marquage de la référence* désigne le recours à la référence syntaxique (anaphorique ou déictique), tel que l'emploi d'un pronom référant à un antécédent dans le but d'assurer la cohésion du discours (voir 2.1). Dans ce mémoire, le premier sens de *référence* se distingue du second par sa présence réservée exclusivement à l'expression *langue de référence*.

² Mon emploi du terme *Sourd·e* réfère à une identité culturelle, sans égard à la condition physiologique de l'audition, ce qui implique l'emploi de la majuscule pour le nom et de la minuscule pour l'adjectif : *un·e Sourd·e*, *une personne sourde*.

L'hypothèse du transfert linguistique permet de comprendre et d'expliquer ces FTS. Selon cette hypothèse, l'écriture sourde constitue en un emploi d'une langue seconde qui s'appuie sur une langue des signes en tant que langue de référence, pouvant générer de ce fait des transferts linguistiques (voir 3.1.1). Ces transferts peuvent être positifs quand les structures entre deux langues sont similaires ou négatifs quand les structures sont différentes.

Dans le cadre de ce mémoire, je m'intéresse à l'emploi du français écrit influencé par les structures de la langue des signes québécoise (LSQ), que je désigne en tant que *français sourd*³. Plus précisément, ce mémoire porte sur l'utilisation du pronom personnel préverbal SE⁴ dans des textes en français sourd.

En français, les pronoms personnels ont pour fonction de marquer la référence syntaxique (définie à la section 2.1). En effet, les pronoms ne peuvent être interprétés sans contexte; leur interprétation dépend d'une association avec un référent. Ce référent peut être soit un élément déictique provenant de la situation d'énonciation (par exemple, ce qui permet habituellement l'interprétation du pronom *je*), soit un antécédent, c'est-à-dire un élément linguistique produit ailleurs dans l'énoncé⁵. L'anaphore, définie comme un élément ayant besoin d'un antécédent dans la phrase pour recevoir son interprétation, permet la reprise sémantique et le remplacement syntaxique d'un élément et, ce faisant, elle participe à la cohésion du discours (Riegel *et al.*, 2009). Cependant, en LSQ, les structures employées pour effectuer une reprise sémantique et pour assurer la cohésion du discours reposent sur des mécanismes différents. Comme mentionné précédemment, une différence de structure entre la langue de référence et la langue cible⁶ peut entraîner des transferts négatifs (voir 3.1.1).

³ Tout au long du texte, j'emploierai le terme *français sourd* pour désigner le français écrit influencé par la LSQ. Toutefois, il est important de noter qu'il s'agit plutôt du *français sourd au Québec*, puisque le recours à différentes langues des signes comme langue de référence entraîne différentes formes à l'écrit; il n'existe pas de *français sourd* unique et universel.

⁴ Les lexèmes du français sont écrits en petites majuscules soulignées (LEXÈMES) et les signes de la LSQ, en petites majuscules (SIGNE). Ainsi, SE réfère au pronom « se » et ses flexions (*me, te, nous, vous*).

⁵ On retrouve habituellement le référent avant l'anaphore, mais il peut être après, dans le cas de cataphores.

⁶ L'expression *langue cible* réfère à la langue seconde dans laquelle l'énoncé est produit.

En ce qui concerne le pronom SE, l'identification des fonctions qu'il remplit peut présenter des difficultés particulières. En effet, le pronom SE a différents emplois qui se regroupent en trois catégories. La première regroupe les SE réfléchis et réciproques, qui remplissent la fonction de complément du verbe (1). La deuxième catégorie regroupe les SE qui ont une fonction passive ou anticausative (2). La dernière catégorie regroupe les SE lexicaux, qui n'ont pas de fonction et qui font partie de l'entrée lexicale du verbe (3).

- (1) Les deux frères se sont parlé toute la soirée.
- (2) La bouteille s'est cassée sous l'impact.
- (3) Le public se souvient du film de ce réalisateur.

Les différentes fonctions (discursives, syntaxiques, sémantiques et lexicales) remplies par le pronom SE ne sont pas toutes actualisées en LSQ et, lorsqu'elles le sont, les structures de la LSQ diffèrent de celles du français. Encore une fois, cette différence entre les structures de la langue de référence et celles de la langue cible (c'est-à-dire la langue seconde employée, étant ici le français) peuvent entraîner des transferts négatifs.

Les études portant sur les particularités du français sourd ont établi que l'emploi des pronoms personnels et, plus particulièrement, l'emploi du SE génèrent des formes typiques sourdes présentant un écart à la norme. Toutefois, peu d'intérêt a été porté sur la compréhension du système d'emploi du SE en français sourd. Dans le but d'apporter des pistes de solution à ce problème, ce mémoire a comme objectif de décrire l'utilisation du SE dans des textes de Sourd·e·s. Cette description se fera selon deux axes. Dans un premier temps, j'identifierai les caractéristiques communes des contextes syntaxiques des énoncés dans lesquels le SE est employé, ce qui permettra une meilleure compréhension des structures utilisées dans ces textes. Afin de préciser la nature de la différence entre l'emploi du SE dans ces textes et son emploi normé, j'observerai dans un deuxième temps les caractéristiques communes des contextes syntaxiques accompagnant les emplois du SE selon leur respect ou non de la norme du français écrit⁷.

⁷ Dans le cadre de ce mémoire, j'adopte la définition suivante de *norme* : « [...] tout ce qui est d'usage commun et courant dans une communauté linguistique ; la norme correspond alors à l'institution sociale que constitue la langue standard » (Dubois *et al.*, 2002, p.330). La norme n'est donc pas établie selon des critères prescriptivistes subjectifs, mais plutôt selon ce qui est jugé comme grammatical par les locuteurs et locutrices, sans restriction de registre.

CHAPITRE 2 : PROBLÉMATIQUE

Le marquage de la référence⁸ repose sur des systèmes intrinsèquement différents en français et en LSQ. Ainsi, le recours à la LSQ comme langue de référence (c'est-à-dire le fait de se baser sur ses connaissances de la LSQ) pour produire une anaphore en français peut générer des transferts négatifs. Étant donné la diversité des rôles associés au pronom SE (fonction grammaticale dans le groupe verbal, passivation, etc.), on peut s'attendre à ce que l'acquisition des verbes pronominaux soit ardue en français langue seconde, notamment pour les Sourd·e·s.

Ce chapitre commence par une définition de la référence (2.1) ainsi qu'une description de son marquage (2.2) par les pronoms personnels en français (2.2.1) et par l'activation de locus en LSQ (2.2.2). Les sections suivantes abordent les fonctions du pronom SE (2.3), puis l'état des connaissances actuelles sur l'emploi des verbes pronominaux dans les textes de Sourd·e·s (2.4), ce qui mène à la formulation de l'objectif général du présent mémoire (2.5), soit la description de l'emploi des verbes pronominaux dans un corpus porteur de caractéristiques du français sourd.

2.1 Définition de la référence

En linguistique, la référence désigne une relation entre une unité linguistique et l'objet ou concept qu'elle dénote (Crystal, 2008). Dans le cadre de ce mémoire, ce terme sera utilisé dans un sens restreint, soit celui de la référence syntaxique – l'établissement d'une relation d'identité entre une unité linguistique et un référent disponible dans le contexte syntaxique ou pragmatique.

Un discours – qu'il soit émis par la voix, signé ou écrit – se doit d'être cohérent. À cette fin, les énoncés qui le composent doivent respecter deux règles : la règle de progression (apporter de l'information nouvelle) et la règle de répétition (avoir des éléments récurrents d'une phrase à l'autre pour créer un fil conducteur) (Riegel *et al.*, 2009). La reprise de l'information dans le texte,

⁸ Le *marquage de la référence* désigne les méthodes employées dans une langue pour établir une relation entre une unité linguistique et le concept qu'elle dénote; cette notion est élaborée à la section 2.1.

nécessaire pour respecter la deuxième règle de cohérence du discours, se fait grâce aux anaphores, c'est-à-dire une expression linguistique qui réactive un référent déjà introduit dans le discours.

Certaines expressions linguistiques (telles que la plupart des pronoms) doivent être associées à un référent dans le discours pour recevoir leur interprétation, et sont donc toujours anaphoriques. Par exemple, dans la phrase « *Carole souffre d'allergies, elle éternue sans cesse* », le pronom *elle* renvoie à son antécédent *Carole* (le mot), et active la même représentation mentale du référent *Carole* (le concept). Dans le cas des langues des signes, des points dans l'espace sont assignés à des référents pour créer des locus référentiels (Dubuisson *et al.*, 1999); leurs réactivations dans des énoncés subséquents génèrent des références syntaxiques (voir 2.2.2).

En français, les anaphores peuvent prendre la forme de pronoms personnels; le lien avec l'antécédent prendra alors la forme d'un partage de traits grammaticaux. Prenons la phrase suivante :

(4) Jérôme et Sandrine vont à l'hôpital, car **elle** est malade.

Dans cette phrase, le seul candidat potentiel pour remplir la fonction d'antécédent du pronom « elle » est le nom « Sandrine », étant donné qu'il est le seul correspondant aux traits réalisés par le pronom : 3^{ème} personne, féminin, singulier. Ainsi, le nom « Jérôme » n'est pas un candidat adéquat, puisqu'il demande un pronom masculin.

En LSQ, le marquage de la référence se fait à l'aide d'associations spatiales. Reprenons la phrase (4), cette fois-ci traduite en LSQ⁹ :

(5) JÉRÔME_(x) SANDRINE_(y) HÔPITAL PTÉ_(z) 6_(x,y)-ALLER-3_(z) POURQUOI MALADE_(y) PTÉ3_(y)

Dans cette phrase, le nom JÉRÔME est signé à un endroit précis de l'espace de signation, désigné par l'indice (x). Un locus référentiel est alors établi : ce point spatial est désormais associé au référent « Jérôme ». Il en est de même pour SANDRINE au locus (y) et l'HÔPITAL au locus (z). Les trois locus sont réactivés dans la flexion verbale, puis un seul est pointé pour spécifier que Sandrine

⁹ L'indice entre parenthèses à la suite d'un signe indique le locus sur lequel le signe est produit.

est malade¹⁰. Le signe POINTÉ, comparativement au pronom *elle*, désigne plus directement son antécédent, puisqu'il désigne l'entité elle-même plutôt que de passer par les traits 3^{ème} personne, féminin et singulier (Liddell, 2000).

La différence entre la modalité vocale du français et la modalité gestuelle de la LSQ a un impact sur la forme que prennent les anaphores (Parisot, 2003), ce qui cause des transferts négatifs dans l'utilisation des pronoms personnels en français sourd (voir 3.1.5 pour une description).

2.2 Structures du marquage de la référence en français et en LSQ

Contrairement au français, la LSQ ne dispose pas d'un paradigme de formes pronominales organisées selon des traits (Liddell, 1990; Parisot, 2003). Tandis que les langues de modalité auditivo-verbale sont limitées à l'articulation séquentielle (et donc à une juxtaposition temporelle des expressions linguistiques), la modalité visuo-gestuelle des langues des signes permet la superposition spatiale d'éléments et, par extension, une certaine permanence de la trace spatiale associant un locus à un référent – c'est-à-dire que cette trace peut être réutilisée plus loin dans le discours, sans que l'on ait à en expliciter le référent de nouveau (Parisot, 2003).

2.2.1 Grammaire du français : les pronoms personnels

Les pronoms personnels du français – nommés ainsi par association au trait grammatical de personne (Riegel *et al.*, 2009) – s'organisent dans un paradigme de formes selon les traits grammaticaux, sémantiques et syntaxiques qu'ils réalisent. Leur rôle est de faire la reprise sémantique d'un référent déjà introduit dans le discours (ou alors présent dans la situation d'énonciation) et de faire le remplacement syntaxique d'un syntagme, le plus souvent nominal (ou prépositionnel, dans les cas des pronoms compléments indirects du verbe). Ils permettent aussi d'alléger le discours comparativement aux autres stratégies de reprise de l'information, telle que la paraphrase. Étant donné que ce sont des mots grammaticaux, leur nombre est limité.

¹⁰ La forme 6(x,y)-ALLER-3(z) est la flexion du verbe ALLER d'une part avec le sujet, SANDRINE(x) et JÉRÔME(y), et d'autre part avec l'objet, l'HÔPITAL(z). Le signe ALLER est produit avec un mouvement allant de (x) et (y) à (z).

Tableau 2.1 – Les pronoms personnels de première et deuxième personne

Traits grammaticaux		Fonction syntaxique				
nb ↓	pers. ↓	sujet de la phrase	complément du verbe			
			complément disjoint	complément en position préverbale		
				c. direct	c. indirect	c. réfléchi
SG	1	je	moi	me		
	2	tu	toi	te		
PL	1	nous				
	2	vous				

Tableau 2.2 – Les pronoms personnels de troisième personne

Traits grammaticaux			Fonction syntaxique				
pers. ↓	nb ↓	genre ↓	sujet de la phrase	complément du verbe			
				complément disjoint	complément en position préverbale		
					c. direct	c. indirect	c. réfléchi
3	SG	M	il	lui	le	lui (animé) y, en (inanimé)	se
		F	elle		la		
	PL	M	ils	eux	les	leur (animé) y, en (inanimé)	
		F	elles				
	indéterminé	on ¹¹	soi	nous			

¹¹ Pour des raisons d'espace, l'utilisation du pronom *on* pour désigner un référent de première personne au pluriel ne sera pas traitée.

Le pronom génère des classes de potentiels référents selon les traits qu'il réalise, et c'est le contexte qui permet de le lier au bon référent. Sa présence est obligatoire : la phrase « **vais à la plage* » est agrammaticale, même si le pronom « *je* » est le seul sujet possible pour cette flexion du verbe *aller*.

2.2.2 Grammaire de la LSQ : l'association et l'activation de locus

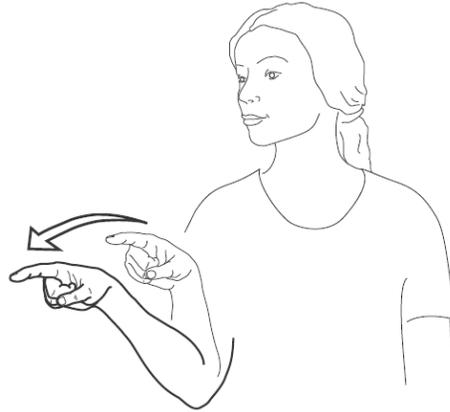
Le marquage de la référence en LSQ se fait par la localisation, c'est-à-dire une marque morphologique variable dont la valeur est déterminée par des contenus discursifs (Cormier *et al.*, 2013). Un locus référentiel est créé lorsqu'un point dans l'espace discursif est associé à un référent, acquérant ainsi une fonction articulatoire (Lillo-Martin et Klima, 1990). Cette trace spatiale est globale et exclusive; il ne pourrait pas avoir d'ambiguïté référentielle avec un autre référent ayant des traits communs (par exemple, deux antécédents 3^{ème} personne singulier masculin). En fait, Sandler et Lillo-Martin (2006) proposent même que l'activation et la réactivation des locus seraient la réalisation morphologique des indices formels de référence (exemples en ASL tirés de Sandler et Lillo-Martin (2006, p. 378))¹² :

- (6) SÉNATEUR_(x:z) TOUS_(x:z) SENTIR PTÉ_(x) PTÉ_(y) PTÉ_(z) INTELLIGENT
 'Chaque sénateur_i pense qu'il_i est intelligent.'
- (7) SÉNATEUR_(x:z) TOUS_(x:z) SENTIR PTÉ_(w) INTELLIGENT
 'Chaque sénateur_i pense qu'il_j est intelligent.'
- (8) SÉNATEUR_(x:z) TOUS_(x:z) SENTIR PTÉ_(x:z) INTELLIGENT
 'Chaque sénateur_i pense qu'ils_i sont intelligents / qu'il_i est intelligent.'

L'association d'un référent à un locus peut se faire de manière manuelle (POINTÉ ou localisation directe) ou non manuelle (inclinaison du tronc ou regard vers le locus) (Parisot, 2003). Par exemple, le signe POINTÉ (figure 2.1) sert à assigner une entité à un locus ou à réutiliser la trace de ce locus pour établir des liens référentiels dans la situation de communication.

¹² La notation (x:z) désigne les locus x, y et z, réactivés d'un seul mouvement en arc.

Figure 2.1 – Signe POINTÉ dirigé vers un locus¹³



PTÉ-3

L'assignation spatiale par le POINTÉ se réalise lorsque le signeur, la signeuse, produit le signe correspondant à un référent donné avant ou en même temps qu'un POINTÉ en direction d'un locus, dorénavant associé à ce référent (9). Les autres stratégies sont la localisation directe – soit la production d'un signe non ancré¹⁴ directement sur un locus – (10), l'inclinaison latérale du tronc vers un locus pendant la production d'un nom (11)¹⁵ et la direction du regard vers un locus pendant qu'un nom est signé (12).

(9) ALEXANDRE PTÉ3_(x) 1-TÉLÉPHONER-3_(x)
 'Je téléphone à Alexandre'

(10) ALEXANDRE_(x) 1-TÉLÉPHONER-3_(x)
 'Je téléphone à Alexandre'

(11) ALEXANDRE PTÉ3_(x) $\frac{\text{T : (y)}}{\text{PROFESSEUR}}$ $\frac{\text{T : (y), r : (x)}}{3_{(y)}\text{-TÉLÉPHONER-3}_{(x)}}$
 'Le professeur téléphone à Alexandre'

¹³ Sauf indication contraire, les figures proviennent de la *Banque visuelle de signes dessinés de la LSQ* (Groupe de recherche sur la LSQ et le bilinguisme sourd, 2013).

¹⁴ Les signes non ancrés sont ceux produits dans l'espace neutre dans leur forme de citation.

¹⁵ Les comportements non manuels sont inscrits au-dessus des signes, et leur portée est indiquée par la ligne. T indique l'inclinaison du tronc; t, la position de la tête et r, la direction du regard.

- (12) R : (x)
ALEXANDRE 3_(x)-MARCHER
'Alexandre marche'

Tandis que l'association spatiale implique de nouvelles informations, la réactivation de trace permet la référence à des éléments déjà introduits dans le discours (Parisot et Rinfret, 2008), assurant ainsi la cohésion du discours. Ainsi, une fois qu'un locus référentiel est établi, il est possible de générer une anaphore en réactivant leur trace spatiale. Les autres stratégies principales d'activation et de réactivation de locus sont l'emploi de classificateurs, le recours à une structure de représentation corporelle, la flexion verbale, les signes de possession et les signes PRO-MÊME.

Les classificateurs, morphèmes incorporés au verbe (ou au nom), prennent la forme d'une configuration manuelle renvoyant à un référent nominal grâce à une relation d'iconicité basée sur des caractéristiques physiques saillantes (par exemple la forme ou la taille, cf. figure 2.2)¹⁶ ou une appartenance à une classe sémantique (figure 2.3) (Voghel, 2016).

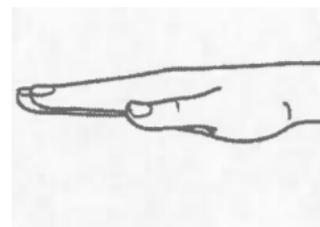
Figure 2.2 – Exemples de classificateurs de forme



Objet de petite taille au contour rond



Objet de forme ronde



Étendue d'une surface plane

¹⁶ Les figures 2.2 et 2.3 sont tirées de Voghel (2016, pp. 93-94).

Figure 2.3 – Exemples de classificateurs sémantiques



Le classificateur effectue une projection du référent dans l'espace, ce qui permet le maintien de la référence d'un nom préalablement introduit dans le discours, en plus d'exprimer des informations sur son orientation, sur la localisation et la direction de son mouvement, ainsi que sur la manière dont ce mouvement se réalise (Voghel, 2016).

Les structures de représentation corporelle correspondent à un changement de perspective effectué quand le signeur, la signeuse, délaisse la structure narrative (figure 2.4)¹⁷ – dans laquelle les locus ont une localisation prototypique dans l'espace neutre – pour adopter le point de vue d'un actant du discours en prenant sa position dans l'espace discursif (figure 2.5) (Saunders, 2016).

Figure 2.4 – Structure de narration

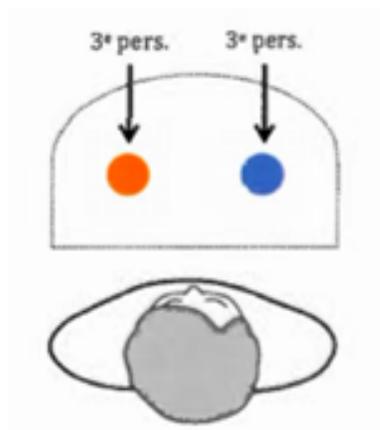
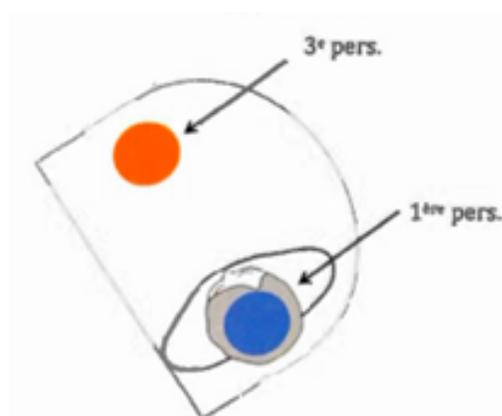


Figure 2.5 – Structure de représentation corporelle



¹⁷ Les figures 2.4 et 2.5 proviennent de Saunders (2016, p. 5).

Une même situation peut être relatée dans une structure de narration (13), ou alors dans une structure de représentation corporelle (14) (exemples tirés de Saunders (2016, p. 6)).

(13) PIERRE_(z) 3_(z)-RÉPONDRE-1 HIER PTÉ3_(z) 3-TRAVAILLER_(z) FORT
 ‘Pierre m’a répondu qu’il avait travaillé fort hier.’

(14) PIERRE_(z) 3_(z)-RÉPONDRE-1 T : droite, r : gauche
 HIER 1-TRAVAILLER FORT
 ‘Pierre m’a répondu « j’ai travaillé fort hier ».’

Bien que les structures de représentation corporelles soient similaires au discours direct (qui implique aussi une modification des référents des pronoms *je* et *tu*), elles ne sont pas limitées à la reprise des paroles de l’actant du discours ¹⁸ (exemple tiré de Saunders (2016, p. 42)) :

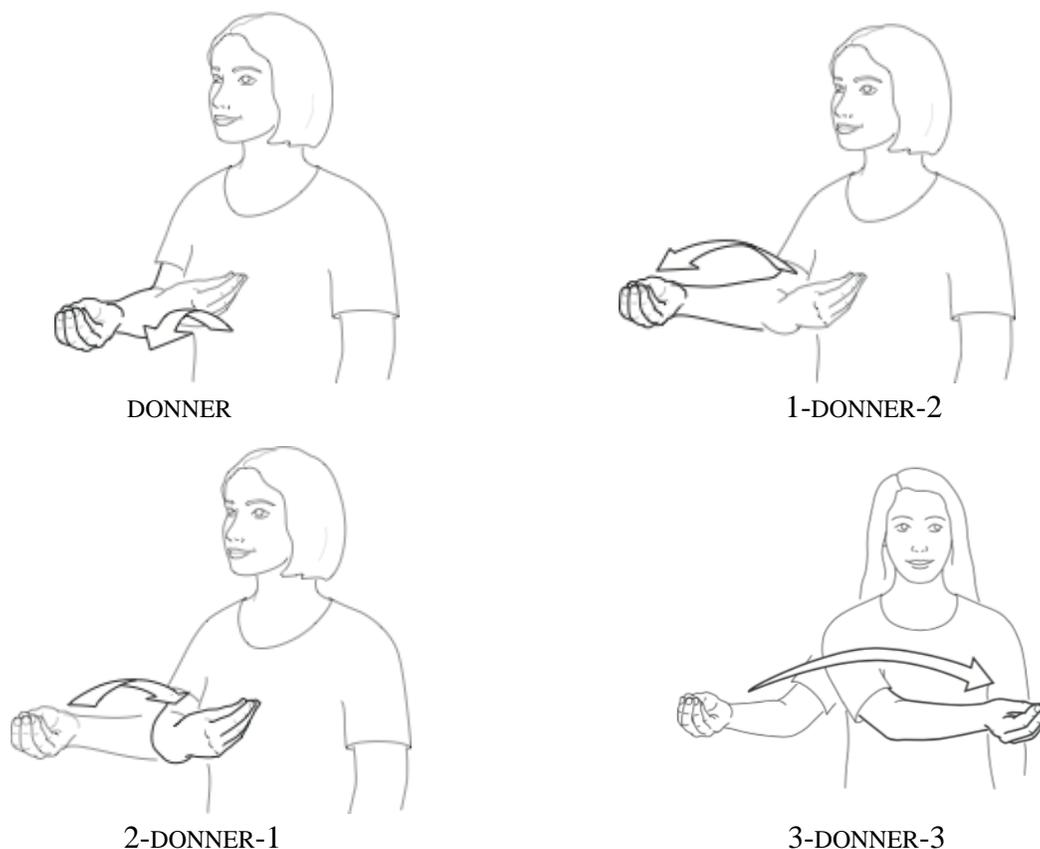
(15) t : avancée, r : vers les mains, exp.f. : intéressé
 HOMME PTÉ-3 LIVRE_(z) LIRE++++_(z)
 ‘Il y a un homme qui lit un livre avec intérêt’

La flexion verbale désigne un changement de directionalité de certaines classes de verbes afin d’associer les rôles thématiques (agent, patient, thème, locatifs) à des locus (Parisot, 2003). La figure 2.6 illustre le signe DONNER dans sa forme de citation ¹⁹ ainsi que différentes flexions possibles.

¹⁸ Voir Saunders (2016) pour une présentation des différentes fonctions discursives et pragmatiques des structures de représentation corporelle.

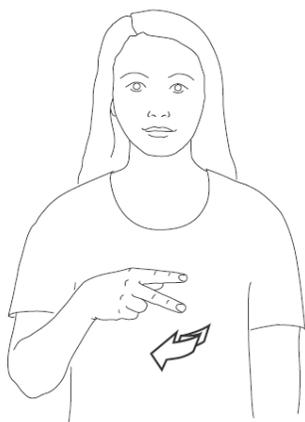
¹⁹ La forme de citation correspond à la forme la plus simple d’un signe, produite en isolation, qui serait celle listée dans un dictionnaire (Voghel, 2016).

Figure 2.6 – Exemples de flexion verbale en LSQ



La réactivation d'un locus peut aussi être effectuée par l'insertion du marqueur de possession APP ou les signes de type PRO-MÊME. Généralement, le signe APP est inséré à la suite d'un nom désignant le possédé et est produit sur le locus du possesseur. Il peut être réalisé avec la configuration manuelle /P/ (figure 2.7) ou avec la configuration manuelle /B/. Les signes de type PRO-MÊME ('moi-même', 'toi-même', 'lui-même', etc.) sont plutôt une marque d'insistance (figure 2.8); ils sont décrits plus en détail dans le cadre théorique (voir 3.1.6).

Figure 2.7 – Signe APPARTENIR



APPARTENIR

Figure 2.8 – Signe MOI-MÊME²⁰



MOI-MÊME

Bref, le système de référence des langues des signes repose sur l'association de référents à des traces spatiales, puis à la réactivation de celles-ci grâce à diverses stratégies. Parmi ces stratégies, on en retrouve certaines qui relèvent du comportement non manuel, et d'autres qui renvoient au référent par un procédé morphologique ou syntaxique. En d'autres mots, l'activation et la réactivation d'un locus référentiel ne sont pas dépendantes d'une insertion d'éléments lexicaux anaphoriques. Cette option demeure disponible; d'ailleurs, les signes POINTÉ, APP et PRO-MÊME sont comparables à l'emploi des pronoms du français. Les signes POINTÉ sont fréquemment désignés dans la littérature comme des « pronoms personnels » ou des « signes pronominaux », bien que cette classification théorique développée pour les langues vocales s'applique moins bien aux langues des signes. En effet, ces signes ne proviennent pas d'un paradigme limité de formes réalisant un ensemble de traits grammaticaux, comme c'est le cas en français. De plus, les POINTÉ n'ont pas de lieu d'articulation fixe, et ont en théorie un nombre infini de points dans l'espace dans lesquels ils peuvent se situer (Liddell, 2000)²¹.

²⁰ La figure 2.8 est tirée de Sandler et Lillo-Martin (2006, p. 374).

²¹ Cela génère le problème de listabilité des morphèmes d'accord (Liddell, 2000), qui alimente le débat quant à la nature du POINTÉ, mais cela dépasse la portée du présent mémoire.

2.3 Le pronom SE

Les verbes pronominaux sont les verbes précédés du pronom SE, qui reçoit les traits grammaticaux du sujet du verbe. Le pronom préverbal SE réalise plusieurs rôles qui peuvent être regroupées en trois grandes catégories (Usito, 2022) : le SE réfléchi, le SE passif et le SE lexical.

La première catégorie, le **SE réfléchi**, désigne les marqueurs de réflexion (16) et de réciprocité (17), qui remplissent nécessairement la fonction grammaticale de complément direct (16a, 17a) ou de complément indirect (16b, 17b) du verbe. Ce pronom est anaphorique et correspond à une pronominalisation du complément verbal lorsque celui-ci est coréférentiel au sujet.

- (16) a. Nathalie se voit dans le miroir. b. Les hôtes se servent une assiette.
(17) a. Ces enfants s'insultent sans cesse. b. Mes parents se parleront demain.

La catégorie des SE réfléchis inclut aussi le pronom indiquant une montée du possesseur dans la possession inaliénable lorsque le possesseur et le sujet de la phrase sont coréférentiels (18). Cette construction n'est pas possible lorsque la possession n'est pas aliénable (19)²².

- (18) Elle lave ses cheveux. → Elle se lave les cheveux.
(19) Elle range son appartement. → *Elle se range l'appartement.

La deuxième catégorie, le **SE passif**, regroupe le SE marqueur du passif d'habitude ou de potentiel (dit SE moyen, 20) et le SE marqueur anticausatif (21) (Antidote, 2022b).

- (20) Ce livre se lit rapidement.
(21) Le verre s'est brisé.

La troisième catégorie, le **SE lexical**, désigne les SE sans fonction syntaxique incorporés dans l'entrée lexicale des « verbes pronominaux purs » (Le Petit Robert, 2022) et de verbes occasionnellement pronominaux dont le SE n'a ni fonction grammaticale, ni rôle de passivation.

²² Par contre, la montée du possesseur ne se limite pas qu'aux cas où le possesseur et le sujet de la phrase sont coréférentiels : *Elle lave mes cheveux.* → *Elle me lave les cheveux.*

Ces pronoms peuvent apporter une information sémantique de processus interne (cf. 22) ou être sémantiquement vides (cf. 23) (Antidote, 2022b).

- (22) Il s'accroupit à côté de l'arbre.
- (23) Ma sœur se souvient de cet endroit.

Ainsi, SE peut réaliser différentes fonctions, et même n'en avoir aucune. Les trois catégories identifiées créant un ensemble disparate, il est difficile d'identifier un élément unifiant tous ces SE, que ce soit une fonction commune ou un sens commun. À cause de cela, les verbes pronominaux sont difficiles à acquérir en français langue seconde, comme nous le verrons dans la section suivante.

2.4 État de la question

De nombreuses études ont été réalisées sur des corpus de textes en français écrits par des Sourd·e·s afin d'en identifier les caractéristiques²³ (voir section 3.1.3), mais aucune ne traite du pronom SE en particulier. On retrouve néanmoins quelques descriptions relativement courtes de l'emploi des verbes pronominaux dans des recherches à portée plus large, notamment Nadeau *et al.* (1991), Dubuisson et Nadeau (1993), Nadeau (1993), Nadeau et Machabée (1998) et Perini (2013).

À ma connaissance, la première mention de l'emploi des verbes pronominaux dans des textes de Sourd·e·s est dans Nadeau *et al.* (1991). Cet article a pour but de proposer une analyse qualitative du français écrit d'étudiant·e·s sourd·e·s. Au sujet des verbes pronominaux, les autrices indiquent que les pronoms réfléchis sont parfois manquants (ex. : *en souvenant*) ou en trop (ex. : *une cloche se sonnait*).

Ensuite, Dubuisson et Nadeau (1993) ont rédigé un rapport pour le Ministère de l'Emploi et de la Solidarité sociale (MESS) sur un projet intitulé « Le français sourd et la LSQ ». Ce projet, qui s'est déroulé de janvier 1991 à mai 1993, avait un objectif triple : a) vérifier l'hypothèse que les difficultés marquées en français écrit chez les apprenant·e·s sourd·e·s étaient sensiblement les

²³ D'ailleurs, analyser les rédactions en français de personnes sourdes était un sujet assez populaire dans les années 90, mais ce l'est moins depuis : l'intérêt se situe plutôt sur l'analyse de la LSQ et de la LSF, ainsi que sur le développement de stratégies d'enseignement du français adaptées à cette population.

mêmes chez tou·te·s les apprenant·e·s, peu importe leur(s) mode(s) de communication et leur niveau scolaire; b) faire une analyse détaillée de la situation; c) proposer des solutions pédagogiques pour favoriser l'apprentissage du français chez les Sourd·e·s.

Afin de donner un exemple d'analyse détaillée d'erreur typique des textes écrits par des Sourd·e·s, les autrices ont dédié un chapitre de ce rapport à la description de l'emploi des verbes pronominaux dans leur corpus de textes. Il s'agit de l'analyse la plus poussée trouvée dans la littérature sur le sujet. Entre autres, les autrices identifient l'inefficacité des classements de verbes pronominaux employés dans les grammaires traditionnelles, tant pour la catégorisation de leurs emplois en français sourd que pour l'enseignement de leur emploi normé à la population sourde, étant donné que ces classements se fient à des connaissances intuitives des locuteur·trice·s du français. Les autrices divisent plutôt les verbes pronominaux dans les 6 catégories suivantes : les verbes essentiellement pronominaux, les verbes transitifs directs, les verbes transitifs indirects, les verbes intransitifs, les verbes au sens passif et les verbes « parties de corps » (qui correspondent aux réfléchis avec trait de possession dans ce mémoire, voir 2.3 et 3.2.1).

Les résultats de Dubuisson et Nadeau (1993) indiquent que la catégorie la plus fréquemment employée est celle des verbes pronominaux transitifs directs (86-69%), suivie des verbes essentiellement pronominaux (13-6%). Le plus haut taux d'erreur est dans l'emploi des verbes pronominaux transitifs directs, qui, selon leur analyse, sont utilisés comme des verbes intransitifs, sans pronoms (ex. : *elle mette à l'abri*²⁴). Elles identifient aussi comme erreurs récurrentes l'emploi du pronom réfléchi avec un verbe ayant un autre complément direct (*le héros doit s'affronter la peur*), ainsi que l'emploi du pronom réfléchi avec un verbe intransitif (*Fido crie « Woof » pour qu'on sait quoi s'arrive*). Le pourcentage d'erreurs dans l'emploi du verbe pronominal diminue quelque peu au fil du temps, mais reste relativement stable : 31% pour les 9-11 ans, 37% pour les 12-14 ans et 26% pour les 15-17 ans. D'ailleurs, la fréquence d'emploi des verbes pronominaux n'augmente pas avec l'âge des scripteurs et scriptrices.

²⁴ Les exemples de verbes employés comme verbes intransitifs de Dubuisson et Nadeau (1993) ne comportent pas de verbes qui, selon mon analyse, sont intransitifs; dans cet exemple, il s'agit plutôt d'un emploi transitif indirect. Il est toutefois possible que leur catégorisation des emplois ne considère que les compléments pronominalisés.

La même année, Nadeau (1993) publie un article cherchant à savoir s'il existe un « français sourd ». Cet article présente des données reprises dans Nadeau et Machabée (1998). Dans les deux publications, les erreurs que l'on retrouve dans des textes de personnes sourdes (dont certain·e·s ne connaissent que le français) sont comparées aux profils d'erreurs de deux groupes, soit des entendant·e·s pour qui le français est la langue maternelle (FL1) ainsi que des entendant·e·s pour qui le français est la langue seconde (FL2). Les résultats indiquent que, de manière générale, le groupe FL1 se distingue des deux autres – c'est-à-dire que les Sourd·e·s font des erreurs similaires à celles du groupe FL2 –, mais certains éléments sont propres au français sourd et ne se retrouvent pas dans les textes des entendant·e·s L1 et L2. Tant Nadeau (1993) que Nadeau et Machabée (1998) indiquent que, dans les textes de Sourd·e·s, le pronom SE est parfois omis, parfois en trop.

Nadeau (1993) précise que l'omission et l'ajout du pronom réfléchi est aussi attesté dans la littérature sur le profil d'erreurs du groupe FL1, mais que ce n'est jamais décrit comme étant une catégorie particulière, tandis que c'est une difficulté récurrente en français sourd, qui persiste jusqu'à l'âge adulte. De plus, pour plusieurs catégories d'erreur en FL1, les occurrences sont circonscrites à certains contextes (par exemple, l'omission de mots dans la deuxième partie d'une coordination) ou alors elles proviennent d'une influence de l'oral, alors qu'au contraire, en français sourd, ces catégories d'erreurs sont généralisées. Cette observation s'applique vraisemblablement à l'emploi des verbes pronominaux.

Toutefois, lorsque le groupe de FS est comparé à celui de FL2, Nadeau (1993) ainsi que Nadeau et Machabée (1998) proposent que certaines catégories d'erreurs, dont l'emploi des verbes pronominaux, sont réellement comparables. Les exemples donnés sont présentés au tableau 2.3 :

Tableau 2.3 – Comparaison des erreurs de verbes pronominaux (FS, FL1 et FL2)

	FS	FL1	FL2
omission	<p>Les garçons souvent bat avec les filles.²⁵ (6-8 ans)</p> <p>Je vais assis sur les chaises. (12-14 ans)</p> <p>Je vais promener au chute au Montmonrency. (15-17 ans)</p>	<p>Ce jeu joue à quatres. (2 à 6^{ème} primaire)</p>	<p>Ils ont mis en colère. (17 ans, L1 chinois)</p> <p>J'arrive pas de reposer. (corpus oral, adultes, L1 variées : arabe, espagnol, portugais, malien)</p> <p>Je suis allé coucher. (13-14 ans L1 anglais ou 5^{ème} primaire, L1 variées : vietnamien, arménien, syrien, espagnol)</p>
pronom en trop	<p>Je m'arrive dans le télémétropole. (12-14 ans)</p> <p>Je me sens que ma mère n'est pas tannée à cause je travaille et elle se repose. (12-14 ans)</p> <p>Un matin d'hiver, Eric vas se réveille Alain Elmalch pour aller à la chasse. (15-17 ans)</p>	<p><i>aucun exemple donné</i></p>	<p>Il se sont tous morts. (corpus oral, adultes, L1 espagnol)</p> <p>Je me sens que je suis pas enrhumé. (17 ans, L1 chinois)</p>

Néanmoins, à mon avis, la ressemblance en FS et en FL2 dans l'emploi des verbes pronominaux proposée par Nadeau (1993) et Nadeau et Machabée (1998) n'est pas expliquée et n'est pas évidente dans ces exemples. En effet, la ressemblance entre ces énoncés me semble n'être que de surface, surtout pour les ajouts (qui impliquent une compréhension différente des fonctions que le pronom SE peut réaliser). L'ajout en FL2 de l'énoncé « *Ils se sont tous morts* » pourrait provenir du verbe *morirse* en espagnol, et donc d'un transfert négatif à partir de la phrase grammaticale *Todos se mueren*; d'ailleurs, on retrouve la forme *se morir* dans certaines références en français, par exemple dans le dictionnaire Antidote (Antidote, 2022a). Pour le deuxième exemple (*Je me sens que je suis pas enrhumé*), il semble provenir d'une confusion entre *se sentir* (verbe attributif) et *sentir que*; plus précisément entre *Je (ne) me sens pas enrhumé* et *Je sens que je (ne) suis pas enrhumé*, deux constructions similaires au sens équivalent. On retrouve le même verbe dans un

²⁵ Nadeau (1993) remarque, avec raison, qu'on ne peut affirmer avec certitude que la cible ici était *Les garçons se battent souvent avec les filles* (omission du SE), et pas *Les garçons battent souvent les filles* (préposition *avec* en trop).

exemple d'ajout en FS (*Je me sens que ma mère n'est pas tannée à cause je travaille et elle se repose*). Or, ici, il n'y a pas de confusion possible entre *sentir que* et *se sentir*; la phrase subordonnée introduite par le verbe *sentir* n'a pas le même sujet que la phrase principale et ne contient pas d'attribut lié sémantiquement au sujet de la phrase principale. Bref, les exemples d'ajout de pronom réfléchi en FL2 semblent avoir une origine claire, mais ce n'est pas le cas pour ceux en FS. Ceci est particulièrement vrai pour le premier exemple (*Je m'arrive dans le télé métropole*), puisque le verbe *arriver* ne peut jamais être pronominal en français et qu'il n'y a pas de transfert évident de la LSQ, puisqu'il n'y a pas d'équivalence exacte au pronom SE en LSQ (voir 3.1.6).

Bref, Nadeau (1993) puis Nadeau et Machabée (1998) soulèvent que, même si l'omission et l'ajout erroné du pronom SE sont attestés dans des corpus d'apprenant·e·s du français L1, cette catégorie d'erreur est davantage caractéristique des apprenant·e·s du français L2 et des apprenant·e·s sourd·e·s.

Hors du Québec – et, plus récemment –, la thèse de Perini (2013), pour laquelle des corpus ont été élaborés en France, contient une section sur l'emploi des verbes pronominaux dans des écrits en français de Sourd·e·s. L'autrice note des omissions du pronom réfléchi tant dans son corpus écrit par des Sourd·e·s que dans celui écrit par des entendant·e·s FL2, ainsi qu'un seul ajout dans le texte d'une Sourde.

En résumé, bien que l'emploi des verbes pronominaux du français dans les textes de Sourd·e·s n'ait jamais été le sujet principal d'une étude, il a tout de même été décrit quelques fois dans la littérature, soit dans des recherches à portée plus large. Il est accepté que, de manière générale, le SE est fréquemment omis ou de trop (Nadeau *et al.*, 1991). Le pourcentage d'erreur diminue légèrement avec l'âge (Dubuisson et Nadeau, 1993), mais l'emploi des verbes pronominaux demeure une difficulté récurrente et persistante (Nadeau, 1993). L'omission ou l'ajout de SE est attesté dans des textes d'entendant·e·s français langue première, mais n'est pas une catégorie d'erreur reconnue pour FL1, tandis que c'est le cas pour les entendant·e·s français langue seconde et pour les Sourd·e·s (Nadeau, 1993; Nadeau et Machabée, 1998; Perini, 2013).

2.5 Objectif général de recherche

À partir des observations de Nadeau *et al.* (1991), de Dubuisson et Nadeau (1993), de Nadeau (1993), de Nadeau et Machabée (1998) et de Perini (2013), je souhaite déterminer si l'emploi des verbes pronominaux pourrait être une forme typique sourde (c'est-à-dire partagée par la communauté sourde, plutôt qu'être idiosyncratique) et identifier les facteurs influençant leur production. Pour ce faire, j'utilise un corpus recueilli par le Groupe de recherche sur la LSQ et le bilinguisme sourd de l'Université du Québec à Montréal (dorénavant « corpus Groupe LSQ 98 »), qui a été utilisé pour la description de plusieurs autres constructions typiques du français sourd. Bien que ce corpus contienne une grande fréquence d'utilisation de formes verbales réfléchies, il n'a pas été étudié en fonction de l'utilisation spécifique du SE en français sourd.

L'objectif général de ce mémoire est donc la description de l'emploi du pronom préverbal SE dans un corpus de textes identifié comme étant porteur des caractéristiques du français sourd, soit le corpus Groupe LSQ 98.

CHAPITRE 3 : CADRE THÉORIQUE

L'objectif général de ce mémoire est la description de l'emploi du pronom préverbal SE dans un corpus de textes identifié comme étant porteur des caractéristiques du français sourd, soit le corpus Groupe LSQ 98. Le cadre théorique se divise en deux volets : le premier porte sur des notions de bilinguisme français – LSQ (3.1), tandis que le deuxième se concentre sur des éléments du français (3.2). Ce chapitre se conclut avec la formulation des questions spécifiques de recherche (3.3).

3.1 L'écriture sourde dans une approche bilingue et biculturelle

Nous avons évoqué précédemment que, selon l'hypothèse du transfert linguistique, l'emploi du français écrit par les Sourd·e·s est influencé par les structures de la langue des signes québécoise (LSQ). Ainsi, des transferts positifs surviennent lorsque les structures de la langue de référence (ici, la LSQ) sont jugées similaires à celles de la langue cible (ici, le français); lorsque les structures diffèrent, les transferts sont négatifs. Cette hypothèse permet d'expliquer l'origine des formes typiques sourdes (FTS) qui forment l'écriture sourde : lorsqu'elles présentent un écart à la norme, les FTS seraient le résultat d'un transfert négatif. Dans cette optique, puisque la fonction des pronoms personnels du français est de marquer la référence syntaxique, les FTS liées à leur emploi proviendraient des structures de marquage de la référence syntaxique en LSQ (emploi de traces spatiales), qui sont bien différentes de celles qui régissent l'utilisation des pronoms en français (emploi d'éléments lexicaux portant les traits d'accord du référent).

Ce premier volet débute par une définition de la notion de transfert linguistique (3.1.1), qui est centrale à la théorie du français sourd (3.1.2). Les caractéristiques du français sourd sont présentées en (3.1.3). Ensuite, on retrouve une comparaison plus détaillée entre les structures de marquage de la référence en français et en LSQ (3.1.4) afin de contextualiser les formes typiques sourdes dans l'emploi des pronoms personnels en français sourd (3.1.5). Une telle comparaison n'est pas possible pour l'emploi du pronom SE en particulier, car il n'existe pas de correspondance directe en LSQ pour les verbes pronominaux du français. Par contre, les éléments de la LSQ correspondant partiellement à ce que les verbes pronominaux permettent d'exprimer (donc de potentielles sources de transfert négatif) sont présentés dans cette section (3.1.6).

3.1.1 Transferts linguistiques dans l'acquisition d'une langue seconde

Lors de l'acquisition d'une langue seconde (dorénavant L2), les langues déjà connues auront un effet sur développement de la L2, parfois positif, parfois négatif. En effet, les études sur le bilinguisme démontrent qu'une personne apprenant une langue seconde se basera sur le bagage linguistique qu'elle a déjà acquis : ses connaissances de sa langue maternelle (L1), celles d'autres langues secondes et celles de la langue cible – même si ces connaissances s'avèrent incomplètes ou fausses (Hummel, 2014; Schachter, 1983).

C'est dans les productions en langue seconde que l'on retrouve les transferts, c'est-à-dire un recours à une structure d'une langue source (telle que la L1) pour la production d'une structure jugée similaire dans la langue cible (Ringbom, 2007). Ces transferts peuvent être positifs et faciliter l'apprentissage de notions de la L2, ou bien être négatifs et causer des interférences (Odlin, 1989). Par exemple, en tentant de produire les phrases en français (24), des locuteur·rice·s natif·ve·s de l'anglais ont produit les formes en (25), présentant un transfert du placement différent des pronoms objet en anglais (26) (exemples tirés de Swain *et al.* (1974)) :

- | | | |
|------|----------------------------|-------------------------|
| (24) | a) Le chien les a mangés. | b) Il les veut encore. |
| (25) | a) *Le chien a mangé les. | b) *Il veut les encore. |
| (26) | a) The dog has eaten them. | b) He wants them still. |

La production en L2 provient d'une série de décisions, lesquelles dépendent de plusieurs processus cognitifs, dont les transferts linguistiques (Gass, 1984). Les transferts ne sont donc pas la seule source d'explication des stades d'acquisition d'une L2, et ils ne correspondent pas à un mécanisme automatique. Ainsi, les apprenant·e·s de l'anglais L2 ayant le français comme L1 ne font pas l'erreur (27), pourtant analogue de (25).

- | | | |
|------|-----------------------|--------------------------|
| (27) | a) *The dog them ate. | b) *He them still wants. |
|------|-----------------------|--------------------------|

De plus, l'association d'une forme erronée dans une L2 à une structure similaire en L1 ne signifie pas nécessairement que l'erreur est le résultat d'un transfert. En effet, il arrive que la même erreur se retrouve dans les productions L2 d'apprenant·e·s n'ayant pas cette structure dans leur L1.

Puisque plusieurs facteurs influencent les productions en L2, il est difficile d'identifier d'un point de vue extérieur la source précise de l'erreur. C'est pourquoi certaines définitions du transfert incorporent un critère d'analyses de fréquence statistiquement significatives pour identifier une forme comme étant le résultat d'un transfert (Gass, 1984).

La notion de transfert est tout de même récurrente dans les études sur le français sourd, puisqu'elle permet d'expliquer adéquatement les formes typiques sourdes (FTS) par les structures de la LSQ²⁶, étant donné que la grammaire de la LSQ est un reflet d'un traitement de l'information propre aux personnes ayant une surdité de naissance (voir 3.1.2).

3.1.2 Le français sourd – théorie

Le recours à la LSQ comme langue de référence pour l'acquisition du français entraîne des transferts négatifs, puisque la grande différence entre les structures de ces langues limite considérablement la possibilité de transferts positifs. Cette proposition a été confirmée par les études qui ont observé les écarts à la norme dans des textes de personnes sourdes. Charrow (1975) s'est penchée sur l'emploi de l'anglais écrit par des Sourd·e·s des États-Unis et a remarqué que, dans la passation d'un test standardisé d'anglais L2, les Sourd·e·s faisaient tou·te·s des erreurs aux mêmes numéros – une régularité qu'on ne retrouvait pas dans les tests des groupes contrôles d'entendant·e·s – et, de plus, produisaient les mêmes formes erronées d'un individu à l'autre. Des professeur·e·s enseignant à des Sourd·e·s ont reconnu ces erreurs comme étant des *deafisms*, des formes récurrentes et communes aux Sourd·e·s. Cela a motivé Charrow (1975) à proposer que les Sourd·e·s ont leur propre dialecte de l'anglais, le *Deaf English*, qui diffère de l'anglais standard. En effet, puisque les *deafisms* sont réguliers et partagés par une communauté linguistique, il convient de ne pas les considérer comme des erreurs individuelles.

Les études portant sur l'emploi du français par des Sourd·e·s au Québec ont observé que les écarts à la norme varient très peu d'un individu à l'autre et qu'il est possible, en observant ces écarts à la norme, d'en faire ressortir des traits caractéristiques, c'est-à-dire des formes typiques à l'ensemble

²⁶ Il faut souligner que la notion de *français sourd* ne désigne pas une réalité unique et universelle : en plus de la LSQ, il existe d'autres langues des signes employées dans des communautés francophones, telles que la langue des signes française (LSF) et la langue des signes française de Belgique (LSFB). Les différences morphosyntaxiques entre ces langues peuvent se refléter dans les variétés de français sourd qui leur sont associées.

de ces textes, mais qui ne sont pas produites par d'autres populations (Lacerte, 1989; Nadeau, 1993). Ces observations ont été réalisées dans plusieurs autres langues, dont le portugais – influencé par la langue des signes brésilienne – (Bernieri-Souza et Vasconcelos, 2021) et le néerlandais – influencé par la langue des signes flamande – (Van Herreweghe et Vermeerbergen, 2003). En français, plusieurs études ont tenté d'identifier des variations individuelles, sans succès, dont Belzile (1975), qui n'a observé aucun effet d'âge de début de la scolarisation, de quotient intellectuel, de degré de perte auditive, ni d'habitude quant au port de prothèse auditive; et Daigle (1995), qui montre qu'il n'y a pas de variation selon le degré d'exposition au français. De plus, ces écarts à la norme ne peuvent s'expliquer que par les limitations physiques dans l'apprentissage de la langue orale ou les difficultés à en faire une lecture labiale (Nadeau *et al.*, 1991).

L'hypothèse de Charrow (1975) a été reprise par Nadeau (1993), qui a proposé de désigner l'emploi du français dans les écrits de Sourd·e·s en tant que *français sourd*. Lacerte (1989) a introduit la notion d'*écriture sourde*, puis Vincent-Durroux (1990) a proposé les *deaf languages* pour désigner de manière générale l'emploi à l'écrit et à l'oral de langues vocales par les personnes sourdes. Plusieurs études ont proposé qu'une surdité de naissance (ou qui survient en bas âge) affecte le traitement de l'information (Dubuisson *et al.*, 1991; Lacerte, 1989; Nadeau *et al.*, 1991); cette différence dans le mécanisme d'encodage et de décodage serait la source des traits typiques aux Sourd·e·s dans l'emploi des langues vocales (langue orales transmises par la voix). Cette proposition est renforcée par le fait que les formes typiques sourdes persistent même chez les personnes ayant un implant cochléaire moderne permettant d'entendre les fréquences de conversation (Vincent-Durroux, 2009). Par ailleurs, la différence entre les textes de Sourd·e·s gestuels (utilisant principalement la LSQ) et ceux d'oralistes (lisant sur les lèvres, ayant le français comme langue maternelle) n'est que quantitative, pas qualitative (Nadeau *et al.*, 1991), ce qui suggère que les formes typiques sourdes ne se limitent pas à des interférences de la LSQ.

Nous avons vu précédemment (3.1.1) que certaines formes employées par des apprenant·e·s bilingues s'expliquent par la théorie des transferts linguistiques – même s'ils ne rendent pas compte de l'entièreté des processus cognitifs impliqués dans la production en L2. La même chose est vraie en français sourd : Lacerte (1989, p. 338) souligne que l'écriture des personnes sourdes « ne relève pas tout à fait du français, ni tout à fait de la LSQ ». Plus précisément, le français sourd est influencé par le traitement de l'information davantage visuel découlant de la surdité, qui lui-même est visible

au-travers de la LSQ, dont la grammaire dérive de cet encodage typique aux Sourd·e·s (Daigle *et al.*, 1994).

3.1.3 Le français sourd – description

De manière générale, les textes de Sourd·e·s sont décrits comme étant plus courts et contenant davantage de phrases simples que ceux d’entendants : Nadeau *et al.* (1991) indiquent que la longueur moyenne d’un texte d’étudiant·e universitaire entendant·e est 50% plus élevée que celle de leurs collègues sourd·e·s, bien que leurs niveaux de compréhension en lecture soient équivalents. Les textes de Sourd·e·s sont composés de phrases courtes, souvent coordonnées ou juxtaposées, présentant une moins grande variété lexicale (Lacerte, 1989). Selon Power (1971), ce n’est qu’à partir de 17 ans qu’un·e Sourd·e pourrait atteindre la longueur moyenne de phrases d’un enfant entendant de 8 ans.

Les phrases agrammaticales des textes de Sourd·e·s sont quantitativement et qualitativement différentes de celles de textes d’entendant·e·s. En effet, elles sont relativement rares dans les textes d’entendant·e·s et sont en majorité une influence de l’oral. Les autres phrases agrammaticales sont circonscrites à une structure précise ou un contexte précis, tel que l’omission de mots lexicaux dans la dernière phrase d’un texte attribuable au relâchement de l’attention (Nadeau, 1993; Nadeau *et al.*, 1991). Au contraire, un des traits distinctifs du français sourd est la présence de phrases dites « incorrigibles » (Daigle *et al.*, 1994; Nadeau, 1993) :

(28) Je jaune le taxi alle les villes. (Nadeau, 1993)

(29) Quand fin manger va voir où manger au tire de sucre deux fois. (Daigle *et al.*, 1994)

Ces phrases sont possiblement le résultat d’une translittération de la phrase équivalente en LSQ, c’est-à-dire les éléments lexicaux des signes sans les structures grammaticales pour les relier, tel que décrit par Van Herreweghe et Vermeerbergen (2003) dans l’emploi du néerlandais influencé par la langue des signes flamande. Tandis que les langues vocales sont linéaires, les langues des signes sont en trois dimensions : elles expriment des relations grammaticales en utilisant l’espace autour des locuteurs, locutrices. En français, ces relations sont indiquées dans l’ordre des mots (*x frappe y* est différent de *y frappe x*) ou via l’utilisation de mots grammaticaux, tels que des

prépositions (*partir au bureau* est différent de *partir du bureau*) (Nadeau *et al.*, 1991). En LSQ, les mouvements (notamment ceux des verbes directionnels et des verbes à classificateurs) rendent évidentes les relations grammaticales : Nadeau (1993) propose qu'il est ardu pour le scripteur sourd ou la scriptrice sourde de passer d'une représentation tridimensionnelle à une description linéaire.

Les formes typiques sourdes se retrouvent à tous les niveaux : organisation du texte, syntaxe de la phrase, choix lexicaux et marques morphologiques. Environ le 2/3 des écarts à la norme relève de la syntaxe et ce sont surtout des omissions de mot (1/2 des écarts); viennent ensuite – par ordre d'importance – la substitution, l'addition et la modification de l'ordre des mots (Lacerte, 1989). Selon Nadeau *et al.* (1991), Lacerte (1989), Niederberger et Berthoud-Papandropoulou (2004)²⁷, Daigle et Dubuisson (1998) et Nadeau (1993), les catégories les plus communes d'écarts à la norme sont les suivantes :

- Déterminants : omissions (surtout pour les déterminants indéfinis), confusions morphologiques, confusions défini – indéfini;
- Prépositions : omissions dans plus du tiers des cas (surtout pour *à* et *en*), choix inadéquat, ajouts, redondances;
- Pronoms : ruptures dans le réseau anaphorique, omissions de pronoms (surtout de clitiques objets, aussi de clitiques sujets²⁸), ajouts;
- Verbes et désinences : omissions de verbes (surtout les copules, les auxiliaires et les semi-auxiliaires), confusions de temps et mode, confusions des auxiliaires *être* et *avoir*;
- Choix lexicaux : utilisations fautives de séquences figées, choix inadéquat de verbe, emplois d'expressions de la LSQ, juxtapositions d'éléments redondants, connaissances lexicales limitées, glissements de catégorie nom – verbe;
- Marques morphologiques incorrectes, dont quelques confusions de genre.

²⁷ Niederberger et Berthoud-Papandropoulou (2004) décrivent le français sourd employé en France; il peut donc présenter des différences structurelles à celui employé au Québec.

²⁸ Lacerte (1989) note un taux d'omission du clitique sujet de 21%, mais ce chiffre comprend les occurrences où le pronom *moi* est employé à la place du pronom *je*.

Les difficultés en français écrit proviennent principalement de l'usage de mots grammaticaux, tels que les prépositions, les pronoms, les déterminants et les auxiliaires. Or, en LSQ, ces catégories sont très peu employées ou n'existent simplement pas (Nadeau *et al.*, 1991). Néanmoins, il y a des limites à dresser le portrait du français sourd à l'aide de catégories d'erreurs, qui découlent de choix méthodologiques lorsque plus d'une correction est possible, surtout lorsque les sources d'agrammaticalités sont multiples (Nadeau, 1993); ce point est davantage discuté en (3.1.5).

Les écarts à la norme dans la conjugaison verbale sont très caractéristiques du français sourd. Selon Nadeau (1993), en français langue première, les erreurs relèvent surtout de la concordance des temps dans une phrase ou de l'homophonie, tandis que les sources d'écarts à la norme sont beaucoup plus variées dans les textes de Sourd·e·s. La grande majorité des verbes est conjuguée au présent, suivi en popularité par l'infinitif et le passé composé. Les autres temps verbaux, bien que rares, sont tout de même utilisés – donc connus (Lacerte, 1989). La flexion verbale ne donne pas d'information sur le temps de l'énoncé : ce rôle est rempli par des marqueurs temporels (tels que des adverbes), le plus souvent en début de phrase, et les phrases se suivent dans un ordre chronologique. Lorsque l'on prend en considération que le temps de l'énoncé est marqué différemment en français sourd, l'usage abondant d'indicateurs temporels (adverbes en début de phrase ou noms précis) fait en sorte qu'il n'y a pas d'ambiguïté dans les textes (Lacerte, 1989).

Une autre caractéristique marquée du français sourd est la grande quantité d'omissions de mots. La majorité de celles-ci peut s'expliquer par un principe d'économie : inutile de répéter un élément déjà mentionné dans les phrases précédentes (Daigle et Dubuisson, 1998). Ainsi, selon Lacerte (1989), tout syntagme nominal, pronom ou verbe est facultatif en français sourd si l'information est jugée redondante, parce que déjà exprimée. Néanmoins, certaines absences ne peuvent pas s'expliquer par un principe d'économie, notamment celles de déterminant et de la copule *être*. Celles-ci sont possiblement le résultat d'un transfert linguistique, étant donné qu'une phrase comme « MAMAN SUPRISE » est tout à fait grammaticale en LSQ.

La syntaxe des textes de Sourd·e·s est aussi marquée par un ordre de mots non normés, surtout dans les syntagmes nominaux. L'adjectif employé dans un syntagme nominal aura un positionnement non normé dans 20% des cas (Lacerte, 1989); on retrouve notamment l'ordre adjectif + déterminant + nom, par exemple dans « *long des poils* » ou « *je regarde belles les*

lumières » (Nadeau, 1993). Parfois le complément du nom est inversé avec le noyau, comme dans « *chocolat petite lapin* » pour ‘petit lapin de chocolat’ (Lacerte, 1989) ou « *la chambre du lit de oreiller* » pour ‘l’oreiller du lit de la chambre’ (Nadeau, 1993).

Le positionnement non normé des mots touche aussi les compléments du verbe et les adverbes. On observe également une sur-généralisation des règles de construction de phrases négatives. La plus courante est la règle *ne + (verbe) + pas*, dans laquelle le verbe est remplacé par les éléments dans la portée de la négation, par exemple « *ma mère dit ne belle cheveux pas* » pour ‘la mère dire que la fille n’est pas belle avec ses cheveux courts’ (Lacerte, 1989). On retrouve aussi *ne pas + (verbe infinitif)*, par exemple dans « *père et mère ne pas beaucoup signe parce que [...]* » pour ‘mes parents ne connaissent pas beaucoup les signes’ (Lacerte, 1989).

En ce qui concerne le lexique, une caractéristique distinctive du français sourd est le glissement de catégorie entre noms et verbes. En LSQ, les noms et les verbes sont fréquemment semblables et ne se différencient que dans la reduplication du signe et dans la vitesse d’exécution du mouvement (Lacerte, 1989). Par ailleurs, la structure argumentale des verbes est souvent simplifiée en français sourd. La plupart des verbes demandant un complément introduit par une préposition sont employés sans préposition, comme dans « *je veux aller vacances* » et « *mère entre un cuisine* »; or, les prépositions sont plus rarement omises lorsque le syntagme prépositionnel n’est pas sélectionné par le verbe : « *Moi aller école Lucien Pagé à Montréal* » (Lacerte, 1989). Certains verbes transitifs deviennent intransitifs, ou alors ont un emploi intransitif : « *[...] et la fille étonne.* » (Lacerte, 1989). Les textes contiennent aussi des séquences de 2-3 mots souvent cooccurents en français, figés et employés dans un sens non normé en français sourd, par exemple l’énoncé « *après à vendre à la maison* » pour ‘après nous allons mettre la maison à vendre’ (Lacerte, 1989). Néanmoins, la création de nouveaux sens pour un mot et la création lexicale sont surtout idiosyncratiques, ou alors elles relèvent d’une interférence avec la LSQ.

3.1.4 La distance entre le français et la LSQ dans la référence syntaxique

Nous venons de voir que, parmi les FTS qui composent le français sourd, on retrouve un emploi des pronoms personnels différent de celui du français écrit normé et de celui des apprenant·e·s entendant·e·s du français (L1 et L2). Tandis qu’en français, les pronoms personnels sont un outil

important dans la cohésion du discours, leur contribution est moindre en français sourd et l'information se retrouve souvent ailleurs. Avant de préciser la description de l'emploi des pronoms personnels en français sourd, cette section vient comparer les structures de marquage de la référence syntaxique du français (2.2.1) à celles de la LSQ (2.2.2) afin d'identifier les transferts négatifs qui pourraient découler de l'utilisation de la LSQ en tant que langue de référence dans l'emploi des pronoms personnels en français.

Il y a deux différences majeures entre les structures du marquage de la référence syntaxique de la LSQ et celles du français : la forme que peut prendre l'anaphore, ainsi que la nature du lien anaphorique. Nous avons vu précédemment (2.1) que la cohésion d'un discours (ou, plus précisément, d'un texte) repose entre autres sur la répétition d'éléments pour assurer la présence d'un fil conducteur. En français, ce mécanisme (souvent appelé la reprise de l'information) s'effectue en grande majorité par l'emploi de syntagmes nominaux (anaphore nominale ou pronominale) – bien que l'anaphore puisse aussi prendre la forme d'un adverbe, d'un verbe ou d'un adjectif (Riegel *et al.*, 2009).

Dans tous les cas, la référence syntaxique requiert en français l'insertion d'un élément lexical tel qu'un pronom (30), sans quoi la reprise de l'information ne peut être réalisée (31) : il n'existe pas de structure syntaxique ou morphologique ayant une fonction anaphorique.

(30) Hier, tu as vu **Camille**. **Elle / celle-ci** enseigne aujourd'hui.

(31) Hier, tu as vu **Camille**. Tu *(**lui**) as téléphoné aujourd'hui.

En LSQ, l'anaphore peut prendre une plus grande variété de formes : elle peut être réalisée par une insertion lexicale, par exemple le signe POINTÉ (32, dont le sens correspond à 30), ou alors par une stratégie morphosyntaxique, telle que la flexion verbale (33, dont le sens correspond à 31).

(32) HIER **CAMILLE** PTÉ_(x) 2-VOIR-3_(x) AUJOURD'HUI **PTÉ3_(x)** ENSEIGNER

(33) HIER **CAMILLE** PTÉ_(x) 2-VOIR-3_(x) AUJOURD'HUI 2-TÉLÉPHONNER-**3_(x)**

Non seulement l'anaphore en LSQ peut prendre des formes qui ne sont pas disponibles en français, aussi entretient-elle un lien différent avec son antécédent. En LSQ, le lien entre le locus et le

référent est global; la référence syntaxique faite via un locus est équivalente à la répétition de l'antécédent en français. En français, pour éviter la lourdeur d'une répétition, il est possible d'employer une anaphore entretenant un lien sémantique avec le référent. Dans le cas d'une anaphore pronominale, ce lien prend la forme d'un partage de traits grammaticaux : le pronom établit une référence syntaxique avec son antécédent en prenant les mêmes traits de personne, de nombre et, pour les pronoms de troisième personne, de genre. Ainsi, dans la phrase *Jérôme et Sandrine vont à l'hôpital car elle est malade*, l'antécédent du pronom *elle* est clairement *Sandrine* (voir 2.1). Prenons maintenant la phrase en (34) :

- (34) Marine et Ashley m'ont appris que leurs voisines n'ont pas encore dit à leurs mères qu'**elles** allaient partir en voyage.

Dans cette phrase, les traits réalisés par le pronom *elles* (3^{ème} personne, féminin, pluriel) ne permettent pas d'identifier sans ambiguïté un référent unique entre a) Marine et Ashley; b) leurs voisines; c) leurs mères ou d) une combinaison des réponses précédentes. En LSQ, il n'y a pas de telle ambiguïté. Le référent est associé dans son entièreté à un locus (trace spatiale), sans passer par une concordance de traits morphologiques. Comme chaque locus ne peut être associé qu'à un seul référent et que chaque référent ne peut être associé qu'à un seul locus; le lien entre le référent et sa reprise sémantique est beaucoup plus explicite que celui en français (Lillo-Martin et Klima, 1990). Si on assigne le locus (w) à *Marine*, le locus (x) à *Ashley*, le locus (y) au référent *leurs voisines* et le locus (z) au référent *leurs mères*, il est possible d'établir une référence univoque en réactivant la trace spatiale du ou des locus appropriés. Ainsi, le lien direct entre le référent et l'unité linguistique de la LSQ limite considérablement la probabilité de générer une ambiguïté.

En LSQ, l'association entre un locus et son référent est en théorie permanente (jusqu'à un potentiel remplacement) et le nombre de locus qu'il est possible d'établir dans un discours est en théorie infini : les limitations ne proviennent pas du système linguistique, mais des facultés cognitives des interlocuteur·trice·s (Schlenker, 2017). Par principe d'économie, il est fréquent que la relation anaphore/antécédent soit réalisée en employant une stratégie de réactivation de trace spatiale plutôt qu'en effectuant une insertion lexicale. Au contraire, en français, dans toutes les relations

anaphoriques, l'anaphore doit être réalisée lexicalement. De plus, la reprise de l'information par l'utilisation d'un pronom peut être ambiguë en français, ce qui est fort moins probable en LSQ²⁹.

3.1.5 L'emploi des pronoms personnels en français sourd

Lacerte (1989) identifie comme caractéristique de l'écriture sourde une faiblesse dans le lien anaphorique permettant l'unité du texte, qu'elle explique par le peu d'utilisation de mots référentiels³⁰ et de pronoms. Nadeau *et al.* (1991) décrivent une « perturbation » du réseau anaphorique dans les textes de Sourd·e·s, causée principalement par le haut taux d'absence de pronoms sujets, de pronoms relatifs et de pronoms clitiques objets. Dans une tâche d'élicitation à partir de vignettes illustrées, Niederberger et Berthoud-Papandropoulou (2004) ont noté un taux de réussite de production écrite de pronoms sujet et pronoms complément du verbe de 32% pour un groupe de Sourd·e·s âgé·e·s de 8 à 14 ans. En faisant fi des erreurs d'accord en genre et en nombre, ce taux montait à 36%, comparativement à 60% chez des enfants entendants de 7 ans et à 75% pour les 9-10 ans.

Ainsi, les textes de Sourd·e·s comportent de manière générale peu de pronoms. En France, Perini (2013) note une préférence pour la répétition de syntagmes nominaux plutôt que l'emploi de la pronominalisation ou d'autres stratégies de reprise de l'information (tels que des synonymes ou des paraphrases). Même à une échelle plus petite, soit dans une seule phrase, Niederberger et Berthoud-Papandropoulou (2004) notent beaucoup plus de répétitions de l'antécédent dans le groupe de Sourd·e·s que dans le groupe d'entendant·e·s. Les mêmes points saillants se dégagent d'une revue de la littérature sur le *Deaf English* (Daigle et Dubuisson, 1998); la pronominalisation serait la structure la plus difficile à maîtriser en anglais pour les Sourd·e·s (Quigley *et al.*, 1977).

Il arrive aussi fréquemment qu'il n'y ait ni répétition, ni pronom : la référence n'est simplement pas réalisée lorsqu'elle est jugée comme étant facile à inférer (Perini, 2013). Cela pourrait provenir

²⁹ Certaines personnes vont jusqu'à affirmer que l'ambiguïté de référent est impossible en LSQ, mais selon Cormier *et al.* (2013), c'est un propos à nuancer. Comme le souligne une évaluatrice, une ambiguïté peut survenir, notamment lorsqu'il est difficile de discriminer les locus (s'ils sont trop nombreux et trop peu distancés dans l'espace), lorsque leur dernière activation est trop éloignée dans le temps ou après un changement de la configuration de l'espace référentiel.

³⁰ Néanmoins, Lacerte (1989) n'indique pas clairement ce que l'expression « mot référentiel » désigne.

d'un transfert négatif de la LSQ, qui permet de construire des relations anaphoriques de manière spatiale plutôt que lexicale (Nadeau *et al.*, 1991). Cette hypothèse est appuyée par le fait que Niederberger et Berthoud-Papandropoulou (2004) ont trouvé un taux de réussite de production de pronom nettement supérieur pour les pronoms sujets (47%) que pour les pronoms compléments du verbe (9%). Or, en LSQ, les compléments du verbe dont le référent a déjà été introduit dans le discours sont bien souvent incorporés dans le mouvement du verbe (dans le cas de verbes souples) ou marqués par des marques morphologiques simultanées.

Les productions en français sourd comprennent aussi des confusions entre la 1^{ère} et la 3^e personne et entre la 2^e et la 3^e personne dans l'emploi des pronoms personnels (Perini, 2013). Nadeau (1993) souligne que la confusion entre les pronoms de la 1^{ère} et de la 3^e personne est aussi relevée dans des textes d'enfants francophones entendants, mais seulement dans un contexte précis, soit quand l'élève s'identifie au personnage d'une histoire et que le texte passe d'un point de vue objectif écrit à un point de vue subjectif écrit au *je*. Or, on retrouve dans des textes de Sourd·e·s des va-et-vient entre les 1^{ère} et 3^e personnes, même dans des narrations d'anecdotes personnelles.

Dans les textes narrés à la première personne, on observe une tendance à employer (*à*) *moi* plutôt que *je* et *me*. Mis à part cela, les erreurs d'accord des pronoms sont rares, étant même moins fréquentes que chez les entendant·e·s (Niederberger et Berthoud-Papandropoulou, 2004). En effet, les enfants entendant·e·s ont tendance à faire des erreurs d'accord lorsque ces traits ne s'entendent pas à l'oral, notamment l'accord du pluriel (par exemple **elle parlent*). Même lorsqu'on prend en considération le plus faible nombre de pronoms personnels dans les textes de Sourd·e·s, on observe qu'il y a proportionnellement moins d'erreurs d'accord en genre et en nombre.

Il demeure toutefois difficile de dresser un portrait précis de l'emploi des pronoms à partir de catégories d'erreurs. Prenons les phrases (35) à (41), lesquelles donnent des exemples d'erreurs de pronom complément en trop dans des textes de Sourd·e·s, d'enfants francophones et d'apprenant·e·s du français langue seconde (exemples de Nadeau (1993)) :

- | | | |
|------|---------------|---|
| (35) | (français L1) | Il porte les cheveux bruns et il y avait un fusil. |
| (36) | (français L1) | Si un joueur trouve une lettre, sa <u>le</u> lui fait 5, 10 ou 15 points. |
| (37) | (français L1) | ... pour que ceux qui <u>en</u> font partie de ce traficage arrêtent. |

- (38) (français L2) Dans un restaurant italien qu'on le nommait Da Giovanni.
- (39) (français L2) Pour qu'elle puisse l'appliquer le produit sur elle.
- (40) (français sourd) Moi, je t'aime le chat parce qu'il est gentil.
- (41) (français sourd) J'ai très hâte de te voir mes parents.

Bien que ces phrases correspondent toutes à la même catégorie d'erreur, elles ne sont pas nécessairement qualitativement équivalentes. En effet, les erreurs des entendants peuvent s'expliquer soit par l'influence d'une cooccurrence fréquente (35 et 36, *il y a* et *ça le fait*), soit par un redoublement du complément verbal (37, 38, 39). Or, les exemples tirés de textes de Sourd·e·s (40 et 41) sont l'ajout d'un référent de deuxième personne qui n'est mentionné ni ailleurs dans la phrase, ni ailleurs dans le texte.

3.1.6 Les formes de la LSQ correspondant aux éléments réalisés par SE en français

Nous avons vu que l'emploi des pronoms est particulier en français sourd, et que cela vient possiblement de transferts négatifs de la LSQ. Cette section se penche maintenant sur les possibilités de transferts de la LSQ dans l'emploi du pronom SE spécifiquement.

Il n'existe pas de correspondance biunivoque entre un élément de la LSQ et les différents rôles que prend le pronom SE. Je tâcherai donc d'identifier des formes de la LSQ correspondant partiellement à certains éléments réalisés par SE selon les fonctions lexicales, syntaxiques et discursives des trois catégories de SE : SE réfléchi, SE passif et SE lexical (voir 2.3).

Il existe une correspondance lexicale entre le SE réfléchi et les signes *PRO-MÊME* ('*soi-même*', '*toi-même*', '*moi-même*'). Les deux sont désignés en tant que *pronom réfléchi* dans le métalangage, mais les fonctions qu'ils réalisent ne sont pas les mêmes. En français, le SE réfléchi marque une relation réflexive ou réciproque, en plus de générer une référence syntaxique et de participer à la structure argumentale du verbe. Les signes *PRO-MÊME*, quant à eux, peuvent marquer l'aspect réfléchi d'un verbe, mais sont plus souvent une marque d'insistance, de manière similaire à ce qu'on retrouve dans « *Benoît a cuisiné ce plat lui-même* »; ils peuvent aussi servir de marqueur phonologique de frontière de phrase.

Le SE passif marque le passif d'habitude ou de potentiel, ou alors un aspect anticausatif. Il existe en LSQ des structures passives, mais elles n'ont pas ces mêmes propriétés sémantiques; le SE passif n'a pas réellement de correspondant lexical ou syntaxique en LSQ. Néanmoins, les structures de représentation corporelle pourraient présenter une correspondance discursive avec le SE passif, étant donné que les deux impliquent un changement de point de vue. Tandis que l'emploi d'un SE passif modifie la structure thématique du verbe (le sujet du verbe passif n'ayant plus le rôle thématique agent), l'emploi d'une structure de représentation corporelle cause l'adoption de la perspective d'un actant du discours.

Quant au SE lexical, il est difficile d'identifier un élément de la LSQ correspondant, même partiellement. En effet, le SE lexical ne réalise aucune fonction syntaxique, discursive ou lexicale. Son apport sémantique soit est nul (*se méfier*), soit correspond à une notion de processus interne (*s'agenouiller*). De telles formes ne sont pas présentes – ou, du moins, pas attestées – en LSQ : ni verbe précédé d'un clitique sans fonction, ni structure demandant l'établissement d'un lien anaphorique sans que le référent ne remplisse de fonction syntaxique dans la phrase, ni élément lexical ayant un apport sémantique de processus interne.

En résumé, je propose que les formes les plus comparables au SE sont les signes PRO-MÊME pour leur correspondance lexicale avec la catégorie des SE réfléchis et les structures de représentation corporelle pour la correspondance de leur fonction discursive à celle des SE passifs.

3.2 Contextes morphosyntaxiques et sémantiques

Les études sur le français sourd et sur le bilinguisme en général indiquent que les productions en langue seconde dépendent en partie de la grammaire de la langue cible. Néanmoins, les connaissances actuelles sur l'emploi des verbes pronominaux dans les textes de Sourd·e·s sont limitées à l'observation d'écarts à la norme prenant la forme d'ajout ou d'omission du pronom SE (voir 2.4), sans que soient précisées ni les constructions dans lesquelles la norme est plus ou moins bien respectée, ni les variables pouvant avoir une influence sur la décision d'employer ou non un verbe pronominal.

Cette section présente divers contextes morphosyntaxiques et sémantiques qui auraient le potentiel de préciser la description de l'emploi du SE dans le cadre d'une étude exploratoire³¹ : les différents types de verbes pronominaux (3.2.1), les paramètres de conjugaison du verbe (3.2.2), le classement aspectuel des procès verbaux (3.2.3), les traits sémantiques des verbes (3.2.4), les différentes structures transitives possibles (3.2.5), les formes possibles du sujet de la phrase (3.2.6), ainsi que les traits sémantiques du sujet (3.2.7).

3.2.1 Verbes pronominaux : définition et classements

Un verbe est dit pronominal lorsqu'il est précédé d'un pronom personnel conjoint ayant les mêmes traits de personne et de nombre que le sujet du verbe (Grevisse et Goosse, 2016). À la troisième personne (singulier ou pluriel), ce pronom a une forme qui lui est propre, le *se*, qui est celle utilisée dans ce mémoire pour désigner ce pronom ainsi que toutes ses flexions. Traditionnellement, les verbes pronominaux sont sous-divisés selon s'ils sont essentiellement pronominaux (c'est-à-dire que ce verbe n'existe pas sans le SE) ou occasionnellement pronominaux, avec ou sans changement de sens. Malheureusement, cette classification s'est révélée être difficile à appliquer par manque de précision (comment définir un « changement de sens »?).

J'ai donc retenu la classification de Riegel *et al.* (2009), avec de légères modifications. On y retrouve six catégories de verbe pronominal : les verbes essentiellement pronominaux (42), les verbes pronominaux autonomes (43), les verbes pronominaux réfléchis (44), les verbes pronominaux réciproques (45), les verbes pronominaux passifs (46) et les verbes pronominaux neutres (47). À cela s'ajoute une septième catégorie pour rendre compte de l'ensemble des verbes du français : les verbes qui ne permettent jamais un emploi pronominal (48).

- (42) Elle se souvient de toi.
- (43) Il s'est sauvé dans les ruelles.
- (44) Je me sers une deuxième assiette.
- (45) Nous nous parlons de nos desserts favoris.

³¹ La fréquence lexicale du verbe est un facteur intéressant, mais la pertinence de l'observer est amenuisée étant donnée l'absence d'une base de données lexicales permettant la recherche multi-mot (permettant donc de différencier *appeler* de *s'appeler*, par exemple).

- (46) Le clocher se voit bien d'ici.
- (47) Le chien se promène sur le quai.
- (48) Ma fille dort comme une bûche.

La classification d'un verbe occasionnellement pronominal varie selon la construction dans laquelle il est employé. Ainsi, le verbe *APERCEVOIR* peut autant être autonome (*Gérard s'est aperçu de son erreur*) que réfléchi (*Je me suis aperçu dans le reflet de la fenêtre*).

Les différents types de verbes pronominaux sont présentés aux sous-sections suivantes : les verbes essentiellement pronominaux en 3.2.1.1, les verbes pronominaux autonomes en 3.2.1.2, les verbes pronominaux réfléchis et réciproques en 3.2.1.3, les pronominaux passifs en 3.2.1.4 et les verbes pronominaux neutres en 3.2.1.5.

3.2.1.1 Verbes essentiellement pronominaux

Les verbes essentiellement pronominaux n'existent pas dans une forme non pronominale. Le SE, n'ayant pas de fonction syntaxique, fait partie du lexème et est indissociable du verbe.

- (49) Dominique se méfie de Marthe.
- (50) *Dominique méfie elle-même de Marthe.
- (51) *Marthe méfie Dominique.

Certains verbes essentiellement pronominaux sont diachroniquement dérivés de verbes non pronominaux, notamment avec les préfixes *s'auto-* et *s'entre-*, qui soulignent respectivement les sens réfléchi et réciproque. Le préfixe *s'en-*, quant à lui, ajoute l'aspect inchoatif au procès verbal, tel que dans *s'envoler* ou *s'endormir*, mais aussi *s'en aller* ou *s'en venir*. Les verbes essentiellement pronominaux peuvent être intransitifs (*s'évanouir*), transitifs indirects (*s'enfuir de*) ou, plus rarement, transitifs directs (*s'arroger*).

3.2.1.2 Verbes pronominaux autonomes

Tout comme les verbes essentiellement pronominaux, les verbes pronominaux autonomes sont précédés d'un SE sans fonction. Par contre, les verbes pronominaux autonomes ont une contrepartie non pronominale; seulement il n'y a pas de lien de sens entre les deux.

- (52) Aglaé se rend à Trois-Rivières.
- (53) *Aglaé rend elle-même à Trois-Rivières.
- (54) Aglaé rend les livres qu'elle avait empruntés.

L'exemple (52) contient un verbe pronominal autonome, car, comme le démontre (53), le SE n'a pas de fonction et il existe une forme non pronominale n'ayant pas de lien de sens (54). Les verbes pronominaux autonomes peuvent être intransitifs (*se dépêcher*), transitifs directs (*s'approprier quelque chose*), transitifs indirects (*s'apercevoir de*) ou attributifs (*se sentir*).

3.2.1.3 Verbes pronominaux réfléchis et réciproques

Les verbes pronominaux réfléchis et réciproques sont les seuls verbes pronominaux dont le SE a une fonction grammaticale dans la phrase. Plus précisément, le verbe conserve le même sens que lorsqu'il n'est pas pronominal, et la présence du SE n'est qu'« une version particulière de la phrase transitive ordinaire où le sujet et le complément sont coréférents » (Riegel *et al.*, 2009, p. 459). Le SE peut être un complément direct du verbe, un complément indirect du verbe, où bien une montée de possesseur dans des possessions inaliénables.

- (55) Makéda se maquille ('Makéda maquille elle-même').
- (56) Vincent et Oscar se ressemblent ('Vincent ressemble à Oscar et Oscar ressemble à Vincent').
- (57) Hudson se tresse les cheveux ('Hudson tresse ses cheveux').
- (58) Youri et Estelle se serrent la main ('Youri serre la main d'Estelle et Estelle serre la main de Youri').

L'exemple (55) contient un verbe réfléchi : *Makéda* est à la fois « la source et le but du procès » *maquiller* (Riegel *et al.*, 2009, p. 457), comparativement à une version non réfléchie telle que

Makéda la maquille. L'exemple (56) montre un verbe pronominal réciproque ayant comme sujet des éléments coordonnés. Les verbes pronominaux réciproques dénotent plusieurs procès symétriques (c'est-à-dire des procès simultanés ou successifs entre plusieurs éléments d'un ensemble) et nécessitent donc des sujets dénotant un référent pluriel, que ce soit un syntagme nominal pluriel, un pronom dont le référent est pluriel, des éléments coordonnés ou un sujet singulier représentant un collectif (par exemple, *une famille* ou *un groupe*). Les exemples (57) et (58), qui correspondent respectivement à un verbe pronominal réfléchi et un verbe pronominal réciproque, présentent un SE considéré dans ce mémoire comme une montée de possesseurs (voir 2.3). La possession inaliénable *les cheveux de Hudson* de la phrase en (57) peut être exprimée soit par un déterminant possessif (*ses cheveux*), soit par un SE « datif de la totalité affectée dans l'une de ses parties par le procès verbal » (Riegel *et al.*, 2009, p. 457), c'est-à-dire par une montée du possesseur dans le syntagme verbal (*se laver les cheveux*). Cette construction est possible tant avec des procès réfléchis qu'avec des procès réciproques, tels qu'en (58). Puisque le SE des verbes pronominaux réfléchis ou réciproques a une fonction grammaticale dans la phrase et est coréférentiel au sujet, ces verbes ne peuvent être intransitifs ou attributifs, seulement transitifs directs ou indirects.

3.2.1.4 Verbes pronominaux passifs

L'une des caractéristiques des verbes pronominaux passifs est d'avoir un sujet ne pouvant pas être associé au rôle thématique agent, puisqu'il est plutôt le point d'aboutissement du procès verbal – il a, selon (Riegel *et al.*, 2009), le rôle sémantique d'objet du verbe³². Les phrases contenant un verbe pronominal passif (59) ont une phrase active équivalente (60), avec laquelle elles entretiennent une relation similaire à celle entre une phrase active et une phrase passive (61). De plus, il existe une analogie entre certaines phrases pronominales passives (62) et les phrases actives à sujet indéterminé (63).

- (59) Ces cantiques se chantent à l'Action de grâces.
- (60) Les curés chantent ces cantiques à l'Action de grâces.
- (61) Ces cantiques sont chantés à l'Action de grâces.

³² Bien que les sujets de verbes pronominaux passifs sont fréquemment inanimés, ce n'est pas obligatoire; ils peuvent même avoir le trait +humain : *un bon conjoint ne se trouve pas facilement*.

- (62) Mimi préfère les vêtements qui se lavent à la machine.
(63) Mimi préfère les vêtements qu'on peut laver à la machine.

Contrairement aux phrases passives (65), les phrases pronominales passives ne peuvent prendre un complément d'agent exprimé (66), bien qu'il soit possible de le sous-entendre dans des compléments adverbiaux (67), instrumentaux (68) ou de manière (69).

- (64) Le jardinier arrose les fleurs tôt le matin.
(65) Les fleurs sont arrosées tôt le matin (par le jardinier).
(66) Les fleurs s'arrosent tôt le matin (*par le jardinier).
(67) Les fleurs s'arrosent délicatement, tôt le matin (*le jardinier est délicat en arrosant*).
(68) Les fleurs s'arrosent tôt le matin avec un arrosoir (*le jardinier utilise un arrosoir pour arroser*).
(69) Les fleurs s'arrosent tôt le matin en effectuant un mouvement de balayage (*le jardinier effectue un mouvement de balayage pour arroser*).

Les phrases pronominales passives contiennent soit un SE anticausatif (70), ou alors un SE moyen (71), qui permet l'expression de vérités générales, d'habitudes, de normes ou de propriétés – en d'autres mots, de procès réitérés ou ayant une permanence dans un intervalle temporel avec un agent générique.

- (70) Soudainement, la porte s'ouvre.
(71) À cette époque, les foulards se portaient d'une autre façon.

Les verbes pronominaux passifs sont similaires aux réfléchis et réciproques dans leur rapport avec la forme non pronominale : le verbe conserve son sens et ses propriétés, mais sa structure transitive est modifiée. Cette modification requiert un verbe non pronominal transitif direct, qui devient un verbe pronominal passif intransitif.

3.2.1.5 Verbes pronominaux neutres

Les verbes pronominaux neutres ressemblent aux pronominaux passifs en ce qu'ils entretiennent une relation de correspondance avec un verbe non pronominal transitif direct, tandis que la forme pronominale ne peut avoir de complément direct, celui-ci étant possiblement « absorbé » par le SE dans les deux cas. Par contre, dans le cas des verbes pronominaux neutres, le sujet peut être un agent et le verbe ne préserve pas son sens. En effet, la plupart des verbes pronominaux neutres présentent un glissement de sens de leur équivalent non pronominal, mais sans qu'il soit possible de dire que les deux verbes ne présentent pas de lien de sens (comme dans le cas des verbes pronominaux autonomes).

La phrase active en (72) et son équivalent passif (73) sont similaires à la phrase (74), qui contient un verbe pronominal neutre, quoiqu'on observe un glissement de sens entre *promener* et *se promener*. Les verbes pronominaux neutres ne peuvent être interprétés comme des pronominaux passifs (l'exemple (75) est acceptable si on parle d'un chien qui tient une laisse dans sa gueule, mais pas dans le même sens qu'*on promène le chien au bout d'une laisse*). De plus, le SE ne peut pas être interprété comme un véritable complément du verbe; excluant ainsi l'interprétation réfléchie (76).

- (72) Ce jeune homme promène un chien.
- (73) Un chien est promené par ce jeune homme.
- (74) Ce jeune homme se promène avec son chien.
- (75) Le chien se promène (#avec une laisse).
- (76) Ce jeune homme se promène (*lui-même).

L'entièreté du procès d'un verbe pronominal neutre a lieu dans la sphère du sujet; le SE servant à marquer la limite finale du procès. Les verbes pronominaux neutres dénotent soit des processus internes de mouvement (77), soit des changements d'état autocausatifs (78), soit l'occurrence du sujet dans un espace (79).

- (77) Ophélie se dirige vers l'école.
- (78) Mon père s'est réveillé avant que son alarme sonne.
- (79) Elle s'est présentée chez moi sans avertissement.

Les verbes pronominaux neutres ont un équivalent non pronominal transitif direct, et ils peuvent être intransitifs (*s'amuser*), transitifs indirect (*s'approcher de*) ou attributif (*s'appeler*).

3.2.2 Temps de conjugaison

Les conjugaisons verbales varient en fonction des traits de temps, de mode, de personne et de nombre (et, dans le cas de participes passés, de genre). Les verbes peuvent être conjugués à un temps simple ou à un temps composé. Les temps simples regroupent les conjugaisons composées seulement d'une racine verbale et de sa flexion morphologique, tandis que les temps composés désignent les conjugaisons constituées d'un auxiliaire de conjugaison (*être* ou *avoir*), lui-même composé d'une racine verbale et d'une flexion morphologique, ainsi que du participe passé du verbe conjugué (Boivin et Pinsonneault, 2020). Les temps simples dénotent l'aspect inaccompli (80) et les temps composés, l'aspect accompli (81).

- (80) Mon chat dormait paisiblement.
- (81) Mon chat a dormi toute la journée.

Dans certaines constructions, on retrouve, au sein d'une même phrase syntaxique, un verbe conjugué (à un temps simple ou composé) suivi d'un verbe infinitif. Ces constructions, appelées dans ce mémoire *verbe + verbe infinitif*, surviennent lorsqu'on emploie un semi-auxiliaire (82) ou une phrase subordonnée infinitive. Les subordonnées infinitives, résultat de la réduction d'une phrase enchâssée lorsque son sujet est coréférentiel au sujet (ou, plus rarement, d'un autre groupe nominal) de la phrase enchâssante, peuvent être complétives (83), complétives interrogatives (84), relatives (85) ou circonstancielles (86).

- (82) Chloé va déménager en Suède.
- (83) Chloé veut déménager en Suède.
- (84) Chloé ne sait pas où ranger sa vaisselle.

- (85) Chloé cherche des boîtes dans lesquelles ranger sa vaisselle.
- (86) Chloé range sa vaisselle avant de déménager en Suède.

3.2.3 Classement aspectuel

Vendler (1957) a proposé une classification aspectuelle des verbes, que Recanati et Recanati (1999) ont légèrement modifiée afin de l'appliquer à des prédicats conjugués plutôt que des verbes lexicaux. En effet, il est possible de diviser les procès ou situations que les verbes expriment en quatre classes : les états (87), les activités (88), les accomplissements (89) et les achèvements (90).

- (87) Aaron connaît Rokiya.
- (88) Marco a tricoté toute la matinée.
- (89) Agnès a rempli son verre d'eau.
- (90) Roméo a gagné la course.

Ces classes aspectuelles sont définies par plusieurs critères. Afin de classer un procès, il faut déterminer s'il est a) progressif ou statique; b) homogène ou non; c) borné ou non; d) télique ou atélique.

Un procès est dit progressif s'il est dynamique; sinon, il est statique. Les états expriment des procès statiques (91), tandis que les activités (92) et les achèvements (93) dénotent des procès progressifs. Les achèvements (94), étant ponctuels, ne sont pas progressifs à proprement dit.

- (91) Fannie sait nager.
- (92) Jean range des bouteilles.
- (93) Ahmed dessine un cercle.
- (94) J'ai trouvé mes chaussettes.

Un procès est homogène s'il reste vrai durant l'entièreté de sa durée, comme c'est le cas des états (95) et des activités (96). Les accomplissements (97) et les achèvements (98), eux, sont hétérogènes.

- (95) Victor était malade la semaine dernière (et il était malade lundi et mardi.)
- (96) Entre 18h et 20h, Mylène regardait un film (et elle regardait un film entre 19h et 19h30).
- (97) Léa a bâti son cabanon en une semaine (#et elle a bâti son cabanon en deux jours).
- (98) Matthew a éternué (*impossible de diviser en sous-intervalles*).

Certains procès ont une durée intrinsèque; ils possèdent naturellement une borne initiale (un début) et une borne finale (une fin). Les états (99) et les activités (100) peuvent durer indéfiniment et sont donc non bornés, contrairement aux accomplissements (101) qui ont une durée intrinsèque. Les achèvements (102) étant ponctuels, ils n'ont pas réellement de borne initiale ou finale.

- (99) (Il a toujours été vrai et il sera toujours vrai que) Samuel aime le jazz.
- (100) (Il a toujours été vrai et il sera toujours vrai que) cet érable grandit.
- (101) Béatrice a débarré la porte.
- (102) Régis est né hier soir.

La télélicité est la propriété d'avoir une fin intrinsèque. Les états (103) et les activités (104) sont atéliques : leur condition de vérité ne dépend pas de l'atteinte d'un point culminant. Les accomplissements (105) et les achèvements (106), eux, sont téliques : il existe un point final avant lequel on ne peut pas dire que le procès est terminé, et après lequel il ne peut pas continuer.

- (103) Bruno est heureux (vrai au moment de l'énoncé, mais aussi avant et après).
- (104) Jocelyn pousse un chariot (vrai au moment de l'énoncé, mais aussi avant et après).
- (105) Nadja traverse la rue (le procès n'est pas terminé avant l'atteinte du trottoir et ne peut pas durer au-delà de ce point).
- (106) Lina a reconnu la chanson au refrain (le procès n'est pas vrai avant l'instant où la chanson est reconnue, et il ne peut pas durer dans le temps, car il est instantané).

3.2.4 Classement sémantique

Il existe diverses classifications des verbes basées sur des propriétés sémantiques. Celle présentée dans *Les verbes français* (dorénavant LVF) de Dubois et Dubois-Charlier (1997) est accompagnée

d'un recensement de plus de 12 000 verbes, correspondant à plus de 25 000 entrées; la classification est déjà appliquée à la majorité des verbes du français, ce qui assure une analyse constante lors d'une analyse de corpus. La classification de LVF se base sur « l'hypothèse qu'il y a adéquation entre les schèmes syntaxiques de la langue et l'interprétation sémantique qu'en font les locuteurs de cette langue » (Dubois et Dubois-Charlier, 1997, p. III). Ainsi, l'ensemble des possibilités de schèmes syntaxiques a été répertorié, puis divisé en classes. Le contenu de LVF est disponible en ligne (Laboratoire de recherche appliquée en linguistique informatique, 2020), ce qui permet la recherche de verbes; la majorité d'entre eux ont plusieurs entrées associées selon les différents sens et les différentes constructions sémantiques possibles.

La classification de LVF comporte 14 classes génériques qui se divisent en 54 classes sémantico-syntaxiques, elles-mêmes divisées en 248 sous-classes. Les sous-classes sont ensuite divisées en sous-types syntaxiques (par exemple « verbe transitif direct à sujet et objet animé »), et finalement en variantes syntaxiques. Les classes génériques, qui sont celles employées dans ce mémoire, sont présentées au tableau 3.1; les classes sémantico-syntaxiques, au tableau 3.2; et un exemple de sous-classes et de sous-types syntaxiques est présenté à la figure 3.1.

Tableau 3.1 – Les classes génériques de LVF

Code	Classe générique	Nombre d'entrées	Nombre de classes sémantico-syntaxiques
C	Verbes de communication	2039 entrées	4 classes
D	Verbes de don ou de privation	953 entrées	3 classes
E	Verbes de mouvements d'entrée / de sortie	2444 entrées	4 classes
F	Verbes de type <i>frapper</i> ou <i>toucher</i>	1727 entrées	4 classes
H	Verbes d'état physique ou de comportement	2079 entrées	4 classes
L	Verbes locatifs	1523 entrées	4 classes
M	Verbes de mouvement	1671 entrées	3 classes
N	Verbes de type <i>munir</i> / <i>démunir</i>	3175 entrées	4 classes
P	Verbes psychologiques	2074 entrées	3 classes
R	Verbes de type <i>réaliser</i>	2289 entrées	4 classes
S	Verbes de saisie ou de possession	1671 entrées	4 classes
T	Verbes de transformation, de mutation, de substitution	2345 entrées	4 classes
U	Verbes d'union, de réunion, de mise en relation	1407 entrées	4 classes
X	Verbes auxiliaires et semi-auxiliaires, impersonnels, <i>exister</i> , inchoatifs ou résultatifs	212 entrées	4 classes

Tableau 3.2 – Les classes sémantico-syntaxiques de LVF

C1 s'exprimer par un son, une parole C2 dire/demander qc C3 montrer qc C4 figuré de C1 et C2	D1 donner qn à qn/qc D2 donner qc à qn/qc D3 figuré de D2	E1 (faire) sortir/venir de qp ou aller qp, sujet hum E2 figuré de F1 E3 (faire) sortir/venir de qp ou aller qp, sujet non-animé E4 figuré de E3
F1 frapper ou toucher qn F2 figuré de F1 F3 frapper ou toucher qc F4 figuré de F3	H1 être/mettre qn dans état physique H2 avoir tel comportement, activité H3 être/mettre qc dans tel état H4 figuré de H3	L1 être ou mettre qn qp L2 figuré de L1 L3 être ou mettre qc qp L4 figuré de L3
M1 faire/imprimer mouvement, hum M2 figuré de M3 faire/imprimer mouvement, non-animé M4 figuré de M3	N1 munir/démunir qn de qc N2 figuré de N1 N3 munir/démunir qc de qc N4 figuré de N3	P1 avoir tel sentiment, telle pensée P2 faire avoir tel sentiment P3 manifester telle pensée sur qn/qc
R1 mettre en état le corps de qn R2 créer qn, un personnage R3 fabriquer qc, mettre qc en un certain état R4 figuré de R3	S1 saisir, arrêter, serrer qn S2 prendre, choisir, abandonner qn S3 saisir, garder qc ; tordre, tirer qc S4 figuré de S3	T1 lier, unir qn à ; détacher qn de T2 figuré de T1 T3 lier, unir qc à ; détacher qc de T4 figuré de T3
U1 lier/unir qn à ; détacher qn de U2 figuré de U1 U3 lier/unir qc à ; détacher qc de U4 figuré de U3	X1 auxiliaires, semi-aux., modaux X2 impersonnels X3 existence X4 inchoatifs, résultatifs	

Figure 3.1 – Les classes sémantico-syntaxiques, les sous-classes et les sous-types syntaxiques de la classe générique X

X: Verbes auxiliaires et semi-auxiliaires, impersonnels, *exister*, inchoatifs ou résultatifs (212 entrées, 4 classes)

X1: Verbes auxiliaires, semi-auxiliaires, modaux (114 entrées, 1 sous-classe)

X1a: *avoir, être, faire, devenir, rester, sembler, rendre, aller, pouvoir, devoir, venir de, échouer à, se mettre à, ...* (114 entrées, 5 sous-types)

X1a.1: *avoir, être, faire* (3 entrées)

X1a.2: *être, devenir, rester, sembler, etc.* (15 entrées)

X1a.3: *rendre* (1 entrée)

X1a.4: *aller, pouvoir, devoir, venir de, etc.* (67 entrées)

X1a.5: *réussir à, échouer à, commencer à, se mettre à, etc.* (28 entrées)

X2: Verbes impersonnels (43 entrées, 1 sous-classe)

X2a: *il faut, il advient, il se peut avec complétive sujet* (43 entrées)

X3: Verbes d'existence (13 entrées, 1 sous-classe)

X3a: *être réel, exister qp, avoir valeur pour qp* (13 entrées, 3 sous-types)

X3a.1: *être réel* (6 entrées)

X3a.2: *être présent qp* (3 entrées)

X3a.3: *avoir telle valeur pour* (4 entrées)

X4: Verbes inchoatifs et résultatifs (42 entrées, 1 sous-classe)

X4a: *commencer ou finir* (42 entrées, 3 sous-types)

X4a.1: *avoir un début ou une fin* (18 entrées)

X4a.2: *avoir un début ou une fin par qc* (20 entrées)

X4a.3: *commencer ou finir qc* (4 entrées)

L'inconvénient de cette classification basée sur les schèmes syntaxiques est qu'elle n'offre pas de définition basée sur des traits sémantiques pour les classes génériques, classes sémantico-syntaxiques et sous-classes; ce n'est qu'en observant les entrées de chacune qu'on peut avoir une bonne idée de ce qu'elle englobe.

3.2.5 Structure transitive

Il existe différentes constructions syntaxiques pour un verbe, correspondant à diverses structures transitives. Ainsi, un verbe peut être intransitif (107), transitif direct (108), transitif indirect (109), ditransitif (110) (111), ou attributif (112) (Riegel *et al.*, 2009).

- (107) Le chien du voisin aboie sans cesse.
- (108) Mon neveu a cuisiné ce repas.
- (109) Uyen rêve à ses prochaines vacances.
- (110) Sophie va poser la question à sa superviseure.
- (111) Édouard parle souvent de ses trouvailles à ses collègues géologues.
- (112) Ce chat est le roi de la maison.

Les verbes intransitifs, comme à l'exemple (107), ne prennent pas de complément direct ou indirect, mais peuvent être accompagnés d'un modificateur, puisqu'ils ne sont pas pris en considération dans la transitivité du verbe. Lorsqu'un verbe a un complément direct ou indirect, il est transitif. Lorsqu'il n'y a qu'un complément direct, c'est-à-dire un syntagme nominal complément du verbe, il s'agit d'un verbe transitif direct, comme en (108). Lorsque le complément a la forme d'un syntagme prépositionnel, il s'agit plutôt d'un complément indirect, comme en (109). Lorsqu'un verbe a deux compléments, il est ditransitif. Le plus souvent, ces verbes ont un complément direct et un complément indirect, tel qu'en (110), mais deux compléments indirects sont aussi possibles, tel qu'en (111). Finalement, lorsqu'un des compléments du verbe dénote une propriété du référent du sujet ou du complément direct, ce verbe est attributif, comme en (112).

3.2.6 Forme du sujet

Le syntagme remplissant la fonction de sujet du verbe peut prendre plusieurs formes. Dans le cadre de ce mémoire, seules les formes que l'on retrouve dans le corpus Groupe LSQ 98 seront considérées, soit des syntagmes nominaux (113) et des pronoms (114) (115), ainsi que les absences de sujet (116) (117).

- (113) Les médias raffolent de ce genre d'histoire.
- (114) Vous devriez plutôt prendre l'autre chemin.
- (115) Les stagiaires qui mangent à la cafétéria proviennent du collège Rosemont.
- (116) Soraya souhaite terminer son projet la semaine prochaine.
- (117) Les livres s'empilent et accumulent de la poussière.

Les sujets pronominaux peuvent être des pronoms personnels (114) ou des pronoms relatifs (115). Les absences de sujet regroupent les occurrences de verbes n'ayant pas de sujet explicite, c'est-à-dire les verbes à l'infinitif (116), à l'impératif ou au participe présent; ainsi que les occurrences d'ellipse du sujet dans une coordination (117).

3.2.7 Traits sémantiques du sujet

Le syntagme nominal sujet transmet des traits morphologiques de personne et de nombre au verbe, mais ces traits peuvent être différents des traits sémantiques du référent du sujet. Prenons les phrases (118) et (119) :

- (118) La famille Beaulieu se rassemble tous les deux mois.
- (119) Toi et moi, on ira au cinéma demain.

À la phrase (118), le sujet du verbe est *la famille Beaulieu*, dont le noyau (*famille*) est un nom singulier. Or, ce syntagme réfère à l'ensemble des membres de cette famille, donc une entité plurielle – comme requis par le verbe *rassembler*, dont le sujet ne peut avoir un référent singulier. Ainsi, le trait morphologique de nombre du syntagme *la famille Beaulieu* ne correspond pas à son trait sémantique de nombre. Dans la phrase (119), le sujet *on* transmet les traits de troisième personne singulier. Par contre, il partage le référent de *toi et moi*, c'est-à-dire un référent de

première personne pluriel. Encore une fois, les traits sémantiques du pronom *on* ne concordent pas avec ses traits morphologiques.

Un autre trait sémantique du sujet est le trait d'animéité, à savoir si le référent est animé (120) ou inanimé (121). Le trait d'animéité peut changer l'interprétation d'un verbe, et certains verbes requièrent un sujet animé (122), voire humain (123).

(120) Des militaires ont envahi le nord du pays.

(121) L'odeur de café a envahi mes narines.

(122) *La roche/l'écureuil dort.

(123) *Le lézard/l'étudiant révise ses notes.

Un autre trait sémantique du sujet est le rôle thématique. L'interprétation d'un prédicat se fait grâce au schéma actanciel du verbe, qui détermine le nombre d'arguments et leur type sémantique. Certains verbes n'ont aucun argument, il s'agit de verbes impersonnels tels que *neiger* ou *falloir* et leur sujet, *il* impersonnel, n'a pas de rôle thématique. La majorité des verbes a au moins un argument, soit leur sujet, qui recevra un rôle thématique. Les définitions de rôles thématiques retenues pour ce mémoire, basées sur celles de Tellier (2016) et de Riegel *et al.* (2009), sont les suivantes :

(124) AGENT – Référent animé, volontaire, dont l'intervention donne lieu au procès.
Rosalie coupe des légumes.

(125) INSTRUMENT – Référent inanimé, pouvant être employé ou contrôlé par un actant volontaire, qui est l'origine du procès; peut être une cause externe.
Le vent a fermé la porte.

(126) PATIENT – Référent animé sur lequel le procès s'exerce; correspond au point d'aboutissement du procès.
Anton a été blessé pendant sa pratique de soccer.

(127) THÈME – Référent inanimé sur lequel le procès s'exerce, correspond au point d'aboutissement du procès.

La fenêtre s'est ouverte facilement.

(128) EXPÉRIENCEUR – Référent animé vivant une émotion ou un état physique.

Mireille s'est fâchée contre son père.

(129) SIÈGE – Référent inanimé où se manifeste un état physique.

Le téléphone vibre à chaque notification.

Les autres rôles thématiques (notamment *but*, *source* et *lieu*) ne pouvant pas être attribués à un syntagme sujet, ils ne sont pas présentés ici.

3.3 Questions spécifiques de recherche

Les notions présentées dans les sections précédentes ont permis de poser les bases théoriques qui me permettront d'atteindre l'objectif général de ce mémoire, soit la description de l'emploi du pronom préverbal SE dans le corpus Groupe LSQ 98.

Il est déjà établi dans la littérature sur le français sourd que l'utilisation des verbes pronominaux génère des écarts à la norme qu'on ne retrouve pas en français langue maternelle, soit des ajouts ou des omissions du pronom SE (voir 2.4). Le cadre théorique a précisé les particularités de l'apprentissage du français pour les Sourd·e·s et les formes typiques sourdes que l'on retrouve en français sourd, dont l'emploi des pronoms personnel; en plus de présenter une variété de structures de phrases possibles en français selon des variables morphosyntaxiques, sémantiques et lexicales. Il serait intéressant de préciser en quoi consistent les écarts à la norme dans l'emploi des verbes pronominaux dans les textes de Sourd·e·s, au-delà des ajouts et des omissions du pronom. Cela nous mène à une première question spécifique de recherche :

1. Quels sont les contextes morphosyntaxiques ou sémantiques dans lesquels on observe une différence entre la norme du français écrit quant à l'utilisation du pronom SE et son emploi dans le corpus Groupe LSQ 98 ?

Le cadre théorique a aussi mis en lumière le fait qu'il existe certaines règles en français sourd qui diffèrent de celles du français standard. Néanmoins, bien que la majorité des formes typiques sourdes présentent une régularité d'une personne sourde à une autre, il existe certains cas pour lesquels les écarts à la norme sont idiosyncratiques. De cela découle une deuxième question spécifique de recherche :

2. Est-ce possible d'observer une « norme sourde » dans l'emploi du SE dans le corpus Groupe LSQ 98? Si oui, quels sont les facteurs qui balisent la présence ou l'absence du SE ?

CHAPITRE 4 :

MÉTHODE

Le cadre théorique a présenté les éléments importants à considérer pour atteindre l'objectif général de ce mémoire, soit la description de l'emploi des verbes pronominaux dans le corpus Groupe LSQ 98, qui a été identifié comme étant porteur de caractéristiques du français sourd. Il a été possible de spécifier ce sur quoi portera la description grâce à la formulation de deux questions spécifiques de recherche :

QR1 : Quels sont les contextes morphosyntaxiques ou sémantiques dans lesquels on observe une différence entre la norme du français écrit quant à l'utilisation du pronom SE et son emploi dans le corpus Groupe LSQ 98 ?

QR2 : Est-ce possible d'observer une « norme sourde » dans l'emploi du SE dans le corpus Groupe LSQ 98? Si oui, quels sont les facteurs qui balisent la présence ou l'absence du SE ?

Le présent chapitre décrit la méthode utilisée pour répondre aux questions spécifiques de recherche. La première section consiste en une présentation du corpus Groupe LSQ 98 et des textes qui le composent (4.1). On y retrouve des données socioculturelles sur les participant·e·s (4.1.1), les procédures qui ont été suivies lors de la création du corpus (4.1.2), ainsi qu'un aperçu sommaire des textes du corpus sous la forme d'une description quantitative et de quelques exemples d'emplois de verbes pronominaux (4.1.3).

La deuxième section de ce chapitre indique la démarche d'identification des occurrences verbales retenues pour l'analyse (4.2). Elle précise les catégories d'occurrences (4.2.1), les motivations derrière les choix méthodologiques pour l'identification des occurrences (4.2.2), le protocole d'analyse grammaticale des phrases du corpus (4.2.3), le protocole d'identification des occurrences (4.2.4) et finalement le protocole pour le jugement de la correspondance ou non à la norme (4.2.5).

La section suivante présente la grille d'analyse descriptive du corpus (4.3). Les colonnes de cette grille correspondent à des variables, regroupées en trois catégories : les variables d'identification (4.3.1), les variables d'observation (4.3.2) et les variables contextuelles (4.3.3).

La quatrième section présente les tests statistiques effectués (4.4) et fournit une description ainsi qu'un guide d'interprétation pour le test du chi-carré (4.4.1), l'analyse factorielle (4.4.2), l'analyse des correspondances multiples (4.4.3) et la description des modalités (4.4.4).

4.1 Présentation du corpus Groupe LSQ 98 et de l'ensemble de données

Le corpus Groupe LSQ 98 a été constitué en 1998 par le Groupe de recherche sur la LSQ et le bilinguisme sourd de l'Université du Québec à Montréal (UQAM). Ce corpus, qui rassemble 96 textes narratifs, a déjà été utilisé pour la description de plusieurs constructions typiques du français sourd. Bien que ce corpus contienne une grande fréquence d'utilisation de verbes pronominaux, il n'avait pas encore été étudié en fonction de l'utilisation spécifique du SE en français sourd.

4.1.1 Participant·e·s

Les textes du corpus Groupe LSQ 98 proviennent de seize participant·e·s, soit six hommes et dix femmes qui ont une surdité profonde ou sévère. L'âge des participant·e·s se situe entre 16 et 65 ans (moyenne de 30 ans, médiane de 20 ans). Plus précisément, douze personnes ont entre 16 et 28 ans, tandis que quatre personnes ont entre 56 et 65 ans. Onze participant·e·s ont une surdité de naissance, trois sont devenu·e·s sourd·e·s entre 6 et 18 mois et les deux autres sont devenues sourdes à 4 et 6 ans. Le niveau de scolarité de la majorité des participant·e·s (onze, plus précisément) est du secondaire; il est du primaire chez deux participant·e·s, collégial chez une participante et universitaire chez deux participantes. Aucun·e participant·e n'a de parent sourd, mais cinq participant·e·s ont une ou des personnes sourdes parmi les membres de leur famille proche.

Tableau 4.1 – Description des participant·e·s

Participant·e	Genre	Âge	Âge d'apparition de la surdité	Niveau de scolarité	Frère ou sœur sourd·e
<i>P00</i>	<i>F</i>	62	<i>6 ans</i>	<i>secondaire</i>	<i>non</i>
P01	H	65	naissance	primaire	oui
P02	H	18	6 mois	secondaire	non
P03	F	17	naissance	secondaire	oui
P04	H	22	naissance	secondaire	oui
P05	F	60	18 mois	primaire	non
P06	F	28	naissance	collégial	non
P07	F	16	naissance	secondaire	non
P08	F	17	11 mois	secondaire	non
P09	F	26	naissance	universitaire	non
P10	H	19	naissance	secondaire	oui
P11	F	56	naissance	secondaire	oui
<i>P12</i>	<i>F</i>	<i>21</i>	<i>4 ans</i>	<i>universitaire</i>	<i>non</i>
P13	H	18	naissance	secondaire	non
P14	F	17	naissance	secondaire	non
P15	H	16	naissance	secondaire	non

Le recrutement s'est déroulé à l'Institut Raymond-Dewar – un centre de réadaptation spécialisé en surdité – et à l'école secondaire Lucien-Pagé, qui offre un programme pour les élèves sourd·e·s. Les personnes désirant participer devaient répondre aux critères suivants : présenter une surdité profonde ou sévère décelée avant l'entrée à l'école; être né·e au Québec ou y être arrivé·e très jeune; utiliser la LSQ comme moyen de communication quotidien et vouloir améliorer sa performance en français. Les textes recueillis ont ensuite fait l'objet d'un tri basé sur des critères

linguistiques additionnels. Plus précisément, les participant·e·s devaient avoir acquis la LSQ avant l'âge de cinq ans, donc avant la scolarisation, et ensuite le français écrit à l'école, donc à partir de 6 ans. Ainsi, les textes de P00 ont été retirés, étant donné son acquisition de la LSQ à un âge trop tardif. Par la suite, les textes de P12 ont aussi été retirés, étant donné qu'ils ne comportent aucun écart à la norme dans l'emploi des verbes pronominaux. Après ces retraits, l'âge le plus tardif d'apparition de la surdité est de 18 mois.

Considérant la nature descriptive de cette étude, je considère cet échantillon représentatif de la population sourde bilingue français-LSQ en regard de son hétérogénéité sociodémographique, malgré que certaines variables concernant les données biographiques des participant·e·s aient des valeurs sous-représentées. En effet, dans ses réflexions sur les complications méthodologiques dans les études traitant du bilinguisme, Bialystok (2001) indique que le développement cognitif et linguistique varie énormément d'un individu bilingue à l'autre selon l'influence d'une « constellation de circonstances sociales, économiques et politiques » (p. 7, traduction libre), ce qui crée des groupes très variés pour lesquels on ne peut pas réellement constituer de groupes monolingues comparables. Bialystok (2001) souligne aussi qu'il existe différentes raisons d'être bilingue et que la maîtrise de chaque langue sera conséquente à un ensemble de besoin et de circonstances. Grosjean (1998), dans son exposition des problèmes méthodologiques récurrents dans les recherches sur les populations bilingues, fait ressortir le besoin de considérer d'autres informations sur les participant·e·s, en plus des données biographiques, telles que l'historique linguistique, la stabilité langagière, les fonctions des langues, etc. Ainsi, selon Grosjean (1998), une description adéquate de participant·e·s bilingues devrait inclure, entre autres, l'âge et la manière d'acquisition de chaque langue, le niveau d'activité récent des différentes compétences linguistiques de chaque langue, les contextes liés à l'emploi de chaque langue, la quantité de temps passé avec d'autres bilingues et des monolingues et le niveau de maîtrise de chaque langue.

D'ailleurs, il importe de rappeler que le présent mémoire adopte une approche exploratoire et inductive, et que les méthodes quantitatives employées ne servent pas à quantifier les écarts à la norme, mais plutôt à en dresser un portrait descriptif. En d'autres mots, même s'il est probable que le texte d'une personne ayant quitté l'école il y a plusieurs années contienne plus d'écarts à la norme que celui rédigé par une personne suivant présentement des cours de grammaire du français, ce n'est pas une variation pertinente à cette étude, dont le but est de décrire la nature de cet écart à

la norme. Or, tel que mentionné dans le cadre théorique (voir 3.1.2), la nature et la fréquence des écarts à la norme dans les textes écrits par des Sourd·e·s varient peu d'un individu à l'autre, peu importe son âge, sa scolarité et ses moyens de communication. Ainsi, le contrôle des variables de l'âge d'acquisition de la langue cible (le français) et de la langue de référence (la LSQ) est suffisant pour valider la représentativité de l'échantillon.

4.1.2 Procédures de création du corpus

Le corpus Groupe LSQ 98 a été constitué dans le cadre d'une étude portant sur l'efficacité de modes d'intervention sur l'amélioration du français chez des personnes sourdes. Pour le volet écriture, les participant·e·s rencontraient à deux reprises un·e membre de l'équipe de recherche et avaient à travailler de manière autonome entre ces rencontres. Leur tâche consistait à produire six textes (un lors de la première rencontre, quatre entre les rencontres et un lors de la dernière rencontre). Les participant·e·s devaient également appliquer des techniques d'amélioration du français écrit, ainsi que tenir un journal de bord.

Pour la rédaction de textes, il a été demandé aux participant·e·s de produire six courts textes à partir de six amorces en consultant, au besoin, le Multidictionnaire et le Bescherelle. Ces amorces, qui ont été présentées sous forme de vidéo en LSQ, sont intitulées « le métro », « l'invitation », « le voyage », « la maison hantée », « le volcan » et « le sport ». Trois de ces amorces vidéo présentaient une histoire complète, une avait un contenu partiel (début d'histoire, sans fin), et les deux autres ne présentaient qu'un thème sans contenu narratif. Dans tous les cas, les participant·e·s ont reçu la consigne d'écrire un texte d'environ une page narrant soit une histoire (vraie ou inventée) en lien avec l'amorce vidéo, soit l'histoire présentée dans la vidéo. Les textes produits ont été anonymisés et transcrits à l'ordinateur. Pour ce mémoire, les participant·e·s ont reçu un nouveau code alphanumérique suivant un ordre aléatoire.

4.1.3 Données du corpus

La longueur moyenne des textes est de 221 mots (écart-type de 85.9), le plus long étant de 563 mots et le plus court, de 65 mots.

Tableau 4.2 – Nombre de mots par participant·e et par texte

	t1	t2	t3	t4	t5	t6	Moyenne
P01	250	168	101	115	206	124	160.7
P02	355	348	232	315	266	316	305.3
P03	354	243	150	151	300	207	234.2
P04	184	119	89	133	87	136	124.7
P05	187	378	145	288	260	257	252.5
P06	375	232	175	200	184	227	232.2
P07	217	221	155	263	186	208	208.3
P08	88	139	152	175	135	200	148.2
P09	272	212	214	212	191	232	222.2
P10	270	270	267	314	331	331	297.2
P11	214	231	143	150	197	177	185.3
P13	563	297	257	173	65	452	301.2
P14	318	294	124	154	105	209	200.7
P15	225	217	188	280	192	227	221.5
Moyenne	276.6	240.6	170.9	208.8	193.2	235.9	221.0

Un total de 3195 verbes (à un temps simple ou composé) a été recensé. Les textes contiennent en moyenne 38 verbes chacun (écart-type de 16.5), le texte le plus long et le plus court contenant respectivement 96 et 9 verbes.

Tableau 4.3 – Nombre de verbes par participant·e et par texte

	t1	t2	t3	t4	t5	t6	Somme	Moyenne
P01	42	25	17	22	31	14	151	25.2
P02	66	60	30	63	27	54	300	50.0
P03	81	47	32	33	58	39	290	48.3
P04	31	23	16	30	17	23	140	23.3
P05	31	55	20	45	21	35	207	34.5
P06	71	36	26	42	31	29	235	39.2
P07	39	33	26	51	31	43	223	37.2
P08	14	28	25	37	24	42	170	28.3
P09	47	35	26	35	27	39	209	34.8
P10	49	52	37	68	55	53	314	52.3
P11	41	44	19	22	29	29	184	30.7
P13	96	54	43	27	9	74	303	50.5
P14	67	54	20	28	21	41	231	38.5
P15	40	43	38	52	26	39	238	39.7
Somme	715	589	375	555	407	554	3195	
Moyenne	51.1	42.1	26.8	39.6	29.1	39.6		38.0

De ces verbes, 341 ont été retenus en tant qu'occurrence à l'étude (en suivant le protocole d'identification d'occurrences présenté en 4.2), ce qui résulte en une moyenne de 4.1 occurrences par texte (écart-type de 2.7). Quatre textes ne comportent aucune occurrence, tandis qu'on retrouve au maximum 13 occurrences dans un même texte. Dans deux autres textes, le pourcentage de verbes étant retenus comme occurrence est de plus de 30%.

Tableau 4.4 – Nombre d’occurrences par participant·e et par texte

	t1	t2	t3	t4	t5	t6	Somme	Moyenne
P01	5	0	2	1	1	1	10	1.7
P02	5	8	2	3	2	4	24	4.0
P03	9	8	2	2	13	4	38	6.3
P04	5	6	5	3	3	4	26	4.3
P05	4	10	5	11	4	10	44	7.3
P06	6	3	8	4	4	3	28	4.7
P07	1	4	3	3	4	4	19	3.2
P08	1	4	3	2	1	6	17	2.8
P09	5	6	3	5	3	4	26	4.3
P10	5	3	3	7	4	5	27	4.5
P11	1	3	3	2	0	3	12	2.0
P13	3	5	8	6	0	11	33	5.5
P14	9	4	0	3	1	2	19	3.2
P15	4	4	2	3	1	4	18	3.0
Somme	63	68	49	55	41	65	341	
Moyenne	4.5	4.9	3.5	3.9	2.9	4.6		4.1

L’entièreté du corpus consiste en 18 564 mots, 3 195 verbes (pouvant correspondre à plusieurs mots) et 341 occurrences retenues pour l’étude (10.7% des verbes). Des occurrences retenues, 40.8% représentent des écarts à la norme. Plus précisément, 53 sont des ajouts du SE (15.5%) et 86 sont des omissions du SE (25.2%); tandis que 71 sont des absences normées (20.8%) et 131 sont des présences normées (38.4%). Ainsi, lorsqu’un SE est produit, il est de trop dans 28.8% des cas.

Voici quelques exemples d'occurrences retenues pour l'analyse :

- (130) Tout à coup, il se leve rapidement et il prend mon sac d'école. (P14t1)
- (131) Je me sauvée courir en pente à bas. (*'Je me sauve en courant en bas de la pente'*, P05t6)
- (132) Les gens se sauvent autre placer. (*'Les gens se sauvent à une autre place'*, P07t5)
- (133) Avant de se rendre au bureau de postes, j'ai pris le métro. (*'Avant de me rendre'*, P01t1)
- (134) Toup à coup la personne s'approchait près de lui et il avait quelque chose dans le main. (P15t4)
- (135) [...] et je veux commander un hot-dog et une salade vert. (P10t2)
- (136) Nous parlons le sujet de la future vie. (*'Nous parlons au sujet de nos futurs'*, P08t2)
- (137) Donc, la porte du métro commence à refermer. (P13t1)
- (138) Je voulais arrêter pour attendre mon amie Line. (PP03t6)
- (139) La porte s'est ouverté par la vieille madame. (P09t4)
- (140) Il arrive chez lui, son coeur se bat très fort et son visage est comme un drap. (P14t4)
- (141) je me trouvais un peu difficiel sa langue d'anglais. (*'Je trouvais un peu difficile (de comprendre) sa langue, l'anglais'*, P04t3)
- (142) Je me suis parti à courir vers le dépanneur dans le métro et ils n'ont raté. (P15t1)
- (143) Un vieux homme se réponde la porte et il dit: « Bonjour qui êtes-vous ? ». (P07t4)
- (144) On fait bronzer sur le plage ensoleillé. (P11t3)
- (145) Je cris, pleure et décourage. (P14t1)
- (146) Quelques minutes plus tard, je reveillais puis, je regardais autout le mur, le planche etc... (P10t6)
- (147) aisni que des arbres avaient endommagé. (P09t5)
- (148) On réunit à chaque deux semaines. (P13t3)

Les exemples (130) à (138) sont des occurrences qui respectent la norme; (130) à (134) représentent des présences normées du SE, tandis que (135) à (138) sont des absences normées facultatives (qui resteraient normées même si l'on ajoutait un SE). Les exemples (139) à (148) présentent des écarts à la norme; (139) à (143) sont des ajouts du SE et (144) à (148) sont des omissions du SE.

4.2 Identification des occurrences

Étant donnée la diversité des types d'occurrences et la présence de formes typiques sourdes, des protocoles ont été mis en place afin d'assurer une analyse consistante des textes du corpus. Cette section présente les différents types d'occurrences retenues pour l'étude (4.2.1), la motivation derrière certains choix méthodologique (4.2.2), ainsi que le protocole d'analyse grammaticale (4.2.3), celui d'identification des occurrences (4.2.4) et celui pour le jugement de la norme (4.2.5).

4.2.1 Catégories d'occurrences

Comme cette étude porte sur l'emploi des verbes pronominaux dans le corpus, ainsi que sur la différence entre cet emploi et celui prescrit par la norme, les occurrences verbales peuvent être décrites selon le croisement de deux variables : a) présence du SE dans l'occurrence (*le prédicat produit contient-il un verbe pronominal?*) et b) présence du SE selon la norme (*considérant le contexte, est-ce que la norme exige un verbe pronominal?*).

Tableau 4.5 – Catégories d'occurrences

	Norme : <u>SE</u> obligatoire	Norme : <u>SE</u> facultatif	Norme : <u>SE</u> interdit
Occurrence : présence du <u>SE</u>	présence normée <i>Je me reposais</i> (P08t2)	présence normée facultative <i>Je me préparais du (de la) soupe</i> (P04t2)	ajout (présence non normée) <i>Son cœur se bat très fort (...)</i> (P14t4)
Occurrence : absence du <u>SE</u>	omission (absence non normée) <i>Nous couchons sur le sable (...)</i> (P10t3)	absence normée facultative <i>Nous parlons (de) notre enfance (...)</i> (P14t1)	absence normée obligatoire <i>Il avait cherché partout (...)</i> P01t4

Les occurrences verbales retenues sont celles présentant une présence normée (facultative ou obligatoire), un ajout, une omission ou une absence normée facultative; les absences normées obligatoires n'ont pas été retenues pour des raisons expliquées en 4.2.2.

4.2.2 Motivation des critères d'identification des occurrences

L'approche adoptée dans ce mémoire est exploratoire, descriptive et inductive plutôt que hypothético-déductive et normative. D'ailleurs, l'emploi des verbes pronominaux relève souvent d'un choix lexical : il serait ardu d'affirmer, par exemple, qu'une phrase agrammaticale correspond nécessairement à une omission du pronom SE, surtout lorsqu'on prend en considération les formes typiques sourdes décrites dans la littérature (voir 3.1.3). Ainsi, l'agrammaticalité de la phrase (149) pourrait être expliquée soit par l'omission du SE dans l'emploi d'un verbe pronominal passif (150), soit par l'omission de l'auxiliaire *être* dans un verbe à la forme passive (151), ou alors par la chute du sujet et à un transfert de la structure objet-verbe de la LSQ (152).

- (149) *Les feuilles ramasser à la pelle.
- (150) Les feuilles se ramassent à la pelle.
- (151) Les feuilles sont ramassées à la pelle.
- (152) Les feuilles, Marie les ramasse à la pelle.

D'ailleurs, des études précédentes ont déjà établi que l'emploi des verbes pronominaux dans les textes de Sourd·e·s diffère de celui d'entendant·e·s francophones (voir 2.4). Plutôt que de partir d'écarts à la norme et d'observer si on peut les classer comme un ajout ou une omission du SE, il semble davantage intéressant de partir du sens de l'énoncé, puis voir comment les participant·e·s ont choisi d'exprimer ce sens. Ainsi, les occurrences d'écarts à la norme représentent des phrases non normées pour lesquelles l'ajout ou le retrait d'un SE résulterait en une forme normée, sans que cela soit la seule ou la meilleure correction possible.

Lors de l'élaboration des critères d'identification des occurrences, il a fallu aussi considérer les tests statistiques à effectuer par la suite. Ainsi, même s'il aurait été intéressant de retenir l'entièreté des verbes produits dans le corpus, la distribution naturelle du français est telle que la majorité des verbes produits ne sont pas pronominaux. L'inclusion de tous les verbes aurait créé une

surreprésentation très importante des absences normées du SE dans l'ensemble des données, rendant les analyses statistiques subséquentes moins intéressantes.

Par contre, retirer l'entièreté des occurrences d'absences normées du SE de l'ensemble des données aurait entraîné la perte d'informations intéressantes sur les contextes où l'on retrouve des absences normées facultatives, c'est-à-dire des occurrences où les participant·e·s avaient le choix d'employer un verbe dans sa forme pronominale tout en respectant la norme, mais ont plutôt utilisé le même verbe dans sa forme de base, sans SE. C'est pour cela que les occurrences d'absence normée facultative du SE ont été retenues, mais pas celles d'absence normée obligatoire (pour le reste du mémoire, le terme « absence normée » fera référence aux absences normées facultatives, à moins d'une mention contraire).

4.2.3 Protocole d'analyse grammaticale

Afin d'assurer une analyse grammaticale consistante tout au long du corpus, un protocole a été établi. De manière générale, l'analyse grammaticale était basée en priorité sur le sens de l'énoncé, plutôt que sur la ponctuation et la syntaxe. Par exemple, l'analyse des prédicats de l'énoncé en (153) n'a pas été faite en priorisant les règles syntaxiques d'énumération – ce qui aurait résulté en l'analyse en (154) –, mais plutôt en priorisant le sens de l'énoncé afin d'identifier des éléments sous-entendus – résultant en l'analyse en (155) :

(153) J'ai visité beaucoup de la musé, le tour de la ville et la culture est très diffirent, je l'intéresse beaucoup. (P05t3)

(154) [P1 J'ai visité [CD visiter beaucoup de la musé], [CD visiter le tour de la ville] et [CD visiter la culture]] / et [P2 la culture_i est [Att. Sujet très diffirent]], [P3 je [CD intéresser l_i?'] intéresse beaucoup].

(155) [P1 J'ai visité [CD visiter beaucoup de la musé]], [P2 (j'ai fait) [CD faire le tour de la ville]] et [P3 la culture_i est [Att. Sujet très diffirent]], [P3 je (m') [CI s'intéresser y_i] intéresse beaucoup].

L'analyse grammaticale du corpus a aussi pris en considération l'incorporation de certaines expressions idiomatiques de la LSQ. Par exemple, on retrouve dans la phrase en (156) un transfert de l'expression TOUCHE FINI, qui signifie 'avoir déjà fait quelque chose'. La phrase a donc été

analysée en considérant *toucher* comme un auxiliaire apportant des informations temporelles et modales, c'est-à-dire que le prédicat qu'il accompagne a déjà été réalisé par le passé.

- (156) C'est une expérience enrichissante de toucher visiter à la maison hantée. (P11t4)
'C'est une expérience enrichissante d'avoir déjà visité une maison hantée.'

4.2.4 Protocole d'identification des occurrences

La première étape de l'identification des occurrences a été de souligner dans l'entièreté du corpus tous les verbes, ainsi que les noms et adjectifs employés en tant que verbes. Les verbes conjugués à un temps composé (auxiliaire *être* ou *avoir* + participe passé) ont été considérés comme un seul verbe, puisque l'auxiliaire se place entre le verbe lexical et le pronom SE (157). Au contraire, les auxiliaires modaux, temporels, aspectuels et actanciels (dont l'auxiliaire *ALLER* du futur périphrastique) ne se placent pas entre le verbe lexical et le pronom SE (158, 159); ils ont donc été considérés comme un verbe à part, à la manière des subordonnées infinitives compléments de verbe (160). Cela reflète la possibilité des auxiliaires d'être eux-mêmes des verbes pronominaux (161), ce qui peut donner lieu à des constructions *auxiliaire* + *verbe lexical* contenant deux verbes pronominaux (162).

- (157) Le couple se marie cet été → Le couple s'est marié l'été passé. (*se marier*)
(158) Le couple (*se) pourra se marier à l'église. (*pouvoir* + *se marier*)
(159) Le couple (*se) va se marier l'été prochain. (*aller* + *se marier*)
(160) Je voudrais me changer avant de sortir. (*vouloir* + *se changer*)
(161) Il s'est laissé bercer par la musique. (*se laisser* + *bercer*)
(162) Elle s'est mise à se méfier des promesses de son patron. (*se mettre à* + *se méfier*)

L'étape suivante a été l'identification du sujet. Si le sujet n'est pas facile à identifier (par exemple s'il y a plus d'un antécédent possible ou si le sujet n'est pas réalisé avant le verbe), la compatibilité sémantique entre le verbe et le sujet prime sur la compatibilité morphosyntaxique.

La troisième étape visait à déterminer si le verbe est précédé du pronom SE ou non, c'est-à-dire si l'occurrence contient un pronom préverbal coréférentiel avec le sujet préalablement identifié. Si

oui, l'occurrence était retenue pour l'analyse. Sinon, l'occurrence pouvait être retenue à condition que l'on puisse y ajouter un SE. Les critères d'identification précis pour les occurrences de présence ou d'absence du SE sont décrits dans les sous-sections suivantes.

a) Protocole d'identification des occurrences de présence du SE

Les occurrences de présence du SE ont été identifiées par la présence d'un pronom préverbal coréférentiel au sujet. Encore une fois, le jugement de cette coréférentialité se base principalement sur le sens, c'est-à-dire s'il est cohérent en contexte que ce pronom fasse référence au sujet. Cette coréférentialité ne peut être partielle, excluant de ce fait les sujets singuliers inclus dans les compléments pluriels et inversement, tel que *je nous achète des rideaux* et *nous m'achetons des rideaux*. Lorsque le pronom SE n'est pas la réalisation d'une référence syntaxique (dans le cas des SE lexicaux et passifs, voir 2.3), le jugement est basé sur la concordance des traits morphologiques.

Étant donné que les occurrences de présence du SE sont identifiées à partir des verbes, les productions de *se* ou *s'* ne sont pas retenues lorsqu'elles ne précèdent pas un verbe ou un mot employé en tant que verbe, tel que dans l'énoncé suivant :

- (163) La prochaine fois, elle sera tenir toujours le dans s'epaule pour un securite. (P05t1)
'Dorénavant, elle tiendra toujours son sac sur son épaule par mesure de sécurité'

Au contraire, lorsque ces pronoms sont produits avant un nom ou un adjectif employé en verbe, ils sont considérés comme une occurrence, tel que dans l'énoncé suivant :

- (164) [...] à cause les filles se bec partout sur la face de mon autre ami. (P07t2)
'Parce que les filles font des becs partout sur la face de mon autre ami'

Chaque occurrence de verbe précédé du pronom SE est retenue pour l'analyse des données.

b) Protocole d'identification des occurrences d'absence du SE

Lorsque le verbe n'est pas précédé de SE, il faut alors déterminer s'il pourrait – ou devrait – l'être : en d'autres mots, si c'est une occurrence d'absence du SE. Ces occurrences peuvent être soit des omissions du SE (165), soit des absences normées (166).

(165) J'ai endormi dans le voyageur (*l'autobus*). (P09t5)

(166) Quelques les seconds plus tard, le tremblement de terre arrêta. (P10t5)

Il importe de rappeler que, si une occurrence présentant un écart à la norme est considérée comme une omission, cela ne signifie pas pour autant que l'ajout d'un SE est la seule ou la meilleure correction possible. Ainsi, pour faire correspondre la phrase en (165) à la norme, on pourrait soit ajouter un SE (167), soit modifier l'auxiliaire (168); le contexte ne permet pas de trancher.

(167) Je me suis endormi dans l'autobus voyageur.

(168) J'étais endormi dans l'autobus voyageur.

De la même manière, l'énoncé en (169) est considéré comme une occurrence d'omission du SE puisque la formulation (170) est normée, même si d'autres modifications sont possibles pour respecter la norme (171).

(169) Je commençais un peu inquiète. (P04t2)

(170) Je commençais un peu à m'inquiéter.

(171) Je commençais à être un peu inquiète.

Ainsi, les occurrences d'omission du SE correspondent à une occurrence d'emploi non normé d'un verbe, pour lequel l'ajout du SE serait un moyen de respecter la norme.

Afin d'identifier les occurrences d'absence du SE, la phrase était modifiée pour y ajouter un SE et le résultat était comparé à des dictionnaires pour juger de la grammaticalité de la modification (voir 4.2.5). Lors de la modification de la phrase, la structure syntaxique des compléments du verbe a été respectée : il ne pouvait pas y avoir de changement, d'ajout ou de retrait de préposition. Ainsi,

le verbe en (172) n'a pas été retenu, car pour être normé, l'ajout d'un SE demanderait l'ajout d'une préposition (173), modifiant la structure transitive du verbe :

(172) [...] j'ai peur de frapper un mat ('*un arbre*'). (P03t6)

(173) [...] j'ai peur de me frapper contre un arbre.

La modification de la phrase pouvait toutefois apporter un changement dans l'auxiliaire de conjugaison (*la fête a fini* ~ *la fête s'est finie*) et dans le type de déterminant dans le cas d'une montée du trait de possession (*je brosse mes cheveux* ~ *je me brosse les cheveux*).

Afin d'être considéré une occurrence d'absence normée du SE, l'ajout du pronom ne devait pas résulter en une formulation d'un registre excessivement soutenu (tel que *il se meurt*), ni être un ajout explétif (tel que *je me fume une clope*); il pouvait néanmoins réaliser un complément sous-entendu facultatif, notamment celui ayant le rôle thématique bénéficiaire (*je dois (m') acheter un billet*).

4.2.5 Protocole de jugement de la norme

Dans ce mémoire, ce qui est désigné comme étant « la norme » correspond au français tel qu'il est employé par les Québécois·e·s francophones sans surdit ; il s'agit d'une norme descriptive plutôt que prescriptive.

Le jugement de la norme sert à déterminer si le verbe d'une occurrence donnée respecte la norme quant à la présence obligatoire, interdite ou facultative du SE. Pour les verbes occasionnellement pronominaux, le changement entre la forme verbale de base et la forme pronominale peut entra ner des diff rences dans la construction syntaxique du syntagme verbal, dans le sens lexical du verbe et dans les restrictions sur les traits s mantiques du sujet, en plus d'apporter des modifications d'aspect verbal. Ainsi, la norme r gissant la possibilit  d'employer un verbe dans sa forme pronominale d pend du contexte dans lequel se trouve le verbe.

La premi re  tape consiste en la d termination de la norme quant à la pr sence ou l'absence du SE avant le verbe de l'occurrence. Pour ce faire, il a fallu identifier le sens lexical du verbe selon le

contexte, en faisant abstraction de la présence ou de l'absence du SE et de potentielles erreurs d'orthographe ou confusions lexicales. Ensuite, trois dictionnaires numériques, soit Le Petit Robert (Dictionnaires Le Robert, 2023), Antidote (Druide informatique, 2022) et Usito (Cajolet-Laganière *et al.*, 2023), ont été consultés afin de vérifier s'il existe une acception répertoriée de ce verbe pour exprimer ce sens. Une attention particulière était portée aux nuances sémantiques (aspect verbal, restriction de traits sémantiques des arguments, etc.) associées aux acceptions. Une fois l'acceptation appropriée identifiée, il était possible de déterminer si, pour ce sens précis, le SE est obligatoire, interdit ou facultatif. La deuxième et dernière étape consiste en la comparaison entre la norme et l'occurrence dans le corpus.

L'application de ce protocole a posé certains défis, dont deux provenant des limitations de l'utilisation de dictionnaires : premièrement, les restrictions sémantiques ne sont pas toujours explicitées dans les ouvrages de référence (par exemple, entre *approcher* et *s'approcher*); deuxièmement, même les dictionnaires québécois ne décrivent pas l'entièreté des emplois courants mais non normés québécois (par exemple, *pratiquer* dans le sens de 's'exercer' est répertorié en tant qu'anglicisme critiqué, mais son emploi transitif – pourtant répandu – ne se retrouve pas dans les dictionnaires : *j'ai pratiqué mon oral toute la fin de semaine*). Lorsque la consultation des dictionnaires ne permettait pas de bien déterminer la norme, le verbe problématique était présenté à un jury composé de 9 personnes ayant le français québécois comme langue maternelle, provenant de 7 régions administratives différentes du Québec et ayant complété des études universitaires dans un domaine lié au français (linguistique, littérature ou communications). Une occurrence donnée était considérée normée si au moins deux juges la trouvaient grammaticale³³, même si elle ne correspondait pas à une acception de dictionnaire.

Un autre défi provenait de l'emploi non normé de verbes, tel que dans la phrase suivante :

- (174) On avait finalement trouver le moyen de recevoir le bonheur avec nos finissants.
'On avait finalement trouvé le moyen de se faire plaisir entre finissants' (P13t3)

³³ Les juges devaient déterminer si ils et elles pouvaient s'imaginer produire cet énoncé (ou alors imaginer une de leurs connaissances le faire), peu importe le contexte, sans que cela sonne maladroit.

Afin d'éviter un jugement de la norme basé sur une correction subjective, ces occurrences ont été considérées comme une acception non normée idiosyncratique du verbe. Ainsi, le verbe *recevoir* de la phrase en (174) n'est pas retenu pour l'étude, bien qu'on puisse le remplacer par un verbe pronominal pour que la phrase soit normée.

4.3 Présentation de la grille d'analyse du corpus

Afin de rendre compte de l'emploi des verbes pronominaux dans le corpus Groupe LSQ 98, une grille d'analyse descriptive a été créée. Cet outil méthodologique a permis de recenser les occurrences retenues en fonction de variables morphosyntaxiques et sémantiques, ainsi que de les classer selon les deux variables principales permettant d'observer l'emploi du SE, soit a) la présence ou l'absence du pronom et b) la correspondance ou non à la norme.

Une fois qu'une occurrence a été identifiée selon le protocole d'identification (voir 4.2), elle était retranscrite dans cette grille d'analyse; chaque rangée correspond à une occurrence. Ensuite, les colonnes étaient remplies selon le contenu des occurrences. Ces colonnes correspondent à des variables qui se divisent en trois catégories : les variables d'identification (4.3.1), les variables d'observation (4.3.2) et les variables contextuelles (4.3.3).

4.3.1 Variables d'identification

Les deux variables d'identification permettent d'identifier, pour chaque occurrence, le ou la participant·e l'ayant produit et le texte dans lequel on la retrouve. La première, [Particip], indique le code alphanumérique du ou de la participant·e, de < P01 > à < P15 >. La deuxième, [Texte], indique le numéro de texte de < T1 > à < T6 >, correspondant aux 6 amorces données aux participant·e·s.

Ces variables permettent de vérifier l'importance de la variation interindividus, ainsi que celle intra-individus qui provient de facteurs non-linguistiques tels que le type d'amorce, le sujet du texte ou le contexte de rédaction.

4.3.2 Variables d'observation de l'emploi du pronom SE

Les quatre variables d'observation servent à répondre aux questions spécifiques de recherche grâce à l'observation de leur distribution en fonction des variables contextuelles.

Tableau 4.6 – Combinaisons possibles des variables d'observation

Occurrence	[PresSe]	[NormePres]	[RspNorme]	[CatOcc]
<i>occurrence 1</i>	< PresO >	< PN_O >	< RN_o >	< PN >
<i>occurrence 2</i>	< PresO >	< PN_N >	< RN_n >	< Ajout >
...	< PresO >	< PN_Fac >	< RN_o >	< PNF >
	< PresN >	< PN_O >	< RN_n >	< Omis >
	< PresN >	< PN_Fac >	< RN_o >	< AN >

La première, [PresSe] (*présence du SE*), indique si l'occurrence en question présente un cas de présence du SE (< PresO >) ou d'absence du SE (< PresN >). L'observation de cette variable permet d'identifier les contextes favorisant la production des verbes pronominaux dans le corpus, ce que l'on peut comparer à la norme du français par la suite.

La deuxième variable, [NormePres], indique la norme quant à la présence du SE pour une occurrence donnée. Les valeurs possibles sont < PN_O > (*présence selon la norme = oui*) pour les occurrences dans lesquelles le SE est obligatoire selon la norme, < PN_N > (*présence selon la norme = non*) pour les SE interdits et < PN_Fac > (*présence selon la norme = facultatif*) pour les cas où la présence du SE est facultative (CD ou CI facultatif, différence minimale entre forme pronominale ou non), ainsi que pour les cas où il est impossible d'identifier la norme à appliquer (*je suis fâché*, impossible de trancher entre construction passive, verbe attributif et verbe *se fâcher* au passé composé avec omission du SE). L'observation de la distribution de [NormePres] à elle seule ne permet pas de décrire l'emploi du SE dans le corpus; elle offre plutôt des informations sur l'utilisation normée du SE. Néanmoins, sa mise en relation avec l'observation de la distribution des autres variables d'observation permet d'identifier l'effet de la distribution naturelle de la langue.

La troisième, [RspNorme] (*respect de la norme*), dépend des deux premières. En croisant la variable de présence ou absence du SE dans l'occurrence et la variable de la norme quant à la présence du SE, il est possible d'identifier si l'occurrence respecte la norme (< RN_o >) ou si elle présente un écart à la norme (< RN_n >). L'observation de la distribution de [RspNorme] permet de voir s'il existe des contextes dans lesquels la norme est mieux respectée, et d'autres où elle l'est moins.

Tableau 4.7 – Combinaisons représentées par la variable [RspNorme]

		[NormePres]		
		PrésNorme_Oui	PrésNorme_Non	PN_Facultatif
[PresSe]	Prés Oui	RN_oui	RN_non	RN_oui
	Prés Non	RN_non	RN_oui	RN_oui

La dernière variable d'observation, [CatOcc] (*catégorie d'occurrence*), est aussi un croisement des deux premières variables, mais avec une catégorisation plus précise. Les écarts à la norme sont divisés en < Ajout > et < Omis >, et les respects de la norme en < PN > (*présence normée*), < PNF > (*présence normée facultative*) et < AN > (*absence normée*).

Tableau 4.8 – Combinaisons représentées par la variable [CatOcc]

		[NormePres]		
		PrésNorme_Oui	PrésNorme_Non	PN_Facultatif
[PresSe]	Prés Oui	PN	Ajout	PNF
	Prés Non	Omis	_ ³⁴	AN

³⁴ Rappel : ces occurrences, trop nombreuses dans le corpus, ne sont pas retenues pour l'analyse.

4.3.3 Variables contextuelles

Les treize variables contextuelles décrivent les contextes dans lesquels se trouve l'occurrence. Les réponses aux questions de recherche sont formulées selon la distribution des variables d'observation en fonction des modalités des variables contextuelles. Toutes ces variables correspondent à des notions présentées dans le cadre théorique (voir 3.2).

La variable [vPro] indique le type de verbe pronominal selon le regroupement suivant : < AutoEP > désigne les verbes pronominaux autonomes ainsi que les verbes essentiellement pronominaux; < PasPro > les verbes jamais pronominaux et ceux ne pouvant pas être pronominaux dans le contexte de l'occurrence; < Neutre > les verbes pronominaux neutres; < Passif > les verbes pronominaux passifs; < RefCD > les réfléchis et réciproques ayant une fonction de complément direct et < RefCI > les réfléchis et réciproques ayant une fonction de complément indirect. Les regroupements de types de verbes pronominaux ainsi que la division des réfléchis selon la fonction du SE ont été décidés en fonction du nombre d'occurrences par modalité et de préanalyses statistiques visant à identifier les regroupements permettant le mieux de décrire la variation contenue dans le corpus.

La variable [Temps] indique si le verbe de l'occurrence est conjugué à un temps simple (< TS >) ou à un temps composé (< TC >).

La variable [VvInf] indique si l'occurrence se retrouve dans une construction verbe conjugué + verbe infinitif (< VinfO >) – c'est-à-dire en présence de semi-auxiliaire ou de phrase subordonnée infinitive – ou non (< VinfN >).

La variable [Aspect] indique le classement aspectuel de l'occurrence, c'est-à-dire si le procès verbal correspond à un achèvement (< aAchv >), à un accomplissement (< aAcmp >), à une activité (< aAct >) ou à un état (< aEtat >); à cela s'ajoute une autre modalité (< AspNA >) pour les verbes ne désignant pas un procès (tel que les semi-auxiliaires) et les occurrences dont la présence de formes typiques sourdes empêchent de procéder aux tests de classement.

La variable [LVF] indique le classement de l'acceptation du verbe selon Les verbes du français (Dubois et Dubois-Charlier, 1997). Les modalités prennent la forme <LVF_A>, où A est remplacé par la lettre assignée à la catégorie sémantico-syntaxique.

La variable [StEmp] indique la structure transitive telle qu'employée dans l'occurrence. L'emploi peut être attributif (< eAtt >), intransitif (< eInt >), transitif direct (< eTD >), transitif indirect (< eTI >) ou bitransitif (< eBiTr >).

La variable [StNorme] indique la structure transitive qui devrait être employée pour que l'occurrence soit normée. L'emploi normé peut être attributif (< nAtt >), intransitif (< nInt >), transitif direct (< nTD >), transitif indirect (< nTI >) ou bitransitif (< nBiTr >); il peut aussi ne pas y avoir d'emploi normé (< nND >), notamment dans le cas d'un nom employé en verbe ou d'un semi-auxiliaire.

La variable [StRsp] indique si la structure transitive employée dans l'occurrence est normée, c'est-à-dire si la modalité de [StEmp] correspond à celle de [StNorme]. La structure transitive employée peut être normée (< RSTo >, *respect de la structure transitive = oui*), ne pas l'être (< RSTn >), ou alors présenter un mauvais choix de la préposition dans le cas de compléments indirects (< RSTchx >).

La variable [Sujet] indique la forme que prend le sujet du verbe de l'occurrence. Il peut être un pronom personnel (< je >, < tu >, < il elle >, < on >, < nous >, < vous >, < ils elles >), un pronom relatif (< ProRel >), un syntagme nominal (< SN >) ou alors être absent – à cause d'une ellipse dans une coordination (< AbsCoord >), sous-entendu avant un verbe infinitif (< SsEnt >) ou encore une absence non normée (< AbsNn >).

Les variables [Pers] et [Nb] indiquent les traits de personne (< 1 >, < 2 > ou < 3 >) et de nombre (< sg > ou < pl >) du référent du sujet. En effet, les préanalyses statistiques laissent paraître que les traits du référent du sujet ont plus d'influence sur l'emploi du SE que les traits morphologiques. Ainsi, le SN *le groupe d'enfants* aurait la modalité < pl >.

La variable [RoleT] indique le rôle thématique du sujet, soit < agent >, expérienceur (< experi >), instrument (< instru >), patient (< pati >), siège (< sieg >) ou thème (< them >); à cela s'ajoute la modalité < RolND > pour les occurrences où le sujet est impersonnel et n'a donc pas de rôle thématique.

La variable [Anime] indique si le sujet est animé (< Ani >) ou inanimé (< Inani >).

4.4 Analyses statistiques

Trois tests statistiques ont été appliqués sur la grille d'analyse du corpus afin d'en fournir une description générale, une description axée sur les variables d'observation, ainsi qu'une description fine de la distribution des variables d'observation. Des tests complémentaires (chi-carré) ont été effectués pour vérifier l'effet des variables d'identification (4.4.1).

Étant donnée l'approche exploratoire et descriptive de cette étude, les données recueillies ont été modélisées grâce à des analyses factorielles, dont les principes de base sont présentés en (4.4.2). Le premier niveau d'analyse, soit la description générale du corpus, prend la forme d'une analyse des correspondances multiples (4.4.3). Ensuite, une analyse factorielle multiple (4.4.4) a permis d'orienter la description sur les variables décrivant l'utilisation des verbes pronominaux. Cette analyse est complétée par une analyse unidimensionnelle de description des modalités (4.4.5), permettant de répondre aux questions de recherche avec plus de précision.

4.4.1 Test du chi-carré

Le test de chi-carré permet de vérifier si deux variables catégorielles sont indépendantes selon la distribution des occurrences. Cette analyse non-paramétrique se base sur le tableau croisé des deux variables testées, dans lequel chaque cellule contient le nombre d'occurrences équivalent à cette combinaison de modalités. Le test de chi-deux calcule le résidu de chaque cellule (la différence entre la distribution réelle et la distribution attendue) et quantifie cet écart, puis un degré de signification est assigné à cette valeur (Yergeau et Poirier, 2021).

4.4.2 Analyse factorielle

L'analyse factorielle désigne plusieurs méthodes d'analyse exploratoire de données multidimensionnelles et descriptives. Ces méthodes permettent la visualisation efficace de données multivariées, prenant habituellement la forme de tableaux trop gros pour en avoir une vue d'ensemble. L'analyse factorielle permet de résumer l'information contenue dans les données en générant des dimensions – c'est-à-dire une combinaison linéaire des variables des données multivariées – et en identifiant celles représentant le maximum de variation des données. Le croisement des dimensions expliquant le plus de la variation du jeu de données permet de représenter ce dernier sous forme de graphique, et ce, en perdant le moins d'informations que possible. Ainsi, l'analyse factorielle facilite l'identification des points saillants cachés dans de gros tableaux, sans recours à une hypothèse à confirmer ou à infirmer. Plus précisément, on peut observer comment les individus se regroupent, quelles sont les relations entre les variables et quelles modalités sont associées. Comme le souligne un évaluateur, il est important de noter que les dimensions des méthodes d'analyse factorielle employées dans ce mémoire sont construites de manière orthogonale, donc que les phénomènes portés par un axe sont indépendants de ce qui est porté par les autres axes. Ainsi, il est possible de passer d'une multitude de variables ayant des modalités corrélées à des dimensions indépendantes, qui condensent l'information contenue dans les variables originales.

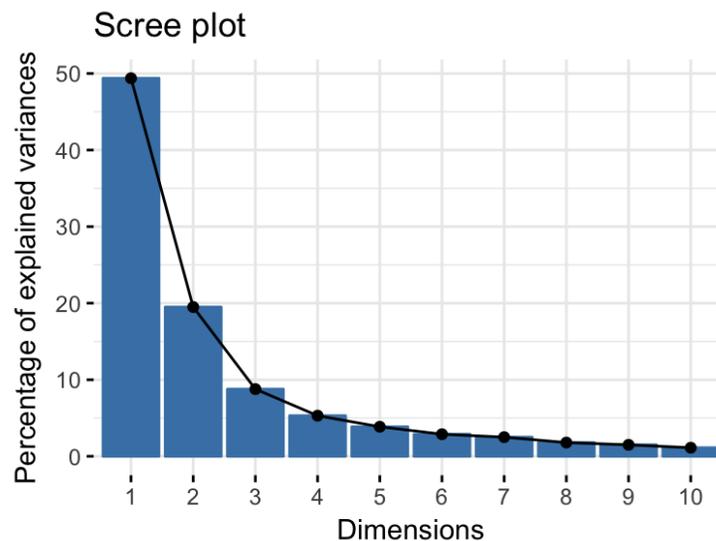
Les méthodes classiques d'analyse factorielle sont l'analyse de composantes principales (ACP) pour les variables continues, l'analyse factorielle des correspondances (AFC) pour les tableaux de contingence et l'analyse des correspondances multiples (ACM) pour les variables catégorielles. Parmi les méthodes avancées, on retrouve l'analyse factorielle multiple (AFM), permettant la mise en relation de groupes de variables quantitatives ou qualitatives. Ces analyses s'effectuent habituellement avec le logiciel R grâce aux packages FactoMineR (Lê *et al.*, 2008) et factoextra (Kassambara et Mundt, 2020).

4.4.3 Analyse des correspondances multiples

Étant donné que les variables de la grille d'analyse sont catégorielles, la variation générale du corpus est modélisée par une analyse des correspondances multiples (ACM). La présente section sert de guide d'interprétation des ACM.

Tel que mentionné précédemment, les analyses factorielles réduisent le nombre de variables en les fusionnant pour créer des dimensions linéaires qui expliquent le maximum de la variation des données; le moins de variables il y a, le plus de variation est contenue dans les premières dimensions (Kassambara, 2017a). Il est possible de visualiser le pourcentage de variation expliqué par chaque dimension grâce à un *scree plot*, tel qu'illustré à la figure 4.1³⁵.

Figure 4.1 – Exemple de *scree plot*

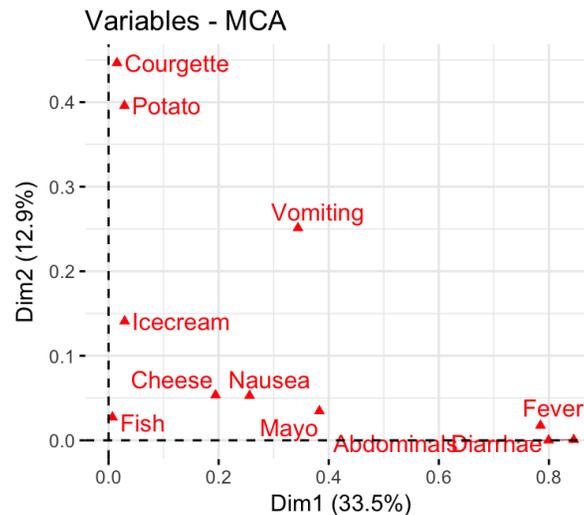


Tant qu'une dimension explique un plus grand pourcentage que celui d'une variable si la variation était aléatoire, elle peut être retenue pour l'étude. En d'autres mots, si un tableau comporte 20 variables, la proportion aléatoire de variation expliquée par chacune serait de 5% (100%/20). Ainsi, une dimension expliquant moins de 5% de la variation serait peu pertinente pour synthétiser l'information.

³⁵ La figure 4.1 est tirée de l'article sur les AFM du site web de STHDA (Kassambara, 2017b).

du corpus). La dimension 2, elle, est surtout définie par les variables consommation de courgettes et consommation de patates.

Figure 4.3 – Exemple de carré des liaisons



Si l'on retourne à l'observation du *biplot*, on remarque que la dimension 1 crée un axe linéaire entre l'absence de fièvre, de diarrhée et de douleurs abdominales au pôle négatif, et la présence de ces symptômes au pôle positif. On peut en conclure que ces variables ont un comportement similaire.

Les variables qui contribuent à la création des dimensions sont les variables actives. Il est également possible de visualiser le positionnement de variables supplémentaires dans les dimensions, sans que ces variables ne participent à la définition de ces dimensions. Cette fonction facultative peut servir à inclure les informations sociodémographiques des participant·e·s.

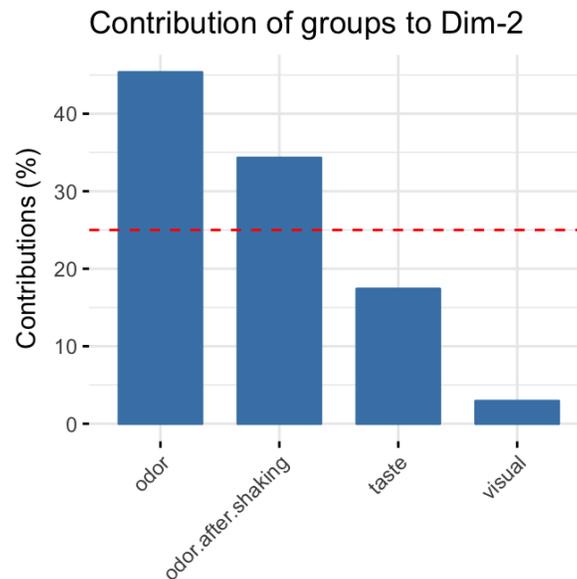
4.4.4 Analyse factorielle multiple

L'analyse factorielle multiple (AFM) a un mécanisme similaire à celui de l'ACM, mais elle permet de regrouper les variables, c'est-à-dire d'étudier simultanément plusieurs jeux de variables ayant les mêmes individus. Dans le cas de variables catégorielles, l'AFM est comparable à la combinaison d'autant d'ACM qu'il y a de groupes de variables. Cette méthode permet d'équilibrer l'influence de chaque groupe de variables (Husson *et al.*, 2016). Par exemple, si un jeu de données

décrivant des plats comporte quatre variables olfactives, neuf variables visuelles et deux variables gustatives, la variation représentée par une ACM sera majoritairement celle décrite par les variables visuelles. La modélisation grâce à l'AFM permet de diviser la variation par les trois groupes de variables, représentant chacune 33,3% de la variation. Ainsi, l'AFM permet d'éviter que des regroupements de variables proportionnellement plus gros que les autres soient surreprésentés dans l'analyse, en plus de rendre possible l'étude des liens entre les groupes de variables.

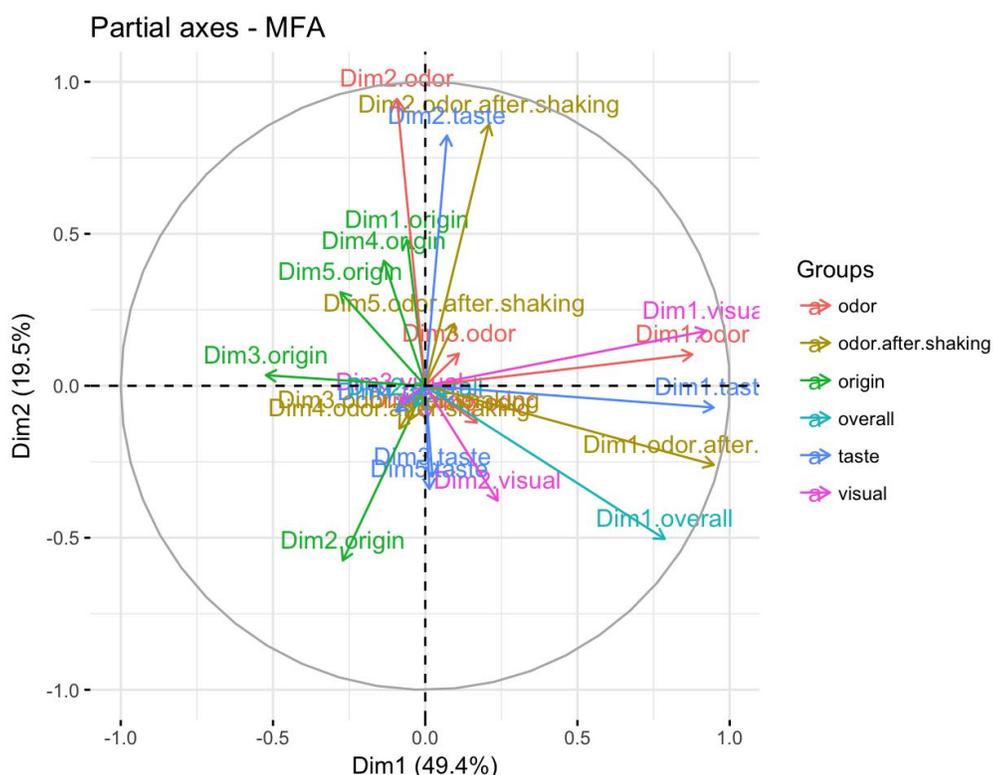
Les graphiques résultant d'une AFM sur des variables catégorielles ainsi que leur interprétation sont les mêmes que pour une ACM. À cela s'ajoutent une représentation de la contribution des groupes à la définition d'une dimension (figure 4.4) ainsi qu'un graphique des axes partiels (figure 4.5)³⁷.

Figure 4.4 – Exemple de contribution d'un groupe de variable



³⁷ Les figures 4.4 et 4.5 sont tirées de l'article sur les AFM du site web de STHDA (Kassambara, 2017b).

Figure 4.5 – Exemple de graphique des axes partiels



Les axes partiels correspondent aux dimensions des ACM (ou autre méthode d'analyse factorielle) effectuées pour chaque groupe de variables (Kassambara, 2017b). À la figure 4.5, on remarque que la première dimension de l'AFM est fortement corrélée aux premières dimensions de quatre groupes de variables (*visual*, *odor*, *taste* et *odor.after.shaking*), tandis que la deuxième est corrélée à la deuxième dimension de trois groupes de variables (*odor*, *taste* et *odor.after.shaking*).

4.4.5 Description des modalités

Les analyses factorielles sont des analyses multidimensionnelles; il est courant de les compléter par des analyses unidimensionnelles afin de caractériser une par une les variables qualitatives présentant un intérêt particulier. Dans le cas d'un jeu de données qualitatives, il est possible de caractériser une variable et les individus définis par ses modalités avec d'autres variables qualitatives (à l'aide de tests de chi-carré) ou avec des modalités de variables catégorielles (Husson *et al.*, 2016). Ces analyses complémentaires sont effectuées par la fonction *catdes* de FactoMineR.

La description des modalités d'une variable par ses liens avec les modalités des autres variables qualitatives se fait par des croisements de variables, de manière similaire au test du chi-carré. La fonction *catdes* sert à comparer la proportion d'individus qui possèdent la modalité A parmi ceux qui possèdent la modalité B à la proportion d'individus qui possèdent la modalité A dans l'entièreté du jeu de données. À partir de cela sont calculées une valeur p et une valeur test indiquant une surreprésentation ou une sous-représentation d'individus présentant la modalité A parmi ceux présentant la modalité B.

Par exemple, dans un jeu de données sur les habitudes de consommation de thé, on génère une description des modalités de la variable *type de thé* (bas de gamme, haut de gamme, variable, etc.) et on observe la caractérisation de ces modalités avec les autres modalités qualitatives du tableau. Pour chaque modalité de *type de thé*, une liste de modalités caractérisantes ayant une valeur p plus petite que 0,05 est générée. Si on observe la description de la modalité *thé haut de gamme* (tableau 4.9³⁸), on remarque que les modalités caractérisant le plus les personnes buvant de ce thé est qu'elles l'achètent en magasin spécialisé (première ligne, association positive) et qu'elles ne l'achètent pas dans des magasins de grande surface (dernière ligne, association négative).

Tableau 4.9 – Exemple de *catdes* d'une modalité

<code>\$category\$t_haut_de_gamme</code>	Cl a/Mod	Mod/Cl a	Global	p.value	v.test
<code>lieuachat=magasin spécialisé</code>	70.00	39.6	10.0	3.16e-11	6.64
<code>forme=vrac</code>	55.60	37.7	12.0	5.59e-08	5.43
<code>variété=noir</code>	28.40	39.6	24.7	1.15e-02	2.53
<code>âge_Q=60 et +</code>	31.60	22.6	12.7	3.76e-02	2.08
<code>ss.effet.santé=ss effet santé</code>	27.30	34.0	22.0	3.81e-02	2.07
<code>ss.effet.santé=Pas.ss effet santé</code>	15.00	66.0	78.0	3.81e-02	-2.07
<code>variété=parfumé</code>	12.40	45.3	64.3	2.86e-03	-2.98
<code>âge_Q=15-24</code>	7.61	13.2	30.7	2.48e-03	-3.03
<code>forme=sachet</code>	8.24	26.4	56.7	1.90e-06	-4.76
<code>lieuachat=GMS</code>	6.25	22.6	64.0	2.62e-11	-6.67

Les résultats d'un test *catdes* prennent la forme d'un tableau pour chaque modalité, tel que le tableau 4.9 pour la modalité *t_haut_de_gamme* (indiqué en haut à droite). La première colonne du tableau indique la modalité caractérisante ainsi que la variable à laquelle elle appartient. La

³⁸ Tableau tiré de Husson *et al.* (2016, p. 154).

deuxième colonne (Cla/Mod) donne le pourcentage d'occurrences de la modalité caractérisante qui présente aussi la modalité décrite dans le tableau, par exemple le pourcentage des occurrences de la catégorie « thé haut de gamme » dans l'ensemble des occurrences correspondant à la modalité « magasin spécialisé » (au tableau 4.9, on observe que de tous les thés achetés dans un magasin spécialisé, 70% correspondent à la catégorie de thé haut de gamme). La troisième colonne (Mod/Cla) indique le pourcentage d'occurrences présentant la modalité décrite ayant aussi la modalité caractérisante, par exemple le pourcentage des occurrences présentant la modalité « magasin spécialisé » dans l'ensemble des occurrences correspondant à la catégorie « thé haut de gamme » (le tableau 4.9 indique que 39.6% des thés de catégorie « haut de gamme » sont achetés dans des magasins spécialisés). La quatrième colonne (Global) donne le pourcentage total des occurrences dans l'entièreté du jeu de données qui correspondent à la modalité caractérisante (par exemple, les thés achetés dans un magasin spécialisé correspondent à 10% de l'ensemble du jeu de données).

CHAPITRE 5 : RÉSULTATS

L'objectif général de ce mémoire est une description de l'emploi des verbes pronominaux dans le corpus Groupe LSQ 98, identifié comme étant porteur des caractéristiques du français sourd. Les réponses aux questions spécifiques de recherche (répétées ci-dessous) sont formulées au moyen de l'observation de la distribution des variables d'observation en fonction des variables contextuelles.

QR1 : Quels sont les contextes morphosyntaxiques ou sémantiques dans lesquels on observe une différence entre la norme du français écrit quant à l'utilisation du pronom SE et son emploi dans le corpus Groupe LSQ 98 ?

QR2 : Est-ce possible d'observer une « norme sourde » dans l'emploi du SE dans le corpus Groupe LSQ 98? Si oui, quels sont les facteurs qui balisent la présence ou l'absence du SE ?

L'effet des variables d'identification est observé à l'aide de tests de chi-carré (5.1), puis la description du corpus se fait en trois niveaux à l'aide de tests statistiques différents : la description générale prend la forme d'une analyse des correspondances multiples comprenant l'ensemble des variables (5.2); la description axée sur l'emploi des verbes pronominaux prend la forme d'une analyse factorielle multiple regroupant les variables contextuelles (5.3); et finalement la description fine est faite à partir d'une *catdes* des variables d'observation (5.4).

5.1 Effet des variables d'identification

Cette section présente les résultats des tests du chi-carrés croisant les variables d'identification ([Particip] et [Texte]) à celles d'observation ([RspNorme] et [PresSe])³⁹.

Les tests du chi-carré de trois croisements sur quatre se sont révélés non significatifs, indiquant que ces paires de variables ne sont pas reliées : [Texte] + [PresSe] (tableau 5.1), [Texte] + [RspNorme] (tableau 5.2), ainsi que [Particip] + [Respect Norme] (tableau 5.3). En d'autres mots, le ratio des occurrences de présence et d'absence pour chaque texte reste similaire à celui de l'ensemble du

³⁹ Pour des raisons d'espace, les variables et modalités seront désignées par leur abréviation seulement dans ce chapitre. Au besoin, consulter la liste des abréviations (p. xii) pour leur nom complet.

corpus, et il en va de même pour le ratio d'occurrences normées et d'écart à la norme. De plus, les participant·e·s ne démontrent pas de tendance individuelle à faire davantage d'écart à la norme.

Tableau 5.1 – Chi-carré de [Texte] + [PresSe]

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	5.047 ^a	5	.410
Likelihood Ratio	5.050	5	.410
Fisher-Freeman-Halton Exact Test	5.011		
N of Valid Cases	341		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 18.88.

Tableau 5.2 – Chi-carré de [Texte] + [RspNorme]

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	8.043 ^a	5	.154
Likelihood Ratio	8.105	5	.151
Fisher-Freeman-Halton Exact Test	8.022		
N of Valid Cases	341		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 16.71.

Tableau 5.3 – Chi-carré de [Particip] + [RspNorme]

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	21.287 ^a	13	.067
Likelihood Ratio	22.783	13	.044
Fisher-Freeman-Halton Exact Test	21.462		
N of Valid Cases	341		

a. 2 cells (7.1%) have expected count less than 5. The minimum expected count is 4.08.

Néanmoins, le test du chi-carré de [Particip] + [PresSe] (tableau 5.4) indique que ces variables ont une corrélation statistiquement significative ($p < 0.01$); les participant·e·s démontrent donc des tendances individuelles dans leur taux de production du pronom SE.

Tableau 5.4 – Chi-carré de [Particip] + [PresSe]

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	37.053 ^a	13	.000
Likelihood Ratio	40.240	13	.000
Fisher-Freeman-Halton Exact Test	37.694		
N of Valid Cases	341		

a. 1 cells (3.6%) have expected count less than 5. The minimum expected count is 4.60.

Lorsque l'on effectue un z-test pour préciser ce résultat dans un tableau croisé (tableau 5.6), on remarque que ces tendances ne proviennent que de trois participant·e·s : P15, qui produit davantage de verbes pronominaux que le reste du groupe; ainsi que P10 et P11, qui en produisent moins. Par contre, dans le cas de P10, ce résultat s'explique par la distribution de [NormePres], c'est-à-dire par ses choix lexicaux. Si l'on croise [NormePres] et [Particip] (dont le test du chi-carré indique une relation significative, voir tableau 5.5) et que l'on observe les résultats du z-test (tableau 5.7), on remarque que P10 a significativement moins d'occurrences correspondant à la modalité < PN_O >. En d'autres mots, P10 a employé moins de verbes pour lesquels la présence du SE est obligatoire, ce qui explique que ses textes contiennent plus d'absences du SE.

Tableau 5.5 – Chi-carré de [Particip] + [NormePres]

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	52.277 ^a	26	.002
Likelihood Ratio	59.296	26	.000
Fisher-Freeman-Halton Exact Test	50.327		
N of Valid Cases	341		

a. 17 cells (40.5%) have expected count less than 5. The minimum expected count is 1.55.

Tableau 5.6 – Tableau croisé de [Particip] + [PresSe]

Particip		PresSe		
		PresN	PresO	Total
P01	Count	2 _a	8 _a	10
	Expected Count	4.6	5.4	10.0
	% within Particip	20.0%	80.0%	100.0%
	% within PresSe	1.3%	4.3%	2.9%
P02	Count	9 _a	15 _a	24
	Expected Count	11.0	13.0	24.0
	% within Particip	37.5%	62.5%	100.0%
	% within PresSe	5.7%	8.2%	7.0%
P03	Count	21 _a	17 _a	38
	Expected Count	17.5	20.5	38.0
	% within Particip	55.3%	44.7%	100.0%
	% within PresSe	13.4%	9.2%	11.1%
P04	Count	9 _a	17 _a	26
	Expected Count	12.0	14.0	26.0
	% within Particip	34.6%	65.4%	100.0%
	% within PresSe	5.7%	9.2%	7.6%
P05	Count	17 _a	27 _a	44
	Expected Count	20.3	23.7	44.0
	% within Particip	38.6%	61.4%	100.0%
	% within PresSe	10.8%	14.7%	12.9%
P06	Count	16 _a	12 _a	28
	Expected Count	12.9	15.1	28.0
	% within Particip	57.1%	42.9%	100.0%
	% within PresSe	10.2%	6.5%	8.2%
P07	Count	6 _a	13 _a	19
	Expected Count	8.7	10.3	19.0
	% within Particip	31.6%	68.4%	100.0%
	% within PresSe	3.8%	7.1%	5.6%
P08	Count	8 _a	9 _a	17
	Expected Count	7.8	9.2	17.0
	% within Particip	47.1%	52.9%	100.0%
	% within PresSe	5.1%	4.9%	5.0%
P09	Count	9 _a	17 _a	26
	Expected Count	12.0	14.0	26.0
	% within Particip	34.6%	65.4%	100.0%
	% within PresSe	5.7%	9.2%	7.6%
P10	Count	18 _a	9 _b	27
	Expected Count	12.4	14.6	27.0
	% within Particip	66.7%	33.3%	100.0%
	% within PresSe	11.5%	4.9%	7.9%
P11	Count	11 _a	1 _b	12
	Expected Count	5.5	6.5	12.0
	% within Particip	91.7%	8.3%	100.0%
	% within PresSe	7.0%	0.5%	3.5%
P13	Count	19 _a	14 _a	33
	Expected Count	15.2	17.8	33.0
	% within Particip	57.6%	42.4%	100.0%
	% within PresSe	12.1%	7.6%	9.7%
P14	Count	10 _a	9 _a	19
	Expected Count	8.7	10.3	19.0
	% within Particip	52.6%	47.4%	100.0%
	% within PresSe	6.4%	4.9%	5.6%
P15	Count	2 _a	16 _b	18
	Expected Count	8.3	9.7	18.0
	% within Particip	11.1%	88.9%	100.0%
	% within PresSe	1.3%	8.7%	5.3%
Total	Count		184	341
	Expected Count		184.0	341.0
	% within Particip		54.0%	100.0%
	% within PresSe		100.0%	100.0%

Each subscript letter denotes a subset of PresSe categories whose column proportions do not differ significantly from each other at the .05 level.

Tableau 5.7 – Tableau croisé de [Particip] + [NormePres]

		NormePres			Total
		PN_Fac	PN_N	PN_O	
P01	Count	2 _a	0 _a	8 _a	10
	Expected Count	2.3	1.6	6.1	10.0
	% within Particip	20.0%	0.0%	80.0%	100.0%
	% within NormePres	2.5%	0.0%	3.8%	2.9%
P02	Count	6 _a	3 _a	15 _a	24
	Expected Count	5.6	3.7	14.7	24.0
	% within Particip	25.0%	12.5%	62.5%	100.0%
	% within NormePres	7.6%	5.7%	7.2%	7.0%
P03	Count	8 _a	3 _a	27 _a	38
	Expected Count	8.8	5.9	23.3	38.0
	% within Particip	21.1%	7.9%	71.1%	100.0%
	% within NormePres	10.1%	5.7%	12.9%	11.1%
P04	Count	6 _{a, b}	9 _b	11 _a	26
	Expected Count	6.0	4.0	15.9	26.0
	% within Particip	23.1%	34.6%	42.3%	100.0%
	% within NormePres	7.6%	17.0%	5.3%	7.6%
P05	Count	10 _{a, b}	13 _b	21 _a	44
	Expected Count	10.2	6.8	27.0	44.0
	% within Particip	22.7%	29.5%	47.7%	100.0%
	% within NormePres	12.7%	24.5%	10.0%	12.9%
P06	Count	11 _a	0 _b	17 _a	28
	Expected Count	6.5	4.4	17.2	28.0
	% within Particip	39.3%	0.0%	60.7%	100.0%
	% within NormePres	13.9%	0.0%	8.1%	8.2%
P07	Count	1 _a	6 _b	12 _{a, b}	19
	Expected Count	4.4	3.0	11.6	19.0
	% within Particip	5.3%	31.6%	63.2%	100.0%
	% within NormePres	1.3%	11.3%	5.7%	5.6%
P08	Count	5 _a	1 _a	11 _a	17
	Expected Count	3.9	2.6	10.4	17.0
	% within Particip	29.4%	5.9%	64.7%	100.0%
	% within NormePres	6.3%	1.9%	5.3%	5.0%
P09	Count	3 _a	4 _a	19 _a	26
	Expected Count	6.0	4.0	15.9	26.0
	% within Particip	11.5%	15.4%	73.1%	100.0%
	% within NormePres	3.8%	7.5%	9.1%	7.6%
P10	Count	11 _a	4 _{a, b}	12 _b	27
	Expected Count	6.3	4.2	16.5	27.0
	% within Particip	40.7%	14.8%	44.4%	100.0%
	% within NormePres	13.9%	7.5%	5.7%	7.9%
P11	Count	4 _a	0 _a	8 _a	12
	Expected Count	2.8	1.9	7.4	12.0
	% within Particip	33.3%	0.0%	66.7%	100.0%
	% within NormePres	5.1%	0.0%	3.8%	3.5%
P13	Count	5 _a	2 _a	26 _a	33
	Expected Count	7.6	5.1	20.2	33.0
	% within Particip	15.2%	6.1%	78.8%	100.0%
	% within NormePres	6.3%	3.8%	12.4%	9.7%
P14	Count	6 _a	4 _a	9 _a	19
	Expected Count	4.4	3.0	11.6	19.0
	% within Particip	31.6%	21.1%	47.4%	100.0%
	% within NormePres	7.6%	7.5%	4.3%	5.6%
P15	Count	1 _a	4 _a	13 _a	18
	Expected Count	4.2	2.8	11.0	18.0
	% within Particip	5.6%	22.2%	72.2%	100.0%
	% within NormePres	1.3%	7.5%	6.2%	5.3%
Total	Count	79	53	209	341
	Expected Count	79.0	53.0	209.0	341.0
	% within Particip	23.2%	15.5%	61.3%	100.0%
	% within NormePres	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of NormePres categories whose column proportions do not differ significantly from each other at the .05 level.

5.2 Analyse générale (ACM)

La description générale du corpus a été réalisée à l'aide d'une analyse des correspondances multiples, qui a permis la modélisation de l'ensemble de la variation du jeu de données. À cette fin, toutes les variables ont contribué à la définition des dimensions, sauf cinq considérées comme variables supplémentaires : celles d'identification ([Particip] et [Texte]) et celles qui interagissent trop avec d'autres variables à l'étude, soit [CatOcc], [NormePres] et [vPro] (ces interactions sont explicitées dans les résultats *catdes* en 5.4).

Le *scree plot* à la figure 5.1 représente les dimensions créées. On y remarque qu'il est difficile de synthétiser l'information contenue dans le corpus ; les dix dimensions les plus importantes n'expliquent que 45.5% de la variation totale du jeu de données. Néanmoins, le pourcentage de variation expliqué par les deux premières dimensions est tout de même plus élevé que si la variation se distribuait aléatoirement parmi les variables à l'étude ; on peut donc considérer ces dimensions dans l'analyse.

Figure 5.1 – Scree plot de l'ACM

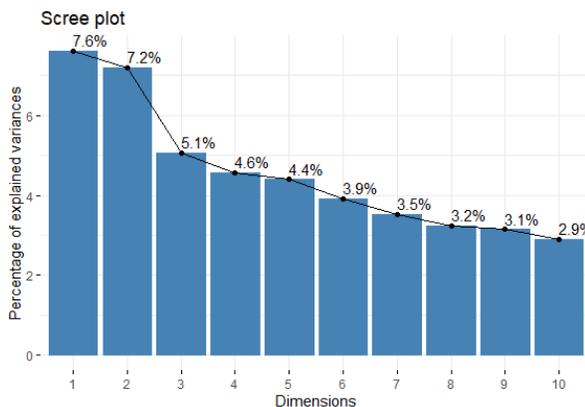
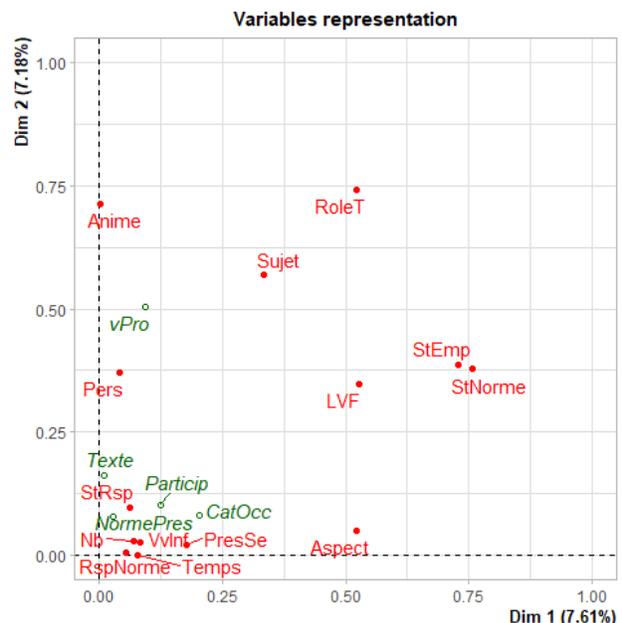


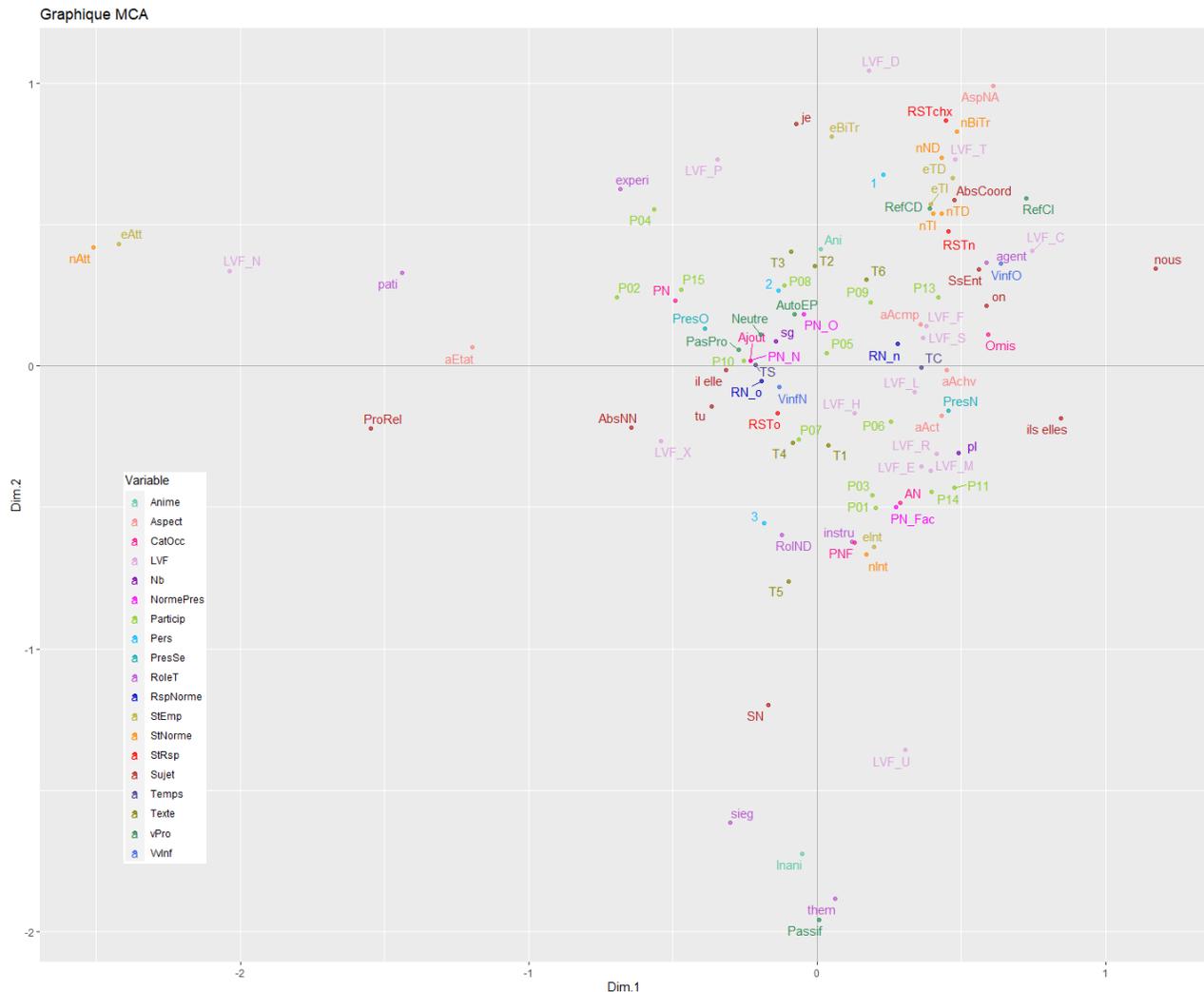
Figure 5.2 – Carré des liaisons de l'ACM



On remarque dans le carré des liaisons (figure 5.2) que de nombreuses variables se retrouvent très près de l'origine : leur contribution à la définition des deux dimensions est minime. La première dimension est définie principalement par la structure transitive (tant celle employée que celle

normée) de la phrase dans laquelle se trouve l'occurrence, tandis que la deuxième dimension est principalement définie par les variables touchant les traits sémantiques du sujet du verbe. Le graphique principal (figure 5.3) permet de préciser ces observations.

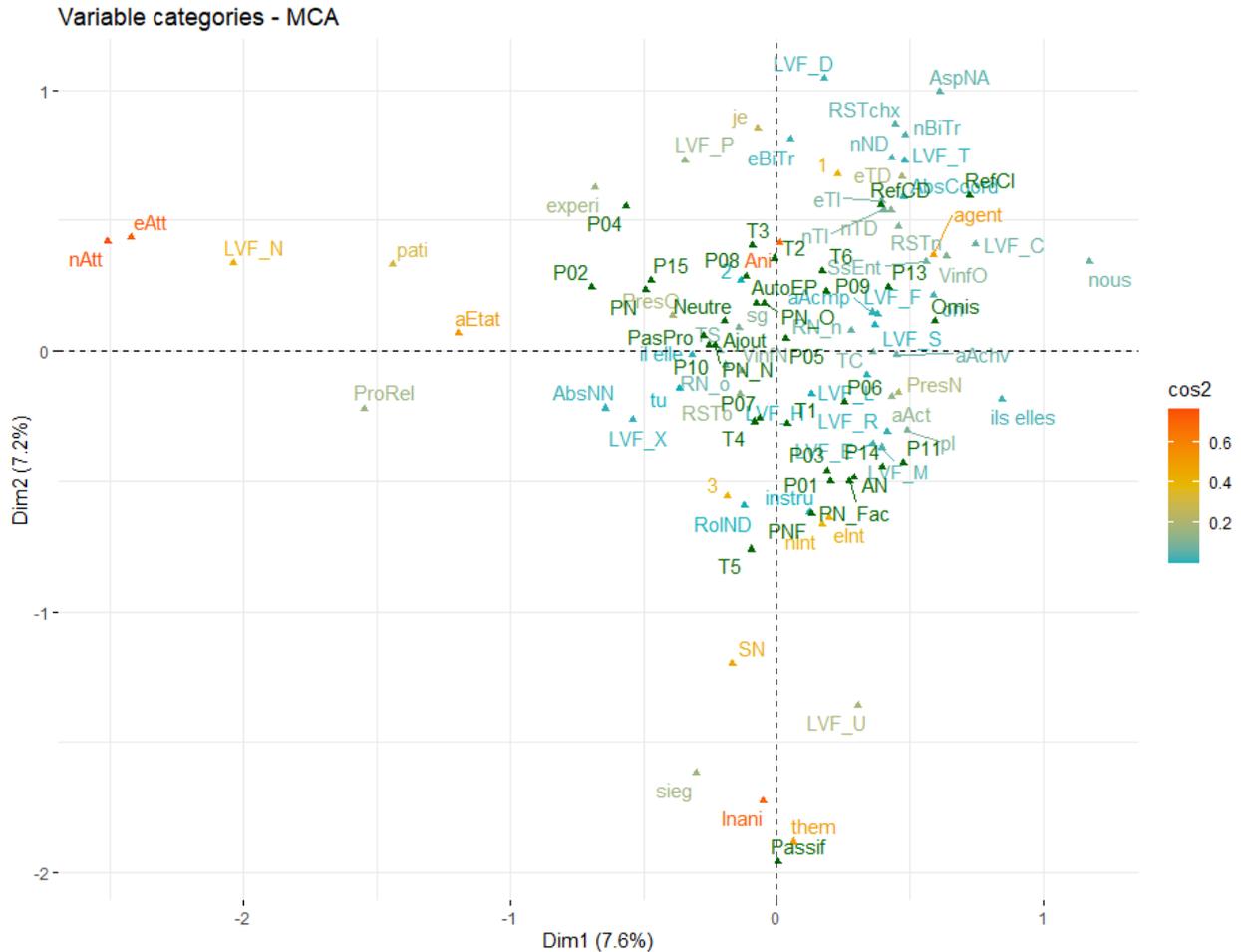
Figure 5.3 – Graphique principal de l'ACM



La première dimension est marquée par la présence de < nAtt > et < eAtt > à l'extrémité de son pôle négatif, en opposition à toutes les autres modalités de structure transitive employée et normée, qui se retrouvent sur le côté positif de la dimension. La deuxième dimension est aussi marquée par les variables sur son pôle négatif : les occurrences dont le sujet prend la forme d'un syntagme nominal ayant le rôle thématique thème ou siège et dont le référent est inanimé se distinguent du reste – ces modalités sont d'ailleurs associées aux constructions pronominales passives. La

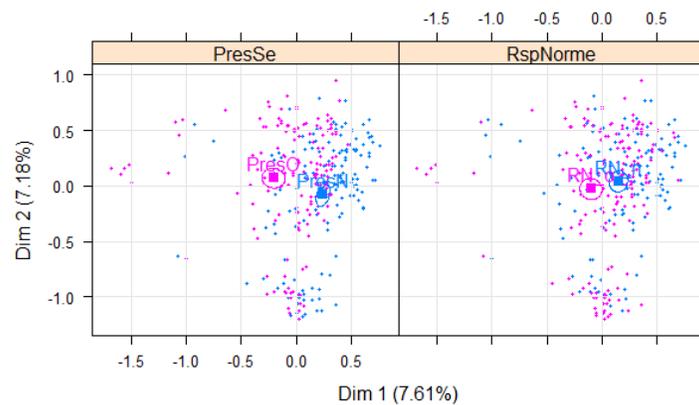
colorisation des modalités en fonction de la qualité de leur représentation (\cos^2) confirme ces observations (les variables supplémentaires sont en vert) :

Figure 5.4 – Qualité de représentation des variables de l’ACM



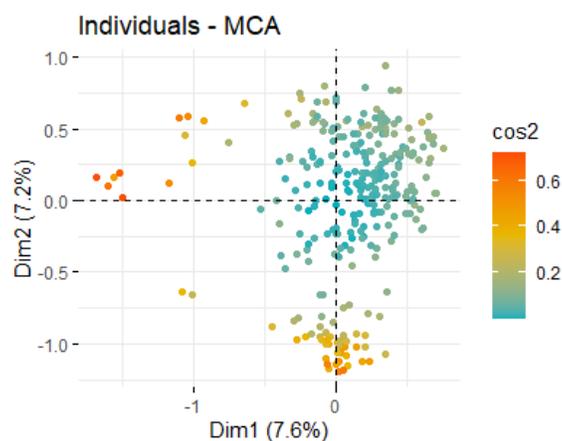
Ce qui ressort principalement de l’ACM est que l’analyse factorielle de l’ensemble des variables du corpus ne permet pas de décrire l’emploi des verbes pronominaux. En effet, les variables de présence ou absence du SE ainsi que du respect ou non de la norme se retrouvent dans le nuage de points, près de l’origine. Leur contribution à la première dimension est minime, et presque nulle pour la deuxième dimension. Comme on peut le voir à la figure 5.5, le placement des individus dans les dimensions ne permet pas un regroupement efficace des modalités des variables [PresSe] et [RspNorme].

Figure 5.5 – Regroupement des individus de l’ACM



En d’autres mots, la variation du corpus dépasse celle générée par l’emploi des verbes pronominaux. La langue naturelle est telle qu’on ne peut pas réellement créer des variables indépendantes sans que des modalités soient interreliées; par exemple il est normal et attendu que les verbes en construction passive aient des sujets ayant le rôle thématique thème. De plus, certaines constructions fréquentes créent une surreprésentation de certaines combinaisons de modalités – par exemple, le côté négatif de la première dimension souligne la grande fréquence de la subordonnée relative « qui s’appelle ‘nom’ ». La figure 5.6 représente la position des individus sur le graphique, coloriés selon la qualité de leur représentation. Ce graphique confirme que l’ACM représente surtout les individus placés aux pôles négatifs des deux dimensions.

Figure 5.6 – Qualité de représentation des individus de l’ACM



Cela a pour effet comparable à celui de « tirer sur la couverture » de la variation du jeu de données : les combinaisons fréquentes représentent un grand pourcentage de la variation totale, et il en reste peu pour les autres combinaisons.

5.3 Analyse orientée sur les variables d'observation (AFM)

L'analyse factorielle multiple permet de limiter l'effet des constructions fréquentes sur la variation et d'attribuer plus de poids aux variables d'emploi des verbes pronominaux grâce au regroupement des variables. L'AFM réalisée pour l'analyse orientée sur les variables d'observation est constituée de quatre groupes. Les deux premiers, nommés « Pres » et « Norme », ne comportent qu'une variable binaire chacun – soit [PresSe] et [RspNorme], respectivement –, tandis que le troisième groupe, « Contexte », est constitué de l'ensemble des variables contextuelles. Cela a pour effet de diminuer l'importance attribuée à ces variables en raison de leur nombre élevé afin d'axer l'analyse sur les variables d'observation. Le quatrième groupe, « Sup », contient les variables supplémentaires.

Le *scree plot* de l'AFM (figure 5.7) démontre une courbe plus marquée que celle de l'AFM. On remarque dans le carré des variables (figure 5.8) que le regroupement des variables contextuelles permet une plus grande contribution des variables d'observation. La première dimension est principalement définie par la variable [PresSe], bien que les deux autres groupes ont aussi une forte contribution. Au contraire, la deuxième dimension est majoritairement définie par les variables contextuelles, avec une contribution moindre des variables d'observation.

Figure 5.7 – *Scree plot* de l'AFM

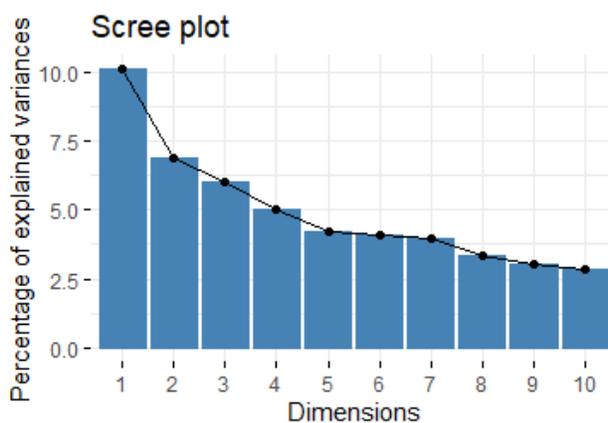
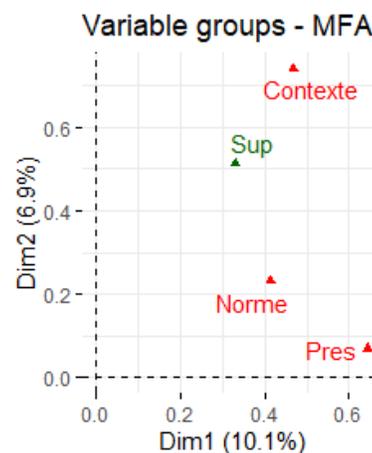
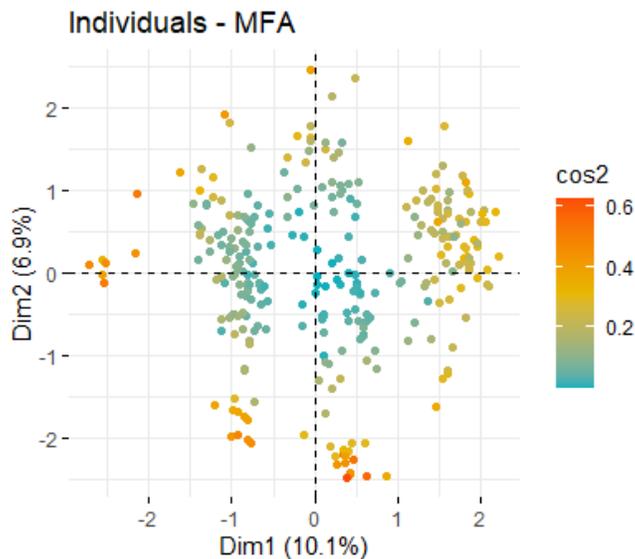


Figure 5.8 – Carré des liaisons de l'AFM



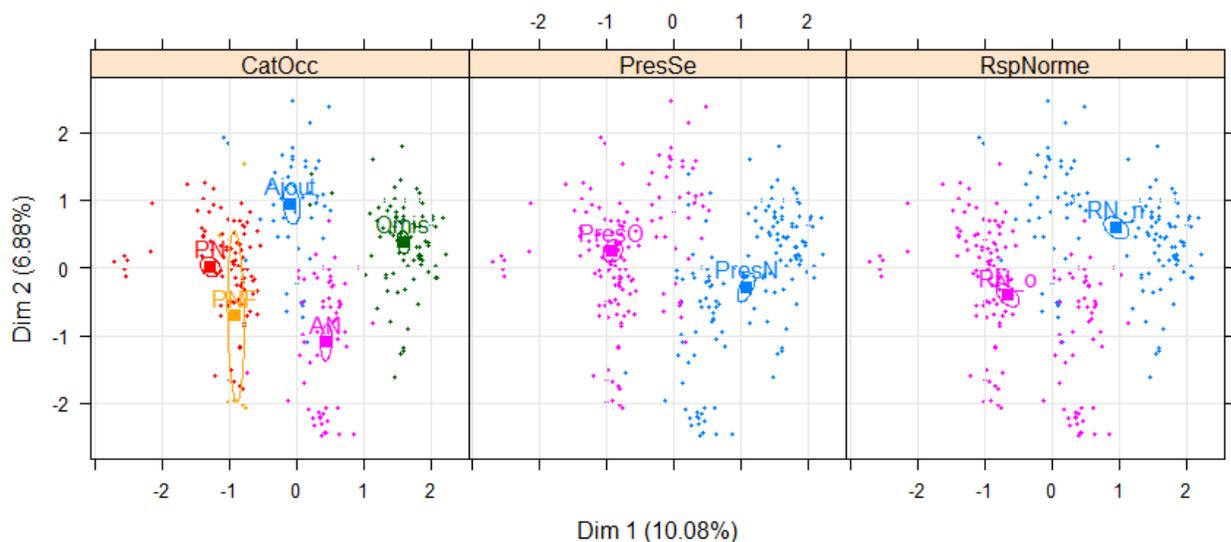
L'équilibre de l'influence des regroupements de variables est bien visible dans le graphique principal de l'AFM (figure 5.9), qui est similaire à celui de l'ACM, mais dont les combinaisons

Figure 5.10 – Qualité de représentation des individus de l'AFM



L'AFM permet aussi un meilleur regroupement des individus en fonction des variables d'observation :

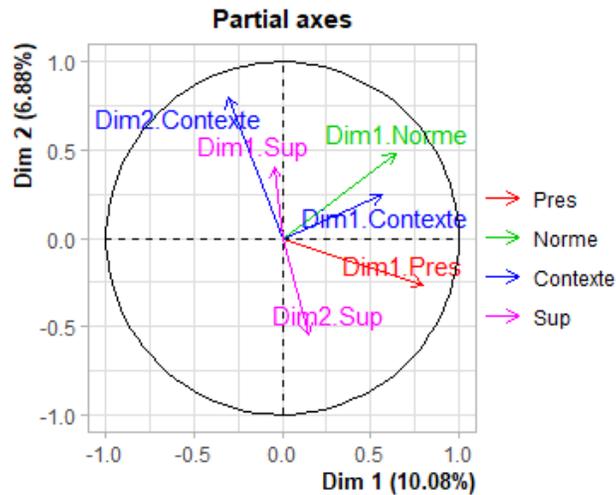
Figure 5.11 – Regroupement des individus de l'AFM



Le graphique des axes partiels (figure 5.12) permet de comparer les groupes de variables en représentant les relations entre les analyses factorielles de chaque groupe (à noter que les groupes « Norme » et « Pres », ne contenant qu'une variable binaire, n'ont qu'une seule dimension créant un axe allant d'une modalité à l'autre). On remarque que les dimensions des variables

supplémentaires sont corrélées avec la deuxième dimension de l'AFM, tant dans l'axe positif que dans l'axe négatif.

Figure 5.12 – Axes partiels de l'AFM



Bien que l'AFM permette de réduire l'importance que prennent les nombreuses variables contextuelles, il demeure que la variation du jeu de données provient principalement des ces variables plutôt que des variables d'observation, comme le démontre le tableau des contributions des groupes de variables aux dix premières dimensions de l'AFM :

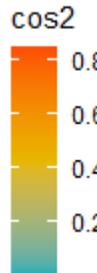
Tableau 5.8 – Contribution (%) des groupes de variables aux dimensions de l'AFM

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5	Dim.6	Dim.7	Dim.8	Dim.9	Dim.10
Pres	42.22	6.72	0.76	23.88	6.24	0.95	0.14	2.33	1.54	0.09
Norme	27.08	22.13	31.44	0.40	5.65	0.70	0.17	0.31	0.01	0.06
Contexte	30.70	71.14	67.81	75.72	88.11	98.34	99.69	97.36	98.45	99.86

Tout de même, les deux premières dimensions de l'AFM offrent une bonne qualité de représentation de la variation des variables [PresSe], [RspNorme] et [CatOcc], comme on le voit à la dernière colonne du tableau 5.9.

Tableau 5.9 – Qualité de représentation des modalités des variables d’observation (AFM)

	Dim.1	Dim.2	D1 + D2
PresN	0.78822	0.05850	0.84672
PresO	0.78822	0.05850	0.84672
RN_n	0.54132	0.20615	0.74747
RN_o	0.54132	0.20615	0.74747
Ajout	0.00171	0.29019	0.29190
AN	0.08266	0.50541	0.58806
Omis	0.84623	0.04727	0.89350
PN	0.88849	0.00013	0.88862
PNF	0.22353	0.13994	0.36347



5.4 Analyse fine des variables d’observation (*catdes*)

Cette section présente les résultats de la fonction *catdes* (description de la catégorie) appliquée aux variables d’observation, soit [PresSe] (5.4.1), [RspNorme] (5.4.2), [CatOcc] (5.4.3) et [NormePres] (5.4.4), en plus de traiter de l’interaction de [NormePres] avec les autres variables d’observation (5.4.5).

Dans les tableaux de cette section, les cases bleues présentent des variables ayant une interaction statistiquement significative avec la catégorie décrite selon le test du chi-carré, les cases vertes contiennent des modalités positivement corrélées avec la modalité décrite et les cases orange indiquent des modalités négativement corrélées (en d’autres mots, la population décrite par une modalité est caractérisée comme présentant les modalités en vert, et comme n’ayant pas les modalités en orange). Les cases grises indiquent des éléments dont l’interaction avec la variable décrite est attendue, que ce soit une conséquence du design de l’étude (par exemple la décision de ne pas retenir les verbes sans SE pour lesquels la présente du SE est interdite) ou alors le résultat de l’interaction entre les différents types de verbes pronominaux et les variables d’observation (par

exemple, il est impossible qu'une occurrence de verbe essentiellement pronominal soit catégorisée comme un ajout du SE). Les lignes grisées indiquent une interaction totale entre les catégories (par exemple, l'absence du SE et les occurrences d'omission), tandis que les lignes à une seule case grisée indiquent une interaction moins complète (par exemple, il est attendu que [vPro] interagisse avec [RspNorme] car < PasPro > est obligatoirement combiné à < RN_n >, mais les autres modalités de [vPro] ne peuvent pas être prédites ainsi).

5.4.1 Résultats *catdes* de [PresSe]

Le tableau 5.10 présente les variables dont le test du chi-carré indique un résultat significatif ($p < 0.05$) pour leur croisement avec [PresSe]. Sont grisées les variables [CatOcc], dont la modalité dépend de la valeur de [PresSe]; [NormePres], pour l'impossibilité de la combinaison < PresN > + < PN_N >; [vPro], pour l'impossibilité de la combinaison < PresN > + < PasPro >; et [RspNorme], étant donnée la décision de ne pas retenir les verbes sans SE ne pouvant être employés avec SE, réduisant ainsi drastiquement le nombre d'occurrences < PresN > + < RN_o >.

Tableau 5.10 – Résultat *catdes* pour [PresSe] (variables)

Variables		
	p.value	df
CatOcc	1.54E-72	4
NormePres	2.99E-24	2
vPro	4.55E-14	5
Temps	2.28E-07	1
RspNorme	1.15E-06	1
StEmp	1.53E-06	4
Nb	5.12E-06	1
LVF	1.13E-05	13
Sujet	6.19E-05	10
StNorme	9.12E-05	5
Particip	4.06E-04	13
RoleT	4.40E-03	6
VvInf	7.18E-03	1
Aspect	4.81E-02	4

Les variables sont classées selon leur valeur p; donc, en faisant fi des variables grisées, la variable ayant l'interaction la plus marquée avec [PresSe] est [Temps].

Le tableau 5.11 présente les modalités caractérisantes des occurrences de présence du SE ainsi que celles des occurrences d'absence du SE (pour lesquelles $p < 0.05$). Ainsi, si on lit la première ligne du tableau pour < PresO >, on apprend que la catégorie d'occurrence de présence normée est caractérisante des présences du SE. La totalité des occurrences présentant la modalité < PN > pour [CatOcc] présente aussi la modalité < PresO > pour [PresSe] (Cla/Mod = 100%). De l'ensemble des occurrences de présence du SE, 66.85% appartiennent à la catégorie de présence normée (Mod/Cla), tandis que cette catégorie d'occurrence ne représente que 36.07% de l'entièreté du corpus (Global).

Il est à noter que, comme [PresSe] est binaire, la description d'une modalité correspond à l'inverse de la description de l'autre modalité.

Tableau 5.11 – Résultat *catdes* pour [PresSe] (modalités)

	PresO (présence)					PresN (absence)					
	Cla/Mod	Mod/Cla	Global	p.value	v.test	Cla/Mod	Mod/Cla	Global	p.value	v.test	
CatOcc=PN	100	66.85	36.07	1.20E-46	14.34	CatOcc=Omis	100	54.78	25.22	2.66E-37	12.76
CatOcc=Ajout	100	28.80	15.54	1.09E-16	8.29	CatOcc=AN	100	45.22	20.82	1.85E-29	11.27
NormePres=PN_N	100	28.80	15.54	1.09E-16	8.29	NormePres=PN_F	89.87	45.22	23.17	2.38E-20	9.24
vPro=PasPro	100	26.63	14.37	2.44E-15	7.92	Temps=TC	64.29	51.59	36.95	2.40E-07	5.17
Temps=TS	64.65	75.54	63.05	2.40E-07	5.17	RspNorme=RN_n	61.87	54.78	40.76	1.20E-06	4.85
RspNorme=RN_o	64.85	71.20	59.24	1.20E-06	4.85	vPro=Passif	81.08	19.11	10.85	5.15E-06	4.56
StEmp=eAtt	89.19	17.93	10.85	1.84E-06	4.77	Nb=pl	68.83	33.76	22.58	5.47E-06	4.55
StNorme=nAtt	88.89	17.39	10.56	3.32E-06	4.65	Sujet=nous	100	9.55	4.40	6.07E-06	4.52
Nb=sg	60.61	86.96	77.42	5.47E-06	4.55	Particip=P11	91.67	7.01	3.52	1.23E-03	3.23
LVF=LVF_N	87.10	14.67	9.09	5.76E-05	4.02	vPro=RefCl	75	13.38	8.21	1.46E-03	3.18
Particip=P15	88.89	8.70	5.28	1.70E-03	3.14	StEmp=elnt	54.55	57.32	48.39	2.38E-03	3.04
StEmp=eBiTr	100	5.43	2.93	1.86E-03	3.11	StNorme=nInt	54.84	54.14	45.45	3.06E-03	2.96
RoleT=experi	71.43	21.74	16.42	3.98E-03	2.88	RoleT=agent	53.41	59.87	51.61	4.98E-03	2.81
Sujet=je	67.05	32.07	25.81	4.23E-03	2.86	VvInf=VinfO	62.07	22.93	17.01	7.77E-03	2.66
CatOcc=PNF	100	4.35	2.35	6.69E-03	2.71	LVF=LVF_T	100	3.82	1.76	9.04E-03	2.61
VvInf=VinfN	57.24	88.04	82.99	7.77E-03	2.66	LVF=LVF_U	66.67	12.74	8.80	1.95E-02	2.34
LVF=LVF_P	68.85	22.83	17.89	9.96E-03	2.58	Particip=P10	66.67	11.46	7.92	2.77E-02	2.20
Aspect=aEtat	64.84	32.07	26.69	1.52E-02	2.43	Sujet=ProRel	23.81	3.18	6.16	3.58E-02	-2.10
NormePres=PN_O	58.85	66.85	61.29	2.34E-02	2.27	RoleT=pati	30.23	8.28	12.61	2.64E-02	-2.22
RoleT=pati	69.77	16.30	12.61	2.64E-02	2.22	NormePres=PN_O	41.15	54.78	61.29	2.34E-02	-2.27
Sujet=ProRel	76.19	8.70	6.16	3.58E-02	2.10	Aspect=aEtat	35.16	20.38	26.69	1.52E-02	-2.43
Particip=P10	33.33	4.89	7.92	2.77E-02	-2.20	LVF=LVF_P	31.15	12.10	17.89	9.96E-03	-2.58
LVF=LVF_U	33.33	5.43	8.80	1.95E-02	-2.34	VvInf=VinfN	42.76	77.07	82.99	7.77E-03	-2.66
LVF=LVF_T	0	0	1.760	9.04E-03	-2.61	CatOcc=PNF	0	0	2.35	6.69E-03	-2.71

VvInf=VinfO	37.93	11.96	17.01	7.77E-03	-2.66	Sujet=je	32.95	18.47	25.81	4.23E-03	-2.86
RoleT=agent	46.59	44.57	51.61	4.98E-03	-2.81	RoleT=experi	28.57	10.19	16.42	3.98E-03	-2.88
StNorme=nInt	45.16	38.04	45.45	3.06E-03	-2.96	StEmp=eBiTr	0	0	2.93	1.86E-03	-3.11
StEmp=eInt	45.45	40.76	48.39	2.38E-03	-3.04	Particip=P15	11.11	1.27	5.28	1.70E-03	-3.14
vPro=RefCl	25	3.80	8.21	1.46E-03	-3.18	LVF=LVF_N	12.90	2.55	9.09	5.76E-05	-4.02
Particip=P11	8.33	0.54	3.52	1.23E-03	-3.23	Nb=sg	39.39	66.24	77.42	5.47E-06	-4.55
Sujet=nous	0	0	4.40	6.07E-06	-4.52	StNorme=nAtt	11.11	2.55	10.56	3.32E-06	-4.65
Nb=pl	31.17	13.04	22.58	5.47E-06	-4.55	StEmp=eAtt	10.81	2.55	10.85	1.84E-06	-4.77
vPro=Passif	18.92	3.80	10.85	5.15E-06	-4.56	RspNorme=RN_o	35.15	45.22	59.24	1.20E-06	-4.85
RspNorme=RN_n	38.13	28.80	40.76	1.20E-06	-4.85	Temps=TS	35.35	48.41	63.05	2.40E-07	-5.17
Temps=TC	35.71	24.46	36.95	2.40E-07	-5.17	vPro=PasPro	0	0	14.37	2.44E-15	-7.92
NormePres=PN_F	10.13	4.35	23.17	2.38E-20	-9.24	CatOcc=Ajout	0	0	15.54	1.09E-16	-8.29
CatOcc=AN	0	0	20.82	1.85E-29	-11.27	NormePres=PN_N	0	0	15.54	1.09E-16	-8.29
CatOcc=Omis	0	0	25.22	2.66E-37	-12.76	CatOcc=PN	0	0	36.07	1.20E-46	-14.34

Ainsi, en faisant abstraction des lignes grisées, la modalité la plus caractérisante des populations définies par la variable [PresSe] est < PN_Fac > de la variable [NormePres], étant donné que sa valeur test est la plus éloignée de 0. Les absences du SE sont caractérisées par le fait que la présence du SE est facultative (valeur test positive), tandis que les présences du SE sont caractérisées par le fait que la présence du SE n'est pas facultative (valeur test négative). Plus précisément, lorsque la présence du SE est facultative, il sera absent dans 90% des cas.

La description par des modalités permet aussi de préciser la description par des variables. Ainsi, l'interaction de [Temps] avec [PresSe] correspond à plus d'occurrences de présence du SE avec les temps simples et plus d'occurrences d'absence du SE avec les temps composés.

5.4.2 Résultats *catdes* de [RspNorme]

La description de la variable [RspNorme] donne moins d'interactions que celle de [PresSe]. Dans le tableau 5.12, les variables grisées sont [CatOcc], dont la modalité dépend de celle de [RspNorme]; [NormePres], car < PN_Fac > correspond de facto à < RN_o >; [vPro], puisque la combinaison < RN_o > et < PasPro > est impossible; et [PresSe], étant donné que la majorité des verbes du corpus sont des absences normées n'ayant pas été retenues pour l'analyse, résultant en une surreprésentation de < PresO > dans la population de < RN_o >.

Tableau 5.12 – Résultat *catdes* pour [RspNorme] (variables)

Variables		
	p.value	df
CatOcc	1.54E-72	4
NormePres	2.94E-29	2
vPro	3.34E-18	5
StRsp	6.50E-07	2
PresSe	1.15E-06	1
LVF	4.95E-05	13

Deux autres variables présentent une interaction statistiquement significative, soit le respect de la structure transitive et la catégorie syntactico-sémantique du LVF. Plus précisément, le tableau 5.13 indique que le respect de la structure transitive est caractérisant des occurrences de respect de la norme, tandis que le non respect de la structure transitive est caractérisant des écarts à la norme. La modalité < LVF_N > est associée aux occurrences de respect de la norme, tandis que < LVF_X >, < LVF_S > et < LVF_P > sont associées aux écarts à la norme.

Tableau 5.13 – Résultat *catdes* pour [RspNorme] (modalités)

	RN_o					RN_n					
	Cla/Mod	Mod/Cla	Global	p.value	v.test	Cla/Mod	Mod/Cla	Global	p.value	v.test	
CatOcc=PN	100	60.89	36.07	9.75E-39	13.02	CatOcc=Omis	100	61.87	25.22	4.33E-44	13.93
NormePres=PN_F	100	39.11	23.17	3.97E-22	9.67	CatOcc=Ajout	100	38.13	15.54	1.64E-24	10.22
CatOcc=AN	100	35.15	20.82	1.45E-19	9.05	NormePres=PN_N	100	38.13	15.54	1.64E-24	10.22
PresSe=PresO	71.20	64.85	53.96	1.20E-06	4.85	vPro=PasPro	100	35.25	14.37	1.91E-22	9.75
StRsp=RSTo	66.03	85.64	76.83	4.53E-06	4.59	StRsp=RSTn	70.31	32.37	18.77	1.38E-07	5.27
LVF=LVF_N	93.55	14.36	9.09	1.13E-05	4.39	PresSe=PresN	54.78	61.87	46.04	1.20E-06	4.85
vPro=RefCl	89.29	12.38	8.21	3.95E-04	3.54	LVF=LVF_X	83.33	7.19	3.52	3.17E-03	2.95
Particip=P06	82.14	11.39	8.21	8.68E-03	2.62	StNorme=nND	66.67	8.63	5.28	2.66E-02	2.22
vPro=Passif	78.38	14.36	10.85	1.10E-02	2.54	Aspect=aAct	48.03	43.88	37.24	3.69E-02	2.09
CatOcc=PNF	100	3.96	2.35	1.43E-02	2.45	LVF=LVF_S	69.23	6.47	3.81	4.17E-02	2.04
StNorme=nAtt	77.78	13.86	10.56	1.55E-02	2.42	LVF=LVF_P	52.46	23.02	17.89	4.34E-02	2.02
vPro=Neutre	66.92	43.07	38.12	2.37E-02	2.26	Particip=P01	10	0.72	2.93	4.46E-02	-2.01
StEmp=eAtt	75.68	13.86	10.85	3.02E-02	2.17	StEmp=eAtt	24.32	6.47	10.85	3.02E-02	-2.17
Particip=P01	90	4.46	2.93	4.46E-02	2.01	vPro=Neutre	33.08	30.94	38.12	2.37E-02	-2.26
LVF=LVF_P	47.54	14.36	17.89	4.34E-02	-2.02	StNorme=nAtt	22.22	5.76	10.56	1.55E-02	-2.42
LVF=LVF_S	30.77	1.98	3.81	4.17E-02	-2.04	CatOcc=PNF	0	0	2.35	1.43E-02	-2.45
Aspect=aAct	51.97	32.67	37.24	3.69E-02	-2.09	vPro=Passif	21.62	5.76	10.85	1.10E-02	-2.54
StNorme=nND	33.33	2.97	5.28	2.66E-02	-2.22	Particip=P06	17.86	3.60	8.21	8.68E-03	-2.62
LVF=LVF_X	16.67	0.99	3.52	3.17E-03	-2.95	vPro=RefCl	10.71	2.16	8.21	3.95E-04	-3.54

PresSe=PresN	45.22	35.15	46.04	1.20E-06	-4.85	LVF=LVF_N	6.45	1.44	9.09	1.13E-05	-4.39
StRsp=RSTn	29.69	9.41	18.77	1.38E-07	-5.27	StRsp=RSTo	33.97	64.03	76.83	4.53E-06	-4.59
vPro=PasPro	0	0	14.37	1.91E-22	-9.75	PresSe=PresO	28.80	38.13	53.96	1.20E-06	-4.85
CatOcc=Ajout	0	0	15.54	1.64E-24	-10.22	CatOcc=AN	0	0	20.82	1.45E-19	-9.05
NormePres=PN_N	0	0	15.54	1.64E-24	-10.22	NormePres=PN_F	0	0	23.17	3.97E-22	-9.67
CatOcc=Omis	0	0	25.22	4.33E-44	-13.93	CatOcc=PN	0	0	36.07	9.75E-39	-13.02

5.4.3 Résultats *catdes* de [CatOcc]

Les variables dont le test du chi-carré donne un résultat significatif avec la variable [CatOcc] sont présentées au tableau 5.14. Les variables grisées sont [NormePres], car la catégorie d'occurrence dépend de la modalité de [NormePres]; [vPro], car des modalités sont incompatibles avec certaines catégories d'occurrences (par exemple < PasPro > et < Ajout >); [PresSe], qui détermine aussi la catégorie d'occurrence; ainsi que [RspNorme], car chaque catégorie d'occurrence ne peut correspondre qu'à une seule modalité de cette variable.

Tableau 5.14 – Résultat *catdes* pour [CatOcc] (variables)

Variables		
	p.value	df
NormePres	5.36E-142	8
vPro	4.81E-76	20
PresSe	1.54E-72	4
RspNorme	1.54E-72	4
LVF	1.14E-15	52
RoleT	1.22E-08	24
StNorme	1.00E-06	20
Temps	2.04E-06	4
Anime	2.05E-06	4
StEmp	3.25E-06	16
StRsp	1.00E-04	8
Nb	2.52E-04	4
Particip	4.34E-04	52
Sujet	4.68E-04	40
Aspect	6.60E-03	16

Le tableau 5.15 présente les modalités caractérisantes des cinq modalités de [CatOcc].

Tableau 5.15 – Résultat *catdes* pour [CatOcc] (modalités)

	Ajout					Omis					
	Cla/Mod	Mod/Cla	Global	p.value	v.test	Cla/Mod	Mod/Cla	Global	p.value	v.test	
NormePres=PN_N	100	100	15.54	1.77E-63	16.82	RspNorme=RN_n	61.87	100	40.76	4.33E-44	13.93
vPro=PasPro	100	92.45	14.37	5.25E-55	15.62	PresSe=PresN	54.78	100	46.04	2.66E-37	12.76
RspNorme=RN_n	38.13	100	40.76	1.64E-24	10.22	NormePres=PN_O	41.15	100	61.29	7.99E-23	9.83
PresSe=PresO	28.80	100	53.96	1.09E-16	8.29	RoleT=agent	34.66	70.93	51.61	3.13E-05	4.16
LVF=LVF_X	66.67	15.09	3.52	6.73E-05	3.99	StRsp=RSTn	45.31	33.72	18.77	9.86E-05	3.89
RoleT=sieg	45	16.98	5.87	1.38E-03	3.20	Nb=pl	40.26	36.05	22.58	9.09E-04	3.32
Temps=TS	20	81.13	63.05	2.36E-03	3.04	Sujet=nous	60	10.47	4.40	4.30E-03	2.86
StNorme=nND	44.44	15.09	5.28	3.01E-03	2.97	Anime=Ani	28	89.53	80.65	1.31E-02	2.48
LVF=LVF_P	27.87	32.08	17.89	6.21E-03	2.74	LVF=LVF_F	50	11.63	5.87	1.55E-02	2.42
LVF=LVF_S	46.15	11.32	3.81	8.92E-03	2.61	Particip=P11	58.33	8.14	3.52	1.59E-02	2.41
Particip=P05	29.55	24.53	12.9	1.15E-02	2.53	vPro=Neutre	32.31	48.84	38.12	1.98E-02	2.33
Sujet=AbsNN	50	9.43	2.93	1.17E-02	2.52	StEmp=eTI	38.46	23.26	15.25	2.21E-02	2.29
Particip=P04	34.62	16.98	7.62	1.28E-02	2.49	Aspect=aAct	32.28	47.67	37.24	2.27E-02	2.28
RoleT=experi	26.79	28.30	16.42	1.74E-02	2.38	Sujet=on	47.62	11.63	6.16	2.39E-02	2.26
RoleT=instru	100	3.77	0.59	2.38E-02	2.26	Particip=P13	42.42	16.28	9.68	2.39E-02	2.26
Aspect=aEtat	23.08	39.62	26.69	2.58E-02	2.23	Temps=TC	31.75	46.51	36.95	3.64E-02	2.09
Nb=sg	17.80	88.68	77.42	2.77E-02	2.20	Sujet=ils elles	54.55	6.98	3.23	4.04E-02	2.05
StRsp=RSTn	25	30.19	18.77	2.83E-02	2.19	VvInf=VinfO	36.21	24.42	17.01	4.13E-02	2.04
Sujet=je	22.73	37.74	25.81	3.76E-02	2.08	Texte=T6	35.38	26.74	19.06	4.23E-02	2.03
Anime=Inani	24.24	30.19	19.35	3.89E-02	2.07	vPro=RefCD	36.36	23.26	16.13	4.51E-02	2.00
LVF=LVF_R	0	0	5.279	4.39E-02	-2.02	Aspect=aAchv	34.21	30.23	22.29	4.64E-02	1.99
Anime=Ani	13.45	69.81	80.65	3.89E-02	-2.07	StNorme=nTI	36.54	22.09	15.25	4.93E-02	1.97
Nb=pl	7.79	11.32	22.58	2.77E-02	-2.2	VvInf=VinfN	22.97	75.58	82.99	4.13E-02	-2.04
StNorme=nTI	5.77	5.66	15.25	2.64E-02	-2.22	Particip=P15	5.56	1.16	5.28	3.88E-02	-2.07
vPro=Passif	2.70	1.89	10.85	1.28E-02	-2.5	Temps=TS	21.40	53.49	63.05	3.64E-02	-2.09
vPro=RefCl	0	0	8.21	7.11E-03	-2.69	RoleT=sieg	5	1.16	5.87	2.28E-02	-2.28
Particip=P06	0	0	8.21	7.11E-03	-2.69	Anime=Inani	13.64	10.47	19.35	1.31E-02	-2.48
LVF=LVF_N	0	0	9.09	4.07E-03	-2.87	Sujet=SN	14.81	13.95	23.75	1.16E-02	-2.53
vPro=RefCD	3.64	3.77	16.13	3.77E-03	-2.9	StNorme=nAtt	8.33	3.49	10.56	9.08E-03	-2.61
Temps=TC	7.94	18.87	36.95	2.36E-03	-3.04	StEmp=eAtt	8.11	3.49	10.85	7.09E-03	-2.69
RoleT=agent	9.09	30.19	51.61	7.04E-04	-3.39	LVF=LVF_N	6.45	2.33	9.09	6.90E-03	-2.70
vPro=AutoEP	0	0	12.32	4.98E-04	-3.48	Nb=sg	20.83	63.95	77.42	9.09E-04	-3.32
NormePres=PN_Fac	0	0	23.17	2.15E-07	-5.19	StRsp=RSTo	20.23	61.63	76.83	2.11E-04	-3.71
vPro=Neutre	0.77	1.89	38.12	2.26E-11	-6.69	Aspect=aEtat	9.89	10.47	26.69	3.53E-05	-4.14
PresSe=PresN	0	0	46.04	1.09E-16	-8.29	vPro=PasPro	0	0	14.37	1.76E-07	-5.22
RspNorme=RN_o	0	0	59.24	1.64E-24	-10.2	NormePres=PN_N	0	0	15.54	4.33E-08	-5.48
NormePres=PN_O	0	0	61.29	5.17E-26	-10.5	NormePres=PN_F	0	0	23.17	2.55E-12	-7.00
						PresSe=PresO	0	0	53.96	2.66E-37	-12.76
						RspNorme=RN_o	0	0	59.24	4.33E-44	-13.93

PN						AN							
	Cl	Mod	Mod/Cl	Global	p.value	v.test		Cl	Mod	Mod/Cl	Global	p.value	v.test
PresSe=PresO	66.85		100	53.96	1.20E-46	14.34	NormePres=PN_F	89.87		100	23.17	8.50E-65	17.00
RspNorme=RN_o	60.89		100	59.24	9.75E-39	13.02	PresSe=PresN	45.22		100	46.04	1.85E-29	11.27
NormePres=PN_O	58.85		100	61.29	5.67E-36	12.52	RspNorme=RN_o	35.15		100	59.24	1.45E-19	9.05
LVF=LVF_N	87.1	21.95		9.09	1.20E-09	6.08	vPro=Passif	62.16	32.39		10.85	5.94E-09	5.82
StNorme=nAtt	75	21.95		10.56	6.50E-07	4.98	vPro=RefCl	64.29	25.35		8.21	2.12E-07	5.19
vPro=Neutre	52.31	55.28		38.12	1.22E-06	4.85	Temps=TC	32.54	57.75		36.95	6.70E-05	3.99
StEmp=eAtt	72.97	21.95		10.85	1.71E-06	4.79	StNorme=nInt	30.32	66.20		45.45	8.89E-05	3.92
Anime=Ani	40.36	90.24		80.65	5.33E-04	3.46	Anime=Inani	37.88	35.21		19.35	3.49E-04	3.58
Temps=TS	42.79	74.80		63.05	6.79E-04	3.40	LVF=LVF_C	50	18.31		7.62	6.40E-04	3.41
vPro=AutoEP	59.52	20.33		12.32	1.08E-03	3.27	RoleT=them	40.48	23.94		12.32	1.94E-03	3.10
Nb=sg	40.15	86.18		77.42	3.17E-03	2.95	StEmp=eInt	27.88	64.79		48.39	1.98E-03	3.09
Particip=P15	66.67	9.76		5.28	8.23E-03	2.64	LVF=LVF_U	40	16.90		8.80	1.26E-02	2.50
StRsp=RSTo	39.69	84.55		76.83	1.04E-02	2.56	Particip=P10	40.74	15.49		7.92	1.48E-02	2.44
RoleT=pati	53.49	18.70		12.61	1.35E-02	2.47	StRsp=RSTo	23.66	87.32		76.83	1.55E-02	2.42
LVF=LVF_E	57.69	12.20		7.62	2.16E-02	2.30	LVF=LVF_T	66.67	5.63		1.76	2.04E-02	2.32
Particip=P01	70	5.69		2.93	3.43E-02	2.12	Particip=P06	39.29	15.49		8.21	2.05E-02	2.32
Sujet=ProRel	57.14	9.76		6.16	4.67E-02	1.99	LVF=LVF_R	44.44	11.27		5.28	2.22E-02	2.29
Particip=P10	18.52	4.07		7.92	4.53E-02	-2.00	Sujet=SN	29.63	33.80		23.75	3.06E-02	2.16
Particip=P11	8.33	0.81		3.52	3.80E-02	-2.07	RoleT=sieg	40	11.27		5.87	4.60E-02	2.00
RoleT=them	21.43	7.32		12.32	3.28E-02	-2.13	StNorme=nBiTr	0	0		3.81	4.52E-02	-2.00
RoleT=sieg	10	1.63		5.87	9.19E-03	-2.60	Texte=T3	10.20	7.04		14.37	4.14E-02	-2.04
LVF=LVF_U	13.33	3.25		8.80	4.83E-03	-2.82	RoleT=experi	10.71	8.45		16.42	3.59E-02	-2.10
LVF=LVF_X	0	0		3.52	4.16E-03	-2.87	LVF=LVF_N	6.45	2.82		9.09	2.95E-02	-2.18
Nb=pl	22.08	13.82		22.58	3.17E-03	-2.95	StRsp=RSTn	10.94	9.86		18.77	2.61E-02	-2.23
StRsp=RSTn	18.75	9.76		18.77	1.02E-03	-3.28	Sujet=je	12.5	15.49		25.81	2.28E-02	-2.28
Sujet=nous	0	0		4.40	1.02E-03	-3.29	vPro=AutoEP	7.14	4.23		12.32	1.36E-02	-2.47
StNorme=nInt	26.45	33.33		45.45	7.32E-04	-3.38	StEmp=eTI	7.69	5.63		15.25	7.40E-03	-2.68
Temps=TC	24.60	25.20		36.95	6.79E-04	-3.40	vPro=Neutre	13.08	23.94		38.12	5.14E-03	-2.80
Anime=Inani	18.18	9.76		19.35	5.33E-04	-3.46	LVF=LVF_E	0	0		7.62	1.77E-03	-3.13
StEmp=eInt	26.67	35.77		48.39	4.72E-04	-3.50	StNorme=nAtt	2.78	1.41		10.56	1.70E-03	-3.14
vPro=Passif	5.41	1.63		10.85	6.76E-06	-4.50	StEmp=eAtt	2.70	1.41		10.85	1.34E-03	-3.21
vPro=PasPro	0	0		14.37	3.27E-11	-6.63	LVF=LVF_P	6.56	5.63		17.89	1.17E-03	-3.25
NormePres=PN_N	0	0		15.54	3.62E-12	-6.95	Anime=Ani	16.73	64.79		80.65	3.49E-04	-3.58
NormePres=PN_F	0	0		23.17	7.42E-19	-8.87	Temps=TS	13.95	42.25		63.05	6.70E-05	-3.99
RspNorme=RN_n	0	0		40.76	9.75E-39	-13.02	vPro=PasPro	0	0		14.37	3.89E-06	-4.62
PresSe=PresN	0	0		46.04	1.20E-46	-14.34	NormePres=PN_N	0	0		15.54	1.27E-06	-4.84
							RspNorme=RN_n	0	0		40.76	1.45E-19	-9.05
							PresSe=PresO	0	0		53.96	1.85E-29	-11.27
							NormePres=PN_O	0	0		61.29	8.44E-37	-12.67

PNF					
	Cla/Mod	Mod/Cla	Global	p.value	v.test
NormePres=PN_F	10.13	100	23.17	6.25E-06	4.52
LVF=LVF_U	13.33	50	8.80	2.86E-03	2.98
vPro=Passif	10.81	50	10.85	6.54E-03	2.72
PresSe=PresO	4.35	100	53.96	6.69E-03	2.71
RoleT=them	9.52	50	12.32	1.07E-02	2.55
RspNorme=RN_o	3.96	100	59.24	1.43E-02	2.45
RspNorme=RN_n	0	0	40.76	1.43E-02	-2.45
PresSe=PresN	0	0	46.04	6.69E-03	-2.71
NormePres=PN_O	0	0	61.29	4.41E-04	-3.51

5.4.4 Résultats *catdes* de [NormePres]

Afin de bien interpréter les tableaux de *catdes* pour les variables d'observation, il est important de prendre en considération la distribution de [NormePres] entre les modalités de ces variables. En effet, [NormePres] permet en quelque sorte d'observer les opportunités qu'ont eu les participants de produire des occurrences d'une modalité donnée d'une variable d'observation. Par exemple, si toutes les occurrences ayant le pronom « vous » comme sujet correspondaient à des présences facultatives du SE, alors il est normal qu'elles correspondent toutes à un respect de la norme : l'effet provient de la distribution de [NormePres] plutôt que du type de sujet.

Tableau 5.16 – Résultat *catdes* pour [NormePres] (variables)

Variables		
	p.value	df
CatOcc	5.36E-142	8
vPro	8.39E-82	10
RspNorme	2.94E-29	2
PresSe	2.99E-24	2
LVF	5.46E-15	26
RoleT	1.40E-09	12
Anime	2.02E-07	2
StNorme	9.63E-06	10
Temps	1.83E-05	2
StEmp	5.54E-04	8
Particip	1.66E-03	26
Sujet	2.25E-03	20
StRsp	2.84E-02	4

Tableau 5.17 – Résultat *catdes* pour [NormePres] (modalités)

PN_N (se interdit)						PN_O (se obligatoire)					
	Cla/Mod	Mod/Cla	Global	p.value	v.test		Cla/Mod	Mod/Cla	Global	p.value	v.test
CatOcc=Ajout	100	100	15.54	1.77E-63	16.82	CatOcc=PN	100	58.85	36.07	5.67E-36	12.52
vPro=PasPro	100	92.45	14.37	5.25E-55	15.62	CatOcc=Omisi	100	41.15	25.22	7.99E-23	9.83
RspNorme=RN_n	38.13	100	40.76	1.64E-24	10.22	vPro=Neutre	84.62	52.63	38.12	8.99E-13	7.15
PresSe=PresO	28.80	100	53.96	1.09E-16	8.29	Anime=Ani	68.36	89.95	80.65	7.79E-08	5.37
LVF=LVF_X	66.67	15.09	3.52	6.73E-05	3.99	vPro=AutoEP	92.86	18.66	12.32	1.27E-06	4.84
RoleT=sieg	45	16.98	5.87	1.38E-03	3.20	LVF=LVF_N	93.55	13.88	9.09	2.99E-05	4.17
Temps=TS	20	81.13	63.05	2.36E-03	3.04	RoleT=agent	70.45	59.33	51.61	3.49E-04	3.58
StNorme=nND	44.44	15.09	5.28	3.01E-03	2.97	StEmp=eTI	80.77	20.10	15.25	1.33E-03	3.21
LVF=LVF_P	27.87	32.08	17.89	6.21E-03	2.74	LVF=LVF_E	88.46	11.00	7.62	2.01E-03	3.09
LVF=LVF_S	46.15	11.32	3.81	8.92E-03	2.61	StNorme=nAtt	83.33	14.35	10.56	3.14E-03	2.95
Particip=P05	29.55	24.53	12.90	1.15E-02	2.53	StNorme=nTI	78.85	19.62	15.25	4.04E-03	2.88
Sujet=AbsNN	50	9.43	2.93	1.17E-02	2.52	StEmp=eAtt	81.08	14.35	10.85	7.53E-03	2.67
Particip=P04	34.62	16.98	7.62	1.28E-02	2.49	Texte=T3	77.55	18.18	14.37	1.05E-02	2.56
RoleT=experi	26.79	28.30	16.42	1.74E-02	2.38	vPro=RefCD	76.36	20.10	16.13	1.12E-02	2.54
RoleT=instru	100	3.77	0.59	2.38E-02	2.26	PresSe=PresO	66.85	58.85	53.96	2.34E-02	2.27
Aspect=aEtat	23.08	39.62	26.69	2.58E-02	2.23	Particip=P13	78.79	12.44	9.68	2.82E-02	2.19
Nb=sg	17.80	88.68	77.42	2.77E-02	2.20	Particip=P04	42.31	5.26	7.62	4.53E-02	-2.00
StRsp=RSTn	25	30.19	18.77	2.83E-02	2.19	Aspect=aEtat	51.65	22.49	26.69	2.95E-02	-2.18
Sujet=je	22.73	37.74	25.81	3.76E-02	2.08	PresSe=PresN	54.78	41.15	46.04	2.34E-02	-2.27
Anime=Inani	24.24	30.19	19.35	3.89E-02	2.07	LVF=LVF_C	38.46	4.78	7.62	1.62E-02	-2.40
LVF=LVF_R	0	0	5.28	4.39E-02	-2.02	LVF=LVF_U	40	5.74	8.80	1.51E-02	-2.43
Anime=Ani	13.45	69.81	80.65	3.89E-02	-2.07	StEmp=eInt	53.33	42.11	48.39	3.64E-03	-2.91
Nb=pl	7.79	11.32	22.58	2.77E-02	-2.20	Sujet=SN	46.91	18.18	23.75	2.77E-03	-2.99
StNorme=nTI	5.77	5.66	15.25	2.64E-02	-2.22	LVF=LVF_X	16.67	0.96	3.52	1.96E-03	-3.10
vPro=Passif	2.70	1.89	10.85	1.28E-02	-2.49	vPro=RefCI	32.14	4.31	8.21	1.37E-03	-3.20
vPro=RefCI	0	0	8.21	7.11E-03	-2.69	RoleT=them	38.10	7.66	12.32	1.33E-03	-3.21
Particip=P06	0	0	8.21	7.11E-03	-2.69	CatOcc=PNF	0	0	2.35	4.41E-04	-3.51
LVF=LVF_N	0	0	9.09	4.07E-03	-2.87	StNorme=nInt	50.97	37.80	45.45	3.79E-04	-3.55
vPro=RefCD	3.64	3.77	16.13	3.77E-03	-2.90	RoleT=sieg	15	1.44	5.87	1.83E-05	-4.28
Temps=TC	7.94	18.87	36.95	2.36E-03	-3.04	vPro=Passif	24.32	4.31	10.85	1.74E-06	-4.78
RoleT=agent	9.09	30.19	51.61	7.04E-04	-3.39	Anime=Inani	31.82	10.05	19.35	7.79E-08	-5.37
vPro=AutoEP	0	0	12.32	4.98E-04	-3.48	vPro=PasPro	0	0	14.37	8.35E-24	-10.06
CatOcc=AN	0	0	20.82	1.27E-06	-4.84	CatOcc=Ajout	0	0	15.54	5.17E-26	-10.55
CatOcc=Omisi	0	0	25.22	4.33E-08	-5.48	CatOcc=AN	0	0	20.82	8.44E-37	-12.67
vPro=Neutre	0.77	1.89	38.12	2.26E-11	-6.69						
CatOcc=PN	0	0	36.07	3.62E-12	-6.95						
PresSe=PresN	0	0	46.04	1.09E-16	-8.29						
RspNorme=RN_o	0	0	59.24	1.64E-24	-10.22						

PN_Facultatif										
	Cl/Mod	Mod/Cl	Global	p.value	v.test					
						Particip=P07	5.26	1.27	5.57	4.66E-02 -1.99
CatOcc=AN	100	89.87	20.82	8.50E-65	17.00	Texte=T3	10.20	6.33	14.37	1.54E-02 -2.42
RspNorme=RN_o	39.11	100	59.24	3.97E-22	9.67	LVF=LVF_N	6.45	2.53	9.09	1.39E-02 -2.46
PresSe=PresN	45.22	89.87	46.04	2.38E-20	9.24	RoleT=experi	10.71	7.59	16.42	1.20E-02 -2.51
vPro=Passif	72.97	34.18	10.85	8.43E-12	6.83	StRsp=RSTn	10.94	8.86	18.77	7.45E-03 -2.68
vPro=RefCI	67.86	24.05	8.21	1.75E-07	5.22	vPro=AutoEP	7.14	3.80	12.32	5.02E-03 -2.81
CatOcc=PNF	100	10.13	2.35	6.25E-06	4.52	vPro=Neutre	14.62	24.05	38.12	2.94E-03 -2.97
Anime=Inani	43.94	36.71	19.35	2.74E-05	4.19	StEmp=eTI	7.69	5.06	15.25	2.19E-03 -3.06
StNorme=nInt	33.55	65.82	45.45	3.73E-05	4.12	LVF=LVF_P	8.20	6.33	17.89	1.15E-03 -3.25
Temps=TC	35.71	56.96	36.95	3.86E-05	4.12	LVF=LVF_E	0	0	7.62	7.79E-04 -3.36
RoleT=them	50	26.58	12.32	5.04E-05	4.05	StNorme=nAtt	2.78	1.27	10.56	6.01E-04 -3.43
LVF=LVF_U	53.33	20.25	8.80	1.93E-04	3.73	StEmp=eAtt	2.70	1.27	10.85	4.58E-04 -3.50
StEmp=eInt	30.91	64.56	48.39	1.08E-03	3.27	Temps=TS	15.81	43.04	63.05	3.86E-05 -4.12
LVF=LVF_C	50	16.46	7.62	2.09E-03	3.08	Anime=Ani	18.18	63.29	80.65	2.74E-05 -4.19
Sujet=SN	34.57	35.44	23.75	7.23E-03	2.69	vPro=PasPro	0	0	14.37	7.65E-07 -4.94
StRsp=RSTo	26.34	87.34	76.83	9.37E-03	2.60	CatOcc=Ajout	0	0	15.54	2.15E-07 -5.19
LVF=LVF_R	50	11.39	5.28	1.20E-02	2.51	CatOcc=Omis	0	0	25.22	2.55E-12 -7.00
LVF=LVF_T	66.67	5.06	1.76	3.05E-02	2.16	CatOcc=PN	0	0	36.07	7.42E-19 -8.87
Particip=P10	40.74	13.92	7.92	3.52E-02	2.11	PresSe=PresO	4.35	10.13	53.96	2.38E-20 -9.24
Particip=P06	39.29	13.92	8.21	4.77E-02	1.98	RspNorme=RN_n	0	0	40.76	3.97E-22 -9.67

5.4.5 Interaction de [NormePres] avec les résultats

Lorsqu'une variable ou une modalité est significativement corrélée avec une modalité d'une variable d'observation, mais aussi avec la modalité associée de [NormePres], il est difficile de déterminer la source de l'influence. Par exemple, le tableau 5.13 indique que 89% des occurrences de verbe pronominal réfléchi CI respectent la norme, mais le tableau 5.17 nous apprend que pour 68% de ce type de verbe pronominal, la présence du SE est facultative. Dans cette situation, il est ardu de déterminer si la corrélation entre < RefCI > et < RN_o > est simplement le résultat de la distribution de [NormePres] ou non.

Néanmoins, certaines variables et modalités ne sont pas corrélées avec [NormePres]. Cette section présente les résultats de *catdes* sur lesquels la distribution de [NormePres] n'apporte aucune piste d'explication.

Si l'on soustrait les variables ayant une interaction significative avec [NormePres] (voir tableau 5.16) de celles en ayant une avec [PresSe] (voir tableau 5.10), on obtient trois variables : [Nb],

[VvInf] et [Aspect]. L'interaction de ces trois variables avec [PresSe] ne peut provenir de la distribution de [NormePres].

Tableau 5.18 – Résultat *catdes* pour [PresSe] (variables, considérant [NormePres])

Variables		
	p.value	df
Nb	5.12E-06	1
VvInf	7.18E-03	1
Aspect	4.81E-02	4

Lorsqu'on se penche sur le croisement des modalités de [NormePres] et de [PresSe], on remarque que < PN_N > correspond toujours à < PresO >, tandis que < PN_Fac > correspond à < PresN > dans 90% des cas. Ainsi, le fait qu'une modalité est caractéristique de < PresO > peut s'expliquer par sa corrélation avec < PN_N >, et il en est de même pour < PresN > et < PN_Fac >. Si l'on reprend le tableau *catdes* décrivant [PresSe] avec des modalités (voir tableau 5.11) et que l'on retire les modalités caractéristiques de < PN_N > de la liste des modalités corrélées positivement avec < PresO > ainsi que celles caractéristiques de < PN_Fac > de la liste des modalités corrélées positivement avec < PresN >, on obtient le tableau suivant :

Tableau 5.19 – Résultat *catdes* pour [PresSe] (modalités, considérant [NormePres])

PresO (présence)						PresN (absence)					
	Cl/Mod	Mod/Cl	Global	p.value	v.test		Cl/Mod	Mod/Cl	Global	p.value	v.test
Particip=P15	88.89	8.70	5.28	1.70E-03	3.14	<i>Nb=pl</i>	68.83	33.76	22.58	5.47E-06	4.55
StEmp=eBiTr	100	5.43	2.93	1.86E-03	3.11	Sujet=nous	100	9.55	4.40	6.07E-06	4.52
VvInf=VinfN	57.24	88.04	82.99	7.77E-03	2.66	Particip=P11	91.67	7.01	3.52	1.23E-03	3.23
NormePres=PN_O	58.85	66.85	61.29	2.34E-02	2.27	<i>RoleT=agent</i>	53.41	59.87	51.61	4.98E-03	2.81
RoleT=pati	69.77	16.30	12.61	2.64E-02	2.22	VvInf=VinfO	62.07	22.93	17.01	7.77E-03	2.66
Sujet=ProRel	76.19	8.70	6.16	3.58E-02	2.10						

L'entièreté des variables corrélées avec [RspNorme] sont aussi corrélées avec [NormePres]. Si l'on soustrait les variables corrélées positivement avec < PN_Fac > de celles corrélées positivement à < RN_o >, et que l'on fait de même pour < PN_N > et < RN_n >, on obtient le tableau suivant :

Tableau 5.20 – Résultat *catdes* pour [RspNorme] (modalités, considérant [NormePres])

	RN_o					RN_n				
	Cl/Mod	Mod/Cla	Global	p.value	v.test	Cl/Mod	Mod/Cla	Global	p.value	v.test
<i>LVF=LVF_N</i>	93.55	14.36	9.09	1.13E-05	4.39	Aspect=aAct				
<i>StNorme=nAtt</i>	77.78	13.86	10.56	1.55E-02	2.42	48.03	43.88	37.24	3.69E-02	2.09
<i>vPro=Neutre</i>	66.92	43.07	38.12	2.37E-02	2.26					
<i>StEmp=eAtt</i>	75.68	13.86	10.85	3.02E-02	2.17					
<i>Particip=P01</i>	90	4.46	2.93	4.46E-02	2.01					

Quatre des variables corrélées avec [CatOcc] ne présentent pas de corrélation significative avec [NormePres] :

Tableau 5.21 – Résultat *catdes* pour [CatOcc] (variables, considérant [NormePres])

Variables		
	p.value	df
Nb	2.52E-04	4
Particip	4.34E-04	52
Sujet	4.68E-04	40
Aspect	6.60E-03	16

Les modalités de [CatOcc] ne peuvent correspondre qu'à une seule modalité de [NormePres] : les présences normées facultatives (< PNF >) et les absences normées (< AN >) correspondent à < PN_Fac >; les présences normées (< PN >) et les omissions (< Omis >) correspondent à < PN_O >; et les ajouts (< Ajout >) correspondent à < PN_N >.

La liste des modalités caractérisantes de < Ajout > est la même que celle de < PN_N >. Pour les autres catégories d'occurrences, si l'on retire de la liste des modalités caractérisantes celles étant corrélées avec la modalité [NormePres] associée, on obtient le tableau 5.22. À noter que la catégorie des présences normées facultatives n'y figure pas, car toutes ses modalités caractérisantes sont corrélées avec < PN_Fac >.

Tableau 5.22 – Résultat *catdes* pour [CatOcc] (modalités, considérant [NormePres])

Omis					
	Cla/Mod	Mod/Cla	Global	p.value	v.test
StRsp=RSTn	45.31	33.72	18.77	9.86E-05	3.89
Nb=pl	40.26	36.05	22.58	9.09E-04	3.32
Sujet=nous	60	10.47	4.40	4.30E-03	2.86
LVF=LVF_F	50	11.63	5.87	1.55E-02	2.42
Particip=P11	58.33	8.14	3.52	1.59E-02	2.41
Aspect=aAct	32.28	47.67	37.24	2.27E-02	2.28
Sujet=on	47.62	11.63	6.16	2.39E-02	2.26
Temps=TC	31.75	46.51	36.95	3.64E-02	2.09
Sujet=ils elles	54.55	6.98	3.23	4.04E-02	2.05
VvInf=VinfO	36.21	24.42	17.01	4.13E-02	2.04
Texte=T6	35.38	26.74	19.06	4.23E-02	2.03
Aspect=aAchv	34.21	30.23	22.29	4.64E-02	1.99
VvInf=VinfN	22.97	75.58	82.99	4.13E-02	-2.04
Particip=P15	5.56	1.16	5.28	3.88E-02	-2.07
Temps=TS	21.40	53.49	63.05	3.64E-02	-2.09
StNorme=nAtt	8.33	3.49	10.56	9.08E-03	-2.61
StEmp=eAtt	8.11	3.49	10.85	7.09E-03	-2.69
Nb=sg	20.83	63.95	77.42	9.09E-04	-3.32
StRsp=RSTo	20.23	61.63	76.83	2.11E-04	-3.71

PN					
	Cla/Mod	Mod/Cla	Global	p.value	v.test
Temps=TS	42.79	74.80	63.05	6.79E-04	3.40
Nb=sg	40.15	86.18	77.42	3.17E-03	2.95
Particip=P15	66.67	9.76	5.28	8.23E-03	2.64
StRsp=RSTo	39.69	84.55	76.83	1.04E-02	2.56
RoleT=pati	53.49	18.70	12.61	1.35E-02	2.47
Particip=P01	70	5.69	2.93	3.43E-02	2.12
Sujet=ProRel	57.14	9.76	6.16	6.67E-02	1.99
Particip=P10	18.52	4.07	7.92	4.53E-02	-2.00
Particip=P11	8.33	0.81	3.52	3.80E-02	-2.07
Nb=pl	22.08	13.82	22.58	3.17E-03	-2.95
StRsp=RSTn	18.75	9.76	18.77	1.02E-03	-3.28
Sujet=nous	0	0	4.40	1.02E-03	-3.29
Temps=TC	24.60	25.20	36.95	6.79E-04	-3.40

AN					
	Cla/Mod	Mod/Cla	Global	p.value	v.test
RoleT=sieg	40	11.27	5.87	4.60E-02	2.00
StNorme=nBiTr	0	0	3.81	4.52E-02	-2.00
Texte=T3	10.20	7.04	14.37	4.14E-02	-2.04
Sujet=je	12.5	15.49	25.81	2.28E-02	-2.28

Dans la liste des modalités caractérisantes des occurrences d'omission, < nAtt > et < eAtt > sont en gras car elles sont corrélées positivement avec < PN_O >, ce qui devrait donner lieu à une corrélation positive avec < Omis >; or c'est une corrélation négative que l'on observe.

CHAPITRE 6 : DISCUSSION

Ce chapitre présente une interprétation et une discussion des résultats exposés au chapitre 5. Tout d'abord, la représentativité des analyses est discutée (6.1), puis sont présentées des tendances générales qui se dégagent des résultats (6.2). La section suivante regroupe les résultats par variables contextuelles pour discuter de leurs effets (6.3). Ensuite sont présentés les résultats ayant trait à la première question de recherche (6.4) et à la deuxième question de recherche (6.5). La dernière section recense quelques hypothèses diverses (6.6).

6.1 Représentativité des analyses

Les tests du chi-carré indiquent que le ratio de présences et d'absences du SE ainsi que le taux de correspondance à la norme ne sont pas corrélés de manière statistiquement significative avec les différents sujets de rédaction. De plus, les participant·e·s ne démontrent pas de tendance individuelle à faire plus ou moins d'écarts à la norme, la proportion restant similaire d'une personne à l'autre. Néanmoins, il y a une corrélation statistiquement significative entre les participant·e·s et la présence ou l'absence du SE; plus précisément, P15 emploie plus fréquemment des verbes pronominaux tandis que les textes de P11 en comportent moins⁴⁰.

L'ACM représente dans un graphique la variation générale du jeu de données. La majorité des occurrences se retrouve dans un nuage de points près de l'origine, et deux regroupements s'en distinguent: l'un sur le côté négatif de la première dimension, l'autre sur le côté négatif de la deuxième dimension⁴¹. Si l'on observe les modalités étant les mieux représentées dans ce graphique, on remarque que la première dimension éloigne les occurrences ayant une structure attributive sur son côté négatif, en plus de diviser les sujets animés entre ceux ayant le rôle thématique patient (sur le côté négatif de l'axe) et ceux ayant le rôle agent (sur le côté positif de l'axe). Le regroupement sur le côté négatif, qui se démarque du reste du corpus, rassemble les

⁴⁰ Le z-test effectué sur [Particip] et [PresSe] indique que P10 aussi produit significativement moins de verbes pronominaux, mais cela est expliqué par la sous-représentation de P10 dans les occurrences où le SE est obligatoire.

⁴¹ Rappelons ici que, puisque les dimensions sont construites de manière orthogonale, les phénomènes décrits par une dimension donnée sont en totale indépendance de ceux portés par les autres dimensions (voir 4.4.2).

occurrences ayant une structure attributive, ce qui inclut la subordonnée fréquente « qui s'appelle X » (sujet pronom relatif, rôle patient, classe lexico-sémantique N). La fréquence élevée de cette subordonnée explique partiellement ce regroupement, mais pas complètement : les occurrences à structure attributive se distinguent du reste du corpus et expliquent une grande partie de la variation du jeu de données. Quant à la deuxième dimension, elle est surtout divisée entre les sujets animés et les sujets inanimés, ces derniers étant éloignés du nuage de points sur le côté négatif. Le côté négatif regroupe aussi les modalités de rôle thématique thème et siège, ainsi que les verbes pronominaux passifs, bien que ceux-ci n'aient pas participé à la définition des modalités. Ceci est possiblement le reflet du grand nombre d'occurrences mentionnant des portes de métro qui s'ouvrent ou qui se ferment : le premier texte consiste en la narration d'une anecdote qui se déroule dans le métro.

Bref, l'ACM fait une analyse générale du jeu de données et modélise l'entière de la variation. Or, ces résultats ne permettent pas d'atteindre l'objectif de ce mémoire ni de formuler de réponse aux questions spécifiques de recherche. Le codage de données utilisé a l'avantage d'avoir des catégories précises, mais présente aussi une limite importante : certaines combinaisons de catégories ne sont pas possibles, par exemple un verbe pronominal réfléchi avec une structure intransitive. Avoir des combinaisons pour lesquelles il n'y a aucune occurrence influence les tests statistiques (puisque les autres combinaisons auront un nombre plus élevé qu'attendu). Ceci, en plus de la surreprésentation de certaines phrases fréquentes, a pour effet de « tirer la couverture » de la distribution : la majorité de la variation du jeu de données est expliquée par ces combinaisons, et cela vient cacher d'autres interactions intéressantes. L'emploi des verbes pronominaux n'est donc pas bien représenté par l'ACM.

Par contre, aucune analyse statistique ne pourra faire une description assez complète des verbes pronominaux du corpus pour en faire un modèle prédictif. Non seulement la variation du corpus dépasse l'emploi des verbes pronominaux, mais aussi les variables contextuelles ne permettent pas une explication complète de la variation des variables de respect de la norme et de la présence du pronom SE. En effet, les études sur le bilinguisme démontrent que les productions en langue seconde dépendent de décisions basées sur trois types de bagage linguistique : la langue maternelle, la langue cible et l'interlangue (voir 3.1.1). Le jeu de données utilisé ne comporte que des variables

observant l'influence de la langue cible; il y a fort à parier que cela explique la difficulté à modéliser l'emploi des verbes pronominaux à partir du jeu de données.

Néanmoins, l'AFM (et les *catdes* qui en découlent) réussit à faire une description de relativement bonne qualité de la variation des variables d'observation. Comme on le voit au tableau 5.9 (voir 5.3), les deux premières dimensions de l'AFM décrivent au moins 75% de la variation de six variables d'observations. Les occurrences d'absence normée ont une moins bonne qualité de représentation et celle des occurrences de présence normée facultative est pire; cela s'explique entre autres par la plus faible fréquence d'occurrences pour lesquelles la présence du SE est facultative. Les occurrences d'ajout sont celles ayant la pire qualité de représentation, ce qui est une conséquence attendue du choix de ne pas retenir toutes les occurrences verbales du corpus Groupe LSQ 98 qui ne sont pas pronominales et pour lesquelles l'ajout du SE ne serait pas normé (c'est-à-dire la majorité des verbes). Ainsi, toutes les occurrences où le SE est interdit selon la norme correspondent à des ajouts : la modalité <PN_N> (SE interdit) de [NormePres] correspond parfaitement à la modalité <Ajout> de [CatOcc].

6.2 Tendances générales

La majorité des verbes du corpus Groupe LSQ 98 sont des verbes non pronominaux ne pouvant pas être précédés du SE dans le contexte. Ceci est représentatif de la distribution naturelle du français : la majorité des verbes d'une production écrite ou orale ne seront pas pronominaux. Étant donné que ces occurrences n'ont pas été retenues pour l'analyse, il n'est pas possible d'établir un taux général de correspondance à la norme pour l'emploi des verbes pronominaux dans le corpus.

Néanmoins, lorsque le contexte rend obligatoire la présence du SE, le pronom est plus souvent produit (63%) qu'absent (37%). D'ailleurs, lorsque le SE est présent, il est normé dans la majorité des occurrences (71%). Lorsqu'on se penche sur les occurrences d'écarts à la norme, on remarque que la majorité représente des omissions (62%) plutôt que des ajouts (38%). Cette observation est compatible avec ce qui ressort de la littérature sur le français sourd, c'est-à-dire que les pronoms personnels sont moins produits qu'en français langue maternelle. Ainsi, le SE est peu réalisé lorsque sa présence est facultative (10%).

6.3 Effet des variables contextuelles

L'analyse générale (l'ACM) donne de l'information sur la variation générale du jeu de données, sans bien représenter l'emploi des verbes pronominaux spécifiquement. Les résultats principaux de l'ACM ayant déjà été discutés, cette section se concentrera plutôt sur les résultats de l'analyse orientée sur les variables d'observation (l'AFM) ainsi que sur les tests *catdes*.

6.3.1 Effet du type de verbe pronominal

L'ACM, qui touche l'entièreté de la variation du corpus, éloigne les verbes pronominaux passifs (liés aux sujets ayant le trait inanimé) du reste des types de verbes pronominaux. Les verbes pronominaux réfléchis s'éloignent du regroupement des verbes pronominaux passifs et des sujets inanimés, tandis que le reste des types de verbes pronominaux se trouve dans le nuage de points près de l'origine. L'éloignement des réfléchis des sujets inanimés est logique, étant donné que la capacité d'entamer un procès verbal réfléchi ou réciproque est normalement limitée aux sujets animés.

L'AFM, qui se concentre davantage sur l'emploi des verbes pronominaux, indique aussi que les verbes pronominaux passifs se démarquent beaucoup dans la variation du jeu de données dans leur association aux sujets inanimés. Néanmoins, cette fois-ci, ce sont les verbes ne pouvant pas être pronominaux qui sont le plus en opposition aux verbes pronominaux passifs. Les quatre autres types sont encore une fois dans le nuage de points; les réfléchis CD s'approchent quelque peu des écarts à la norme, les réfléchis CI des absences de SE, les pronominaux neutres ainsi que les autonomes et essentiellement pronominaux des présences de SE. Aucun type de verbe pronominal n'est en proximité du point de respect de la norme.

Néanmoins, lorsqu'on regarde les *catdes* sans interaction possible avec la variable [NormePres], le seul résultat statistiquement significatif touchant la variable des types de verbes pronominaux est la caractérisation des occurrences de respect de la norme par les verbes pronominaux neutres. Par contre, il faut prendre ce résultat avec un grain de sel, étant donné qu'il y a une modalité de la variable de types de verbes pronominaux dont toutes les occurrences sont nécessairement non normées, soit la modalité des verbes ne pouvant pas prendre une forme pronominale en contexte.

L'impossibilité de la combinaison [vPro]<PasPro> + [RspNorme]<oui> contribue à la fréquence plus haute que celle attendue de la combinaison [vPro]<Neutre> + [RspNorme]<oui>.

Les types de verbes pronominaux dont la présence est le plus souvent obligatoire dans le corpus sont les réfléchis CD, les neutres ainsi que les autonomes et les essentiellement pronominaux. On s'attend donc à ce que ces trois modalités de [vPro] soient surreprésentées dans les catégories d'occurrences de présence normée et d'omission. C'est le cas pour les neutres, mais les autonomes et essentiellement pronominaux ne caractérisent que les présences normées (et sont donc moins souvent omis), tandis que les réfléchis CD ne caractérisent que les omissions.

6.3.2 Effet du temps de conjugaison

L'ACM n'apporte pas d'information de qualité sur la variable de temps de conjugaison, sinon que la subordonnée fréquente « qui s'appelle » est surtout utilisée à un temps simple. L'AFM indique que les occurrences de verbe à un temps composé se rapprochent des absences de SE, tandis que celles à un temps simple se rapprochent des présences de SE.

Les *catdes* indiquent que la variable du temps de conjugaison est corrélée à la variable de présence du SE et à la catégorie d'occurrence, mais pas au respect de la norme. D'ailleurs, la modalité de temps composé caractérise les occurrences où la présence du SE est facultative; on aurait pu s'attendre à ce qu'elle caractérise aussi les occurrences de respect de la norme (puisque les occurrences de SE facultatif sont toujours normées), mais ce n'est pas le cas. Ce qui ressort le plus des *catdes* est que la modalité de temps composé est liée à l'absence du SE, et plus spécifiquement aux omissions; tandis que le temps simple est lié à la présence du SE, et plus spécifiquement à sa présence normée. À partir de cette observation, on peut formuler l'hypothèse que l'emploi de temps de verbes composés demande un plus grand effort cognitif, ce qui fait en sorte que la forme pronominale – qui représente une difficulté supplémentaire – sera moins souvent employée, même lorsque sa présence est obligatoire. Une autre hypothèse, que partage une évaluatrice, est que le SE et les auxiliaires de conjugaison seraient proches cognitivement, notamment à cause de leur position préverbale et, dans le cas de l'auxiliaire ÊTRE, de la similarité de leurs graphies. En d'autres mots, le SE et les auxiliaires de conjugaison pourraient parfois être amalgamés, donc avoir une distribution presque complémentaire.

Les modalités de la variable de présence de la construction *verbe + verbe à l'infinitif* ne sont pas bien représentées dans l'ACM; on en déduit simplement que cette construction est moins souvent employée dans la subordonnée fréquente « qui s'appelle ». Dans l'AFM, la modalité d'absence de construction avec un verbe infinitif est proche de l'origine, mais s'approche légèrement de respect de la norme; tandis que la présence de ce type de construction s'approche du point des écarts à la norme.

Les *catdes* indiquent que, même considérant la distribution de [NormePres], la variable de construction *verbe + verbe à l'infinitif* est liée à celle de présence du SE : lorsque cette construction est présente, le SE est moins souvent réalisé, tandis que lorsqu'elle n'est pas employée, le SE est davantage utilisé. De plus, il y a moins d'occurrences d'omission lorsqu'il n'y a pas la construction *verbe + verbe à l'infinitif*, tandis que sa présence entraîne un plus grand nombre d'omissions du SE obligatoire.

Ces observations sont compatibles avec celles de la variable du temps de conjugaison (temps simple ou composé). Les constructions *verbe + verbe à l'infinitif*, similairement à l'emploi du temps composé, pourraient demander plus d'effort cognitif, entraînant une baisse de fréquence d'utilisation des verbes pronominaux.

6.3.3 Effet de la classe aspectuelle

Dans l'ACM, seule la classe aspectuelle d'état a une bonne représentation : son point se situe loin de l'origine, près du regroupement correspondant à la subordonnée fréquente « qui s'appelle ». Les autres classes aspectuelles se situent près de l'origine, mis à part la modalité de classe aspectuelle non applicable, c'est-à-dire les occurrences sur lesquelles on ne peut pas passer les tests de classement aspectuel : les formes typiques sourdes et les semi-auxiliaires. Cette modalité s'éloigne des deux regroupements principaux (subordonnée « qui s'appelle » et phrases à verbe pronominal passif et sujet inanimé).

L'AFM indique que le point des occurrences d'état est très proche de celui de présence du SE, tandis que celui des occurrences ne pouvant avoir de classe aspectuelle est près du point des occurrences d'ajout (ce qui est attendu, puisque rares sont les semi-auxiliaires pronominaux et que les formes

typiques sourdes retenues pour l'analyse sont celles ayant un SE préverbal). La classe des accomplissements tend vers le point de respect de la norme, tandis que celle des activités et celle des achèvements, très proches l'une de l'autre, s'approchent du point d'absence du SE.

Les *catdes* indiquent que la classe d'état caractérise les occurrences de présence du SE et d'ajout non normé, mais que cela est possiblement dû à la surreprésentation des états dans les occurrences où la présence du SE est interdite selon la norme. En prenant en considération la variable [NormePres], il en ressort que les activités et les achèvements caractérisent les occurrences d'omission du SE et que les activités caractérisent les écarts à la norme.

6.3.4 Effet du classement lexico-sémantique

Dans le graphique de l'ACM, la classe lexico-sémantique la mieux représentée est la classe N (verbes de type *munir/démunir*), qui s'approche des occurrences à structure attributive : cette classe est celle qui correspond à l'entrée du verbe *appeler* dans LVF et le regroupement des occurrences à structure attributive comprend entre autres la subordonnée fréquente « qui s'appelle ». Les autres classes lexico-sémantiques ayant une qualité de représentation modérée sont la classe P (verbes psychologiques), placée en opposition au regroupement des verbes pronominaux passifs, ainsi que la classe U (verbes d'union, de réunion et de mise en relation), qui est en périphérie du regroupement des occurrences à sujet inanimé et verbe pronominal passif (ce qui reflète la mention fréquente de portes de métro s'ouvrant et se fermant dans le premier texte). Sinon, la classe C (verbes de communication) et la classe T (verbes de transformation, de mutation, de substitution) s'éloignent des deux regroupements (surtout des occurrences attributives et des occurrences passives, respectivement), la classe D (verbes de don ou de privation) est en opposition au regroupement des verbes pronominaux passifs et des sujets inanimés, et la classe X (verbes auxiliaires et semi-auxiliaires, impersonnels, *exister*, inchoatifs ou résultatifs) s'approche légèrement des deux rassemblements. Les autres classes sont dans le nuage de points, près de l'origine.

Dans le graphique de l'AFM, la classe N est située près des occurrences de présence normée, la classe U est proche du point des absences normées et la classe P se trouve très près du point des occurrences d'ajout. La classe R (verbes de type *réaliser*) approche des absences normées; la classe

T et la classe F (verbes de type *frapper* ou *toucher*), des omissions; et la classe X, des ajouts. Plus près de l'origine, on retrouve la classe D dans la zone des ajouts, la classe C, la classe H (verbes d'état physique ou de comportement) et la classe L (verbes locatifs) près des absences du SE, la classe S (verbes de saisie ou de possession) dans l'axe du point des écarts à la norme, la classe M (verbes de mouvement) et la classe E (verbes de mouvement d'entrée/sortie) dans la zone des absences normées.

Lorsqu'on ne prend pas en considération la distribution de [NormePres], les *catdes* indiquent que la variable de classification lexico-sémantique est significativement corrélée aux variables de présence du SE, de respect de la norme et de catégorie d'occurrence. Deux modalités caractérisent les occurrences d'absence du SE : la classe T, qui caractérise aussi les absences normées, et la classe U, qui caractérise aussi les absences normées et les présences normées facultatives, quoi qu'elle est corrélée négativement aux présences normées. Les modalités caractérisant les occurrences de présence du SE sont la classe N, qui caractérise aussi le respect de la norme et les présences normées en plus d'être corrélée négativement avec les ajouts et les omissions, ainsi que la classe P, qui caractérise aussi les écarts à la norme (plus précisément les ajouts) et est corrélée négativement avec les absences normées. Deux autres modalités caractérisent les écarts à la norme: la classe S, qui caractérise aussi les ajouts, ainsi que la classe X, corrélée positivement aux ajouts et négativement aux présences normées. Les autres corrélations significatives ne touchent que les catégories d'occurrences : la classe F caractérise les omissions; la classe C, les absences normées; la classe R, les absences normées tout en ayant une corrélation négative avec les ajouts; et la classe E, les présences normées tout en ayant une corrélation négative avec les absences normées.

Si on retire toute corrélation pouvant être expliquée (du moins partiellement) par la distribution de [NormePres], il ne reste que deux résultats significatifs : la classe N caractérise le respect de la norme (reflétant la maîtrise de la subordonnée « qui s'appelle », très fréquente dans le corpus), tandis que la classe F caractérise les omissions du SE.

6.3.5 Effet de la structure transitive

Les variables d'emploi et de norme de structure transitives sont les deux plus importantes dans la définition de la dimension 1 de l'ACM. Cette dimension différencie les phrases ayant une structure

attributive du reste du corpus. Les structures intransitives se rapprochent du regroupement des phrases à sujet inanimé et verbe pronominal passif sur la dimension 2, tandis que le reste des modalités des variables de structure transitive se retrouvent dans le nuage de points.

Dans l'AFM, on observe que les occurrences de structure attributive sont sur le même axe que le point des présences du SE. Les occurrences d'emploi d'une structure bitransitive (deux compléments verbaux) sont près du point de présence du SE, et sont dans la zone générale des ajouts, tandis que les occurrences où la norme demande une structure bitransitive sont plus près du point des ajouts. Le point <nND>, c'est-à-dire les occurrences pour lesquelles la norme quant à la structure transitive n'est pas disponible, est près du point des ajouts, ce qui est attendu étant donné que ce sont des occurrences de FTS; ce point est d'ailleurs très près du point des occurrences pour lesquelles la classe aspectuelle n'est pas applicable, qui regroupe aussi les FTS. Les occurrences de structure intransitive se situent près des absences normées, tandis que les transitives directes et indirectes se rapprochent des écarts à la norme.

Les *catdes*, si l'on ne prend pas en considération la distribution de [NormePres], indiquent que les occurrences de structure intransitive caractérisent les absences du SE (plus spécifiquement les absences normées) et caractérisent négativement les présences normées, tandis que les occurrences de structure transitive indirecte caractérisent les omissions. Si l'on divise emploi et norme de structure transitive, l'emploi transitif indirect caractérise aussi négativement les absences normées, et les occurrences pour lesquelles la structure normée est transitive indirecte caractérisent négativement les ajouts. Les structures attributives caractérisent les occurrences de présence du SE (plus spécifiquement de présences normées) et caractérisent négativement les absences normées. Néanmoins, toutes ces observations sont expliquées, sinon partiellement, par la distribution de [NormePres].

Lorsque l'on prend [NormePres] en considération, il reste des caractérisations significatives pour les structures bitransitives et pour les attributives. L'emploi d'une structure bitransitive caractérise la présence du SE, tandis que les occurrences où la structure normée est bitransitive sont négativement corrélées avec les occurrences d'absence normée du SE. Ces deux observations semblent être le reflet du fait que le SE a parfois une fonction syntaxique dans le syntagme verbal, et que les absences normées du SE représentent fréquemment un complément verbal sous-entendu,

par exemple *elle (s')achète un nouveau livre*. En ce qui concerne les structures attributives, elles caractérisent les occurrences de respect de la norme et sont corrélées négativement avec les occurrences d'omission. Ce résultat est étonnant, puisque les structures attributives caractérisent les occurrences où le pronom SE est obligatoire; on s'attendrait donc à ce qu'elles soient surreprésentées dans les deux catégories d'occurrences associées à la présence obligatoire du SE, soit les présences normées et les omissions. Or, les structures attributives sont surreprésentées dans les présences normées, mais sous-représentées dans les omissions. Il semblerait donc que de toutes les structures transitives, les attributives sont celles qui correspondent le plus à la norme dans le corpus. Cela provient possiblement de la fréquence élevée de la subordonnée « qui s'appelle », qui semble être bien maîtrisée par les participant·e·s.

Dans l'ACM, les trois modalités de la variable du respect de la structure transitive normée sont dans le nuage de points; le non-respect de la structure transitive s'éloignant davantage des regroupements de la subordonnée fréquente « qui s'appelle » et des phrases avec un sujet inanimé et un verbe pronominal passif. La modalité du mauvais choix de la préposition se trouve encore plus loin des regroupements, surtout des phrases pronominales passives, ce qui est attendu étant donné que ce type de verbe pronominal n'accepte pas de préposition.

Les modalités de la variable du respect de la structure transitive normée sont plus éloignées de l'origine dans l'AFM. Les occurrences de respect de la structure transitive se rapprochent du point du respect de la norme dans l'emploi des verbes pronominaux; tandis que les occurrences de non-respect de la structure transitive sont près du point des écarts à la norme. Le point des occurrences présentant un mauvais choix de la préposition se situe près du point des occurrences de l'ajout non normé du SE.

Les *catdes* révèlent que la variable du respect ou non de la structure transitive n'est pas liée significativement à la variable de présence ou d'absence du SE; elle est l'est néanmoins avec les variables du respect de la norme de l'emploi du SE et de la catégorie d'occurrence : le respect de la structure transitive caractérise les occurrences de respect de la norme (les présences et les absences normées) tandis que le non-respect de la structure transitive caractérise les occurrences d'écart à la norme (les ajouts et les omissions).

6.3.6 Effet de la forme du sujet

La variable de la forme du sujet a une grande contribution à la définition de la deuxième dimension de l'ACM, en troisième place d'importance derrière les variables du rôle thématique du sujet et du trait d'animéité du sujet. Dans le graphique de l'ACM, on observe que la forme du sujet ayant la meilleure qualité de représentation est le syntagme nominal, qui est en périphérie du regroupement des sujets inanimés et des verbes pronominaux passifs. Les deux autres formes de sujet étant relativement bien représentées sont les pronoms relatifs, qui se rapproche du regroupement des structures attributives (notamment la subordonnée « qui s'appelle »), ainsi que le pronom *je*, qui est en opposition aux sujets inanimés et aux verbes pronominaux passifs. À l'extérieur du nuage de point, on retrouve le pronom *nous* qui s'éloigne des deux regroupements, surtout celui des structures attributives. Le reste des formes de sujet sont dans le nuage de points autour de l'origine.

Dans le graphique de l'AFM, la modalité de la variable de forme du sujet la plus éloignée de l'origine est le pronom *nous*, qui s'approche des occurrences d'omissions. Ensuite vient le pronom relatif, qui se situe près des présences normées. Les syntagmes nominaux et le pronom *tu* se situent près des absences normées, tandis que les pronoms *ils* et *elles* sont près des occurrences d'absence du SE. Le pronom *je* se situe près des ajouts, et le *on* tend vers les omissions. Plus près de l'origine, on retrouve les absences non normées du sujet qui penchent vers les occurrences de présence du SE, les absences normées du sujet dans une coordination qui s'approchent des ajouts, ainsi que les sujets sous-entendus (tels que les sujets d'infinitifs, de participes présents ou de verbes impératifs) qui tendent vers les occurrences d'écart à la norme. La dernière forme de sujet, les pronoms *il* et *elle*, se trouve tout près de l'origine.

Les *catdes* indiquent que la forme du sujet n'est pas corrélée avec la variable de respect de la norme dans l'emploi des verbes pronominaux, mais l'est avec les variables de présence du SE et de catégorie d'occurrence. La distribution de [NormePres] ne peut pas expliquer toutes les corrélations, mais peut en expliquer certaines : que les syntagmes nominaux caractérisent les absences normées et sont corrélés négativement avec les omissions, que les absences non normées du sujet caractérisent les ajouts et que le sujet *je* caractérise les présences de SE ainsi que les ajouts. Même en prenant compte de l'influence de [NormePres], le pronom *je* sujet est négativement corrélé avec les absences normées. Les pronoms *on*, *il* et *elle* ainsi que *nous* caractérisent les

occurrences d'omissions; le *nous* caractérise aussi les absences de SE en général, en plus d'être corrélé négativement aux occurrences de présence normée. Les pronoms relatifs, quant à eux, caractérisent les présences de SE, plus particulièrement les présences normées.

6.3.7 Effet des traits sémantiques du sujet

La variable d'animéité ne contribue pas du tout à la définition de la première dimension de l'ACM, mais c'est la deuxième plus importante dans la définition de la deuxième dimension (légèrement derrière la variable de rôle thématique du sujet). En effet, la dimension 2 éloigne les occurrences ayant un sujet inanimé (et celles qui ont un verbe pronominal passif) du reste du corpus. Leur position est presque identique dans le graphique de l'AFM, quoique les points sont un peu plus près de l'origine.

Les *catdes* indiquent que la variable d'animéité n'est corrélée qu'avec les catégories d'occurrences. La modalité de sujet inanimé est caractéristique des occurrences d'ajout et d'absence normée, tandis que celle de sujet animé est caractéristique des occurrences d'omission et de présence normée. Néanmoins, ces corrélations pourraient n'être que le résultat de la distribution de [NormePres].

La variable du rôle thématique du sujet est la plus importante dans la définition de la deuxième dimension de l'ACM, en plus d'avoir un apport considérable pour la dimension 1. Les modalités les mieux représentées sont celles du rôle thématique thème, dont le point est situé au cœur du regroupement des occurrences au sujet inanimé et verbe pronominal passif; du rôle thématique patient, qui s'approche du regroupement des occurrences à structure attributive; et du rôle thématique agent, qui est situé légèrement en opposition à ces deux regroupements. Deux autres rôles thématiques ont une représentation de qualité modérée, soit siège, situé près du regroupement des verbes pronominaux passifs et sujets inanimés; et expérienceur, qui s'oppose au regroupement des sujets inanimés et s'approche quelque peu des verbes attributifs. Les deux modalités restantes, soit le rôle thématique instrument et les occurrences pour lesquelles le sujet n'a pas de rôle thématique (verbes impersonnels, auxiliaires, etc.), penchent vers le regroupement des verbes pronominaux passifs et sujets inanimés.

Dans le graphique de l'AFM, le rôle thématique patient se situe proche des occurrences de présence du SE, et celui agent tend vers les occurrences d'écart à la norme. Les occurrences de thème, celles de siège ainsi que celles pour lesquelles le sujet n'a pas de rôle thématique sont à proximité du point des occurrences d'absence normée. Les rôles expérienceur et instrument se rapprochent davantage aux occurrences d'ajout du SE.

Les *catdes* indiquent que la variable du rôle thématique du sujet n'a aucune corrélation significative avec la variable du respect de la norme; elle en a néanmoins avec les variables de présence du SE et de catégorie d'occurrence. Le rôle thématique agent caractérise les occurrences d'absence du SE, plus particulièrement les omissions, et est corrélé négativement avec les ajouts. Le rôle thématique patient, quant à lui, caractérise les occurrences de présence du SE, plus spécifiquement les présences normées. Le rôle thématique thème caractérise négativement les présences normées, et positivement les présences normées facultatives et les absences normées. Le rôle thématique siège caractérise positivement les ajouts et les absences normées et est corrélé négativement avec les omissions et les présences normées. Le rôle thématique expérienceur caractérise la présence du SE et son ajout non normé, et est corrélé négativement avec les absences normées. Finalement, le rôle thématique instrument est corrélé avec les occurrences d'ajout.

Lorsque l'on considère la distribution de [NormePres] dans les *catdes*, il subsiste certaines corrélations : celle entre le rôle thématique agent et l'absence du SE, celle entre le rôle thématique patient et la présence du SE – et plus spécifiquement de sa présence normée – et celle entre le rôle thématique siège et l'absence normée du SE.

La variable du trait de personne du sujet ne contribue presque pas à la première dimension de l'ACM, et sa contribution à la deuxième dimension est modérée. Les sujets de première et de troisième personne ont une moyenne qualité de représentation, tandis que ceux à la deuxième personne, très peu nombreux, en ont une mauvaise. La troisième personne tend vers le regroupement des sujets inanimés et des verbes pronominaux passifs tout en étant relativement près de l'origine : les sujets inanimés sont tous de la troisième personne, mais ce ne sont pas tous les sujets à la troisième personne qui sont inanimés. Les points de première et de deuxième personnes se rapprochent davantage du point de sujet animé, avec la première personne un peu plus loin sur l'axe par opposition au regroupement des sujets inanimés et verbes pronominaux passifs.

Dans l'AFM, la première personne se situe dans la zone générale des ajouts, en étant plus près du point d'écart à la norme en général. La deuxième personne s'approche des occurrences de présence du SE, tandis la troisième personne est près du point des occurrences respectant la norme, dans la zone générale des absences normées. Par contre, aucune de ces corrélations n'est statistiquement significative, étant donné qu'on ne retrouve pas cette variable dans les tableaux de *catdes*.

La variable du trait de nombre du sujet contribue peu aux définitions des dimensions de l'ACM. La modalité du trait singulier est près de l'origine, quoiqu'un peu plus près du regroupement des occurrences à structure transitive et un peu plus éloignée du regroupement des occurrences à verbe pronominal passif. La modalité du pluriel, plus éloignée de l'origine, se rapproche davantage du regroupement des verbes pronominaux passifs et s'éloigne du regroupement des verbes attributifs.

Les positions des traits singulier et pluriel sont presque identiques dans le graphique de l'AFM. Le trait singulier est dans l'axe du point des occurrences de présence du SE, et le trait pluriel s'approche du point des occurrences d'absence du SE, tout en étant dans la zone générale des absences normées.

Les *catdes* indiquent que le trait de nombre n'a pas de corrélation significative avec la variable du respect de la norme, mais qu'il en a une avec la variable de présence (variable la plus corrélée avec [PresSe], si on ignore celles qui sont expliquées par [NormePres]) et avec la variable de catégorie d'occurrence. Le trait singulier caractérise les occurrences de présence du SE, spécifiquement celles d'ajout, mais cette corrélation peut provenir de la distribution de [NormePres]. Sinon, le trait singulier caractérise les occurrences de présence normée, tandis que le trait pluriel caractérise les occurrences d'absence du SE, particulièrement des omissions. Il est possible de faire un parallèle avec l'hypothèse émise quant aux résultats *catdes* des variables de temps verbal : plus un élément est complexe cognitivement, moins le pronom SE est employé.

6.4 Observations liées à la question de recherche 1 (respect de la norme)

La première question spécifique de recherche est la suivante : Quels sont les contextes morphosyntaxiques ou sémantiques dans lesquels on observe une différence entre la norme du français écrit quant à l'utilisation du pronom SE et son emploi dans le corpus Groupe LSQ 98 ? En

d'autres mots, y a-t-il des contextes pour lesquels la norme de l'emploi des verbes pronominaux est mieux maîtrisée ?

Avant toute chose, il a fallu vérifier s'il existe des corrélations entre les variables d'identification et [RspNorme], c'est-à-dire si le taux de respect de la norme dépend d'une tendance individuelle ou du sujet de rédaction. Or, les tests du chi-carré indiquent qu'il n'y a pas de corrélation statistiquement significative entre l'une ou l'autre de ces variables et le respect de la norme; les participant·e·s sont comparables, tout comme les différents textes.

Ensuite, il a été possible d'observer la distribution de la variable de respect de la norme en fonction des variables contextuelles. L'analyse orientée sur les variables d'observation offre une assez bonne qualité de représentation des modalités de respect de la norme et d'écart à la norme – les deux premières dimensions de l'AFM permettent de modéliser près de 75% de leur variation.

La suite de cette section se divise comme suit : une première sous-section traite des modalités caractérisant les occurrences respectant la norme (6.4.1), la suivante présente celles caractérisant les occurrences d'écart à la norme (6.4.2), et la dernière se penche plus particulièrement sur les occurrences d'ajout non normé du SE (6.4.3).

6.4.1 Description des occurrences respectant la norme d'emploi des verbes pronominaux

Selon les résultats des tests *catdes*, quatre modalités caractérisent les occurrences de respect à la norme sans que cela puisse être expliqué par la distribution de [NormePres] : les verbes de la classe lexico-sémantique N (verbe de type *munir/démunir*), les occurrences ayant une structure attributive, les occurrences dont la structure normée est attributive, ainsi que les verbes pronominaux neutres. À cela s'ajoutent trois modalités, elles aussi caractérisant les occurrences de respect à la norme, mais étant aussi surreprésentées dans les occurrences où la présence du SE est facultative (c'est-à-dire dont la distribution de [NormePres] explique la surreprésentation dans les occurrences normées) : les occurrences pour lesquelles la structure transitive est normée, les verbes pronominaux réfléchis et réciproques dont le SE a la fonction complément indirect du verbe, ainsi que les verbes pronominaux passifs.

De toutes les occurrences retenues pour l'analyse, 31 correspondent à la classe lexico-sémantique N. Il n'y a aucune occurrence d'ajout pour cette classe et, lorsque la présence du SE est obligatoire, il est présent 93% du temps (contrairement à 63% pour l'ensemble du corpus). Parmi les 31 occurrences de classe N, on retrouve 20 occurrences des verbes *s'appeler* ou *se nommer*, ainsi que huit occurrences du verbe *s'habiller*. Ces verbes fréquents semblent bien maîtrisés : de ces 28 occurrences, 27 présentent une présence normée du SE (le pronom réfléchi du verbe *s'habiller* ayant été omis une fois). Les trois autres occurrences appartenant à la classe N correspondent à deux absences normées dans des structures passives (*se mouiller* et *se vider*) et à une omission (*se venger*). Ainsi, il se pourrait que le fait que la classe N caractérise les occurrences normées ne provienne pas des caractéristiques sémantiques associées à cette classe, mais provienne plutôt de la surreprésentation de verbes fréquents dans les occurrences; tout comme il semblerait être le cas avec les occurrences à structure attributive. Il aurait été intéressant de pouvoir séparer l'effet de la fréquence d'un verbe des autres variables contextuelles; malheureusement, aucune base de données lexicales ne comportait de recherche multimot, nécessaire pour les verbes pronominaux.

Les verbes employés dans une structure attributive sont au nombre de 37, tandis qu'il y a 36 occurrences dont la structure transitive normée est celle attributive; un total de 38 occurrences ont une structure attributive soit employée, soit normée. Encore une fois, plus de la moitié de ces occurrences (20) correspond aux verbes *s'appeler* / *se nommer*, qui ont un taux de présence normée de 100%. On retrouve aussi neuf occurrences du verbe *se sentir*, qui se divisent en six présences normées et trois omissions, ainsi que quatre occurrences d'ajout non normé avec le verbe *être*. Nous avons vu précédemment que la structure attributive caractérise aussi les occurrences de présence du SE en général (voir 6.3.5), ainsi que les occurrences où la présence du SE est obligatoire. La fréquence lexicale du verbe *s'appeler* ne peut expliquer à elle seule ces résultats. En effet, même lorsqu'on enlève les occurrences de *s'appeler* et de *se nommer*, les occurrences attributives se démarquent du reste du corpus : quand le SE est obligatoire, il est produit dans 73% des cas (contre 63% dans l'ensemble du corpus); les écarts à la norme présentent davantage d'ajouts que d'omissions et, lorsque le SE est présent, il n'est normé que dans 57% des cas (contre 71% dans l'ensemble du corpus). Il semblerait donc que la structure attributive soit avant toute chose corrélée à la présence du SE, et que la surreprésentation de cette structure transitive dans les occurrences où le SE est obligatoire fait en sorte qu'elle caractérise le respect à la norme.

Le type de verbe pronominal le plus employé est celui des verbes pronominaux neutres, avec 130 occurrences. Parmi celles-ci, 70 sont des présences normées (dont deux facultatives), 42 sont des omissions, 17 sont des absences normées et une seule est un ajout. La proportion de présences normées et d'omissions lorsque la présence du SE est obligatoire est similaire à celle du corpus. Néanmoins, lorsque le SE est présent, il est normé dans 99% des cas (contre 71% dans l'ensemble du corpus). Les omissions correspondent à 98% des écarts à la norme des occurrences de verbes pronominaux neutres (contre 62% pour l'ensemble du corpus). En d'autres mots, les occurrences de ce type de verbe pronominal se distinguent des autres par la sous-représentation des ajouts non normés du SE. Ainsi, il ne semblerait pas avoir beaucoup de confusion entre la forme pronominale et celle non pronominale de verbes pronominaux neutres, même si leurs sens sont similaires; on ne trouve d'ailleurs aucune occurrence d'ajout construite sur le modèle de « je me promène mon chien ». Il semble peu probable que le haut taux de respect à la norme proviennent du changement de sens, puisque la catégorie des verbes pronominaux autonomes et des verbes essentiellement pronominaux ne caractérisent pas les occurrences respectant la norme. D'ailleurs, des onze occurrences de verbes essentiellement pronominaux du corpus, quatre présentent une omission – ce qui correspond au pourcentage d'omission en moyenne dans le corpus dans les cas où le SE est obligatoire. Finalement, la catégorie des verbes pronominaux neutres est incontestablement la moins connue du public général et celle qu'on retrouve le plus rarement dans les grammaires scolaires, étant généralement regroupée avec les autonomes, les réfléchis, les réciproques et les passifs sous l'étiquette de verbes « occasionnellement pronominaux ». De prime à bord, on aurait pu s'attendre à ce que cette catégorie méconnue de verbes pronominaux soit moins bien maîtrisée que les autres, mais c'est le contraire que l'on observe.

Deux autres catégories de verbes pronominaux caractérisent les occurrences respectant la norme, soit les verbes pronominaux réfléchis et réciproques CI ainsi que les verbes pronominaux passifs. Concernant les verbes réfléchis et réciproques CI, eux aussi se distinguent des autres types de verbes pronominaux par la sous-représentation des ajouts non normés de SE – il n'y a aucune occurrence de cette catégorie dans les verbes réfléchis et réciproques CI. Ceci est un résultat attendu, étant donné qu'il n'était pas possible de déterminer si les SE ajoutés remplissaient une fonction de complément verbal ou non. De plus, la corrélation entre les verbes pronominaux réfléchis et réciproques CI et les occurrences de respect de la norme s'explique par la

surreprésentation des occurrences où le SE est facultatif dans cette catégorie de verbes pronominaux. En effet, les occurrences pour lesquels le SE est facultatif représentent 23% du corpus entier, mais 68% des occurrences de verbe pronominal réfléchi ou réciproque CI. Or, ces occurrences sont toujours normées.

Avec les verbes réfléchis et réciproques CI, les verbes passifs sont les deux types de verbes pronominaux caractérisant les occurrences pour lesquelles le SE est facultatif. En effet, pour 73% des occurrences de verbes pronominaux passifs (27 occurrences sur 37), la forme pronominale est facultative, tout en respectant la norme.

Finalement, le respect de la norme pour la structure transitive employée caractérise les occurrences de respect de la norme de l'emploi des verbes pronominaux. Cela semble logique : plus le syntagme verbal est construit de manière à respecter la norme du français écrit (donc plus que la production s'approche de la langue cible), moins il y aura d'occurrences d'omission ou d'ajout non normé du SE – et ce, même si la variable de respect de la structure transitive ne dépend pas de la présence ou de l'absence du SE. Néanmoins, il n'est pas possible d'affirmer que cela n'est pas seulement un effet de la distribution de [NormePres], étant donné que les occurrences dont la structure transitive est normée (77% du corpus) sont surreprésentées dans les occurrences pour lesquelles la présence du SE est facultative (représentant 87% de ces occurrences).

6.4.2 Description des occurrences présentant un écart à la norme d'emploi des verbes pronominaux

De toutes les modalités caractérisant les écarts à la norme, seule une ne peut être expliquée par la distribution de [NormePres], soit la classe aspectuelle des activités. Toutes les autres (les classes lexico-sémantiques X, S et P; l'emploi d'une structure transitive non normée ainsi que les occurrences pour lesquelles on ne peut déterminer une norme de structure transitive, soit les FTS et les semi-auxiliaires) sont caractérisées par une surreprésentation des occurrences pour lesquelles le SE est interdit.

La classe aspectuelle des activités représente 37% du corpus et 44% des écarts à la norme. Au sein des occurrences d'activité, 48% sont des écarts à la norme, tandis que 41% des occurrences du

corpus en sont. Le ratio de respect ou écart à la norme dans les occurrences de présence du SE est similaire à celui pour l'ensemble du corpus, tout comme l'est celui des ajouts et omissions dans les occurrences d'écart à la norme. Là où la classe des activités se distingue davantage est dans le taux de production et d'absence dans les cas où le SE est obligatoire : il est produit dans 49% des cas (contre 63% pour l'ensemble du corpus) et omis dans 51% des cas (contre 37% pour l'ensemble du corpus). C'est donc par leur plus haut taux d'omissions que les occurrences de la classe aspectuelle activité se distinguent des autres classes aspectuelles; il est plus ardu de formuler une hypothèse à savoir pourquoi.

Des douze occurrences correspondant à la classe lexico-sémantique X (verbes auxiliaires et semi-auxiliaires, impersonnels, *exister*, inchoatifs ou résultatifs), seules deux respectent la norme. Plus précisément, les occurrences appartenant à la classe X correspondent à 8 ajouts, 2 omissions et 2 absences normées. Cette classe représente beaucoup d'ajouts, ce qui est contraire à l'observation que les pronoms personnels sont moins produits en français sourd, tandis que les verbes de cette classe peuvent rarement être pronominaux. Lorsqu'on regarde les douze occurrences de plus près, on remarque que trois d'entre elles (une omission et deux ajouts) semblent provenir d'une confusion entre deux semi-auxiliaires à sens similaire : **mettre à* a pu être utilisé plutôt que *se mettre à* par confusion avec *commencer à* ou *partir à*, et inversement pour **se partir* qui serait le reflet de cette confusion; et *s'aller* semble être un mélange entre *aller* et *s'en aller*. Six autres occurrences non normées sont liées à des procès verbaux dénotant un état ou un changement d'état : quatre occurrences d'ajout correspondent au verbe *être*, et deux autres ajouts correspondent à *se recommencer à* + un état⁴². Il est possible de lier cette observation à l'utilisation de la forme pronominale pour marquer un aspect inchoatif (la différence entre *voler* et *s'envoler* notamment) ainsi que pour marquer les changements d'état (tel que dans *le ciel s'assombrit*) (Labelle, 1992). Il ne nous reste qu'un écart à la norme appartenant à la classe X, soit l'omission dans l'occurrence [...] *ça faisait pas très facile* [...] ('ça se faisait difficilement').

61 occurrences appartiennent à la classe lexico-sémantique P (verbes psychologiques), dont 32 présentent des écarts à la norme. De ces écarts, plus de la moitié correspond à des ajouts, tandis que les ajouts représentent 38% des occurrences totales. De plus, parmi les occurrences de verbes

⁴² À noter ici qu'il s'agit d'un état dans le sens sémantique du terme, non pas en tant que classe aspectuelle.

psychologiques, lorsque le SE est présent, il est normé dans 60% des cas et un ajout dans 40% (comparativement aux moyennes générales du corpus de 71% et 29%, respectivement). Ainsi, les verbes de la classe P sont fréquemment accompagnés d'un ajout non normé du SE, ce qui, encore une fois, va à l'encontre de la tendance générale en français sourd à produire moins de pronoms personnels. Il serait intéressant d'explorer cette piste, et déterminer si cette tendance aux ajouts pour les verbes psychologiques pourrait provenir d'un lien entre le SE et le concept de *soi*, de personne⁴³.

La dernière classe lexico-sémantique à caractériser les écarts à la norme est la classe S (verbes de saisie ou de possession). Des 13 occurrences appartenant à cette classe, 69% sont des écarts à la norme, soit six ajouts et trois omissions. Dans l'ensemble du corpus, lorsque la présence du SE est obligatoire, il est omis dans 37% des occurrences; cette proportion s'élève à 60% pour les occurrences de la classe S – néanmoins, ce pourcentage provient de très peu d'occurrences. De plus, lorsqu'un SE est présent avec un verbe de la classe S, il s'agit d'un ajout non normé dans 75% du temps (contre 29% pour l'ensemble du corpus). Le 2/3 des écarts à la norme sont des ajouts, tandis que dans l'ensemble du corpus les ajouts ne représentent que 38% des écarts à la norme. Néanmoins, il est ardu de formuler une hypothèse à savoir pourquoi les verbes de cette catégorie se comportent différemment. À mon avis, cette classe lexico-sémantique a une définition difficile à cerner; le sens de saisie ou de possession est parfois très concret (*serrer quelqu'un dans ses bras*), mais d'autres fois plus métaphorique (*prendre un tel chemin* ou *arrêter son mouvement*), voire très abstrait (*goûter quelque chose* ou *se sauver la vie*).

Deux autres modalités caractérisent les occurrences d'écart à la norme, provenant toutes deux de variables qui concernent la structure transitive. La structure transitive est observée dans trois variables : a) la structure transitive telle qu'employée dans l'occurrence, b) la structure transitive normée pour le verbe en concordance avec le contexte de l'occurrence, et c) la correspondance entre les deux premières variables, c'est-à-dire le respect ou l'écart à la norme dans la structure transitive employée; c'est les deux dernières variables (celle de la structure transitive normée et celle du respect de la norme) dont il est question. Ainsi, les occurrences d'écart à la norme dans l'emploi du SE sont caractérisées par une surreprésentation des occurrences présentant également

⁴³ Voir 6.6 pour une réflexion à ce sujet en lien avec les propositions sur le morphème *s-* de Kayne (2000).

un écart à la norme dans la structure transitive, ainsi que les occurrences pour lesquelles il n'est pas possible d'identifier une structure transitive normée, c'est-à-dire lorsque l'occurrence est soit un semi-auxiliaire, soit la FTS de glissement de catégorie, dans ce cas un nom employé comme verbe. La corrélation entre un écart à la norme dans la structure transitive et un écart à la norme dans l'emploi du SE ne surprend pas, étant donné que le français sourd est décrit comme comportant fréquemment des phrases accumulant de nombreuses erreurs de syntaxe. La corrélation entre les écarts à la norme et les occurrences pour lesquelles on ne peut identifier de structure transitive normée (FTS et semi-auxiliaires) est quelque peu attendue. En effet, les FTS touchant le choix du verbe (tel que l'emploi d'un nom en tant que verbe) ayant été retenues pour l'analyse correspondent presque exclusivement à des ajouts: le choix lexical du verbe n'étant pas normé, il n'y a pas de norme quant à son emploi et donc le SE ne peut être facultatif ou obligatoire. Néanmoins, lorsqu'on ne regarde que les semi-auxiliaires, les occurrences d'ajout demeurent surreprésentées (correspondant à 23% des occurrences de semi-auxiliaires comparativement à 16% du corpus en général), mais il n'est pas possible de déterminer si ce résultat est statistiquement significatif.

6.4.3 Description des occurrences d'ajout non normé du SE

Dans l'analyse des résultats des tests *catdes*, la distribution de [NormePres] a été prise en compte, puisqu'elle pouvait expliquer (partiellement ou complètement) la présence de certaines modalités dans les résultats. Toutefois, il reste pertinent d'observer les modalités caractérisant les occurrences d'ajout non normé du SE, même s'il s'avère que celles-ci sont surreprésentées dans la modalité de présence interdite du SE de la variable [NormePres]. En effet, les verbes n'étant pas précédés du SE dans un contexte où la norme interdit la forme pronominale n'ont pas été retenus pour l'analyse, puisque trop nombreux. La discussion des occurrences d'ajouts non normés du SE aurait été plus complète avec une comparaison à ces verbes n'ayant pas été retenus, mais elle demeure intéressante.

Les modalités caractérisant les occurrences d'ajout sont : les classes lexico-sémantiques X, P et S; les rôles thématiques siège, expérimenteur et instrument; la conjugaison du verbe à un temps simple; les structures transitives pour lesquelles on ne peut établir de norme (FTS et semi-auxiliaires) ainsi que les écarts à la norme dans la structure transitive employée; le pronom *je* comme sujet et les absences non normées du sujet; les sujets singuliers et inanimés; et la classe aspectuelle d'état. Les

occurrences d'ajouts pour les modalités liées à la structure transitive et celles de classes sémantiques ont été discutées à la sous-section précédente (6.4.2).

Les 20 occurrences dont le sujet a le rôle thématique siège comprennent neuf ajouts. Parmi celles-ci, sept introduisent un attribut du sujet ou dénotent un changement d'état – d'ailleurs, deux semblent provenir d'une confusion entre SE et le verbe *être*. Les deux autres ajouts, présentés en (175) et (176), désignent un procès verbal similaire à ceux d'action interne que l'on retrouve dans les SE lexicaux tel que *s'accroupir* (voir 2.3).

(175) Le téléphone se sonne (P14t2).

(176) Son coeur se bat (P14t4).

Les occurrences dont le sujet a le rôle thématique instrument ne sont qu'au nombre de deux ; il est surprenant que cela ait été suffisant pour obtenir un résultat statistiquement significatif. Dans les deux cas, il y a un ajout non normé du SE et le sujet a la forme d'un SN portant le trait sémantique d'inanimé. De plus, dans les deux cas, l'occurrence serait normée si le SE était remplacé par le pronom *nous* complément direct ou indirect du verbe.

Le troisième rôle thématique caractérisant les occurrences d'ajout est celui expérienceur, qui correspond à 56 occurrences – dont 15 ajouts. Les occurrences dont le sujet a le rôle thématique expérienceur représentent 16% du corpus, mais 28% des ajouts. Il est possible ici de faire un lien avec les occurrences d'état, de structure attributive et de verbes psychologiques.

Les occurrences conjuguées à un temps simple, au nombre de 215, représentent 63% des occurrences retenues pour l'étude et 81% des ajouts. De manière générale, la distribution des catégories d'occurrences est similaire à celle de l'ensemble du corpus, mis à part une plus grande quantité d'ajouts que celle attendue dans les écarts à la norme (48% plutôt que 38%). Il semblerait que cette modalité caractérise les occurrences d'ajout surtout car elle caractérise aussi la présence du SE (ce qui est discuté en 6.5).

Les 88 occurrences ayant le pronom *je* comme sujet représentent 26% du corpus et 38% des ajouts. Ces occurrences se divisent en 20 occurrences d'ajout, 39 de présence normée, 11 d'absence

normée et 18 d'omission. Lorsqu'on compare cette distribution à celle de l'ensemble du corpus, on remarque que les occurrences d'ajout sont surreprésentées tant dans leur proportion dans les occurrences de présence du SE que dans leur proportion dans les occurrences présentant un écart à la norme.

L'autre forme de sujet qui caractérise les occurrences d'ajout est l'absence non normée du sujet, qui correspond à une FTS fréquemment discutée dans la littérature (voir 3.1.3). Un aspect intrigant des dix occurrences d'absence non normée du sujet (et possiblement un simple hasard) est que les cinq occurrences d'ajouts ont toutes une structure intransitive; la sixième occurrence d'absence de sujet non normée du sujet avant un verbe intransitif étant une absence normée du SE. Considérant que les écarts à la norme dans la structure transitive caractérisent aussi les occurrences d'ajout, les résultats suggèrent que les écarts à la norme ne viennent pas seuls et que, plus la construction syntaxique d'une phrase sera éloignée de la norme de syntaxe du français, moins l'emploi des verbes pronominaux sera normé.

Les occurrences dont le sujet est au singulier, au nombre de 264, représentent la grande majorité du corpus. De celles-ci, 47 correspondent à des ajouts. Des 102 écarts à la norme, 46% correspondent à des ajouts (contre 38% dans l'ensemble du corpus). Par contre, lorsqu'on se penche sur les 134 présences du SE, 84% sont une présence normée (contre 71% dans l'ensemble du corpus); il se pourrait donc que la caractérisation des ajouts par les sujets au singulier provienne plutôt de la surreprésentation des occurrences de présence du SE dans les occurrences ayant un sujet au singulier.

Les occurrences d'ajouts sont aussi caractérisées par les occurrences dont le sujet est sémantiquement inanimé. De tout le corpus, 66 occurrences ont un sujet inanimé, dont 16 présentent un ajout du SE. Les ajouts sont surreprésentés dans les occurrences de présence du SE (dont 50% sont normées, comparativement à 71% pour l'ensemble du corpus) et dans les écarts à la norme (dont 64% correspondent à des ajouts, comparativement à 38% pour l'ensemble du corpus). Néanmoins, il est difficile d'établir un lien théorique entre les sujets sémantiquement inanimés et l'ajout non normé du SE.

La dernière modalité caractérisant les occurrences d'ajout est la classe aspectuelle d'état, qui correspond à 91 occurrences, dont 21 ajouts. Lorsqu'on observe les occurrences où le SE est présent, il est normé dans 64% des cas (contre 71% pour l'ensemble du corpus), et les occurrences d'écart à la norme sont majoritairement des ajouts, bien que cette catégorie d'occurrences représente 38% des écarts à la norme dans l'ensemble du corpus.

Les modalités qui caractérisent négativement les occurrences d'ajout sont les verbes pronominaux neutres, les pronominaux autonomes et essentiellement pronominaux, les réfléchis et réciproques (complément direct et indirect), et les pronominaux passifs; le rôle thématique agent; la conjugaison à un temps composé; les classes lexico-sémantiques N (verbes de type *munir* ou *démunir*) et R (verbes de type *réaliser*); la structure transitive indirecte comme construction normée; ainsi que les sujets au pluriel et ceux ayant le trait animé. Ces corrélations ne sont pas discutées en détail, mais il est important de préciser que la présence de plusieurs types de verbes pronominaux s'explique en grande partie par le fait que l'entièreté des ajouts correspond à la modalité *pas pronominal dans ce contexte*, résultant en une sous-représentation des ajouts dans toutes les autres modalités de type de verbe pronominal.

6.5 Observations liées à la question de recherche 2 (présence du SE)

La deuxième question spécifique de recherche est la suivante: est-ce possible d'observer une « norme sourde » dans l'emploi du SE dans le corpus Groupe LSQ 98 et, si tel est le cas, quels sont les facteurs qui balisent la présence ou l'absence du SE ?

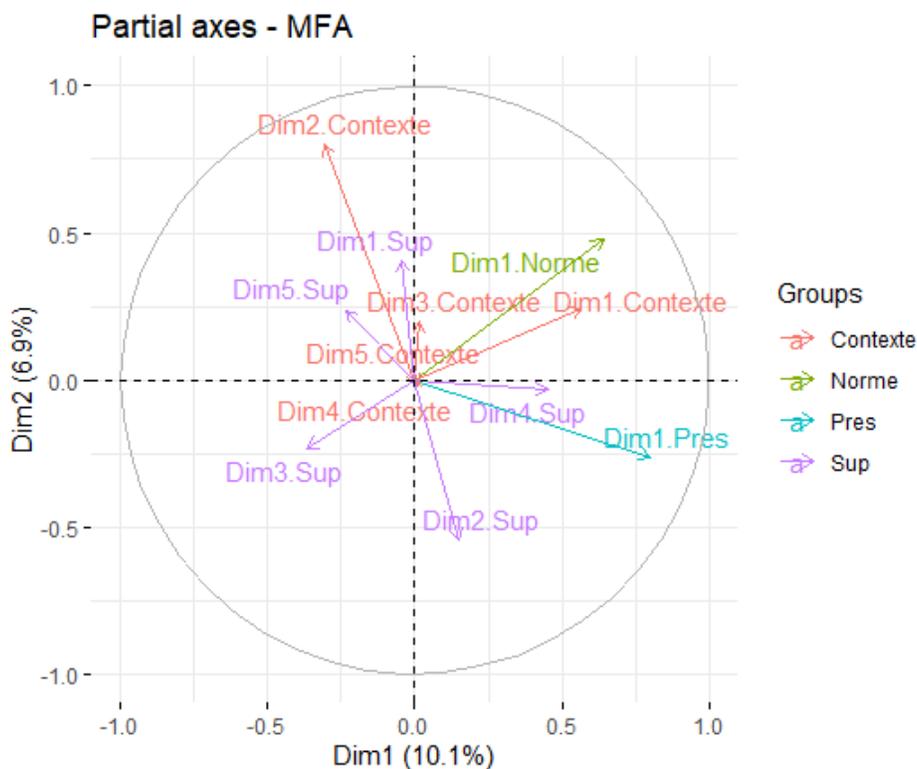
Afin de répondre à cette question, la variation de la présence ou de l'absence du SE a été observée en fonction de variables contextuelles, après une vérification de l'influence des variables d'identification, soit le sujet de rédaction et le numéro de participant·e.

Les tests du chi-carré indiquent que les participant·e·s ne s'équivalent pas tou·te·s dans la fréquence d'emploi du SE, et que l'utilisation de verbes pronominaux dépend à un certain point d'une tendance individuelle. Cette corrélation est toutefois limitée : des quatorze participant·e·s, seul·e·s trois se distinguent du groupe, soit P15 avec davantage de présences du SE ainsi que P10 et P11 avec davantage d'absences du SE. De plus, la distribution de la variable [NormePres] indique

que P10 avait significativement moins d'occurrences où le SE est obligatoire, expliquant son plus bas taux de production.

La variation de la variable de présence ou absence du SE a été observée en fonction des variables contextuelles. L'analyse factorielle multiple offre une bonne qualité de représentation des modalités de présence ou absence du SE, les deux premières dimensions permettant de représenter près de 85% de leur variation. Néanmoins, le graphique des axes partiels indique que les variables contextuelles ne sont pas les plus adéquates pour prédire la présence ou l'absence du SE (figure 6.1). En effet, même en considérant cinq dimensions de variables contextuelles et d'identification, la dimension de présence ou absence du SE (la flèche bleue) n'est pas en proximité de dimension d'autres regroupements de variables, contrairement à la dimension de respect ou non de la norme (la flèche verte), qui est assez proche de la première dimension du regroupement de variables contextuelles.

Figure 6.1 – Graphique des axes partiels de l'AFM (5 dimensions)



Ce que l'on comprend de ce graphique est que les variables et modalités ayant joué un rôle important dans la description de la première dimension des variables contextuelles sont aussi importantes pour distinguer les occurrences respectant la norme ou non. En d'autres mots, les variables contextuelles sont pertinentes pour déterminer si l'occurrence correspond à la norme, mais pas pour savoir si le SE sera produit ou non. Il se pourrait donc que les facteurs déterminants si un verbe sera pronominal ou non ne proviennent pas du français, mais plutôt de transferts de la LSQ.

Ainsi, la variable de présence du SE est corrélée significativement à une variable d'identification (participant·e), en plus de ne pas présenter de corrélation intéressante avec les cinq premières dimensions des variables contextuelles. Considérant ceci, en plus de l'absence de données comparables en français langue maternelle ou langue seconde de participant·e·s entendant·e·s, les résultats ayant trait à la deuxième question de recherche ne seront pas discutés en détail; les tests *catdes* sur la variable de présence du SE doivent être considérés plutôt comme des pistes à explorer que comme des résultats concluants.

Il est tout de même possible de formuler des observations générales. Déjà, la majorité des verbes du corpus Groupe LSQ 98 n'a pas été retenue pour l'étude, étant donné qu'il n'y avait pas de SE et que son ajout n'aurait pas été normé. En d'autres mots, les verbes pronominaux sont minoritaires et la tendance générale est de ne pas produire de SE avant un verbe. Ainsi, lorsque la norme interdit la présence du SE, il est rarement produit, tout comme lorsque sa présence est facultative. Néanmoins, lorsque sa présence est obligatoire, il est plus souvent présent (59%) qu'absent (41%).

6.6 Hypothèses diverses

Une piste qui semble intéressante à explorer est l'apparente confusion fréquente entre le pronom SE et le verbe *être* – il est d'ailleurs pertinent de noter ici que ce verbe n'existe pas en LSQ. Cette confusion résulte en occurrences de phrases dénotant un état ou un changement d'état avec un ajout du SE, tel qu'en (177) et en (178).

(177) Rachel se soulage (P03t1, pour 'Rachel est soulagée').

(178) Le bois se danger (P05t6, pour 'Le bois est dangereux').

Or, ce genre de construction existe en français, et il est possible d'utiliser tant *Jérôme est inquiet* que *Jérôme s'inquiète*; néanmoins cette possibilité dépend de propriétés lexicales et est difficile à prédire, et donc d'enseigner en français langue seconde.

De plus, il existe en français une alternance entre différentes formes de pronoms de troisième personne. Tandis que les pronoms et possessifs de première personne au singulier commencent presque exclusivement en *m-* (*me, mon, moi*) et que ceux de deuxième personne au singulier débutent par *t-* (*te, ton, toi*), ceux de troisième personne sont parfois en *l-* et parfois en *s-* (*le, la, les, lui, se, son, leurs, soi*). Kayne (2000) indique que cette alternance distingue les pronoms de troisième personne de ceux de première ou deuxième personne; le SE – et, par extension, *soi* – semblent avoir un sens abstrait de personne ou d'identité qui n'est pas encore bien défini. Par conséquent, il est difficile de l'enseigner, ce qui pourrait résulter en la surreprésentation de présence du SE et d'ajouts dans de nombreuses modalités, dont les structures transitives, les verbes d'état, les verbes psychologiques, et ainsi de suite.

Finalement, on aurait pu s'attendre à l'utilisation de stratégies d'évitement, tel que *prendre une douche* ou *mettre ses vêtements* plutôt que *se doucher* ou *s'habiller*, mais ce genre de formulation n'a pas semblé être plus fréquent que dans un corpus de français langue maternelle. Comme le souligne une évaluatrice, cela peut s'expliquer par le fait que le recours à de telles stratégies requiert une bonne connaissance des collocations abstraites du français, dont l'apprentissage est difficile pour les apprenant·e·s sourd·e·s.

CHAPITRE 7 :

CONCLUSION

Les textes écrits par des personnes sourdes comportent des formes typiques sourdes, dont certains écarts à la norme qu'on retrouve dans l'ensemble des apprenant·e·s du français langue seconde et d'autres qui sont propres à la population sourde (Nadeau et Machabée, 1998). Cette observation a été faite dans plusieurs langues, ce qui appuie l'hypothèse de l'écriture sourde (Charrow, 1975). Selon cette hypothèse, les formes typiques sourdes s'expliquent par un transfert linguistique des structures d'une langue des signes employée comme langue de référence dans une production dans une langue seconde lorsque les structures des deux langues sont jugées similaires. Les formes typiques sourdes caractérisant le français sourd sont bien documentées dans la littérature, mais l'emploi des verbes pronominaux n'est pas très bien compris : on indique seulement que le pronom SE est parfois omis, parfois de trop. Or, le pronom SE remplit une grande variété de rôles, ce qui rend l'identification d'une structure correspondante en LSQ plus difficile.

La recherche menée dans le cadre de ce mémoire a voulu combler quelque peu la lacune dans la littérature en offrant, dans une démarche exploratoire, une description de l'emploi du pronom préverbal SE dans un corpus de textes identifié comme étant porteur des caractéristiques du français sourd, soit le corpus LSQ 98. Les 14 participant·e·s dont les textes ont été retenus sont Sourd·e·s de naissance ou devenu·e Sourd·e à 18 mois au plus tard. Afin d'atteindre l'objectif général, deux questions spécifiques de recherche ont été formulées : 1) Quels sont les contextes morphosyntaxiques ou sémantiques dans lesquels on observe une différence entre la norme du français écrit quant à l'utilisation du pronom SE et son emploi dans le corpus Groupe LSQ 98 ? 2) Est-ce possible d'observer une « norme sourde » dans l'emploi du SE dans le corpus Groupe LSQ 98? Si oui, quels sont les facteurs qui balisent la présence ou l'absence du SE ?

Les 84 textes ont été analysés afin d'identifier les occurrences de verbes retenus pour l'étude, c'est-à-dire les verbes précédés du SE (présences normées et ajouts), ceux qui devraient l'être selon la norme (omissions) ainsi que ceux qui pourraient l'être, mais ne le sont pas (absences normées). Les verbes n'étant pas précédés du SE et qui ne pouvaient pas l'être selon le contexte n'ont pas été retenus à cause de leur trop grand nombre. La réponse aux questions spécifique de recherche a été

formulée à partir de la variation des variables de présence ou absence du SE et de respect ou non de la norme en fonction de variables d'observations : le type de verbe pronominal, le temps de conjugaison, la présence d'un verbe infinitif, le classement aspectuel du procès, la classe sémantico-syntaxique du verbe, la structure transitive, la forme du sujet, le rôle thématique du sujet, l'animéité du sujet ainsi que ses traits de nombre et de personne. L'analyse s'est faite en trois temps : une analyse des correspondances multiples (ACM) pour observer la variation générale du jeu de données, une analyse factorielle multiple (AFM) orientée sur la variation des variables décrivant l'emploi du SE, ainsi que des tests *catdes* afin de préciser les résultats de l'AFM.

Les résultats indiquent que, lorsqu'il est obligatoire, le SE est produit 63% du temps, mais il l'est seulement dans 10% des occurrences où sa présence est facultative. Lorsqu'il est produit, sa présence est normée dans 71% des cas, et les écarts à la norme comprennent davantage d'omissions que d'ajouts. Ces observations sont compatibles avec la littérature sur le français sourd, qui indique que les pronoms personnels sont moins fréquemment employés dans les textes de Sourd·e·s.

La première question de recherche cherche à savoir s'il y a des contextes dans lesquels la norme de l'emploi du SE sera davantage ou moins respectée. Les analyses effectuées ne permettent pas d'identifier de contexte qui entraîne systématiquement un écart à la norme, mais plusieurs contextes sont significativement corrélés à la variable du respect de la norme. Les principaux contextes caractérisant les occurrences de respect de la norme sont l'emploi de verbes pronominaux neutres, autonomes ou essentiellement pronominaux; la classe sémantico-syntaxique N (verbes de type *munir / démunir*); ainsi que la structure attributive. Les principaux contextes caractérisant les écarts à la norme sont l'emploi de verbes pronominaux réfléchis CD (dont le SE est fréquemment omis); la conjugaison du verbe à un temps composé (bien que ce contexte soit surreprésenté dans les occurrences où la présence du SE est facultative et qui respectent la norme, peu importe si le SE est présent ou non); les classes aspectuelles d'activité et d'achèvement; les classes sémantico-syntaxique F (verbes de type *frapper* ou *toucher*), P (verbes psychologiques), S (verbes de saisie ou de possession), et X (verbes auxiliaires et semi-auxiliaires, impersonnels, *exister*, inchoatifs ou résultatifs); les occurrences où la structure transitive employée n'est pas normée; ainsi que la présence d'une construction *verbe + verbe infinitif*.

La deuxième question spécifique de recherche cherche à savoir si on peut identifier une norme sourde dans l'emploi du SE dans le corpus et, plus particulièrement, les contextes qui influencent la décision d'employer ou non le SE. Néanmoins, les résultats indiquent que les variables observées ne permettent pas de déterminer s'il s'agit d'une norme sourde ou bien d'un emploi aléatoire, puisque les variables contextuelles observées ne permettent pas de caractériser adéquatement la variable de présence ou absence du SE, bien que des contextes soient significativement corrélés avec cette variable. Les principaux contextes caractérisant la présence du SE sont la conjugaison du verbe à un temps simple, l'absence de la construction *verbe + verbe infinitif*, la classe aspectuelle d'état, la structure attributive et celle bitransitive, le rôle thématique patient ou expérienceur pour le sujet, un sujet dont le référent est singulier, l'emploi d'un pronom relatif en sujet, ainsi que les classes sémantico-syntaxiques P (verbes psychologiques) et N (verbes de type *munir / démunir*). Les principaux contextes caractérisant l'absence du SE sont les classes aspectuelles d'activité et d'achèvement, un sujet ayant le rôle thématique agent et l'emploi du pronom *nous* comme sujet.

Les résultats ont permis la formulation d'hypothèses, dont deux principales :

- 1) Plus un contexte contient des éléments pouvant entraîner une plus grande charge cognitive (tels qu'un sujet dont le référent est pluriel, la conjugaison du verbe à un temps composé et la présence d'une construction *verbe + verbe infinitif*), moins le SE est produit et plus il y aura d'écarts à la norme.
- 2) Les scripteurs et scriptrices sourd·e·s produisent davantage de SE dans les contextes qui sont plus associé au concept du *soi*, à la personne et à ses émotions (tels que les changements d'état, les verbes à structure attributive, le verbe *être*, les verbes psychologiques et les sujets ayant le rôle thématique expérienceur).

Une limite importante du présent mémoire est que les variables observées ne touchent que le contexte en français; aucune variable ne considère la LSQ, malgré l'hypothèse qu'elle est la langue de référence des participant·e·s lors de la production des textes. En effet, l'incorporation de variables permettant l'observation de possibles transferts de la LSQ demande une démarche méthodologique dépassant les limites de cette étude. Or, les résultats indiquent que les variables observées ne sont pas suffisantes pour prédire si le SE sera présent ou absent, ce qui pourrait impliquer que la décision à employer un verbe dans une forme pronominale ou non dans une phrase

donnée provient d'un transfert de la phrase correspondante en LSQ. Ainsi, il serait très intéressant, dans une étude future, de traduire des phrases du corpus en LSQ et d'en analyser les structures de marquage de la référence. De plus, l'inclusion de tous les verbes du corpus – même ceux qui ne sont pas précédés du SE et qui ne peuvent l'être – ainsi qu'une comparaison à des corpus comparables en français langue maternelle et en français langue seconde permettraient de dresser un portrait complet de l'emploi des verbes pronominaux en français sourd.

RÉFÉRENCES

- Antidote. (2022a). Mourir. Dans *Dictionnaire Antidote*. Druide Informatique.
- Antidote. (2022b). Se. Dans *Dictionnaire Antidote*. Druide informatique.
- Belzile, T. (1975). *Analyse syntaxique de l'expression écrite chez des enfants présentant des problèmes auditifs*. [Thèse de doctorat, Université du Québec à Montréal].
- Bernieri-Souza, R. et Vasconcelos, C.E.B. (2021). Ensino e aprendizagem de português como segunda língua para surdos. *Trem de Letras*, 8(1), Article e021006-e021006.
- Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press.
- Boivin, M.-C. et Pinsonneault, R. (2020). *La grammaire moderne: description grammaticale du français*. Chenelière éducation.
- Cajolet-Laganière, H., Martel, P., Masson, C.-É. et Mercier, L. (2023) *Usito*. Université de Sherbrooke. <https://usito.usherbrooke.ca/>
- Charrow, V.R. (1975). A Psycholinguistic Analysis of “Deaf English”. *Sign language studies*, 1007(1), 139-150. <https://doi.org/10.1353/sls.1975.0015>
- Cormier, K., Schembri, A. et Woll, B. (2013). Pronouns and pointing in sign languages. *Lingua*, 137, 230-247.
- Crystal, D. (2008). *A dictionary of linguistics and phonetics*. Wiley.
- Daigle, D. (1995). *L'influence du contact avec le français sur la performance des sourds en français écrit* [Mémoire de maîtrise, Université du Québec à Montréal].
- Daigle, D. et Dubuisson, C. (1998). Que peut-on conclure des recherches portant sur l'écriture? Dans C. Dubuisson et D. Daigle (dir.), *Lecture, écriture et surdité: visions actuelles et nouvelles perspectives* (p. 131-151). Les Éditions Logiques.
- Daigle, D., Dubuisson, C. et Germain, C. (1994). « Je travaillent balaie la porpre. » Comment faire pour comprendre? *Revue de l'Association canadienne de linguistique appliquée (ACLA)*, 16(2), 39-48.
- Dictionnaires Le Robert. (2023). *Le Petit Robert - Version numérique*, version 5.8. www.lerobert.com
- Druide informatique. (2022). *Antidote 11*, version 3.1.
- Dubois, J. et Dubois-Charlier, F. (1997). *Les verbes français*. Larousse-Bordas.

- Dubois, J., Giacomo, M., Guespin, L., Marcellesi, C., Marcellesi, J. B. et Mével, J. P. (2002). Norme. Dans *Dictionnaire de linguistique* (p. 330). Larousse-Bordas.
- Dubuisson, C., Miller, C. et Lelièvre, L. (1999). *Grammaire descriptive de la LSQ: Le comportement manuel et le comportement non manuel* (Tome 1, édition revue et augmentée) Université du Québec à Montréal, Groupe de recherche sur la LSQ.
- Dubuisson, C. et Nadeau, M. (1993). *Le français sourd et la LSQ* [Rapport présenté au MESS]. Université du Québec à Montréal, Groupe de recherche sur la LSQ et le bilinguisme sourd.
- Dubuisson, C., Vincent-Durroux, L. et Nadeau, M. (1991). L'enseignement de la langue maternelle aux déficients auditifs. *Glossa*, 32-37.
- Gass, S. (1984). A review of interlanguage syntax: Language transfer and language universals. *Language Learning*, 34(2), 115-132.
- Grevisse, M. et Goosse, A. (2016). *Le bon usage*. De Boeck Supérieur.
- Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and cognition*, 1(2), 131-149.
- Groupe de recherche sur la LSQ et le bilinguisme sourd. (2013). *Banque visuelle de signes dessinés de la LSQ*.
- Hummel, K.M. (2014). *Introducing second language acquisition*. Wiley Blackwell.
- Husson, F., Lê, S. et Pagès, J. (2016). *Analyse de données avec R*. Presses universitaires de Rennes.
- Kassambara, A. (2017a). *ACM - Analyse des correspondances multiples avec R : L'essentiel (Méthodes des composantes principales dans R : Guide pratique)*. STHDA Statistical Tools for High-Throughput Data Analysis. <http://www.sthda.com/french/articles/38-methodes-des-composantes-principales-dans-r-guide-pratique/75-acm-analyse-des-correspondances-multiples-avec-r-l-essentiel/>
- Kassambara, A. (2017b). *AFM - Analyse factorielle multiple avec R : L'essentiel (Méthodes des composantes principales dans R : Guide pratique)*. STHDA Statistical Tools for High-Throughput Data Analysis. <http://www.sthda.com/french/articles/38-methodes-des-composantes-principales-dans-r-guide-pratique/77-afm-analyse-factorielle-multiple-avec-r-l-essentiel/>
- Kassambara, A. et Mundt, F. (2020). Package 'factoextra'.
- Kayne, R.S. (2000). Person morphemes and reflexives in Italian, French, and related languages. Dans *Parameters and universals*. Oxford University Press.
- Labelle, M. (1992). Change of state and valency. *Journal of linguistics*, 28(2), 375-414.

- Laboratoire de recherche appliquée en linguistique informatique. (2020). *Versions informatisées du LVF et du DEM de Jean Dubois et Françoise Dubois-Charlier*. Recherche appliquée en linguistique informatique. <http://rali.iro.umontreal.ca/rali/?q=fr/versions-informatisees-lvf-dem>
- Lacerte, L. (1989). L'écriture sourde québécoise. *Revue québécoise de linguistique théorique et appliquée*, 8(3-4), 303-345.
- Le Petit Robert. (2022). Se. Dans *Dictionnaire Le Petit Robert de la langue française - Édition abonnés* (Éditions Le Robert éd.)
- Lê, S., Josse, J. et Husson, F. (2008). FactoMineR: A package for multivariate analysis. *Journal of Statistical Software*, 25(1), 1-18.
- Liddell, S.K. (1990). Four functions of a locus: Reexamining the structure of space in ASL. Dans C. Lucas (dir.), *Sign language research: Theoretical issues* (p. 176-198). Gallaudet University Press.
- Liddell, S.K. (2000). Indicating verbs and pronouns: Pointing away from agreement. Dans K. Emmorey et H. L. Lane (dir.), *The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima* (p. 303-320). Psychology Press.
- Lillo-Martin, D. et Klima, E.S. (1990). Pointing out differences: ASL pronouns in syntactic theory. Dans S. D. Fischer et P. Siple (dir.), *Theoretical issues in sign language research* (Vol. 1, p. 191-210). The University of Chicago Press.
- Nadeau, M. (1993). Peut-on parler de français sourd? *Revue de l'Association canadienne de linguistique appliquée (ACLA)*, 15(2), 97-117.
- Nadeau, M., Dubuisson, C. et Gélinas-Chebat, C. (1991). Le français écrit des étudiants sourds: une analyse qualitative. *Bulletin de l'ACLA*, 13(1), 89-103.
- Nadeau, M. et Machabée, D. (1998). Dans quelle mesure les erreurs des Sourds sont-elles comparables à celles des entendants? Dans C. Dubuisson et D. Daigle (dir.), *Lecture, écriture et surdité: visions actuelles et nouvelles perspectives* (p. 131-151). Les Éditions Logiques.
- Niederberger, N. et Berthoud-Papandropoulou, I. (2004). Utilisation des pronoms personnels en français écrit par des enfants sourds bilingues. Un parcours spécifique d'apprentissage? *Lidil*, 30(30), 27-38.
- Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge University Press.
- Parisot, A.-M. (2003). *Accord et cliticisation: l'accord des verbes à forme rigide en LSQ*. [Thèse de doctorat, Université du Québec à Montréal].

- Parisot, A.-M. et Rinfret, J. (2008). Description des formes de l'assignation spatiale en langue des signes québécoise. *Silexicales 5, Syntaxe, interprétation, lexicque des langues signées*, 91-101.
- Perini, M. (2013). *Que peuvent nous apprendre les productions écrites des sourds? Analyse de lectures écrites de personnes sourdes pour une contribution à la didactique du français écrit en formation d'adultes*. [Thèse de doctorat, Paris 8].
- Power, D.J. (1971). Deaf children's acquisition of the passive voice. *Journal of Speech and Hearing Research*, 16(1), 5-11.
- Quigley, S., Powers, D. et Steinkamp, M. (1977). Language structure of deaf children. *Volta Review*, 79, 73-84.
- Recanati, C. et Recanati, F. (1999). La classification de Vendler revue et corrigée. *Cahiers Chronos*, 4, 167-184.
- Riegel, M., Pellat, J.-C. et Rioul, R. (2009). *Grammaire méthodique du français*. (4e éd. ent. rev. éd.). Quadrige/PUF.
- Ringbom, H. (2007). *Cross-linguistic similarity in foreign language learning* (Vol. 21). Multilingual Matters.
- Sandler, W. et Lillo-Martin, D. (2006). *Sign language and linguistic universals*. Cambridge University Press.
- Saunders, D. (2016). *Description des structures de représentation corporelle en langue des signes québécoise chez des locuteurs sourds langue première et langue seconde*. [Mémoire de maîtrise, Université du Québec à Montréal].
- Schachter, J. (1983). A new account of language transfer. *Language transfer in language learning*, 2, 98-111.
- Schlenker, P. (2017). Sign language and the foundations of anaphora. *Annual Review of Linguistics*, 3, 149-177.
- Swain, M., Naiman, N. et Dumas, G. (1974). *Aspects of the learning of French by English-speaking five year olds*. Bilingual Education Project, Modern Language Center, Ontario Institute for Studies in Education.
- Tellier, C. (2016). *Éléments de syntaxe du français : méthode d'analyse en grammaire générative*. (3^e édition). Chenelière éducation.
- Usito. (2022). Se. Dans *dictionnaire Usito*. Récupéré de https://usito.usherbrooke.ca/définitions/se_1
- Van Herreweghe, M. et Vermeerbergen, M. (2003). Flemish Sign Language and Dutch: Syntactic differences due to a different modality as exemplified in reference tracking. Dans S.

Granger, J. Lerot et S. Petch-Tyson (dir.), *Corpus-based approaches to contrastive linguistics and translation studies* (p. 91-103) Brill.

Vendler, Z. (1957). Verbs and times. *The philosophical review*, 66(2), 143-160.

Vincent-Durroux, L. (1990). *Problèmes d'acquisition du langage chez des adolescents déficients auditifs profonds oralistes, anglophones et francophones* [Thèse de doctorat, Université de Sherbrooke].

Vincent-Durroux, L. (2009). Deaf languages: does the hypothesis still apply? *Corela. Cognition, représentation, langage*(7-2).

Voghel, A. (2016). *Le système des verbes à classificateur de la langue des signes québécoise*. [Thèse de doctorat, Université du Québec à Montréal].

Yergeau, E. et Poirier, M. (2021) *Test de Chi-2*. SPSS à l'UdeS.
<https://spss.espaceweb.usherbrooke.ca/test-de-chi-2/>