

**LA VARIATION SOCIOLINGUISTIQUE DU FRANÇAIS ÉCRIT PAR DES  
INTERNAUTES AU QUÉBEC DANS L'ESPACE PUBLIC DE FACEBOOK**

FIONA PATTERSON

A DISSERTATION SUBMITTED TO  
THE FACULTY OF GRADUATE STUDIES  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN FRANCOPHONE STUDIES  
YORK UNIVERSITY  
TORONTO, ONTARIO

JUNE 2024

© Fiona Patterson, 2024

## Résumé

Les technologies numériques deviennent un moyen omniprésent de communiquer par écrit et représentent un nouvel usage social de la langue qui ressemble aux conversations orales. Puisque la sociolinguistique a pour but d'étudier la langue dans son contexte social, cette perspective se prête bien à l'observation des échanges dans des espaces numériques. En effet, les sociolinguistes notent l'importance d'étudier cette nouvelle forme langagière, qui contient des éléments des codes écrit et oral, en plus de symboles. Ce nouvel emploi social de la langue interroge la dichotomie oral-écrit, et remet en question les notions de variétés de langue. De plus, les traces écrites d'interactions informelles offrent l'occasion d'accéder au vernaculaire, sans l'intervention du chercheur. Tandis que les recherches variationnistes dans ce domaine pour l'anglais et d'autres langues se multiplient, elles sont moins nombreuses pour l'emploi du français dans le Web 2.0, et plus rares pour le français québécois. Dans ce cadre, cette thèse examine la variation langagière du français québécois à partir d'un corpus original constitué de commentaires écrits sur des pages publiques Facebook par des internautes qui déclarent être situés au Québec. Elle tente de répondre à deux grandes questions : 1. La communication numérique écrite produite sur Facebook par des internautes au Québec ressemble-t-elle plutôt à la langue standard écrite québécoise, ou à la langue orale québécoise ? 2. Quelle est la variabilité des emplois standards et non standards d'un commentaire à l'autre ? Adoptant une approche systématique cette étude examine ainsi l'emploi de huit variables sociolinguistiques du français québécois qui font l'objet d'études depuis des décennies (en l'occurrence des éléments lexicaux, verbaux, syntaxiques et pronominaux). Dans un premier temps, des analyses de concordance permettent de dégager les fréquences d'emploi de ces marqueurs sociolinguistiques, pour ensuite les comparer aux fréquences attestées à l'oral et dans d'autres contextes numériques. Dans un deuxième temps, des analyses de collocation cherchent à déterminer à quel point les internautes emploient des formes non standards dans leurs écrits numériques. Les résultats de cette thèse mènent à une réflexion sur la théorie sociolinguistique et sur l'avenir du français québécois et des langues minoritaires dans l'ère du numérique.

## **Dédicace**

*À la mémoire de mes grands-parents, Otto et Gerda.*

## Remerciements

Je tiens tout d'abord à exprimer ma profonde reconnaissance envers mon directeur de thèse, le Professeur Jerzy Kowal, pour son accompagnement expert, son soutien continu et ses précieux conseils qui m'ont aidé à façonner ce travail de recherche. Il m'a fait confiance, me laissant la liberté de découvrir mon chemin, tout en me guidant à focaliser et à préciser mes idées et mon objet d'étude. Je tiens également à remercier les membres de mon comité de thèse, Mme Sylvie Rosiński Pellerin et Mme Catherine Lamaison. Leur appui, leurs perspectives et leurs commentaires ont contribué énormément à mes réflexions et à la qualité de cette thèse

J'adresse également mes sincères remerciements à M. Raymond Mougeon, Professeur émérite à l'Université York, pour son expertise en sociolinguistique et ses remarques perspicaces qui ont grandement enrichi ce travail. Sous sa direction durant mes études de maîtrise il y a maintenant quinze ans, j'ai eu la chance d'apprendre de manière approfondie ses méthodes de recherche en linguistique variationniste. Cette formation a joué un rôle primordial dans mon orientation et ma démarche pour le présent projet.

Mon collaborateur de recherche M. James Corcoran du département de langues, littératures et linguistique à l'Université York mérite également ma gratitude pour son mentorat inestimable et ses conseils qui ont contribué de manière significative à mon cheminement académique et surtout mes capacités en rédaction scientifique.

À ma famille au Canada et en Allemagne, je suis profondément reconnaissant pour leur soutien, leur compréhension et leur encouragement tout au long de cette aventure académique exigeante. À mes proches et mes amis, qui ont joué un rôle, direct ou indirect, dans la réalisation de cette thèse, je vous remercie de m'avoir toujours encouragé, d'avoir montré une curiosité envers mes recherches, bref, d'avoir été une source de positivité et d'inspiration pendant un processus parfois solitaire. Leur contribution a été essentielle et je leur en suis extrêmement reconnaissante.

## TABLE DES MATIÈRES

<b>Résumé</b> .....	<b>ii</b>
<b>Dédicace</b> .....	<b>iii</b>
<b>Remerciements</b> .....	<b>iv</b>
<b>TABLE DES MATIÈRES</b> .....	<b>v</b>
<b>LISTE DES FIGURES</b> .....	<b>xv</b>
<b>INTRODUCTION</b> .....	<b>1</b>
0.1 Introduction.....	1
0.2 Définition du sujet.....	3
0.3 La communication numérique écrite (CNE) : terminologies.....	5
0.4 Contexte sociolinguistique et culturel.....	7
0.5 Problématique .....	13
0.6 Méthodologie .....	14
0.7 Objectifs et organisation de la thèse .....	15
<b>CHAPITRE 1 : LA COMMUNICATION NUMÉRIQUE ÉCRITE ET LA SOCIOLINGUISTIQUE VARIATIONNISTE</b> .....	<b>17</b>
1.0 Introduction.....	17
1.1 La CNE et l'étude de la langue dans son contexte social .....	17
1.1.1 La CNE et la sociolinguistique variationniste .....	19
1.2 La CNE et la quête du vernaculaire .....	20
1.2.1 L'accès au vernaculaire en contexte numérique .....	21
1.2.1.1 L'étude du français parlé au Québec en contexte numérique .....	22
1.3 La CNE et la question de la norme .....	23
1.3.1 La CNE et le jeu avec la norme .....	24
1.4 Les caractéristiques générales de la CNE .....	24
1.5 Les paramètres communicatifs de la CNE.....	25
1.5.1 Les dispositifs de la CNE.....	26
1.5.2 Les plateformes de la CNE .....	27
1.5.3 Les situations de communication de la CNE .....	28
1.5.4 L'hybridité de la CNE.....	31
1.5.4.1 Exemples de l'hybridité de la CNE .....	32

1.6 La CNE et les dimensions de la variation linguistique .....	33
1.6.1 La CNE et la variation diaphasique .....	33
1.6.1.1 Le contexte numérique et la variation diaphasique.....	34
1.6.1.2 Preuves de la variation diaphasique en contexte numérique .....	35
1.6.2 La CNE et la variation diamésique .....	36
1.6.2.1 La médiatisation technonumérique et la variation diamésique.....	36
1.6.2.2 Les plateformes socionumériques et la variation diamésique.....	37
1.6.2.3 L'hybridité de la CNE comme fonction de la variation diamésique .....	38
1.6.2.3.1 L'oralité comme mesure de l'hybridité.....	39
1.7 Conclusion .....	39
<b>CHAPITRE 2 : ÉTAT DE LA RECHERCHE ET PERTINENCE DU PROJET .....</b>	<b>41</b>
2.0 Introduction.....	41
2.1 Bref historique des études sur la CNE .....	41
2.1.1 Les registres de langue en anglais dans le Web 1.0.....	42
2.2 L'étude sociolinguistique de la CNE dans le Web 2.0 .....	44
2.3 L'étude variationniste de la CNE dans différentes langues .....	44
2.4 L'étude variationniste de la CNE en anglais dans le Web 2.0.....	47
2.4.1 Les messageries instantanées .....	47
2.4.2 Les médias socionumériques .....	47
2.5 L'étude variationniste de la CNE en français dans le Web 2.0 .....	48
2.5.1 Le français européen et québécois dans le (t)chat .....	49
2.5.2 Le français européen et les appareils numériques.....	50
2.5.3 Le français européen dans les SMS .....	51
2.5.4 Les variétés de français dans les médias socionumériques.....	52
2.6 L'étude variationniste du français québécois dans le Web 2.0 .....	53
2.6.1 L'emploi du <i>ne</i> dans le (t)chat .....	54
2.6.2 L'emploi des pronoms dans les SMS.....	54
2.6.3 L'emploi du futur dans les SMS .....	55
2.6.4 L'emploi des formules interrogatives dans les SMS .....	56
2.6.5 Variation et normes dans les SMS .....	56
2.7 Synthèse de l'état de la recherche.....	57
2.8 La pertinence du projet .....	59
2.9 Les questions de recherche .....	59
2.10 Conclusion .....	61

<b>CHAPITRE 3 : MÉTHODOLOGIE : L'ÉTUDE VARIATIONNISTE DE LA LANGUE DANS UN TERRAIN SOCIONUMÉRIQUE .....</b>	<b>63</b>
3.0 Introduction.....	63
3.1 La CNE comme objet d'étude.....	63
3.2 L'espace public de Facebook comme terrain d'étude.....	64
3.3 La linguistique de corpus et la CNE .....	66
3.4 La constitution d'un corpus à partir de la plateforme Facebook .....	67
3.4.1 Les fonctionnalités de la plateforme Facebook .....	69
3.4.1.1 Les espaces de la CNE des internautes sur Facebook.....	69
3.4.1.1.1 Les pages publiques .....	70
3.4.1.1.1.1 Les commentaires des publications sur les pages publiques .....	71
3.5 Le choix des données .....	78
3.6 La situation de communication à l'étude .....	79
3.7 L'extraction, le stockage et l'anonymisation des données.....	81
3.7.1 L'éthique : la sécurité, l'anonymat et le pseudonymat .....	83
3.8 Le contexte variable et la recherche des variantes .....	85
3.8.1 Le logiciel #LancsBox et la concordance .....	86
3.8.2 Le logiciel #LancsBox et la collocation.....	90
3.8.3 L'analyse descriptive statistique des données et le logiciel GoldVarb .....	92
3.8.4 Limites possibles de la méthodologie .....	93
3.9 Résultats attendus et hypothèses .....	94
3.10 Conclusion .....	95
<b>CHAPITRE 4 : RÉSULTATS DES ANALYSES DE CONCORDANCE.....</b>	<b>97</b>
4.0 Introduction.....	97
4.1 Un corpus de CNE québécoise .....	97
4.2 Présentation des résultats .....	99
4.3 Résultats de l'analyse.....	100
4.3.1 Analyse des marqueurs lexicaux .....	100
4.3.1.1 Analyse des marqueurs de contrainte .....	101
4.3.1.1.1 Analyse des données des marqueurs de contrainte .....	102
4.3.1.1.2 Analyse quantitative/statistique descriptive des données des marqueurs de contrainte.....	104
4.3.1.1.3 Comparaison aux fréquences d'emploi attestées à l'oral.....	105
4.3.1.2 Analyse des marqueurs de conséquence .....	108

4.3.1.2.1	Analyse des données des marqueurs de conséquence.....	109
4.3.1.2.2	Analyse quantitative/statistique descriptive des données des marqueurs de conséquence .....	110
4.3.1.2.3	Comparaison aux fréquences d'emploi attestées à l'oral.....	112
4.3.2	Analyse des marqueurs verbaux .....	114
4.3.2.1	Analyse de l'emploi variable du futur .....	115
4.3.2.1.1	Analyse des données de l'emploi variable du futur .....	116
4.3.2.1.2	Analyse quantitative/statistique descriptive des données de l'emploi variable du futur.....	117
4.3.2.1.3	Comparaison aux fréquences d'emploi attestées à l'oral et dans les SMS .....	118
4.3.2.2	Analyse de l'emploi variable des auxiliaires <i>avoir</i> et <i>être</i> au passé composé ....	120
4.3.2.2.1	Analyse des données de l'emploi variable des auxiliaires <i>avoir</i> et <i>être</i> .....	121
4.3.2.2.2	Analyse quantitative/statistique descriptive des données de l'emploi variable des auxiliaires <i>avoir</i> et <i>être</i> .....	122
4.3.2.2.3	Comparaison aux fréquences d'emploi attestées à l'oral.....	125
4.3.3	Analyse des marqueurs syntaxiques .....	128
4.3.3.1	Analyse de l'emploi variable de la particule <i>ne</i> dans la négation.....	129
4.3.3.1.1	Analyse des données de l'emploi variable de la négation .....	131
4.3.3.1.2	Analyse quantitative/statistique descriptive des données de l'emploi variable de la négation.....	133
4.3.3.1.3	Comparaison aux fréquences d'emploi attestées à l'oral et dans le (t)chat .....	134
4.3.3.2	Analyse de l'emploi variable des formules interrogatives.....	136
4.3.3.2.1	Analyse des données de l'emploi variable des formules interrogatives .....	137
4.3.3.2.2	Analyse quantitative/statistique descriptive des données de l'emploi variable des formules interrogatives .....	138
4.3.3.2.3	Comparaison aux fréquences d'emploi des formules interrogatives attestées à l'oral.....	140
4.3.4	Analyse des marqueurs pronominaux .....	142
4.3.4.1	Analyse de l'emploi variable des pronoms <i>on</i> et <i>nous</i> .....	143
4.3.4.1.1	Analyse des données de l'emploi variable des pronoms <i>on</i> et <i>nous</i> .....	144
4.3.4.1.2	Analyse quantitative/statistique descriptive des données sur l'emploi variable des pronoms <i>on</i> et <i>nous</i> .....	144
4.3.4.1.3	Comparaison aux fréquences d'emploi attestées à l'oral et dans les SMS .....	146
4.3.4.2	Analyse de l'emploi variable des pronoms simples et composés non clitiques .	148
4.3.4.2.1	Analyse des données des pronoms simples et composés non clitiques .....	149

4.3.4.2.2 Analyse quantitative/statistique descriptive des données des pronoms simples et composés non clitiques .....	150
4.3.4.2.3 Comparaison aux fréquences d'emploi attestées à l'oral et dans les SMS .....	151
4.4 Conclusion .....	153
<b>CHAPITRE 5 : RÉSULTATS DES ANALYSES DE COLLOCATION .....</b>	<b>155</b>
5.0 Introduction .....	155
5.1 Présentation des résultats de collocation.....	155
5.2 Cooccurrence de formes non standards avec les variables à quatre variantes .....	156
5.2.1 Cooccurrences de formes non standards et des marqueurs de contrainte .....	156
5.2.2 Cooccurrences de formes non standards et des marqueurs de conséquence .....	159
5.2.3 Cooccurrences de formes non standards et des formules interrogatives .....	161
5.3 Cooccurrences de formes non standards et des variables à deux variantes .....	164
5.3.1 Cooccurrences de formes standards et non standards dans les occurrences du futur .....	164
5.3.2 Cooccurrences de formes standards et non standards de la négation .....	167
5.4 Conclusion .....	169
<b>CHAPITRE 6 - DISCUSSION DES RÉSULTATS .....</b>	<b>171</b>
6.0 Introduction.....	171
6.1 Synthèse et discussion des résultats de concordance .....	171
6.2 Synthèse et discussion des résultats de collocation .....	177
6.3 Un corpus québécois .....	180
6.3.1 Le numérique et les langues minoritaires .....	181
6.3.2 Le numérique comme lieu d'échange linguistico-culturel.....	182
6.3.3 Le numérique comme lieu de valorisation du parler régional .....	183
6.4 Les axes de la variation linguistique .....	184
6.4.1 Le numérique et la question de la diamésie .....	184
6.4.2 La médiatisation technologique et la variation linguistique .....	186
6.5 Limites de l'étude .....	189
6.6 Conclusion .....	191
<b>CONCLUSION .....</b>	<b>193</b>
Conclusions générales.....	193
Pistes futures de recherche.....	196
<b>Bibliographie .....</b>	<b>201</b>

**Annexe : Bilan des résultats d'études antérieures ..... 245**

## LISTE DES TABLEAUX

Figure 0.4 Les différences situationnelles entre trois modalités numériques .....	9
Figure 0.4.2 Activité sur les réseaux socionumériques au Québec.....	12
Tableau 1.5.3 - Schéma de classification de la situation de communication de CNE .....	30
Figure 3.4.1.1.1 - Publication sur une page publique .....	71
Figure 3.4.1.1.1.1 - La boîte de saisie de texte pour les commentaires publics.....	72
Figure 3.4.1.1.1.2 - Les commentaires publics sur une publication .....	74
Figure 3.4.1.1.1.3 - La modification et la suppression des commentaires.....	75
Figure 3.4.1.1.1.4 - La visualisation des commentaires.....	76
Figure 3.4.1.1.1.5 - Les options de visualisation des commentaires .....	77
Figure 3.4.1.1.1.6 L'ajout d'éléments non textuels aux commentaires .....	77
Tableau 3.6.1 Les caractéristiques sociales .....	79
Tableau 3.6.2 Les caractéristiques technologiques .....	80
Tableau 3.7 - Les statistiques du corpus .....	83
Figure 3.8.1 L'écran principal du logiciel #LancsBox .....	87
Figure 3.8.1.1 L'écran KWIC du logiciel #LancsBox.....	88
Tableau 3.8.1.1 Échantillon d'occurrences stockées et codées .....	89
Figure 3.8.2 L'écran Text du Logiciel #LancsBox.....	91
Tableau 3.8.2 Échantillon de commentaires stockées et codées pour la collocation.....	92
Tableau 4.2 - Total des occurrences dans le corpus par variable à l'étude.....	100
Tableau 4.3.1.1a - Exemples de l'emploi des marqueurs de contrainte dans le corpus.....	102
Tableau 4.3.1.1b - Graphies cherchées des marqueurs de contrainte .....	102
Tableau 4.3.1.1.2 - Analyse quantitative/statistique descriptive des données des marqueurs de contrainte.....	104
Tableau 4.3.1.1.3 - Comparaison des fréquences d'emploi des marqueurs de contrainte attestées à l'oral et sur Facebook.....	106
Figure 4.3.1.1.3 - Comparaison des fréquences d'emploi des marqueurs de contrainte attestées à l'oral et sur Facebook.....	106
Tableau 4.3.1.2a - Exemples de l'emploi des marqueurs de conséquence dans le corpus .....	109
Tableau 4.3.1.2b - Graphies cherchées des marqueurs de conséquence.....	109

Tableau 4.3.1.2.2 - Analyse quantitative/statistique descriptive des données des marqueurs de conséquence .....	110
Tableau 4.3.1.2.3 - Comparaison des fréquences d'emploi des marqueurs de conséquence attestées à l'oral et sur Facebook .....	112
Figure 4.3.1.2.3 - Comparaison des fréquences d'emploi des marqueurs de conséquence attestées à l'oral et sur Facebook.....	113
Tableau 4.3.2.1a - Exemples de l'expression du futur dans le corpus .....	115
Tableau 4.3.2.1b - Graphies cherchées de l'expression du futur .....	116
Tableau 4.3.2.1.2 - Analyse quantitative/statistique descriptive des données de l'emploi variable du futur.....	117
Tableau 4.3.2.1.3 - Comparaison des fréquences d'emploi du futur attestées à l'oral, dans les SMS et sur Facebook .....	118
Figure 4.3.2.1.3 - Comparaison des fréquences d'emploi du futur attestées à l'oral, dans les SMS et sur Facebook .....	119
Tableau 4.3.2.2a - Exemples de l'emploi du passé composé dans le corpus .....	121
Tableau 4.3.2.2b - Graphies cherchées du passé composé .....	121
Tableau 4.3.2.2.2a - Analyse quantitative/statistique descriptive des données de l'emploi variable de l'auxiliaire avoir .....	123
Tableau 4.3.2.2.2b - Analyse quantitative/statistique descriptive des données de l'emploi variable de l'auxiliaire être .....	123
Tableau 4.3.2.2.3 - Comparaison des fréquences d'emploi de l'auxiliaire avoir avec les verbes intransitifs à l'oral et sur Facebook.....	126
Figure 4.3.2.2.3 - Comparaison des fréquences d'emploi de l'auxiliaire avoir avec les verbes intransitifs à l'oral et sur Facebook.....	127
Tableau 4.3.3.1a - Exemples de l'emploi de la négation dans le corpus .....	130
Tableau 4.3.3.1b - Graphies cherchées de la négation.....	131
Tableau 4.3.3.1.2 - Analyse quantitative/statistique descriptive des données de l'emploi variable de la négation, fragments exclus.....	134
Tableau 4.3.3.1.3 - Comparaison des fréquences d'emploi attestées à l'oral, dans le (t)chat, et sur Facebook.....	135

Figure 4.3.3.1.3 - Comparaison des fréquences d'emploi attestées à l'oral et dans le (t)chat, et sur Facebook.....	135
Tableau 4.3.3.2 - Exemples de l'emploi des formules interrogatives dans le corpus.....	137
Tableau 4.3.3.2.2 - Analyse quantitative/statistique descriptive des données de l'emploi variable des formules interrogatives.....	138
Tableau 4.3.3.2.3 - Comparaison des fréquences d'emploi des formules interrogatives à l'oral, dans les SMS et sur Facebook.....	140
Figure 4.3.3.2.3 - Comparaison des fréquences d'emploi des formules interrogatives à l'oral, dans les SMS et sur Facebook.....	141
Tableau 4.3.4.1a - Exemples de l'emploi des pronoms on et nous dans le corpus.....	143
Tableau 4.3.4.1b - Graphies cherchées des pronoms on et nous.....	143
Tableau 4.3.4.1.2 - Analyse quantitative/statistique descriptive des données sur l'emploi variable des pronoms on et nous.....	144
Tableau 4.3.4.1.3 - Comparaison des fréquences d'emploi des pronoms on et nous attestées à l'oral, dans les SMS et sur Facebook.....	146
Tableau 4.3.4.2a - Exemples de l'emploi des pronoms pluriels non clitiques dans le corpus....	148
Tableau 4.3.4.2b - Graphies cherchées des pronoms pluriels non clitiques.....	149
Tableau 4.3.4.2.2.a - Analyse quantitative/statistique descriptive des données de l'emploi variable des pronoms nous et nous-autres.....	150
Tableau 4.3.4.2.2.b - Analyse quantitative/statistique descriptive des données de l'emploi variable des pronoms vous et vous-autres.....	150
Tableau 4.3.4.2.2.c - Analyse quantitative/statistique descriptive des données de l'emploi variable des pronoms eux et eux-autres.....	151
Tableau 4.3.4.2.2.d - Analyse quantitative/statistique descriptive des données de l'emploi variable des pronoms simples et composés non clitiques.....	151
Tableau 4.3.4.2.3 - Comparaison des fréquences d'emploi des pronoms non clitiques simples et composés à l'oral, dans les SMS et sur Facebook.....	151
Figure 4.3.4.2.3 - Comparaison des fréquences d'emploi des pronoms non clitiques simples et composés à l'oral, dans les SMS et sur Facebook.....	152
Tableau 5.2.2 Cooccurrence des marqueurs de conséquence avec éléments non standards.....	160
Tableau 5.2.3 Cooccurrence des formules interrogatives avec éléments non standards.....	161

Tableau 5.3.1 Cooccurrence du futur avec éléments non standards .....	164
Tableau 5.3.2 Cooccurrences des variantes de la négation.....	167
Tableau 6.1 Les fréquences d'emploi de cinq variables à l'étude .....	172
Figure 6.1 Graphique synthétique des fréquences d'emploi dans trois contextes .....	173
Tableau 6.1.1 Les variables à l'étude selon leur longueur et leur fréquence d'emploi .....	176
Tableau 6.2 Pourcentage d'occurrences accompagnées d'éléments non standards .....	178
Figure 6.2 Pourcentage d'occurrences accompagnées d'éléments non standards .....	178
Tableau 6.4.2 Les caractéristiques sociales et technologiques de la CNE.....	187
Figure 6.4.2 Les axes de la variation linguistique .....	188
Figure 6.4.2.1 Les axes de la variation linguistique dans l'ère du numérique.....	188

## LISTE DES FIGURES

Figure 0.4 Les différences situationnelles entre trois modalités numériques .....	9
Figure 0.4.1 Utilisation des réseaux socionumériques au Québec.....	11
Figure 0.4.2 Activité sur les réseaux socionumériques au Québec.....	12
Figure 3.4.1.1.1 - Publication sur une page publique .....	71
Figure 3.4.1.1.1.1 - La boîte de saisie de texte pour les commentaires publics.....	72
Figure 3.4.1.1.1.2 - Les commentaires publics sur une publication .....	74
Figure 3.4.1.1.1.3 - La modification et la suppression des commentaires.....	75
Figure 3.4.1.1.1.4 - La visualisation des commentaires.....	76
Figure 3.4.1.1.1.5 - Les options de visualisation des commentaires .....	77
Figure 3.4.1.1.1.6 L'ajout d'éléments non textuels aux commentaires .....	77
Figure 3.8.1 L'écran principal du logiciel #LancsBox .....	87
Figure 3.8.1.1 L'écran KWIC du logiciel #LancsBox .....	88
Figure 3.8.2 L'écran Text du Logiciel #LancsBox.....	91
Figure 4.3.1.1.3 - Comparaison des fréquences d'emploi des marqueurs de contrainte attestées à l'oral et sur Facebook.....	106
Figure 4.3.1.2.3 - Comparaison des fréquences d'emploi des marqueurs de conséquence attestées à l'oral et sur Facebook.....	113
Figure 4.3.2.1.3 - Comparaison des fréquences d'emploi du futur attestées à l'oral, dans les SMS et sur Facebook .....	119
Figure 4.3.2.2.3 - Comparaison des fréquences d'emploi de l'auxiliaire avoir avec les verbes intransitifs à l'oral et sur Facebook.....	127
Figure 4.3.3.1.3 - Comparaison des fréquences d'emploi attestées à l'oral et dans le (t)chat, et sur Facebook.....	135
Figure 4.3.3.2.3 - Comparaison des fréquences d'emploi des formules interrogatives à l'oral, dans les SMS et sur Facebook .....	141
Figure 4.3.4.1.3 - Comparaison des fréquences d'emploi des pronoms on et nous attestées à l'oral, dans les SMS et sur Facebook.....	147
Figure 4.3.4.2.3 - Comparaison des fréquences d'emploi des pronoms non clitiques simples et composés à l'oral, dans les SMS et sur Facebook .....	152

Figure 6.1 Graphique synthétique des fréquences d'emploi dans trois contextes .....	173
Figure 6.2 Pourcentage d'occurrences accompagnées d'éléments non standards .....	178
Figure 6.4.2 Les axes de la variation linguistique .....	188
Figure 6.4.2.1 Les axes de la variation linguistique dans l'ère du numérique.....	188

## INTRODUCTION

« La mer n'a pas de structure avant que quelqu'un ne la parcoure : c'est la navigation qui donne un sens et met en rapport des points différents. L'espace virtuel est un espace qui se structure au fur et à mesure qu'on y navigue. »  
(Vitali-Rosati, 2014 : 4.2.3)

### 0.1 Introduction

Il va sans dire que la technologie, et en particulier le numérique, ont bouleversé nos manières de communiquer. Par exemple, l'imprimerie de Gutenberg a changé la trajectoire du partage d'informations et des connaissances, et celle de la littérature et des langues vernaculaires de l'Europe. Pareillement, les avancées dans les technologies numériques au cours des dernières décennies ont radicalement transformé les habitudes sociales, y compris la communication. Au passé, l'imprimerie de Gutenberg « a influé sur d'autres – un prototype pour la production de masse – et [...] a eu un impact direct sur le monde des idées en rendant largement accessible les savoirs et en créant un espace dans lequel de nouvelles formes d'expression pouvaient florir » (Crowley et Heyer, 2011 : 64, ma traduction). Dans le même ordre d'idées, de nos jours, le numérique mène à « la désintégration de la distinction masse/interpersonnel, la résurgence de l'amateur, la modularité et l'opacité des logiciels, la forme fluide de la connaissance en réseau et les pratiques de signification sociale connectées latéralement » (Gillespie *et al.*, 2014 : 5, ma traduction). En d'autres mots, les technologies numériques ont bouleversé la structure-même que prend la communication.

Alors que ces deux technologies, l'imprimerie de Gutenberg et les technologies numériques, ont eu un effet de décentralisation (Bernstein, 2013 : 158 ; Gillespie *et al.*, 2014 : 5) et de démocratisation sur l'accès à l'information, le numérique opère ce même effet sur la communication elle-même (van Dijck, 2013 : 10). Plus précisément, Internet, création plutôt sociale que technique (Crystal, 2001 : 1), « a été fondée empiriquement à partir d'expérimentations sociales, en opposition au modèle hiérarchique vertical » (Perriault, 2012 : 153) permettant une socialisation virtuelle à travers laquelle les internautes sont à la fois récipiendaires et consommateurs, producteurs et participants de la culture (van Dijck, 2013 : 32-33). Le contenu des médias sociaux, par exemple, est donc « moins réglementé, plus fluide et plus diversifié »

(Sergeant and Tagg, 2014 : 4) que des moyens de communication traditionnels. Le numérique efface donc les limites du temps et de l'espace, comme l'avait prédit Marshall McLuhan (1994), permettant des échanges sans bornes.

En même temps, les technologies numériques nous éloignent physiquement les un.e.s des autres, mettant en place comme médiateur une machine. Étant de la génération qui est passée de l'analogique au numérique, et qui a grandi en parallèle aux développements dans les technologies de la communication, j'ai toujours été fascinée par l'observation des changements sociétaux que je vivais en temps réel. Ces technologies, qui ont transformé la communication entre individus, ouvrent également une boîte de pandore en ce qui concerne les questions sur les enjeux sociologiques, linguistiques, psychologiques, politiques, pédagogiques, économiques, etc. Par exemple, les réseaux socionumériques en particulier « forment une nouvelle dimension à travers laquelle les gens organisent leurs vies ». Ces « plateformes influe[nt] sur l'interaction humaine, sur le plan individuel et communautaire, ainsi que sur le plan social, tandis que les mondes du numérique et du hors numérique deviennent de plus en plus entremêlés » (van Dijck, 2013 : 2, ma traduction). Quel est donc l'impact sur nos relations sociales, sur nos habitudes culturelles et sur notre comportement linguistique, de cette médiatisation des interactions humaines par les machines ? Comment cette culture numérique opère-t-elle sur nos vies ?

Cette thèse s'inspire ainsi d'un livre paru en 2019 de la linguiste canadienne Gretchen McCulloch, intitulé *Because Internet : Understanding the new rules of language*, et des propos du penseur canadien Marshall McLuhan (1994) dans *Understanding media : The extensions of man*. La variété de français parlée aujourd'hui au Québec, alors qu'elle est en situation minoritaire au Canada, témoigne d'une richesse historique et culturelle. Étant donné ce contexte unique, j'ai donc voulu examiner l'état actuel de cette variété de langue dans la tradition variationniste des recherches sur le français québécois. J'ai voulu observer son évolution et sa progression dans l'ère du numérique, tout en explorant les enjeux discutés par McCulloch et McLuhan. En même temps, j'ai voulu appliquer la théorie et la méthode de la sociolinguistique variationniste, qui ont le mérite de fournir une approche quantifiable et ainsi assez objective.

## 0.2 Définition du sujet

Les technologies numériques de la communication ont ouvert un nouveau champ d'étude et de recherche pour les sciences humaines et sociales, en particulier pour la sociolinguistique, qui se veut d'examiner l'emploi de la langue dans son contexte social. Comme le notent Christine Develotte et Marie-Anne Paveau, « [l]a plupart des communications et relations humaines s'élaborent désormais [...] dans les dispositifs numériques [...] à partir d'appareils connectés » (2017 : 160). Laetitia Bibié-Emerit le souligne elle aussi : « cette offre "pluri-communicationnelle" proposée par les supports de communication du Web 2.0 plonge [l'internaute] dans une nouvelle ère discursive » (2015 : 44-45). Ceci s'explique d'autant plus que la communication par mode virtuel devient omniprésente, surtout depuis le début de la pandémie COVID en 2020, qui a bouleversé les habitudes sociales. Les réseaux socionumériques représentent donc une nouvelle sphère d'activité à travers laquelle les gens organisent leurs vies.

Ce sont ainsi les communications et relations humaines de cette nouvelle ère discursive qui encadrent la présente recherche. Cette analyse du français québécois (abrégié *FQ*) écrit dans l'espace public de la plateforme Facebook (abrégié *FB*) se situe en effet au carrefour des cadres théoriques de l'étude du numérique et de la sociolinguistique variationniste. À l'instar de Michel Marcochia (2016), j'emploie la terminologie « Communication numérique écrite » (abrégié *CNE*) pour désigner toute production écrite par le biais de la technologie et dans l'espace du Web 2.0.

L'évolution du numérique est pertinente pour l'étude de l'emploi de la langue, puisque cette technologie, qui était au départ un simple réseau permettant le partage d'information, est devenu un lieu social et relationnel. L'opposition pertinente ici est celle entre le Web 1.0, constitué de sites Web, et le Web 2.0, constitué de plateformes de communication et de réseautage interactives. Alors que le Web 1.0 est paru en premier, notons que ces deux réseaux existent maintenant simultanément. Plus précisément,

les médias réseautés du Web 1.0 représent[ent] des services auxquels on [peut] s'inscrire pour créer des groupes et recevoir des nouvelles, mais ne permett[ent] pas la connexion automatique aux autres. En revanche, au tournant du millénaire, les services du Web 2.0 sont devenus des véhicules interactifs et bidirectionnels pour la socialité en réseau (van Dijck, 2013 : 5, ma traduction).

Ainsi, le Web 2.0, comparativement au Web 1.0 qui offre aux internautes un accès passif aux sites Web, représente une technologie de réseautage engendrant une « culture participative » (Cardon, 2019). Le Web 1.0, constitué de sites Web statiques, est principalement une source de recueil d'informations produites par des professionnels, pour consommation par des internautes qui ne peuvent que naviguer, lire et imprimer. En effet, le Web 1.0 « est devenu la grande encyclopédie universelle, où on trouve tous les renseignements désirés (plus ou moins vérifiés !). Il garantit la libre circulation de l'information » (Marchessault, 2014 : 34). Les sites Web constituant le Web 1.0 ne permett(ai)ent donc pas aux internautes d'échanger entre eux, mais seulement un accès statique à des renseignements. Maintenant, le Web 2.0, constitué d'applications qui permettent la participation, telles que le courriel, la messagerie et les réseaux sociaux, encourage plutôt la contribution des internautes (Androustopoulos, 2010 : 207). Cette technologie représente ainsi « l'avènement du *writable Web* communément appelé Web contributif » ce qui veut dire que « [l']individu devient [...] capable d'émission en plus de la réception » (Quoniam et Boutet, 2008 : 133). En somme, grâce au Web 2.0, les internautes peuvent depuis un certain temps s'engager dans des conversations informelles à l'écrit, représentant une nouvelle forme discursive sans précédent congruent. Cette transition du Web 1.0 au Web 2.0 représente donc un moment tournant dans les activités sociales et communicatives au 21<sup>e</sup> siècle, et c'est la tâche de la sociolinguistique d'étudier la langue dans tous ses contextes.

En effet, selon le linguiste anglais David Crystal, la discipline de la linguistique se doit d'étudier la langue utilisée dans le numérique :

là, où nous trouvons la langue, nous trouvons des linguistes. C'est le rôle des linguistes : de trouver, de décrire et d'analyser les manifestations de la langue partout. Il s'en suit que lorsque nous rencontrons la plus grande base de données de productions langagières que le monde n'ait jamais vue, on s'attendrait à retrouver des linguistes qui l'explorent, pour voir ce qui se passe. Cela a déjà commencé. Par conséquent, un nouveau domaine émerge : la linguistique d'Internet (Crystal, 2011 : 11, ma traduction).

Ainsi, le rôle de la linguistique est d'aller là où se situe la langue, afin de l'étudier. Crystal met l'accent sur Internet comme lieu d'enquête, et la linguistique d'Internet a donc pour but l'analyse synchronique du langage dans tous les lieux d'activité dans le numérique (Crystal, 2005).

Si la langue employée sur les réseaux socionumériques est un sujet d'intérêt pour domaine de la (socio)linguistique, c'est que cette technologie permet de « nouveaux modes et canaux de communication » (Marchal, 2014 : 4), tout en faisant ressortir une quatrième dimension de la communication, celle du médium électronique ou numérique (Crystal, 2011 : 16). En effet, cet emploi de la langue démontre « une grande capacité d'innovation linguistique » ainsi qu'« un langage abrégé et phonétique » (Marchal, 2014 : 4). Les écrits numériques s'avèrent donc un sujet d'étude pertinent, surtout dans le contexte des variétés et des registres de langue.

### **0.3 La communication numérique écrite (CNE) : terminologies**

Les technologies modernes de la communication permettent une grande diversité de types de productions langagières, tant à l'écrit qu'à l'oral. Il importe donc de rappeler que cette étude porte plus particulièrement sur la *communication numérique écrite* (CNE), c'est-à-dire sur la production de messages écrits par le biais d'un appareil ou d'un moyen technologique et numérique (dorénavant *technonumérique*) tel que l'ordinateur, le téléphone mobile et la tablette électronique. J'adopte le terme technonumérique pour désigner toutes les technologies de la communication qui ont une composante numérique, afin de les distinguer d'autres technologies de la communication qui sont analogiques, telles que le téléphone traditionnel (non cellulaire), le télégraphe, ou moins interactives tels que le télécopieur, etc.

La CNE englobe ainsi toute production langagière écrite diffusée par une diversité de plateformes *socionumériques* du Web 2.0, telles que le SMS, le courriel, le blogue, les médias socionumériques, le forum, le (t)chat ou la messagerie instantanée (Marriccia, 2016). J'adopte le terme socionumérique pour désigner les plateformes du Web 2.0 permettant de l'interaction sociale, afin de distinguer la plateforme du dispositif ou de l'appareil technologique, et *média socionumérique* pour décrire toute plateforme de réseautage social, comme Twitter, Facebook, Instagram, etc. On pourrait parler aussi de *réseau social* ou de *média social*.

Ce choix terminologique exclut donc toute communication diffusée sur le Web 1.0, c'est-à-dire les sites Web, ou encore les médias numériques de masse, car la CNE « réfère à des situations de communication interpersonnelle » (Marcoccia, 2016 : 7). D'autre part, elle ne s'applique pas aux communications numériques non écrites, telles que les communications synchrones ou asynchrones audiovisuelles, par exemple, des communications multimodales accomplies par le biais de plateformes telles que Skype, Zoom, ou par le moyen d'enregistrement audio envoyé par l'intermédiaire d'un dispositif technonumérique.

La terminologie CNE est préférable à l'option CMO, *communication médiatisée par ordinateur* – adaptation française du CMC ou « *computer mediated communication* » de Susan Herring (1996) –, puisque le concept CNE est à la fois plus large et plus précis que celui de CMO, qui se limite aux productions langagières transmises seulement par ordinateur, et exclut d'autres moyens technonumériques. Dans la présente étude, je m'intéresse à la CNE produite par des internautes, et non par d'autres sources telles que des organismes, des institutions, ou des influenceurs. Ces internautes sont donc des individus appartenant à une communauté linguistique, c'est-à-dire des locuteurs et locutrices qui communiquent par écrit dans le Web 2.0.

Pour cerner la CNE, Michel Marcoccia souligne l'importance de tenir compte des « très nombreuses situations [...] “avec de l'écrit” à l'occasion desquelles s'échangent à la fois des textes et des images » (Marcoccia, 2016 : 16). Textes et images parsèment en effet fréquemment les conversations dans les médias socionumériques ou par messagerie instantanée. Par ailleurs, la CNE inclut les symboles et icônes présentes dans la production écrite, tels que les émoticônes, les émojis, les chiffres, et les marques de ponctuation, c'est-à-dire les signes non alphabétiques.

#### 0.4 Contexte sociolinguistique et culturel

L'emploi de la langue en ligne par les internautes fait l'objet d'étude en sociolinguistique depuis déjà trois décennies. Au début, ces recherches se focalisaient principalement sur l'anglais (Collot et Belmore, 1996 ; Ferrara *et al.*, 1991 ; Paolillo, 2001 ; Tagliamonte et Denis, 2008 ; Thurlow, 2003 ; Yates, 1996). Plus récemment elles s'intéressent aussi à d'autres langues et variétés de langues comme l'allemand (Androutsopoulos, 2007 ; Storrer, 2014), le gallois (Cunliffe, 2007 et 2009), les langues africaines (Deumert et Vold Lexander, 2013), le japonais et le coréen (Fouser, Inoue et Lee, 2000), et l'arabe (Palfreyman et al Khalil, 2003), ainsi qu'à celles de la francophonie (Atifi, 2007 ; Cougnon, 2008 ; Marcoccia, 2012 ; Panckhurst, 1998 ; Stark, 2012 et 2014 ; van Compernelle et Williams, 2007). L'étude de la CNE dans les variétés de français au sein de la francophonie est donc un domaine florissant et prometteur.

Mais pourquoi s'intéresser à la variété du français parlée dans la province du Québec ? Les études menées depuis plusieurs décennies sur le sujet par les sociolinguistes (Dessureault-Dober, 1974 ; Laberge, 1977 ; Mougeon, 1976 ; Sankoff et Thibault, 1977 ; Sankoff et Vincent, 1977) ont surtout des motivations historiques et culturelles. Tout d'abord, l'histoire de cette variété de langue est particulière : c'était au début un mélange de variétés du français et de patois des premiers colons arrivés en Nouvelle-France. Subséquemment, tandis qu'en 1759 il y a eu unification de ces variétés de langue, de l'autre côté de l'Atlantique, l'unification s'est faite plus tard autour d'une variété de langue différente (Martel, 2017). Par conséquent, le français en Nouvelle-France « étant complètement isolé de la mère patrie, a évolué différemment du français de France » (Martel et Cajolet-Laganière, 2007 : 1). Ces facteurs historiques, tout comme le contact avec l'anglais, langue hégémonique, ont mené à la variété du français qui est parlée aujourd'hui dans la province du Québec.

Le français parlé au Québec fait ainsi preuve d'écarts avec le français standard ou de référence, étant marqué par « une accumulation de conservatismes [et d'archaïsmes] de France et d'innovations québécoises [...] au niveau du lexique et de la prononciation » (Martel et Cajolet-Laganière, 1996 : 70). Par ailleurs, « la situation linguistique du Québec est unique à l'intérieur de la francophonie : le Québec est le seul endroit en dehors de l'Europe [et les territoires français d'outre-mer], où le français est la langue maternelle de la population » (Martel, 2017 : 16). Ainsi, de manière traditionnelle, le français parlé au Québec est une variété langagière qui présente à l'oral des traits linguistiques d'intérêt pour les sociolinguistes. Dans l'ère du numérique, la question se pose donc : comment cette variété de langue distinctive est-elle transposée à l'écrit par les individus dans des plateformes socionumériques ?

Dans son *Plan stratégique 2018-2023*, l'Office québécois de la langue française (OQLF) note que les nouvelles technologies de la communication « ont des répercussions importantes dans le paysage linguistique québécois » (Gouvernement du Québec, 2020 : 2). C'est pourquoi il vise à effectuer des études « sur les usages linguistiques », y compris dans l'espace public afin de « dresser l'état de la situation linguistique au Québec et d'en apprécier l'évolution par rapport aux données antérieures » (Gouvernement du Québec, 2020 : 8). L'objet de la présente étude s'aligne alors avec celui de l'organisme qui a pour mandat la surveillance du français québécois.

Pour ce qui est de la variation dans la CNE produite par des internautes au Québec, à date, seulement un petit nombre de recherches se sont concentrées sur l'emploi de certains marqueurs sociolinguistiques. Rémi van Compernelle (2010) s'est notamment penché sur l'étude de l'emploi variable de la négation dans le (t)chat. Ensuite, l'équipe d'Hélène Blondeau et Mireille Tremblay (2014 et 2019) a étudié l'emploi des pronoms et du futur dans les SMS, ainsi que l'étude plus générale de la variation et des normes communautaires dans les SMS (2022). Enfin, Anne-Sophie Bally (2022) s'est focalisée sur l'étude des formules interrogatives dans les SMS. Le manque de recherches portant sur le français québécois en ligne, comparativement à la quantité d'études portant sur l'emploi de l'anglais, ou encore sur celui des variétés du français hexagonal, représente donc l'une des raisons d'être de la présente étude.

Le français québécois dans la CNE est en outre un sujet fort pertinent en termes des effets éventuels du numérique sur le changement linguistique. Suivant l'approche de l'équipe de recherche de Raymond Mougeon (Martineau et Mougeon, 2003 et Mougeon, 2013) qui s'est penchée sur des textes anciens pour étudier l'histoire du français parlé au Québec et au Canada, ma recherche, quant à elle, se penche sur des textes numériques pour accéder au vernaculaire contemporain. En effet, alors que d'habitude, « les textes écrits ont l'inconvénient de ne pas toujours refléter la langue parlée, lieu privilégié des changements » (Poplack et St-Amant, 2009 : 511), comme le notent Hélène Blondeau et Mireille Tremblay, « il est fort probable que l'on assiste à un déplacement des pratiques scripturales informelles vers d'autres plateformes tout aussi propices à l'expression vernaculaire » (Blondeau et Tremblay, 2022b : 141). Puisque les cinq recherches susmentionnées portant sur le français québécois en ligne traitent uniquement des modalités du (t)chat et du SMS – sans mentionner le fait que deux des cinq recherches datent déjà de plus de dix ans – la présente étude pose donc un regard sur les activités linguistiques dans l'espace public de la plateforme Facebook. Il s'agit notamment de commentaires sur des publications par cinq pages de médias de presse québécoises. Non seulement cette modalité n'a pas encore été étudiée dans le cadre de la variété du français employée au Québec, mais il existe aussi des différences situationnelles et technologiques entre le (t)chat, le SMS et l'espace public de Facebook. La Figure 0.4 ci-dessous présente quelques-unes des distinctions entre ces trois modalités :

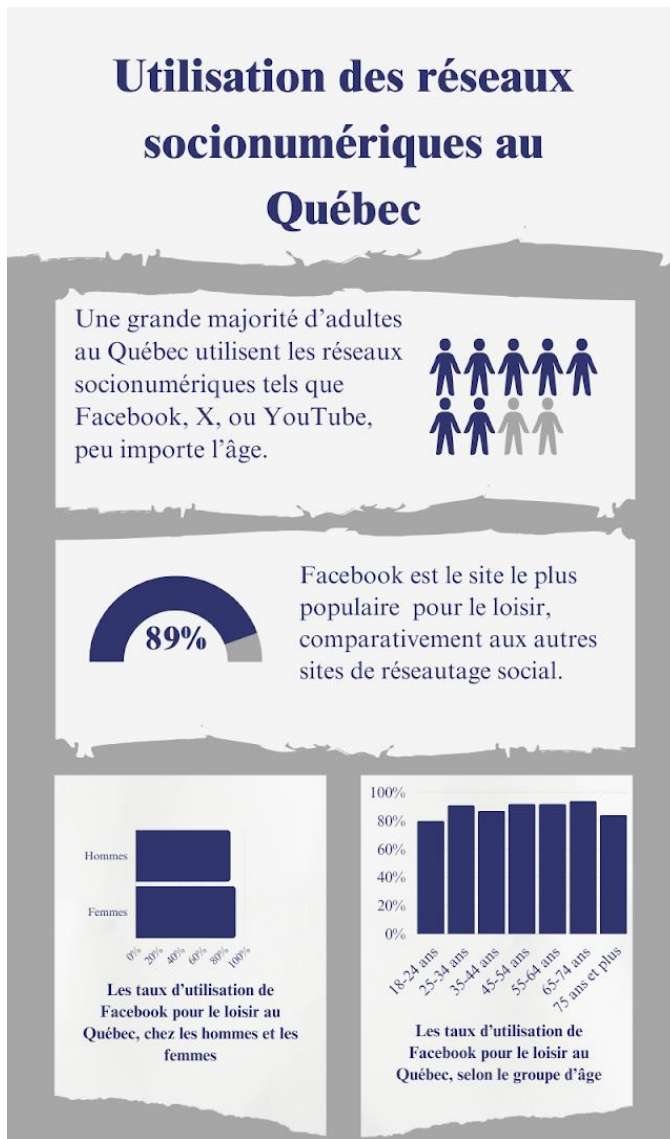
**Figure 0.4 Les différences situationnelles entre trois modalités numériques**

(t)chat	SMS	Facebook
<ul style="list-style-type: none"> <li>• public</li> <li>• anonyme</li> <li>• taille illimitée</li> <li>• synchrone</li> </ul>	<ul style="list-style-type: none"> <li>• privé</li> <li>• individu connu</li> <li>• taille limitée</li> <li>• synchrone/asynchrone</li> </ul>	<ul style="list-style-type: none"> <li>• public</li> <li>• anonyme</li> <li>• taille illimitée</li> <li>• asynchrone</li> </ul>

Alors que le (t)chat et l'espace de Facebook à l'étude sont publics et anonymes, les SMS sont envoyés en privé entre individus qui se connaissent. Ensuite, le (t)chat et Facebook ne limitent pas la longueur des messages, tandis qu'un SMS a une limite de caractères. Enfin, tandis que le (t)chat est en mode synchrone, la communication par SMS et sur Facebook peut se faire de manière synchrone bien que plus souvent asynchrone. D'après David Crystal (2011) et Susan Herring (2002), ces différences de modalité pourraient avoir un impact sur la CNE produite.

En outre, selon un recensement récent, les Québécois.es passent une bonne partie de leur temps en ligne sur Facebook. La Figure 0.4.1 ci-dessous présente quelques données pertinentes de la dernière enquête effectuée en 2023 par l'Académie de la transformation numérique, intitulée *NETendances : Actualités en ligne, réseaux sociaux et balados*.

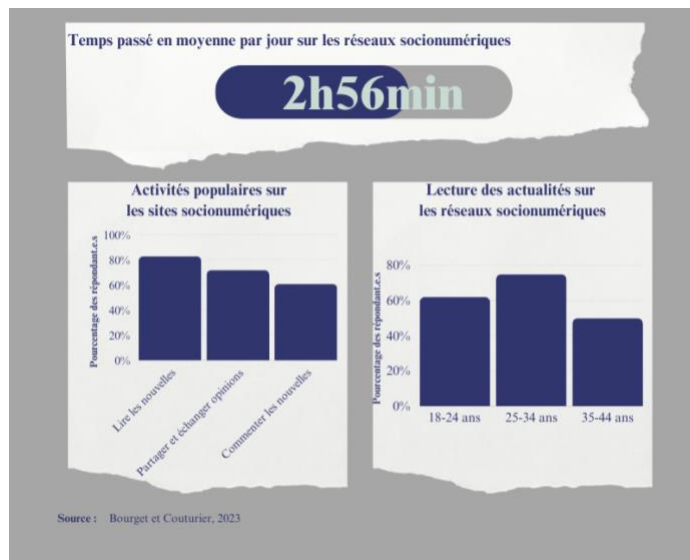
Figure 0.4.1 Utilisation des réseaux sociaux numériques au Québec



Plus précisément, cette enquête a déterminé qu'une grande majorité d'adultes de tous les groupes d'âges et sexes au Québec, soit 78 %, « ont utilisé un ou plusieurs réseaux sociaux comme Facebook, X/Twitter ou YouTube [et] que cette proportion était relativement stable depuis deux ans alors qu'elle se situait à 78 % en 2021 et à 79 % en 2022 ». Ensuite, « comme c'était le cas en 2022, Facebook demeure en 2023 la plateforme de réseaux sociaux la plus utilisée pour le loisir avec un pourcentage d'utilisation de 89 % » parmi les répondant.e.s (Bourget et Couturier, 2023 : 16-17).

Il s'avère donc que Facebook est la plateforme de choix des Québécois.es, et la Figure 0.4.2 ci-dessous confirme que les Québécois.es passent plusieurs heures par jour sur les réseaux sociaux où ils et elles s'engagent dans une variété d'activités, et qu'en effet les adultes de 18 à 44 ans se servent en grand nombre des réseaux sociaux pour consulter les actualités.

**Figure 0.4.2** Activité sur les réseaux sociaux au Québec



Il est donc clair que les activités et la communication sur les réseaux sociaux se sont insérées de façon importante dans la vie quotidienne des Québécois.es, et que Facebook est leur plateforme préférée. De plus, alors que Facebook n'est pas une nouvelle plateforme, en comparaison à d'autres plus récentes telles que TikTok ou Discord, c'est un lieu textuellement riche, et populaire auprès de la communauté francophone au Québec (Bourget *et al.*, 2022) fournissant donc un terrain fertile pour l'étude de la langue écrite.

## 0.5 Problématique

À l’instar des recherches susmentionnées sur la variation sociolinguistique dans la CNE produite par des internautes dans le (t)chat (van Compernelle, 2010) et les SMS (Bally, 2022 ; Blondeau *et al.*, 2014 ; Blondeau et Tremblay, 2022 ; Tremblay *et al.*, 2019), la présente étude se focalise sur la variation sociolinguistique dans le terrain public de la plateforme Facebook. Mon étude s’intéresse aux questions liées à la notion de la norme, à la variation sociolinguistique, et à la relation entre l’oral et l’écrit : Le contenu de la CNE dans l’espace public de Facebook suit-il les règles du standard écrit, ou diverge-t-il en ayant recours à des marqueurs sociolinguistiques typiques du français québécois à l’oral ? Si cette CNE contient bel et bien des marqueurs de l’oral informel, quelles sont alors leurs fréquences d’emploi, et comment ces fréquences se comparent-elles à celles qui sont attestées à l’oral ? À quel point les internautes mélangent-ils et elles des formes standards et non standards dans leurs commentaires ? Plus précisément, cette recherche tentera de répondre à deux grandes questions :

- I. La CNE québécoise sur FB ressemble-t-elle plutôt à la langue standard écrite québécoise, ou à la langue orale québécoise ?
  
- II. Quelle est la variabilité des emplois standards et non standards d’un commentaire à l’autre ?

Cette analyse se fera en adoptant des méthodes variationnistes pour déterminer les fréquences d’emploi de marqueurs sociolinguistiques attestés à l’oral, notamment de deux éléments de chacune des catégories suivantes : lexicale – marqueurs de contrainte et de conséquence ; verbes – emploi du futur et des auxiliaires *avoir* et *être* au passé composé ; syntaxe – emploi variable des formules interrogatives et de la négation ; et pronoms – alternance des pronoms *on* et *nous*, et des pronoms non clitiques simples et composés. Cette problématique et les questions de recherche apparaîtront de manière plus détaillée dans le deuxième chapitre.

Le premier grand résultat attendu est que les marqueurs sociolinguistiques choisis se trouveront bel et bien dans le corpus étudié. Ensuite, de façon plus générale, les résultats attendus sont que la CNE sur Facebook suivra les tendances linguistiques communautaires, extérieures au numérique, mais aussi que les internautes suivront une série de normes différente. En outre, on peut présumer qu'en raison des différences situationnelles entre l'oral, le (t)chat, le SMS et Facebook, les résultats de cette étude divergeront tant du standard écrit que de l'oral, et des résultats des études du FQ dans les autres modalités numériques. Par exemple, il est vraisemblable que les marqueurs sociolinguistiques les plus courts seront aussi les plus fréquents, suivant le principe de l'économie du langage. Ce principe repose sur la notion qu'« il est plus économique de recourir à un minimum de traits pertinents » et, en effet, « [l]es systèmes d'écriture évoluent vers une économie de plus en plus grande » (Dubois *et al.*, 2002 : 463-467). J'examine dans le Chapitre 1 l'impact possible de cette tendance vers l'économie sur la CNE. Les résultats attendus et les hypothèses sont discutés plus en détails dans le troisième chapitre.

Il importe aussi de mentionner qu'à la suite du projet de loi C-18 (Loi sur les nouvelles en ligne, 2023) instauré par le gouvernement fédéral du Canada, Facebook ne permet plus depuis le premier juin 2023 (Meta, 2023) aux organismes journalistiques de partager des publications sur l'actualité avec un hyperlien direct vers l'article. Il n'existe donc plus cet espace de discussion publique sur Facebook autour des sujets des articles de presse, espace d'où les données de mon corpus ont été extraites. Mon étude devient alors non seulement une première à étudier la CNE sur Facebook, mais aussi une contribution exclusive aux connaissances dans le domaine. De futures recherches pourront bien se focaliser sur d'autres espaces publics sur Facebook, mais ce corpus sera à ma connaissance le seul à être constitué d'écrits numériques provenant de ce lieu numérique précis.

## **0.6 Méthodologie**

Cette étude utilise des méthodes de la sociolinguistique variationniste, dans la tradition initiée par William Labov (1976), comme c'est le cas des études susmentionnées dans la section précédente. Cette approche applique les principes de l'enquête et de l'observation de preuves scientifiques à l'étude de la langue. Notamment, elle vise à employer des méthodes de concordance, soit le calcul du nombre d'occurrences d'un mot ou d'une expression, de collocation, soit l'étude des

cooccurrences de deux mots ou expressions, ainsi que des analyses statistiques pour calculer les fréquences d'emploi relatives des marqueurs étudiés.

Les analyses de concordance permettront de répondre à la première question de recherche concernant la proximité de cette CNE au standard écrit, à l'oral et aux autres contextes numériques. Les analyses de collocation permettront quant à elles de déterminer la variabilité dans l'emploi de formes standards et non standards de commentaire en commentaire. Ces méthodes sont alors à la fois qualitatives et quantitatives : le qualitatif sert à attester la présence d'un marqueur, et de coder son environnement linguistique, et le quantitatif sert à déterminer les fréquences d'emploi relatives, et donc l'effet possible des facteurs internes et externes sur la présence d'un marqueur. Cette démarche méthodologique est explicitée dans le troisième chapitre.

Les limites éventuelles de la méthode et de l'étude elle-même résident dans la manière dont les données ont été recueillies. Plus précisément, il s'agit des limites temporelles et matérielles du corpus lui-même, du manque de renseignements au sujet des internautes, de la récence (ou non) des données antérieures sur lesquelles s'appuie cette étude et des différences en termes de collectes de données des différentes études. Ces limites seront discutées de manière plus détaillée dans les troisième et sixième chapitres.

## **0.7 Objectifs et organisation de la thèse**

La présente étude a pour but de circonscrire la variété du français au Québec, telle qu'elle s'emploie dans l'espace public de Facebook. En se focalisant sur un nouveau terrain d'étude, cette recherche comblera les lacunes en ajoutant aux connaissances dans le domaine de la sociolinguistique variationniste sur l'emploi du français dans les réseaux sociaux numériques. De plus, cette recherche fait le bilan du système linguistique, en traitant de plusieurs parties du système linguistique, comparativement aux recherches antérieures qui se sont penchées sur une ou deux variables sociolinguistiques seulement.

Le premier chapitre présente le cadre théorique, en précisant comment la sociolinguistique variationniste s'applique à l'étude de la CNE produite par des internautes au Québec dans l'espace

public de Facebook. Le deuxième chapitre expose la pertinence et les objectifs de l'étude, tout en explicitant les questions de recherche et les hypothèses. Le troisième chapitre présente les fondements et la démarche méthodologiques, en décrivant la collecte et l'analyse statistique des données, selon la méthode variationniste. Les quatrième et cinquième chapitres présentent les résultats des analyses de concordance et de collocation des données. Les fréquences d'emploi sur Facebook présentées dans le Chapitre quatre sont également comparées à d'autres fréquences d'emploi dans des contextes oraux et socionumériques. Le sixième chapitre, après une présentation synthétique des résultats de concordance et de collocation, discute des enjeux théoriques à la lumière des résultats, en proposant un nouveau schéma des axes de la variation linguistique. Enfin, la conclusion aborde brièvement les conséquences et les limites des résultats, et suggère des pistes de recherche futures.

# **CHAPITRE 1 : LA COMMUNICATION NUMÉRIQUE ÉCRITE ET LA SOCIOLINGUISTIQUE VARIATIONNISTE**

## **1.0 Introduction**

Ce chapitre présente le cadre théorique de l'étude de la production langagière telle qu'elle se manifeste dans le numérique, sous l'angle de la sociolinguistique variationniste. L'objet de ce chapitre est donc de montrer comment la sociolinguistique s'applique à l'analyse de la CNE, et comment les particularités de la CNE rendent complexes certaines notions théoriques de la sociolinguistique traditionnelle. Il s'agira, dans un premier temps, de montrer la pertinence de la CNE pour l'étude de la variation linguistique selon une perspective sociolinguistique. Dans un deuxième temps seront présentés les paramètres et les caractéristiques de la CNE qui font de cette production langagière un objet d'étude d'intérêt pour la sociolinguistique. Dans un troisième temps, il s'agira de faire ressortir les liens entre ces caractéristiques de la CNE et l'étude de la variation linguistique. Enfin, l'apport de ces réflexions théoriques à la présente étude sera abordé.

## **1.1 La CNE et l'étude de la langue dans son contexte social**

La communication par le biais des technologies numériques forme maintenant une partie intégrante de nos vies sociales. Comparativement aux moyens de communication écrite traditionnels tels que la lettre et le billet télégraphique, la CNE représente une série de nouvelles habitudes sociales et langagières qui découlent des multiples lieux et manières de communication écrite créés par la technologie numérique. Comme le fait remarquer Marie-Anne Paveau (2019 : 9), « [u]ne grande partie de nos activités sociales, qui passent par le langage, et nous font entrer en contact avec d'autres, se font en ligne : nos relations sociolangagières sont connectées ». L'analyse de ces pratiques numériques et sociales est par conséquent d'intérêt pour les spécialistes en sciences sociales, notamment les linguistes et les sociolinguistes.

La sociolinguistique, domaine contemporain qui marie l'étude du système linguistique à celle du contexte social, se prête en effet bien à l'analyse de ces pratiques que Marie-Anne Paveau (2019) qualifie d'ailleurs de *sociotechnolangagières*. Depuis l'avènement de l'Internet, et surtout du Web 2.0, de nombreux sociolinguistes proposent ainsi de prendre l'espace virtuel comme nouveau terrain d'étude (Androutsopoulos, 2006 ; Seargeant et Tagg, 2014 ; Stanford, 2016 ; Friginal *et al.*, 2017 ; Tatossian, 2020 ; Tremblay *et al.*, 2019 ; Blondeau et Tremblay, 2022b.). C'est devenu un lieu d'enquête incontournable pour les chercheurs en sociolinguistique.

Les pratiques dans les espaces numériques, et en particulier celles dans les réseaux sociaux numériques, méritent d'être étudiées sous l'angle de cette perspective qui s'intéresse à l'emploi de la langue dans son contexte social (Staehr, 2014 ; Seargeant et Tagg, 2014). D'ailleurs, d'après Deborah Tannen (2013), la plupart de ce qui se passe dans les échanges en ligne ressemble à ce qui s'est toujours passé dans la conversation à l'oral. David Crystal précise ainsi qu'une sociolinguistique de l'espace numérique doit se donner pour but d'étudier les caractéristiques et les variétés de la langue utilisées à travers diverses plateformes techno- et socio numériques (Crystal, 2001 : 60-61), comme le feraient les sociolinguistes à l'oral.

En raison de ces ressemblances entre les écrits numériques et l'oral, la perspective sociolinguistique offre dès lors la possibilité de suivre l'évolution de la langue dans les médias sociaux numériques tout en considérant son contexte social. Une approche purement sociologique mettrait de côté d'importantes considérations théoriques et conceptuelles sur les particularités des dispositifs techno- et socio numériques qui influent sur la production langagière, telles que les notions de variation et de médium. Une approche purement linguistique, quant à elle, écarterait de l'étude l'importance non seulement de l'aspect social du contexte, mais aussi de l'interaction dans les choix langagiers des internautes en réaction à ces dispositifs. La sociolinguistique permet donc de tenir compte à la fois des productions langagières et des spécificités sociales dans lesquelles elles sont produites.

### 1.1.1 La CNE et la sociolinguistique variationniste

La production écrite dans les espaces du Web 2.0 s'avère très variable, et une grande partie de cette CNE naît dans des situations informelles (Marcoccia, 2016 : 74). En effet, selon Geoffrey Nunberg, il existe plus de variété dans le répertoire langagier en virtuel que dans l'écrit traditionnel — des vernaculaires informels comme d'autres genres de langages informels ayant leurs propres « règles stylistiques prégnantes » (Nunberg, 1999 : 117). Les lieux de la CNE sont donc bien des lieux de variation(s) linguistique(s) (Bibié-Emerit, 2015 ; Crystal, 2011 ; Herring, 2007 ; Marcoccia, 2016).

Si la sociolinguistique s'intéresse à l'étude de la langue (et de sa variation) dans son contexte social, la sociolinguistique variationniste en particulier cherche à mesurer la variation linguistique de manière quantitative. Traditionnellement, les (socio)linguistes examinent la variation langagière à l'oral. Toutefois, d'après Bulot et Blanchet (2013 : 63), elle peut aussi se réaliser à l'écrit. Or dans le cadre des technologies numériques de la communication, comme le souligne Michel Marcoccia (2016 : 65), la CNE peut justement diverger dans son rapport au standard écrit. Elle offre ainsi un nouveau terrain d'étude de la variation sociolinguistique, permettant notamment d'examiner la variation à l'écrit.

La variation linguistique est un phénomène mesurable. En effet, les sociolinguistes variationnistes perçoivent la variation linguistique non pas « comme une manifestation de compétence linguistique limitée de la part des locuteurs [et des locutrices], mais comme une propriété fondamentale de l'exercice du langage » (Boutet et Gadet, 2003 : 18). En outre, cette variation est contrainte par une régularité ou prédictibilité (Encrevé, 1976 : 16). En d'autres termes, la variation elle-même suit des tendances souvent prévisibles. De plus, tandis que la sociolinguistique variationniste traditionnelle, c'est-à-dire non numérique, se préoccupe plutôt des facteurs sociaux tels que l'âge, le sexe et le statut social de la personne (Labov, 1976), dans la sphère du numérique, le médium et la situation de communication sont aussi des dimensions attribuables à la variation linguistique (Herring, 2004 : 363). La variation se présente également à travers les modalités – le courriel, le clavardage, le SMS, les réseaux socionumériques – de communication (Herring, 2002 : 139). Il importe donc d'étudier cette variation dans tous ses contextes numériques.

Dans ce cadre, on ne perçoit pas les écrits numériques comme de la production langagière avec ou sans *fautes*, mais plutôt comme un comportement langagier révélateur des paramètres techno-, socio-numériques et situationnels de la CNE. L'étude variationniste se penche alors sur les fréquences d'emploi de certains éléments langagiers dans le but de comprendre, de décrire et de distinguer ces facteurs externes, et ainsi de rendre compte de l'effet de ces facteurs et de ces dimensions de la variation. La perspective variationniste, tout en tenant compte de la nature sociale de la communication, offre donc la possibilité d'examiner la CNE de manière objective et quantitative, sans porter de jugement sur sa *qualité*. En somme, cette perspective permet de repenser la notion de norme(s) linguistique(s) et les sources de la (ou des) variation(s) linguistique(s).

## **1.2 La CNE et la quête du vernaculaire**

Les sociolinguistes (et les variationnistes) s'intéressent souvent à la variété de langue dite vernaculaire. Il s'agit du « style où l'on accorde le minimum d'attention à la surveillance de son propre discours » (Labov, 1976 : 289). Cette variété, qui représenterait le *vrai* usage par les membres d'une communauté linguistique, témoigne d'une variation parfois surprenante par rapport au standard.

Dans le contexte du français parlé dans la province de Québec, le standard serait « la variété socialement valorisée que la majorité des Québécois[.es] francophones tendent à utiliser dans les situations de communication formelle » (Reinke et Ostiguy, 2016 : 11, citant AQP, 1977 : 11), ou encore « une langue de qualité associé[e] à l'usage formel » (Reinke et Ostiguy, 2016 : 124). Le standard écrit est alors une variété de langue très proche du français de référence (35), qui est celle « illustrée dans des outils de référence [...] (grammaires, dictionnaires [...]) » (127), qui suit donc les règles grammaticales, orthographiques, morphosyntaxiques, etc. En effet, cette forme standard s'apparente à celle du français hexagonal, car plusieurs études (De Villers, 2004 et 2005 ; Tremblay, 2016) ont démontré une « très grande similitude » entre le contenu linguistique de journaux québécois et français. Cela « appuie la thèse selon laquelle la variété standard écrite serait la même en France et au Canada » (Tremblay, 2020 : 85). Le vernaculaire est donc une forme langagière qui a tendance à diverger de ce standard écrit, et d'après Davy Bigot « [l]e cadre

théorique et méthodologique [de la sociolinguistique variationniste] est [...] le plus adapté à l'étude du français parlé au Québec et aux phénomènes de variation qui lui sont propres (2021 : 2).

En effet, d'une part, le vernaculaire représente une source précieuse d'indices sur les mécanismes du langage humain ainsi que sur les comportements et les motivations des locuteurs et locutrices (Siemund, 2008). D'autre part, c'est dans cette variété de langue que l'on peut trouver « de nombreuses innovations linguistiques [...] ce style donne ainsi l'indication la plus fiable de l'état actuel de la langue » (Coveney, 2016 : 1). Enfin, selon Labov, pour les sociolinguistes variationnistes, c'est en observant le vernaculaire que l'on obtient les données les plus systématiques, nécessaires à l'analyse de la structure linguistique (Labov, 1976 : 289). Les sociolinguistes se doivent alors de trouver une situation de communication qui permette d'observer de manière systématique le vernaculaire.

Comment accéder à cette variété familière, sans la compromettre ? Si le vernaculaire est la variété la moins surveillée par le sujet parlant, cela veut dire logiquement que si ce dernier sait qu'il est observé, il fera plus attention à son discours. Voici le *paradoxe de l'observateur* — les sociolinguistes veulent observer la production langagière des sujets quand ils ou elles ne font pas attention à leur production langagière, mais la présence même des sociolinguistes fait que les sujets surveilleront leur production langagière (Labov, 1976). Le vernaculaire représente alors une variété essentiellement mythique pour les sociolinguistes dans les contextes traditionnels de l'étude de terrain, qui prennent la forme surtout d'entrevues semi-dirigées.

### **1.2.1 L'accès au vernaculaire en contexte numérique**

L'espace numérique offre aux chercheurs un accès à de nouvelles sources de production langagière non standard. Par exemple, étant donné la nature informelle des situations de communication dans le Web 2.0, d'après Gretchen McCulloch (2019 : 21), ce terrain permet d'accéder à la production langagière inconsciente sur Internet puisque les linguistes ont accès à d'innombrables exemples de langage public et informel. Ainsi, « le discours électronique offre un terrain fertile pour l'exploration de la transposition du vernaculaire à l'écrit » (Blondeau et Tremblay, 2022b : 121)

car cette situation « donne la possibilité d’observer “en cachette” de très nombreux discours et échanges écrits naturels, en évitant d’être pris dans le fameux paradoxe de l’observateur » puisque « les auteurs des messages ignorent notre présence » (Marcoccia, 2016 : 37). La CNE est donc non seulement un objet d’étude pertinent à la sociolinguistique variationniste, mais elle offre également la possibilité d’examiner de manière mesurable et en secret le vernaculaire. En outre, Andreas Staehr (2014 : 25) précise que les plateformes des médias socionumériques, en offrant l’occasion d’employer la langue à l’écrit quotidiennement, entraînent un rapprochement entre les styles vernaculaires oraux et écrits. L’étude de la CNE présente alors l’occasion d’examiner la langue d’une nouvelle manière, inaccessible aux linguistes avant les échanges dans le Web 2.0.

### **1.2.1.1 L’étude du français parlé au Québec en contexte numérique**

Pour ce qui est du français parlé au Québec, cette variété de langue fait l’objet d’études variationnistes depuis plusieurs décennies, surtout en raison de l’écart entre le FQ d’usage et le FQ standard. Plus précisément, comme nous l’avons décrit dans l’introduction, au Québec, le FQ oral familial s’écarte de manière importante du standard (Martel et Cajolet-Laganière, 1995) sur plusieurs plans. L’étude de la CNE en FQ permet donc d’examiner cette tension entre le français standard et le français d’usage au Québec dans des contextes écrits. En effet, d’après Hélène Blondeau et Mireille Tremblay, la CNE offre « l’occasion exceptionnelle de revoir la dichotomie traditionnelle entre le français écrit normé et le français oral vernaculaire » (2022b : 122). Les sociolinguistes ont commencé à examiner ce phénomène dans des contextes SMS et (t)chat, mais il doit encore être examiné sur une plateforme de média socionumérique.

À l’instar de cinq travaux variationnistes plus ou moins récents (Bally, 2022 ; Blondeau *et al.*, 2014 ; Blondeau et Tremblay, 2022 ; Tremblay *et al.*, 2019 ; van Compernelle, 2010), la présente recherche élargit ainsi le champ d’études vers un nouveau terrain. Si l’on souhaite comprendre l’évolution, l’état présent et l’avenir du FQ, la logique demande d’étudier l’emploi de cette langue dans *tous* les contextes sociaux. Hélène Blondeau et ses collègues (2014 : 137) renchérissent sur ce point dans leur étude sociolinguistique des SMS de francophones montréalais :

Un nouvel examen de la configuration sociolinguistique de Montréal requiert non seulement de prendre en compte les productions linguistiques orales des Montréalais, ce qui correspond à la pratique traditionnelle des

sociolinguistes, mais également d'élargir l'analyse aux nouvelles formes de communication médiatisées par ordinateurs (CMO) [...] de nos jours partie prenante des échanges communicationnels quotidiens.

Une étude sociolinguistique variationniste a donc le mérite de quantifier la variation dans la CNE, afin de mieux comprendre le comportement linguistique de la communauté, surtout dans ce nouveau contexte techno- et socionumérique. Cet examen de la CNE produite par des internautes qui déclarent résider au Québec dans un espace du Web 2.0 intouché, permettra de contribuer aux connaissances sur le comportement linguistique des internautes à travers différentes modalités.

### **1.3 La CNE et la question de la norme**

La variation inhérente à la CNE oblige à repenser les réflexions sur la notion de norme. La variation linguistique repose en effet sur cette idée d'une norme linguistique : la variation est alors toute divergence de la norme. Une communauté linguistique, explique William Labov (1976 : 187), « se définit [...] par une participation conjointe à un ensemble de normes ». Ainsi, d'après Kristin Reinke et Luc Ostiguy (2016 : 112), c'est aux sociolinguistes que revient la tâche de décrire une norme objective, en observant comment communiquent « réellement les membres d'une communauté ». Cette étude a donc pour but de dégager les normes d'usage dans cette nouvelle situation communicative qu'est la CNE.

En dehors de la sphère CNE, les sociolinguistes proposent autant de perspectives que de normes. Pour Garvin (1983), la norme représenterait le pouvoir institutionnel ou national, tandis que Françoise Gadet, citant Labov, distingue « “norme endogène” (évaluation) » et « “norme institutionnelle” (la norme) » ou encore, les « dominantes » et les « dominées [...] qui expriment la solidarité de groupe » (Gadet, 1995 : 21-22). Enfin, Robert Vézina (2009 : 1) distingue « norme de l'usage » et « norme de bon usage », notant que

[I]a norme de l'usage est celle de la conformité à ce qui est utilisé le plus fréquemment et ce qui est généralement partagé par l'ensemble des locuteurs de la langue [...] L'usage est en effet multiple, dynamique, traversé par plusieurs tendances [...] à l'image de la société qui parle la langue en question.

L'étude des écrits numériques est donc importante pour comprendre quels genres de normes sous-tendent cette situation de communication.

### 1.3.1 La CNE et le jeu avec la norme

La CNE ouvre justement une dialectique sur la possibilité de nouvelles normes linguistiques communautaires, surtout en ce qui concerne les différents types de normes susmentionnés. En particulier, puisque les internautes utilisent la langue sous forme de CNE dans de nouveaux contextes sociaux- et technonumériques, sans intervention d'une autorité externe, ce sont eux qui déterminent et coconstruisent la ou les normes d'usage(s). Les linguistes ont déjà abordé cette question surtout dans le contexte du français hexagonal. Par exemple, Michel Marcoccia (2012 : 167), dans son étude sociolinguistique du forum *Ados.fr*, a constaté qu'il « n'y a pas encore dans ce collectif de partage d'un ensemble de normes mais, au contraire, de nombreuses négociations sur ces normes ». Dans une autre étude, à travers l'analyse de souhaits d'anniversaire sur Facebook, Laetitia Bibié-Emerit a trouvé « que les locuteurs jouent avec la norme et s'en servent en dehors des cadres prescriptifs : la norme devient source de créativité » (Bibié-Emerit, 2015 : 334). Enfin, l'analyse sociolinguistique d'un corpus SMS réalisée par Amélie Cougnon indique qu'en

ce qui concerne la perte des règles [...] la majorité des usagers a [...] pleinement connaissance de la norme dont elle s'écarte. Sous l'effet d'un contexte « relâché », d'une complicité voulue avec l'interlocuteur, d'un désir d'abrèger l'acte de communication [...] elle jongle avec de nouvelles règles graphiques qu'elle semble maîtriser, comme un nouveau sous-langage, en plus des règles traditionnelles. (2010 : 409).

Pour ce qui est de la norme et du FQ, d'après Hélène Blondeau et Mireille Tremblay (2022b : 122) « la nouvelle scripta vernaculaire » des SMS « répond à des normes communautaires ». Étudier la CNE produite sur Facebook par des internautes au Québec selon une perspective sociolinguistique variationniste contribue alors aux recherches sur les normes communautaires qui se construisent par les internautes, et, plus largement, aux réflexions sur l'établissement des normes linguistiques.

### 1.4 Les caractéristiques générales de la CNE

La CNE fait preuve de variabilité à travers le Web 2.0. D'après Michel Marcoccia (2016 : 69), « on observe en fait une assez grande hétérogénéité des discours numériques écrits ». Comme dans toute situation de communication, le contexte et ses paramètres vont déterminer les choix langagiers des interlocuteurs. Casey Jewitt (2016 : 70) note lui aussi que les technologies numériques rendent possible toute une gamme de modes, elles bouleversent et refont les genres,

et souvent elles remodelent les pratiques et l'interaction. Qu'il s'agisse du (t)chat, du forum de discussion ou encore du courrier électronique, le style de langue diffère selon la plateforme et selon la situation. Rémi van Compernelle et Lawrence Williams (2007), dans leur étude de trois situations de CNE – notamment le (t)chat, le forum et le clavardage – ont trouvé par exemple que la variation dans l'orthographe diffère selon la plateforme. Cette variation peut s'expliquer par trois facteurs, sur lesquels je reviendrai au cours de cette étude : 1. la diversité des dispositifs et des plateformes ; 2. les situations engendrées par ces plateformes ; 3. la manipulation des codes langagiers qui caractérisent la CNE. Une approche sociolinguistique, qui en soi prend en considération la nature sociale de la communication, peut rendre compte de l'importance de ces paramètres pour la production langagière.

### **1.5 Les paramètres communicatifs de la CNE**

La variation touche tous les niveaux de la CNE, comme les variations graphiques qui « sont susceptibles de se faire concurrence » puisqu'une personne peut « utiliser une forme graphique » tandis que son homologue « en choisit une autre » (Feuillard, 2017 : 227). Un regroupement d'énoncés équivalents, « *two or more ways of saying the same thing* » (Sankoff, 1980 : 55) est une variable linguistique, et les variantes sont chacune des alternances possibles d'une variable. Par exemple, les mots suivants, quatre variantes d'une variable, évoquent tous le sens de véhicule personnel : *automobile*, *voiture*, *char* et *bagnole*. Ces variables et variantes se retrouvent à plusieurs niveaux du système linguistique : phonologique, sémantique, lexical, verbal, syntaxique, etc. Par exemple, dans les SMS écrits en français, on constate de la variation lexicale et syntaxique, ou encore de la variation par rapport à une norme orthographique (Cougnon, 2010 : 397). La variation se présente ainsi de différentes manières.

Selon la perspective sociolinguistique, dans le cas de la communication traditionnelle (c'est-à-dire non CNE), la variation linguistique peut être le résultat d'un ou de plusieurs facteurs internes ou externes à la langue. Par exemple, l'environnement linguistique (phonologique ou grammatical) d'une variable détermine les facteurs internes (Walker, 2010). Ensuite, les paramètres de la situation de communication ainsi que les caractéristiques sociales des sujets parlants eux-mêmes déterminent les facteurs externes (Bulot et Blanchet, 2013). Dans le cas de la CNE, de nouveaux

paramètres s'ajoutent à ces dimensions traditionnelles de la variation linguistique, autant internes qu'externes. Pour ce qui est des facteurs internes, il s'agit de mots et d'éléments entourant la variante dans un syntagme qui peuvent opérer sur le choix de cette variante. Dans le cas de la CNE, un exemple serait les choix orthographiques écrits en ligne pouvant être conditionnés par la phonologie et la phonétique (Eisenstein, 2015). En ce qui concerne les facteurs externes, la nature publique ou privée de la communication, le type d'auditoire, l'environnement technologique et l'appareil utilisé, font partie de facteurs externes ayant un effet possible sur la production langagière dans le numérique (Herring, 2007). Je discuterai des impacts possibles des nouveaux facteurs internes et externes à la lumière des résultats de l'étude dans le quatrième et le cinquième chapitres.

### **1.5.1 Les dispositifs de la CNE**

Une variété de dispositifs technonumériques s'offre aux internautes auteurs de CNE. Les plus connus sont l'ordinateur, le téléphone mobile ou intelligent, et la tablette électronique. Chacun de ces appareils possède ses propres modalités, fonctionnalités et contraintes.

Prenons les différents types de clavier. Le clavier d'ordinateur est physique, tandis que celui du téléphone intelligent et de la tablette électronique est aujourd'hui le plus souvent virtuel ou numérique. Dans le cas de l'ordinateur portable, il s'agit d'un clavier qui fait partie intégrante de l'appareil, tandis qu'avec l'ordinateur personnel, c'est un appareil indépendant avec ou sans fil. Le clavier virtuel, apparaissant sur l'écran de l'appareil quand nécessaire, n'a pas « toutes les fonctionnalités du clavier complet » (Poirier et Schadle, 2004 : 134). Par ailleurs, les claviers, qui sont de formats variés, n'ont pas tous le même nombre de touches. Ces différences peuvent avoir un effet sur la production langagière. Comme le fait remarquer Franck Poirier, « la difficulté de saisie d'information dépend directement de la taille de l'objet. Plus l'objet est petit, plus il est difficile voire impossible d'avoir une fonctionnalité de saisie de texte » (Poirier, 2017 : 69). D'autre part, les appareils portables « sont des objets nomades qui sont employés parfois debout, en marchant, posés sur les genoux, à une main » et la fonction de leur mini-clavier « n'est pas de saisir de longs textes » (Poirier et Schadle, 2004 : 133). Clairement, chaque type d'appareil offre ses propres possibilités communicatives, tout en posant des contraintes pour l'internaute,

comme la taille et la rapidité avec laquelle il peut faire parvenir le message. Ces contraintes vont sans doute avoir un impact sur la nature de la production langagière écrite.

L'ordinateur, le téléphone mobile ou intelligent et la tablette électronique offrent aussi la fonction parole texte, permettant à la personne non pas de taper, mais de vocaliser son message, que le logiciel interprétera et transcrira par la suite. En effet, prévu en partie pour les personnes qui sont dans l'incapacité de taper, « [l']assistant vocal est une alternative au clavier pour les tout petits appareils qui n'offrent pas de clavier ou un clavier très réduit qui induit inévitablement une charge cognitive et motrice importante » (Poirier, 2017 : 71). Alors que l'assistant vocal représente une adaptation permettant l'accessibilité, il implique aussi deux conséquences contradictoires pour l'entrée des données langagières. Premièrement, le logiciel n'utilisera ni raccourcis (p. ex. les abréviations et l'orthographe phonétique), ni ponctuation sans que l'utilisateur le précise. Deuxièmement, il risque de ne pas transcrire avec exactitude la production langagière entrée, entraînant surtout des fautes lexicales et grammaticales. Enfin, les logiciels d'autocorrection de texte, censés supprimer ou « corriger » les « fautes » (intentionnelles ou non) produites par l'internaute, pourraient eux aussi contribuer à la variation linguistique présente dans la CNE.

### **1.5.2 Les plateformes de la CNE**

La liste des différentes plateformes socionumériques dans le Web 2.0 ne cesse de croître. Que ce soit un forum, un média socionumérique, ou une modalité de (t)chat, chaque plateforme détermine comment le langage peut être rentré, réglant souvent la longueur du message, la vitesse des échanges, etc. Certaines plateformes, telles que le SMS et X (Twitter), ont des limites de caractères, tandis que d'autres, telles que Facebook ou le courrier électronique, permettent à l'internaute d'écrire autant qu'il ou elle le veut. À l'heure actuelle, la plupart des plateformes proposent un mélange de vidéos, d'images et de messages textuels, comme c'est le cas sur Instagram, X (Twitter) et Facebook, par exemple. Avec d'autres plateformes, l'internaute ne peut qu'entrer de la production écrite, comme c'est le cas pour les commentaires sur YouTube. La plupart des plateformes fournissent également une variété de situations de communication, tant en public qu'en privé. Elles permettent même à l'internaute de régler ses paramètres de sécurité pour rendre son compte plus ou moins public. Enfin, les plateformes offrent la possibilité de communiquer en

mode synchrone, asynchrone, ou les deux (Crystal, 2011 : 33). Il existe donc une panoplie de plateformes, chacune avec ses propres contraintes et fonctionnalités.

Christine Develotte et Marie-Anne Paveau font remarquer que les médias sociaux sont des « espaces multimodaux » dans lesquels coexistent « des éléments graphiques, iconiques et picturaux, statiques ou mobiles, et des fonctions de communication synchrones et asynchrones. Ces espaces discursifs conditionnent [...] ce qu'il pourra être possible d'énoncer et de produire dans l'interaction ». Il s'agit alors d'« un espace d'exposition discursive qui configure les productions discursives en fonction de leurs affordances » (2017 : 206). L'existence de ces plateformes numériques, telles que les réseaux sociaux, le clavardage et le SMS, permet aux internautes de communiquer tout en éliminant l'importance du contexte spatio-temporel. Simultanément, la nature même de chaque plateforme opère sur les types de situations de communication et sur la nature de la CNE qui résulte des interactions entre internautes.

### **1.5.3 Les situations de communication de la CNE**

D'après David Crystal (2011 : 15), la première étape dans une *linguistique d'Internet* est d'établir les propriétés qui conditionnent l'expérience et le comportement langagier de ses internautes, en distinguant ce monde des mondes connus du langage oral et écrit. En raison de la diversité des appareils et des plateformes dont disposent les internautes auteurs de CNE, les types de situations de communication possibles sont multiples. Marie-Anne Paveau, spécialiste en textualités numériques, précise notamment que,

pour étudier les productions langagières en ligne, la connaissance des contextes numériques est nécessaire. En effet, Internet, et le Web en particulier, ne constituent pas de simples supports pour une production scripturale qui s'y transporterait, mais bien des environnements qui configurent structurellement les écritures de manière spécifique (2015 : 1).

L'approche traditionnelle en sociolinguistique a été d'étudier la situation de communication pour pouvoir étudier le comportement linguistique de la communauté (Hymes, 2001 : 9), et de rendre compte des facteurs de la variation linguistique. Une adaptation de cette approche est donc nécessaire dans le cadre d'une étude sur la communication dans l'espace numérique.

Le tableau synthétique ci-dessous, créé à partir des schémas de classification de Susan Herring (2007) et de Michel Marcoccia (2016), permet de rendre compte de cette variété de situations et de classer une situation de communication CNE selon ses caractéristiques technonumériques et sociales. Tandis que le schéma des facettes sociales et technologiques de la CMO de Susan Herring s'inspire quant à lui du schéma SPEAKING de Dell Hymes (1974), celui de Michel Marcoccia est tiré de la typologie des outils de communication numérique de Jacques Anis (1999) et de la notion de persistance des messages de Thomas Erikson (1999).

**Tableau 1.5.3 - Schéma de classification de la situation de communication de CNE**

<b>Caractéristiques sociales</b>	<b>Description</b>	<b>Caractéristiques technologiques</b>	<b>Description</b>
Scripteur	Le nombre ou le type de scripteur(s), ses (leurs) caractéristiques	Instrument d'entrée	Clavier ordinateur, téléphone ou tablette, saisie vocale...
Lecteur, auditoire, cadre participatif	L'auditoire (anonyme ou connu, limité ou illimité)	Support lecture	Écran d'ordinateur, de téléphone ou de tablette
Lecteur réponse	Le type de réponse de la part du lecteur	Code	Code (langagier ou autre, variété de langue) utilisé (oral, écrit, iconique)
Répondeur	Le nombre ou le type de scripteur(s) qui répond(ent) au message	Médium, canal	Nom du médium, de la plateforme, du site Web 2.0
But des messages, objectifs	Le but, les fins du message	Mode de transmission	Décrit les caractéristiques des messages (mots, caractères, limites) ; calibrage, longueur, granularité
Sujets, thèmes	Le(s) sujet(s), le(s) thème(s) du message	Temporalité, synchronicité	Synchronicité ou asynchronicité de l'échange communicatif
Ton des messages	Le ton du message, le niveau de formalité	Persistance du message, adaptations	Message persistant ou qui disparaît, qui peut être filtré ou non, qui peut être cité ou non.
Environnement social	La nature publique ou non de l'environnement social, et si l'environnement est modéré ou non	Environnement technologique	Décrit comment les messages sont vus et lus par les scripteurs et par l'auditoire, apparence physique, format de présentation
Normes sociales, linguistiques	Les normes sociales ou linguistiques, organisationnelles, qui encadrent la production du message		

D'après Michel Marcoccia (2016 : 28), un tel schéma permet de « décrire très précisément les situations de communication numérique que l'on désire analyser », ce qui est nécessaire dans le cadre d'une étude de la variation linguistique. Je reprendrai ce schéma dans le troisième chapitre, qui porte sur la méthodologie, pour l'appliquer à la situation de communication à l'étude dans ma recherche, en l'occurrence celle des commentaires publics sur FB. Cette précision me permet de

délimiter les facteurs externes possibles de la variation linguistique retrouvée dans la CNE en FQ sur FB.

#### **1.5.4 L'hybridité de la CNE**

L'autre caractéristique de la CNE qui pourrait être d'importance à la sociolinguistique variationniste est celle de l'hybridité langagière. Au début des technologies numériques de la communication, les chercheurs et les pédagogues se plaignaient que les internautes négligeaient leur production écrite, contribuant par là même à l'appauvrissement de la langue. Or, de nos jours, les linguistes se penchent avec beaucoup d'intérêt sur un phénomène que la société voyait auparavant comme étant fautif : on constate un nouvel emploi de la langue, qui mélange l'oral et l'écrit, ainsi que l'iconique. Alors que pour certains (Anis, 1998 ; Collot et Belmore, 1996 ; Fouser, Inoue et Lee, 2000 ; Panckhurst, 1998 ; Yates, 1996), la CNE est un style de langue « marqué par l'oralité », Michel Marcoccia (2004 : 1) voit plutôt cette communication comme faisant « du face à face avec de l'écrit ». Selon lui et Nathalie Gauducheau, la CNE serait « un type de discours hybride, entre l'oral et l'écrit [...] qui brouille les frontières entre ces deux types » (2007 : 39). Enfin, d'après David Crystal (2011), la CNE est un langage écrit qui tire en quelque sorte dans la direction de l'oral, mais qui ne correspond ni à la catégorie traditionnelle de l'écrit, ni à celle de l'oral. Il s'agit donc d'une forme langagière combinant des caractéristiques des deux modalités langagières traditionnelles, l'oral et l'écrit, ainsi que des éléments typographiques et iconiques.

Quelles que soient les nuances de sa description, il est certain que l'hybridité de la CNE fournit à la sociolinguistique variationniste une nouvelle source de données quantifiables, qui peuvent ensuite être comparées aux tendances trouvées à l'oral ou dans d'autres formes écrites, numériques ou autres. Ces nouvelles pistes d'enquêtes permettent de mieux comprendre le comportement linguistique des internautes lorsqu'ils communiquent à l'écrit par le biais de la technologie numérique.

### 1.5.4.1 Exemples de l'hybridité de la CNE

Il suffit d'ouvrir une plateforme de réseautage socionumérique, ou un logiciel de messagerie sur son téléphone, pour constater des exemples de l'hybridité de la CNE. La manipulation orthographique, l'emploi d'émojis et de symboles non alphabétiques, l'abréviation, la troncation, l'emploi d'acronymes et l'emploi de termes qui proviennent de l'oral, ne sont que quelques exemples des traces de l'oralité dans la CNE (Marruccia, 2016). Le message ci-dessous, tiré de l'étude effectuée par Hélène Blondeau et ses collègues (2014 : 138) sur les SMS envoyés par des Montréalais, illustre bien ces stratégies communicatives et démontre la qualité hybride de la CNE :

l'éventail des ressources inscrites au répertoire d'un des texteurs du corpus :

(1) Ha okay ! Merci... ms jchekait mon agenda pi j'ai comme coudon c6 kan ki va nous donne les autre livre ? Lol :P (Texto4Science, 112257)

Traduction non CNE : *Ha, okay ! Merci... Mais je checkais mon agenda pis j'étais comme : « coudon, c'est quand qu'il va nous donner les autres livres ? » Lol :P*

Cet extrait contient en effet plusieurs exemples de la mixité orale et écrite de la CNE, notamment « des marques discursives » et des « anglicismes lexicaux » typiques du FQ oral, « un introducteur de discours en émergence [...] trois élisions [...] une abréviation [...] quatre québécoisismes [...] trois cas d'orthographe phonétique », ainsi que des « déviations par rapport à la norme écrite », (Blondeau *et al.*, 2014 : 138), notamment l'accord verbal et l'accord nominal. De plus, on constate des interjections telles que « Ha okay ! », l'emploi de lettres et de chiffres pour représenter les sons phonétiques, comme dans « c6 » (le chiffre 6 représente vraisemblablement le *e* avec accent aigu). Enfin, cette personne utilise deux stratégies pour représenter des éléments paralinguistiques : des points de suspension pour marquer une pause, comme on le ferait à l'oral, ainsi que la combinaison d'un signe de ponctuation et d'une lettre (:P) pour représenter le visage qui tire la langue. Cet extrait témoigne de la grande variation dans le comportement langagier des scripteurs, et met en évidence la pertinence de l'hybridité de la CNE, puisqu'elle bouleverse la notion traditionnelle de variation linguistique.

## 1.6 La CNE et les dimensions de la variation linguistique

Après avoir esquissé les paramètres propres aux contextes CNE qui sont pertinents dans le cadre d'une étude sociolinguistique variationniste, il importe de s'attarder sur trois aspects clés : la diversité des dispositifs et des plateformes techno- et socionumériques, la variabilité des situations de communication engendrées par la technologie numérique, et l'hybridité oral-écrit de la CNE, dans le cadre des dimensions diaphasique (variation selon la situation) et diamésique (variation selon le médium, vue de manière traditionnelle comme étant les différences entre l'oral et l'écrit) de la variation linguistique.

### 1.6.1 La CNE et la variation diaphasique

La variation de la CNE s'explique en partie par la gamme de situations communicatives engendrées par les fonctionnalités des plateformes socionumériques : il s'agit alors d'une variation diaphasique. En effet, selon Géraldine Blattner et Lawrence Williams (2011),

la communication électronique n'est pas homogène et se manifeste à travers différents styles et genres qui sont généralement guidés par des facteurs humains tels que le but de l'échange et les caractéristiques du locuteur et de son audience [...] l'environnement électronique est une source qui a une influence sur le registre de discours sélectionné par les internautes.

Toute plateforme socionumérique qui permet la CNE aura ses propres paramètres sociaux et situationnels. D'après David Crystal (2011 : 77), les contraintes situationnelles suffisent à distinguer les variétés de CNE. Les paramètres technologiques, par exemple, peuvent rester très ouverts et fluides, comme dans Facebook qui offre une variété d'options de production écrite sans limites de longueur. Dans d'autres cas, la plateforme est très contraignante : X/Twitter, par exemple, limite la longueur des messages, et Reddit est organisé par sujet. Quant aux paramètres sociaux, tout est possible. Pour ne citer que deux exemples, le SMS est un message envoyé en privé à un individu connu, souvent un proche, tandis que le tweet peut s'adresser à un auditoire vaste et anonyme. Le contexte précis dans lequel une communication numérique écrite se produit aura ainsi sans doute un effet sur le contenu linguistique de cette CNE.

La situation de communication « est constituée de l'ensemble des éléments extralinguistiques qui conditionnent le déroulement de l'interaction » c'est-à-dire de « tous les éléments qui préexistent à l'interaction » (André, 2019). On appellera donc variation diaphasique, ou variation stylistique ou situationnelle, celle qui est déterminée par la situation de communication (ex. la proximité des sujets parlants, le ton de la conversation, etc.). En effet, le degré de formalité dans laquelle une interlocution a lieu détermine souvent le style de langue utilisé.

### **1.6.1.1 Le contexte numérique et la variation diaphasique**

La notion de variation diaphasique est importante pour l'étude sociolinguistique, puisque les individus ont la faculté de varier leur langage selon la situation de communication et ses paramètres extralinguistiques (Moreau, 1997 ; Wüest, 2009). On peut également parler de variété soutenue (formelle) et de variété familière (informelle) (Reinke et Ostiguy, 2016 : 30). Dans le contexte de la CNE, un internaute pourra, par exemple, modifier son style de langue selon qu'il envoie un texto à son proche, ou qu'il écrit une publication sur un site de réseautage professionnel tel que LinkedIn.

Comme dans toute situation de communication, le contexte et les paramètres encadrant la production de la CNE vont déterminer les choix langagiers des internautes. Le Web 2.0 peut effacer le temps et la distance, et rendre l'auditoire limité et privé ou illimité, public et anonyme. Ainsi, la diversité des plateformes socionumériques qui créent une infinitude de situations de communication rend plus complexe la question de la formalité et de la variation diaphasique. En effet, d'après Michel Marcoccia (2016 : 24),

analyser la communication numérique écrite oblige à en distinguer les différentes variétés, déterminées à la fois par des critères technologiques et situationnels (ou communicationnels) : on n'écrit pas de la même manière un texto adressé à un ami et un courrier électronique professionnel, pour des raisons liées à la fois à l'outil utilisé et à la situation (par exemple à son degré de formalité).

Tandis que certains contextes CNE sont très semblables à des contextes écrits hors Internet, tels que les écrits académiques, juridiques et journalistiques, d'autres contiennent plus ou moins d'éléments typiques de l'oral, tels que le SMS ou le chat (écrit aussi (t)chat) souvent informels et instantanés, comme le serait une conversation (Crystal, 2011 : 20). On pourrait même concevoir

un continuum de styles, car les recherches sur la CNE à travers différentes plateformes suggèrent l'existence de plus d'*un* style CNE. En effet, les recherches du sociolinguiste anglais Douglas Biber indiquent qu'il y aurait un certain nombre de variétés ou de registres langagiers à travers Internet, donc Web 1.0 (Biber *et al.*, 2017), et il est vraisemblable que le même principe s'appliquerait aux différentes situations de communication qui existent dans le Web 2.0.

### **1.6.1.2 Preuves de la variation diaphasique en contexte numérique**

Globalement dans la CNE, on trouve « à la fois des marques d'oralité (le style “informel” des courriers électroniques) et des marques d'écrit standard (le “ne” de la négation, par exemple) » (Marcoccia, 2004 : 1). Dans les listes de diffusion, la CNE est surtout d'un registre de langue soutenu et le recours au style familier y est assez rare (Marcoccia, 2016 : 69). Ceci peut s'expliquer par la relation des participants à la conversation ainsi que par le but de la communication ; une telle forme de communication est moins dialogique, mais plutôt informative.

Pour ce qui est des messages envoyés par courriel, le contenu peut varier énormément. En effet, selon David Crystal (2011 : 21) certains individus se contentent d'envoyer leurs messages sans faire de révision, ni se soucier des anomalies éventuelles dans leur écriture numérique. D'autres se donnent la peine de revoir leurs messages, comme ils le feraient dans des contextes communicatifs hors Internet. Cette variation s'explique par la diversité des contextes sociaux dans lesquels un courriel peut être envoyé ; un message dans un contexte professionnel contiendra vraisemblablement un style beaucoup plus formel que des messages envoyés entre proches ou au sein d'un groupe d'ami.e.s.

En ce qui concerne les différences stylistiques selon la situation, Jacques Maurais (2003 : 11) a constaté que « les courriels ne sont pas une simple transcription d'un oral graphié, comme on peut le voir dans les “chats” ou “clavardages”, dont l'une des caractéristiques est la temporalité — synchrone — des échanges, alors que les courriels sont asynchrones ». La langue semble en effet également varier selon l'existence ou non d'une modération de la situation de communication : dans le (t)chat non modéré, les « formes oralisées non traditionnelles pullulent », tandis que dans le (t)chat modéré ou dans le forum, « il n'y a presque pas de variation » (Marcoccia, 2016 : 69).

On voit alors comment l'adaptation des schémas de Susan Herring (2007) et de Michel Marcoccia (2016) est importante pour pouvoir décrire les facettes technologiques et sociales de la situation de communication de la plateforme utilisée qui peuvent avoir un effet sur la variation dans la CNE.

### **1.6.2 La CNE et la variation diamésique**

La notion de variation diamésique est remise en question étant donné le processus de médiatisation technonumérique de la CNE. La sociolinguistique a traditionnellement considéré la variation diamésique comme représentant les « différences de moyens utilisés à mesure pour communiquer » (Wüest, 2009 : 147, citant Mioni, 1983 : 508) c'est-à-dire les « adaptations nécessaires au passage de productions orales à des productions écrites » (Lacaze, 2022). Autrement dit, les différences de médium peuvent elles aussi être à l'origine des variations dans une même langue (Gregory et Carroll, 1978 : 38).

Mais les particularités de la CNE rendent plus complexe la notion de variation diamésique en sociolinguistique : quels sont ces différents *moyens* ? Le *médium* est-il le code langagier (oral ou écrit), le mode de production (clavier d'ordinateur, téléphone), ou encore la plateforme par laquelle le message est transmis (logiciel de messagerie SMS ou (t)chat, site Web, média socionumérique, etc.) ? Quelles *adaptations* entrent en jeu, avec un langage hybride oral-écrit (et iconique) ? Ces questions seront abordées plus en détails dans les prochaines sections, et ensuite à la lumière des résultats de la présente étude, dans le sixième chapitre et la conclusion.

#### **1.6.2.1 La médiatisation technonumérique et la variation diamésique**

Dans le cadre de la production de la CNE, la variation diamésique (c'est-à-dire la variation du médium) peut inclure les fonctionnalités et les contraintes des différents dispositifs, et les différents appareils dont les internautes se servent. Comme mentionné précédemment, la CNE peut être le résultat d'encodage par ordinateur, par téléphone intelligent, ou par tablette électronique, avec différents types de claviers et même à travers la transcription de l'oral en mode écrit par saisie vocale.

Nous avons vu que chaque appareil a ses propres fonctionnalités et/ou limites. Un.e internaute qui écrirait par clavier d'ordinateur, assis à son bureau, ne produirait sans doute pas une CNE comme le ferait un.e internaute qui écrirait avec les doigts seulement, sur son téléphone intelligent, debout dans le métro. En effet, une représentation concrète de cette idée est fournie par Michel Marcoccia (2016 : 87-88), qui souligne que les particularités de la CNE sont « en partie déterminé[e]s par l'environnement technologique » puisque ces environnements « produisent à la fois une restriction et une ouverture des actions (d'écriture) possibles ». Plus précisément, en ce qui concerne la production langagière, « [l]a technologie utilisée pour rédiger un message numérique peut [...] déterminer quelques caractéristiques de l'écriture » (18). Il en découle que la variation dont témoigne la CNE est en partie due à la variation diamésique inhérente aux variétés de médiatisation technonumérique.

### **1.6.2.2 Les plateformes socionumériques et la variation diamésique**

La deuxième contrainte sur la production de la CNE est celle créée par la plateforme socionumérique du Web 2.0 elle-même. En effet, d'après David Crystal (2011), la production langagière sur Internet varie selon la modalité en question, que ce soit le courriel, le chat, le SMS, ou le médium socionumérique. Par exemple dans les SMS, d'après Amélie Cougnon, les individus modifient la typologie surtout par l'abréviation, « pour répondre à des besoins d'économie d'espace et de temps » (Cougnon, 2010 : 397). En outre, d'après Laetitia Bibié-Emerit, « Facebook met à disposition des [internautes] des outils technologiques » tels que des boîtes de saisie permettant texte et images par exemple, « qui métissent leur discours et en font des matières technolangagières [...] Les évolutions technologiques du site ont donc un impact direct sur la matérialité des énoncés » (2015 : 260).

Du côté de X/Twitter, la contrainte de longueur pourrait avoir un effet sur les caractéristiques linguistiques utilisées dans les tweets, surtout comparé aux publications sur Facebook. Par exemple, non seulement la limite de longueur, mais également l'emploi de hashtags dans un tweet, pourrait limiter le montant de langage inclus. Ainsi, l'on constate souvent des mots raccourcis, des abréviations et l'omission de pronoms dans les tweets, ce qui pourrait limiter les choix linguistiques de l'internaute (Friginal *et al.*, 2017 : 348). Dans le même ordre d'idées, David

Crystal (2011 : 77) se demande s'il existe une seule variété langagière des médias sociaux, avec des sous-variétés pour chaque plateforme, ou si en revanche les différences de fonctionnalité entre les plateformes sont assez significatives pour qu'on puisse accorder à chacune sa propre variété langagière. Tous ces exemples montrent que la variation diamésique dans le cadre du numérique ne résulte plus simplement des différences entre les codes langagiers (oral, écrit et iconique), mais aussi des différences entre les plateformes socionumériques.

### **1.6.2.3 L'hybridité de la CNE comme fonction de la variation diamésique**

L'hybridité dont témoigne la CNE amène tout d'abord à remettre en question la vraie source de la variation diamésique. Considérons la variation évidente entre l'oral et l'écrit : l'oral se constitue de sons phoniques qui sont combinés pour créer un sens, tandis que l'écrit est un ensemble de marques scripturales qui sont combinées pour former un sens. Ces marques scripturales sont des signes arbitraires qui représentent les sons de l'oral. D'autre part, étant donné que l'oral et l'écrit ne s'emploient pas dans les mêmes situations, les deux codes ont souvent des fonctions différentes, ce qui aura sans doute un effet sur la production langagière dans l'un ou l'autre cas. La variation diamésique traditionnelle s'explique alors par une différence inhérente dans la nature des deux codes. Plus précisément, l'oral et l'écrit représentent deux pôles sur un continuum langagier de « degrés intermédiaires » (Wüest, 2009 : 147), entre lesquels se situent plusieurs variétés de langue (Boutet et Gadet, 2003 : 19).

Dans le cadre du numérique, David Crystal (2011 : 19) note que les productions langagières dans le numérique varient selon leurs ressemblances à l'oral et à l'écrit. Il ne faut pas non plus oublier que la CNE contient ou accompagne d'autres types de codes – les icônes, les symboles, les émojis, les images, les gifs, les vidéos etc. – qui rendent encore plus complexe la notion d'hybridité. D'après Michel Marcoccia (2016 : 72), grâce à son hybridité, la CNE se situerait en position médiane du continuum langagier. En outre, cette hybridité et cette position sont le résultat de critères de formalité (situation formelle ou informelle) et de « dialogalité » (situation de monologue ou de dialogue).

### 1.6.2.3.1 L'oralité comme mesure de l'hybridité

Nous venons de le voir, c'est plutôt la situation, et/ou la plateforme qui détermine le degré d'oralité, et donc le degré d'hybridité, de la CNE. Si la variation diamésique s'explique par les différences entre codes langagiers, et que la CNE est un langage hybride constitué d'éléments oraux et écrits (sans oublier les composantes iconiques), à quoi ressemble cette variation ? La CNE, n'a-t-elle alors pas une variation inhérente ? Et comment situer la CNE sur le continuum langagier ? Quel rôle joue cette hybridité, dans la variation diamésique ? La variation diamésique traditionnelle (variation entre l'oral et l'écrit) est-elle même pertinente ?

Comme l'explique Marcoccia, « [l]es écrits numériques se distinguent ainsi du point de vue des écarts qu'on peut y observer par rapport à l'écrit standard. Par exemple, le Tchat non modéré s'éloigne beaucoup des écrits standard alors que le forum de discussion et le courrier électronique en sont plus proches » (Marcoccia, 2016 : 69). Par conséquent, l'hybridité de la CNE est un résultat de la diamésie. Il faut aussi noter que la diamésie (la plateforme, le dispositif) aura également un impact sur les paramètres de la situation de communication. Cette dimension diaphasique joue aussi sans doute un rôle dans le degré d'hybridité de la CNE.

Une analyse sociolinguistique variationniste de la CNE devient d'autant plus pertinente, puisque « “le français ne connaît pas de forme dévolue à l'oral ou à l'écrit”, mais seulement des différences statistiques quant à la fréquence d'emploi de certaines formes » (Gadet, 2003 : 10, citant Wüest, 2009 : 154). Une telle approche permet alors de quantifier l'hybridité, et le degré d'oralité de la CNE, et ainsi de rendre compte de la variation diamésique, qui est non plus seulement le résultat de la différence entre l'oral et l'écrit, mais de la différence entre chaque médium *techno et/ou socionumérique*.

## 1.7 Conclusion

La CNE, en tant que langage hybride oral-écrit, émerge de l'environnement techno- et/ou socionumérique, comprenant à la fois l'appareil et la plateforme. Mon étude se concentre sur la quantification et la mesure de ce langage particulier. Je m'intéresse à la CNE sur Facebook puisque

cette plateforme se distingue de manière technologique et situationnelle des autres contextes numériques qui ont été étudiés dans le cadre du FQ, notamment le (t)chat et le SMS. Par exemple, les internautes peuvent entrer leur CNE à partir d'une diversité d'appareils, et il s'agit d'une situation plus publique que le SMS et moins rapide que le (t)chat. À l'image des projets de constitution de corpus *texto4science* (Langlais et Drouin, 2012) et *sosweet* (Magué et al., 2015-2020), le présent projet se veut d'établir un corpus Facebook de taille significative, fournissant ainsi une source de recherches futures.

Les résultats de ma recherche contribuent ainsi aux connaissances sur la CNE, se classant avec les résultats des études variationnistes susmentionnées sur le FQ. Les données des études d'Anne-Sophie Bally (2022), Hélène Blondeau *et al.* (2014), de Mireille Tremblay *et al.* (2019), de Rémi van Compernelle (2010), et d'Hélène Blondeau et Mireille Tremblay (2022b), ainsi que les données des études antérieures sur le FQ oral (voir le tableau synthétique en Annexe), fournissent à mon étude des points de repère et de comparaison. Comparer la CNE du FQ sur Facebook aux données du FQ par SMS et par (t)chat permet de placer la CNE de Facebook sur un continuum de langage socionumérique par rapport aux deux autres médias. Comparer la CNE du FQ sur Facebook aux données antérieures de l'oral permet de placer la CNE de Facebook sur un continuum oral-écrit, ou encore, sur un continuum vernaculaire-standard. Ces comparaisons permettent aussi de connaître l'état du FQ, et la relation qu'entretiennent les Québécois.es avec leur langue.

## CHAPITRE 2 : ÉTAT DE LA RECHERCHE ET PERTINENCE DU PROJET

### 2.0 Introduction

Dans ce chapitre, pour situer le présent projet dans son cadre de recherche, je me propose, dans un premier temps, de faire le bilan des études pionnières en sociolinguistique sur la CNE, qui se sont focalisées sur l'effet de la modalité technologique sur la variation linguistique et l'hybridité. Ces travaux ont ouvert le domaine des études sur la CNE, et ont fourni des pistes préliminaires de recherche et de réflexion théorique et méthodologique. Ensuite, je passerai en revue les études variationnistes qui ont analysé la CNE dans le (t)chat, dans les SMS et dans les médias socionumériques, et ce dans plusieurs langues, outre l'anglais et le français. Ce regroupement de recherches, tout en appliquant une méthode quantitative d'analyse de la CNE, montre comment cette dernière varie selon la modalité et selon la variété régionale. Ensuite, j'aborderai cinq études variationnistes sur le français québécois dans l'espace du Web 2.0, notamment dans le SMS et le (t)chat. Enfin, j'explicitai les buts de la présente étude, y compris les questions de recherche et les variantes sociolinguistiques qui figurent dans mon analyse. Ma recherche, qui se pose donc comme un élargissement et un approfondissement des travaux précédents, s'appuie sur les cinq dernières études variationnistes de la CNE, et s'inspire des recherches susmentionnées portant sur d'autres langues et d'autres modalités socionumériques.

### 2.1 Bref historique des études sur la CNE

Les premières études sur la CNE attestent de l'existence de la variation linguistique en ligne. Certaines se sont focalisées sur l'idée de nouveauté graphique ou linguistique, comme celle de Jacques Anis (1998, 1999) qui a introduit la notion de *néographie*, ou celle de David Crystal (2001), avec le concept de *Netspeak*, c'est-à-dire une forme langagière particulière à la communication sur Internet. D'autres linguistes se sont intéressé.e.s à l'effet d'Internet sur la langue et sur la communication, sur les relations interpersonnelles et sur les dynamiques de groupe, aussi bien qu'à l'émergence de structures et de normes sociales, et aux impacts de la communication par Internet sur la société (Herring, 2002 :111). La publication de *Because*

*Internet : Understanding the New Rules of Language* par la sociolinguiste Gretchen McCulloch en 2019 illustre l'évolution de la perception de l'emploi de la langue en ligne. Les linguistes conçoivent maintenant que ce qui s'écrit par le biais du numérique n'est pas une nouvelle langue, mais un nouvel usage de la langue, avec ses propres *règles*. En somme, les recherches montrent que la CNE témoigne de variation langagière depuis le début de ces technologies et que la production langagière par les internautes représente une transgression des normes et de la variation par rapport au standard. Les études de la CNE nous renseignent alors sur la relation entre les différentes modalités techno- et socionumériques et sur la production langagière issue de chacune.

### **2.1.1 Les registres de langue en anglais dans le Web 1.0**

La présente étude s'inspire en partie des recherches sur les registres de langue en anglais dans le Web 1.0 par l'équipe de Douglas Biber (Biber et Egbert, 2016 ; Biber *et al.*, 2017 ; Biber et Egbert, 2018). Ces travaux sur la variation linguistique sur Internet ont cherché à identifier et à classifier les registres de langue dans le Web 1.0. L'équipe a exploité la fonctionnalité de Google pour collecter du contenu textuel de sites Web d'une variété de catégories et de genres. Une fois les sites repérés, l'équipe de Biber a codé et annoté chaque document dans le corpus selon ses caractéristiques communicatives et situationnelles (2016 : 95). Elle a ensuite cherché les fréquences d'emploi de plusieurs éléments du système linguistique, comme dans la présente étude. À travers leurs analyses multidimensionnelles, Douglas Biber et Jesse Egbert (2018 : 215) ont noté que les tendances de la variation linguistique sur le Web 1.0 diffèrent de celles des textes écrits hors Internet. Ils ont constaté des registres narratifs, des registres d'opinion, de persuasion et de conseil, aussi bien que des registres relevant de la production orale.

À la lumière de leurs résultats, Biber et Egbert décrivent le Web 1.0 comme un espace linguistique et situationnel continu. Plus précisément, le Web 1.0 offrirait une variété de registres changeant selon le type de texte et selon la situation communicative. Il est concevable que ce genre de variation à travers différents types de textes et différentes situations s'applique aussi bien aux espaces techno- et socionumériques du Web 2.0. Si l'espace du Web 1.0 offre déjà des preuves de l'existence de différents registres, l'examen du Web 2.0 avec sa composante sociale devient

d'autant plus digne d'intérêt. On ne peut cependant tirer de conclusions sans décrire tout d'abord la variation dans tous les espaces et à travers toutes les modalités du Web 2.0.

L'intérêt des travaux de Douglas Biber réside dans leur design méthodologique – repris en partie dans la présente étude – qui consiste à chercher plusieurs éléments du système linguistique, mais aussi à prendre en compte les caractéristiques communicatives et situationnelles des textes. En entretien avec Eric Frignal, Douglas Biber note que la question fondamentale suivante est à l'origine de sa méthode d'analyse multi-dimensionnelle : « quelles sont les similarités et les différences linguistiques entre l'oral et l'écrit ? » (2013 : 143, ma traduction). Cette question évoque l'un des points d'inspiration pour le présent projet, et l'approche systématique et multidimensionnelle de Biber a alimenté ma démarche.

J'ai ainsi sélectionné la méthode de Biber pour deux raisons principales. Tout d'abord cette optique systématique, qui examine plusieurs parties du système linguistique, telles que le lexique, la syntaxe, et les verbes, fournit une étendue de données compréhensive qui, d'après Biber, permet de « dévoiler des tendances linguistiques importantes » (Frignal, 2013: 143, ma traduction). Par exemple, sa méthode « d'analyse multi-dimensionnelle s'est montrée utile pour prédire la probabilité que des traits linguistiques spécifiques vont varier à travers différents textes » (138, ma traduction). Cette approche a donc le potentiel de faire en sorte que mon étude contribue de manière importante aux recherches sur l'emploi du français parlé au Québec dans le numérique.

Alors que d'autres recherches sur la CNE en sociolinguistique se sont focalisées sur une ou deux variables, la présente étude cherche à présenter une vue d'ensemble de la production langagière par des internautes au Québec, et non l'emploi d'un élément en particulier. En deuxième lieu, il importe de tenir compte du cadre dans lequel la CNE dans l'espace public de Facebook est produite, surtout étant donné la comparaison qui sera faite à d'autres contextes numériques et oraux. Ces questions seront abordées lors de l'analyse des résultats dans le Chapitre 4. Bien que n'examinant pas différents types de textes, mon travail cernera les paramètres de la situation de communication particulière à l'étude afin de rendre compte de ces paramètres dans l'analyse et dans la discussion des résultats. Cette question sera traitée plus en détail dans le Chapitre 3.

## 2.2 L'étude sociolinguistique de la CNE dans le Web 2.0

Les premières recherches sociolinguistiques sur l'emploi de la langue dans le Web interactif datent déjà de plusieurs années. Dès 1991, Kathleen Ferrara et ses collègues ont déterminé que le discours écrit interactif (« *Interactive written discourse* ») était un registre naturel, ou encore une variété de langue hybride qui affiche des caractéristiques de l'oral et de l'écrit. Ces linguistes ont également émis l'hypothèse que les normes d'usage de la langue sont en train de devenir codifiées (Ferrara *et al.*, 1991 : 10). Ensuite, dans son étude des SMS, Crispin Thurlow (2003) a de son côté conclu que ces messages, qui sont à la fois linguistiquement ordinaires et communicativement adeptes, servent les maximes sociolinguistiques de (a) la vitesse et l'économie, (b) la restitution paralinguistique et (c) l'approximation phonologique. Enfin, Michel Marcoccia (2012 : 167) a examiné comment les jeunes emploient la norme scripturale dans des forums de discussion, notant que cette communauté ne partage pas « un ensemble de normes » mais qu'il existe « au contraire, de nombreuses négociations sur ces normes ». Ces premières études sociolinguistiques sur la CNE dans le Web 2.0 ont suggéré les tendances linguistiques en ligne, et ont conduit à des recherches plus précises.

## 2.3 L'étude variationniste de la CNE dans différentes langues

Les sociolinguistes de tous les coins du monde emploient des méthodes quantitatives variationnistes pour étudier la CNE, dans plusieurs langues, dans le Web 1.0 et le Web 2.0. Par exemple, dans son étude quantitative et comparative du contenu linguistique du site Web du journal *Die Zeit* et de sites Wikipédia en allemand, Angelika Storrer (2014) a analysé les différentes variantes de *lol*, une abréviation qui remplacerait l'expression « *laughing out loud* » de l'anglais. Son étude a déterminé que les internautes varient leur style de langue selon le contexte, et surtout selon le but de la communication, et que les internautes s'engagent dans des stratégies de verbalisation à l'écrit marquées par la langue parlée (2014 : 191).

Du côté du Web 2.0, deux études sur les langues de l'Amérique du Sud ont examiné des phénomènes de la communauté linguistique telles que manifestées sur Twitter. Bailey Charland (2010) a examiné la fréquence d'emploi d'éléments verbaux en espagnol par des internautes des

capitales de l'Amérique centrale et de l'Amérique du Sud, tandis qu'Elizabeth Herring (2015) a étudié le contact de l'espagnol et du guarani au Paraguay par le biais de l'emploi des interrogatives. Charland (2010) a trouvé que la formule *tener que* (traduction : *devoir*) était beaucoup plus fréquente que *necesitar* (traduction : *nécessiter, avoir besoin*), peu importe la région géographique. Charland a également examiné la collocation de ces deux expressions avec d'autres verbes, sans tirer de conclusions sur ses résultats par rapport au contexte numérique. Si l'étude a examiné l'emploi d'éléments verbaux sur Twitter, elle n'a pas abordé la fréquence d'emploi de ces verbes dans la communauté linguistique. Quant à l'étude de Herring (2015), les résultats ont indiqué que le guarani sur Twitter varie dans son contenu d'éléments de la langue espagnole en ce qui concerne les marqueurs interrogatifs. Alors que ces deux études montrent la possibilité d'utiliser un réseau socionumérique dans le cadre d'une étude variationniste, elles n'ont pas traité en détail de l'importance du contexte numérique.

En Europe, quelques études variationnistes ont examiné l'alternance stylistique en allemand et en flamand. Tout d'abord, Beat Siebenhaar (2008) a analysé la variété de langue utilisée par des internautes écrivant en allemand suisse, et Kristin Luckhardt (2009) la variation stylistique en allemand, les deux dans le (t)chat. L'étude de Siebenhaar a montré la présence d'une variation stylistique qui reflète les tendances dans la communauté linguistique. Son enquête, comme le fait la présente étude, a également examiné la variation linguistique sur le plan individuel, c'est-à-dire au niveau microlinguistique, et a constaté qu'en dépit d'une variation au niveau individuel, en général les internautes tendent à favoriser une seule variété de langue. Luckhardt (2009), pour sa part, a soulevé dans ses résultats des tendances d'économie, d'expression émotionnelle et de solidarité dans les choix stylistiques des internautes allemand.e.s. Plus récemment, l'équipe de Lisa Hilte *et al.* (2018) a étudié les pratiques langagières d'adolescent.e.s flamand.e.s dans les médias socionumériques. Tout en examinant l'emploi du vernaculaire, leurs analyses ont dévoilé l'impact important de la classe sociale sur les pratiques langagières en contexte numérique. Ce sont des questions dont tient compte la présente étude, surtout dans les Chapitres 4, 5 et 6 lors de la présentation et de la discussion des résultats.

Sur le continent africain, Ana Deumert et Kristin Vold Lexander (2013) ont analysé la variation orthographique dans les SMS d'internautes plurilingues de cinq pays africains, et en Asie, Yi Zhang (2020) a examiné l'emploi du japonais sur un site socionumérique multimodal chinois. Deumert et Lexander (2013) ont trouvé que les internautes, tout en démontrant une « dextérité linguistique », puisaient dans des ressources linguistiques locales et internationales en écrivant leurs messages. L'étude de Zhang (2020), quant à elle, a montré que la CNE sur le site socionumérique en question était une langue hybride, incluant des éléments du japonais dans le texte chinois. Dans le cadre d'une étude plus globale sur l'emploi des émojis, l'équipe de Xuan Lu *et al.* (2016) a examiné la CNE d'internautes de 212 pays, ainsi que la co-occurrence d'émojis par pays. Ils ont déterminé que l'emploi des émojis variait selon le pays et selon la culture, et que quelques groupes de pays partageaient la cooccurrence de certains émojis. Leur étude a alors renforcé l'idée que la culture joue un rôle dans l'emploi des émojis dans la CNE. Alors que la présente étude ne se préoccupe pas de l'emploi des émojis, la question de la culture et de la production langagière en ligne sera reprise dans le Chapitre 6 lors de la discussion des résultats.

Enfin, plus récemment, Benedikt Szmrecsanyi et Alexandra Engel (2022) ont analysé l'emploi de trois éléments grammaticaux dans la CNE en anglais et en néerlandais afin d'examiner l'effet de la modalité et de la formalité sur le registre de langue. Ces derniers ont trouvé que les registres formels sont plus complexes que les registres informels. Le travail de Szmrecsanyi et Engel est particulièrement pertinent dans le cadre de la présente étude, qui compare le contenu linguistique dans des contextes technologiques et non technologiques.

Tous ces travaux illustrent divers aspects de l'analyse quantitative variationniste mise en place pour étudier la CNE dans plusieurs langues, et pas seulement en anglais et en français. Ils confirment aussi que la présente étude s'insère clairement dans un champ et une méthode de recherches florissantes, non seulement au Canada ou en Amérique du Nord, mais aussi sur le plan mondial. Enfin, le bilan des travaux dans le domaine montre qu'il y a un besoin d'effectuer une recherche plus étendue du système linguistique, ce qui est fait ici. À part un petit nombre d'enquêtes (Biber et Egbert, 2016 et 2018 ; Friginal *et al.*, 2017 ; Szmrecsanyi et Engel, 2022 ; Tagliamonte et Denis, 2008), la majorité des travaux sur la CNE se focalisent sur un ou deux éléments linguistiques seulement.

## **2.4 L'étude variationniste de la CNE en anglais dans le Web 2.0**

Les recherches dont traite cette section ont également employé des méthodes variationnistes pour étudier la CNE, dans ce cas en anglais. Leur intérêt réside dans le fait qu'elles rejoignent les bases méthodologiques de la présente étude. Elles ont chacune dégagé et calculé les fréquences d'emploi relatives de plusieurs éléments du système linguistique, pour en tirer des conclusions sur le comportement linguistique des internautes.

### **2.4.1 Les messageries instantanées**

Du côté de l'anglais au Canada, dans le cadre de leur étude de messages instantanés (MI), Sali Tagliamonte et Derek Denis (2008) ont examiné des éléments grammaticaux ainsi que des éléments de la graphie SMS. Ils ont conclu que le MI est un nouveau registre unique, qui témoigne d'une combinaison de toute la variation de la communauté linguistique — formelle, informelle et extrêmement vernaculaire (Tagliamonte et Denis, 2008 : 3). De plus, ils ont noté que la langue des MI reflète la même variation et les mêmes processus d'évolution linguistique actuels dans les variétés modernes de l'anglais. Nous verrons dans les sections qui suivent que les études sur le français québécois soutiennent cette conclusion. Il s'agira alors de vérifier dans quelle mesure les mêmes constats s'appliquent à la CNE produite par des internautes au Québec sur Facebook.

### **2.4.2 Les médias socionumériques**

Une autre étude notable est celle d'Eric Friginal et de ses collègues (2017), qui ont comparé les tendances dans la variation microlinguistique en anglais sur Facebook et sur Twitter. Cette étude s'est intéressée surtout aux différences de modalité, donc aux paramètres technologiques qui opèrent sur la production langagière. Comme dans la présente étude, l'équipe de Friginal a constitué un corpus de publications accessibles publiquement, afin de chercher plusieurs éléments grammaticaux, syntaxiques et sémantiques. Elle a ensuite employé la méthode d'analyse variationniste multidimensionnelle de Douglas Biber (Biber et Egbert, 2016 et 2018) afin d'examiner la CNE selon le sujet de publication.

L'étude de Friginal et de ses collègues touche à la question de la variation diamésique – c'est-à-dire la variation selon le médium, vu de manière traditionnelle comme étant les différences entre l'oral et l'écrit – engendrée par les particularités de chaque plateforme socionumérique. Friginal et son équipe ont ainsi examiné comment les fonctionnalités des deux plateformes avaient un impact sur les tendances linguistiques. L'étude a permis de constater que contrairement à Facebook, qui est moins contraignante, la limite de longueur du message sur Twitter pourrait déterminer quels éléments linguistiques sont utilisés dans les tweets. Plus précisément, la recherche a noté l'emploi plus fréquent de troncations et d'abréviations, et de l'omission de pronoms dans les tweets, pouvant mener, selon Friginal et ses collègues, à limiter les choix linguistiques (2017 : 348). De plus, leurs résultats ont montré des différences notables entre les publications sur les deux médias socionumériques, suggérant que chaque plateforme contient un registre bien distinct. Cette étude est importante dans le cadre du présent travail, puisque cela suggère que la variation du FQ trouvée sur Facebook différera de la variation retrouvée dans le (t)chat et dans les SMS. De plus, l'équipe de Friginal a prouvé que les méthodes d'analyse sociolinguistique variationnistes quantitatives peuvent s'appliquer à la CNE d'origine publique Facebook.

Les analyses variationnistes de la CNE en anglais dans le Web 2.0 sont révélatrices de parallèles entre les tendances linguistiques dans la communauté et celles que l'on relève dans le Web 2.0. Elles suggèrent aussi une réorganisation des normes linguistiques pour correspondre aux paramètres technologiques et sociaux du Web 2.0. J'explorerai dans les sections suivantes comment ces mêmes tendances semblent se reproduire dans la CNE en français.

## **2.5 L'étude variationniste de la CNE en français dans le Web 2.0**

Cette section fait le recensement des études de la CNE en français dans différentes modalités de clavardage, se focalisant surtout sur deux éléments du système linguistique, notamment l'emploi ou non de la particule *ne* dans la négation et l'alternance des pronoms *on* et *nous*. Il est à noter que les études de van Compernelle et Williams (2007) et de van Compernelle (2008a et 2008c) ne précisent pas de quelle variété du français il s'agit. Dans le cas de la première étude, les données ont été extraites de six sites Web différents, dont un, *Cyberpresse.ca*, de provenance canadienne.

Cela laisserait croire qu'une partie des données serait alors en français québécois, mais cette nuance n'est pas indiquée dans les détails de l'étude. L'étude de Compernelle (2008a), quant à elle, ne précise pas la source des données analysées.

### 2.5.1 Le français européen et québécois dans le (t)chat

Pour ce qui est de l'étude de la négation, Rémi van Compernelle et Lawrence Williams ont examiné la variation linguistique dans le (t)chat, le forum de discussion et le clavardage modéré (t)chat, trouvant que la particule *ne* dans la négation « n'apparaît que très rarement », c'est-à-dire dans moins de 5 % des cas (2007 : 68). Dans une autre étude, Rémi van Compernelle (2008a : 325) a constaté un taux d'emploi de la particule *ne* dans la négation d'environ 12 %. En revanche, une étude de Lawrence Williams (2009) sur l'emploi du *ne* dans les (t)chats modérés et non modérés français et FQ révèle que la particule *ne* était conservée dans presque tous les cas. Les 11 instances dans lesquelles la particule était omise provenaient de la composante québécoise du corpus. Toutefois, Williams doute de la validité de ses résultats, parce qu'il n'a recueilli aucun renseignement sur l'âge ou la région géographique des internautes. En effet, Williams avoue que le contenu sociolinguistique de la plupart des commentaires suggère que les internautes provenaient d'une variété de pays francophones :

*[t]his slightly lower ne retention rate in the Cyberpresse data set is not unexpected, given the much less frequent use of ne reported in previous studies using data from Québec compared to data from France; however, it is far from certain that ne deletion has anything to do with the geographic origin of participants since no biographical information is available. In fact, many of the comments and statements in the Cyberpresse data set suggest that there were participants from a variety of francophone countries. The lack of verifiable information about participants severely limits the analysis of external social factors (2009 : 484).*

Géraldine Blattner et Lawrence Williams ont également entrepris une étude variationniste de l'emploi du *ne* dans la CNE en français européen provenant de la modalité de (t)chat sur le site Web *europnet.org* (2011). D'après leurs analyses, les internautes de la cinquantaine ont tendance à maintenir la particule *ne* plus souvent que les plus jeunes. La première tranche d'âge conserve le *ne* dans environ 34 % des cas, tandis que les plus jeunes dans la vingtaine, la trentaine et la quarantaine utilisent le *ne* entre 15 et 20 % du temps environ. Blattner et Williams avancent que l'« environnement électronique est favorable à la chute du *ne* tout comme l'est la conversation

orale » (2011 : 117). La présente étude se donne donc pour objectif de mettre à l'épreuve cette hypothèse dans le contexte de la CNE produite par des internautes au Québec.

Enfin, pour ce qui est de l'alternance *on* et *nous*, van Compernelle (2008c) a déterminé que le pronom *on* est préféré dans le (t)chat à un taux d'environ 95 %, comparativement au pronom *nous*. Il est à noter que son étude s'est penchée uniquement sur les données provenant d'un site Web français européen, *epiknet.org*, un des sites également utilisé dans l'étude susmentionnée de van Compernelle et Williams (2007). Il est donc douteux que ce site contienne de la CNE écrite en français québécois. D'après van Compernelle, ces résultats s'alignent sur les tendances de l'oral, et il souligne qu'afin de mieux comprendre la CNE, il importe de se pencher davantage sur la variation sociolinguistique dans d'autres modalités technonumériques. La présente étude répond donc à ce besoin d'étudier d'autres contextes numériques.

En somme, les paramètres de chaque modalité de (t)chat, ainsi que les paramètres sociaux des internautes, produisent des fréquences d'emploi différentes. Les études existantes témoignent aussi de la variabilité du contenu sociolinguistique de la CNE en français, et même de résultats parfois contradictoires dans différentes modalités du Web 2.0. Elles se focalisent sur l'oralité de la CNE sans expliquer en détail l'importance du contexte techno- et siconumérique. L'objectif de la présente étude est donc de déterminer comment ces tendances se traduisent dans toutes les parties du système linguistique. De plus, la présente recherche répond au besoin d'analyser plus profondément l'emploi du français québécois en ligne, en délimitant le facteur externe de la géographie.

### **2.5.2 Le français européen et les appareils numériques**

Une autre étude récente a traité de la variation linguistique en français dans le (t)chat. En raison des avancées technologiques et de la multiplication des types d'appareils numériques, Géraldine Blattner et Lawrence Williams (2022) ont voulu réexaminer la CNE dans le (t)chat européen et la comparer à la CNE de leur corpus de 2008. Ils notent que même avec les développements technologiques des dernières années, il existera sans doute toujours des contraintes sur la production des caractères sur chaque plateforme et dans tous les contextes (2022 : 224). En effet,

leur analyse diachronique a révélé que l'emploi des accents est beaucoup plus favorisé chez les internautes dans la vingtaine, qu'il ne l'était dix ans plus tôt. Blattner et Williams présument que cela est en raison du développement de nouveaux appareils tels que les téléphones et montres intelligents ou les tablettes électroniques, et à l'amélioration de différentes technologies telles que l'autocorrection et les logiciels de saisie vocale (2022 : 240). Blattner et Williams font alors allusion à l'importance probable des paramètres technologiques et communicatifs de la CNE, sans en tirer des conclusions concrètes et sans s'interroger sur la question de la diamésie.

Toutes ces études sur la CNE dans le (t)chat ont examiné la variation linguistique d'un point de vue quantitatif et parfois qualitatif. Outre l'étude de Blattner et Williams (2011), elles n'ont pas analysé les facteurs sociaux tels que l'âge, le sexe, etc., puisque souvent ces données sont inaccessibles ou peu fiables dans des espaces Web 2.0. C'est aussi le cas pour la présente étude qui s'inscrit cependant dans un champ d'études et une méthode déjà bien établis.

### **2.5.3 Le français européen dans les SMS**

En ce qui a trait aux SMS, Elisabeth Stark a examiné les fréquences d'emploi des variantes de la négation (2012) et des pronoms clitiques (2014) à partir des données Franco-suissees dans le corpus *sms4science*, cette composante d'un corpus plus large ayant été élaborée par Stark *et al.* (2009-2015). La première étude de Stark (2012) a révélé une fréquence d'emploi du *ne* dans la négation d'environ 24 %, ce qui représente un taux plus élevé qu'à l'oral et que dans le (t)chat (van Compernelle, 2008a). Cela pourrait alors indiquer que l'influence du standard écrit a, dans ce cas, un impact sur la CNE. Sa deuxième étude a constaté que « les variantes de l'orthographe standard sont [...] les variantes graphiques par défaut » (2014 : 145). En revanche, son analyse a également « mené à des résultats contradictoires » entre l'emploi des pronoms clitiques sujets et des pronoms clitiques objets dans les SMS (144). Stark a conclu ses deux études avec peu de certitude, proposant de futures recherches sur l'emploi du français dans les SMS.

Dans une troisième étude, Elisabeth Stark et Isabelle Riedel (2012) ont examiné l'emploi du participe passé dans les SMS franco-suisse. Elles ont trouvé que l'accord du participe passé suivait le standard écrit à un taux d'environ 90 %. En revanche, avec l'auxiliaire *avoir*, le participe passé était accordé suivant le standard dans seulement 77 % des cas. Alors que Stark et Riedel ont conclu que « le “degré de déviance” de l'orthographe du français standard contemporain est minime », elles ont également noté « le caractère particulièrement conservateur du corpus suisse » (2012 : 134). On voit alors un décalage entre les fréquences d'emploi du *ne* dans les SMS comparativement au (t)chat, qui est certainement dû aussi aux variétés régionales.

Les résultats des études de Stark et de ses collègues font également ressortir une variabilité dans l'emploi des formes standards et non standards au sein d'un même corpus. On pourrait proposer l'idée d'une variabilité de l'hybridité dans les SMS, qui s'expliquerait par plusieurs facteurs sociaux et technologiques. Ce sont des questions que va examiner la présente étude dans le cadre de la CNE sur Facebook, une plateforme qui permet la saisie de texte par téléphone autant que par clavier d'ordinateur : les analyses de collocation permettront de cerner plus précisément cette variabilité linguistique.

#### **2.5.4 Les variétés de français dans les médias socionumériques**

Si les analyses variationnistes dans le Web 2.0 sont peu nombreuses, celles portant sur la CNE en français dans les médias socionumériques sont encore plus rares. À ma connaissance, la seule étude qui s'intéresse au FQ dans un médium socionumérique est celle de Gretchen McCulloch et Jeffrey Lamontagne (2020). Cette équipe a exploré la représentation de la phonologie dans trois variétés du français sur Twitter, notamment les français laurentien (variété parlée au Québec), septentrional (variété parlée au nord de la France) et méridional (variété parlée au sud de la France). Cette étude avait notamment pour objectif d'analyser l'emploi répétitif et exagéré de consonnes et de voyelles, afin « de déterminer si des différences dialectales sont communiquées à l'écrit » (171). Cette recherche a alors comparé, à travers une analyse statistique, les fréquences de répétition des lettres. McCulloch et Lamontagne ont ainsi pu constater que « la répétition des lettres sur Twitter sert à éclaircir des différences entre des variétés de français » (184). Bien que la méthode employée dans la recherche de McCulloch et Lamontagne ne soit pas tout à fait dans la même tradition que la

présente étude, elle suggère cependant que les internautes du même dialecte vont suivre des tendances linguistiques de la communauté dans leur CNE. Il s'agira donc ici d'examiner de quelle manière les internautes emploient le FQ sous forme de CNE et dans quelle mesure ils et elles suivent les tendances linguistiques de leur parler régional.

## **2.6 L'étude variationniste du français québécois dans le Web 2.0**

Les seules recherches sociolinguistiques variationnistes qui étudient uniquement la CNE en français québécois, à ma connaissance, consistent en un nombre restreint d'études entreprises entre 2010 et 2022 (Bally ; 2022 ; Blondeau *et al.*, 2014 ; Blondeau et Tremblay, 2022b ; Tremblay *et al.*, 2019 ; van Compernelle, 2010). À part la plus récente par Blondeau et Tremblay (2022b), qui a fait un survol global d'éléments linguistiques, ces travaux ont examiné une ou deux parties seulement du système linguistique. Les chercheurs et chercheuses ont tous comparé leurs résultats aux fréquences d'emploi attestées à l'oral et certains, sans faire de comparaison quantitative, ont discuté généralement des tendances trouvées dans d'autres études sur la CNE. Par exemple, l'étude de 2014 par Blondeau *et al.* a fait allusion aux procédés repérés dans le clavardage en anglais par l'étude de Tagliamonte et Denis (2008). Cette série d'études représente une suite à la tradition des recherches variationnistes sur le FQ menées depuis les années 1960 au Canada. La présente étude, tout en s'inscrivant dans cette tradition, répond au besoin de dresser le tableau du système linguistique du FQ en ligne, en cherchant et en analysant un plus grand nombre de variables sociolinguistiques, en faisant des analyses de collocation, et en examinant un terrain autre que le SMS ou le (t)chat. En outre, les résultats des études discutées dans cette section 2.6 seront repris dans le Chapitre 4 lors de la comparaison des résultats de la présente étude à ceux d'autres contextes numériques.

### 2.6.1 L'emploi du *ne* dans le (t)chat

Partant de la prémisse que les fréquences d'emploi de la particule *ne* varient selon le temps, l'endroit et le contexte communicatif (447), Rémi van Compernelle (2010) a étudié son emploi en français montréalais dans le (t)chat. Il a ensuite comparé ses résultats aux fréquences d'emploi présentes dans d'autres modalités de la CNE. Notamment, cette comparaison s'est faite au (t)chat français européen (voir van Compernelle, 2007), et à des annonces de rencontres en ligne de la région de Montréal (voir van Compernelle, 2008b). Ensuite, il a fait une comparaison aux fréquences attestées à l'oral en français canadien (Poplack et St-Amand, 2007 ; Sankoff et Vincent, 1977). Ce chercheur a déterminé que l'emploi du *ne* dans le (t)chat montréalais (5 %) était moins fréquent que dans d'autres variétés du français en ligne (16 % dans le (t)chat européen et 78 % dans les annonces), mais plus fréquent qu'en français montréalais parlé, soit entre 0,2 et 0,5 %. Cette étude est importante dans le cadre du présent travail, puisqu'elle a comparé l'emploi du *ne* dans différentes modalités technonumériques et dans différents contextes communicatifs, tant en ligne qu'à l'oral. Cependant, van Compernelle n'ayant ni analysé ni discuté ces différences en grand détail, il s'agira donc d'établir des comparaisons afin de discuter davantage de la problématique langue et technologie.

### 2.6.2 L'emploi des pronoms dans les SMS

Tout en puisant dans le corpus *Texto4Science* (Langlais et Drouin, 2012), Hélène Blondeau, Mireille Tremblay et Patrick Drouin (2014) se sont penchés sur deux éléments pronominaux, notamment l'alternance *on* vs *nous*, et l'alternance des pronoms simples et composés au pluriel – *nous* vs *nous-autres*, *vous* vs *vous-autres*, *eux* vs *eux-autres* –, qui varient de manière importante à l'oral en français montréalais. En particulier, cette équipe note que

[ ]a comparaison du comportement de ces deux variables dans les textos et à l'oral guide notre réflexion sur la place singulière qu'occupe ce nouveau genre de communications sur les diverses dimensions variationnelles et, plus particulièrement, sur le continuum entre l'oral et l'écrit. L'analyse met en relief comment les scripteurs exploitent les ressources linguistiques, ici différentes variantes d'une variable linguistique, dans le cadre des interactions orales ou médiatisées par ordinateur (2014 : 139).

Ces propos s'alignent sur la démarche de la présente étude qui prend en compte les dimensions de la variation linguistique dans le cadre des modalités de la CNE. Les résultats de l'étude de Blondeau *et al.* (2014) révèlent qu'en ce qui concerne l'alternance des pronoms *on* et *nous*, le premier était employé à une fréquence d'environ 96 %, une fréquence très proche de celle qui est attestée à l'oral, en l'occurrence 99 % (Laberge, 1977). Pour ce qui est de l'alternance des formes simples et composées des pronoms au pluriel, l'étude a déterminé que ce sont les formes simples qui sont préférées dans les SMS. En effet, la forme composée apparaît une fois seulement sur un total de 79 occurrences. Ces résultats suggèrent, d'après Blondeau et ses collègues « que le changement en cours à l'oral a continué d'avancer [...] au point où les formes en *-autres* deviennent marginales » (2014 : 155). En somme, Blondeau *et al.* ont trouvé que « les textos [se] rapprochent considérablement » de l'oral, mais qu'on constate certaines différences. Ces divergences s'expliqueraient en partie par une préférence pour des formes plus simples en raison de l'économie de place lors des échanges SMS (Blondeau *et al.*, 2014 : 161). La présente étude se propose donc d'examiner comment ces tendances se traduisent dans une plateforme moins restreinte comme Facebook autant sur le plan de l'entrée textuelle que sur le plan du contexte social public.

### **2.6.3 L'emploi du futur dans les SMS**

Plus récemment, Mireille Tremblay, Hélène Blondeau et Emmanuelle Labeau (2020) ont examiné la variation dans l'expression du futur dans les SMS de francophones de la Belgique et du Québec – notons qu'il s'agit du même corpus que celui de l'étude effectuée en 2014, *Texto4Science* (Langlais et Drouin, 2012). Plus précisément, il s'agit dans cette étude de l'alternance entre le futur synthétique (futur simple - *j'irai*) et analytique (futur proche - *je vais aller*), ainsi que de l'emploi du présent (*je vais*) pour représenter le futur. Comme c'est le cas pour l'étude de 2014 par Blondeau *et al.*, cette équipe a comparé les fréquences d'emploi dans les SMS aux fréquences attestées à l'oral. L'étude a trouvé que les individus emploient le présent comme futur beaucoup plus fréquemment dans les SMS (40 % du temps), qu'à l'oral (10 % des cas). D'autre part, le futur analytique est préféré dans les SMS aux dépens du futur synthétique, ce qui rejoint également les tendances des situations informelles à l'oral (2019 : 92). Enfin, les résultats de cette étude suggèrent que les messages SMS ne suivent pas les règles du standard écrit, mais adoptent plutôt différentes normes qui se rapprochent de la langue orale (Tremblay *et al.*, 2019 :

92). Plus tard, d'après ces mêmes données, Mireille Tremblay a identifié « quatre points de convergence entre les messages textes et l'oral familier », notamment sur le plan de la variation diatopique, la variation diastratique, la nature conversationnelle des échanges et la proximité entre ceux participant dans la conversation (Tremblay, 2020 : 88). Je discute des convergences de la CNE sur Facebook avec les SMS et l'oral dans le Chapitre 6.

#### **2.6.4 L'emploi des formules interrogatives dans les SMS**

Une étude récente entreprise par Anne-Sophie Bally (2022), puisant dans le corpus *Texto4Science* (Langlais et Drouin, 2012), a examiné la variation des formules interrogatives directes ou totales en français québécois dans les SMS. Il s'agit plus précisément de l'alternance entre les formes suivantes : l'inversion du sujet et du verbe, l'utilisation de *est-ce que*, la forme sans inversion, et l'ajout du suffixe *-tu* après le verbe comme dans *Il veux-tu bien écouter ?*. Comme dans la présente étude, elle a comparé les fréquences d'emploi des formules interrogatives dans les SMS aux fréquences attestées à l'oral, et à celles trouvées par van Compernelle (2010), Blondeau *et al.* (2014) et Tremblay *et al.* (2019). Bally a déterminé que tout comme à l'oral, la forme sans inversion est la plus fréquente et la forme en *est-ce que*, la moins fréquente. Quant aux deux autres formes, la forme en *-tu* serait plus fréquente à l'oral que dans les SMS, et la fréquence d'emploi de la forme de l'inversion du sujet et du verbe serait à peu près similaire à l'oral et dans les SMS. Bally (2022 : 13) conclut que « l'écrit SMS est une forme hybride de message écrit, dans laquelle les scripteurs se permettent d'introduire des formes de l'oral, c'est-à-dire de transgresser en partie le code écrit. » La question se pose alors : qu'en est-il de ces tendances dans l'espace public et moins contraignant de Facebook ?

#### **2.6.5 Variation et normes dans les SMS**

Une dernière étude récente sur le FQ menée par Hélène Blondeau et Mireille Tremblay (2022b) puisant également dans le corpus *texto4science* (Langlais et Drouin, 2012) s'est penchée, à travers un survol de plusieurs éléments linguistiques, sur la variation et les normes communautaires dans les SMS. Ces dernières notent que « la nouvelle scripta vernaculaire répond à des normes communautaires. Très variable, elle est empreinte de marques sociostylistiques de toutes

sortes » (121). Elles concluent également que « les textos reflètent une grammaire vernaculaire formellement distincte » (130) du français de référence. Elles discutent de la notion de variabilité, c'est-à-dire du fait que la CNE du SMS elle-même fait preuve de variation dans son emploi de variantes tantôt formelles, tantôt informelles. Elles précisent que la CNE témoigne d'une hétérogénéité dans son contenu de variantes standards et vernaculaires ; dans certains cas, les variantes standards sont préférées, et dans d'autres, les internautes tendent vers l'oralité.

Ces deux aspects méthodologiques de l'étude de Blondeau et Tremblay s'alignent sur la démarche adoptée dans la présente étude, qui analyse plusieurs niveaux du système linguistique – dans ce cas de manière quantitative en utilisant par ailleurs des analyses de collocation afin de mieux cerner cette *variabilité*. Quelques recherches seulement se sont penchées sur la collocation dans le cadre des études sur le français (Billiez et Buson, 2013 ; Coveney, 2011 ; Mougeon et Rehner, 2015), mais c'est une méthode dont se servent des sociolinguistes à l'Université de Lancaster (Březina, 2014 ; Březina *et al.*, 2015 ; Březina et Fox, 2021 ; Březina et Pořízka, 2021 ; Oksuz *et al.*, 2021 ; Williams, 1998 et 2002), dans le cadre de leurs études sur des corpus de langue anglaise écrite. Ainsi, la présente étude élargit les recherches sur le sujet et contribue aux connaissances sur la production langagière par des locuteurs du FQ dans l'espace du Web 2.0.

## **2.7 Synthèse de l'état de la recherche**

Ce recensement des écrits témoigne de l'intérêt pour la variation linguistique qui existe en anglais, en français et dans d'autres langues, et ce pour toute une gamme de modalités technologiques, que ce soit le Web 1.0 ou le Web 2.0. Les recherches ont examiné la langue écrite dans des contextes présentant différents paramètres et contraintes techno- et socionumériques. La diversité des résultats montre à quel point le contenu linguistique de la CNE varie du point de vue social et stylistique, même au sein d'une modalité particulière. La plupart des études antérieures évoquent également les notions de négociation et de coconstruction de normes linguistiques dans le Web 2.0 pour traduire une apparente incertitude des internautes vis-à-vis des normes linguistiques à adopter dans le Web 2.0.

Les recherches citées dans ce chapitre invitent à plus de réflexions sur le comportement linguistique dans le Web 2.0. Tout d'abord, si elles mentionnent le rôle possible joué par la technologie dans ces productions langagières, elles n'ont pas examiné en grand détail les conséquences de la diamesie. Elles n'ont pas non plus décortiqué la signification possible de l'hybridité dans ces communications. Le terrain le moins examiné est manifestement l'espace des médias socionumériques, en particulier Facebook, surtout en ce qui concerne le français québécois. En outre, les cinq études variationnistes sur le FQ dans le Web 2.0 ont seulement traité d'une ou deux parties du système linguistique ; c'est pour cela que la présente étude examinera un éventail de marqueurs sociolinguistiques de plusieurs parties du système linguistique. De plus, aucune de ces cinq études n'a effectué de comparaison avec les résultats provenant d'autres contextes CNE. Enfin, seule l'étude de Blondeau et Tremblay (2022b) fait allusion à la collocation, c'est-à-dire à la cooccurrence de différentes variantes dans la CNE d'un même scripteur. Ce sont par conséquent des questions qui doivent être examinées plus en détails, et sur lesquelles se penche donc la présente étude.

Ma recherche s'aligne sur l'analyse systématique de la CNE dans des textes numériques, comme celles de Tagliamonte et Denis (2008), Biber *et al.* (2016, 2017 et 2018) et Friginal *et al.* (2017), pour mieux comprendre le comportement linguistique des internautes qui écrivent en FQ. Cette démarche fournit une vue d'ensemble du système macrolinguistique, ce qui va au-delà des recherches existantes sur la CNE en FQ, qui, nous l'avons vu, se focalisent sur un ou deux éléments seulement du système linguistique. De plus, elle permet de tirer davantage de conclusions sur l'apport de la diamesie, étant donné les points de repère de l'oral et des autres médiums comme le SMS. L'examen de collocations de variantes chez les individus facilite aussi la quantification de la variabilité linguistique en ligne. Enfin, étant donné les résultats des études de la CNE en FQ dans le (t)chat et par SMS, il importe de voir ce qu'il en est sur le plan sociolinguistique sur Facebook. Par exemple, le thème récurrent dans toutes ces études est que la CNE semble suivre les tendances de la communauté linguistique à l'oral. Est-ce que cela s'applique à la CNE en FQ sur Facebook ? La CNE sur Facebook contient-elle autant de traces de l'oral que d'autres modalités telles que le (t)chat et le SMS ?

## **2.8 La pertinence du projet**

Le but global de la présente étude est donc d'examiner le comportement linguistique des internautes qui communiquent en FQ à l'écrit sur une plateforme socio- et technonumérique précise, notamment Facebook. Les données qualitatives permettront d'attester l'emploi, ou non, de ces marqueurs sociolinguistiques sur Facebook, et les données quantitatives permettront de déterminer combien de ces constructions se retrouvent dans le corpus (Rosenbach, 2007). Pour ce faire, j'effectue une analyse de la variation sociolinguistique de plusieurs éléments du système linguistique, tant au niveau macro- qu'au niveau microsociolinguistique. Plus précisément, cette étude a pour but de déterminer à quelles fréquences les membres de la communauté linguistique FQ utilisent les variantes sociolinguistiques dans leurs communications écrites sur Facebook, comparativement à l'oral. Ceci permettra de déterminer le degré d'oralité de la CNE en FQ sur Facebook, en comparant les résultats aux données antérieures sur le FQ oral. En effet, d'après Michel Marcoccia (2016 : 55) « dans les travaux pionniers, on trouve souvent la question suivante : les écrits numériques sont-ils plus proches de l'écrit ou de l'oral ? ».

## **2.9 Les questions de recherche**

Comme je l'ai mentionné dans l'introduction, cette étude tente de répondre à deux grandes questions :

- I. La CNE québécoise sur Facebook ressemble-t-elle plutôt à la langue standard écrite québécoise, ou à la langue orale québécoise ?
- II. Quelles sont les tendances dans la variabilité des emplois standards et non standards ?

Plus précisément, l'étude répondra aux questions ci-dessous, sur la CNE écrite sur Facebook par des internautes au Québec :

1. Quelles sont les fréquences d'emploi de huit variables sociolinguistiques lexicales, verbales, syntaxiques et pronominales ?

Dans le cadre de cette question, les variables suivantes feront l'objet d'analyse :

- a. Le système lexical
  - L'emploi variable des marqueurs de contrainte : *Juste, seulement, rien que, ne ... que*
  - L'emploi variable des marqueurs de conséquence : *Alors, donc, fait que, so*
- b. Le système verbal
  - L'emploi variable du futur analytique (proche) et synthétique (simple)
  - L'emploi variable des auxiliaires *avoir* et *être* au passé composé
- c. Le système syntaxique
  - L'emploi variable de la particule *ne* dans la négation
  - L'emploi variable des formules interrogatives : *Est-ce que*, inversion, *-tu*, aucune inversion
- d. Le système pronominal
  - L'emploi variable des pronoms *on* et *nous*
  - L'emploi variable des pronoms simples et composés au pluriel : *Nous, nous-autres ; vous, vous-autres ; eux, eux-autres*

Pour ce qui est de la deuxième grande question, qui est de nature plus exploratoire – Quelles sont les tendances dans la variabilité des emplois standards et non standards ? – je tenterai de répondre à la question suivante :

2. Quelles sont les cooccurrences des variables lexicales, syntaxiques, et de l'emploi du futur, avec des formes non standards ?

Pour répondre à la première grande question, j'effectue une analyse quantitative de la CNE en FQ sur Facebook, permettant de dégager les fréquences d'emploi des marqueurs, ou variantes, sociolinguistiques. Ensuite, je compare ces fréquences à celles attestées à l'oral, dans le but de déterminer à quelle fréquence les Québécois.es, dans leur production écrite au sein de l'espace public sur Facebook, utilisent les variantes typiques de l'oral ou celles qui s'alignent sur le standard écrit. Dans chacun des regroupements de marqueurs ci-dessus certains sont, selon les études antérieures, typiques du FQ oral, et « s'écartent fortement de la norme prescrite » (Barbarie, 1982 : 145), tandis que d'autres s'associent « nettement au contexte de formalité » (Blondeau, 2000 : ii)

et sont alors associés à la langue standard et à la norme écrite. Enfin, je compare mes résultats à ceux d'autres contextes numériques (Bally, 2022 ; Blondeau *et al.*, 2014 ; Blondeau et Tremblay, 2022b ; Tremblay *et al.*, 2019 ; van Compernelle, 2010).

La deuxième grande question a pour but d'explorer la variabilité entre les formes standards et non standards de commentaire en commentaire. Cette analyse plus restreinte sert à explorer de manière plus précise les tendances de groupe en examinant le contenu sociolinguistique de commentaires individuels. Suivant l'approche systématique de Douglas Biber (1988), par le moyen d'analyses de cooccurrences, j'examinerai de plus près les commentaires qui contiennent une variante à l'étude, afin de déterminer si le commentaire est uniforme dans son contenu, c'est-à-dire, si le contenu est uniquement de type oral, de type standard, ou un mélange des deux. D'après Biber, les tendances de cooccurrence sont centrales à la variation, et ce genre d'analyse est bien conçu pour la description des relations entre les genres écrits et oraux d'une langue (Biber, 1988 : 56). Peu de recherches en sociolinguistique ont traité de la cooccurrence de variantes en français et en FQ. Aussi mon analyse de cooccurrences permettra-t-elle d'apporter des précisions sur ce sujet, en faisant des constatations sur les choix des individus, et en examinant la variation au niveau du commentaire. Cette analyse fera la lumière sur la notion de norme(s) communautaire(s) chez les internautes lorsqu'ils communiquent en FQ dans le Web 2.0.

## **2.10 Conclusion**

Ce chapitre a présenté le bilan des recherches en sociolinguistique qui se sont focalisées sur la communication par le biais des technologies numériques. Dans un premier temps, il a été question de revoir les études préliminaires sur le Web 1.0 et le Web 2.0, qui ont ouvert la voie au domaine florissant que nous connaissons actuellement. Dans un deuxième temps, la présentation de l'analyse multidimensionnelle des écrits dans le Web 1.0 par l'équipe de Douglas Biber (2016 et 2018), celle des premières études en sociolinguistique sur la communication dans le Web 2.0, ainsi que le survol des études variationnistes sur le plan mondial ont permis de placer la présente étude dans son espace de recherche. Les sections suivantes du chapitre se sont concentrées sur les résultats des études variationnistes du français hexagonal et québécois, afin de montrer en quoi la présente étude ajoute aux connaissances, en élargissant l'étendue des données et en traitant d'un

terrain inexploré. Faire le tour des recherches les plus pertinentes dans le domaine a donc permis d'arriver aux buts principaux de la présente étude et aux questions de recherche auxquelles elle tente de répondre, pour combler les lacunes dans l'ensemble des recherches existantes.

## **CHAPITRE 3 : MÉTHODOLOGIE : L'ÉTUDE VARIATIONNISTE DE LA LANGUE DANS UN TERRAIN SOCIONUMÉRIQUE**

### **3.0 Introduction**

Dans ce troisième chapitre, je présente les précisions méthodologiques de ce travail. Tout d'abord, il s'agit d'exposer l'objet d'étude, notamment la communication numérique écrite, et ensuite de décrire le terrain, notamment celui de l'espace public du médium socionumérique Facebook. Ensuite, je montre comment la linguistique de corpus s'applique à l'analyse de la CNE, prenant comme terrain Facebook. Plus précisément, je décris les fonctionnalités du site Facebook, et les particularités des commentaires publics sur cette plateforme. Par la suite, je justifie le choix des données sur les pages publiques, et j'expose la situation de communication d'où elles proviennent. Enfin, je présente les démarches méthodologiques que j'ai suivies dans mon étude, notamment l'extraction, le stockage et l'analyse des données. Le chapitre se conclut en faisant un retour sur les questions de recherche et mes hypothèses.

### **3.1 La CNE comme objet d'étude**

Prendre la CNE comme objet d'étude, c'est étudier la langue dans un nouveau contexte, une forme de langue qui diffère d'autres formes langagières. En l'occurrence, il s'agit d'une forme écrite qui relève de l'hybridité, combinant lettres, symboles, icônes, chiffres, marques de ponctuation, et qui peut incarner des représentations de l'oral. Toutefois, c'est un code langagier créé par un humain qui transmet du sens au sein d'une interaction sociale. En effet, les lieux de production de la CNE sont le plus souvent des contextes sociaux, permettant de la conversation et du dialogue entre interlocuteurs et interlocutrices par écrit. La CNE est alors tout simplement le produit d'un type de comportement linguistique.

La communication numérique écrite représente une nouvelle et précieuse source de données pour les sociolinguistes. D'après Susan Herring (2004 : 338), la CNE se prête bien à une étude empirique, puisqu'en communiquant, les internautes laissent une trace textuelle. Cela rend leur

comportement linguistique plus facilement accessible que ne l'est la communication orale, qui est éphémère et qui doit être enregistrée pour ensuite être transcrite. De plus, au lieu de *créer* une situation d'entrevue et d'enregistrer les productions langagières pour ensuite les transcrire à l'écrit, les linguistes ont *directement* accès à ces productions langagières sous forme de CNE, issues d'interactions sociales permises par la technologie. En outre, cela permet aux linguistes d'employer des méthodes empiriques au niveau des productions linguistiques individuelles afin de renseigner sur des phénomènes de niveau communautaire. Pour la sociolinguistique variationniste, prendre la CNE comme objet d'étude fournit alors des données linguistiques authentiques déjà transcrites à l'écrit par l'internaute. Prendre la CNE comme objet d'étude sous-entend alors une adaptation des méthodes aux propriétés des médias socionumériques (Tannen et Trester, 2013 : 4). Ainsi, les méthodes d'étude de terrain et de cueillette de données dans un contexte numérique sont réorientées.

### **3.2 L'espace public de Facebook comme terrain d'étude**

Prendre l'espace public du réseau socionumérique Facebook comme terrain d'étude implique la prise de certaines décisions de la part des sociolinguistes. D'après Josiane Boutet, « [l]a sociolinguistique peut être considérée comme une forme de linguistique de terrain. [...] Les données sont recueillies dans des situations sociales réelles » (2002 : 569). Elle ajoute que les sociolinguistes doivent construire leur terrain en décidant de l'endroit, et des paramètres régissant l'observation et la collecte des données (Boutet, 2002 : 569). Tandis que le terrain de la sociolinguistique traditionnelle est le monde dans lequel les humains interagissent dans des situations sociales face à face, celui des médias socionumériques offre toutes sortes de nouveaux types d'interactions sociales. Par exemple, avec les dispositifs technologiques et socionumériques, les gens peuvent

accéder sur commande à leurs amis et à l'information, où qu'ils soient. Il existe une possibilité de présence continue et d'une conscience envahissante des autres dans le réseau. Les séparations physiques des personnes dans le temps et l'espace sont moins importantes (Rainie et Wellman, 2012 : 12, cité et traduit par Grossetti, 2014 : 195).

Ainsi, Facebook est une plateforme d'étude propice non seulement en raison de sa popularité et de la rareté des recherches la concernant, mais parce que ce site permet de clairement délimiter le terrain. En effet, les différents lieux et les nombreuses possibilités d'échange qu'elle offre aux

internautes fournissent des niches de recherche. Chacune de ces occasions de communication entraîne ses propres paramètres situationnels, sociaux, et technologiques.

De plus, Facebook est une plateforme avec un contenu textuel plus riche et engageant, contrairement à d'autres d'ordre plus visuel comme Instagram (Shahbaznezhad *et al.*, 2021). J'ai alors choisi une fonctionnalité et un lieu précis sur Facebook, ce qui m'a permis de cibler, de cerner, et de délimiter les données, sans intervenir auprès des internautes. Dans ce sens, prendre un lieu de communication précis sur Facebook comme terrain m'a permis d'établir les paramètres d'observation et de collecte de données, puisque la plateforme elle-même (et donc le terrain lui-même) est déjà délimitée. De plus, cet espace sur Facebook permet d'accéder à des communications faites en public, et d'après Jacques Maurais (2003 : 1), puisque « la politique linguistique du gouvernement du Québec a été redéfinie en fonction du concept de français, langue commune [...] il est nécessaire d'étudier [...] la langue lorsque celle-ci sert dans les communications publiques ». Dès lors, l'espace de Facebook que j'ai choisi comme terrain répond au besoin d'accéder aux communications en contexte public.

Ce lieu précis m'a également permis d'accéder à des données issues de *situations sociales réelles*, et non créées par moi-même. Dans l'espace public de Facebook, les internautes s'engagent dans des dialogues écrits, à deux ou à plusieurs. Ce sont alors en effet des situations sociales réelles, dans le sens que ce sont des interactions authentiques, et non des situations créées par une autorité externe. C'est ainsi un terrain qui permet aux linguistes d'observer sans participer, et dans la plupart des cas, sans même que leur présence n'interfère dans les interactions sociales. De plus, c'est un terrain qui par sa nature, comme d'autres espaces numériques, permet un accès à des données quantifiables, dans mon cas les données écrites sous forme de CNE. Pour la sociolinguistique variationniste, le fait de pouvoir quantifier les données est primordial, puisque c'est par l'analyse quantitative et systématique des données que l'on tire des conclusions sur le comportement linguistique dans la communauté (Labov, 1976). Ainsi, la CNE produite dans l'espace public de Facebook est une source de données authentiques et ouvertes à l'étude sociolinguistique.

En effet, la méthode en sociolinguistique variationniste traditionnelle est d'accéder à la langue la moins surveillée par le sujet parlant à travers l'entrevue semi-dirigée. D'après Labov, les sociolinguistes cherchent le style de langue « où l'on accorde le minimum d'attention à la surveillance de son propre discours. C'est en l'observant que l'on obtient les données les plus systématiques, nécessaires à l'analyse de la structure linguistique » (Labov, 1976 : 288). En revanche, avec cette méthode, comme nous l'avons vu au Chapitre 1, il fallait considérer le *paradoxe de l'observateur*, puisque la présence du chercheur ou de la chercheuse dans l'entrevue crée une situation dans laquelle le sujet pourrait se sentir surveillé. Bref, l'accessibilité des données publiques sur Facebook atténue ce paradoxe.

Internet, et Facebook en particulier, offrent un vaste répertoire de production langagière informelle, ou encore un vernaculaire déjà transcrit, qui permet d'observer les échanges en ligne tout en évitant d'y participer. Il est alors possible de neutraliser ce paradoxe de l'observateur, dans le cadre d'une étude sur le Web 2.0. D'après Jannis Androutsopoulos, il est même possible de collectionner ses données sans aucun contact avec les internautes (Androutsopoulos, 2014 : 75). En particulier, certains types de sites Web, surtout ceux qui sont accessibles publiquement, sont encore plus fructueux, car « l'anonymat offert par ce type de communication » permet « d'observer clandestinement » (van Compernelle et Williams, 2007 : 58) ou « en cachette » (Maccoccia, 2016 : 38) le discours des internautes. Dans le cadre d'une recherche sociolinguistique variationniste, le fait de prendre Facebook comme terrain diminue ainsi une limitation inhérente à la méthode traditionnelle.

### **3.3 La linguistique de corpus et la CNE**

La sociolinguistique variationniste repose sur une linguistique de corpus, et donc sur l'accès à un corpus, ou sinon, à la possibilité de constituer un corpus. Tandis qu'Internet et les médias socationumériques ne sont pas en soi des corpus (Page *et al.*, 2014 : 156), les linguistes conviennent que le Web est une source ou une plateforme permettant la constitution ou la construction de corpus (Hundt *et al.*, 2007 : 3 ; Friginal et Hardy, 2014 : 195). Le terrain devient ainsi la source directe du corpus éventuel, sans intermédiaire. Comme pour la délimitation du terrain, les sociolinguistes

doivent décider des détails de la constitution du corpus. En effet, d'après Céline Poudat et ses collègues,

[c]onstituer un corpus de CMR nous semble être une opération particulièrement ardue, peut-être d'abord du fait de la particularité de la **situation communicative**, dont nous n'avons jamais eu autant de traces matérielles [...] la construction du corpus dépend [alors] de la visée de la recherche (2020 : 1).

Cette prise de décisions est importante, surtout dans le cadre d'une étude des communications sur les médias socionumériques, en raison des métadonnées et des « spécificités du dispositif de communication » (Poudat *et al.*, 2020 : 1). Il est possible, par exemple, de choisir de construire son corpus en direct, au fur et à mesure que les communications sont produites. L'étude des souhaits d'anniversaire sur Facebook par Laetitia Bibié-Emerit (2015), par exemple, a employé une telle démarche. Pour ce faire, Bibié-Emerit a créé un compte Facebook dédié au corpus, puisqu'étant donné la nature ethnographique de son étude, elle voulait un « corpus évolutif » (19). Un autre moyen de constituer un corpus CNE est de demander à des volontaires de faire parvenir leurs communications. Il s'agissait notamment de la démarche entreprise par le projet *sms4science*, qui a demandé à des locuteurs de différentes régions de la francophonie de faire parvenir leurs messages textes. Ce corpus a par la suite été utilisé par une variété d'études sociolinguistiques.

Il est aussi possible de construire un corpus à partir de communications préexistantes sur un site Web. Plusieurs linguistes (Halté, 2020 ; Magué *et al.*, 2020) et sociolinguistes (Friginal *et al.*, 2017 ; McCulloch et Lamontagne, 2020) ont opté pour cette méthode, qui comprend l'extraction de données CNE produites antérieurement par des internautes avant le début de la collecte des données. La présente étude adopte cette dernière approche, prenant de la CNE déjà produite sur Facebook, sur des pages précises, et délimité dans le temps.

### **3.4 La constitution d'un corpus à partir de la plateforme Facebook**

À ma connaissance, nous l'avons vu précédemment, les sociolinguistes variationnistes n'ont pas encore examiné la production langagière en FQ sur Facebook. Il n'existe alors aucun corpus FQ extrait de ce terrain : la première étape dans mon processus de recherche était donc de constituer une collection de production écrite. Le corpus, « un échantillon de la langue » constitué « à partir d'un ensemble d'énoncés que l'on soumet à l'analyse » doit être « réellement représentatif »

(Dubois *et al.*, 2002 : 123-4). J'ai donc choisi de constituer un corpus à partir de communications préexistantes sur des espaces publics de la plateforme Facebook. D'après Michel Marcoccia, ce genre de corpus serait un corpus « naturel » (Marcoccia, 2016 : 39), puisqu'il n'y a aucune interférence de la part du ou de la linguiste. Il s'agit notamment de commentaires écrits par des internautes en réaction à des publications sur cinq pages publiques Facebook gérées par des médias de presse.

Comme ce sont des communications librement accessibles sans compte FB, les linguistes ne doivent ni inviter les internautes à partager leurs communications ni les avertir que leurs communications seront observées. De plus, d'après William Labov, du moins dans les entrevues semi-dirigées, les sujets parlants sont plus aptes à employer la variété de langue moins surveillée lorsqu'ils sont interrogés sur des sujets provoquant des émotions (1976 : 290). Or, dans le cas des publications sur les pages Facebook d'organismes journalistiques, les articles commentés traitent souvent de sujets régionaux ou politiques, et pourraient ainsi susciter une réaction forte de la part des internautes. On se rapprocherait alors de la variété de langue la moins soutenue.

En outre, prendre l'espace public et anonyme de FB éclaire sur la façon dont les internautes utilisent réellement et publiquement la langue. Cette démarche permet d'examiner l'emploi de la langue dans l'espace public, et évite que le contenu linguistique soit impacté par la conscience prescriptive du sujet parlant. En revanche, des volontaires qui partageraient leurs messages privés avec les linguistes pourraient vouloir faire attention à leur emploi de la langue, sachant que leurs messages font partie d'une étude. C'est le cas du corpus *sms4science* utilisé par Blondeau *et al.* (2014), qui est constitué de messages SMS envoyés volontairement par les participants. De plus, contrairement aux communications publiques sur Facebook, les SMS sont le plus souvent envoyés en privés entre des personnes qui se connaissent. L'autre avantage de choisir les données de l'espace public de Facebook est qu'il est possible d'inclure seulement la CNE des internautes qui se présentent comme étant dans la région géographique de la province du Québec, en consultant les renseignements qu'ils ont choisi d'afficher ouvertement sur leur profil public.

### **3.4.1 Les fonctionnalités de la plateforme Facebook**

La plateforme Facebook contient une gamme de possibilités de communication. Établi au départ pour la population étudiante de Harvard en 2004 et ouvert au public en 2005 (Bainbridge, 2020 : 27), ce site Web permet le réseautage social. Il est aussi aujourd'hui disponible sous application pour téléphone et tablette mobiles. Il offre à l'internaute plusieurs situations communicatives, la plupart par écrit. Facebook permet également des options de communication audiovisuelle. Afin de pouvoir communiquer sur Facebook, l'internaute doit créer un compte et doit s'y connecter pour pouvoir être actif sur le site. La création d'un compte est gratuite, toutefois, l'internaute doit accepter ses conditions de service<sup>1</sup>. De plus, certains espaces du site Facebook sont publiquement accessibles de manière passive, c'est-à-dire en mode lecture seulement, sans devoir se connecter à un compte. C'est grâce à cet accès public et passif que j'ai extrait mes données.

#### **3.4.1.1 Les espaces de la CNE des internautes sur Facebook**

Facebook offre plusieurs options plus ou moins privées à l'internaute pour produire de la CNE : une publication sur son propre « mur », qui sera visible à l'auditoire de son choix ; une conversation (t)chat en synchrone ou en asynchrone par message direct, soit à deux ou en groupe ; une publication dans un groupe, dans lequel les membres échangent sur un thème particulier ; un commentaire de publication sur une page publique, entre autres. Dans les trois premiers cas, la situation peut être plus ou moins publique ou privée, selon les préférences de l'internaute et/ou selon le (t)chat ou le groupe dans lequel il ou elle communique (Milette, 2012 : 19).

J'ai choisi d'extraire des données de la troisième catégorie, celle des publications commentées. Dans les trois premiers cas, il serait impossible de déterminer avec certitude s'il s'agit d'une situation entièrement privée ou entièrement publique. De plus, on risquerait de pénétrer dans la vie privée des internautes, entraînant des questions d'ordre éthique. Prendre les commentaires par des

---

<sup>1</sup> Les conditions de service de Facebook sont disponibles à URL [https://www.facebook.com/legal/terms?paipv=0&eav=AfYU\\_ZIPKINYN8wjB1tTDiU4mwYLuRSW\\_4-B9tUVpPdsXTyqr49zflt36cM\\_F\\_bfr0M&\\_rdr](https://www.facebook.com/legal/terms?paipv=0&eav=AfYU_ZIPKINYN8wjB1tTDiU4mwYLuRSW_4-B9tUVpPdsXTyqr49zflt36cM_F_bfr0M&_rdr)

internauts sur les pages publiques permet ainsi d'accéder à des données authentiques tout en respectant les principes de l'éthique de la recherche en sciences sociales.

#### **3.4.1.1.1 Les pages publiques**

Facebook permet autant à des individus qu'à des organismes de créer un compte. L'individu aura alors un profil, avec lequel il créera son réseau, et un « mur » sur lequel il écrira et échangera avec les membres de son réseau. Il pourra aussi accéder aux autres espaces plus ou moins publics de Facebook (van Dijck, 2013 : 47). Notons aussi que les utilisateurs et utilisatrices qui apportent des commentaires sur des publications par pages ne doivent pas être réseautés avec les autres utilisateurs afin de voir les commentaires des autres ni de leur répondre.

La présente étude s'intéresse notamment aux commentaires écrits par des internautes sur des pages gérées par des organismes journalistiques. En particulier, ces types de pages, le plus souvent publiquement accessibles, et sur lesquelles sont partagées des mises à jour et des renseignements, offrent la possibilité aux internautes ayant des comptes personnels d'interagir dans les commentaires de ces publications. Les pages publiques contiennent des renseignements au sujet de la page, ainsi que des photos, d'autres ressources et des publications que les internautes peuvent commenter. Ces publications peuvent contenir uniquement du texte, mais peuvent aussi être accompagnées d'une photo ou d'une vidéo, et contiennent le plus souvent un hyperlien, sur lequel les internautes peuvent cliquer pour accéder à l'article au complet. Les commentaires publics sur Facebook se situent directement sous la publication ; ce sont alors ces commentaires sur la page Facebook qui font l'objet de mon étude, et non ceux qui se trouvent parfois directement sur le site Web de l'organisme journalistique.

En ce qui concerne les commentaires sur les pages publiques, les fonctionnalités et le design de l'application de téléphone et de tablette sont presque identiques aux descriptions qui suivent dans les prochaines sections. Les internautes ont alors en général les mêmes options de fonctionnalité indépendamment de l'appareil qu'ils utilisent. En revanche, comme nous l'avons vu au Chapitre 1, l'appareil pourrait avoir un effet sur le contenu de la CNE produit par le scripteur.

Voici un exemple de publication sur une page publique gérée par la chaîne de télévision TVA :

**Figure 3.4.1.1.1 - Publication sur une page publique**



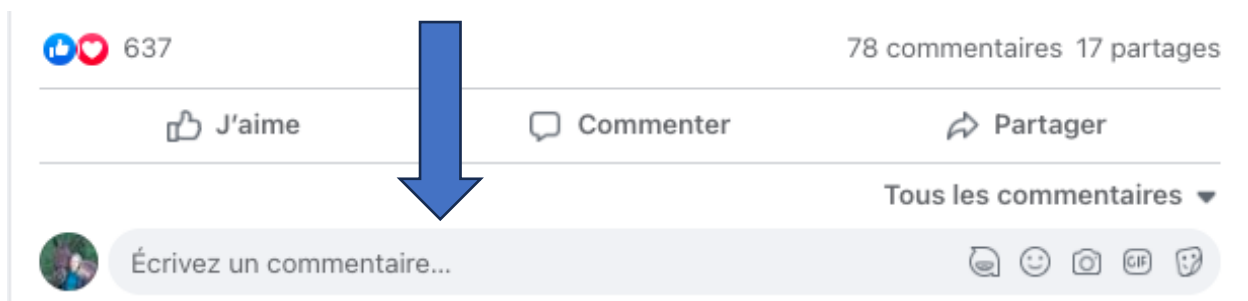
#### **3.4.1.1.1 Les commentaires des publications sur les pages publiques**

Les publications par des pages publiques permettent aux internautes d’interagir avec des organismes ainsi qu’avec autrui. En plus d’« aimer » ou de « partager », les internautes peuvent également commenter les publications partagées par ces pages. Cette fonction commentaire, dans ce cas des pages gérées par des organismes de presse, permet aux internautes – des gens ordinaires – d’échanger avec l’organisme et avec le public de l’organisme, et de s’engager dans

des discussions et des interactions. En effet, la plupart des organismes de presse gèrent des pages Facebook afin de partager des nouvelles sous forme d'articles, et d'inviter les utilisateurs à les commenter, dans le but de guider les internautes vers leurs sites Web et d'augmenter le nombre de lecteurs de leurs publications (Thiele et Turnsek, 2022 : 143). Ces entreprises publient du contenu qui peut être aimé, partagé ou commenté par l'auditoire (Hille et Bakker, 2014) ; ainsi, les internautes qui visitent ces pages peuvent ajouter des commentaires sur chaque publication. Ces commentaires, qui sont liés à un compte personnel, peuvent être lus par tous les internautes de FB (Hille et Bakker, 2014), mais plus largement aussi, par n'importe quel.e internaute avec une connexion Internet. Cet espace fonctionne donc comme une sorte de square public.

Pour ce qui est des fonctionnalités de l'entrée textuelle, la flèche bleue dans la Figure 3.4.1.1.1.1 indique la boîte de saisie dans laquelle l'internaute peut écrire, sous la publication.

**Figure 3.4.1.1.1 - La boîte de saisie de texte pour les commentaires publics**

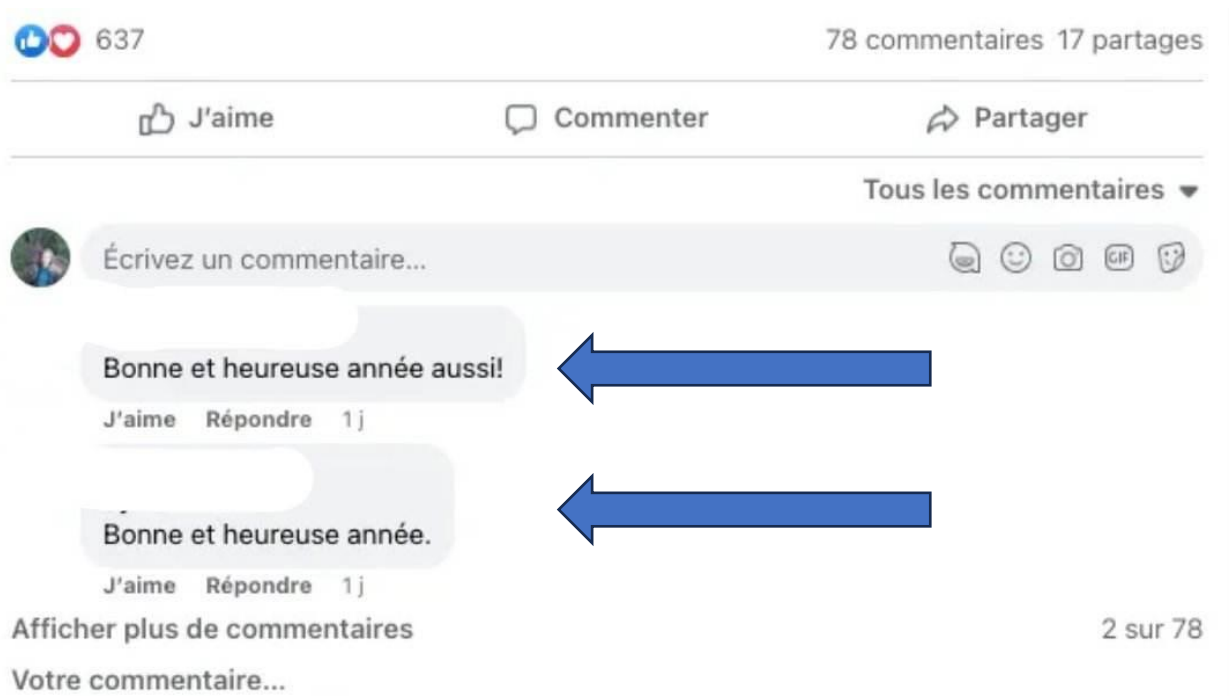


Une fois écrit, le commentaire peut être vu par tout internaute, que cette personne se connecte à un compte Facebook ou pas. Par exemple, n'importe qui peut accéder à ces pages publiques et aux commentaires qu'elles contiennent par le biais d'un moteur de recherche tel que Google. Ce commentaire est alors disponible à un auditoire inconnu, anonyme et illimité. C'est grâce à cet accès ouvert que j'ai extrait mes données. En revanche, depuis 2023 les publications par des organismes journalistiques sur Facebook ne sont plus visibles pour les personnes résidant au Canada (Meta, 2023), en raison du projet de loi C-18 (Gouvernement du Canada, 2023). Les implications de ce changement législatif sont discutées dans le cadre du Chapitre 6 et de la conclusion.

Les internautes qui commentent les publications sur ces pages publiques peuvent ensuite choisir de répondre aux commentaires des autres, ou de créer leur propre commentaire. Les commentaires des autres se retrouvent en dessous de la boîte de saisie de texte. Parfois ces échanges se font presque de manière synchrone, s'il s'avère que les deux internautes sont connectés en même temps. Dans le cadre de mon étude, j'ai extrait non seulement les commentaires, mais également les réponses aux commentaires. En revanche, je ne tiens pas compte des minuties de ces interactions, ni des liens entre les commentaires et les réponses. Cette possibilité d'échange quasi synchrone et la relation entre les réponses pourraient avoir un impact sur le contenu linguistique, et ainsi sur le niveau de formalité de la CNE. Par exemple, des échanges plus rapides pourraient produire de la CNE qui contient plus de traits de l'oral et qui témoigne de plus d'économie du langage par le moyen d'abréviations. En effet, d'après Rémi van Compernelle et Lawrence Williams, dans des situations plutôt synchrones, les internautes ont « tendance à utiliser des formes abrégées et des syllabogrammes afin de gagner du temps. La situation asynchrone des forums de discussion [...] aboutit à un système orthographique plus ou moins traditionnel. » En revanche, ces derniers ont constaté dans leur étude d'échanges dans un chat modéré que le public visé jouait un plus grand rôle que la synchronicité, sur le plan de la variation langagière. (2007 : 65-66). C'est donc une question qui mériterait plus d'attention à l'avenir.

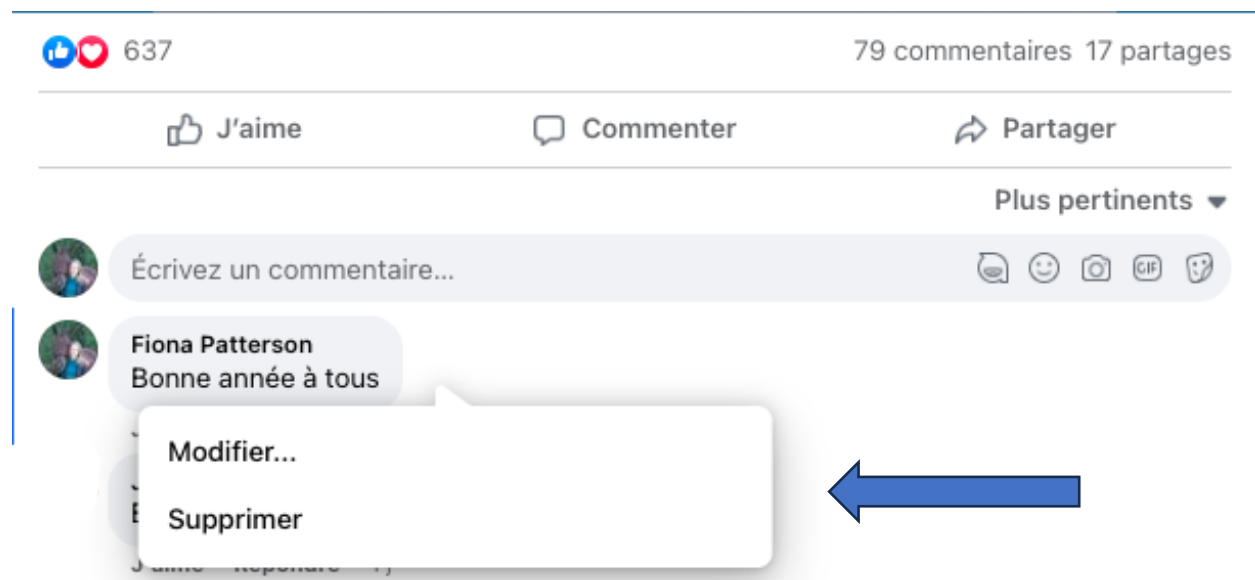
Les flèches bleues dans la Figure 3.4.1.1.2 illustrent à quoi ressemblent des commentaires publics par des internautes sur une publication :

**Figure 3.4.1.1.2 - Les commentaires publics sur une publication**



Il importe de rappeler que, contrairement à d'autres plateformes socionumériques telles que X/Twitter ou le SMS, il n'y a aucune limite de texte pour un commentaire sur une page publique Facebook. Cela contribue à cette richesse textuelle de Facebook, puisque les internautes ne se sentent pas contraints dans la rédaction de leur commentaire. L'internaute avec un compte Facebook peut ensuite modifier ou supprimer son commentaire ; la Figure 3.4.1.1.3 montre ces fonctionnalités de modification et de suppression, indiquées par la flèche bleue :

**Figure 3.4.1.1.3 - La modification et la suppression des commentaires**



Dans mon étude, je ne tiens pas compte des modifications qui auraient été apportées à un commentaire, mais il importe de noter que les commentaires extraits auraient pu être écrits et ensuite modifiés avant que je ne les aie extraits. Des commentaires supprimés avant mon extraction ne peuvent par ailleurs pas figurer dans mon analyse.

Les commentaires sont visibles sous trois rubriques : la rubrique de défaut est « Plus pertinents », illustrée par la flèche bleue dans la Figure 3.4.1.1.4, qui affiche les commentaires que la plateforme Facebook a considérés comme étant les plus pertinents pour l'internaute. Il est également possible de changer de rubrique, afin de voir les commentaires les plus récents, ou encore, tous les commentaires. Lors de la collecte des données, j'ai profité de la fonctionnalité « Tous les commentaires », illustrée dans la Figure 3.4.1.1.5, afin de ne pas limiter la taille ni le contenu du corpus, et d'éviter que les paramètres de défaut de Facebook n'excluent des données potentielles.

Figure 3.4.1.1.4 - La visualisation des commentaires



Figure 3.4.1.1.5 - Les options de visualisation des commentaires



De plus, l'internaute peut parfois, dépendamment des paramètres établis par la page, inclure d'autres éléments non textuels, comme les émojis, une photo ou une image animée, par exemple une image sous format .gif. La flèche bleue dans la Figure 3.4.1.1.6 montre les fonctionnalités non textuelles disponibles aux internautes :

Figure 3.4.1.1.6 L'ajout d'éléments non textuels aux commentaires



Le commentaire peut même contenir uniquement des éléments non textuels, c'est-à-dire qu'il n'est pas obligatoire d'écrire du contenu textuel pour pouvoir ajouter un « commentaire ». Dans le cadre de cette étude, j'examine seulement les commentaires qui ont du contenu textuel en français. Si ce contenu non textuel ne figure pas dans l'analyse des données de la présente étude, il mérite cependant qu'on s'y attarde dans une recherche future.

### **3.5 Le choix des données**

Pour cette étude, j'ai ciblé les commentaires d'internautes sur les cinq pages Facebook suivantes d'organismes de presse québécois : *TVA Nouvelles*, *La Presse*, *Le Devoir*, *Le Journal de Montréal*, *Le Journal de Québec*. Comme nous l'avons vu, il s'agit d'une connexion réseautée et très publique avec un auditoire vaste et anonyme. J'ai également choisi ces pages afin de délimiter le plus possible la production langagière comme étant de résident.e.s québécois.es, puisque ce sont des pages d'intérêt régional. De plus, les internautes sont très actifs dans les commentaires sur ces cinq pages, ce qui permet de constituer un corpus de taille importante. Par exemple, j'ai décidé d'exclure de mon extraction les commentaires sur d'autres pages, comme celle de la revue québécoise *L'actualité*, en raison de leur très petit nombre. J'ai ciblé les commentaires sur les publications créées par ces cinq pages sur une durée de 24 heures, notamment le 1 juillet, 2022. Il importe de noter que cette date a été choisie au hasard, et que c'est un congé en raison de la fête nationale du Canada, mais également la journée québécoise du déménagement. Ainsi, ces facteurs auraient pu impacter les données, puisque d'une part, il se pourrait qu'étant en congé, plus de personnes que d'habitude sont en ligne, mais inversement, moins de gens pourraient se brancher en raison d'un déménagement.

Cette démarche permet de délimiter le contexte. D'une part, elle établit des paramètres temporels précis, tout en conservant des renseignements sur le contexte (Androutsopoulos, 2018 : 235). D'autre part, les publications commentés à travers ces cinq pages portent en grande partie sur les mêmes thèmes et événements, donc la nature du sujet discuté reste assez homogène. Ensuite, j'ai extrait les commentaires provenant exclusivement d'internautes qui indiquent publiquement sur

leur profil Facebook qu'ils habitent au Québec, c'est-à-dire que l'utilisateur ou l'utilisatrice indique soit sa ville de résidence, soit la région dans laquelle il ou elle se situe.

### 3.6 La situation de communication à l'étude

Dans le cadre de cette étude, la situation de communication en question est alors celle que reflètent les commentaires sur des publications des cinq pages susmentionnées sur le site Web Facebook. Le Tableau 3.6.1 ci-dessous décrit la situation précise régissant la production de chaque commentaire dans son unité :

**Tableau 3.6.1 Les caractéristiques sociales**

Caractéristiques sociales	Description	Commentaires publics sur Facebook
Scripteur	Le nombre ou le type de scripteur(s), ses (leurs) caractéristiques	Individu affichant publiquement sur son profil son lieu de résidence comme étant au Québec
Lecteur, auditoire, cadre participatif	L'auditoire (anonyme ou connu, limité ou illimité)	Anonyme, illimité
Lecteur réponse	Le type de réponse de la part du lecteur	Varié. Aucune réponse nécessaire, mais plusieurs réponses possibles par différents individus.
Répondeur	Le nombre ou le type de scripteur(s) qui répond(ent) au message	Illimité
But des messages, objectifs	Le but, les fins du message	Varié
Sujets, thèmes	Le(s) sujet(s), le(s) thème(s) du message	Selon la publication commentée
Ton des messages	Le ton du message, le niveau de formalité	Varié
Environnement social	La nature publique ou non de l'environnement social, et si l'environnement est géré par un modérateur ou non	Public (d'intérêt québécois), modéré
Normes sociales, linguistiques	Les normes sociales ou linguistiques, organisationnelles, encadrant la production du message	Normes sociales - à déterminer Normes linguistiques - à déterminer

**Tableau 3.6.2 Les caractéristiques technologiques**

<b>Caractéristiques technologiques</b>	<b>Description</b>	<b>Commentaires publics sur Facebook</b>
Instrument d'entrée	Clavier ordinateur, téléphone ou tablette, saisie vocale...	Varié
Support lecture	Écran d'ordinateur, de téléphone ou de tablette	Varié
Code	Code (langagier ou autre, variété de langue) utilisé (oral, écrit, iconique)	Production écrite en français (québécois), texte et symboles
Médium, canal	Nom du médium, de la plateforme, du site Web 2.0	Facebook
Mode de transmission	Les caractéristiques des messages (mots, caractères, limites) ; calibrage, longueur, granularité	Commentaires qui contiennent un nombre illimité de mots, phrases. Variabilité de la longueur.
Temporalité, synchronicité	Synchronicité ou asynchronicité de l'échange communicatif	Asynchrone, avec possibilité de quasi synchronicité
Persistance du message, adaptations	Message persistant ou qui disparaît, qui peut être filtré ou non, qui peut être cité ou non.	Message persistant, filtrable, searchable, qui peut être supprimé, modifié, et auquel d'autres peuvent répondre
Environnement technologique	Comment les messages sont vus et lus par les scripteurs et par l'auditoire, apparence physique, format de présentation	Les commentaires apparaissent dans l'ordre de publication, mais la visibilité peut être gérée par chaque individu de l'auditoire. Les réponses apparaissent en alinéa au commentaire d'origine.

Chaque commentaire à l'étude s'inscrit donc dans une situation sociale et technologique précise. Les commentaires se retrouvent uniquement sur la plateforme Facebook, mais sont accessibles et consultables par le biais d'un moteur de recherche également, sans passer par une connexion à un compte Facebook. Dans mon cas, je me suis servie du moteur de recherche Google pour accéder aux commentaires et aux profils publics. Il s'agit alors d'un environnement complètement public, mais qui est souvent modéré par l'organisme gérant la page en question. Alors que les commentaires peuvent être modérés ou supprimés, les normes sociales, linguistiques, etc., entourant cette situation de communication sont, pour l'instant, inconnues. L'instrument d'entrée et le support de lecture varient selon l'internaute ; contrairement aux SMS, la diversité des instruments d'entrée pourrait avoir un effet sur la CNE. Pour ce qui est de l'effet du support de lecture, il est difficile de voir comment cela pourrait influencer sur la CNE. Le commentaire peut n'être lu par personne ou être vu par un auditoire anonyme et illimité. Il peut rester sans réponse

ou recevoir un nombre illimité de réponses. Dans le cadre de cette recherche, les relations et les interactions entre commentaire et réponse, ou entre individus ne seront pas examinées, mais comme nous l'avons vu, il reste que ce sont des échanges qui se font en public, et non en privé comme c'est le cas des SMS.

Cette situation se distingue alors des (t)chats (étudiés par Blattner et Williams, 2011, 2022 ; van Compernelle et Williams 2007 ; van Compernelle, 2008a, 2008c, 2010; Williams, 2009) puisque les commentaires sont persistants, ou encore, permanents, et puisque l'internaute peut ajouter son commentaire à n'importe quel moment : ce n'est donc pas uniquement une situation synchrone, mais asynchrone aussi. La situation sur Facebook est alors moins rapide et moins structurée sur le plan temporel que celle des (t)chats. En revanche, le fait que ces commentaires se font en réaction à des articles journalistiques crée potentiellement une situation plus structurée sur le plan thématique que dans le (t)chat. La situation se distingue également des SMS (étudiés par Bally, 2022 ; Blondeau *et al.*, 2014 ; Blondeau et Tremblay, 2022b; Stark 2012, 2014 ; Stark et Riedel, 2012 ; Tremblay *et al.*, 2019) puisque ces derniers sont envoyés en privé et seulement par le biais d'un téléphone mobile, tandis qu'un commentaire sur Facebook est public et peut être écrit à partir d'une diversité d'appareils. De plus, souvent le SMS a une limite de longueur, contrairement aux commentaires à l'étude dans la présente recherche.

### **3.7 L'extraction, le stockage et l'anonymisation des données**

J'ai suivi un processus rigoureux et systématique pour extraire, stocker et anonymiser les données. J'ai extrait de manière manuelle les commentaires écrits, tout d'abord en consultant chacune des pages cibles, en naviguant vers les publications de la date en question, et ensuite en copiant et collant les commentaires pertinents. Alors qu'il existe des logiciels tels que Facepager (Jünger et Keyling, 2019), permettant l'extraction automatique de contenu textuel sur Facebook, l'extraction manuelle a permis d'exclure immédiatement les commentaires non pertinents : par exemple, les commentaires écrits par des internautes n'affichant pas publiquement leur lieu de résidence, ou les commentaires sous forme d'image uniquement.

J'ai utilisé une combinaison de logique et d'expérience personnelle avec les médias sociaux afin d'exclure la CNE provenant de comptes douteux. Plus précisément, les noms d'utilisateurs au nom suspect ou suggérant un pseudonyme ou un compte illégitime, qui ne correspondaient pas à la formule *Nom Prénom* ou *Prénom Nom*, ont été écartés des analyses. Si le nom ou le profil public contenait même un élément qui ne correspondait pas à mes critères, sa CNE était exclue. Par exemple, si le nom n'était pas clairement formulé d'un prénom et un nom de famille, ou si l'utilisateur indiquait qu'il habitait à Tombouctou, son commentaire était exclu de l'extraction. Dans le même ordre d'idées, des commentaires de genre « spam » ont été exclus. Par exemple, la CNE d'un internaute appelé *Bon Homme* ou *Ma Dame*, et un commentaire incohérent avec du contenu ou des « demandes indésirables, [...] de la publication excessive de liens ou d'images [ou des] invitations à des personnes » (Meta, 2024) inconnues l'étaient également. D'ailleurs, dans la grande majorité des cas, si le nom était douteux, le lieu de résidence n'était pas affiché, ou était un endroit autre que le Québec.

Au total, j'ai extrait la CNE de 1600 internautes au Québec, ce qui représente un corpus d'environ 46 000 mots. Le Tableau 3.7 indique les statistiques pour chaque composante du corpus :

**Tableau 3.7 - Les statistiques du corpus**

Composante du corpus (page Facebook)	Nombre approximatif de mots	Nombre approximatif de commentaires	Nombre d'internautes
<i>La Presse</i> <sup>2</sup>	23000	970	615
<i>TVA Nouvelles</i> <sup>3</sup>	10800	625	455
<i>Le Journal de Québec</i> <sup>4</sup>	3500	274	224
<i>Le Devoir</i> <sup>5</sup>	6800	279	192
<i>Le Journal de Montréal</i> <sup>6</sup>	2700	166	114
Total du corpus	46800	2314	1600

Les commentaires de la page de *La Presse* proviennent de 615 internautes, *TVA Nouvelles* a fourni 455 internautes, le *Journal de Québec* 224 internautes, *Le Devoir* 192 et le *Journal de Montréal* 114 internautes. Quelques personnes ont commenté des publications provenant de plus d'une page ; par exemple, 26 ayant commenté des publications par la page Facebook de *La Presse* ont aussi commenté des publications d'autres pages Facebook cibles. Un internaute en particulier a commenté des publications des pages Facebook de *La Presse*, du *Journal de Montréal* et du *Journal de Québec*.

### 3.7.1 L'éthique : la sécurité, l'anonymat et le pseudonymat

Les internautes qui détiennent un compte Facebook peuvent régler les paramètres de sécurité de leurs comptes pour rendre leurs communications plus ou moins privées ou publiques. Ils ou elles peuvent régler l'accès à leur compte : les renseignements personnels sur le profil public, la

<sup>2</sup> Disponible à URL <https://www.facebook.com/LaPresseFB/>

<sup>3</sup> Disponible à URL <https://www.facebook.com/TVAnouvelles/>

<sup>4</sup> Disponible à URL [https://www.facebook.com/JdeQuebec/?locale=fr\\_CA](https://www.facebook.com/JdeQuebec/?locale=fr_CA)

<sup>5</sup> Disponible à URL <https://www.facebook.com/ledevoir/>

<sup>6</sup> Disponible à URL [https://www.facebook.com/jdemontreal/?locale=fr\\_CA](https://www.facebook.com/jdemontreal/?locale=fr_CA)

visibilité du compte, la communication avec autrui par message direct. Si un internaute choisit de commenter dans les espaces publics de Facebook, comme les groupes et les pages publiques, ses écrits sont publiquement visibles, mais son compte reste privé. Afin de respecter la vie privée des internautes, dans le cadre de mon étude, les commentaires provenant de comptes privés sont exclus des données recueillies.

En termes de procédures d'éthique, le consensus chez les spécialistes en sciences du langage est d'utiliser son bon jugement, étant donné le contexte particulier à l'étude. C'est-à-dire qu'il faut tenir compte de la plateforme, du degré de confidentialité de l'objet d'étude, de la vulnérabilité des internautes et de la possibilité de préjudice que pourrait causer l'étude aux internautes (Boyd, 2010 ; Varis, 2016 : 60). Il est à noter que l'étude présente a passé l'approbation de la Faculté des études supérieures de l'Université York, et ne requiert pas d'examen par le comité de l'éthique puisque les données proviennent de sujets qui ont déjà partagé leur production langagière dans un espace public (i.e. ce sont des données préexistantes ; voir ESRC, 2015 ; Paolillo, 2016 : 50 ; Spilioti, 2017 : 2). Pour me guider, j'ai également suivi l'*Énoncé de politique des trois conseils : Éthique de la recherche avec des êtres humains* du Conseil de recherche en sciences humaines du Gouvernement du Canada (2018).

Alors que dans le cadre de la présente étude il n'y a aucune interaction directe entre le chercheur et l'internaute, puisque l'important est de ne pas mettre l'internaute à risque, j'ai choisi de prendre toutes les précautions possibles. J'ai anonymisé les données de chaque personne, afin d'éviter que l'individu soit identifiable (Marriccia, 2016 : 46 ; Androutsopoulos, 2018 : 242 ; Franzke *et al.*, 2020 : 10), en assignant à chaque internaute un numéro distinct et arbitraire. Les désignations anonymes vont de 1111 à 2711. Enfin, j'ai annoté chaque composante du corpus, pour distinguer la CNE par des internautes de la CNE produite par la page Facebook elle-même. Une fois que je les ai anonymisées et annotées, j'ai stocké les données dans des fichiers de format .txt ; un fichier par page Facebook, donc cinq fichiers .txt au total. Le format .txt est nécessaire pour l'analyse des données par le biais du logiciel #LancsBox. Garder une version du corpus dans Excel permet également de maintenir une base de données consultable et filtrable.

### 3.8 Le contexte variable et la recherche des variantes

Avant de pouvoir commencer l'analyse des données, il est nécessaire de définir le contexte variable. Rappelons qu'une variable linguistique est un regroupement d'énoncés équivalents, deux ou plusieurs manières de dire la même chose, et les variantes sont chacune des alternances possibles d'une variable. Tout d'abord, les sociolinguistes identifient une variable linguistique, souvent d'après les recherches antérieures, et ensuite, pour la définir, « il faut a) établir l'éventail de tous les contextes linguistiques où elle apparaît ; b) définir autant de variantes [...] qu'on peut raisonnablement en distinguer ; c) élaborer un indice quantitatif qui permette d'en mesurer les valeurs » (Labov, 1976 : 129). Il faut également veiller à exclure les contextes dans lesquels une variante est catégorique (Tagliamonte, 2006 : 13), ou figée, comme dans les expressions idiomatiques. Les sociolinguistes doivent aussi s'assurer que la source des données soit fiable elle aussi – elle doit représenter des données authentiques dans une diversité de contextes (Tagliamonte, 2006 : 72). Ainsi, les variables et variantes linguistiques ne sont pas choisies au hasard par les sociolinguistes, qui s'orientent à partir « des descriptions et remarques déjà faites par les grammairien[.ne.]s contemporain[.ne.]s et antérieur[.e.]s » (Marcellesi et Gardin, 1974 : 119). Cette démarche, celle de délimiter le contexte variable, ou encore de définir l'enveloppe de la variation, est nécessaire afin de répondre au *principe de la comptabilité*, qui requiert que chaque variante d'une variable donnée soit prise en compte. Ainsi, il ne suffit pas d'étudier les formes les plus intéressantes ou les plus novatrices, mais de tenir compte de toutes les formes à valeur référentielle équivalentes qui varient dans un contexte donné (Tagliamonte, 2006 : 13).

Dans le cadre de mon étude, j'ai cherché les instances de chaque variante, et j'ai écarté les formes qui ne correspondent pas au sens ou au contexte variable. Plus précisément, j'ai exclu les formes ayant un sens sémantique différent du contexte variable, et les formes figées et catégoriques. Pour illustrer ces différences d'ordre sémantique, dans le cas des marqueurs de restriction, la variante *juste* peut se confondre avec l'autre sens de ce mot. Par exemple, *C'est juste le début* versus *Ce n'est pas juste* : dans le premier énoncé, *juste* a un sens restrictif ou de contrainte, et pourrait être remplacé par *seulement*, tandis que dans le deuxième, *juste* a le sens de *équitable*, et donc ne correspond pas aux sens des marqueurs de restriction. En ce qui concerne les formes catégoriques, ce sont celles qui n'ont jamais de variation, par exemple dans le cas de la négation, l'expression

*n'est-ce pas* prend toujours la particule *ne*. Ainsi, puisqu'il n'y a jamais de variation pour ces formes, elles sont exclues de l'analyse. Je me suis servie de l'éventail des recherches variationnistes antérieures portant sur le français québécois pour me guider. En particulier, j'ai utilisé les méthodes de recherche, d'organisation et d'analyse des données employées dans le cadre de mon travail d'assistantat de recherche (de 2008 à 2010) sous la direction du sociolinguiste Raymond Mougeon (Mougeon, 2013).

En outre, j'ai cherché et inclus les graphies atypiques, comme était la démarche dans l'étude de Mougeon (2013) sur un corpus écrit de pièces de théâtre françaises. J'ai donc tenté de trouver toutes les occurrences possibles en cherchant une diversité d'orthographe, en consultant les recherches antérieures et des dictionnaires du français québécois (Bergeron, 1980 et 1982). Par exemple, j'ai trouvé plusieurs instanciations possibles de la variable *fait que* (p.ex. *faitque, faque, faq, fack*, etc.). De plus, j'ai conservé et codé les occurrences d'une variable qui étaient novatrices, ou qui ne s'alignaient ni sur des occurrences connues ni sur des formes figées/catégoriques. Par exemple, dans le cas de la négation, un nombre important d'occurrences figuraient dans un fragment de phrase (p.ex. « *NON, pas assez de protéines* »), ou en début de phrase non verbale (p.ex. « *pas capable je sais pas comment François fait ...* »).

### 3.8.1 Le logiciel #LancsBox et la concordance

Les analyses de concordance et de collocation ont été effectuées grâce au logiciel #LancsBox<sup>7</sup>, développé par Vaclav Březina et son équipe (2021) à Lancaster University, qui est un outil d'analyse de données linguistiques et de corpus à libre accès. J'ai choisi d'adopter ce logiciel car il permet plusieurs fonctionnalités, notamment la constitution d'un corpus et/ou le téléchargement d'un corpus existant, la visualisation graphique des données linguistiques, plusieurs moyens d'analyse de données linguistiques dans n'importe quelle langue, et l'annotation des parties du discours. En particulier, #Lancsbox est un logiciel précieux dans le cadre de la présente étude puisqu'il offre la possibilité de faire des analyses de concordance et de collocation. Une fois

---

<sup>7</sup> Le logiciel, ainsi que le guide utilisateur, sont disponibles en ligne à télécharger gratuitement, à URL <http://corpora.lancs.ac.uk/lancsbox/download.php>.

téléchargé, ce logiciel gratuit offre une gamme de fonctions accessibles à partir de l'écran principal, illustrées dans la Figure 3.8.1 :

**Figure 3.8.1 L'écran principal du logiciel #LancsBox**

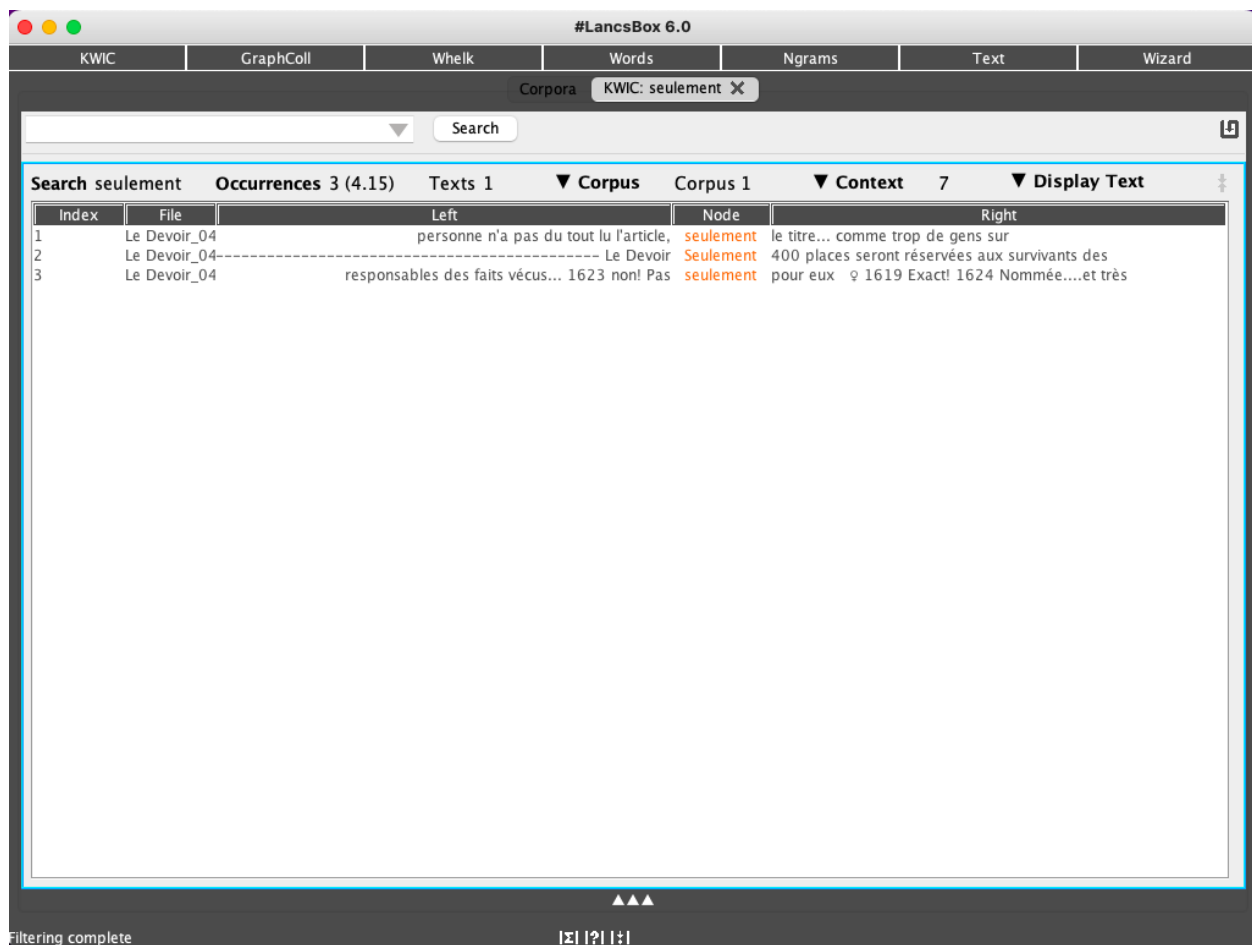


Cet écran principal permet la création d'un nouveau corpus ou le (télé)chargement d'un corpus préexistant. Dans mon cas, j'ai utilisé la fonction « Corpus » sous la rubrique « Load data » dans le menu à gauche, puisque j'avais déjà extrait mes propres données manuellement. Il importe de noter que #Lancsbox ne permet pas d'extraire des données directement de Facebook. Une fois le corpus téléchargé dans #LancsBox, les fonctions du menu horizontal, KWIC, GraphColl, Whelk, Words, Ngrams, Text, Wizard, permettent de l'analyser d'une variété de manières.

Une concordance est une analyse des données qui crée une liste de toutes les instances d'un mot (ou d'une expression) dans son contexte discursif (Tagliamonte, 2006 : 67). J'ai alors cherché les

variantes à l'étude dans le corpus grâce à la fonction « KWIC », qui permet d'effectuer des recherches de concordance, c'est-à-dire de retrouver et de dégager les formes récurrentes. Ce logiciel permet d'effectuer des recherches très précises, comme on le ferait avec un moteur de recherche. La Figure 3.8.1.1 ci-dessous montre l'exemple des occurrences de la forme *seulement* cherchées dans le fichier .txt des données extraites de la page Facebook du journal *Le Devoir*.

**Figure 3.8.1.1 L'écran KWIC du logiciel #LancsBox**



Le logiciel trouve chaque occurrence, souvent appelé en sociolinguistique un token, la conservant dans son environnement linguistique. Il est également possible de préciser la taille du contexte linguistique à garder, notamment entre un et plus de 30 mots à gauche et à droite du token cherché. Dans la capture d'écran ci-dessus, il est possible de constater que j'ai gardé sept mots à gauche et à droite de l'occurrence. En outre, on voit que la deuxième occurrence fait partie de la CNE publiée par la page *Le Devoir*, comme l'indique les symboles venant directement à gauche de la variante,

qui représentent les annotations que j’ai portées au corpus. Dans ce cas, l’occurrence est exclue des données qui seront analysées puisqu’il ne s’agit pas de CNE provenant d’un internaute, mais plutôt de l’organisme gérant la page publique. Une fois les occurrences trouvées dans #LancsBox, elles sont copiées, collées et codées dans un fichier Microsoft Excel. Ce système de codage est nécessaire pour l’analyse dans GoldVarb. Le Tableau 3.8.1.1 montre un échantillon de quelques tokens stockés et codés dans Microsoft Excel, selon la variante, son environnement linguistique, et la page d’où elle provient :

**Tableau 3.8.1.1 Échantillon d’occurrences stockées et codées**

Texte à gauche	Occurrence	Texte à droite	Variante	Environnement linguistique	Page
<i>bien raison. Bitcoins, jamais de riche moi.</i>	<b>Juste</b>	<i>du monde qui ont perdu de l</i>	J	N	P
<i>...souligner cette situation de représentation mais s’en</i>	<b>seulement</b>	<i>à ça est manquer le coeur de...</i>	S	R	P
<i>de cochonneries jetables lors de leur séjour,</i>	<b>ne</b>	<i>serait ce que la fameuse bouteille d’eau</i>	N	C	P
<i>y de quoi a etre inquiet!!! 1490</i>	<b>Rien que</b>	<i>le fait qu’ils aient attendu d’avoir quitté</i>	R	N	D

Les codes attribués dans les colonnes *Variante*, *Environnement linguistique* et *Page* sont ensuite utilisés par le logiciel GoldVarb dans l’analyse statistique. Dans la colonne *Variante*, le token reçoit un code *J* pour *juste*, *S* pour *seulement*, *N* pour *ne que* et *R* pour *rien que*. On constate dans la colonne *Environnement linguistique* l’élément qui suit le token : *N* pour désigner un groupe nominal, *R* pour désigner une préposition, et *C* pour désigner un verbe conjugué au conditionnel. Ensuite, dans la colonne *Page*, le token est codé pour la page d’où elle provient : *P* pour *La Presse*, *D* pour *le Devoir*, *Q* pour *Journal de Québec*, *M* pour *Journal de Montréal*, et *T* pour *TVA Nouvelles*. Ces renseignements sont utiles lors de la discussion des résultats de chaque variable, car l’environnement linguistique pourrait jouer un rôle dans le choix de la variante.

### 3.8.2 Le logiciel #LancsBox et la collocation

La collocation est la coprésence systématique de mots ou d'expressions dans les textes et dans le discours (Xiao, 2015 : 106). La collocation peut être décrite comme « l'interaction entre mots » (Williams, 2006 : 154) ou l'association d'un élément linguistique avec d'autres au sein d'un énoncé (Dubois *et al.*, 2002 : 91). Par exemple, lorsque la particule *ne* de la négation est employée pour mettre de l'emphase, elle est souvent accompagnée d'un adverbe renforçant l'emphase (Mougeon et Rehner, 2013 : 4). Ces genres de coprésences peuvent être analysées et ensuite visualisées graphiquement (Březina, 2016 : 90). Normalement, les analyses de collocation sont faites afin de dégager les cooccurrences les plus fréquentes, et d'établir ensuite des réseaux de collocation. Dans le cadre de mon étude, l'analyse de collocations sert à déterminer à quel point les internautes privilégient l'écrit standard aux formes plutôt informelles. J'adopte alors une méthode simplifiée de l'analyse des collocations (Di Vito, 2015 ; Xiao, 2015 : 119), puisqu'il s'agit d'une recherche restreinte et exploratoire, comparée à la totalité du corpus.

J'ai dégagé les collocations en analysant seulement les commentaires contenant les variables suivantes : les marqueurs lexicaux de contrainte et de conséquence, les formules interrogatives, l'expression du futur, et ceux contenant aux moins deux instances de la négation. J'ai effectué cette recherche en utilisant la fonction « Text » de #LancsBox, qui surligne dans leur intégrité les commentaires qui contiennent la variante recherchée, permettant de plus facilement vérifier la présence ou non d'une deuxième variante ou d'un deuxième élément non standard dans le même commentaire. La Figure 3.8.2 montre l'écran de la fonction « Text » du logiciel #LancsBox. Les commentaires contenant le token recherché sont surlignés en gris foncé, et le terme recherché, dans ce cas *pas*, est surligné en orange. De cette manière, la fonction « Text » permet d'isoler et d'extraire seulement les commentaires pertinents.

**Figure 3.8.2 L'écran Text du Logiciel #LancsBox**



Je me suis limitée à l'examen des cooccurrences des cinq variables susmentionnées afin d'en simplifier l'analyse, et de rendre celle-ci la plus objective possible. J'ai copié et collé dans un fichier Excel chaque commentaire contenant l'une des variantes en question. Ensuite, j'ai assigné à chaque variante et à chaque élément dans le commentaire un code, selon son appartenance à la langue non standard (N), ou à la langue standard (S). Le commentaire recevait un code N lorsqu'une variante non standard, c'est-à-dire la variante la plus fréquente à l'oral informel, était présente dans le commentaire et/ou lorsqu'il y avait la présence d'un élément divergeant du standard écrit, qu'il s'agisse d'un émoji, de lexique, d'une orthographe, d'un accord, d'une ponctuation ou accentuation non standards, de l'alternance codique ou d'un emprunt de l'anglais, d'interjections ou de québécoisismes provenant de l'oral.

Si le commentaire contenait uniquement des variantes ou des éléments S, il recevait un code final de S. Si un commentaire contenait uniquement des variantes N, il recevait un code final de N. Si le commentaire contenait un mélange de variantes (N et S), il recevait un code final de M. Par exemple, le Tableau 3.8.2 montre trois commentaires codés selon la présence de variantes standards et non standards de la négation. Le premier commentaire dans la série contient uniquement des formes standards (S), c'est-à-dire que toutes les occurrences de la négation dans ce commentaire contiennent la particule *ne* ; ce commentaire reçoit donc un code final de S. Le

deuxième commentaire contient une forme standard de la négation, ainsi qu'une forme non standard, et reçoit donc un code final de M, et le troisième, qui contient uniquement des formes non standards, la particule *ne* étant omise dans les deux cas, reçoit un code final de N.

**Tableau 3.8.2 Échantillon de commentaires stockés et codés pour la collocation**

Commentaire	Page	1	2	3	4	5	Code final
<i>Si vous <b>n'en comprenez pas</b> l'importance, ça <b>ne veut pas</b> dire que ce <b>n'est pas</b> pertinent; l'ouverture d'esprit sur le monde, ça vous dit qqchose ? sinon, passez donc par-dessus ce qui <b>ne vous intéresse pas</b> en vous gardant une p'tite gêne. 😊</i>	D	S	S	S	S	S	S
<i>vous oublier le pâté croche un mets bien de chez nous qui est en péril. La soupe borsh est même fait ici mais eux <b>ne font pas</b> de pâté croche... <b>c'est pas</b> juste faudrait que Justin et Mélanie leurs explique que ça va dans les deux directions 😊</i>	D	S	N				M
<i>Marc Dionne la première partie de ta réponse <b>je suis pas</b> d'accord. Mais pour la loi c11 c'est vrai mais <b>les gens veulent pas</b> savoir</i>	D	N	N				N

Ainsi, il a été possible de calculer la fréquence des commentaires homogènes (N ou S) et hétérogènes (M) dans leur contenu linguistique. Cette analyse permet de se focaliser plus précisément sur la variation sociolinguistique dans la CNE d'individus.

### 3.8.3 L'analyse descriptive statistique des données et le logiciel GoldVarb

Pour ce qui est de l'analyse statistique, il faut naturellement tenir compte non seulement des occurrences, mais également des non-occurrences d'une variable. La fréquence d'emploi relative de chaque variante peut ensuite être calculée à partir du total des occurrences de toutes les variantes de cette variable. Il s'agit alors d'une analyse proportionnelle (Tagliamonte, 2006 : 72), d'un calcul des fréquences relatives d'emploi des variantes à l'intérieur d'une enveloppe de la variation.

Si une variation donnée n'est pas du tout conditionnée par des facteurs internes ou externes, les fréquences relatives d'emploi, donc le taux de la variation, ne changeront pas selon le contexte, et

il s'agira d'une variation complètement *libre*. Ceci représente l'*hypothèse nulle* – la variation est complètement le fruit du hasard et est alors non conditionnée par des facteurs internes ou externes. Ces facteurs externes peuvent être, par exemple, les caractéristiques du locuteur ou de la locutrice comme l'âge, le sexe et le statut socioéconomique, ou la région comme le dialecte. Les facteurs internes peuvent être par exemple l'environnement linguistique, c'est-à-dire les autres éléments linguistiques entourant la variante à l'étude, qui peut conditionner ou non une certaine variante. Dans le cadre de cette étude, je me focalise sur les facteurs externes de la diamésie, donc de la plateforme, et de la situation de communication, tout en tenant compte du facteur interne de l'environnement linguistique.

Ainsi, si la variation retrouvée dans la CNE sur Facebook ne diverge pas de celle du standard écrit, c'est-à-dire si les internautes n'utilisent pas les variantes de l'oral dans leur CNE, cela suggérerait que la variation n'est pas conditionnée par le contexte socio- ou technonumérique. En revanche, si la variation dans la CNE retrouvée sur Facebook est différente du standard écrit, c'est-à-dire si les internautes emploient des variantes de l'oral à des fréquences non négligeables, cela suggérerait qu'il existe des facteurs internes ou externes, par exemple la diamésie, influant sur la production langagière.

C'est le logiciel GoldVarb<sup>8</sup> qui m'a permis de calculer les statistiques de récurrence. Tout d'abord, j'ai codé chaque occurrence dans le fichier Excel avec une lettre pour les facteurs dépendants et indépendants pertinents : son alternance, son environnement linguistique et la page Facebook dont elle provenait. J'ai ensuite sauvegardé ces codes dans un fichier token *.tkn* pour analyse statistique par GoldVarb. Ce logiciel calcule non seulement les fréquences d'emploi relatives, mais affiche aussi les fréquences selon les codes attribués représentant le contexte (page Facebook dont provient le commentaire) et l'environnement linguistique (élément qui suit la variante). Cela permet de déterminer si le contexte a un impact sur le choix de variante.

### **3.8.4 Limites possibles de la méthodologie**

---

<sup>8</sup> Le guide utilisateur de GoldVarb est disponible à URL <https://albuquerque.bioinformatics.uottawa.ca/goldvarb/goldmanual.dir/gvmanual.html>

Alors que les méthodes ethnographiques fournissent des descriptions et des interprétations de la langue en contexte, le variationnisme est à sa base quantitatif, et génère des descriptions statistiques de tendances dans les traits linguistiques de groupes d'individus (Dray et Drummond, 2019 : 242). En particulier, la méthode variationniste évite la subjectivité, étant ancrée dans l'objectivité et la réduction de tout effet du chercheur (Dray et Drummond, 2019 : 243). De plus, en utilisant des données publiquement accessibles, la méthode d'extraction des données employée, nous l'avons vu, évite que les participants ne soient influencés par la présence de la personne menant l'étude.

En revanche, ces méthodes variationnistes deviennent paradoxalement moins ethnographiques, et moins axées sur les caractéristiques, les motivations et les interactions des individus. Il est par exemple impossible de savoir exactement qui a contribué aux données de ce corpus. Ainsi, on ne peut pas savoir pourquoi un.e internaute aurait décidé d'employer une variante aux dépens d'une autre. De plus, en raison du manque de renseignements au sujet des internautes dont la production langagière figure dans le corpus, il n'est pas possible de faire une analyse socialement stratifiée des données – c'est-à-dire des fréquences d'emploi des marqueurs selon l'âge, le sexe, le statut socio-économique, le niveau d'éducation. Cette limite ne change pas la substance des résultats, mais empêche de connaître plus précisément quels internautes font quels choix linguistiques dans la production de leur CNE. Ensuite, ma méthode d'échantillonnage a exclu une grande quantité de contenu textuel, puisque je n'ai pas tenu compte de la production écrite des internautes n'affichant pas publiquement leur lieu de résidence. On peut se contenter de connaître le grand pourcentage de Québécois.es de toutes tranches d'âge qui se branchent sur Facebook.

### **3.9 Résultats attendus et hypothèses**

L'hybridité de la CNE utilisée dans le Web 2.0 a été attestée par d'innombrables linguistes et sociolinguistes. Les recherches portant sur le FQ dans le (t)chat et dans les SMS montrent qu'en général, les tendances linguistiques trouvées en ligne suivent celles de la communauté extérieure, mais aussi que les internautes suivent une nouvelle série de normes. Étant donné que la situation communicative dans les commentaires publics sur Facebook diffère d'autres contextes technologiques qui ont été étudiés, je m'attends à des résultats qui divergent tant du standard écrit

que de l'oral, et des résultats des études du FQ dans les autres modalités numériques. En particulier, je me propose de vérifier les hypothèses suivantes, au cours de mes analyses:

1. Au niveau macrolinguistique, les tendances de la CNE en FQ s'aligneront surtout sur les tendances de l'écrit standard du FQ, avec seulement quelques traces de l'oral :
  - a. Comme la longueur des messages sur Facebook n'est pas limitée (diamésie), et que la communication a lieu dans un espace public (diaphasie), les internautes auront tendance à utiliser les formes standards dans leur CNE. On constatera alors une prépondérance de formes standards.
  - b. En revanche, en raison de la tendance vers l'économie du langage, ainsi que la nature conversationnelle de la CNE (diaphasie), les internautes privilégieront les formes qui sont plus courtes/plus faciles à écrire.
2. Au niveau microlinguistique, le comportement linguistique différera d'un commentaire à l'autre, mais le standard écrit sera privilégié.
  - a. En raison de la taille illimitée des messages sur Facebook, ainsi que la nature publique et anonyme de cette situation de communication, la majorité des internautes utiliseront des formes standards dans leur CNE. D'autres facteurs externes pourraient également jouer un rôle (p.ex. L'âge, le niveau de scolarité, le statut socioéconomique etc.) dans la variation sociolinguistique d'internaute en internaute, mais en raison de la manière dont les données ont été recueillies, ces renseignements ne sont pas disponibles.

### **3.10 Conclusion**

Ce chapitre a présenté les démarches méthodologiques de mon étude, et a décrit comment les méthodes établies de la linguistique de corpus s'adaptent à l'analyse de la CNE dans un terrain sicionumérique. Mon étude, qui emploie des méthodes quantitatives, suit la tradition méthodologique des recherches antérieures en sociolinguistique variationniste sur le FQ. Dans ce chapitre, j'ai décrit les objectifs précis de ma recherche, les variables qui seront examinées, ainsi que les outils utilisés pour recueillir et analyser les données. De plus, ce chapitre a expliqué les

particularités de la plateforme Facebook, ainsi que les caractéristiques sociales et technologiques de la situation de communication à l'étude. J'ai présenté les deux méthodes d'analyse employées dans cette étude, notamment celles des concordances et des collocations, qui permettront de dégager les tendances tant au niveau macrolinguistique (dans la communauté) qu'au niveau microlinguistique (chez des individus). Cette combinaison de méthodes représente, à ma connaissance, une première dans les recherches variationnistes sur le français québécois dans le numérique, et permet de mettre à l'épreuve mes hypothèses sur le comportement linguistique de la communauté et de l'individu. Ce choix de méthodes permet ainsi d'examiner de manière empirique la variation sociolinguistique et ainsi le degré d'hybridité et d'oralité de la CNE produite par des internautes sur Facebook situés au Québec.

## CHAPITRE 4 : RÉSULTATS DES ANALYSES DE CONCORDANCE

### 4.0 Introduction

Dans ce chapitre, je présente les résultats des analyses statistiques des données textuelles que j'ai recueillies de la plateforme Facebook. Dans un premier temps, je décris le contenu linguistique du corpus, avec plusieurs exemples de production écrite, pour montrer le degré de variabilité et la nature québécoise de cette CNE. Deuxièmement, je présente les résultats des analyses de concordance pour les éléments lexicaux, verbaux, syntaxiques et pronominaux qui font l'objet de ma recherche. C'est-à-dire que j'expose les fréquences d'emploi relatives à l'intérieur de chaque variable. Ensuite, je compare mes résultats à ceux retrouvés dans des contextes oraux et dans des contextes CNE, lorsque possible. Enfin, il s'agit d'une discussion globale des résultats de concordance, et des implications possibles.

### 4.1 Un corpus de CNE québécoise

Alors que la délimitation géographique du présent corpus repose sur les pages ciblées et les coordonnées affichées par les internautes, le corpus est manifestement de nature québécoise. En particulier, tout en témoignant des caractéristiques de la CNE, le contenu linguistique du corpus reflète également celles de la variété de français employée au Québec. D'une part, les commentaires extraits de Facebook contiennent des abréviations, des marques d'oralité, des émojis et l'emploi de signes de ponctuation, entre autres. D'autre part, le corpus présente, outre les éléments linguistiques que j'examine plus loin, maintes traces du français québécois, notamment l'élision du [l], comme dans *y va toujours avoir des experts (1830, La Presse)*, des traits lexicaux comme *pantoute* et *icitte*, des tournures québécoises comme *astheure* pour *à cette heure*, l'interrogation en suffixe *-tu*, et *tsé* pour *tu sais*. Les extraits suivants illustrent la nature hybride et québécoise du contenu du corpus. Dorénavant, le numéro entre parenthèses suivant l'extrait est le code anonymisé et arbitraire qui représente l'internaute ayant produit le commentaire. La page Facebook dont provient l'extrait est également indiquée entre parenthèses.

Alors **pkoi** dites-vous des niaiseries si vous vous en sacrez ? (1450, *Le Devoir*)

Encore dans **l'champ... y a d'la** carence en culture générale; vous devriez vous abstenir de commenter, pas du tout à votre avantage **entk**... (1450, *Le Devoir*)

cest **bizz** pareille **lol** (1644, *TVA Nouvelles*)

**omg** jen veux **pu** 😞 (1242, *Journal de Québec*)

Ah oui la bonne eau fraîche du port... **C'est tu** juste moi **ouuu** ? 😊 (1300, *Journal de Québec*)

**Tabarouette** que le «jem'enfoutisme» s'exprime sans gêne ni honte **astheure**... (1450, *Le Devoir*)

**coudons** vos journaliste **dorment tu** au gaz à R.C. (1543, *Le Devoir*)

**Tser** quand tu as la chance de **pogner** la tourista... à Québec 😞😞😞😞😞 (1323, *Journal de Québec*)

Dans les extraits précédents, on constate plusieurs éléments typiques de la CNE, tels que des abréviations et troncations de mots français comme *pkoi*, *bizz*, *entk* et d'expressions anglaises comme *omg*, et *lol*, la représentation de la voix à travers la répétition de lettres comme dans *ouuu* et enfin l'emploi d'émojis. Ces extraits illustrent également la saveur québécoise du corpus : on trouve des exemples de l'écriture phonétique comme dans *pu*, *l'champ*, *d'la*, *tser*, ainsi que plusieurs expressions et tournures et éléments lexicaux québécois comme l'interrogation en suffixe *-tu*, *tabarouette*, *astheure*, *coudons*, *pogner*, typiques de l'oral.

Ce qui surprend, c'est que certaines de ces expressions ne sont pas plus rapides à dactylographier que d'autres formules possibles. Cela indiquerait que le besoin de rapidité et l'économie du langage ne sont pas les seuls facteurs qui entrent en jeu, lorsque les internautes au Québec écrivent sur Facebook. En revanche, on remarque aussi dans les extraits précédents que certains internautes choisissent d'exclure des apostrophes ou des signes de ponctuation, notamment les virgules ; ces faits suggèrent que l'internaute au Québec fait bel et bien des choix délibérés en ce qui concerne sa production écrite sur Facebook. Ces extraits montrent aussi que la variation dans cette CNE se situe à plusieurs niveaux : variation stylistique, variation de codes standard et non standard/jargon Internet, variation de langue (anglais, français) et variation de codes écrits et iconiques.

## 4.2 Présentation des résultats

Dans les prochaines sections sont présentées les analyses statistiques des fréquences d'emploi relatives pour chacune des variables à l'étude, à partir des analyses de concordance. Une fois les résultats présentés, je les compare aux fréquences attestées à l'oral, et le cas échéant, aux fréquences dans d'autres contextes numériques.

J'analyse dans ce chapitre des éléments linguistiques qui font l'objet d'études variationnistes depuis plusieurs décennies, surtout dans des contextes à l'oral, tant formels qu'informels, mais plus récemment aussi dans des contextes écrits numériques (Bally, 2022 ; Blondeau *et al.*, 2014 ; Blondeau et Tremblay, 2022 ; Tremblay *et al.*, 2019 ; van Compernelle, 2010). Chaque variable linguistique que j'examine est alors attestée en français québécois. Dans le cadre de mon étude, j'analyse les fréquences d'emploi relatives des variantes de chaque variable dans un contexte numérique qui, à ma connaissance, n'a pas été examiné. Il importe de noter que dans mon corpus d'environ 46000 mots, chacune des variables que j'ai choisies d'examiner est présente – certaines beaucoup plus fréquentes que d'autres, et le marqueur de conséquence *so* est la seule variante absente du corpus. Le Tableau 4.2 présente le total des occurrences par variable :

**Tableau 4.2 - Total des occurrences dans le corpus par variable à l'étude**

<b>Variable</b>	<b>Nombre d'occurrences</b>
Marqueurs de contrainte ( <i>juste, seulement, ne...que, rien que</i> )	133
Marqueurs de conséquence ( <i>alors, donc, fait que, so</i> )	116
Expression du futur (analytique et synthétique)	374
Emploi variable des auxiliaires <i>avoir</i> et <i>être</i> au passé composé	340
Négation (emploi et omission de la particule <i>ne</i> )	841
Interrogation directe ( <i>est-ce que, inversion, sans inversion, -tu</i> )	205
L'emploi des pronoms <i>on</i> et <i>nous</i>	399
L'emploi des pronoms simples et composés	329
<b>Total des occurrences de toutes les variantes</b>	<b>2737</b>

Au total, toutes les occurrences des variantes représentent environ 6 % de la taille du corpus. Les variantes les plus fréquentes sont l'emploi de la négation, les variantes verbales et les variantes pronominales, tandis que les variantes les moins fréquentes sont les variantes lexicales. En fait, les variantes lexicales représentent seulement 249 occurrences, comparativement aux variantes verbales avec 714 occurrences, pronominales, avec 728 occurrences et syntaxiques avec 1046 occurrences. Ces différences seront discutées lors de la présentation des résultats dans ce chapitre, et dans le Chapitre 6, lors de la discussion générale des résultats.

### **4.3 Résultats de l'analyse**

#### **4.3.1 Analyse des marqueurs lexicaux**

Dans cette première section, je présente les résultats des analyses de concordance des marqueurs sociolinguistiques lexicaux de contrainte et de conséquence, qui font depuis plusieurs années l'objet d'étude surtout au Canada (Blondeau *et al.*, 2017 et 2019 ; Dessureault-Dober, 1974 ; Dostie, 2006 ; Falkert, 2006 ; Léard, 1983 ; Massicotte, 1984 et 1986 ; Mougeon *et al.*, 2016 ; Nadasdi et Keppie, 2004 ; Rehner et Mougeon, 1998 ; Thibault et Daveluy, 1989) mais également

ailleurs dans la Francophonie (Bolly et Degand, 2009 ; Ferre, 2011 ; Konan, 2013 ; Leeman, 2004 ; Mosegaard-Hansen, 1997). Tout d'abord, il s'agira de présenter les résultats portant sur les marqueurs de contrainte, et ensuite ceux portant sur les marqueurs de conséquence. Notons que la présence/fréquence de ces marqueurs en français québécois n'a pas encore été examinée dans un autre contexte numérique, donc mes comparaisons se limiteront à l'oral.

#### **4.3.1.1 Analyse des marqueurs de contrainte**

Les marqueurs de contrainte à l'étude que j'ai cherchés dans le corpus sont *juste*, *seulement*, *ne...que* et *rien que*. Ces quatre expressions expriment toutes le sens sémantique de *seulement* en français québécois. D'après Mougeon *et al.* (2016 : 115) ce sont « des formes qui remontent à plus ou moins loin dans l'histoire de la langue » (121). En effet, *juste* serait le marqueur de restriction le plus récent, tandis que les emplois attestés des expressions *seulement*, *ne...que*, et *rien que* en français remonteraient à plusieurs siècles. À partir de corpus oraux, les études antérieures des années 1980 (Massicotte 1986 ; Thibault et Daveluy, 1989) ont trouvé que *juste* était plus fréquent que *rien que* et *seulement* en français québécois (Rehner et Mougeon, 1998). Quant à *ne...que*, cette formule est très rarement employée dans les corpus oraux (Mougeon *et al.*, 2016 : 122). Étant donné que la présente étude est la première à examiner ces marqueurs dans la CNE, comment ces tendances de l'oral se traduisent-elles dans le numérique ?

À la suite de ces études du français québécois parlé, j'ai alors cherché les quatre mêmes expressions de restriction dans le corpus Facebook : le Tableau 4.3.1.1a présente des exemples tirés du corpus de l'emploi de chacun des marqueurs.

**Tableau 4.3.1.1a - Exemples de l'emploi des marqueurs de contrainte dans le corpus**

<b>Variante</b>	<b>Exemple</b>	<b>Page</b>
<i>Juste</i>	Pensez-vous vraiment que c'est <b>juste</b> les jeunes ? (2124) Les riches c est <b>juste</b> des profiteurs (1203)	<i>La Presse</i> <i>Journal de</i> <i>Montréal</i>
<i>Seulement</i>	les cies aeriennes oui mais pas <b>seulement</b> eux... (2453) Il n'y a pas <b>seulement</b> des coliformes dans l'eau du fleuve (1408)	<i>La Presse</i> <i>Journal de</i> <i>Québec</i>
<i>Ne...que</i>	Ce <b>n'est que</b> de la bêtise. (2532) ...nous voilà dans cette inflation hors control....et ce <b>n'est que</b> le début (1187)	<i>La Presse</i> <i>TVA Nouvelles</i>
<i>Rien que</i>	Pas <b>rien que</b> ça c'est rendu que des professeurs... (1501) <b>Rien que</b> le fait qu'ils aient attendu d'avoir quitté... (1490)	<i>Le Devoir</i> <i>Le Devoir</i>

Lors des analyses de concordance, j'ai cherché les graphies standards, autant que non standards, souvent en utilisant le symbole \* qui permet de chercher toutes les occurrences peu importe ce qui suit le symbole \*. Par exemple, en cherchant « ju\* » j'ai pu trouver toutes les occurrences de mots qui commencent par « ju », y-compris, *juste*, aussi bien que *just*. Le Tableau 4.3.1.1b montre les graphies cherchées pour chaque variante.

**Tableau 4.3.1.1b - Graphies cherchées des marqueurs de contrainte**

<b>Variante</b>	<b>Graphies cherchées</b>
<i>Juste</i>	Juste, ju*
<i>Seulement</i>	Seulement, seulement, slmt, smnt, seu*
<i>Ne...que</i>	Ne...que, ne ...qu', n' ...que
<i>Rien que</i>	Rien que, rien qu', rienque, rien*, rinqe, inque, yinqe

#### **4.3.1.1.1 Analyse des données des marqueurs de contrainte**

Le corpus compte un total de 133 occurrences de marqueurs de contrainte, et la variante la plus fréquente est *juste*, avec 96 occurrences, et ce dans chacune des composantes du corpus. En

d'autres termes, le marqueur *juste* est présent dans les commentaires tirés des cinq pages Facebook desquelles j'ai extrait des données. Notons que dans toutes les occurrences des marqueurs de contrainte, j'ai constaté une seule graphie non standard, notamment un emploi de *just* dans la composante du *Journal de Québec*:

*Fait just 40ans que le quai a cote se vide la mais bon* (1353)

La deuxième variante la plus fréquente est *ne ... que*, qui apparaît 18 fois, et ce, chez 18 internautes distincts. Contrairement à *juste*, qui reflète sans doute l'influence de l'anglais (Mougeon *et al.*, 2016) et représente un usage plutôt informel, la formule *ne...que* représente un niveau de langue très soutenu, donc il est surprenant que cette variante soit la deuxième plus fréquente dans un corpus de CNE. Sur ce point, Raymond Mougeon et ses collègues notent qu'à l'oral « [d]ans les corpus de Montréal, *ne... que* n'est employé que par une faible minorité de locuteurs hautement éduqués qui maintiennent systématiquement la particule *ne* » (2016 : 122). En effet, la fréquence des occurrences de *ne...que* est quelque peu inattendue, surtout dans un contexte informel en ligne. Cette formule apparaît dans trois des cinq composantes du corpus, étant absente des commentaires extraits des pages du *Journal de Québec* et du *Journal de Montréal*. Ceci pourrait s'expliquer par la plus petite taille de ces composantes du corpus, comparativement aux trois autres. Ce mélange dans le corpus de deux variantes qui se situent aux deux extrêmes du continuum informel-formel, *juste* et *ne...que*, témoigne de la nature hybride, sur le plan stylistique, de la CNE par les internautes québécois.es sur Facebook. Il se pourrait aussi que ce soit une diversité d'individus, de différents groupes d'âge et avec différents niveaux d'éducation, qui écrivent dans la section des commentaires, et donc qui contribuent à cette variation entre deux extrêmes stylistiques. Par exemple, d'après Raymond Mougeon et ses collègues, *juste* « est hautement favorisé par les locuteurs les plus jeunes » (2016 : 135) tandis que *ne...que*, marginal dans la plupart des corpus, « est une forme d'hyper-norme » (141).

Ensuite, le marqueur *seulement* apparaît 17 fois, et ce à travers les cinq composantes du corpus, comme c'est le cas pour *juste*. Il importe de noter que j'ai exclu des analyses une occurrence de *seulement* qui apparaît en fin d'énoncé, et n'est alors dans ce cas pas équivalent aux trois autres marqueurs de restriction, qui ne peuvent pas syntaxiquement apparaître en finale.

Occurrence exclue : *Si il est réélu seulement!* (2178, *La Presse*)

Pour ce qui est de la variante *rien que*, elle n'apparaît que deux fois dans le corpus de deux internautes distincts – et ce dans les commentaires tirés de la page Facebook du journal *Le Devoir* uniquement (voir les extraits ci-dessus dans le Tableau 4.2.1.1a). Le petit nombre d'occurrences de cette variante, qui s'aligne plutôt sur l'oral, comparativement au nombre inattendu d'occurrences de la variante *ne...que*, qui est une variante avec une fréquence faible à l'oral (Mougeon *et al.*, 2002), est quelque peu perplexé. De plus, ces deux variantes contrastives, *rien que* et *ne ... que*, se trouvent toutes les deux dans les commentaires des publications de la page *Le Devoir*, mais pas par les mêmes internautes. Cela suggère que des individus provenant d'une diversité d'âges et de backgrounds socio-économiques commentent les publications de cette page en particulier.

#### 4.3.1.1.2 Analyse quantitative/statistique descriptive des données des marqueurs de contrainte

Le Tableau 4.3.1.1.2 présente les occurrences et les fréquences d'emploi relatives des marqueurs de contrainte.

**Tableau 4.3.1.1.2 - Analyse quantitative/statistique descriptive des données des marqueurs de contrainte**

Variante	Nombre d'occurrences	Fréquence d'emploi relative
<i>Juste</i>	96	72 %
<i>Seulement (que)</i>	17	13 %
<i>Ne...que</i>	18	13,5 %
<i>Rien que</i>	2	1,5 %
TOTAL	133	100

Les résultats montrent que le marqueur *juste* apparaît dans 71,6 % des occurrences de contrainte, tandis que *seulement* et *ne...que* sont utilisés 13 % et 13,5 % du temps, respectivement ; enfin, *rien que* n'apparaît que dans 1,5 % des cas. L'emploi le plus fréquent de *juste* pourrait s'expliquer tant par le principe de l'économie du langage (Crystal, 2011) ou économie temporelle (Marcoccia, 2016) que par l'influence de l'anglais (Rehner et Mougeon, 1998). C'est la variante qui serait la plus rapide et la plus simple à dactylographier, comparativement aux trois autres – *juste* comporte seulement cinq lettres, aucun espace, et aucune ponctuation – et c'est celle qui a un équivalent analogue en anglais.

Pour ce qui est de l'environnement linguistique, les marqueurs de restriction sont suivis le plus souvent d'un nom, soit 42 occurrences, ou 31 % des cas, ou d'une préposition, notamment 34 occurrences, ou 25 % des cas. En revanche, les marqueurs *juste* et *ne...que* sont les seuls à être employées avec un verbe : la formule *ne...que* est la seule à être employée avec un verbe à l'imparfait, au conditionnel ou au subjonctif, et *juste* est la seule à être suivie d'un adverbe et d'un adjectif. De plus, *juste* et *ne...que* sont les seules à être employées avec un verbe à l'infinitif. Notons en même temps que le petit nombre d'occurrences totales ne permet pas de faire de remarques conclusives sur l'effet de l'environnement linguistique sur le choix d'une variante ou d'une autre.

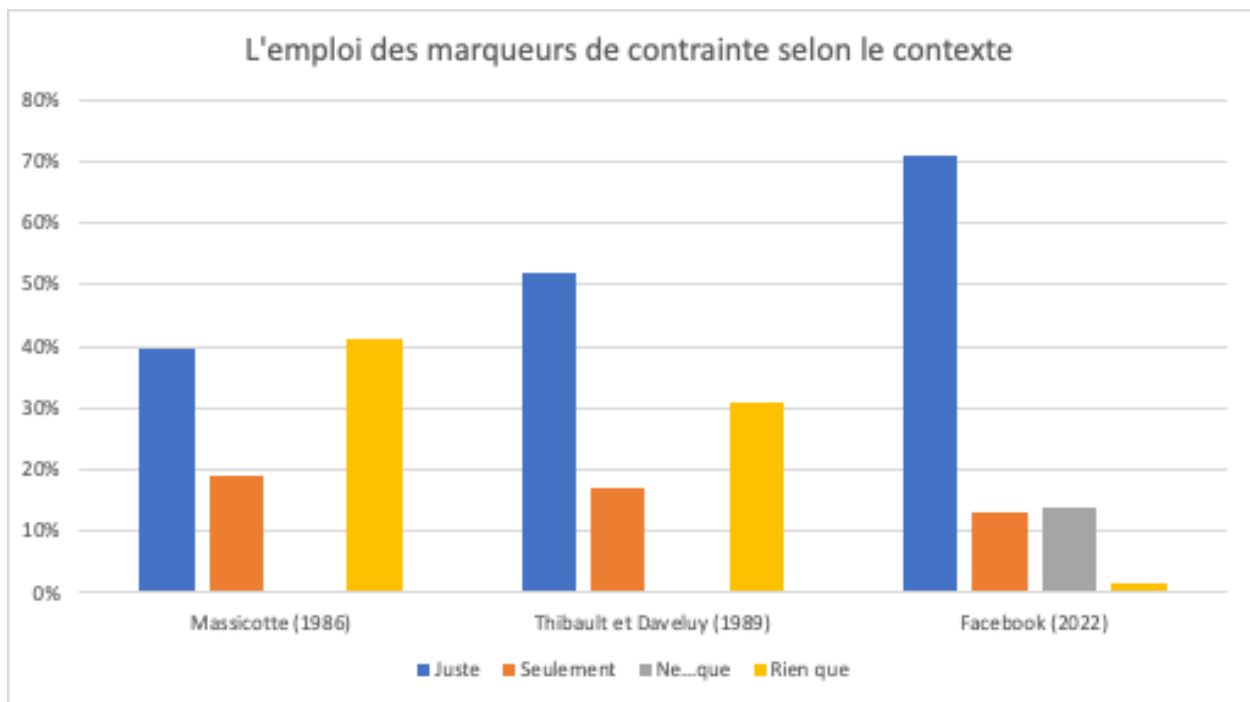
#### **4.3.1.1.3 Comparaison aux fréquences d'emploi attestées à l'oral**

Le Tableau 4.3.1.1.3 présente les résultats de recherches antérieures ayant entrepris des analyses statistiques de l'emploi des marqueurs de contrainte dans des contextes oraux, et y appose les résultats de mon analyse. La Figure 4.3.1.1.3 montre quant à elle une représentation visuelle des données du Tableau 4.3.1.1.3. Il importe de noter que les études de Massicotte (1986) et de Thibault et Daveluy (1989) n'ont pas examiné l'emploi de la formule *ne...que* dans leurs analyses, en raison de la présence marginale de cette variante dans leurs corpus.

**Tableau 4.3.1.1.3 - Comparaison des fréquences d'emploi des marqueurs de contrainte attestées à l'oral et sur Facebook**

Variante	Massicotte (1986)	Thibault et Daveluy (1989)	Facebook (2022)
<i>Juste</i>	39,7 %	52 %	71,6 %
<i>Seulement</i>	19 %	17 %	13 %
<i>Ne...que</i>	exclu	exclu	13 %
<i>Rien que</i>	41,3 %	31 %	1,5 %

**Figure 4.3.1.1.3 - Comparaison des fréquences d'emploi des marqueurs de contrainte attestées à l'oral et sur Facebook**



La Figure 4.3.1.1.3 illustre de manière nette les tendances dans l'emploi des quatre marqueurs de contrainte. Tout d'abord, sur Facebook, *juste* est de loin le marqueur de contrainte utilisé le plus souvent par les internautes au Québec. On constate qu'effectivement, les emplois sur Facebook de *juste* et de *seulement* suivent généralement les tendances de l'oral, tandis que l'emploi du marqueur *rien que*, qui est la variante en deuxième rang à l'oral, est presque négligeable dans les commentaires sur Facebook. Les données de 1986 et de 1989 suggèrent une tendance décroissante dans l'emploi de cette variante, et donc il se peut que la rareté de son emploi sur Facebook suive ce parcours. La fréquence d'emploi de la formule *ne...que* est la seule qui ne semble pas suivre les

tendances communautaires, mais il est difficile de faire de comparaison fiable puisque les études antérieures, datent d'une part déjà de quelques décennies, et ont d'autre part exclu cette variante de leurs analyses.

La fréquence de *ne...que* sur Facebook pourrait s'expliquer en partie par les taux d'utilisation de cette plateforme, qui sont plus élevés chez les Québécois.es âgés de plus de 75 ans, qui, selon le dernier recensement provincial, utilisent Facebook à un taux de 98 %, tandis que chez les 55-64 ans, c'est 92 %, les 65-74 ans, 93 %, et chez les groupes d'âge plus jeunes (18-54 ans), les taux sont entre 68 % et 91 % (Bourget *et al.*, 2022 : 7). Il est possible aussi que les 19 internautes qui ont choisi d'employer cette variante aient un niveau d'éducation élevé, suivant les résultats de l'étude de Massicotte (1986) et les propos de Mougeon *et al.* (2016). Puisqu'il n'est pas possible de tirer des conclusions définitives sur cette variante en particulier, ainsi que la formule *rien que*, les deux méritent plus d'attention dans des recherches futures.

Alors que les marqueurs de contrainte n'ont pas encore été analysés dans un autre contexte numérique, la comparaison des fréquences d'emploi à l'écrit sur Facebook à celles de l'oral permet de constater que les emplois de *juste* et de *seulement* sur Facebook suivent généralement les tendances de la communauté linguistique à l'oral : *juste* est le plus fréquent dans l'étude de Thibault et Daveluy (1989), comme dans la nôtre. De même, la fréquence de *seulement* sur Facebook (13 %) se rapproche des fréquences à l'oral de 19 % et 17 %. En revanche, les fréquences d'emploi de la formule *rien que* sont l'inverse sur Facebook de ce qu'elles sont à l'oral : *rien que* est le marqueur le moins fréquemment employé sur Facebook, tandis qu'il est l'un des plus fréquents, ou le plus fréquent, dans les corpus oraux.

On pourrait alors postuler qu'en général, dans la CNE sur Facebook produite par les internautes au Québec, on assiste à une simplification du système qui comporte certaines ressemblances à l'oral. On constate la préférence pour *juste* qui l'emporte sur la formule *rien que* comparativement à l'oral. Cela est peut-être en raison de l'effort nécessaire pour écrire *rien que*, formule qui a été pratiquement écartée du système des marqueurs de contrainte puisqu'elle ne représente que 1,5 % des cas sur Facebook. Ce qui surprend, c'est que le taux de fréquence d'emploi de la formule *ne...que* est pratiquement identique au taux d'emploi de *seulement*. Il est fort probable que les

internauts font des choix, délibérés ou inconscients, pour atteindre des buts spécifiques lorsqu'ils emploient ces variantes. Comme piste future, il serait intéressant d'explorer cette question plus en détail, par le biais d'un questionnaire auprès d'internautes au Québec, au sujet de leur comportement et de ce qui motive leurs choix langagiers dans la sphère publique numérique.

#### 4.3.1.2 Analyse des marqueurs de conséquence

Les marqueurs de conséquence à l'étude sont *alors*, *donc*, *fait que* et *so*, ce dernier étant un emprunt de l'anglais. Selon le dictionnaire *Usito*, *donc* et *alors* remontent à plusieurs siècles, tandis que *(ça) fait que* apparaît dans le *Dictionnaire de la langue québécoise* de Léandre Bergeron (1980) sous la forme *ça fa que*, et signifie « de sorte que, c'est pourquoi ». La formule *so* n'apparaît pas dans ces deux dictionnaires. Les recherches antérieures du français québécois parlé ont montré que la formule *(ça) fait que* est employée le plus souvent dans des situations informelles, mais aussi que son emploi a augmenté dans les 30 dernières années, pour devenir la forme la plus fréquente à l'oral, aux dépens de *donc* et *alors* (Beaulieu *et al.*, 2019). Il importe de noter que « [d]ans les communautés francophones hors Québec, la variante anglaise *so* est par ailleurs attestée comme concurrente directe de *(ça) fait que* » (Beaulieu *et al.*, 2019 : 3). En d'autres mots, *donc* et *alors* sont ainsi les formules plutôt standards, tandis que *(ça) fait que* représente le vernaculaire (Blondeau et Tremblay, 2022a), et *so* relève plutôt du français parlé en dehors de la province du Québec.

En m'appuyant sur les études du français québécois parlé, j'ai alors cherché ces quatre mêmes expressions de conséquence dans le corpus Facebook : le Tableau 4.3.1.2a présente des exemples de l'emploi de chacun des marqueurs.

**Tableau 4.3.1.2a - Exemples de l'emploi des marqueurs de conséquence dans le corpus**

Variante	Exemple	Page
<i>Alors</i>	à quoi servent les taxes sur l'essence <b>alors</b> ? ? ? ? (2171) Le mot existe, il doit <b>alors</b> être prononcé pour expliquer les choses (1497)	<i>La Presse</i> <i>Le Devoir</i>
<i>Donc</i>	comme tu as <b>donc</b> raison (1841) Rester <b>don</b> chez vous ceux et celles qui ... (2666)	<i>TVA Nouvelles</i> <i>La Presse</i>
<i>Fait que</i>	<b>Fack</b> la guerre est pas du tout se que vous montrer... (2633) Situation qui était prévisible <b>fac</b> les seuls personnes à blâmer c'est soi-même (1716)	<i>La Presse</i> <i>TVA Nouvelles</i>
<i>So</i>		

Il importe de noter que lors des analyses de concordance, j'ai cherché les graphies standards, autant que non standards, souvent en utilisant le symbole \* qui permet de chercher toutes les occurrences peu importe ce qui suit le symbole \*. Par exemple, chercher « don\* » permet de trouver toutes les occurrences de mots qui commencent par « don », y compris, *donc*, aussi bien que *don*. Le Tableau 4.3.1.2b montre les graphies cherchées pour chaque variante.

**Tableau 4.3.1.2b - Graphies cherchées des marqueurs de conséquence**

Variante	Graphies cherchées
<i>Alors</i>	Alors, alor*
<i>Donc</i>	Donc, don*
<i>Fait que</i>	Fait que, fait*, fac*, fak*, faq*, ça fa que, faitque, faque
<i>So</i>	So*

#### 4.3.1.2.1 Analyse des données des marqueurs de conséquence

Le corpus compte seulement 104 occurrences de marqueurs de conséquence, et la variante la plus fréquente est *donc*, avec 60 occurrences, et ce dans chacune des composantes du corpus. Notons qu'il y a dans le corpus deux instances seulement de la même graphie non standard, en l'occurrence, *don*. La deuxième variante la plus fréquente est *alors*, qui apparaît 40 fois. Je n'ai

constaté aucune graphie non standard de la variante *alors*. En revanche, notons qu’il y a dans le corpus 12 occurrences de la formule *alors que* ; ces occurrences ont été exclues de l’analyse puisqu’elles ont un sens contrastif, tandis que le mot *alors* a un sens de conséquence. Ensuite, la variante *fait que* est utilisée quatre fois seulement, par quatre internautes distincts, avec trois graphies non standards distinctes, *fac*, *fack* et *faque*. Ces quatre occurrences se trouvent uniquement dans les deux plus grandes composantes du corpus, notamment *TVA Nouvelles* et *La Presse*. Il est à noter que je n’ai trouvé aucune occurrence de *fait que* précédée du pronom démonstratif *ça*. Enfin, le marqueur *so* n’apparaît pas comme marqueur de conséquence dans le corpus, et une fois seulement comme adverbe emprunté de l’anglais, dans la composante *TVA Nouvelles* : « *So delusional, une vraie honte.* » Cette variante est donc exclue des analyses statistiques des marqueurs de conséquence.

#### 4.3.1.2.2 Analyse quantitative/statistique descriptive des données des marqueurs de conséquence

Le Tableau 4.3.1.2.2 présente les occurrences et les fréquences d’emploi relatives des marqueurs de conséquence.

**Tableau 4.3.1.2.2 - Analyse quantitative/statistique descriptive des données des marqueurs de conséquence**

Variante	Nombre d’occurrences	Fréquence d’emploi relative
<i>Alors</i>	40	38 %
<i>Donc</i>	60	58 %
<i>Fait que</i>	4	4 %
<i>So</i>	0	0 %
TOTAL	104	100 %

Les analyses statistiques montrent que le marqueur *donc* apparaît dans 58 % des occurrences de contrainte, suivi du marqueur *alors* qui apparaît dans 38 % des cas. L’emploi de *fait que* est

marginal, à 4 % des cas, et le marqueur *so*, comme je l'ai indiqué, n'est pas présent dans le corpus comme marqueur de conséquence. Les emplois plus fréquents de *alors* et *donc* pourraient s'expliquer tant par le principe de l'économie du langage que par l'influence du standard écrit. Ce sont les variantes qui seraient les plus rapides et les plus simples à dactylographier, comparativement à *fait que*. En même temps, *alors* et *donc* sont aussi les variantes qui relèvent plutôt du standard (Blondeau et Tremblay, 2022a).

Pour ce qui est de l'environnement linguistique, les marqueurs de conséquence *alors* et *donc* peuvent être en finale d'énoncé ; tandis que *fait que* est toujours en début d'énoncé et uniquement suivi d'un élément nominal ou pronominal dans le corpus. *Alors*, pour sa part, peut être suivi d'un pronom, d'un adverbe, d'une préposition, d'un l'infinitif ou d'un verbe conjugué au présent. *Donc*, en revanche, peut soit être suivi d'un adjectif, d'un adverbe, d'un nom, d'un pronom, d'une préposition ou d'un verbe à l'infinitif, soit être conjugué au présent ou au passé composé. La variante *donc* est ainsi plus syntaxiquement flexible que la variante *alors* dans ce corpus.

En somme, tandis que syntaxiquement des différences légères se présentent, ces trois marqueurs ont le même sens sémantique, et globalement, les marqueurs de conséquence sont suivis 31 % du temps d'éléments nominaux et pronominaux, comme dans *Alors ça aurait un impact...* et *Enlevons donc la lettre N*, et 23 % du temps d'adjectivaux et adverbiaux, comme dans *Alors là je ne comprends plus rien* et *donc facilement 20\$/mois*. Alors que le marqueur *donc* est le plus flexible, *Fait que* est le plus syntaxiquement limité – mais, rappelons-le, très rare dans le corpus avec seulement quatre occurrences – étant suivi uniquement d'éléments nominaux ou pronominaux, comme dans *Fac on laisse la Russie faire ...* et *fac les seules personnes à blâmer*. Notons aussi encore le petit nombre d'occurrences des marqueurs de conséquence : il serait pertinent d'examiner la fréquence de ces marqueurs dans d'autres situations communicationnelles en contexte numérique.

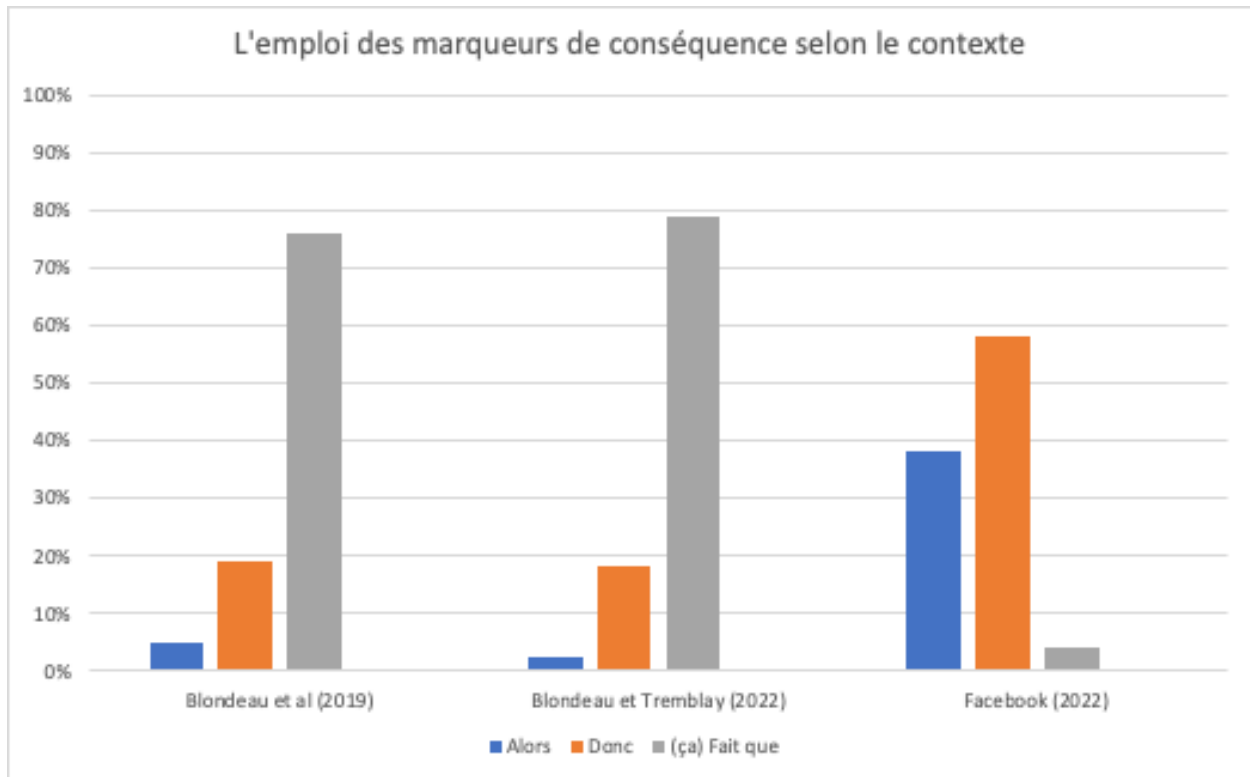
#### 4.3.1.2.3 Comparaison aux fréquences d'emploi attestées à l'oral

Le Tableau 4.3.1.2.3 présente les résultats de recherches antérieures, qui ont entrepris des d'analyses statistiques de l'emploi des marqueurs de conséquences dans des contextes oraux, et reprend les résultats de mon analyse. Il importe de noter qu'à ma connaissance, ces marqueurs n'ont pas encore été examinés en français québécois dans un autre contexte socionumérique. La Figure 4.3.1.2.3 est une représentation visuelle des données du Tableau 4.3.1.2.3

**Tableau 4.3.1.2.3 - Comparaison des fréquences d'emploi des marqueurs de conséquence attestées à l'oral et sur Facebook**

<b>Variante</b>	<b>Blondeau <i>et al.</i> (2019)</b>	<b>Blondeau et Tremblay (2022a)</b>	<b>Facebook (2022)</b>
<i>Alors</i>	5 %	2,6 %	38 %
<i>Donc</i>	19 %	18,4 %	58 %
<i>(ça) fait que</i>	76 %	78,9 %	4 %
<i>so</i>	0 %	exclu	0 %

**Figure 4.3.1.2.3 - Comparaison des fréquences d'emploi des marqueurs de conséquence attestées à l'oral et sur Facebook**



On constate que les tendances de l'emploi des marqueurs de conséquence sur Facebook vont à l'encontre des tendances attestées à l'oral : tandis qu'à l'oral *fait que* est de loin le plus fréquent, son emploi est quasi négligeable sur Facebook, et *donc* et *alors* prennent la relève. Le marqueur *alors* est employé beaucoup plus souvent sur Facebook qu'à l'oral, notamment presque neuf fois plus souvent que dans le corpus de Blondeau *et al.* (2019) et 17 fois plus souvent que dans le corpus de Blondeau et Tremblay (2022a). Ensuite, *donc* est employé environ trois fois plus souvent sur Facebook que dans les deux corpus oraux de 2019 et 2022. Les seuls points communs sont que *donc* est plus fréquent que le marqueur *alors*, et l'absence de *so* des corpus du français québécois numériques et oraux.

Ici on peut argumenter pour une convergence de deux facteurs : celui du standard écrit et celui du principe de la simplification (Thibault, 1991). Par exemple, il est logiquement plus rapide et plus simple d'écrire *alors* ou *donc* que d'écrire *fait que*, que ce soit sur un clavier d'ordinateur ou sur celui d'un téléphone. En revanche, dans des situations informelles à l'oral *fait que* est

monosyllabique, et se prononce simplement /fɛk/ (Blondeau et Tremblay, 2022a), et donc ce marqueur à l'oral est aussi rapide que ses homologues *alors* et *donc*, sur le plan de la simplicité phonétique. D'ailleurs, les quatre occurrences de *fait que* dans le corpus Facebook sont sous forme de graphie non standard, et ont été fusionnées en un mot : *fac* (qui apparaît deux fois), *faque*, et *fack*. Mais comment expliquer l'absence de *so* ? On voit dans les corpus oraux du FQ que *so* n'est presque pas du tout utilisé même à l'oral, donc dans cette instance, cela suggère que la tendance sur Facebook suit celle de l'oral informel au Québec.

En somme, pour ce qui est du système lexical, il semblerait que les internautes au Québec qui écrivent sur Facebook ne suivent pas de près les tendances de l'oral informel, mais se rapprochent plutôt du standard écrit, surtout pour ce qui est des marqueurs de conséquence *alors* et *donc*. La seule exception est l'emploi du marqueur de contrainte *juste* ; bien que son emploi soit plus fréquent sur Facebook qu'à l'oral, il est, dans les deux contextes, le marqueur le plus fréquemment employé. De plus, étant donné le petit nombre d'occurrences de ces deux variables comparativement aux autres variables à l'étude, on peut se demander si les internautes ne sentent pas le besoin d'utiliser ces types formules, qui sont à la base des marqueurs de relation. Une étude se focalisant sur l'emploi de tels marqueurs de relation en général dans la CNE pourrait éclaircir cette question.

#### **4.3.2 Analyse des marqueurs verbaux**

Plusieurs études au fil des années se sont penchées sur les deux alternances verbales à l'étude, dans les variétés du français parlé au Canada (Bigot, 2008 ; Blondeau, 2006 ; Deshaies et Laforge, 1981 ; Réa 2016 et 2020 ; Sankoff et Thibault, 1977 ; Villeneuve, 2017) et ailleurs au monde francophone (Abouda et Skrovec, 2015 et 2017 ; Emirkanian et Sankoff, 1986 ; Jeanjean, 1988 ; Mbey Makang, 2014). Ces marqueurs, notamment l'emploi variable de l'expression du futur et l'emploi variable des auxiliaires *avoir* et *être* au passé composé, sont approximativement aussi fréquents que les marqueurs pronominaux, plus fréquents que les marqueurs lexicaux mais moins fréquents que les marqueurs syntaxiques dans le corpus. En tout, le corpus contient 721 occurrences des deux variables ; de ce chiffre, 374 sont des formules du futur, et 347 du passé

composé. Tout d’abord, il s’agira de présenter les résultats portant sur l’emploi variable du futur, et ensuite l’emploi des auxiliaires au passé composé.

#### 4.3.2.1 Analyse de l’emploi variable du futur

Les marqueurs représentant l’emploi variable du futur à l’étude sont le futur analytique, aussi appelé le futur proche, et le futur synthétique, aussi appelé le futur simple. Afin de simplifier les analyses, je n’ai pas tenu compte de l’emploi du présent pour exprimer le futur. Les chercheurs s’accordent sur le fait que le futur analytique, qui relève surtout de la langue vernaculaire, devient plus fréquent en usage à l’oral que le futur synthétique, qui lui, relève surtout du standard écrit et de la langue surveillée (Blondeau et Labeau, 2016 ; Tremblay *et al.*, 2019).

Suivant la tradition des études sur l’emploi du futur, j’ai alors cherché toutes les occurrences du futur analytique et du futur synthétique dans le corpus Facebook : le Tableau 4.3.2.1a présente des exemples tirés directement du corpus qui illustrent l’emploi de chacune de ces formules verbales.

**Tableau 4.3.2.1a - Exemples de l’expression du futur dans le corpus**

Variante	Exemple	Page
<i>Analytique</i>	JAMAIS <b>je vais mettre</b> mes pieds là (1301) Ce n est qu un début <b>ça va crasher</b> solide malheureusement (2107)	<i>Journal de Québec La Presse</i>
<i>Synthétique</i>	... pas graves <b>j’irai</b> en Ontario (2391) ...y’a rien de gratuit au Québec et <b>on commencera</b> pas par les logements!!! (1164)	<i>La Presse Journal de Montréal</i>

Il importe de noter que lors des analyses de concordance, j’ai cherché les graphies standards, autant que non standards. Pour ce qui est du futur analytique, j’ai commencé par chercher les conjugaisons du verbe *aller* au présent, pour toutes les personnes, gardant seulement ces occurrences qui étaient suivies d’un verbe à l’infinitif (peu importe la graphie de ce deuxième élément). Quant au futur synthétique, en cherchant « \*rai\* » j’ai pu par exemple trouver toutes les occurrences de mots se terminant en « rai », ce qui m’a permis de trouver toutes les formules du

futur synthétique pour la première personne du singulier. Cette stratégie a permis de trouver le maximum d'occurrences, même si le verbe n'était pas conjugué suivant le standard, par exemple si l'internaute avait écrit *j'irais* pour *j'irai*. Le Tableau 4.3.2.1b montre les graphies recherchées pour chaque variante.

**Tableau 4.3.2.1b - Graphies recherchées de l'expression du futur**

Variante	Graphies recherchées
<i>Futur analytique</i>	Vais, va, vas, allon*, all*, von*, mas, m'as, ma, m'a
<i>Futur synthétique</i>	*rai*, *ras, *ra, *rons, *rez, *ront

#### 4.3.2.1.1 Analyse des données de l'emploi variable du futur

En tout, le corpus contient 374 occurrences du futur exprimé soit par le futur analytique soit par le futur synthétique. Le futur analytique est la forme la plus fréquente, avec 238 occurrences, et ce dans chacune des composantes du corpus. J'ai constaté quelques graphies non standards : en particulier, j'ai trouvé 22 cas où la terminaison du verbe ne suit pas le standard, un cas de la formule vernaculaire *je m'en vas*, écrit *ma*, qui représente une agglutination du pronom avec le verbe *aller*:

*J'espère que tu va faire la file aujourd'hui pour te baigner dans de l'eau ... (1384,*

*Journal de Québec)*

*On vas savoir pour qui voter (2192, La Presse)*

*me coute le double d'essence [...] ce ptit 20\$ la ma le prendre [je m'en vais le prendre]*

*certain (2180, La Presse)*

*jvai pt laisser faire le japon (1770, TVA Nouvelles)*

Par contraste, le corpus contient un total de 136 occurrences du futur synthétique à travers les cinq pages composantes. J'ai également constaté cinq divergences des terminaisons standards, et une graphie non standard du futur synthétique :

*fait moi un live quand tu y sera comme ça si y'a un pépin j'arrive... (1361, Journal de*

*Québec)*

*Le Legault Quebécois baisseras pas les taxes (2336, La Presse)*

...en tuka jespere tu le *fra* pas (2211, *La Presse*)

#### 4.3.2.1.2 Analyse quantitative/statistique descriptive des données de l'emploi variable du futur

Le Tableau 4.3.2.1.2 présente les occurrences et les fréquences d'emploi relatives des marqueurs de l'emploi du futur.

**Tableau 4.3.2.1.2 - Analyse quantitative/statistique descriptive des données de l'emploi variable du futur**

Variante	Nombre d'occurrences	Fréquence d'emploi relative
Futur analytique	238	63,6 %
Futur synthétique	136	36,4 %
TOTAL	374	100 %

L'emploi du futur analytique est alors de loin la formule la plus fréquemment utilisée par les internautes au Québec sur Facebook, et ce dans environ 64 % des cas. En revanche, le futur synthétique est employé dans seulement environ 36 % des cas du futur dans ce corpus. Il importe aussi de noter que c'est la troisième personne du singulier qui est la forme la plus fréquente des conjugaisons du futur analytique, avec 137 des 238 cas, soit 58 % des occurrences. La même tendance se produit dans l'emploi du futur synthétique, la majorité, avec 84 des 136 cas, soit 62 %, étant conjugués à la troisième personne du singulier.

Pour ce qui est de l'environnement linguistique, les études antérieures sur l'emploi variable du futur en français ont suggéré l'importance du type de phrase comme facteur dans l'emploi du futur. En effet, toutes les études variationnistes se focalisant sur le français québécois ont trouvé que la polarité opère sur le choix de variante, les contextes négatifs favorisant le futur synthétique (Tremblay *et al.*, 2019 : 78). Les résultats présents renforcent ce principe : dans les 252 emplois

du futur analytique, seulement quatre cas, soit 2 %, sont à la forme négative, tandis que dans les 138 emplois du futur synthétique, 35 cas, soit 26 %, sont à la forme négative. Cela suggère qu'une phrase négative entraîne plus souvent l'emploi du futur synthétique que du futur analytique dans les écrits par internautes au Québec dans l'espace public de Facebook, comme à l'oral.

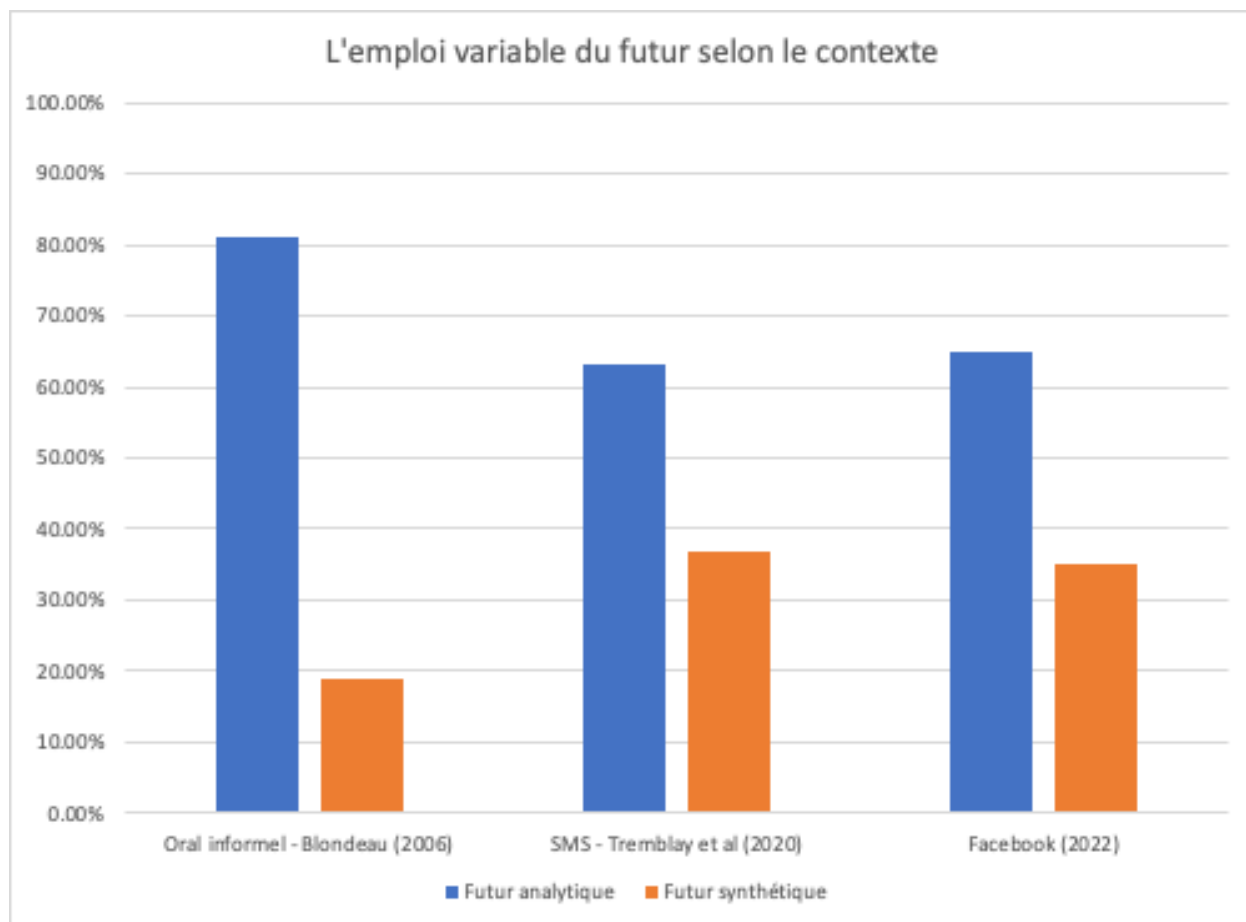
#### 4.3.2.1.3 Comparaison aux fréquences d'emploi attestées à l'oral et dans les SMS

Le Tableau 4.3.2.1.3 présente d'une part les résultats de recherches antérieures, qui ont entrepris des analyses statistiques de l'emploi variable du futur analytique et synthétique dans des contextes oraux et dans les SMS, et reprend d'autre part les résultats de mon analyse. La Figure 4.3.2.1.3 est une représentation visuelle des données du Tableau 4.3.2.1.3.

**Tableau 4.3.2.1.3 - Comparaison des fréquences d'emploi du futur attestées à l'oral, dans les SMS et sur Facebook**

<b>Variante</b>	<b>Oral informel - Blondeau (2006)</b>	<b>SMS - Tremblay <i>et al.</i> (2019)</b>	<b>Facebook (2022)</b>
Futur analytique	81 %	63 %	63,6 %
Futur synthétique	19 %	37 %	36,4 %

**Figure 4.3.2.1.3 - Comparaison des fréquences d'emploi du futur attestées à l'oral, dans les SMS et sur Facebook**



On constate premièrement que les tendances de l'emploi variable du futur sur Facebook suivent généralement les tendances attestées à l'oral dans des contextes informels : dans les deux cas, le futur analytique est la forme la plus fréquente, mais sur Facebook, le futur synthétique est environ 15 % plus fréquent qu'à l'oral, et le futur analytique environ 15 % moins fréquent sur Facebook qu'à l'oral. Deuxièmement, en comparant l'emploi dans les SMS aux emplois sur Facebook, on constate que les fréquences sont pratiquement identiques. Cela suggère qu'il existe des ressemblances entre la CNE par SMS et la CNE sur Facebook. Alors que cela prendrait plus d'effort d'écrire un temps composé qu'un temps simple, peu importe le type de clavier, la préférence du futur analytique peut s'expliquer tant par l'influence de l'oral (Tremblay *et al.*, 2019) que par le principe de l'économie de l'effort (Crystal, 2011). Plus précisément, cette forme comporte moins de particularités et d'exceptions grammaticales : les déclinaisons du futur

analytique, c'est-à-dire le verbe *aller* au présent, sont plus simples que celles du futur synthétique, et par conséquent il ne faut pas se soucier des verbes irréguliers avec le futur analytique.

Les résultats suggèrent d'autre part que les internautes, lorsqu'ils écrivent sur Facebook, choisissent plutôt de suivre les normes communautaires de l'oral informel, au moins pour ce qui est de l'emploi du futur, mais aussi qu'ils semblent utiliser le futur synthétique légèrement plus fréquemment qu'à l'oral. En effet, Hélène Tremblay et ses collègues (2019 : 92), d'après leurs analyses de l'emploi variable du futur dans les SMS au Québec, ont conclu que les messages SMS, au lieu d'être fautifs ou incorrects, obéissent à une différente série de normes, se rapprochant plutôt de la langue orale. Étant donné que « ce qui est exprimé au futur proche est considéré comme moins hypothétique » (OQLF, 2023), il sera utile à l'avenir d'examiner si les sujets des publications commentés par les internautes sur Facebook ont un impact sur le choix entre le futur analytique et le futur synthétique, et d'examiner plus précisément les contextes dans lesquels les deux temps sont employés.

#### **4.3.2.2 Analyse de l'emploi variable des auxiliaires *avoir* et *être* au passé composé**

Les marqueurs de l'emploi variable du passé composé à l'étude sont l'emploi de l'auxiliaire *avoir* et celui de l'auxiliaire *être*. Le standard écrit, ou la norme prescrite, exige au passé composé l'emploi de l'auxiliaire *être* pour les verbes intransitifs tels que *aller*, *venir*, *sortir* et *partir* et l'auxiliaire *avoir* pour tous les autres verbes. Mais Béatrice Réa (2016 : 5) note qu'à travers la francophonie, les « locuteurs natifs emploient à l'oral *avoir* et *être* avec des verbes qui, en théorie, ne peuvent être utilisés qu'avec *être* ». En effet, ce phénomène a été observé dans plusieurs variétés du français, y compris « dans le français de Montréal » (Sankoff et Thibault, 1977).

Suivant ces analyses antérieures, j'ai alors cherché toutes les occurrences du passé composé dans le corpus Facebook : le Tableau 4.3.2.2a illustre des exemples de chacune de ces formules verbales, tirés directement du corpus.

**Tableau 4.3.2.2a - Exemples de l'emploi du passé composé dans le corpus**

Variante	Exemple	Page
<i>Avoir</i>	Semaine dernière à Ottawa <b>j'ai payé</b> 1.96 le litre à 1 coin de là il était 1.91 le litre (2201) <b>Trudeau a fait</b> du Canada le pays de l'hypocrisie...(1462)	<i>La Presse</i> <i>Le Devoir</i>
<i>Être</i>	Le mois dernier <b>je suis allée</b> aux urgences, avec tellement de douleurs...(1792) <b>Nous sommes revenu</b> aujourd'hui...(2604)	<i>TVA Nouvelles</i> <i>La Presse</i>

Il importe de noter que lors des analyses de concordance, j'ai cherché les graphies standards, autant que non standards. Pour ce qui est de l'emploi des auxiliaires *avoir* et *être*, j'ai commencé par chercher les conjugaisons du verbe *avoir* et *être* au présent, pour toutes les personnes, gardant seulement les occurrences qui étaient suivies d'un participe passé, peu importe sa graphie. Par exemple, parfois le participe passé se terminait en *-er* au lieu du *é*. Je me suis aussi assurée de chercher les cas en *j'*, *t'*, *n'*. Le Tableau 4.2.2.2b montre les graphies recherchées pour chaque variante.

**Tableau 4.3.2.2b - Graphies recherchées du passé composé**

Variante	Graphies recherchées
<i>Avoir</i>	J'ai, j'ai, ai, t'a, t'as, a, ya, avon*, ave*, ont, ons, on,
<i>Être</i>	Suis, jsuis, j'suis, sus, chuis, chus, chu, es, t'es, tes, est, c'est, s'est, somme*, ête*, ete*, sont, son,

#### 4.3.2.2.1 Analyse des données de l'emploi variable des auxiliaires *avoir* et *être*

En tout, le corpus contient 340 occurrences du passé composé, mais seulement 58 cas de verbes intransitifs, exprimé soit avec l'auxiliaire *avoir*, soit avec l'auxiliaire *être*. Je n'ai trouvé aucune occurrence avec *être* provenant de la composante *Journal de Québec*, et une seule provient de la composante *Le Devoir*. Quant aux participes passés, des 58 cas de verbes intransitifs, 11, soit 19 % d'entre eux, ont une terminaison du participe passé qui diverge du standard ; en voici deux exemples :

*Mon beau frère est restez coincé dans la toilette. Hahaha (1950, TVA Nouvelles)*  
*Les derniere troupes russes sont partit la queue entre les jambes... (1807, TVA Nouvelles)*

#### **4.3.2.2.2 Analyse quantitative/statistique descriptive des données de l'emploi variable des auxiliaires *avoir* et *être***

Globalement alors, l'emploi des auxiliaires dans le corpus suit le standard ; des 58 cas de verbes intransitifs, je n'ai trouvé que cinq cas dans lesquels l'auxiliaire *avoir* est employé là où le standard écrit demanderait l'auxiliaire *être*. Dans un autre cas, le verbe transitif est accompagné de l'auxiliaire *être*, contrairement au standard. Les Tableaux 4.3.2.2.a et b présentent les occurrences et les fréquences d'emploi relatives des emplois non standards des auxiliaires *avoir* et *être*.

**Tableau 4.3.2.2a - Analyse quantitative/statistique descriptive des données de l'emploi variable de l'auxiliaire avoir**

<b>Verbe</b>	<b>Nombre d'occurrences avec <i>avoir</i></b>	<b>Total des occurrences</b>	<b>Fréquence d'emploi relative</b>
Aller	0	5	0 %
Arriver	0	3	0 %
Décéder	0	2	0 %
Descendre	1	1	50 %
Devenir	0	6	0 %
Rendre	0	22	0 %
Rester	0	1	0 %
Revenir	0	5	0 %
Partir	1	5	20 %
Passer	3	3	100 %
Sortir	0	1	0 %
Venir	0	3	0 %

**Tableau 4.3.2.2b - Analyse quantitative/statistique descriptive des données de l'emploi variable de l'auxiliaire être**

<b>Verbe</b>	<b>Nombre d'occurrences avec <i>être</i></b>	<b>Total des occurrences</b>	<b>Fréquence d'emploi relative</b>
Avoir	1	14	7 %

Les cinq cas dans lesquels l’auxiliaire *avoir* est employé à la place de l’auxiliaire *être*, sont avec les verbes *descendre*, *partir* et *passer* :

*je parle avec ma cousine ainsi qu’un ami qui demeure en Ontario et l’essence n’a pas descendu il est à 2.16\$ (2417, La Presse)*

*J’ai parti a 17 ans de chez mes parent...pi a date pas besoin d’aide (1848, TVA Nouvelles)*

*J’ai passé hier! (1717, TVA Nouvelles)*

*on a passé par là ... pas de trafic juste rendu sur la 40 (1719, TVA Nouvelles)*

*...a failli mourir d’une crise cardiaque au Mexique parce qu’un crabe lui a passé entre les pieds (1760, TVA Nouvelles)*

Le seul cas dans lequel l’auxiliaire *être* est employé à la place de l’auxiliaire *avoir* est avec le participe passé du verbe *avoir*, et provient de la composante *La Presse* du corpus :

*Le Québec ont ce fait fourrer a tour de bras pi le monde d’orme debout parce qu’il sont eu 500\$ dollars il Von voté pour Lui pour ce faire fourrer 5fois plus (2423)*

Il est fort possible que cette instance représente une faute de frappe, puisque le *s-* au début de l’auxiliaire *sont* devrait en effet se placer à la fin du pronom *il* précédant le verbe.

En général, les internautes au Québec sur Facebook suivent le standard écrit en ce qui concerne les auxiliaires *avoir* et *être* au passé composé, dans presque tous les cas. Quant aux verbes intransitifs, qui prendraient l’auxiliaire *être* selon le standard écrit, dans cinq des 58 cas, c’est-à-dire dans environ 10 % des cas, ils emploient l’auxiliaire *avoir*. Plus précisément, selon mes résultats, le verbe *passer* est employé 100 % du temps de manière non standard au passé composé, tandis que les verbes *descendre* et *partir* sont employés de manière non standard dans 50 % et 20 % des cas respectivement.

Il importe de noter les propos de l’Office québécois de la langue française (2023) au sujet de l’emploi des auxiliaires au passé composé avec le verbe *passer*. L’organisme note qu’en pratique « *passer* s’emploie aujourd’hui le plus souvent avec *être*, du moins à la forme intransitive, c’est-

à-dire sans complément d'objet. L'auxiliaire *avoir* [...] reste [...] possible, mais il peut être senti comme vieilli ». Les résultats au sujet de l'emploi de l'auxiliaire *avoir* avec le verbe *passer* au passé composé rejoignent donc par exemple les résultats pour les marqueurs de restriction, où la formule *ne...que*, qui est vieillie, apparaît aussi fréquemment que *seulement*. Quant aux verbes *aller, arriver, décéder, devenir, rendre, rester, revenir, sortir* et *venir*, ils sont employés suivant le standard dans tous les cas. En revanche, nous avons vu aussi que dans environ 20 % des cas de verbes intransitifs, le participe passé ne suit pas le standard écrit. Étant donné le très petit nombre d'occurrences de ces variantes, il n'est pas possible de tirer des conclusions sur cette alternance. Il faudrait donc examiner son emploi dans d'autres contextes et situations numériques.

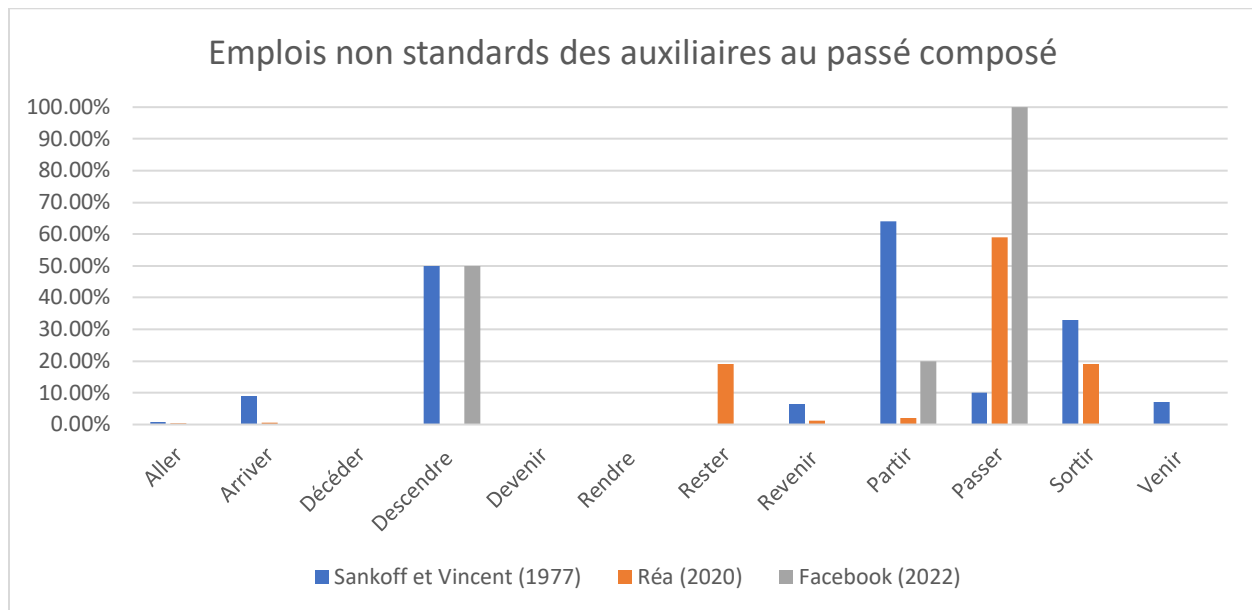
#### **4.3.2.2.3 Comparaison aux fréquences d'emploi attestées à l'oral**

Le Tableau 4.3.2.2.3 présente les fréquences d'emploi relatives non standards, selon les résultats de recherches antérieures, qui ont entrepris des analyses statistiques de l'emploi variable des auxiliaires *avoir* et *être* au passé composé dans des contextes oraux informels, et reprend les résultats de mon analyse. La Figure 4.3.2.2.3 est une représentation visuelle des données du Tableau 4.3.2.2.3.

**Tableau 4.3.2.2.3 - Comparaison des fréquences d'emploi de l'auxiliaire *avoir* avec les verbes intransitifs à l'oral et sur Facebook**

<b>Verbe</b>	<b>Sankoff et Vincent (1977)</b>	<b>Réa (2020)</b>	<b>Facebook (2022)</b>
Aller	0,7 %	0,4 %	0 %
Arriver	9 %	0,6 %	0 %
Décéder	N/A	N/A	0 %
Descendre	50 %	32,4 %	50 %
Devenir	N/A	N/A	0 %
Rendre	N/A	N/A	0 %
Rester	N/A	19,4 %	0 %
Revenir	6,5 %	1,3 %	0 %
Partir	64 %	2,1 %	20 %
Passer	10 %	59 %	100 %
Sortir	33 %	19 %	0 %
Venir	7 %	0 %	0 %

**Figure 4.3.2.2.3 - Comparaison des fréquences d'emploi de l'auxiliaire *avoir* avec les verbes intransitifs à l'oral et sur Facebook**



Les données du Tableau et de la Figure 4.3.2.2.3 permettent d'observer la progression dans le temps de l'emploi non standard de l'auxiliaire *avoir* avec les verbes intransitifs. Plus précisément, à l'oral, de 1977 à 2020, on constate une chute générale dans les fréquences d'emploi de l'auxiliaire *avoir* avec des verbes intransitifs. La seule instance qui a augmenté à l'oral, c'est l'emploi de l'auxiliaire *avoir* avec le verbe *passer*, montant d'environ 50 % entre 1977 et 2020. Les écrits sur Facebook suivent cette tendance communautaire : à part le verbe *passer*, la fréquence d'emploi de l'auxiliaire *avoir* avec des verbes intransitifs est négligeable.

On notera d'autre part, qu'avec le futur, les internautes préfèrent la forme privilégiée à l'oral et dans les SMS, alors qu'au passé composé ils et elles suivent de près le standard écrit, sauf pour le verbe *passer*. Les résultats des analyses de concordance des auxiliaires *avoir* et *être* au passé composé permettent alors uniquement d'attester la présence de cette variable dans un corpus d'écrits numériques sur Facebook par internautes au Québec. Le petit nombre d'occurrences non standards indiquerait que pour cette variable, les internautes tendent à suivre la règle verbale. Un corpus plus large, et une étude sur d'autres plateformes et d'autres contextes socionumériques plus propices à l'emploi du passé, pourraient éclaircir cette question.

Il importe aussi de préciser que la taille de mes données au passé composé, comparativement aux données des études antérieures de Sankoff et Vincent (1977) et Réa (2020), est assez petite ; il serait alors difficile de tirer des conclusions. Il se pourrait que le contexte communicatif sur Facebook n'incite pas à l'emploi du passé composé, puisque, vraisemblablement, les internautes s'engagent dans des discussions portant sur les sujets des publications, c'est-à-dire des actualités dans le présent. On constate d'ailleurs que dans le système verbal, l'emploi du futur, notamment 374 occurrences, par les internautes au Québec sur Facebook est plus légèrement fréquent que l'emploi du passé composé, soit 340 occurrences.

En somme, pour ce qui est des marqueurs verbaux à l'étude, tandis que l'emploi du futur sur Facebook suit exactement les tendances retrouvées dans les SMS, et ressemble aux tendances à l'oral, l'emploi variable des auxiliaires *avoir* et *être* est une question nuancée, qui mérite plus d'attention, surtout étant donné le petit nombre d'occurrences de verbes intransitifs au passé composé dans le corpus Facebook. Plus de recherches à partir d'un plus grand corpus sont alors nécessaires pour comprendre les tendances dans les emplois des marqueurs verbaux sur Facebook.

#### **4.3.3 Analyse des marqueurs syntaxiques**

Les chercheurs en sociolinguistique se penchent depuis longtemps sur les deux alternances syntaxiques à l'étude, qui sont fréquentes dans plusieurs variétés du français (Ashby, 1981 et 2001 ; Auger et Villeneuve, 2021 ; Barbarie, 1982 ; Bigot, 2008 ; Burnett et Tremblay, 2015 ; Corblin et Tovenà, 2003 ; Coveney, 1997 ; Halvorsen, 2021 ; Hansen et Malderez, 2000 ; Larrivé et Guryev, 2021 ; Meisner, 2017 ; Poplack et St-Amant, 2009). Dans ce corpus, l'emploi variable de la particule *ne* de la négation et l'expression variable de l'interrogation directe sont les variables les plus fréquentes, en particulier celles de la négation. En tout, le corpus contient 1046 occurrences des deux variables, dont 841 sont des formules de négation, et 205 de l'interrogation directe. Dans un premier temps, il s'agira de présenter les résultats portant sur la négation, suivie de celles portant sur les formules interrogatives.

#### 4.3.3.1 Analyse de l'emploi variable de la particule *ne* dans la négation

Les marqueurs de négation faisant l'objet d'analyse ici sont tous des occurrences non figées de la négation avec ou sans la particule *ne*, incluant tous les adverbes qui suivent : les formules en (*ne*) ... *pas*, la négation partielle en (*ne*)...*jamais*, (*ne*)...*rien*, et ainsi de suite. La négation a évolué au fil du temps : alors qu'au passé, c'était la particule *ne* qui exprimait la négation, cette particule n'est produite maintenant que rarement à l'oral, et c'est plutôt la particule *pas* (et les autres adverbes) qui sont venus s'implanter (Mougeon et Rehner, 2013). Tandis que les formules où la particule *ne* est conservée relèvent plutôt du standard et des contextes formels ou écrits, l'absence du *ne* serait caractéristique de presque toutes les variétés du parler informel à travers la francophonie.

En effet, Poplack et St-Amand (2007) considèrent l'emploi du *ne* dans la négation comme étant un marqueur de formalité. Qui plus est, étant donné la chute de l'emploi du *ne* à l'oral, les chercheurs tendent actuellement à analyser plutôt les contextes où le *ne* est conservé dans des contextes dits productifs (Sankoff et Vincent, 1977), spécialisés (Poplack et St-Amand, 2007) ou emphatiques (Fonseca-Greber, 2007 ; van Compernelle, 2009) et qui auraient des fonctions pragmatiques ou stylistiques en français moderne (van Compernelle, 2010). Comment ces tendances et fonctions se traduisent-elles dans la CNE ?

Suivant ces recherches antérieures, j'ai donc cherché toutes les instances de négation dans le corpus Facebook : le Tableau 4.3.3.1a présente plusieurs exemples de négation totale et partielle tirés directement du corpus.

**Tableau 4.3.3.1a - Exemples de l'emploi de la négation dans le corpus**

<b>Variante</b>	<b>Exemple</b>	<b>Page</b>
<i>ne...aucun</i>	la plupart du monde autour d'eux le font et montre <b>aucun</b> respect envers ceux qui continuent à manger de la viande. (1676)	<i>TVA Nouvelles</i>
<i>ne...aucunement</i>	<b>C'est aucunement</b> en lien avec le fait d'être vegan...(1700)	<i>TVA Nouvelles</i>
<i>ne...jamais</i>	Il <b>ne</b> faut <b>jamais</b> croire un avocat...(1667)	<i>TVA Nouvelles</i>
<i>Ne...nulle part</i>	On s'en va <b>nulle part</b> y a plus assez d'avions (2351)	<i>La Presse</i>
<i>ne...pas</i>	Ça <b>n'a pas</b> d'allure les salaires qu'ils se paient. 😊 😊 😊 (2097)	<i>La Presse</i>
<i>ne...personne</i>	...déjà il <b>n'y a personne</b> qui gouverne au fédérale. (2289)	<i>La Presse</i>
<i>ne...plus</i>	Même les canards <b>n'y allaient plus</b> ... (1380)	<i>Journal de Québec</i>
<i>ne...rien</i>	y'a <b>rien</b> de gratuit au Québec... (1169)	<i>Journal de Montréal</i>

Il importe de noter que lors des analyses de concordance, j'ai cherché les graphies standards, autant que non standards, de chacune des formules possibles, telles qu'indiquées dans le Tableau 4.3.3.1b. J'ai cherché la particule *ne*, et ensuite les adverbes qui peuvent former une négation, en m'assurant d'éliminer les doublages. Notons que je n'ai trouvé aucune occurrence de *ne...guère*, *ne...goutte*, ni *ne...point* (peu importe la présence ou l'absence de la particule *ne*).

**Tableau 4.3.3.1b - Graphies cherchées de la négation**

<b>Variante</b>	<b>Graphies cherchées</b>
<i>ne...aucun(e), aucunement</i>	<i>Aucun, aucune</i>
<i>ne...guère</i>	<i>Guère, guere</i>
<i>ne...goutte</i>	<i>goutte</i>
<i>ne...jamais</i>	<i>Jamais, jms</i>
<i>Ne...nulle part</i>	<i>Nullepart, nulle part</i>
<i>ne...pas</i>	<i>pas</i>
<i>ne...personne</i>	<i>personne</i>
<i>ne...plus</i>	<i>Plus, pu, pus</i>
<i>ne...point</i>	<i>point</i>
<i>ne...rien</i>	<i>rien</i>

#### **4.3.3.1.1 Analyse des données de l'emploi variable de la négation**

Tout d'abord, une recherche générale avec la fonction « Words » dans #LancsBox pour les fréquences des mots *ne* et *pas* révèle que *pas* est le cinquième mot le plus fréquent dans le corpus, avec 863 occurrences pour 10 000 mots, tandis que *ne* est le 24<sup>e</sup> mot le plus fréquent avec 273 occurrences pour 10 000 mots. Cela montre la prépondérance globale de ces deux termes dans le corpus.

Pour ce qui est de mes analyses, le corpus contient 841 occurrences de négation, après avoir exclu les expressions figées ou dans lesquelles le *ne...pas* est catégorique, telles que *n'est-ce pas*. Par exemple, dans le cas de la formule *ne...* sans adverbe (p.ex. [...] *on ne peut être sur la même longueur d'onde* (1483)), l'emploi de la particule *ne* est catégorique ; il s'agit alors d'un contexte

sans variation, qui a été exclue des analyses statistiques. En revanche, il faut aussi s'interroger sur les 18 instances, provenant de 15 internautes distincts, de cette formule très soutenue, pareillement à la formule de contrainte *ne...que*, qui était marginale dans les études antérieures, mais apparaît 13 fois dans ce corpus d'écrits numériques.

Ensuite, il importe de noter qu'environ 10 % des occurrences de la négation repérées dans le corpus représentent un fragment de phrase, c'est-à-dire une phrase où il manque un élément clé ; souvent, dans ces cas, il manque un verbe, et/ou la phrase commence par *pas*. Apparaissent aussi des occurrences de fragments avec d'autres adverbes qui imitent un style plutôt télégraphique, comme s'il s'agissait d'interjections.

*NON, pas assez de protéines ... (1663, TVA Nouvelles)*

*En plus, pas de Noël, pas de rassemblement. (2513, La Presse)*

*Pas capable je sais pas comment français fait (1797, TVA Nouvelles)*

*aucun sens logique me semble (2100, La Presse)*

*Pas capable de rien faire de bon. RIEN. (1148, La Presse)*

*bien raison. Bitcoins, jamais de riche moi. (2102, La Presse)*

*JAMAIS (1624, Le Devoir)*

*Oui, pouf, pu là! 🌀 (2548, La Presse)*

Dans ces fragments, je n'ai trouvé aucune occurrence de la négation dans laquelle la particule *ne* est employée, c'est-à-dire que dans les fragments écrits par les internautes au Québec sur Facebook, le *ne* est omis à un taux de 100 %. Par conséquent, ces occurrences ont également été exclu des analyses statistiques.

Notons que pour les occurrences conservées dans l'analyse, j'ai constaté plusieurs graphies non standards, surtout avec les adverbes *plus*, et *rien*, mais aussi avec *pas* ; quelques exemples sont présentés ci-dessous :

*Pas de danger que LEGO fasse ça ici mais **cpas** grave il va gagner quand même*  
(2284, *La Presse*)

*fait plus d'un mois qui **pus** de taxe en alberta et le gas a pas amonter*  
(1227, *La Presse*)

*viens pas chialer après quand on aura **pu** de service dans rien*  
(2467, *Journal de Québec*)

*les municipales non **riens** fait non aux lieux de percevoir un taxe*  
(1875, *TVA Nouvelles*)

La formule la plus fréquente est le *(ne)...pas*, qui apparaît 778 fois sur 841, ensuite *(ne)...plus* apparaît 59 fois, *(ne)...rien* 58 fois, *(ne)...jamais* 40 fois, *(ne)...aucun(e)* 35 fois, *ne...* sans adverbe 18 fois, *(ne)...personne* six fois, et *(ne)...aucunement* quatre fois et enfin *(ne)...nulle part* une fois. La majorité des occurrences proviennent des pages de *La Presse* et *TVA Nouvelles*, qui sont les deux plus grandes composantes du corpus, car seulement 278 des 778 occurrences de la négation proviennent des pages *Le Devoir*, *Journal de Montréal* et *Journal de Québec*.

#### **4.3.3.1.2 Analyse quantitative/statistique descriptive des données de l'emploi variable de la négation**

Le Tableau 4.3.3.1.2 présente les occurrences et les fréquences d'emploi relatives de l'alternance entre l'emploi et l'omission de la particule *ne* dans les formules négatives.

**Tableau 4.3.3.1.2 - Analyse quantitative/statistique descriptive des données de l'emploi variable de la négation, fragments exclus**

Variante	Nombre d'occurrences	Fréquence d'emploi relative
Emploi du <i>ne</i>	447	53 %
Omission du <i>ne</i>	394	47 %
TOTAL	841	100 %

Après l'exclusion des occurrences sous forme de fragments ainsi que les formes en *ne sans pas* et en *n'est-ce pas*, on constate que la particule *ne* est employée légèrement plus fréquemment qu'elle n'est omise, étant employée dans 447 ou 53 % des cas, et omise dans 394 ou 47 % des cas.

En ce qui concerne l'environnement linguistique, l'instance dans laquelle le *ne* est employé beaucoup plus fréquemment qu'il n'est omis, est celle de la formule *ne...plus* – ici, le *ne* est employé dans 40 cas sur 59, soit 68 % des cas. On se demande pourquoi la particule *ne* serait employée plus souvent avec l'adverbe *plus* qu'avec les autres adverbes. Étant donné le manque de contexte et de la voix dans les situations communicatives en ligne, il se peut que ces individus, en employant la particule *ne* suivie de l'adverbe *plus*, veulent empêcher que *plus* soit confondu avec le sens d'ajout que ce mot peut transmettre, qui est communiqué à l'oral par la prononciation, ou non, du son [s] en final.

#### **4.3.3.1.3 Comparaison aux fréquences d'emploi attestées à l'oral et dans le (t)chat**

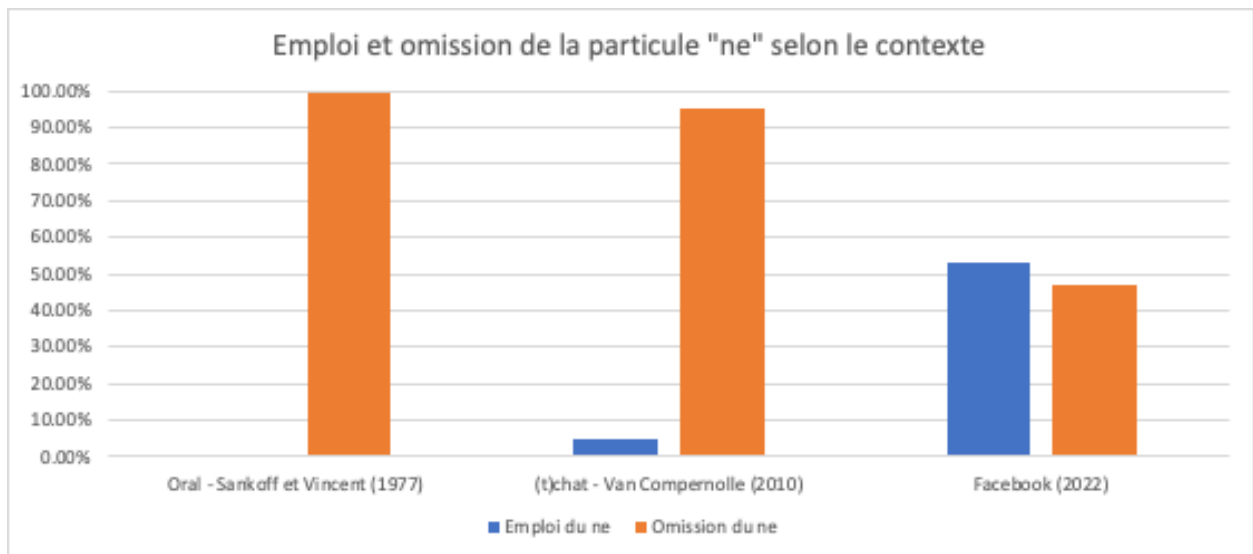
Le Tableau 4.3.3.1.3 présente les fréquences de l'emploi variable de la négation attestées à l'oral et dans le (t)chat, comparativement aux fréquences trouvées sur Facebook. La Figure 4.3.3.1.3 présente ensuite ces données de manière visuelle.

**Tableau 4.3.3.1.3 - Comparaison des fréquences d'emploi attestées à l'oral, dans le (t)chat, et sur Facebook**

Variante	Oral - Sankoff et Vincent (1977)	(t)chat - Van Compernelle (2010)	Facebook (2022)
Emploi du <i>ne</i>	0,5 %	5,1 %	53 %
Omission du <i>ne</i>	99,5 %	94,9 %	47 %

On constate que les internautes sur Facebook situés au Québec alternent entre l'emploi et l'omission de la particule *ne* à une proportion presque égale, ce qui va à l'encontre du standard écrit, mais également des tendances attestées à l'oral et dans le (t)chat.

**Figure 4.3.3.1.3 - Comparaison des fréquences d'emploi attestées à l'oral et dans le (t)chat, et sur Facebook**



Comparativement à ce que l'on constate à l'oral et dans le (t)chat, l'emploi variable de la négation est plus partagé sur Facebook ; les internautes ont donc peut-être plus tendance à suivre le standard écrit. On constate en effet que la particule *ne* est employée beaucoup plus fréquemment sur Facebook qu'à l'oral et que dans le (t)chat. En même temps, les résultats suggèrent une influence importante de l'emploi non standard ou vernaculaire, puisque la particule *ne* est omise presque aussi souvent sur Facebook qu'elle est employée. La différence des taux d'emplois dans le (t)chat et sur Facebook suggère aussi que les variétés et les registres diffèrent selon la plateforme, le

contexte, ou la modalité. Il serait important à l'avenir d'étudier les raisons que les internautes choisissent d'employer ou d'omettre cette particule, par le biais de recherches plus ethnographiques permettant par exemple un dialogue avec l'internaute pour l'interroger sur ses choix linguistiques.

#### **4.3.3.2 Analyse de l'emploi variable des formules interrogatives**

Les formules interrogatives directes à l'étude sont celles avec *Est-ce que*, avec inversion du sujet et du verbe, sans inversion, et avec le suffixe *-tu*, qui est une tournure exclusivement québécoise. Ce sont donc les formules posant une question fermée, qui demandent une réponse oui/non. D'après Yves Barbarie (1982 : 145), « l'énoncé interrogatif offre une variété de formes; certaines de ces formes s'écartent fortement de la norme prescrite » et d'autres « sont rarement rencontrées dans la langue parlée ». Plus précisément, d'après Stefan Halvorsen, tandis que les grammairiens indiquent que « l'interrogation totale peut être exprimée [...] par inversion [...] par intonation et [...] par *est-ce que* », en français québécois « nous trouvons une construction supplémentaire – la construction qui comporte la particule *-tu*. Cette structure interrogative peut être considérée à la fois archaïque et innovatrice » (2021 : 20-23). Notons que je n'ai pas cherché les formules interrogatives formées par *comment*, *quand*, *pourquoi* etc., ni les questions indirectes comme *Je voulais vous demander si ...*, afin de simplifier l'analyse et la comparaison.

Suivant ces propos, j'ai alors cherché toutes les occurrences de l'interrogation directe dans le corpus Facebook : le Tableau 4.3.3.2 illustre des exemples de chaque formule, tirés directement du corpus.

**Tableau 4.3.3.2 - Exemples de l'emploi des formules interrogatives dans le corpus**

Variante	Exemple	Page
<i>Est-ce que</i>	<b>Est-ce que</b> à cette époque les Whoppers étaient faits de viande de mammouth ? (1944)	<i>TVA Nouvelles</i>
Inversion	<b>SOMMES NOUS RENDU DES ROBOTS ?</b> (1502)	<i>Le Devoir</i>
Sans inversion	Vous êtes bien informé vous ? ? (1844)	<i>TVA Nouvelles</i>
Suffixe <i>-tu</i>	coudons vos journaliste dorment <b>tu</b> au gaz à R.C. (1483)	<i>Le Devoir</i>

Il importe de noter que pour les formules interrogatives, j'ai pu faire une recherche dans #LancsBox pour quelques-unes des formules en *est-ce que*, et en « *-tu* » et d'inversion, en cherchant le trait d'union entouré d'étoiles \*-\*, tenant toujours compte de graphies standards et non standards. Mais puisqu'il n'est pas possible de faire une recherche pour le symbole « ? » dans ce logiciel, il a fallu plusieurs parcours manuels du corpus afin de repérer toutes les occurrences de l'interrogation. Il se peut donc que quelques occurrences aient été manquées, mais j'ai atténué cette limite en faisant trois parcours du corpus de manière manuelle et par le biais de #Lancsbox.

#### 4.3.3.2.1 Analyse des données de l'emploi variable des formules interrogatives

En tout, le corpus contient seulement 205 occurrences d'interrogation directe. La formule sans inversion est de loin la plus fréquente, avec 156 occurrences, et ce dans chacune des cinq composantes du corpus. Quant aux trois autres formules, les questions formulées par inversion apparaissent 25 fois, les questions avec le suffixe *-tu* apparaissent 13 fois, dans les deux cas, dans toutes les composantes du corpus sauf la page du *Journal de Montréal*, et enfin celles en *est-ce que* apparaissent 11 fois, et ce dans les cinq composantes du corpus. En général, les formules interrogatives directes sont donc plus ou moins réparties à travers les cinq composantes du corpus. J'ai constaté une panoplie de différentes graphies non standards comme le montrent les exemples ci-dessous :

*y'a pt des raison du pk sa ne lui convien pas le secteur qui lui on donner proposer est ce que tu le connaît tu connaît son histoire ect ect pense pas (1834, TVA Nouvelles)*

*... comprend tu ? (2120, La Presse)*

*Stu moi qui es weird ou une decennie....c en 2012 ca....cest loin du vintage d'habitude. (2015, TVA Nouvelles)*

*Plusieurs ont quitte depuis quelques temps et c est maintenant qu ils s inquietent ? ? (1515, Le Devoir)*

#### **4.3.3.2 Analyse quantitative/statistique descriptive des données de l'emploi variable des formules interrogatives**

Le Tableau 4.3.3.2.2 présente les occurrences et les fréquences d'emploi relatives des formules d'interrogation directe.

**Tableau 4.3.3.2.2 - Analyse quantitative/statistique descriptive des données de l'emploi variable des formules interrogatives**

<b>Variante</b>	<b>Nombre d'occurrences</b>	<b>Fréquence d'emploi relative</b>
<i>Est-ce que</i>	11	5,4 %
Inversion	25	12,1 %
Sans inversion	156	76 %
Suffixe <i>-tu</i>	13	6,3 %
TOTAL	205	100 %

On constate tout d'abord que la formule *est-ce que* est la moins fréquente, avec une fréquence d'emploi relative de 5,4 %, et celle en *-tu* est légèrement plus fréquente, à 6,3 %. Ensuite, les

formules avec inversion représentent 12,1 % des cas, tandis que les formules sans inversion sont les plus fréquentes, représentant 76 % des cas.

Pour ce qui est de l'environnement linguistique, la grande majorité des formules interrogatives directes sont sous forme de phrase positive, seulement 34 occurrences ou 16 % des cas sont sous forme de phrase négative. Ensuite, 38 occurrences, ou 18 % des cas, sont sous forme de fragments de phrase, c'est-à-dire sans verbe, pour la plupart. Dans tous les cas, il s'agit de formules sans inversion – en voici deux exemples :

*Les chiffre svp ? ? (2098, La Presse)*

*et les camionneurs, les gros transporteurs, les autobus, les tracteurs ? ? ? ? (2164, La Presse)*

Ces fragments, ainsi qu'un bon nombre de cas sans inversion, écrits sous forme de question directe, représentent une reprise d'information de l'article ou d'un commentaire par un autre internaute. En particulier, un bon nombre de ces occurrences sans inversion expriment de la surprise ou de l'incrédulité, indiquée par un ou plusieurs points d'interrogation en fin d'énoncé.

Enfin, se focalisant sur l'emploi de la ponctuation, la formule interrogative en suffixe *-tu* entraîne moins souvent le point d'interrogation que les trois autres formules interrogatives : ce symbole est utilisé dans toutes les occurrences de l'interrogation sans inversion (100 %), dans 23 des 25 occurrences (92 %) avec inversion, dans 10 des 11 occurrences (91 %) en *est-ce que*, mais dans seulement 9 des 13 occurrences (69 %) de l'interrogation exprimées avec le suffixe *-tu*. Ces résultats suggèrent que le suffixe *-tu* en soi serait un marqueur interrogatif, et que le point d'interrogation devient facultatif. De plus, la seule différence dans la CNE entre une phrase interrogative sans inversion et une phrase déclarative est effectivement le point d'interrogation. Cette observation est sans doute le résultat de l'adaptation par l'internaute au contexte, qui ne permet pas l'utilisation de la prosodie pour communiquer qu'il s'agit d'une question. Pour mieux cerner ces différences entre CNE sur Facebook et d'autres situations de communication, passons à la comparaison des résultats de la présente étude à ceux dans d'autres contextes.

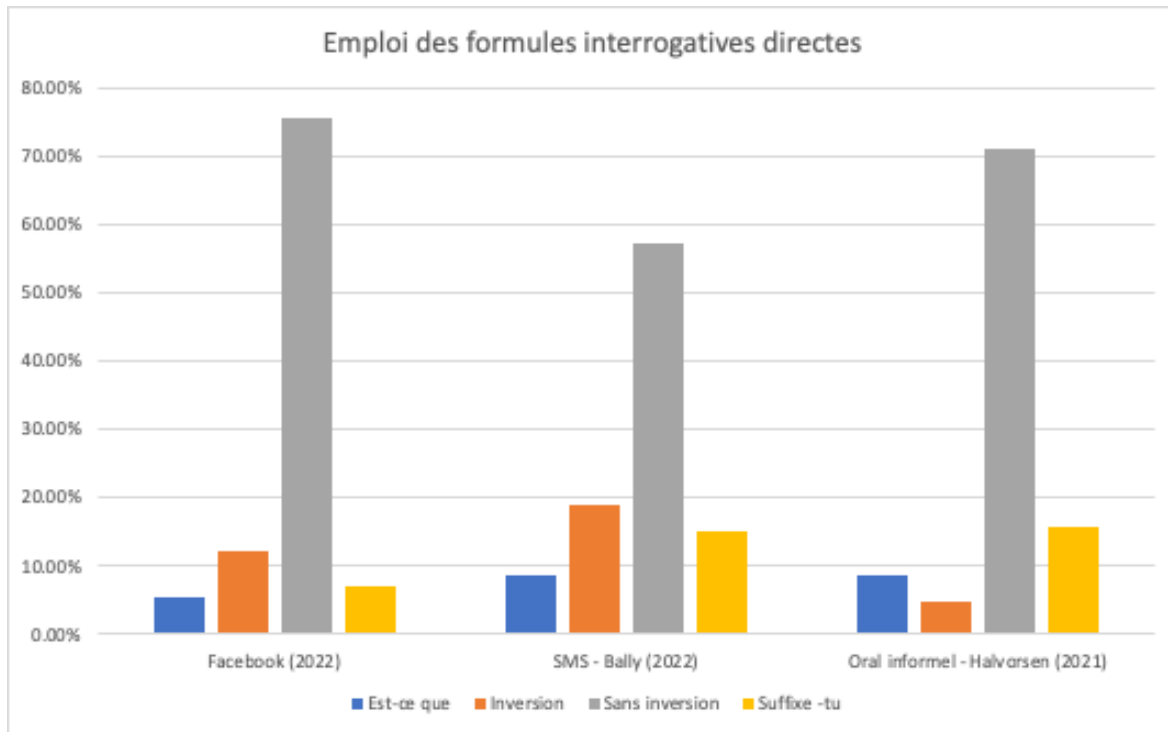
#### 4.3.3.2.3 Comparaison aux fréquences d'emploi des formules interrogatives attestées à l'oral

Le Tableau 4.3.3.2.3 présente les résultats de recherches antérieures, qui ont entrepris des d'analyses statistiques de l'emploi des formules interrogatives directes dans des contextes oraux formels et informels, et reprend les résultats de mon analyse. Ensuite la Figure 4.3.3.2.3 montre une représentation visuelle des données du Tableau 4.3.3.2.3.

**Tableau 4.3.3.2.3 - Comparaison des fréquences d'emploi des formules interrogatives à l'oral, dans les SMS et sur Facebook**

<b>Variante</b>	<b>Oral informel - Halvorsen (2021)</b>	<b>SMS - Bally (2022)</b>	<b>Facebook (2022)</b>
<i>Est-ce que</i>	8,5 %	8,5 %	5,3 %
Inversion	4,8 %	19,1 %	12,1 %
Sans inversion	71,1 %	57,3 %	76 %
Suffixe <i>-tu</i>	15,6 %	15,1 %	6,3 %

**Figure 4.3.3.2.3 - Comparaison des fréquences d'emploi des formules interrogatives à l'oral, dans les SMS et sur Facebook**



D'après le Tableau et la Figure 4.2.3.2.3 on constate tout d'abord que les fréquences sur Facebook s'alignent plus ou moins sur les fréquences récentes attestées dans des contextes de l'oral informel (Halvorsen, 2021) et dans les SMS (Bally, 2022). La forme avec inversion est légèrement plus fréquente dans les SMS qu'à l'oral informel et sur Facebook, et la forme en *-tu* est légèrement moins fréquente sur Facebook qu'à l'oral et dans les SMS. Alors que cette formule en suffixe *-tu* est la deuxième plus fréquente à l'oral, elle représente seulement environ 16 % des occurrences dans ce contexte, ce qui surprend étant donné qu'il s'agit d'une « forme essentiellement orale » (Bally, 2022 : 13) et plutôt emblématique du français québécois. Toutefois, les fréquences d'emploi des formules interrogatives directes sur Facebook semblent suivre de manière générale les tendances communautaires, que ce soit à l'oral ou dans un contexte CNE. Peut-on alors conclure que cette variable syntaxique est plus ou moins stable, peu importe le contexte ? Il serait pertinent aussi de faire une analyse de la progression de l'emploi de ces formules dans le temps, en s'appuyant sur les recherches antérieures dans différents contextes à l'oral (Bigot, 2011 ; Lyster, 1996 ; Villeneuve, 2017).

Ce qui surprend, c'est la prépondérance des formes sans inversion sur Facebook, étant donné que le contexte est à l'écrit, puisqu'à l'oral l'interrogation sans inversion serait normalement exprimée par l'intonation de la voix. Au Chapitre 5, mes analyses de collocation examinent la fréquence à laquelle les formules interrogatives sont accompagnées d'autres éléments non standards, entre autres, ceux relevant de l'oral, afin de mieux comprendre ce phénomène de mixité de styles langagiers.

En somme, pour ce qui est des marqueurs syntaxiques, les résultats sont partagés. D'une part, les fréquences d'emploi des formules interrogatives sur Facebook suivent les tendances attestées dans des contextes informels à l'oral, et celles attestées dans les SMS. En revanche, les fréquences d'emploi de la négation avec ou sans la particule *ne*, ne s'alignent ni sur les fréquences à l'oral, ni sur celles du (t)chat. Tout de même, on a vu que la particule *ne* est employée presque aussi fréquemment qu'elle est omise par les internautes au Québec sur Facebook, ce qui suggère une concurrence entre le standard et le non standard. Une analyse plus pointue, se focalisant sur les facteurs internes et l'environnement linguistique pourrait examiner en plus de profondeur ces contextes dans lesquels la particule *ne* est conservée et omise sur Facebook, afin de comprendre les choix linguistiques des internautes.

#### **4.3.4 Analyse des marqueurs pronominaux**

Cette prochaine section présente les résultats portant sur l'analyse des marqueurs pronominaux dans le corpus, notamment tout d'abord l'alternance entre les pronoms *on* et *nous* et ensuite l'emploi variable des pronoms simples et composés non clitiques. Les études sur ces alternances dans les variétés de français parlées au Canada (Blondeau, 2000, 2001 et 2011 ; Blondeau *et al.*, 2002 ; Deshaies, 1991 ; Laberge, 1977) et ailleurs dans la francophonie (Boutet, 1986 ; Coveney, 2000 ; Grafström, 1969), s'accumulent. Les marqueurs pronominaux à l'étude ici étaient parmi les plus fréquents dans le corpus, comparativement aux éléments verbaux et lexicaux. En tout, le corpus contient 728 occurrences des deux groupes de marqueurs pronominaux ; de ce chiffre, 399 sont de l'alternance *on-nous*, et 329 des pronoms simples et composés au pluriel.

#### 4.3.4.1 Analyse de l'emploi variable des pronoms *on* et *nous*

Les premiers marqueurs pronominaux à l'étude sont *on* et *nous*. Dans les deux cas, le pronom représente la première personne du pluriel. Traditionnellement, l'emploi du pronom *on* à la place du *nous* était considéré comme appartenant « à la langue non standard de niveau populaire », mais, l'emploi du *on* est devenu plus courant, « à tel point que l'usage de *nous* est de plus en plus réservé à la langue écrite ou aux registres les plus soignés de l'oral » (Thomas, 2015 : 126). En particulier, Laberge (1977) précise que le pronom *nous* est devenu une variante de prestige, employée par les locuteurs pour des buts stylistiques (van Compernelle, 2008c). Les études du français parlé au Québec montrent alors que le pronom *on* est de loin le plus fréquent, représentant environ 98 % des occurrences (Blondeau *et al.*, 2014).

Suivant ces propos, j'ai alors cherché toutes les occurrences des pronoms *on* et *nous* dans le corpus Facebook : le Tableau 4.3.4.1a illustre des exemples de chaque emploi, tirés directement du corpus.

**Tableau 4.3.4.1a - Exemples de l'emploi des pronoms *on* et *nous* dans le corpus**

Variante	Exemple	Page
<i>On</i>	<b>On</b> sait vivre. Et <b>on</b> sait voir la différence. (2119)	<i>La Presse</i>
<i>Nous</i>	Je ne peux pas croire que <b>nous</b> en sommes rendus là! (2128)	<i>La Presse</i>

Il importe de noter que lors des analyses de concordance, j'ai cherché les graphies standards, autant que non standards. Dans ce cas, l'ajout du symbole \* avant et après le mot a permis de trouver les occurrences avec trait d'union ou apostrophe, par exemple *veut-on* et *qu'on*. Le Tableau 4.3.4.1b montre les graphies recherchées pour chaque variante.

**Tableau 4.3.4.1b - Graphies recherchées des pronoms *on* et *nous***

Variante	Graphies recherchées
<i>On</i>	On, ont, ons, *on*
<i>Nous</i>	Nous, nou, ns, *nous*

#### 4.3.4.1.1 Analyse des données de l'emploi variable des pronoms *on* et *nous*

En tout, le corpus contient 399 occurrences des pronoms sujets *on* et *nous*, et la variante de loin la plus fréquente est *on*, avec 352 occurrences, et ce dans chacune des composantes du corpus. Le pronom *nous* apparaît alors seulement 47 fois comme pronom sujet ; il n'y a aucune occurrence provenant de la composante *Journal de Montréal*, et une instance seulement de la page du *Journal de Québec*. Notons que j'ai constaté un bon nombre de graphies non standards, 34 cas ou environ 10 % des cas, et ce seulement dans l'emploi du pronom *on*, qui est souvent écrit *ont* dans le corpus.

*CHEZ PAS!!!! ? ? ? ? Ont est pas surpris pantoute (2105, La Presse)*

[Je ne sais pas! ? On n'est pas du tout surpris.]

*Ont sais toute que dieu existe pas (1927, TVA Nouvelles)*

[On sait tous que dieu n'existe pas.]

#### 4.3.4.1.2 Analyse quantitative/statistique descriptive des données sur l'emploi variable des pronoms *on* et *nous*

Le Tableau 4.3.4.1.2 présente les occurrences et les fréquences d'emploi relatives de l'alternance des pronoms *on* et *nous*.

**Tableau 4.3.4.1.2 - Analyse quantitative/statistique descriptive des données sur l'emploi variable des pronoms *on* et *nous***

Variante	Nombre d'occurrences	Fréquence d'emploi relative
<i>On</i>	352	88 %
<i>Nous</i>	47	12 %
TOTAL	399	100 %

Les résultats montrent que le pronom *on* apparaît dans 88 % des occurrences de cette alternance, tandis que *nous* est utilisé 12 % du temps seulement. L'emploi le plus fréquent de *on* pourrait s'expliquer par le principe de l'économie du langage. C'est la variante qui serait la plus rapide et la plus simple à dactylographier, comparativement à *nous*. De même pour la forme verbale correspondante, la déclinaison du pronom *on* est plus simple que celle du *nous*.

En effet, j'ai exclu des analyses 13 instances où le pronom *nous* est soit sans verbe accordé, soit suivi par *on*, soit suivi par un autre groupe nominal à la troisième personne du pluriel avec un verbe accordé à ce deuxième groupe nominal, et non au *nous* :

*Votre population est à 4 millions et nous 10 millions. (1789, TVA Nouvelles)*

*Et nous on avait payé pour céder la division ferroviaire à Alston... (1806, La Presse)*

*nous le peuples en ont ras le bolt.de vous promener gratuit (2288, La Presse)*

Cela suggère encore une préférence pour les formes plus simples chez les internautes au Québec sur Facebook. Ce phénomène ne se produit pas lorsque *on* est employé dans ce corpus ; c'est-à-dire que le pronom *on* est toujours accordé directement avec le verbe qui le suit, et il n'y a pas d'ajout de pronom ou de groupe nominal.

Ensuite, les pronoms *on* et *nous* ne se retrouvent pas à des taux proportionnels avec les mêmes temps verbaux ; ainsi, le pronom *on* est le seul à apparaître avec un verbe au subjonctif, notamment dans 10 occurrences, et un verbe au passé simple, notamment une occurrence ; le pronom *on* apparaît plus souvent que le pronom *nous* avec le futur proche et le conditionnel ; *on* et *nous* apparaissent à un taux similaire avec le futur simple et l'imparfait ; et le pronom *nous* apparaît plus souvent avec le passé composé comparativement au pronom *on*. Ces constatations suggèrent que ces deux pronoms ne se produisent pas nécessairement dans les mêmes environnements linguistiques. Encore plus surprenante est la coprésence du pronom *on*, qui relève plutôt des tendances à l'oral, avec le passé simple, qui relève de l'écrit littéraire. Une étude plus pointue

d'occurrences spécifiques de l'emploi de ces deux pronoms pourrait révéler l'apport de l'environnement linguistique dans le choix du pronom.

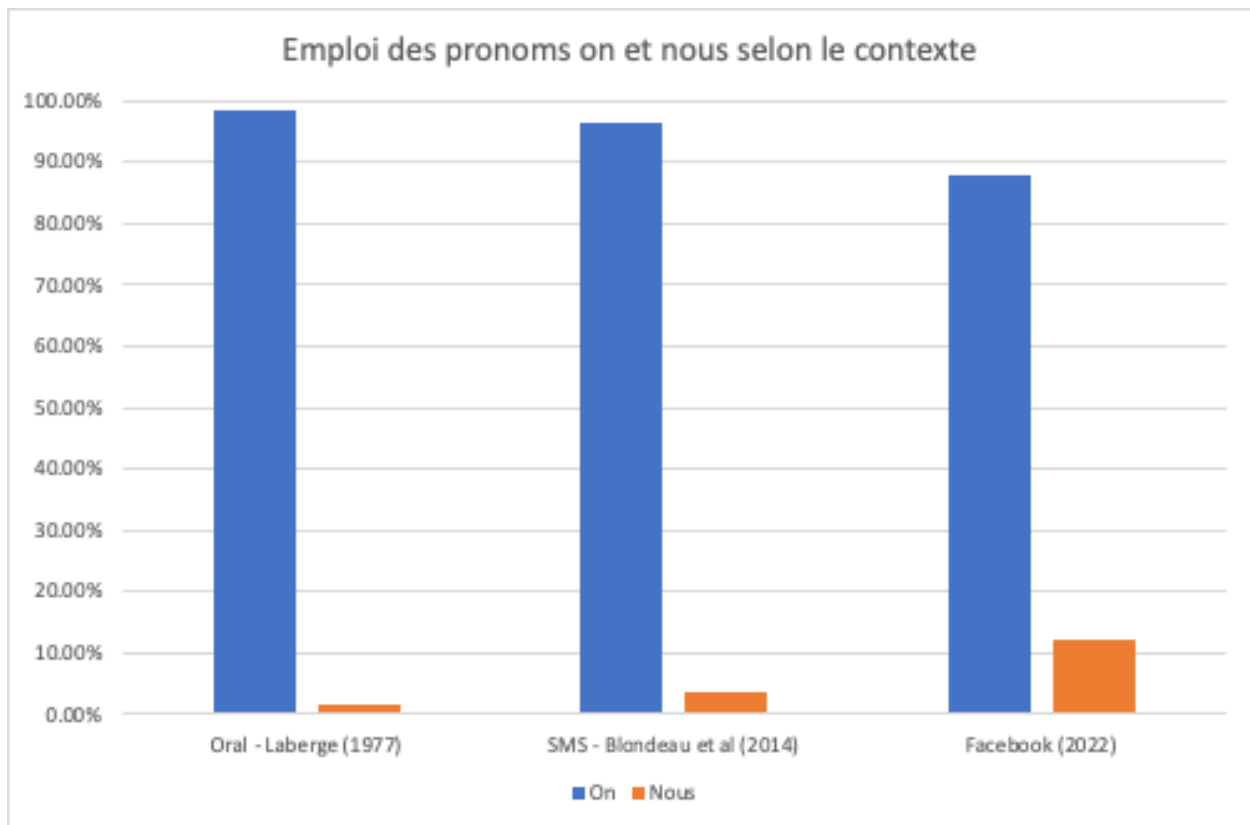
#### **4.3.4.1.3. Comparaison aux fréquences d'emploi attestées à l'oral et dans les SMS**

Le Tableau 4.3.4.1.3 présente les résultats de recherches antérieures, qui ont entrepris des analyses statistiques de l'alternance d'emploi des pronoms *on* et *nous* dans des contextes oraux et dans les SMS, et reprend les résultats de mon analyse. Ensuite la Figure 4.3.4.1.3 montre une représentation visuelle des données du Tableau 4.3.4.1.3.

**Tableau 4.3.4.1.3 - Comparaison des fréquences d'emploi des pronoms *on* et *nous* attestées à l'oral, dans les SMS et sur Facebook**

<b>Variante</b>	<b>Oral - Laberge (1977)</b>	<b>SMS - Blondeau <i>et al.</i> (2014)</b>	<b>Facebook (2022)</b>
<i>On</i>	98,4 %	96,3 %	88 %
<i>Nous</i>	1,6 %	3,7 %	12 %

**Figure 4.3.4.1.3 - Comparaison des fréquences d'emploi des pronoms *on* et *nous* attestées à l'oral, dans les SMS et sur Facebook**



On constate que les internautes sur Facebook semblent suivre les tendances communautaires, car le *on* est préféré. En même temps, le *nous* est employé sept fois plus fréquemment sur Facebook qu'à l'oral et trois fois plus fréquemment que dans les SMS, suggérant que l'influence du standard écrit joue un rôle légèrement plus prononcé sur Facebook, que dans les SMS. Comme c'est le cas pour la plupart des variables à l'étude, il semblerait que c'est la forme la plus simple qui prime, ce qui, dans ce cas, s'aligne sur la forme privilégiée à l'oral.

#### 4.3.4.2 Analyse de l'emploi variable des pronoms simples et composés non clitiques

Les pronoms simples et composés non clitiques à l'étude sont *nous*, *nous-autres*, *vous*, *vous-autres*, *eux* et *eux-autres*. Ce ne sont alors pas des pronoms sujets, mais des pronoms facultatifs. Bien que cette alternance soit « typique du français montréalais » (Blondeau, 2000 : 4), « les formes composées, anciennes formes par défaut, sont concurrencées par les formes simples dans certains contextes » à l'oral, et à l'écrit, les formes composées sont rares aujourd'hui (Blondeau *et al.*, 2014 : 146).

Suivant ces propos, j'ai alors cherché toutes les occurrences des pronoms non clitiques à la première, deuxième et troisième personne du pluriel : le Tableau 4.3.4.2a illustre des exemples tirés directement du corpus, sauf pour *eux-autres*, car il n'y avait aucune occurrence de cette variante.

**Tableau 4.3.4.2a - Exemples de l'emploi des pronoms pluriels non clitiques dans le corpus**

Variante	Exemple	Page
<i>Nous</i>	ben justement, il serait peut être temps que ça soit le gouvernement qui se serre la ceinture et non <b>nous</b> ! (2174)	<i>La Presse</i>
<i>Nous-autres</i>	Pis <b>nous autre</b> au Québec, on se laisse faire (2383)	<i>La Presse</i>
<i>Vous</i>	Aller vous faire piquer si vous voulez voyagé commencer vous à réaliser que c'est <b>vous</b> le problème (2486)	<i>La Presse</i>
<i>Vous-autres</i>	continuez de vous entretenir entre <b>vous autres</b> et haïr les ZanglA pi les ZamaricaiN (2706)	<i>La Presse</i>
<i>Eux</i>	...mais habituellement ce ne sont pas <b>eux</b> qui braillent (2400)	<i>La Presse</i>

Il importe de noter que lors des analyses de concordance, j'ai cherché les graphies standards, autant que non standards, souvent en utilisant le symbole \* qui permet de chercher toutes les occurrences peu importe ce qui suit le symbole \*. Par exemple, chercher « nous\* » permettrait de trouver toutes les occurrences de mots qui commencent par « nous », y compris, *nous-autres*, *nous autres*, tandis que la recherche de « \*nous » permet de trouver les occurrences où une lettre ou un symbole précède le mot *nous*, comme *allons-nous*. Notons aussi que j'ai exclu des analyses les formes

figées telles que *s'il vous plaît*. Le Tableau 4.3.4.2b montre les graphies recherchées pour chaque variante.

**Tableau 4.3.4.2b - Graphies recherchées des pronoms pluriels non clitiques**

Variante	Graphies recherchées
<i>Pronoms simples</i>	Nous*, *nous, ns, vous*, *vous, vs, *eux, eux*
<i>Pronoms composés</i>	*autre*

#### 4.3.4.2.1 Analyse des données des pronoms simples et composés non clitiques

En tout, le corpus contient 329 occurrences de l'alternance des pronoms simples et composés non clitiques aux première, deuxième et troisième personnes du pluriel. Les pronoms simples sont de loin les variantes les plus fréquentes. Le pronom *vous* est le plus fréquent, avec 205 occurrences, ensuite *nous* apparaît 76 fois, et *eux* apparaît 45 fois. Pour ce qui est des pronoms composés, on ne constate que trois occurrences dans tout le corpus : *nous-autres* apparaît deux fois, et *vous-autres* une fois. Le corpus ne contient aucune occurrence du pronom composé *eux-autres*. Il importe aussi de noter que les trois cas composés viennent tous de la composante *La Presse* du corpus, qui est d'une plus grande taille que les quatre autres composantes.

Il importe de noter que mes recherches ont soulevé quelques graphies non standards : dans le cas des pronoms composés, les internautes omettent le trait d'union, mettant seulement un espace, et pour ce qui est des pronoms simples, j'ai constaté six graphies non standards du pronom *vous* – la moitié provenant d'un même internaute – et deux graphies non standards du pronom *eux*, mais aucune divergence pour le pronom *nous* :

*Vs aviez RIEN COMPRIS. TOUT EST PLANIFIER JUSTEMENT. POUR VS  
 ECOEUREZ ...alors vs allez ACCEPTER LEUR SOLUTION=IDENTIFICATION  
 NUMÉRIQUE (2452, La Presse)*

*eu il sens colis c'est nous autres qui payes leurs déplacements (2288, La Presse)*

[Eux, ils s'en câlissent, c'est nous-autres qui payent leurs déplacements.]

Dans les deux cas, on peut se demander si ces troncations sont le résultat du principe de l'économie, ou si d'autres facteurs internes ou externes jouent un rôle dans le choix de graphie. L'autre fait intéressant est que dans le deuxième extrait ci-dessus, l'internaute alterne entre l'emploi de pronoms simples et composés : en début de phrase la personne choisit la forme simple, *eux*, à la place de *eux-autres*, et plus loin dans la phrase elle emploie *nous-autres*, à la place de *nous*. Ces genres d'alternances par individus seront discutées et exposées plus en détail dans le prochain chapitre.

#### **4.3.4.2.2 Analyse quantitative/statistique descriptive des données des pronoms simples et composés non clitiques**

Les Tableaux 4.3.4.2.2a, b, c et d présentent les occurrences et les fréquences d'emploi relatives des pronoms simples et composés du pluriel.

**Tableau 4.3.4.2.2.a - Analyse quantitative/statistique descriptive des données de l'emploi variable des pronoms *nous* et *nous-autres***

<b>Variante</b>	<b>Nombre d'occurrences</b>	<b>Fréquence d'emploi relative</b>
<i>Nous</i>	76	97 %
<i>Nous-autres</i>	2	3 %
<b>TOTAL</b>	78	100 %

**Tableau 4.3.4.2.2.b - Analyse quantitative/statistique descriptive des données de l'emploi variable des pronoms *vous* et *vous-autres***

<b>Variante</b>	<b>Nombre d'occurrences</b>	<b>Fréquence d'emploi relative</b>
<i>Vous</i>	205	99,5 %
<i>Vous-autres</i>	1	0,5 %
<b>TOTAL</b>	200	100 %

**Tableau 4.3.4.2.2.c - Analyse quantitative/statistique descriptive des données de l'emploi variable des pronoms *eux* et *eux-autres***

Variante	Nombre d'occurrences	Fréquence d'emploi relative
<i>Eux</i>	45	100 %
<i>Eux-autres</i>	0	0 %
TOTAL	45	100 %

**Tableau 4.3.4.2.2.d - Analyse quantitative/statistique descriptive des données de l'emploi variable des pronoms simples et composés non clitiques**

Variante	Nombre d'occurrences	Fréquence d'emploi relative
<i>Simple</i> s	326	99 %
<i>Composés</i>	3	1 %
TOTAL	329	100 %

Les analyses statistiques montrent qu'en effet, les pronoms simples sont de loin les plus fréquents, représentant 97 % à 100 % des occurrences. La prépondérance des formes en *vous* s'explique par le fait que les internautes échangeant sur Facebook s'engagent dans des discussions et répondent à d'autres internautes. La préférence des formes simples peut s'expliquer par l'économie du langage, et le désir de s'exprimer le plus rapidement possible.

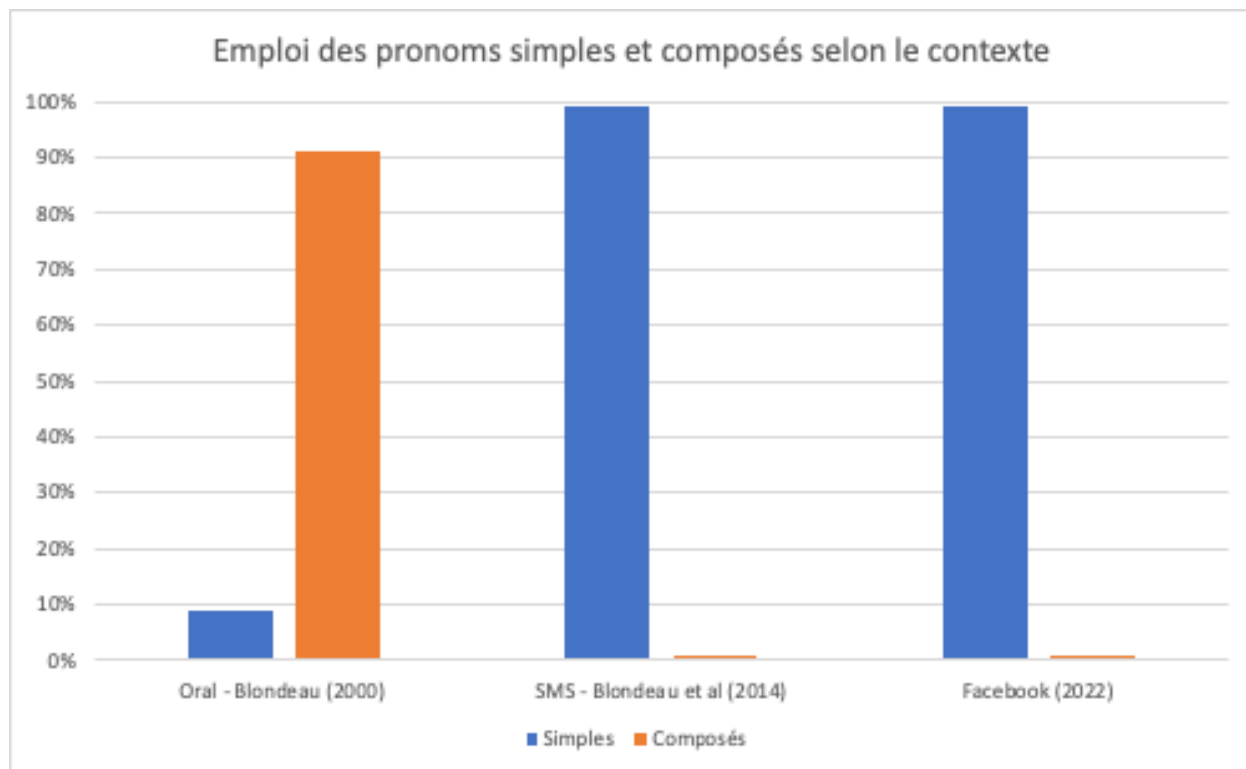
#### **4.3.4.2.3 Comparaison aux fréquences d'emploi attestées à l'oral et dans les SMS**

Le Tableau 4.3.4.2.3 présente les résultats de recherches antérieures, qui ont entrepris des analyses statistiques de l'alternance d'emploi des pronoms simples et composés du pluriel à l'oral et dans les SMS, et reprend les résultats de mon analyse. Ensuite la Figure 4.3.4.2.3 montre une représentation visuelle des données du Tableau 4.3.4.2.3.

**Tableau 4.3.4.2.3 - Comparaison des fréquences d'emploi des pronoms non clitiques simples et composés à l'oral, dans les SMS et sur Facebook**

Variante	Oral - Blondeau (2000)	SMS - Blondeau <i>et al.</i> (2014)	Facebook (2022)
<i>Simple</i> s	9 %	99 %	99 %
<i>Composés</i>	91 %	1 %	1 %

**Figure 4.3.4.2.3 - Comparaison des fréquences d'emploi des pronoms non clittiques simples et composés à l'oral, dans les SMS et sur Facebook**



Ces représentations montrent d'une part les écarts entre l'oral et la CNE, et d'autre part la similitude des données dans les SMS comparativement à Facebook. En effet, mes résultats s'alignent exactement sur ceux de Blondeau *et al.* (2014) en ce qui concerne l'emploi des pronoms simples et composés du pluriel dans les SMS. D'après Blondeau *et al.* (2014), cette chute des pronoms composés aux dépens des pronoms simples suit les tendances communautaires : n'oublions pas non plus que les données de l'oral datent de déjà plus de 20 ans. Enfin, on constate encore la préférence pour les formes simples, ce qui pourrait s'expliquer par le désir de rapidité, comme c'est le cas pour plusieurs variables susmentionnées dans ce chapitre.

En somme, pour ce qui est des marqueurs pronominaux, les internautes au Québec, lorsqu'ils écrivent dans l'espace public de Facebook, emploient le plus souvent les formes simples,

notamment le pronom *on* aux dépens du pronom *nous*, et les pronoms simples non clitiques *nous*, *vous* et *eux*, aux dépens des formes composées *nous-autres*, *vous-autres* et *eux-autres*. Ces tendances s'alignent à la fois sur l'évolution des tendances communautaires du français parlé au Québec, et sur les tendances retrouvées dans les SMS québécois.

#### 4.4 Conclusion

Ce quatrième chapitre a présenté les résultats des fréquences d'emploi des marqueurs lexicaux, verbaux, syntaxiques et pronominaux dans la production écrite par des internautes au Québec sur Facebook. Tout d'abord, ces analyses montrent que les variables les plus fréquentes dans le corpus sont les marqueurs syntaxiques de négation et d'interrogation, et les moins fréquentes, les marqueurs lexicaux. En particulier, l'emploi de la négation est de loin la variable la plus fréquente, avec plus de 800 occurrences dans le corpus, comparativement aux autres variables qui ont toutes moins de 400 occurrences chacune.

Pour ce qui est des fréquences relatives d'emploi, parmi les marqueurs lexicaux, la variante *juste* est de loin la plus fréquente des formules de contrainte, et les variantes *donc* et *alors* sont les plus fréquentes parmi les marqueurs de conséquence. Ensuite, le futur analytique est de loin privilégié comparativement au futur synthétique, et le corpus contient très peu d'instances d'emplois non standards des auxiliaires *avoir* et *être* au passé composé. Quant aux marqueurs syntaxiques, tandis que l'emploi de la particule *ne* est légèrement plus fréquent que son omission, la formule interrogative sans inversion est de loin la plus fréquente parmi les quatre variantes. Enfin, pour ce qui est des marqueurs pronominaux, le pronom *on* est beaucoup plus fréquent que le pronom *nous*, et l'emploi des pronoms composés non clitiques est marginal, comparativement aux pronoms simples *nous*, *vous* et *eux*.

Ce qui surprend, c'est que parmi ces fréquences d'emploi retrouvées sur Facebook, certaines d'entre elles s'alignent sur les tendances communautaires à l'oral et/ou dans d'autres contextes numériques, tandis que d'autres divergent de manière assez significative de ces autres contextes. Comparativement aux fréquences attestées à l'oral, les fréquences relatives d'emploi sur Facebook suivantes s'en rapprochent : les variantes du futur, l'alternance *on-nous* et les formules

interrogatives. Par contraste, les emplois sur Facebook des marqueurs de contrainte et de conséquence, des auxiliaires *avoir* et *être* au passé composé, de la négation, et des pronoms composés non clitiques divergent de façon importante des fréquences attestées à l'oral. En revanche, pour la grande majorité des variantes sur Facebook, leurs fréquences d'emploi s'alignent étroitement sur celles attestées dans les SMS et le (t)chat.

La seule exception flagrante est l'alternance entre l'emploi et l'omission de la particule *ne* dans les formules négatives : l'emploi du *ne* est environ dix fois plus fréquent sur Facebook qu'il ne l'est dans le (t)chat. En bref, ces résultats suggèrent que la production écrite par les internautes au Québec sur Facebook, alors qu'elle montre quelques ressemblances avec le français parlé au Québec dans des situations informelles, ressemble beaucoup à la production écrite par les Québécois.es dans d'autres contextes numériques.

Somme toute, l'emploi des marqueurs sociolinguistiques à l'étude dans la CNE écrite par les internautes au Québec sur Facebook suit généralement les tendances communautaires. Le Chapitre 5, qui présente les résultats des analyses de collocation, examine encore plus en détails dans quelle mesure cette CNE sur Facebook suit (ou pas) les tendances du standard écrit.

## CHAPITRE 5 : RÉSULTATS DES ANALYSES DE COLLOCATION

### 5.0 Introduction

L'analyse des fréquences d'emploi de marqueurs sociolinguistiques du français québécois dans le Chapitre 4 a permis de mettre en lumière les similitudes entre la communication numérique écrite par des internautes sur Facebook et les pratiques linguistiques des Québécois.es à l'oral et à l'écrit dans d'autres contextes en ligne. Les résultats de collocation présentés dans ce chapitre offrent une mesure plus précise de la variabilité et de l'hybridité de cette production écrite sur le plan de commentaires individuels. S'inspirant des études de collocation en français par Gaétane Dostie (2013), Raymond Mougeon et Katherine Rehner (2015) et Kristen Terry Kennedy (2017), ce chapitre a pour objectif d'examiner les cooccurrences d'éléments standards et non standards, afin d'explorer à quel point les internautes québécois.es sur Facebook adhèrent aux normes de l'écrit standard ou non.

### 5.1 Présentation des résultats de collocation

D'après Michel Marcoccia (2016 : 74), la CNE contient des éléments non standards qui représentent

les fonctionnalités des marqueurs paraverbaux et non verbaux de l'oral [...] Il s'agit des émoticônes, ou de l'utilisation des majuscules pour donner de l'emphase, de l'allongement ou de la répétition de caractères pour simuler des effets de prononciation ou de la représentation de vocalisations, comme « Hum », par exemple.

Ainsi, dans le but de découvrir à quel point ces types de fonctionnalités imprègnent le corpus, ce chapitre présente les résultats de cooccurrence d'éléments non standards avec cinq variables à l'étude. Plus précisément, j'analyse les cooccurrences d'éléments standards et non standards dans un seul commentaire d'internaute. Tout d'abord, il s'agit d'examiner les variables à quatre variantes, notamment les marqueurs lexicaux de contrainte et de conséquence et les formules interrogatives. Ensuite, je montre les résultats de cooccurrence avec des variables à deux variantes, notamment l'emploi du futur et de la négation. Il importe de noter aussi que pour l'analyse de la négation je me suis limitée aux occurrences contenant aux moins deux instances de la négation.

Dans le même ordre d'idées, je ne me focalise que sur les cinq variables susmentionnées par un souci de simplicité, le nombre d'occurrences non standards des auxiliaires et des pronoms composés étant minime, et le nombre d'occurrences des pronoms *nous* et *on* étant assez vaste.

Les éléments non standards pris en compte dans le cadre de ces analyses sont les suivants :

- les marqueurs sociolinguistiques à l'étude non standards comme *fait que*,
- les émojis et les hashtags
- les accords non standards comme *aller* pour *allez*, *les table* pour *les tables*
- l'absence ou la surutilisation de la ponctuation ou de l'accentuation comme *!! ? ? ?*,
- l'alternance codique par des emprunts de l'anglais, comme *checker*, et *booker*,
- l'interjection ou un québécoisisme provenant de l'oral comme *coudon* et *tabarouette*.

## **5.2 Cooccurrence de formes non standards avec les variables à quatre variantes**

Dans cette section je présente les taux des cooccurrences d'éléments non standards avec chaque occurrence des marqueurs lexicaux de contrainte et de conséquence, ainsi qu'avec chaque occurrence des formules interrogatives. Cela permet de déterminer dans quelle mesure les éléments non standards sont présents dans la CNE produite par des internautes au Québec sur Facebook.

### **5.2.1 Cooccurrences de formes non standards et des marqueurs de contrainte**

Pour ce qui est de la cooccurrence d'éléments non standards avec les marqueurs de contrainte, le Tableau 5.2.1 montre le nombre d'occurrences de chaque variante, accompagnées d'éléments non standards dans le même commentaire.

**Tableau 5.2.1 Cooccurrence des marqueurs de contrainte avec éléments non standards**

<b>Variante</b>	<b>Commentaires (occurrences) contenant des éléments non standards</b>	<b>Total des occurrences de cette variante</b>	<b>Pourcentage des occurrences de la variante contenant des formes non standards</b>
<i>Juste</i>	73	96	76 %
<i>Seulement</i>	9	18	50 %
<i>Ne...que</i>	4	18	22 %
<i>Rien que</i>	1	2	50 %
<b>TOTAL</b>	<b>87</b>	<b>134</b>	<b>65 %</b>

Les données dans le tableau ci-dessus permettent de constater que dans la majorité des occurrences de la variante *juste*, le commentaire contient des éléments non standards, y compris les émojis. En particulier, 73 soit environ 76 % des commentaires contenant *juste*, contiennent également des formes non standards, tels que présentés dans la liste de la section 5.1. Quant aux commentaires contenant *seulement*, neuf occurrences, soit 50 % des commentaires contiennent des éléments non standards. Voici les deux exemples les plus évidents, de deux internautes distincts, provenant de la composante *La Presse* du corpus. Dorénavant, les éléments non standards seront en caractères italiques.

on parle ici de taxes *provibcial* donc *oui seulement* 80 *piassent* et *oui* la moyenne des auto *c est* 8 *a* 9 *litre* au 100 pas *tous* le monde qui *ont* des *pick up* (2183)

[On parle ici de taxes provinciales donc oui seulement 80 piasses et oui la moyenne des autos c'est 8 à 9 litres au 100, pas tout le monde qui a des camionnettes.]

**Seulement** 400 ? *Ben voyons criss*, c'est plus que raisonnable. (2542)

Dans les deux cas, l'internaute emploie un québécoisisme – *piassent* et *ben voyons criss* – et dans le premier, des marques de la conversation – *oui* – un emprunt à l'anglais – *pick up* – des fautes de frappe et accords non standards ou manquants, un accent manquant, et aucune ponctuation. Cela

témoigne d'une part du besoin de rapidité et d'autre part de la variété des pratiques langagières chez ces internautes.

Certaines cooccurrences indiquées dans le Tableau 5.2.1 surprennent, par exemple celles en conjonction avec la formule *ne...que*. En effet, tandis que la fréquence de la formule *ne...que* est inattendue dans la CNE puisqu'il s'agit d'un usage très soutenu, son emploi dans 22 % des cas est accompagné d'autres éléments non standards avant et/ou après dans le même commentaire. En voici deux exemples :

 *la* les gens **ne** demandent **qu'**à voyager et les *cies* diminuent (2667, *La Presse*)

[Là, les gens ne demandent qu'à voyager et les compagnies diminuent.]

Radio cacanne **ne** *s'inquiète* de la *liberte* d'expression **que** lorsqu'ils... (1525, *Le Devoir*)

[Radio Canada ne s'inquiète de la liberté d'expression que lorsqu'ils...]

Pour ce qui est des éléments typiques de la CNE, le premier commentaire contient un emoji et une abréviation. Ensuite, dans les deux commentaires, il manque des accents, et dans d'autres cas, des espaces remplacent les apostrophes, et/ou il manque de la ponctuation. Ainsi, alors que certains internautes prennent le soin d'utiliser la formule *ne...que*, ils ou elles font moins attention à d'autres éléments dans leurs commentaires.

Paradoxalement, alors que la variante *rien que* est plutôt informelle, son emploi sur Facebook ne l'est pas nécessairement. Voici les deux seules occurrences de *rien que*, provenant des commentaires sur la page *Le Devoir*:

**Rien que** le fait qu'ils aient attendu d'avoir quitté pour s'en inquiéter en dit long! (1490)

Pas **rien que** ça *c'eat* rendu que des professeurs d'université *donne leur opinion* et se font suspendre *sens* solde maintenant le contrôle de l'information dans les media *pus* personne peut parler de rien débattre d'idée! C'est triste pour les *media traditionnelle* car *a suivre* *se* chemin les condamne *a disparaître!!!!* (1501)

[Pas rien que ça, c'est rendu que des professeurs d'université donnent leurs opinions et se font suspendre sans solde, maintenant le contrôle de l'information dans les médias, plus personne ne peut parler de rien débattre d'idée! C'est triste pour les médias traditionnels car suivre ce chemin les condamne à disparaître!]

Plus précisément, la première occurrence de *rien que* ne comporte aucun autre élément non standard, tandis que dans la deuxième, l'internaute, moins attentif à son style, a fait un usage très limité de la ponctuation. De plus, dans les deux cas les internautes concluent leur commentaire avec un ou plusieurs points d'exclamation.

En résumé, bien que les variantes *ne...que* et *seulement* sont des formes plutôt standards, elles sont accompagnées respectivement de 25 % et 50 % du temps d'éléments non standards. De manière plus générale, on observe avec les marqueurs de contrainte la présence non négligeable d'éléments non standards. Ces constatations méritent d'être examinées de manière approfondie dans le cadre de futures recherches, afin de comprendre pourquoi les internautes alternent entre le standard et le non standard. Par exemple, il serait pertinent d'interroger les internautes sur l'intention, et sur le rôle de l'affect dans le choix de tournures et de ponctuation. Une telle recherche apporterait quelques réponses sur l'importance et la fonction sociale de ces éléments non standards.

### **5.2.2 Cooccurrences de formes non standards et des marqueurs de conséquence**

En ce qui concerne les cooccurrences de formes non standards avec chacun des marqueurs de conséquence, le Tableau 5.2.2 présente le nombre d'occurrences de chaque variante accompagnée d'éléments non standards. Une première observation réside dans la similitude du taux d'éléments non standards accompagnant les marqueurs de conséquence avec celui observé pour les marqueurs de contrainte, soit 65 %. Notons que la variante *so* a été exclue de cette analyse puisqu'elle n'apparaît pas dans le corpus.

**Tableau 5.2.2 Cooccurrence des marqueurs de conséquence avec éléments non standards**

Variante	Occurrences contenant des éléments non standards	Total des occurrences de cette variante	Pourcentage des occurrences de la variante contenant des formes non standards
<i>Alors</i>	25	40	63 %
<i>Donc</i>	37	60	62 %
<i>Fait que</i>	4	4	100 %
TOTAL	66	104	64 %

Les trois exemples suivants montrent la variété d'éléments non standards accompagnant les marqueurs de conséquence, notamment des troncations et des abréviations, l'emploi répétitif de ponctuation, de symboles et de voyelles, les émojis, un manque d'accents, des accords non standards et des québécoisismes :

*Alors pkoï* dites-vous des niaiseries si vous vous en sacrez ? D'ailleurs, *pkoï* êtes-vous sur le site du Devoir au départ ? ? ? étrange... 🤔 🤔 🤔 (1450, *Le Devoir*)

😁 aaaaaa oui la belle époque. [...] **Donc** pas de belles variétés d'aliments comme aujourd'hui. Pas de carte soleil, [il] fallait payer pour les soins et *c'était pas le meme* niveau qu'aujourd'hui dans l'avancement des *recherche*. [...] (2134, *La Presse*)

*Ouain faque...* Ça va être un trio frites au poulet *stp!* 😊 😊 \*\*\*\*\* (2027, *TVA Nouvelles*)

En résumé, pour ce qui est des marqueurs de conséquence, bien que *alors* s'aligne plutôt sur l'oral et *donc* plutôt sur l'écrit (Blondeau et Tremblay, 2022a), les deux variantes se retrouvent avec des éléments non standards 63 % et 62 % du temps respectivement. Ensuite, logiquement, 100 % des occurrences du marqueur *fait que*, qui relève de l'oral, contiennent d'autres éléments non standards. En tout, les marqueurs de conséquence sont accompagnés d'éléments non standards dans 64 % des cas, ce qui illustre la prépondérance d'éléments non standards dans les écrits dans l'espace public de Facebook par ces internautes au Québec.

### 5.2.3 Cooccurrences de formes non standards et des formules interrogatives

Pour ce qui est des cooccurrences de formes non standards avec chacune des formules interrogatives, le Tableau 5.2.3 présente le nombre d'occurrences, de chaque variante, accompagnées d'éléments non standards. Une première remarque est que le taux d'éléments non standards accompagnant ces variantes est légèrement plus bas, 54 %, comparativement aux marqueurs lexicaux, à 64 % pour les marqueurs de conséquence et 65 % pour les marqueurs de contrainte.

**Tableau 5.2.3 Cooccurrence des formules interrogatives avec éléments non standards**

<b>Variante</b>	<b>Occurrences contenant des éléments non standards</b>	<b>Total des occurrences de cette variante</b>	<b>Pourcentage des occurrences de la variante contenant des formes non standards</b>
<i>Est-ce que</i>	4	11	36 %
<i>Inversion</i>	13	25	52 %
<i>Sans inversion</i>	80	156	51 %
<i>-tu</i>	14	15	93 %
<b>TOTAL</b>	111	207	54 %

Tout d'abord, on constate que la variante typiquement québécoise relevant de l'oral, la formule interrogative avec le suffixe *-tu*, représente le taux le plus élevé de cooccurrence avec des formes non standards, à 93 % des occurrences. Ce taux est proche de celui observé pour *fait que*, un marqueur de conséquence relevant également de l'oral, qui se retrouve dans 100 % de ses occurrences dans le corpus avec des éléments non standards. Dans les deux cas, ce sont alors des formes non standards typiquement québécoises, contrairement aux autres variantes, qui se retrouvent très souvent avec d'autres éléments non standards. Ensuite, la forme la plus formelle ou standard, *est-ce que*, affiche le taux le plus bas de formes non standards l'accompagnant, avec seulement 36 % des occurrences. Quant aux variantes interrogatives d'inversion et sans inversion, elles sont accompagnées d'autres éléments non standards environ 50 % du temps.

Ces observations suggèrent que la plupart des internautes qui choisissent une variante interrogative plus standard optent généralement pour un style plus standard, tandis que ceux et celles qui choisissent une variante non standard relevant de l'oral tendent à adopter un style non standard dans leur CNE. Cette observation s'aligne dans l'ensemble sur les résultats de cooccurrence des marqueurs de contrainte et de conséquence.

Les cinq exemples suivants illustrent la cooccurrence d'éléments non standards avec chacune des formules interrogatives, ainsi que l'unique occurrence de la formule en suffixe *-tu* qui est en cooccurrence avec des formes standards uniquement :

**Est-ce qu** il y a encore la section fumeur et non fumeur ? Si *c'est pas* le cas, ce n'est pas vraiment *vintage* 😊 😊 \*\*\*\*\* (2012, *TVA Nouvelles*)

*Ah oui* la bonne eau fraîche du port... **C'est tu juste** moi *ouuu* ? 😊 (1300, *Journal de Québec*)

toi **t** *mieux* que tout le monde ? 🙄 🙄 🙄 (2151, *La Presse*)

*Coudonc*, **avez-vous** lu et compris le sens de l'article ?... inutile de répondre. (1450, *Le Devoir*)

**j'ai tu** l'air d'avoir besoin et envie d'y aller ? (1266, *Journal de Québec*)

Les quatre premiers extraits révèlent les éléments non standards suivants : une variante sociolinguistique non standard, notamment l'omission de la particule *ne*, des émojis, un emprunt à l'anglais, notamment *vintage*, des interjections, notamment *Ah* et *Coudonc*, la répétition de symboles ou de lettres comme \*\*\*\*\* et *ouuu*, et une orthographe phonétique, comme *t* pour représenter *t'es*. Le premier extrait en particulier met en évidence le mélange de codes et de styles, une hybridité écrite, orale et iconique : l'internaute emploie la formule plutôt formelle *est-ce que*, alterne entre l'omission et l'emploi de la particule *ne* pour la négation, utilise des émojis et des

symboles, ainsi que de l'anglais, le tout dans le même commentaire. Enfin, le dernier extrait illustre un autre exemple surprenant de ce va et vient entre le standard et le non standard : l'internaute commence son commentaire de manière non standard, puis choisit de s'aligner sur le standard écrit dans la suite de son commentaire.

Notons aussi, comme nous l'avons vu au Chapitre 4, que les formules interrogatives avec et sans inversion sont accompagnées d'un point d'interrogation en fin d'énoncé dans 92 % et 100 % des occurrences, respectivement. Il est alors intéressant que les internautes produisent un mélange d'éléments standards et non standards environ 50 % du temps, mais se forcent à utiliser la ponctuation en finale presque catégoriquement. Cela appuie la thèse que les internautes font des choix linguistiques conscients, dans des buts précis.

En somme, ces résultats indiquent que les marqueurs interrogatifs sont accompagnés un peu plus de 50 % du temps d'éléments non standards. Des analyses plus granulaires de ces cooccurrences pourraient dévoiler le rôle de l'environnement linguistique dans les choix stylistiques, et un questionnaire auprès d'internautes pourrait fournir des renseignements sur ce qui motive ces choix.

Les résultats présentés dans les trois sections ci-dessus montrent le taux important de cooccurrence de variantes et d'éléments non standards avec les variantes de contrainte, de conséquence et d'interrogation. Globalement, pour ces trois variables, les taux de cooccurrence d'éléments non standards sont de 65 %, 63 % et 54 % respectivement. C'est-à-dire que plus de 50 % des commentaires contenant une variante lexicale ou interrogative à l'étude contiennent du langage et/ou des éléments qui divergent du standard écrit, en ce qui concerne ces trois variables sociolinguistiques. Je propose que ces chiffres soient une mesure du degré d'hybridité de la CNE d'internautes au Québec sur Facebook.

### 5.3 Cooccurrences de formes non standards et des variables à deux variantes

Dans cette section je présente les taux des cooccurrences de formes standards et non standards pour les occurrences de la négation et du futur. Comme c'était le cas dans les sections précédentes, j'ai également tenu compte de la présence des éléments qu'illustrés dans la liste dans la section 5.1.

#### 5.3.1 Cooccurrences de formes standards et non standards dans les occurrences du futur

Pour ce qui est de la cooccurrence des instances du futur, j'ai examiné tous les commentaires pour leur contenu d'éléments non standards. Le Tableau 5.3.1 montre le nombre d'occurrences de chaque variante qui sont accompagnées d'éléments non standards.

**Tableau 5.3.1 Cooccurrence du futur avec éléments non standards**

<b>Variante</b>	<b>Occurrences contenant des éléments non standards</b>	<b>Total des occurrences de cette variante</b>	<b>Pourcentage des occurrences de la variante contenant des formes non standards</b>
Futur analytique	183	238	77 %
Futur synthétique	68	136	50 %
<b>TOTAL</b>	<b>251</b>	<b>374</b>	<b>67 %</b>

Tout d'abord, des 374 commentaires contenant un verbe conjugué au futur (synthétique ou analytique), 251 d'entre eux contenaient un élément non standard ; cela équivaut à 67 % des commentaires contenant des éléments non standards. Ensuite, l'analyse des occurrences du futur analytique comparativement à celles du futur synthétique indique que les formes non standards sont en cooccurrence avec le futur analytique plus souvent qu'avec le futur synthétique ; en effet, tandis que 77 % des cas du futur analytique se retrouvent en cooccurrence avec des éléments non standards, 50 % des cas du futur synthétique sont accompagnés d'éléments non standards.

Les trois extraits suivants illustrent la cooccurrence d'éléments non standards avec l'expression du futur analytique :

ton ami Duhaime **va** nous **regler** ça lui 😊😊😊😊 (2197, *La Presse*)

*oui mais ses nous on* décide qui **va être** p.m et Legault *ses* dehors (2225, *La Presse*)

[Oui, mais c'est nous, on décide qui va être le premier ministre et Legault c'est dehors.]

Wow une grosse économie à la pompe sur le plein *sa!!!!* 2.25\$ de plus pour s'acheter une *bar* de chocolat *loll!* *Criss* vous voulez *aidés* les gens.... 1\$/L et *sa va faire* une *différence* et lâcher de nous *rire dans face!* (2406, *La Presse*)

[Wow une gross économie à la pompe sur le plein ça! 2,25\$ de plus pour s'acheter une barre de chocolat \*mort de rire\*! Christ, vous voulez aider les gens ... 1\$/L et ça va faire une différence et lâcher de nous rire dans le visage!]

On remarque les mêmes genres de procédés que dans les sections précédentes : l'emploi d'émojis, soit la répétition de signes de ponctuation soit un manque de majuscules en début de phrase, d'accents et de ponctuation, des accords non standards, comme *vous voulez aidés*, des orthographes phonétiques comme *ses* pour *c'est* et *sa* pour *ça*, l'acronyme *loll*, des québécoisismes comme *criss* et *rire dans face*, et d'autres procédés de l'oral tel que l'emphase par la répétition des pronoms *nous* et *on*.

Les extraits suivants illustrent la cooccurrence d'éléments non standards avec l'expression du futur synthétique. Les procédés ressemblent à ceux déjà explicités ; on constate qu'ils sont plutôt grammaticaux, comme les accords inattendus et un manque de ponctuation, mais on observe aussi des émojis, des interjections de l'oral comme *Ben oui* et des québécoisismes comme *chialer*, une variante non standard, notamment l'omission de la particule *ne*, et de l'écriture phonétique, comme *pu* pour *plus*. En revanche, on ne voit pas dans ces extraits le procédé de répétition de lettres ou de signes de ponctuation.

Legault j'espère sincèrement que les *québécois se réveilleront* et *l'automne tu seras* au chômage. Tu as détruit le Québec et la vie de beaucoup de monde 😏😏 (2264, *La Presse*)

*Ben oui soyons dans une dictature comme la Russie, tu seras content* (2573, *La Presse*)

Viens pas *chialer* après quand on *aura pu* de service dans *rien #societedelache* (2467, *La Presse*)

De plus, il importe de noter que 15 commentaires contenant le futur synthétique contenaient également un verbe conjugué au futur analytique, c'est-à-dire que les internautes mélangent les deux variantes dans un même commentaire, dans environ 4 % des cas de l'expression du futur, ou 11 % des cas du futur synthétique. L'exemple suivant illustre cette alternance entre futur analytique et futur synthétique dans un même commentaire.

Au Québec jamais ça *va arrivé* pas *graves j'irai* en Ontario (2391, *La Presse*)

En somme, les résultats de collocation de la variable du futur s'alignent plus ou moins sur ceux des marqueurs de contrainte, de conséquence et d'interrogation, sur le plan du taux de cooccurrence d'éléments non standards, qui oscille entre 54 % et 67 % pour ces quatre variables. Cela montre à quel point des éléments non standards sont présents dans ce corpus de production écrite dans l'espace public de Facebook des internautes au Québec, et qu'il est donc impossible de dire que cette CNE suit de très près le standard écrit.

Ensuite, le futur analytique, relevant plutôt de l'oral comparativement au futur synthétique, se retrouve plus souvent avec des éléments non standards que son homologue. Cela semble suivre la tendance constatée pour les autres variables. Par exemple, les variantes les plus formelles, *ne...que* et *est-ce que* se retrouvent le moins souvent avec des éléments non standards, notamment dans 22 % et 36 % des cas respectivement. En revanche, les marqueurs moins formels, notamment *juste*, *fait que* et l'interrogation avec le suffixe *-tu* se retrouvent le plus souvent avec des éléments non standards, dans 76 %, 100 % et 93 % des cas, respectivement. Pour ce qui est des autres variantes,

elles se retrouvent avec des éléments non standards entre 50 % et 62 % du temps. En d'autres termes, aucune variante, même si elle est hyper-formelle, ne se retrouve uniquement dans des contextes à 100 % standards ; le non standard est donc manifeste dans ce corpus.

### 5.3.2 Cooccurrences de formes standards et non standards de la négation

En ce qui concerne les cooccurrences des deux variantes de la négation, notamment l'emploi ou l'omission de la particule *ne*, je me suis restreinte à l'examen des commentaires qui contenaient au moins deux occurrences de la négation. Cela équivaut à un peu plus de 10 % de la totalité des occurrences de la négation qui ont été examinées dans le Chapitre 4. Cette analyse se limite alors à la cooccurrence des variantes standards et non standards de la négation, mais plus loin j'explicité la présence d'autres éléments non standards dans cet échantillon. Le Tableau 5.3.2 présente le nombre de commentaires contenant au moins deux instances de la négation : plus précisément, les commentaires contenant uniquement la variante standard, soit l'emploi du *ne*, ceux contenant uniquement la variante non standard soit l'omission du *ne*, et ceux contenant un mélange des deux variantes, c'est-à-dire dans lequel l'internaute a alterné entre l'emploi et l'omission de la particule *ne* dans son commentaire.

**Tableau 5.3.2 Cooccurrences des variantes de la négation**

Variante	Nombre de commentaires avec 2+ instances de la négation	Pourcentage des commentaires avec 2+ instances de la négation
Emploi du <i>ne</i>	55	39 %
Omission du <i>ne</i>	44	31 %
Alternance des deux variantes	43	30 %
TOTAL	142	100 %

En tout, 142 commentaires contenaient deux ou plusieurs instances de la négation ; de ces commentaires, 55 ou 39 % contenaient uniquement la variante standard, l'emploi du *ne*, 44 ou 31 % contenaient uniquement des variantes non standards, l'omission du *ne*, et un autre 43 ou 30 % des commentaires contenaient un mélange de variantes standards et non standards. Ainsi, tandis que 39 % des commentaires contiennent uniquement la variante standard, 61 % contiennent

au moins une instance de variante non standard. Ces résultats s'alignent sur ceux présentés par Hélène Blondeau et Mireille Tremblay (2022b : 134) dans leur survol des pratiques du vernaculaire dans les SMS, qui ont également trouvé « des cas qui attestent de l'absence et de la présence de *ne* dans un même énoncé » et donc de la « variation inhérente chez un même texteur ». Cette pratique de variation stylistique sur le plan individuel pourrait alors représenter une caractéristique des écrits numériques.

Par ailleurs, en examinant les 55 commentaires contenant uniquement la variante standard, 35 ou 64 % d'entre eux contiennent au moins un autre élément non standard tel qu'explicité dans la section 5.1 de ce chapitre, et ce par 30 internautes distincts. Bref, des 142 commentaires contenant au moins deux instances de la négation, seulement 20 commentaires écrits par 17 internautes distincts contiennent uniquement de la production écrite qui suit le standard. En d'autres termes, 14 % des commentaires contenant au moins deux occurrences de la négation suivent le standard écrit et 86 % des commentaires contiennent au moins un élément non standard. Les trois exemples ci-dessous illustrent cette cooccurrence d'éléments non standards avec l'expression de la négation suivant le standard écrit :

Il **ne** s'agit **pas** de votre nombril ici, mais de *qqchse* d'important pour d'autres dont vous **ne** mesurez **pas** ni **n**'en comprenez l'ampleur!! *Tabarouette* que le «jem'enfoutisme» s'exprime sans gêne ni honte *astheure...* Gardez-vous donc une *p'tite* gêne, votre «boutade» est tout à fait inutile, nombriliste et puérole. 🙄🙄🙄 (1450, *Le Devoir*)

*ils vient* explicitement pour voir les victimes... s'il **n**'y avait **pas** *eut* de crime, le *PAPE* **ne** se serait **pas** déplacé. (2540, *La Presse*)

*donc* selon ce que tu *dit ont* **ne** devrait **pas** entendre les *lgbtq+* ou *feministe* dire “Hommes *blanc* hétéro” *vue qu'il* le disent pour *insulté*. *Ont* **ne** devrait **pas** non plus entendre de personnes *noir* dire “*white boy*” ou “*white girl*” qui est cité comme une insulte... Un peu de Tom Mcdonald vous *ferais* du bien 🙄🙄 (2120, *La Presse*)

On remarque dans le premier exemple une abréviation, *qqchose*, des québécoisismes, *tabarouette* et *astheure*, et une troncation, *p'tite*, ainsi que des émojis. Dans le deuxième, on constate des formes non standards : une terminaison verbale, *ils vient*, un participe passé, *eut*, et l'emploi non standard de majuscules, *PAPE*. Enfin le troisième exemple illustre les mêmes genres d'éléments non standards : le manque de majuscule en début de phrase ; l'accord verbal, *tu dit*, *vue*, *insulté*, et *ferais* ; l'accord nominal et adjectival, *les féministe*, *Hommes blanc* ; le pronom, *ont* ; de l'anglais, *white boy* et *white girl* ; et un émoji.

Les résultats portant sur les cooccurrences des variantes standards et non standards de la négation suivent donc généralement la tendance des autres variables examinées dans ce chapitre, notamment que le corpus regorge d'éléments non standards. Plus précisément, les formes non standards de la négation sont en cooccurrence dans 61 % des cas, et les éléments non standards y compris la variante non standard sont en cooccurrence dans 86 % des cas de cet échantillon du corpus, donc des commentaires contenant au moins deux occurrences de la négation.

## 5.4 Conclusion

En somme, les résultats présentés dans ce chapitre démontrent que les cinq variables en question sont accompagnées d'éléments non standards entre 50 % et 86 % du temps. La majorité des commentaires contiennent ainsi un mélange de formes standards et non standards. Cela prouve que les internautes au Québec ne suivent pas le standard écrit à 100 % dans leur CNE sur Facebook. Ces résultats suggèrent en effet une mixité ou une hybridité dans cette production écrite, tout comme c'est le cas dans d'autres contextes numériques (Crystal, 2011 ; Marcoccia, 2016).

Plus précisément, les variantes qui sont les plus fréquentes à l'oral informel, tendent à avoir la plus haute fréquence de cooccurrence avec des éléments non standards dans cette CNE. Par exemple, le marqueur de conséquence *fait que*, qui est de loin le plus fréquent à l'oral informel, est accompagné d'éléments non standards sur Facebook dans 100 % des cas. Cette tendance se présente également pour l'emploi du futur analytique, le marqueur de contrainte *juste*, la formule interrogative en suffixe *-tu*, et l'omission de la particule *ne* dans la négation, toutes ces variantes étant les plus fréquentes ou les deuxièmes plus fréquentes à l'oral informel. Ce qui surprend, c'est

que même les variantes relevant plutôt du standard écrit, peu fréquentes à l’oral, montrent des cooccurrences non négligeables avec des éléments non standards sur Facebook, entre 25 % et 64 % du temps.

Un autre élément surprenant est la coprésence de deux formes d’une même variable, par exemple la cooccurrence du futur analytique et synthétique, ou de l’emploi et de l’omission de la particule *ne* dans un même commentaire. Dans leur étude sur les textos québécois, Hélène Blondeau et Mireille Tremblay (2022b), sans faire d’analyse quantitative, ont également relevé des exemples de cette alternance entre l’emploi et l’omission du *ne* dans un même message SMS. Il est donc possible que ce procédé d’alternance stylistique constitue une caractéristique de la CNE, quel que soit le contexte ou la plateforme. Dans le chapitre suivant, il s’agit de faire la synthèse et de tirer des conclusions sur les résultats présentés dans les Chapitres 4 et 5.

## CHAPITRE 6 - DISCUSSION DES RÉSULTATS

### 6.0 Introduction

Cette étude avait pour but d'examiner la variation sociolinguistique du français écrit par des internautes au Québec dans l'espace public de Facebook afin de déterminer si les écrits sur Facebook s'alignent sur l'écrit standard ou plutôt sur la production linguistique dans d'autres contextes. Le deuxième objectif était de déterminer le degré de variabilité de cette CNE, c'est-à-dire dans quelle mesure les internautes utilisent dans leurs commentaires des éléments non standards, que ce soit sur le plan linguistique, iconique ou typographique. La première partie de ce chapitre présente une synthèse et une discussion des résultats de concordance et de collocation, et la deuxième une discussion des implications théoriques des résultats et les limites de l'étude.

### 6.1 Synthèse et discussion des résultats de concordance

Dans un premier temps, à l'exemple des études variationnistes récentes sur le français québécois écrit en contexte numérique qui ont comparé les données de l'oral à celles des SMS (Bally, 2022 ; Blondeau et Tremblay, 2022b), et à l'instar de John Mansfield *et al.* (2023) qui ont examiné les divergences linguistiques entre des dialectes selon le contexte, je propose une vue d'ensemble comparative de plusieurs des marqueurs sociolinguistiques à l'étude.

Afin de présenter un bilan des fréquences d'emploi sur Facebook, à l'oral et dans d'autres contextes numériques, j'adopte un graphique synthétique qui montre les fréquences d'emploi des variables à l'étude selon le contexte de communication. Le Tableau et la Figure 6.1 illustrent les fréquences d'emploi de quatre variables avec une alternance binaire, soit l'alternance du futur analytique et synthétique, l'emploi variable de la négation, l'alternance *on-nous* et l'alternance des pronoms non clitiques simples et composés, ainsi que les quatre variantes de l'emploi variable des formules interrogatives, et ce, selon le contexte, soit à l'oral, sur (t)chat ou SMS, et sur Facebook.

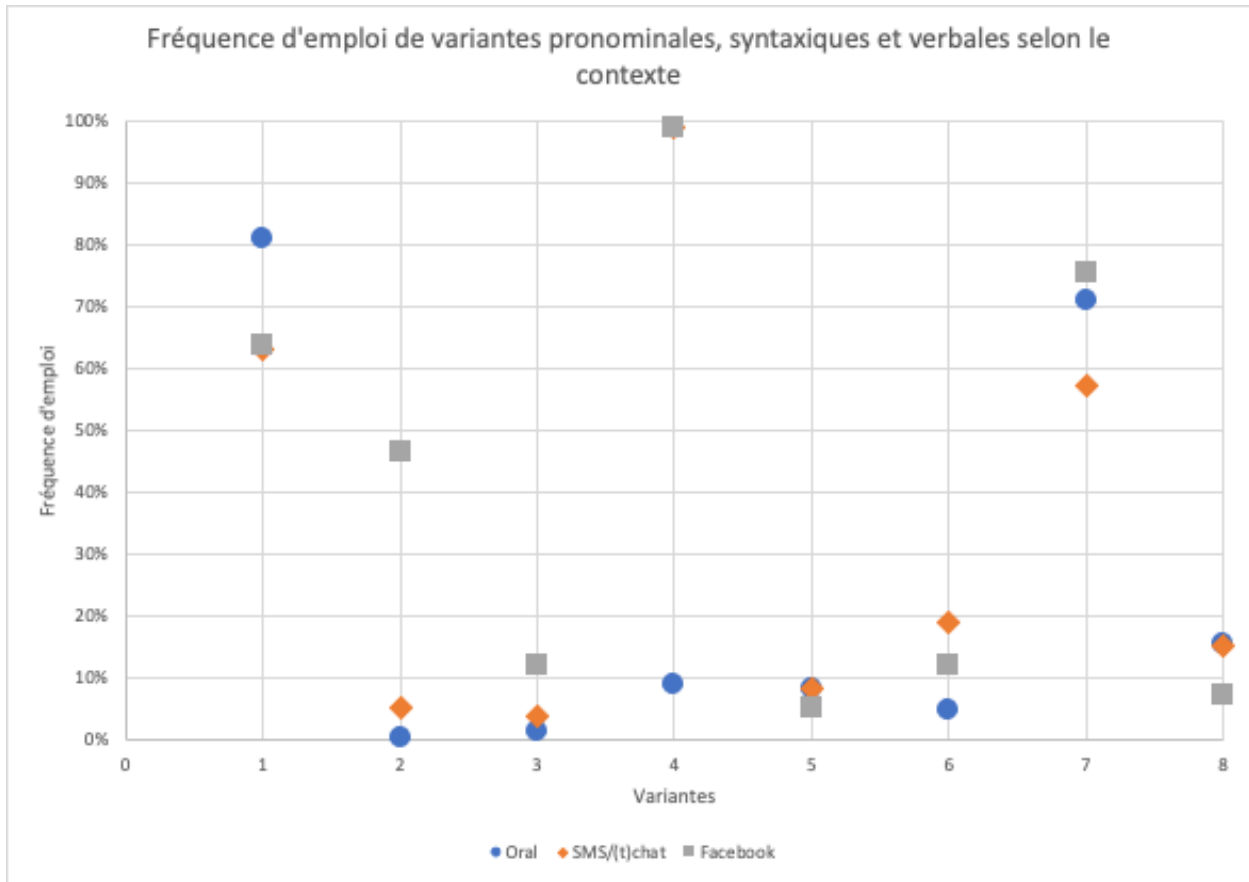
Notons les recherches antérieures dont proviennent ces fréquences d'emploi : l'expression du futur à l'oral par Blondeau (2006) et dans les SMS par Tremblay *et al.* (2019) ; l'emploi de la négation à l'oral par Sankoff et Vincent (1977) et dans le (t)chat par van Compernelle (2010) ; les pronoms *on* et *nous* à l'oral par Laberge (1977) et dans les SMS par Blondeau *et al.* (2014), les pronoms simples et composés à l'oral par Blondeau (2000) et dans les SMS par Blondeau *et al.* (2014), et les formules interrogatives à l'oral par Halverson (2021) et dans les SMS par Bally (2022).

La Figure 6.1 est une représentation graphique des fréquences du Tableau 6.1, et montre de manière visuelle les convergences et divergences entre l'emploi de ces variables à l'oral, dans des contextes numériques et sur Facebook. Afin de rendre la présentation synthétique plus claire, j'ai exclu les marqueurs de contrainte et de conséquence et l'emploi variable des auxiliaires *avoir* et *être* au passé composé, dont la présence reste de toute façon marginale dans le corpus. D'ailleurs, les marqueurs qui ne figurent ni dans le Tableau ni dans la Figure synthétique 6.1 n'ayant pas encore, à ma connaissance, été étudiés dans d'autres contextes numériques, la comparaison serait incomplète.

**Tableau 6.1 Les fréquences d'emploi de cinq variables à l'étude**

<b>Variante</b>	<b>Oral</b>	<b>SMS/(t)chat</b>	<b>Facebook</b>
Futur analytique	81 %	63 %	63,60 %
Emploi du <i>ne</i>	0,50 %	5,10 %	46,50 %
<i>nous</i>	1,60 %	3,70 %	12 %
Pronoms simples	9 %	99 %	99 %
<i>est-ce que</i>	23 %	8,50 %	5,30 %
inversion	24,40 %	19,10 %	12,10 %
sans inversion	36,50 %	57,30 %	75,40 %
<i>-tu</i>	16,50 %	15,10 %	7,20 %

**Figure 6.1 Graphique synthétique des fréquences d'emploi dans trois contextes**



**Légende de l'axe horizontal :** 1. Futur ; 2. Négation ; 3. *On-nous* ; 4. Pronoms non clitiques ; 5. Interrogation - *est-ce que* ; 6. Interrogation - inversion ; 7. Interrogation - sans inversion ; 8. Interrogation - suffixe *-tu*

Le graphique synthétique de la Figure 6.1 montre à quel point cette CNE ressemble à d'autres contextes numériques ainsi qu'à la production orale dans des situations informelles. On constate que les fréquences d'emploi sur Facebook se rapprochent de celles des SMS et du (t)chat pour six variantes : l'emploi du futur, l'emploi des pronoms *on* et *nous* et des pronoms simples et composés, et l'emploi des formules interrogatives en *est-ce que*, avec inversion et en suffixe *-tu*. Les seules divergences majeures entre les résultats de Facebook et ceux des autres contextes numériques sont l'emploi de la négation et la formule interrogative sans inversion. Puisque les résultats antérieurs sur la négation en ligne ont trait au (t)chat et non au SMS, les présents résultats montrent à quel point la CNE sur Facebook ressemble à la CNE dans les SMS, fait surprenant étant donné les différences technologiques et situationnelles entre ces contextes numériques.

En revanche, on constate trois divergences plus importantes entre les fréquences de l'oral et celles qui sont attestées sur Facebook. Plus précisément, on note des différences très significatives de 20 % à 90 % sur le plan de l'emploi du futur, de la négation, et de l'emploi des pronoms simples et composés. N'oublions pas cependant que les données de l'oral portant sur l'emploi de la négation et des pronoms simples et composés datent de plus de 20 ans, et que cette comparaison doit donc se faire avec prudence.

Alors que la variation sociolinguistique sur Facebook diverge donc de l'oral d'environ 30 %, elle diverge d'autres contextes CNE d'environ 12 % seulement. De plus, si l'on exclut les résultats portant sur la négation dans le (t)chat, et si l'on tient compte uniquement des fréquences dans les SMS et sur Facebook, la divergence est de seulement 6,4 %. Ces résultats portant sur la transposition à l'écrit de pratiques de l'oral s'alignent alors de façon générale sur les tendances trouvées dans les recherches sur le numérique, entre autres, celles de Blattner et Williams (2011), de Ferrara *et al.* (1991), Storrer (2014). Ce qui devient clair, c'est que cette CNE produite par des internautes au Québec sur Facebook ne suit pas complètement le standard écrit, puisqu'elle contient des éléments de l'oral tout en se rapprochant des tendances des SMS.

Revenons aux hypothèses présentées dans le Chapitre 3 au sujet de la variation sociolinguistique au niveau du système et de la communauté. Tout d'abord, en raison de la taille illimitée des messages et du contexte public de la situation de communication, j'avais estimé que la variation trouvée dans la production écrite des internautes au Québec sur Facebook s'alignerait sur les tendances de l'écrit standard, avec seulement quelques traces de l'oral. J'avais ensuite avancé que ce seraient les formes les plus courtes qui seraient privilégiées en raison de la tendance du langage vers l'économie (Thurlow, 2003). Étant donné les résultats de cette étude sur l'étendue du système linguistique, la première partie de l'hypothèse est intenable – la Figure 6.1 démontre la proximité entre les fréquences d'emploi des marqueurs sociolinguistiques sur Facebook et celles qui sont attestées à l'oral et dans d'autres écrits numériques. Il est donc impossible de dire que les écrits sur Facebook suivent de près le standard écrit. Quant à la préférence des variantes plus courtes, rappelons tout d'abord que le principe de l'économie du langage se définit comme le recours « à un minimum de traits pertinents » (Dubois *et al.*, 2002 : 163), conséquence de la « recherche permanente de l'équilibre entre [les] besoins communicatifs d'une part, inertie mémorielle et

inertie articulatoire d'autre part » (Martinet, 1980 : 177). Ce principe est donc « un critère [...] qui requiert, entre autres, qu'une analyse [par le locuteur ou la locutrice] doive viser à être aussi courte et à utiliser le moins de termes possibles » (Adegbokou, 2011 : 27), ou « une réduction graphique d'un mot ou d'une suite de mots exécutée [...] en vue de gagner de la place dans un texte sur papier ou tout autre support » (Faudaous, 1990 : 7). Le besoin d'économiser le temps et l'effort, est un principe linguistique fondamental qui représente le noyau de l'évolution linguistique (Vicentini, 2003 : 38).

Appliquons maintenant ce principe à l'analyse des variantes de chacune des variables à l'étude : le Tableau 6.1.1 présente les variantes de chaque variable selon sa longueur et sa fréquence d'emploi. Ce tableau exclut la variable du passé composé, qui ne peut pas contribuer à cette analyse, en raison de la variabilité de longueur des occurrences du passé composé.

**Tableau 6.1.1 Les variables à l'étude selon leur longueur et leur fréquence d'emploi**

Variable	Variante	Longueur/formulation	Fréquence d'emploi
Contrainte	<b>Juste</b>	<b>5, simple</b>	<b>72 %</b>
	<i>Seulement</i>	9, simple	13 %
	<i>Ne ... que</i>	5, composé	13 %
	<i>Rien ... que</i>	7, composé	1,5 %
Conséquence	<i>Alors</i>	5, simple	44,8 %
	<b>Donc</b>	<b>4, simple</b>	<b>51,7 %</b>
	<i>Fait que (fac, fack et faque)</i>	3 à 5, simplifié	3,4 %
Futur	<b>Analytique</b>	<b>Temps composé (plus long)</b>	<b>63,6 %</b>
	Synthétique	Temps simple (plus court)	36,4 %
Négation	<b>Ne...pas (plus, rien etc.)</b>	<b>5+, composé</b>	<b>53 %</b>
	<i>...pas</i>	3+, simple	47 %
Interrogation	<i>Est-ce que</i>	9, composé	5,4 %
	Inversion	Variable	12,1 %
	<b>Sans inversion</b>	<b>Variable</b>	<b>76 %</b>
	Suffixe -tu	2	6,3 %
Alternance <i>On-Nous</i>	<b>On</b>	<b>2</b>	<b>88 %</b>
	<i>Nous</i>	4	12 %
Pronom simple ou composé	<b>Simple</b>	<b>3 à 4</b>	<b>99 %</b>
	Composé	9 à 10	1 %

Les variantes les plus fréquentes sont en caractères gras. On constate que pour les marqueurs lexicaux de contrainte et de conséquence et les alternances pronominales *on-nous* et simple-composé, ce sont les variantes plus courtes qui sont les plus fréquentes. Ensuite, pour les formules interrogatives, la variante la plus fréquente est la forme sans inversion ; on pourrait avancer que c'est celle qui requiert le moins d'ajouts ou de transformations comparativement aux trois autres formules, puisque l'ajout du point d'interrogation en fin de phrase suffit ; ce serait la « plus économe du point de vue des réalisations linguistiques » et du travail cognitif qu'elle appelle (Guignard, 2011 : 11) puisqu'elle réduirait « l'effort mental ou physique » (Fafa, 2017 : 37).

Ce sont donc uniquement pour les variantes de l'emploi variable du futur et de la négation que la variante la plus simple ou la plus courte n'est pas la plus fréquente. En revanche, comme je l'ai suggéré dans le Chapitre 4, la préférence du futur analytique pourrait s'expliquer par une plus petite charge cognitive. Cette conjugaison verbale comprend moins d'exceptions, puisqu'il suffit de connaître la conjugaison du verbe *aller* au présent de l'indicatif, alors que pour le futur synthétique il faut connaître les terminaisons du futur synthétique en plus des exceptions des racines de certains verbes. Rappelons aussi le grand nombre d'occurrences de la négation sous forme fragmentée, qui ont été exclues de cette analyse – cela indiquerait aussi le désir de rapidité chez ces internautes. Bref, à part la négation, les variantes les plus fréquemment employées par les internautes au Québec sur Facebook, sont aussi les plus courtes ou les plus simples, suivant le principe de l'économie du langage. Ces résultats portant sur l'influence du principe de l'économie du langage dans des contextes numériques s'alignent sur ceux de Thurlow (2003), Luckhardt (2009), et Blondeau *et al.* (2014). Ainsi, pour ce qui est de la première hypothèse, la prépondérance du standard écrit est donc réfutée, mais le principe de l'économie du langage tient.

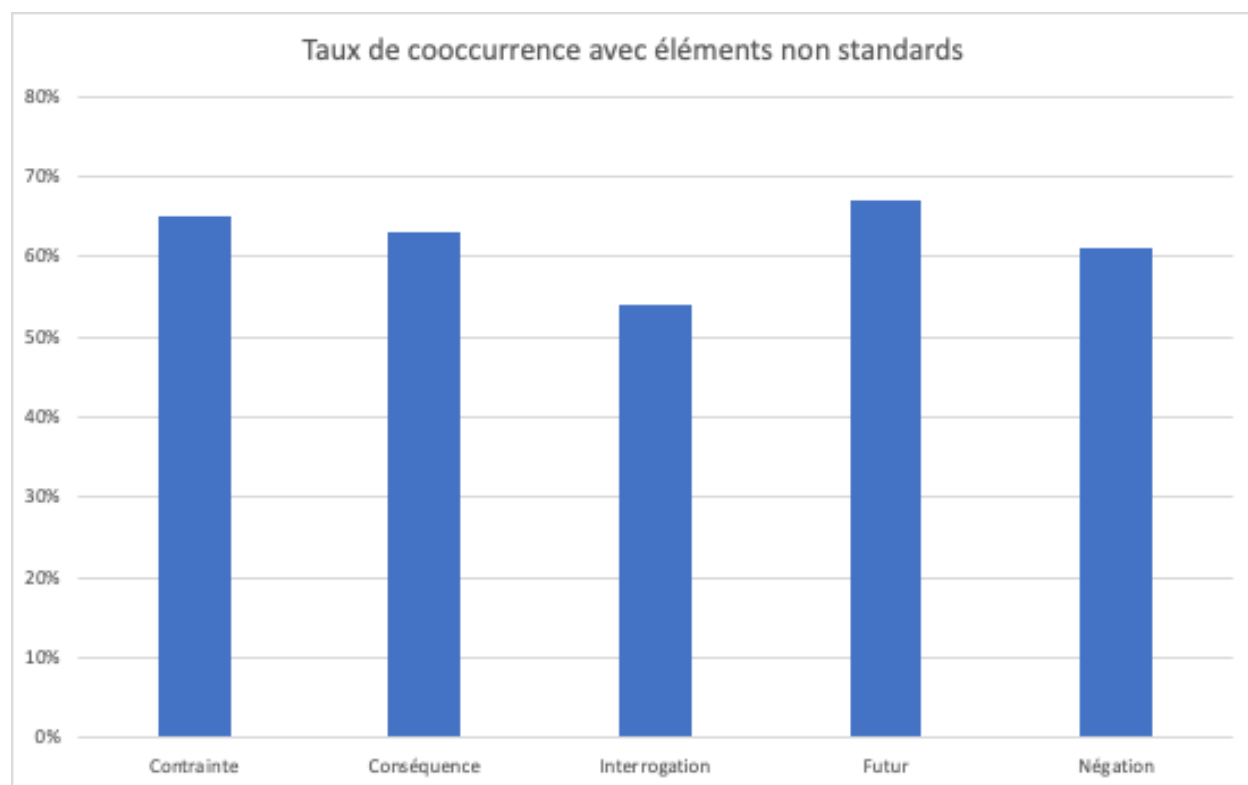
## **6.2 Synthèse et discussion des résultats de collocation**

Examinons maintenant les résultats des analyses exploratoire de collocation d'éléments linguistiques et iconiques avec les marqueurs lexicaux de contrainte et de conséquence, les formules interrogatives, l'emploi variable du futur et la négation. Le Tableau 6.2 et la Figure 6.2 montrent, pour chaque variable, c'est-à-dire toutes les variantes à l'intérieur de la variable, le nombre total d'occurrences qui sont accompagnées d'autres éléments non standards dans un même commentaire.

**Tableau 6.2 Pourcentage d’occurrences accompagnées d’éléments non standards**

Variable	Taux de cooccurrence avec formes non standards
Contrainte	65 %
Conséquence	63 %
Interrogation	54 %
Futur	67 %
Négation	61 %

**Figure 6.2 Pourcentage d’occurrences accompagnées d’éléments non standards**



En général, les commentaires contenant ces cinq variables contiennent des éléments non standards dans plus de 50 % des cas ; alors que les 374 occurrences du futur synthétique et analytique sont accompagnées le plus souvent d’éléments non standards, à une fréquence moyenne de 67 %, les 207 formules interrogatives le sont le moins souvent, à une fréquence moyenne de 54 %. En particulier, ce sont les variantes relevant de l’oral – notamment *juste, rien que, fait que,*

l'interrogation sans inversion et en suffixe *-tu*, le futur analytique, et l'omission du *ne* – qui sont accompagnées d'éléments non standards plus souvent que les autres variantes, à une fréquence moyenne de 78 %. En outre, même les variantes relevant du standard écrit sont accompagnées d'éléments non standards à une fréquence moyenne de 50 %.

Somme toute, les internautes se servent très souvent d'éléments non standards si l'on en juge par cet échantillon d'écrits numériques. Il s'agit donc d'une production langagière qui contient « de nombreux exemples de français vernaculaire intégrant des traits [du français de référence (autrement dit le français standard)], et ce, à tous les niveaux de la grammaire » (Blondeau et Tremblay, 2022b : 136). Cela suggère que « la grammaire du français vernaculaire [est] perméable à l'influence » du français standard (Blondeau et Tremblay, 2022b : 137), et renforce d'autant plus cette notion d'hybridité de la CNE.

Revenons à l'hypothèse présentée au Chapitre 3, au sujet de la variabilité dans les commentaires individuels. J'avais postulé que la production langagière et le degré de variation sociolinguistique allaient différer de commentaire en commentaire, mais que la majorité des internautes allaient employer le standard écrit. D'après la synthèse comparative présentée dans la section 6.1, et la fréquence de formes non standards du corpus, telle que présentée dans le Tableau et la Figure 6.2, ce sont en fait une minorité de commentaires qui contiennent systématiquement le standard écrit, car la majorité des commentaires examinés contiennent un mélange d'éléments standards et non standards. Nous constatons que souvent même les internautes qui choisissent d'employer les variantes hyper-formelles comme le *ne...que* incorporent aussi du lexique du vernaculaire, des émojis, de la ponctuation insolite, et omettent les accents. Ces conclusions s'alignent sur celles de David Crystal, qui indique que le contenu textuel des blogues et des courriels peut varier de manière significative d'un individu à l'autre, par rapport aux normes de la langue standard écrite (2011 : 21). Ce principe s'applique alors également à la CNE produite dans l'espace public de Facebook par des internautes au Québec.

Des analyses quantitatives plus pointues pourraient cerner la concentration d'éléments non standards par commentaire, et une étude suivant des volontaires qui indiqueraient par quel instrument d'entrée et quel moyen de saisie ils produisent leur CNE pourrait nous renseigner quant à l'effet de l'appareil et du type de clavier sur le contenu linguistique.

### 6.3 Un corpus québécois

Revenons sur la nature québécoise du contenu textuel de ce corpus extrait de communications publiques sur la plateforme Facebook. D'après Martel et Cajolet-Laganière (1996 : 125), « [i]l y a une manière québécoise d'utiliser le français », et ce corpus en témoigne. Il s'agit d'un corpus qui confirme la présence de traits de l'oral, et plus précisément une fréquence non négligeable de certains marqueurs sociolinguistiques du parler familier. Les analyses de concordance ont attesté de la présence de chacune des variantes pour chaque variable à l'étude (sauf le marqueur de conséquence *so*, qui est plus typique du français parlé au Canada en dehors de la province du Québec). Comme le soulèvent Blondeau et Tremblay, « on remarque très souvent [dans les SMS] les mêmes variantes qu'à l'oral et, par conséquent, une étendue beaucoup plus large du répertoire de variantes [qu'en français standard ou de référence] » (2022b : 134). En revanche, la présence de certaines d'entre elles est moins conclusive, comme l'emploi variable des auxiliaires *avoir* et *être* au passé composé, et la rare présence des pronoms composés en *-autres*. De même, la fréquente présence inattendue de la particule *ne* de la négation représente un autre résultat difficile à expliquer. Impossible alors de tirer des conclusions catégoriques, mais impossible aussi de nier que cette CNE est imprégnée de traits du français parlé au Québec.

En ce qui concerne l'état de cette variété de langue, le gouvernement du Québec et les organismes de protection linguistique tels que l'Office québécois de la langue française (abrégé OQLF) s'acharnent à promouvoir la langue française au Québec et au Canada, et se soucient de l'emploi de plus en plus dominant de l'anglais surtout dans les communications numériques puisque « l'échange de contenus numériques, comme l'information ou la culture, et la mondialisation des échanges commerciaux intensifient l'utilisation de la langue anglaise » (Gouvernement du Canada, 2023). En particulier, dans son plan stratégique, l'OQLF note les « répercussions importantes dans le paysage linguistique québécois » résultant du numérique, et souligne qu'il est « essentiel de

promouvoir l'utilisation du français » (Gouvernement du Québec, 2020 : 2). Les résultats de mon étude s'insèrent donc dans cet effort de surveiller l'état et l'évolution de cette variété du français. La présente étude démontre que cette variété de langue, alors qu'elle est « en situation de vive concurrence » avec l'anglais au Canada, (Woehrling, 2001 : 196) est employée de manière importante par les internautes dans une situation de communication publique qui n'est pas modérée par une autorité linguistique. On peut donc supposer que le français québécois est en bonne santé sur le plan socioculturel et sociotechnonumérique.

### **6.3.1 Le numérique et les langues minoritaires**

Cette recherche fait ainsi la lumière sur les pratiques langagières des internautes au Québec, et suggère qu'ils ou elles voient la valeur d'employer leur langue dans le numérique, un lieu qui est souvent dominé par l'anglais. Elle montre aussi, comme l'ont fait d'autres études dans le domaine (Blondeau et Tremblay, 2022 ; Ferrara *et al.*, 1991 ; Tremblay *et al.*, 2019), que les normes communautaires sont reflétées dans la CNE en ligne et que les échanges sur les sites de réseautage social ouvrent une nouvelle possibilité de vitalité et de florissement des langues minoritaires (Jones, 2013 ; Prado, 2016). En effet, alors qu'auparavant, « Internet était perçu comme servant uniquement à répandre le pouvoir d'une langue dominante aux dépens des langues en danger [...] il est devenu clair que les réseaux socionumériques permettent l'épanouissement d'une nouvelle génération d'utilisateurs et d'apprenants de la langue, ainsi contribuant au succès d'inverser le glissement linguistique » (Jongbloed-Faber *et al.*, 2017 : 174, ma traduction). Ces espaces numériques créent du réseautage qui fait renaître les communautés linguistiques des langues minoritaires (Crystal, 2000 ; Cunliffe, 2007, 2009) en contribuant à la transcription du vernaculaire à l'écrit (Blondeau et Tremblay, 2022b). En somme, d'après Hassan Atifi, « [e]n rupture avec la situation orale, le média électronique écrit permet donc l'émergence, la mixité et l'alternance de plus en plus de langues ou dialectes minoritaires » (2007 : 124). Le réseau socionumérique Facebook ouvre ainsi un genre de place publique virtuelle permettant à l'internaute au Québec d'« écrire son vernaculaire » (Blondeau et Tremblay, 2022b).

### 6.3.2 Le numérique comme lieu d'échange linguistico-culturel

Le contexte linguistique engendré par le numérique, qui « offre une grande liberté de création aux usagers » comme c'est le cas pour les SMS, (Cougnon, 2008), représente « un tremplin vers les jeux sur la langue » (Cougnon, 2010) où « le recours à une langue régionale tiendrait une place prépondérante » (Cougnon, 2008). Plus précisément, les discussions publiques sur les publications des médias journalistiques sont une forme de discours public vernaculaire qui peut relever du langage standard, mais qui est particulièrement propice à de multiples variétés de traits sociolinguistiques, en plus de caractéristiques du parler quotidien (Johansson, 2017). En outre, Josh Iorio (2015) souligne que l'on peut voir un site de réseautage social tel que Facebook comme un endroit riche en pratiques du vernaculaire, puisque lecture et écriture sont volontaires et autonomes. Il prend comme exemple les publications par des organismes journalistiques, écrites suivant les normes journalistiques et prescriptives, mais commentées par les internautes suivant leurs propres pratiques d'écriture du vernaculaire. Le contexte de la CNE semble effectivement représenter un espace de pratiques du vernaculaire (Blondeau et Tremblay, 2022b), une source de créativité, d'invention et d'originalité (Barton et Lee, 2012) ou encore, un lieu d'échange linguistico-culturel, qui pourrait être important pour la survie et la vitalité continue de cette variété de langue.

L'importance du numérique pour le développement culturel au Québec ne s'applique pas seulement à la langue. Par exemple, les échanges *peer-to-peer* facilités par le Web sont régulièrement utilisés « au Québec pour transmettre des œuvres du patrimoine national (musique, films, livres, émissions de télévision, etc.), dont plusieurs sont non disponibles autrement ». De même, d'autres ont noté « le rôle structurant joué par le numérique dans le développement du rap québécois [...] une culture en marge de l'industrie ». Plus récemment une étude transversale de la découverte de contenus culturels en ligne au Québec a trouvé que « les supports numériques [y compris les médias socionumériques] font partie intégrante des processus de découverte de contenus culturels d'une grande majorité de participant.e.s » (Poirier *et al.*, 2023 : 262). Enfin, l'Organisation internationale de la Francophonie souligne « les potentialités offertes par le numérique pour produire, diffuser, promouvoir et valoriser les contenus francophones tout en

contribuant au développement de nouveaux modes d'expression » (OIF, 2015 : 3). Le numérique est alors clairement un lieu important d'échanges et d'innovations linguistiques et culturelles.

### **6.3.3 Le numérique comme lieu de valorisation du parler régional**

Les productions langagières dans les espaces publics numériques sont importantes pour plusieurs raisons. Tout d'abord, Florian Coulmas (2013) note que cette technologie a facilité l'écriture de langues qui restaient non écrites ; d'après David Barton et Carmen Lee, ces pratiques vernaculaires dans le Web 2.0 ont plus de valeur que dans d'autres sphères (2012). En effet, de manière traditionnelle, le vernaculaire possède moins de valeur sociale que les variétés linguistiques standards et dominantes qui sont appuyées par l'éducation et les institutions. Mais dans les communications en ligne, les pratiques vernaculaires et régionales prennent de l'ampleur, surtout en ce qui concerne le lien entre le vernaculaire, l'identité et l'auditoire (Barton et Lee, 2012). Ensuite, dans les tendances de l'évolution linguistique, lorsqu'une langue est écrite, elle devient importante, légitimée (Coulmas, 2013). Le penseur Pierre Bourdieu ne soutenait-il pas que « [l]e code [...] régit la langue écrite [...] par opposition à la langue parlée » qui est « implicitement tenue pour inférieure » (Bourdieu, 2001 : 76) ? D'après lui, la langue légitime, standard, écrite, a donc plus de *capital linguistique*. En effet, l'écrit est un discours avec du poids, qui ne peut pas être questionné ou réfuté directement, contrairement à l'oral. L'écrit, tout en élargissant le potentiel du langage, convertit des dialectes oraux en dialectes graphiques, et donne donc à une forme scripturale un pouvoir bien plus important qu'un dialecte oral (Ong, 1982). Par exemple, dans le cas de l'histoire de l'anglais, l'invention de la presse d'imprimerie a mené à la préservation, à la codification, et parfois même à la création et à l'enrichissement de langues vernaculaires (Eisenstein, 1980).

Ainsi, tout comme l'émergence de la presse d'imprimerie a engendré une transition vers le vernaculaire de manière extrêmement décentralisée (Bernstein, 2013), l'émergence des technologies de la communication permettent une nouvelle « *explosion typologique* qui étend les esprits et les voix des individus pour reconstituer le dialogue humain » (McLuhan, 1994, ma traduction). Enfin, avec le numérique, la frontière entre le standard et le non standard, et le vernaculaire et l'institutionnel, deviendra de plus en plus floue (Iorio, 2015).

À la lumière de ces réflexions sur l’oral, l’écrit et la légitimité linguistique, quel sera l’avenir du français québécois vernaculaire, qui témoigne d’autant de vitalité et de variation dans le Web 2.0 ? Les internautes au Québec qui écrivent leur vernaculaire dans les espaces publics du Web 2.0 sont-ils et elles en train de légitimer leur parler régional ? Dans son étude portant sur les souhaits d’anniversaire sur Facebook, Laetitia Bibié-Emerit (2015 : 144) soutenait que les internautes, « [e]n réinventant leur langue, [...] se la sont appropriée : ils l’ont “recréée” et s’en sentent responsables ». Peut-on soutenir cette même thèse dans le cadre de la CNE produite dans l’espace public de Facebook par les internautes au Québec ?

#### **6.4 Les axes de la variation linguistique**

Revenons à quelques questions théoriques posées dans le premier chapitre autour des axes traditionnels de la variation linguistique. Il y a notamment la question de la diamésie, qui est vue de manière traditionnelle comme étant la variation inhérente des codes oral et écrit, utilisés pour communiquer. Étant donné la nature hybride de la CNE, comme le montrent les résultats de la présente étude, la question avancée par Jakob Wüest (2009) quant à la nécessité de la notion de diamésie mérite d’être posée à nouveau.

##### **6.4.1 Le numérique et la question de la diamésie**

Pour reprendre la définition d’Albert Mioni, la diamésie est « l’opposition entre l’écrit et le parlé » sous forme de « continuum de degrés intermédiaires » (Wüest, 2009 : 147, citant Mioni, 1983 : 508). Wüest soulève également le fait qu’à l’oral, on distingue différentes situations de communication selon la spontanéité et le nombre d’interlocuteurs, mais « [q]uand on passe à l’expression écrite, la perspective change. Il ne s’agit plus de la situation de communication dans laquelle le texte est produit, mais de la question de savoir si le texte est conçu pour être lu ou non » (Wüest, 2009 : 147). Une situation orale est soit monologique, soit dialogique, tandis qu’à l’écrit, un texte peut être écrit sans même que son auteur sache d’avance si son texte sera lu. Par exemple, dans le cas d’un commentaire public écrit sur une plateforme socionumérique, il se peut que le

message soit lu par plusieurs inconnus, ou pas lu du tout. À l'oral, on sait à qui on s'adresse ; ceci n'est pas toujours le cas à l'écrit.

Dans le cadre de la CNE, la notion du médium est aussi importante que celle de la situation. Les paramètres communicatifs de l'oral ne sont pas identiques à ceux de l'écrit ; la grande « différence entre la langue parlée spontanée et les textes écrits [est que] la communication spontanée connaît ses propres règles, comme l'organisation des tours de parole, [mais] elle ne se coule pas dans le moule de types de texte » (Wüest, 2009 : 155-156). En outre, dans le cas de la CNE, les paramètres technologiques et sociaux, tels que la plateforme utilisée, la nature privée ou publique du contexte, etc., auront également un impact sur la situation de communication, et ainsi sur les choix langagiers des internautes.

C'est autant la plateforme que la situation qui a un effet sur la production langagière. D'une part, les « plateformes de réseautage social ont créé un nouveau moyen de communication avec ses propres contraintes et affordances, façonnant le comportement langagier et créant de nouveaux sous-registres » (Friginal *et al.*, 2017 : 342, ma traduction). D'autre part, la situation aura un impact sur la CNE : « la rapidité des échanges [...] va amener les écrits numériques à se rapprocher plus ou moins de l'écrit ou de l'oral » (Marcoccia, 2016 : 72). Par exemple, « en situation de communication asynchrone, les participants disposent de temps pour produire le message, peuvent le modifier et l'améliorer » (54). On voit alors dans le cadre des communications en ligne, que les notions de diamésie, relevant du médium, et de diaphasie, relevant de la situation, sont étroitement liées.

De plus, ces contraintes et affordances vont également permettre le choix de code langagier, oral ou écrit. Sur une plateforme comme Snapchat, par exemple, on aurait plus tendance à communiquer par moyen audio-visuel, tandis qu'un site tel que Reddit ne permet que de l'écrit, avec l'option de partage d'URL vers une vidéo. Sur Internet, à l'oral, les internautes ne vont pas se servir d'émojis : l'oral est auditif et phonétique, l'écrit est visuel et graphique. À l'oral, du sens est transmis à travers la prosodie, à l'écrit le sens n'est transmis que de manière iconique. Jakob Wüest résume cette question en soulignant sa complexité :

Or, la communication orale est au fond d'une nature plus complexe que la communication écrite [...] L'opposition entre l'écrit et le parlé est complexe. Elle relève en partie de la diaphasie et en partie des conditions de production propres aux deux codages, ainsi que, dans une moindre mesure, du codage graphique lui-même (Wüest, 2009 : 156-159).

La distinction entre oral et écrit en ligne demeure alors importante ; ainsi, la notion de diaphasie est nécessaire, mais elle ne peut pas englober la notion du médium technologique.

#### **6.4.2 La médiatisation technologique et la variation linguistique**

En ce qui concerne l'idée d'une nouvelle forme langagière, soulevée par plusieurs chercheurs dans le domaine (Blondeau et Tremblay, 2022b ; Crystal, 2011 ; Marcoccia, 2016), elle implique que ce qui s'écrit en ligne se distingue des autres formes langagières. Or, les résultats de la présente étude montrent que la CNE dans l'espace public de Facebook n'est pas si différente de celle des SMS et du (t)chat – du moins au niveau des marqueurs sociolinguistiques à l'étude. Dans le même ordre d'idées, « plusieurs phénomènes linguistiques des textos s'apparentent à ceux observés en contexte de clavardage » (Tatossian, 2020: 109). Cela suggère une influence moins importante dans les différences de plateforme et de situation, et une plus forte influence de la médiatisation technologique englobant ces facteurs.

Pour éclaircir la question, revenons au schéma des paramètres technologiques et sociaux des communications en ligne, présenté dans le Tableau 6.4.2. À part quelques caractéristiques sociales, la plupart sont le résultat de la médiatisation technologique comme celles de lecteur, de réponse, d'environnement social et technologique, et de toutes les caractéristiques technologiques.

**Tableau 6.4.2 Les caractéristiques sociales et technologiques de la CNE**

Caractéristiques sociales	Caractéristiques technologiques
Scripteur	Instrument d'entrée
Lecteur, auditoire, cadre participatif	Support lecture
Lecteur réponse	Code
Répondeur	Médium, canal
But des messages, objectifs	Mode de transmission
Sujets, thèmes	Temporalité, synchronicité
Ton des messages	Persistance du message, adaptations
Environnement social	Environnement technologique
Normes sociales, linguistiques	

En fin de compte, c'est la médiatisation technologique qui opère surtout sur les paramètres diaphasique de la situation et diamésique du choix de code. Je propose ainsi une modification aux axes de la variation linguistique, pour rendre compte de la dimension technologique influant la diamésie et la diaphasie : il faudrait en effet un axe technologique qui expliquerait les caractéristiques communes aux différentes productions selon la modalité et la situation, puisque c'est une adaptation aux affordances et aux contraintes technologiques

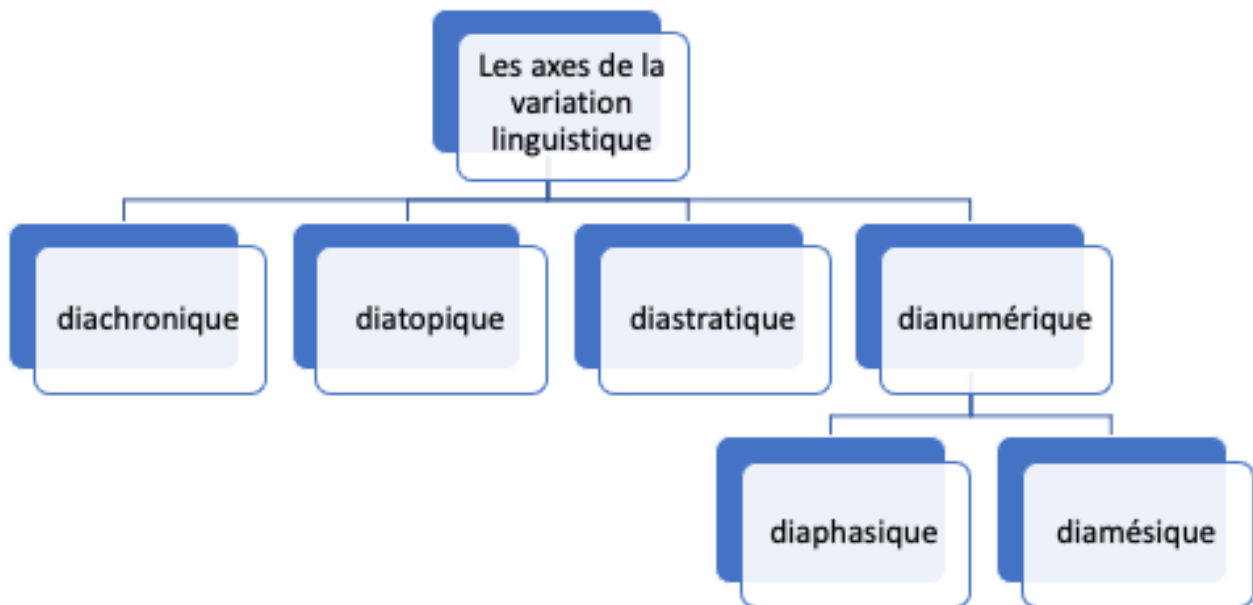
La Figure 6.4.2 représente le schéma traditionnel des axes de la variation linguistique, c'est-à-dire les facteurs externes, comme l'époque (diachronie), les caractéristiques du sujet parlant (diastratie), la géographie (diatopie), le code (diamésie) et la situation (diaphasie), influant sur la production langagière. Cette configuration rend compte de la diamésie traditionnelle, qui fait la distinction entre langage oral et langage écrit, mais ne permet pas de prendre en considération l'impact du numérique et des différences de médium technologique.

Figure 6.4.2 Les axes de la variation linguistique



La Figure 6.4.2.1 ci-dessous présente un schéma alternatif, s'appliquant à l'analyse de la langue produite lors de communications médiatisées par la technologie, qui incorpore un axe *dianumérique*, avec les axes diamésique et diaphasique découlant du premier.

Figure 6.4.2.1 Les axes de la variation linguistique dans l'ère du numérique



Je propose que cette deuxième version du schéma des axes de la variation linguistique illustrée dans la Figure 6.4.2.1, clarifie la notion de diamésie, en faisant la distinction entre le médium langagier, donc l'axe diamésique, et le médium technologique, donc l'axe *dianumérique*.

L'axe *dianumérique*, c'est-à-dire les contraintes et les affordances des appareils et des plateformes technologiques, aura surtout un effet sur le choix de code langagier (oral ou écrit) ainsi que sur son degré d'hybridité, y compris l'emploi d'éléments iconiques. Ensuite, l'axe *dianumérique* aura aussi un impact sur l'axe diaphasique, puisque souvent la technologie en question entraîne une situation de communication précise. Ainsi, ce schéma rend compte de l'impact des paramètres technologiques sur la production langagière en situation numérique. Quelle sera alors l'évolution de ces productions langagières en ligne, étant donné la rapidité des développements technologiques tels que l'intelligence artificielle et le rôle de plus en plus important des algorithmes dans nos interactions dans l'espace numérique ?

### **6.5 Limites de l'étude**

En plus des limites méthodologiques mentionnées dans le Chapitre 3, il importe de soulever certaines limites sur le plan du corpus, qui auraient pu avoir un impact sur les données et sur la comparaison des contextes oral et numérique. Tout d'abord, le corpus ne peut pas englober la production écrite de chaque internaute au Québec sur Facebook. Le corpus lui-même est alors limité par sa taille et par le temps. J'ai tenté d'atténuer cette limite en faisant l'extraction de données écrites provenant de cinq pages Facebook distinctes, mais la taille de chaque composante du corpus est inégale. Les données de cette étude ne permettent donc pas de tirer des conclusions générales sur la population au Québec, mais plutôt sur le comportement de ces internautes, qui, dans cette situation précise sur Facebook, contribuent à ce corpus par leur production langagière. Un corpus plus vaste aurait pu révéler des données plus significatives pour certains marqueurs peu fréquents, tels que les marqueurs lexicaux, l'emploi du passé composé, et l'emploi des pronoms composés.

Les résultats qui font l'objet de la comparaison des contextes oral, SMS, (t)chat et Facebook ont eux aussi leurs limites. D'une part, les quatre contextes ont des paramètres situationnels et sociaux différents, ce qui fait qu'une comparaison des données ne peut qu'être approximative. D'autre part, les méthodes de cueillette des données peuvent différer de contexte en contexte, ce qui peut également diminuer le degré d'exactitude de la comparaison. Par exemple, les données de ces quatre contextes n'ont pas été recueillies en même temps, et il existe en fait de grands décalages

temporels entre certaines données de l'oral et les données de contextes numériques. Ainsi, quelques-unes des séries de données de l'oral sont récentes, notamment pour les variantes de l'interrogation (Halvorsen, 2021), les marqueurs de conséquence (Blondeau *et al.*, 2019 et Blondeau et Tremblay, 2022a), et l'emploi des auxiliaires *avoir* et *être* (Réa, 2020), mais la plupart de ces données datent de plus de dix ans : il s'agit des données portant sur les marqueurs de contrainte (Massicotte, 1986 ; Thibault et Daveluy, 1989), sur l'emploi variable du futur (Blondeau, 2006), sur la négation (Sankoff et Vincent, 1977), sur l'alternance *on-nous* (Laberge, 1977) et sur l'emploi des pronoms non clitiques (Blondeau, 2000). Le même genre de situation se présente du côté des données SMS et (t)chat : les deux corpus, notamment le corpus SMS *texto4science* (Langlais et Drouin, 2012) et le corpus (t)chat créé par Rémi van Compernelle (2010), datent déjà de plus de dix ans eux aussi. Il serait alors préférable d'effectuer une comparaison avec des données plus récentes.

Il importe de rappeler également que les données de l'oral et celles des études sur le SMS ont été recueillies auprès de sujets parlants et scriptants qui se sont portés volontaires, tandis que les données (t)chat et les données Facebook proviennent de contenus publiquement accessibles. Dans le cas des données de l'oral et des SMS, l'effet de l'observateur pourrait alors jouer un rôle dans la production orale et écrite. En revanche, la nature anonyme du présent corpus a éliminé la possibilité de faire des analyses stratifiées par âge, sexe, statut socioéconomique et de connaître l'appareil utilisé par l'internaute dans la production de sa CNE. Par conséquent, cette étude ne peut pas, par exemple, déterminer si ce sont les femmes ou les hommes qui choisissent d'utiliser une forme plutôt qu'une autre, ni l'impact du niveau d'éducation ou du statut socio-économique sur le choix de niveau de langue (Labov, 1976).

Rappelons enfin que l'objectif principal de cette étude pionnière était d'attester des marqueurs de l'oral, et de déterminer leurs fréquences d'emploi sur Facebook : cette étude a effectivement montré que des marqueurs de l'oral à travers le système linguistique sont employés par ces internautes, et que certaines fréquences d'emploi ressemblent aux fréquences dans d'autres contextes à l'oral et dans le numérique.

## 6.6 Conclusion

Ce chapitre a discuté de manière synthétique et plus théorique des résultats, de plusieurs enjeux découlant des résultats, et des limites possibles de la présente étude. À la lumière des résultats de la présente étude, plusieurs questions théoriques se posent, par rapport au rôle du numérique tant dans l'emploi et la valorisation potentielle des langues minoritaires et régionales, que dans les dimensions de la variation linguistique. Voilà pourquoi il sera important à l'avenir d'élargir ce genre de recherche à encore d'autres contextes de la communication numérique écrite tels que X/Twitter, YouTube et Reddit, afin de mieux comprendre l'impact des espaces numériques sur l'évolution langagière. En effet, la transformation numérique a créé une démocratisation de la communication et « l'amplification de la performance individuelle ». Les « internautes [...] peuvent désormais se mettre en scène et faire spectacle avec l'illusion de la liberté » (Ben Amor et Granget, 2011 : 114).

Plus largement, les résultats de la présente étude des écrits numériques sur Facebook d'internautes au Québec viennent étayer un certain nombre de notions théoriques en (socio)linguistique. Tout d'abord, nous avons vu que la communication numérique écrite sur Facebook par des internautes au Québec témoigne de maintes marques d'oralité, ce qui suggère la primauté de l'oral. En effet, Marshall McLuhan (1994), bien avant l'ère du numérique, avait prédit que grâce à sa capacité d'étendre notre système central nerveux, la technologie électronique semblait favoriser la parole inclusive et participative par rapport à l'écrit spécialisé. Le numérique nous permettrait alors de revenir à nos origines langagières et de nous lancer dans de nouvelles pratiques de l'oralité, dans ce contexte, transcrites à l'écrit.

Ensuite, les résultats de cette étude s'alignent plutôt sur les pratiques langagières numériques dans d'autres langues, en ce qui concerne la variation et l'hybridité de codes et de styles ; c'est-à-dire que ce qui se passe en ligne en français au Québec ressemble à ce qui se passe sur le plan mondial. Par exemple, nous avons vu des similarités avec les pratiques attestées en anglais par Tagliamonte et Denis (2008), Friginal *et al.*, (2017), ainsi que des pratiques dans d'autres langues (Luckhardt, 2009 ; Siebenhaar, 2008 ; Zhang 2020) et d'autres variétés du français (Blattner et Williams, 2011 ; van Compernelle, 2008 ; Stark, 2012). Ces ressemblances entre les pratiques dans diverses

langues en ligne renforcent l'idée que la variation est inhérente à l'emploi de la langue par une communauté ; en effet, d'après Labov, l'absence de systèmes de communication complexes et de variation stylistique seraient dysfonctionnelle (1973 : 203).

Enfin, ces ressemblances de pratiques numériques à travers les langues du monde viennent soutenir l'idée d'une universalité de pratiques langagières. Ainsi, tandis que nos habitudes pourraient s'adapter aux transformations technologiques opérant sur la communication, comme il existe une régularité et même une prévisibilité dans la variation linguistique (Labov, 1976), il se pourrait que cette régularité et cette prévisibilité s'appliquent plus largement à nos manières d'adapter nos comportements linguistiques communautaires à travers le temps.

## CONCLUSION

### Conclusions générales

Cette recherche, une première sur le plan du terrain étudié, notamment Facebook, et sur le plan de l'étendue des données analysées, avait deux buts principaux. Tout d'abord, il s'agissait d'attester de la présence et de la fréquence d'emploi relative de marqueurs de l'oral dans la CNE sur Facebook écrite par des internautes au Québec. Ensuite, il s'agissait de comparer ces résultats aux productions langagières dans d'autres contextes oraux et numériques. Ces démarches avaient pour but de déterminer le degré d'oralité et d'hybridité de cette forme linguistique utilisée dans l'espace public d'un réseau socionumérique. Tout en appliquant la théorie et la méthode sociolinguistique variationniste à l'étude du contexte socio- et technonumérique engendré par les nouvelles technologies de la communication, j'ai tenté de combler les lacunes des recherches antérieures sur le français québécois, qui s'étaient focalisées jusqu'à présent seulement sur deux terrains – les SMS et le (t)chat –, et seulement sur des éléments restreints du système linguistique.

Revenons aux résultats de cette étude : les analyses de concordance et de collocation des huit variables à l'étude ont dévoilé des résultats partagés. Pour ce qui est des marqueurs lexicaux de contrainte et de conséquence, qui sont les moins fréquents dans le corpus comparativement aux autres types de marqueurs analysés, les analyses de concordance dégagant les fréquences d'emploi suggèrent par exemple l'influence de l'anglais et celle de l'économie du langage. Du côté des marqueurs verbaux, les fréquences d'emploi des futurs analytique et synthétique s'alignent presque parfaitement sur celles qui ont été relevées dans les SMS, et se rapprochent des tendances de l'oral. En revanche, celles de l'emploi variable des auxiliaires du passé composé ne permettent pas de tirer de conclusions sur leurs fréquences, en raison du petit nombre d'occurrences. Ensuite, l'analyse des marqueurs syntaxiques suggère que les fréquences d'emploi des formules interrogatives sont stables, à travers différents contextes, tandis que les résultats portant sur la négation montrent de grands écarts entre les fréquences sur Facebook comparativement à d'autres contextes. Enfin, les résultats indiquent que les fréquences d'emploi des deux variables pronominales, notamment l'alternance *on-nous* et l'alternance des pronoms

simples et composés, suivent de près les tendances communautaires. Pour ce qui est des analyses de collocation, elles montrent que les marqueurs relevant de la langue non standard sont accompagnés plus fréquemment d'autres éléments non standards dans un même commentaire que les marqueurs standards. De plus, toutes les variantes des cinq variables pour lesquelles les analyses de collocation ont été effectuées sont accompagnées d'éléments non standards.

Les résultats de cette recherche suggèrent d'une part que malgré la différence de situation de communication sur Facebook et dans les SMS, les internautes s'engagent dans les mêmes genres de pratiques linguistiques. Ainsi, l'importance des tendances communautaires dans les choix langagiers est renforcée. D'autre part, le caractère non concluant de certains résultats de cette étude montre que plus de recherches sont nécessaires afin de comprendre la motivation des internautes dans leurs choix linguistiques en ligne.

Pour résumer, ce corpus de CNE produite par des internautes au Québec dans l'espace public de Facebook témoigne non seulement de variations sociolinguistique et sociostylistique, mais également d'innovations des pratiques linguistiques qui remettent en question la dichotomie traditionnelle entre l'oral et l'écrit. En effet, les fréquences d'emploi des éléments à l'étude sont similaires à celles de l'oral et de la production écrite dans d'autres contextes numériques, mais ne s'alignent pas exactement sur l'un ni sur l'autre. En outre, les écrits sur Facebook ne suivent pas non plus à la lettre le français québécois standard écrit. Il s'agirait donc d'une pratique langagière distincte et hybride, mélangeant des variantes sociolinguistiques formelles et informelles.

Les fréquences d'emploi relevées sur Facebook semblent par conséquent suivre les tendances communautaires, comme l'ont suggéré d'autres chercheurs en sociolinguistique ayant étudié la langue dans des contextes numériques (Blondeau et Tremblay, 2022b ; Cunliffe *et al.*, 2013 ; McCulloch et Lamontagne, 2020 ; Siebenhaar, 2008 ; Tagliamonte et Denis, 2008). Ainsi, les pratiques langagières dans cette CNE ressemblent à celles que l'on constate dans d'autres variétés du français et dans d'autres langues, sur le plan de la variation (Atifi et Marcoccia, 2006 ; Friginal *et al.*, 2017 ; Herring, 2015 ; Storrer, 2014) et de l'hybridité oral-écrit (Bibié-Emerit, 2015 ; Ferrara *et al.*, 1991 ; Zhang, 2020). Par exemple, Michel Marcoccia précise que dans la CNE, on « trouve de nombreux procédés de verbalisation du matériau non-verbal et paraverbal, comme les

émoticônes ou les marques de ponctuation expressive » (2016 : 73), ce qui rejoint les données recueillies dans le cadre de la présente étude.

Plus généralement, on peut aussi conclure que la production écrite sur Facebook par des internautes au Québec représente un langage hybride, qui mélange des formes relevant du standard, aussi bien que du vernaculaire en plus d'autres éléments, tels que les émojis et les symboles. De plus, cette CNE dans un espace public n'est pas homogène : certains internautes suivent le standard écrit, d'autres non, et d'autres encore utilisent un mélange des deux. Il y a une diversité de pratiques langagières sur le plan individuel. En effet, étant donné que les fréquences sur Facebook divergent des fréquences attestées à l'oral de seulement 30 % environ, et que les internautes utilisent des formes non standards dans leurs commentaires plus de 50 % du temps, on peut conclure que cette CNE s'apparente plus à l'oral et au non standard qu'au standard écrit.

Les résultats de cette étude s'alignent alors sur les propos du linguiste anglais David Crystal (2011), qui a suggéré que les écrits dans le numérique ne ressemblent ni à la langue écrite, ni à la langue orale, mais plutôt à une nouvelle forme de langage. En effet, d'après Jacques David et Harmony Goncalves,

nous assistons au développement d'une nouvelle variété du français écrit, dont les objectifs évoluent vers une écriture immédiate, plus libre, désormais affranchie [...] mais aussi vers une écriture plus accessible [...], plus affective (par l'expression des sentiments, des émotions, de la corporéité du langage), plus inventive (par l'invention de néographies, de jeu[x] de mots et de signes) et plus socialisante [...] (2007 : 46).

Plus récemment, les chercheurs en sociolinguistique variationniste ont renchéri sur ce point. Mireille Tremblay (2020 : 96) suggère par exemple que la langue écrite par les Québécois.es dans le numérique, « [inspirée] de l'oral [vient] suppléer à l'écrit standard, qui apparaît inadéquat ou insuffisant dans ce contexte ». Pour sa part, Anaïs Tatossian (2020 : 121) note que ce contexte technonumérique « génère de nouveaux besoins communicationnels, qui exigent immédiateté et spontanéité de l'expression » et donc que les internautes tentent de « reproduire les paramètres d'une conversation en face à face ». Il s'avère alors que les pratiques écrites des internautes, transgressant souvent les règles du standard écrit, seraient le résultat d'un besoin communicatif et expressif.

En somme, ces résultats suggèrent que les pratiques langagières sur Facebook ne sont pas si différentes de celles qui sont avérées dans d'autres contextes numériques, ce qui valide les prémisses de la sociolinguistique par rapport à l'importance de la communauté linguistique dans l'évolution de l'emploi de la langue (Hymes, 2001), et par rapport au caractère fondamental de la variation linguistique inhérente à toute communauté linguistique (Labov, 1976). D'autre part, cela renforce la notion que les tendances communautaires l'emportent sur *le bon usage*, c'est-à-dire qu'il y aurait des *normes d'usage technonumérique*. Ainsi, les internautes « jouent avec la norme et s'en servent en dehors des cadres prescriptifs : la norme devient source de créativité » (Bibié-Emerit, 2015 : 334).

Les résultats de cette étude soulignent enfin la primauté de l'oral, qui se faufile dans la langue écrite. La communication numérique écrite n'est donc pas une écriture fautive ou une source de dégradation de la qualité de la langue faisant preuve de « difficultés d'expression, une grammaire et une orthographe fautives, des phénomènes d'hésitation dans la production, [...] un nivellement de la culture lié à un appauvrissement de la langue maternelle » (Marchal, 2014 : 4). Le numérique serait plutôt un contexte dans lequel les internautes peuvent pratiquer et performer leur « dextérité linguistique » (Deumert et Vold Lexander, 2013). Il s'agirait donc d'un endroit propice à la manipulation de l'outil langagier et à l'évolution linguistique.

### **Pistes futures de recherche**

Cette recherche ouvre la voie à une multitude de sujets d'étude potentiels. En fait, les résultats de cette étude suscitent plus de questions que de réponses, et fournissent un point de départ à de futures recherches en sociolinguistique et dans d'autres domaines tels que les communications, les sciences politiques, le marketing, la sociologie et la psychologie.

Sur le plan des objets d'études sociolinguistiques possibles, il serait intéressant de se focaliser sur l'effet du sujet de la publication sur la production langagière dans les commentaires en ligne. Par exemple, les recherches par l'équipe de Douglas Biber (2017) sur les écrits dans le Web 1.0 ont trouvé une corrélation entre le sujet du texte numérique et le style de langue utilisé. Ensuite, une approche plus ethnographique pourrait examiner l'importance des relations et des interactions

entre les internautes dans de tels espaces numériques et publics, sous l'optique de la théorie des types de réseaux sociaux (Milroy, 1987). Enfin, un plus grand éventail de sites journalistiques pourrait mener à un corpus plus vaste et plus riche, et à une analyse de régionalismes québécois.

Pour ce qui est de l'analyse plus granulaire de ces productions langagières, un certain nombre d'éléments linguistiques et iconiques qui n'ont pas pu être analysés dans le cadre de cette étude méritent de l'attention, notamment l'alternance codique, l'emploi des émojis, et d'autres marqueurs sociolinguistiques typiques du français québécois tels que l'élision du *l*. Par exemple, d'après les analyses de Jannis Androutsopoulos des pratiques langagières dans des forums (2007), sur Facebook (2014 ; 2015), et sur YouTube (Androutsopoulos et Tereick, 2016) chez des internautes multilingues en Allemagne, l'alternance codique est un phénomène qui révèle les motivations pragmatiques des internautes. En effet, le présent corpus est assez riche en alternance codique français-anglais, une pratique qui « serait un élément central du discours des bilingues » et abondant au Québec (Kärki, 2017, citant Nguyen, 2015 : 2).

Les extraits suivants tirés du corpus attestent de l'existence de tels procédés de *translanguaging* (Garcia et Otheguy, 2020) :

*c'est **obviously** polonais, donc pour une opinion objective sur la Russie on repassera ☺*  
(2581, La Presse)

*Un **shutdown** de cet envergure se chiffre en milliards.* (2590, La Presse)

*Et faire des meubles **cheap en presswood** tout le monde est capable.* (2596, La Presse)

*sauf si le **deal** était de payer dans une devise autre que la leur...* (2598, La Presse)

*j'hais ça du pas possible. J'suis **safe*** (1751, TVA Nouvelles)

***Good job** les propriétaires abusifs* (1165, Journal de Montréal)

*sa fait 25 ans que le monde dit que toute est trop cher dans vie mais **anyway** chialler est plus facile que de trouver des solutions* (1167, Journal de Montréal)

*pas juste les gaz ou vous penser que les bateaux **flush** leur toilettes* (1264, Journal de Québec)

*Tu vas sortir de là **bleaché!!!** (1314, Journal de Québec)*

*la bouche ouverte **of course** 🙄 (1352, Journal de Québec)*

Ces exemples montrent la diversité de contextes de l’alternance codique. Les internautes empruntent de l’anglais des noms comme *shutdown*, *deal* et *password* ; des verbes comme *flush* ; des adjectifs comme *cheap*, *bleaché* et *safe* ; des adverbes comme *obviously* et *anyway* ; et des expressions courantes comme *good job* et *of course*. C’est une question qui mérite plus d’attention : Katja Kärki (2017) et Jade Toivonen (2023) ont examiné l’alternance codique en français québécois écrit sur Facebook et YouTube, respectivement, et ce phénomène représente donc une piste de recherche féconde.

D’autres recherches pourraient étudier les attitudes des internautes québécois.es par rapport à l’emploi de la langue dans le Web 2.0, pour déterminer par exemple si les individus pensent qu’il est important de suivre, ou pas, le standard écrit. Comme je l’ai mentionné précédemment, une étude pointue, suivant l’approche de Pierre Halté (2020) sur l’emploi des émojis, pourrait éclairer les sociolinguistes sur les fonctions pragmatique, sémantique et syntaxique de ces signes, et une analyse comparative de l’emploi d’émojis dans différentes régions de la francophonie serait essentielle pour découvrir le rôle culturel de ces binettes.

L’étude de la langue en contexte numérique pourra aussi servir à avancer les connaissances dans d’autres domaines des sciences sociales. Tout d’abord, des recherches axées sur la communication pourraient se focaliser sur l’impact de l’auditoire (Bell, 1984), sur l’importance de la représentation de soi (Androutsopoulos, 2007 ; Georges *et al.*, 2009 ; Perea, 2010), et sur le contenu sociolinguistique et/ou sociostylistique de la CNE (Johnson, 2013). Par exemple, tandis que Monique Lebrun et Nathalie Lacelle (2012) ont examiné les choix langagiers comme geste identitaire chez de jeunes internautes québécois.es à partir d’un questionnaire, Ian Johnson, dans son étude sociolinguistique de l’emploi du gallois et de l’anglais sur Twitter, a trouvé que les internautes bilingues alternaient la langue selon l’auditoire (2013 : 114). Ainsi, il serait pertinent d’examiner cette problématique en puisant dans un corpus d’écrits numériques par des internautes au Québec.

Dans le domaine des sciences politiques, une étude de la relation entre la CNE et les attitudes politiques permettrait par exemple de comprendre l'importance des communications sur les réseaux sociaux lors des élections (Schwörer *et al.*, 2020 ; Svidronova *et al.*, 2019). Les chercheurs en marketing, qui se sont déjà intéressés à la production écrite des consommateurs et aux interactions entre clients et entreprises en ligne (Johann *et al.*, 2021), pourraient se pencher plus particulièrement sur les habitudes des internautes au Québec, au service des industries québécoises. Les sociologues pourraient se focaliser sur les attitudes des internautes envers la CNE qu'ils lisent en ligne (Lacerte, 2006) et sur leur propre comportement linguistique lorsqu'ils écrivent, et ce, selon la plateforme. Des études relevant du domaine de la psychologie pourraient même s'intéresser aux processus cognitifs (Klimova et Pikhart, 2020) qui sous-tendent l'alternance entre les codes écrits, oraux et iconiques dans la CNE. Il serait par exemple pertinent d'explorer ce qui motive l'alternance entre l'emploi et l'omission de la particule *ne* par un même internaute. Bref, les objets de recherche sont illimités dans le cadre des sciences sociales.

Étant donné la problématique de l'effet de la médiatisation technologique sur la production langagière, il serait par ailleurs pertinent d'examiner l'emploi de la variété de français parlé au Québec écrite sur diverses plateformes socio- et technonumériques, y compris les données de messages SMS plus récentes, ainsi que dans les différents contextes communicatifs engendrés par ces sites Web. Afin de pouvoir suivre l'évolution du français québécois, il faudra ainsi surveiller la situation dans tous les contextes communicatifs ; et pour déterminer l'impact des écrits en ligne sur la langue dans d'autres contextes, il faudra continuer à analyser le langage utilisé par les médias journalistiques à l'oral et à l'écrit. Voici un travail important tant pour les chercheurs en linguistique que pour des institutions comme l'Office québécois de la langue française.

Notons enfin les contributions potentielles issues de la collaboration des (socio)linguistes, qui s'intéressent à la communication par le biais du numérique, et des spécialistes en linguistique computationnelle, en sciences des données, en traitement du langage naturel et en intelligence artificielle générative. Par exemple, tandis qu'une étude récente par A.J. Alvero (2023) applique la théorie sociolinguistique à l'apprentissage automatique par les machines, celles de Wilkinson Daniel Wong Gonzales (2024) et Danang Satria Nugraha (2024), appliquent des méthodes computationnelles *Deep Learning* et *Big Data* respectivement, dans le cadre de leurs recherches

sociolinguistiques. Ces études montrent l'apport de la combinaison des compétences et des perspectives de domaines divers, ce qui ajoute aux connaissances en (socio)linguistique, mais qui pourrait également faire avancer les technologies de la communication numérique elles-mêmes.

Mon étude ajoute donc à un champ de recherche florissant en (socio)linguistique qui cherche à quantifier et à mesurer la communication numérique écrite à travers la diversité d'appareils et de plateformes socionumériques. Comme le souligne Susan Herring, cette trajectoire d'enquête « révélera les principes sous-jacents, les dimensions systématiques de la variation, qui rendent compte de la relation entre les paramètres des médias de la communication et de la communication humaine en général » (2002 : 152, ma traduction). Afin de comprendre l'impact de la transformation numérique sur la société et sur nos vies, il faut continuer d'examiner comment les individus, les groupes et les communautés se comportent dans ces situations technologiques et sociales. En fin de compte, la capacité d'utiliser le langage comme outil interactionnel, et en particulier la communication numérique écrite, est un phénomène exclusif à notre civilisation : son étude permettra ainsi de mieux comprendre la condition humaine.

## Bibliographie

ABOUDA, Lofti et Marie SKROVEC (2015), « Du rapport entre formes synthétique et analytique du futur. Étude de la variable modale dans un corpus oral micro-diachronique », *Revue de Sémantique et Pragmatique*, 38, p. 35-57. [En ligne]. URL : <https://journals.openedition.org/rsp/512> [Site consulté le 15 novembre 2023].

ABOUDA, Lofti et Marie SKROVEC (2017), « Du rapport micro-diachronique futur simple/futur périphrastique en français moderne. Étude des variables temporelles et aspectuelles », *Corela*, HS-21. [En ligne]. URL : <http://journals.openedition.org/corela/4804> [Site consulté le 24 novembre 2023].

ADEGBOKU, Dele (2011), « De l'économie en langue ou dans le langage : une linguistique “des temps qui pressent” », *Synergies Afrique Centrale et de l'Ouest* n° 4, p. 25-37. [En ligne]. URL : <https://gerflint.fr/Base/Afriqueouest4/dele.pdf> [Site consulté le 20 novembre 2023].

ALVERO, A.J. (2023), « Sociolinguistic Perspectives on Machine Learning with Textual Data », dans BORCH, Christian et Juan Pablo PARDO-GUERRA [dir.], *The Oxford Handbook of the Sociology of Machine Learning*, Oxford Academic. [En ligne]. URL : <https://osf.io/preprints/socarxiv/b5w6y> [Site consulté le 15 janvier 2024].

ANDRÉ, Virginie (2019), « Pourquoi faire de la sociolinguistique des interactions verbales avec des enseignants et des apprenants de Français langue étrangère ? », *Linx* 79. [En ligne]. URL : <http://journals.openedition.org/linx/3694> [Site consulté le 21 mars 2022].

ANDROUTSOPOULOS, Jannis (2006), « Introduction: Sociolinguistics and computer-mediated communication », *Journal of Sociolinguistics*, 10(4), p. 419–438. [En ligne]. URL : <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9841.2006.00286.x> [Site consulté le 20 août 2022].

ANDROUTSOPOULOS, Jannis (2007), « Language choice and code switching in German-based diasporic web forums », dans DANET, Brenda et Susan HERRING [dir.], *The Multilingual Internet Language, Culture, and Communication Online*, Oxford University Press. [En ligne]. URL : <https://doi.org/10.1093/acprof:oso/9780195304794.003.0015> [Site consulté le 15 août 2022].

ANDROUTSOPOULOS, Jannis (2010), « Localizing the global on the participatory web », dans COUPLAND, Nikolas [dir.], *The Handbook of Language and Globalisation*, Oxford, Wiley-Blackwell, p. 203–231. [En ligne]. URL : <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781444324068.ch9> [Site consulté le 23 août 2022].

ANDROUTSOPOULOS, Jannis (2014), « Linguaging when contexts collapse: Audience design in social networking », *Discourse, Context & Media*, Volumes 4–5, p. 62-73. [En ligne]. URL : <https://www.sciencedirect.com/science/article/abs/pii/S2211695814000257> [Site consulté le 17 juillet 2022].

ANDROUTSOPOULOS, Jannis. (2015), « Networked multilingualism: Some language practices on Facebook and their implications », *The International Journal of Bilingualism : Cross-Disciplinary, Cross-Linguistic Studies of Language Behavior*, 19(2), p. 185–205. En ligne]. URL : <https://journals.sagepub.com/doi/10.1177/1367006913489198> [Site consulté le 20 juillet 2022].

ANDROUTSOPOULOS, Jannis (2018), « Online data collection », dans MALLISON, Christine, Becky CHILDS et Gerard VAN HERK [dir.], *Data Collection in Sociolinguistics*, New York, Routledge, p. 233-244.

ANDROUTSOPOULOS, Jannis et TEREICK, Jana (2016), « YouTube: Language and discourse practices in participatory culture », dans GEORGAKOPOULOU, Alexandra et Tereza SPILIOTI [dir.], *The Routledge Handbook of Language and Digital Communication*, Routledge.

ANIS, Jacques (1998), *Texte et ordinateur. L'écriture réinventée ?*, De Boeck, Bruxelles.

ANIS, Jacques (1999), *Internet, communication et langue française*, Paris, Hermès science.

ASHBY, William (1981), « The loss of the negative particle *ne* in French », *Language*, 57(3), p. 674-687. [En ligne]. URL : <https://www.jstor.org/stable/414345> [Site consulté le 10 janvier 2024].

ASHBY, William (2001), « Un nouveau regard sur la chute du *ne* en français parlé tourangeau : s'agit-il d'un changement en cours ? », *Journal of French Language Studies*, 11(1), p. 1-22. [En ligne]. URL : <https://doi.org/10.1017/S0959269501000114> [Site consulté le 10 janvier 2024].

ASSOCIATION QUÉBÉCOISE DES PROFESSEURS DE FRANÇAIS (1977), « Le congrès du dixième anniversaire. Les résolutions de l'Assemblée générale », *Québec français*, 28, p. 10-12.

ATIFI, Hassan (2007), « Continuité et/ou rupture dans l'Internet multilingue : quelle langue parler dans un forum diasporique ? », *Glottopol : Revue de sociolinguistique en ligne*, 2007, 10, p. 113-127. [En ligne]. URL : [http://glottopol.univ-rouen.fr/telecharger/numero\\_10/gp110\\_08atifi.pdf](http://glottopol.univ-rouen.fr/telecharger/numero_10/gp110_08atifi.pdf) [Site consulté le 15 juin 2023].

ATIFI, Hassan et Michel MARCOCCIA (2006), « Communication médiatisée par ordinateur et variation culturelle : analyse contrastive de forums de discussion français et marocains », *Les Carnets du Cediscor* 9. [En ligne]. URL: <http://journals.openedition.org/cediscor/629> [Site consulté le 23 décembre 2023].

AUGER, Julie, et Anne-José VILLENEUVE (2021), « Étude comparative des particules interrogatives en picard et dans deux variétés de français parlées au Canada », *Langue française*, 4, p. 57-74. [En ligne]. URL : <https://www.cairn.info/revue-langue-francaise-2021-4-page-57.htm> [Site consulté le 15 décembre 2023].

BAINBRIDGE, William Sims (2020), *The Social Structure of Online Communities*, Cambridge, Cambridge University Press. [En ligne]. URL : <https://doi.org/10.1017/9781108615211> [Site consulté le 10 décembre 2022].

BALLY, Anne-Sophie (2022), « Les interrogatives totales en français québécois dans l'écrit SMS : à la croisée de l'oral et de l'écrit », *Congrès Mondial de Linguistique Française - CMLF 2022, SHS Web of Conferences 138*. [En ligne]. URL : [https://www.shs-conferences.org/articles/shsconf/abs/2022/08/shsconf\\_cmlf2022\\_12006/shsconf\\_cmlf2022\\_12006.html](https://www.shs-conferences.org/articles/shsconf/abs/2022/08/shsconf_cmlf2022_12006/shsconf_cmlf2022_12006.html) [Site consulté le 15 septembre, 2023].

BARBARIE, Yves (1982), « Analyse sociolinguistique de la syntaxe de l'interrogation en français québécois », *Revue québécoise de linguistique*, 12(1), p. 145–167. [En ligne]. URL : <https://www.erudit.org/fr/revues/rql/1982-v12-n1-rql2952/602496ar.pdf> [Site consulté le 15 janvier 2023].

BARTON, David et Carmen LEE (2012), « Redefining vernacular literacies in the age of Web 2.0 », *Applied Linguistics*, 33(3), p. 282–298. [En ligne]. URL : <https://doi.org/10.1093/applin/ams009> [Site consulté le 15 juin 2023].

BEAULIEU, Suzy, Davy BIGOT et Anne-José VILLENEUVE (2019), « L'expression de la conséquence en français québécois : recherche empirique et pistes pour l'enseignement de l'oral », *Correspondance*, 25(4). [En ligne]. URL : <https://correspo.ccdmd.qc.ca/wp-content/uploads/2019/12/Correspondance-lexpression-de-la-consequence-en-francais-quebecois-recherche-empirique-et-pistes-pour-lenseignement-de-loral-2.pdf> [Site consulté le 15 décembre 2023].

BÉDARD, Édith et Jacques MAURIS (1983), *La norme linguistique*, Québec, Conseil supérieur de la langue française.

BELL, Allan (1984), « Language style as audience design », *Language in Society* 13, p.145-204. [En ligne]. URL : <https://www.jstor.org/stable/4167516> [Site consulté le 15 janvier 2022].

BEN AMOR, Samy et GRANGET, Lucia (2011), « L'identité numérique. De la construction au suicide en 52 minutes », *Les Cahiers du numérique*, 7(1), p. 103-115. [En ligne]. URL : <https://www.cairn.info/revue-les-cahiers-du-numerique-2011-1-page-103.htm> [Site consulté le 20 décembre 2023].

BERGERON, Léandre (1980), *Dictionnaire de la langue québécoise*, vlb éditeur.

BERGERON, Léandre (1982), *The Québécois Dictionary*, James Lorimer Limited.

BERNSTEIN, William J. (2013), *Masters of the Word : How Media Shaped History from the Alphabet to the Internet*. New York, Grove Press.

BIBER, Douglas (1988), *Variation across Speech and Writing*, Cambridge University Press.

BIBER, Douglas et Jesse EGBERT (2016), « Register Variation on the Searchable Web: A Multi-Dimensional Analysis », *Journal of English Linguistics*, 44(2), p. 95–137. [En ligne]. URL : <https://doi.org/10.1177/0075424216628955> [Site consulté le 20 décembre 2022].

BIBER Douglas et Jesse EGBERT (2018), *Register Variation Online*. Cambridge University Press.

BIBER, Douglas, Jesse EGBERT et Meixiu ZHANG (2017), « Using corpus-based analysis to study register and dialect variation on the searchable web », dans FRIGINAL, Eric [dir.], *Studies in Corpus-Based Sociolinguistics*, New York, Routledge. [En ligne].

URL : <https://doi.org/10.4324/9781315527819> [Site consulté le 21 août 2021].

BIBIÉ-EMERIT, Laetitia (2015), *Description du discours numérique : étude des bouleversements linguistiques du Web 2.0 au travers de l'exemple des souhaits d'anniversaire sur Facebook*, Thèse de doctorat, Linguistique, Université Michel de Montaigne — Bordeaux III. [En ligne]. URL : <https://tel.archives-ouvertes.fr/tel-01442467/document> [Site consulté le 20 août 2021].

BIGOT, Davy (2008), *Le point sur la norme grammaticale du français québécois oral*, Thèse de doctorat, UQAM. [En ligne]. URL : <https://archipel.uqam.ca/1305/> [Site consulté le 18 août 2021].

BIGOT, Davy (2011), « De la norme grammaticale du français parlé au Québec », *Arborescences*, 1. [En ligne]. URL : <https://www.erudit.org/fr/revues/arbo/2011-n1-arbo1515298/1001939ar/> [Site consulté le 20 août 2021].

BIGOT, Davy (2021), *Le bon usage québécois. Étude sociolinguistique sur la norme grammaticale du français parlé au Québec*. Québec, Presses de l'Université Laval.

BILLIEZ, Jacqueline et Laurence BUSON (2013), « Perspectives diglossique et variationnelle - Complémentarité ou incompatibilité ? Quelques éclairages sociolinguistiques », *Journal of French Language Studies*, 23(1), p. 135-149. [En ligne]. URL : <https://doi.org/10.1017/S0959269512000397> [Site consulté le 25 août 2021].

BLATTNER, Géraldine et Lawrence WILLIAMS (2011), « L'emploi variable du *ne* dans le discours électronique synchrone : une étude variationniste en temps apparent », *Langage et société*, 138, p. 109-129. [En ligne]. URL : <https://www.cairn.info/revue-langage-et-societe-2011-4-page-109.htm?ref=doi> [Site consulté le 15 juin 2022].

BLATTNER, Géraldine et Lawrence WILLIAMS (2022), « The variable use of diacritics in synchronous computer-mediated French discourse: Replication research », *Journal of French Language Studies*, 32(2), p. 216-242. [En ligne]. URL : <https://doi.org/10.1017/S0959269522000011> [Site consulté le 20 août 2023].

BLONDEAU, Hélène (2000), « Parcours d'un marqueur sociolinguistique. Les pronoms non clitiques du pluriel en français parlé de Montréal », thèse de doctorat, Université de Montréal. [En ligne]. URL : <https://www.collectionscanada.gc.ca/obj/s4/f2/dsk2/ftp02/NQ53289.pdf> [Site consulté le 20 août 2021].

BLONDEAU, Hélène (2006), « La trajectoire de l'emploi du futur chez une cohorte de Montréalais francophones entre 1971 et 1975 », *Revue de l'Université de Moncton*, 37, p.73–98. [En ligne]. URL : <https://www.erudit.org/fr/revues/rum/2006-v37-n2-rum1731/015840ar.pdf> [Site consulté le 22 août 2021].

BLONDEAU, Hélène (2011), *Cet "autres" qui nous distingue : tendances communautaires et parcours individuels dans le système des pronoms en français québécois*, Presses de l'Université Laval.

BLONDEAU, Hélène et Emmanuelle LABEAU (2016), « La référence temporelle au futur dans les bulletins météo en France et au Québec : regard variationniste sur l'oral préparé », *Canadian Journal of Linguistics*, 61(3), p. 240–258. [En ligne]. URL : <https://doi.org/10.1017/cnj.2016.26> [Site consulté le 20 décembre 2023].

BLONDEAU, Hélène, Raymond MOUGEON et Mireille TREMBLAY (2017), « Divergent vs convergent patterns of variation and change in Montréal and Welland French: The case of consequence markers », Communication présentée à *New Ways of Analyzing Variation* (NWAV) 46, University of Wisconsin, Madison, Wisconsin.

BLONDEAU, Hélène, MOUGEON, Raymond et Mireille TREMBLAY (2019), « Analyse comparative de *ça fait que*, *alors*, *donc* et *so* à Montréal et à Welland: mutations sociales, convergences, divergences en français laurentien », *Journal of French Language Studies*, 29(1), p.35–65. [En ligne]. URL : <https://doi.org/10.1017/S0959269518000169> [Site consulté le 18 décembre 2022].

BLONDEAU, Hélène, Gillian SANKOFF et Anne CHARITY (2002), « Parcours individuels dans deux changements linguistiques en cours en français montréalais », *Revue québécoise de linguistique* 31(1), p. 13–38. [En ligne]. URL : <https://www.erudit.org/fr/revues/rql/2002-v31-n1-rql538/006843ar/> [Site consulté le 15 juillet 2023].

BLONDEAU, Hélène, et Mireille TREMBLAY (2022a), « The hidden dimensions of a change from below: Consequence markers in Montreal French », *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 67(1-2), p. 22–52. [En ligne]. URL : <https://doi.org/10.1017/cnj.2022.3> [Site consulté le 15 juillet 2023].

BLONDEAU, Hélène et Mireille TREMBLAY (2022b), « Écrire son vernaculaire : variation et normes communautaires dans les messages textes en français québécois », *Journal of French Language Studies*, 32(2), p. 120–144. [En ligne]. URL : <https://doi.org/10.1017/S0959269522000096> [Site consulté le 20 juillet 2023].

BLONDEAU, Hélène, Mireille TREMBLAY et Patrick DROUIN (2014), « Hybridité et variation dans les SMS : Le corpus Texto4Science et l’oralité en français montréalais », *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 59(1), p. 137-165. [En ligne]. URL : doi:10.1017/S0008413100000189 [Site consulté le 22 août 2021].

BOLLY, Catherine et Liesbeth DEGAND (2009), « Quelle (s) fonction (s) pour donc en français oral ? : Du connecteur conséquentiel au marqueur de structuration du discours », *Linguisticae Investigationes*, 32(1), p. 1-32. [En ligne]. URL : <https://benjamins.com/catalog/li.32.1.01bol> [Site consulté le 22 décembre 2023].

BOURDIEU, Pierre (2001), *Langage et pouvoir symbolique*. Éditions du Seuil.

BOURGET, Claire, Jonathan COUTURIER et Guillaume DUBÉ (2022), *Actualités en ligne, réseaux sociaux et balados*, NETendances, 13(9), Académie de transformation numérique. [En ligne]. URL : <https://transformation-numerique.ulaval.ca/wp-content/uploads/2023/06/netendances-2022-actualites-en-ligne-reseaux-sociaux-et-balados.pdf> [Site consulté le 15 juin 2023].

BOURGET, Claire et Jonathan COUTURIER (2023), *Actualités en ligne, réseaux sociaux et balados*, NETendances, 14(3), Académie de transformation numérique. [En ligne]. URL : <https://transformation-numerique.ulaval.ca/enquetes-et-mesures/netendances/actualites-en-ligne-reseaux-sociaux-et-balados-2023/> [Site consulté le 15 décembre 2023].

BOUTET, Josiane (1986), « La référence à la personne en français parlé : le cas de “on” », *Langage et société*, n°38, p. 19-49. [En ligne]. URL : [https://www.persee.fr/doc/lso\\_0181-4095\\_1986\\_num\\_38\\_1\\_2070](https://www.persee.fr/doc/lso_0181-4095_1986_num_38_1_2070) [Site consulté le 20 décembre 2023].

BOUTET, Josiane (2002), « Terrain », dans CHARAUDEAU, Patrick et Dominique MAINGUENEAU [dir.], *Dictionnaire d'analyse du discours*, Paris, Éditions du Seuil, p. 568-570.

BOUTET, Josiane et Françoise GADET (2003), « Pour une approche de la variation linguistique », *Le français aujourd'hui*, 143, p. 17-24. [En ligne]. URL : <https://www.cairn.info/revue-le-francais-aujourd-hui-2003-4-page-17.htm> [Site consulté le 12 octobre 2022].

BOYD, danah (2010), « Privacy and publicity in the context of big data », WWW, Raleigh, North Carolina, 29 avril. [En ligne]. URL : <http://www.danah.org/papers/talks/2010/WWW2010.html> [Site consulté le 16 janvier 2023].

BŘEZINA, Vaclav (2014), « Effect sizes in corpus linguistics: Keywords, collocations and diachronic comparisons », *35th ICAME conference*, Nottingham, 30 avril – 4 mai.

BŘEZINA, Vaclav (2016), « Collocation networks : exploring associations in discourse », dans BAKER, Paul et Jesse EGBERT [dir.], *Triangulating Methodological Approaches in Corpus-Linguistic Research*, New York, Routledge, p. 90-107. [En ligne]. URL : <https://doi.org/10.4324/9781315724812> [Site consulté le 6 janvier 2023].

BŘEZINA, Vaclav et Petr POŘÍZKA (2021), « Collocation graphs and networks using #LancsBox: Applications in English and Czech », *Casopis pro Moderní Filologii*, 103(1), p. 36-59. [En ligne]. URL : <https://casopispromodernifilologii.ff.cuni.cz/en/2021-1-2/> [Site consulté le 6 janvier 2023].

BŘEZINA, Vaclav et Lorrae FOX (2021), « Adjective + noun collocations in L2 and L1 speech: Evidence from the trinity Lancaster corpus and the spoken BNC2014 », dans GRANGER, Sylviane [dir], *Perspectives on the L2 Phrasicon: The View from Learner Corpora*, Bristol, Multilingual Matters, p. 152-177. [En ligne]. URL : <https://doi.org/10.21832/9781788924863-008> [Site consulté le 6 janvier 2023].

BŘEZINA, Vaclav, Pierre WEILL-TESSIER et Tony McENERY (2020), #LancsBox v. 5.x. [logiciel]. [En ligne]. URL : <http://corpora.lancs.ac.uk/lancsbox> [Site consulté le 6 janvier 2023].

BŘEZINA, Vaclav, Matt TIMPERLEY et Tony McENERY (2018), #LancsBox v. 4.x [software]. [En ligne]. URL : <http://corpora.lancs.ac.uk/lancsbox> [Site consulté le 6 janvier 2023].

BŘEZINA, Vaclav, Tony McENERY et Stephen WATTAM, (2015), « Collocations in context: A new perspective on collocation networks », *International Journal of Corpus Linguistics*, 20(2), p.139-173. [En ligne]. URL : <https://benjamins.com/catalog/ijcl.20.2.01bre> [Site consulté le 6 janvier 2023].

BULOT, Thierry et Philippe BLANCHET (2013), *Une introduction à la sociolinguistique : pour l'étude des dynamiques de la langue française dans le monde*, Paris, Éditions des archives contemporaines.

BURNETT, Heather, et Mireille TREMBLAY (2015), « L'expression de la négation en français québécois: pantoute, polarité négative et mots-N1 », dans REMYSEN, Wim [dir.], *Les Français d'ici*, Les Presses de l'Université de Laval, p. 261-290.

CARDON, Dominique (2019), *Culture numérique*, Armand Colin.

CHARLAND, Bailey (2010), « Spanish modals of obligation : different uses of *tener que* and *necesitar* », Rapport de recherche de maîtrise, Kansas State University. [En ligne]. URL : <https://krex.k-state.edu/server/api/core/bitstreams/52e7733a-0ffc-4389-887d-5a7cf22922eb/content> [Site consulté le 15 avril 2023].

CHOMSKY, Noam (1986), *Knowledge of Language: Its Nature, Origin, and Use*, Greenwood Publishing Group.

COLLOT, Milena et Nancy BELMORE (1996), « Electronic language : A new variety of English », dans HERRING, Susan [dir.], *Computer-Mediated Communication. Linguistic, Social and Cross-Cultural Perspectives*, Amsterdam/Philadelphia, John Benjamins, p. 13-28.

GOUVERNEMENT DU CANADA (2018), *Énoncé de politique des trois Conseils : Éthique de la recherche avec des êtres humains*, Secrétariat sur la conduite responsable de la recherche, Conseil de recherches en sciences humaines, conseil de recherches en sciences naturelles et en génie du Canada, instituts de recherche en santé du Canada. [En ligne]. URL : <https://ethics.gc.ca/fra/documents/tcps2-2018-fr-interactive-final.pdf> [Site consulté le 16 janvier 2023].

CORBLIN, Francis et Lucia TOVENA (2003), « L'expression de la négation dans les langues romanes », *Les langues romanes: problèmes de la phrase simple*, 64, p. 242-279. [En ligne]. URL : <https://www.linguist.univ-paris-diderot.fr/~tovenapapers/CorblinTovena03NC.pdf> [Site consulté le 12 décembre 2023].

COVENEY, Aidan (1997), « L'approche variationniste et la description de la grammaire du français: le cas des interrogatives », *Langue française*, 115, p. 88-100. [En ligne]. URL : [https://www.persee.fr/doc/lfr\\_0023-8368\\_1997\\_num\\_115\\_1\\_6224](https://www.persee.fr/doc/lfr_0023-8368_1997_num_115_1_6224) [Site consulté le 18 décembre 2023].

COVENEY, Aidan (2000), « Vestiges of nous and the 1st person plural verb in informal spoken French », *Language Sciences (Oxford)*, 22(4), p. 447–481. [En ligne]. URL : [https://doi.org/10.1016/S0388-0001\(00\)00014-0](https://doi.org/10.1016/S0388-0001(00)00014-0) [Site consulté le 18 décembre 2023].

COVENEY, Aidan (2011), « A language divided against itself ? Diglossia, code-switching and variation in French », dans MARTINEAU, France et Terry NADASDI [dir]. *Le français en contact: Hommages à Raymond Mougeon*, Les Presses de l'Université Laval, p. 51-85.

COVENEY, Aidan (2016), « La quête du vernaculaire dans l'étude de la variation grammaticale », *SHS Web of Conferences*, 27, 1003–. [En ligne]. URL : <https://doi.org/10.1051/shsconf/20162701003> [Site consulté le 18 décembre 2023].

COUGNON, Louise-Amélie (2008), « Le français de Belgique dans l'écrit spontané : Approche syntaxique et phonétique d'un corpus de 30 000 SMS », *Travaux du cercle Belge de linguistique*, 3. [En ligne]. URL : <https://sites.uclouvain.be/bkl-cbl/wp-content/uploads/2014/08/cou2008.pdf> [Site consulté le 23 décembre 2023].

COUGNON, Louise-Amélie (2010), « Orthographe et langue dans les SMS : conclusions à partir de quatre corpus francophones », *Éla. Études de linguistique appliquée*, 160, p. 397-410. [En ligne]. URL : <https://www.cairn.info/revue-ela-2010-4-page-397.htm> [Site consulté le 23 décembre 2023].

COULMAS, Florian (2013), *Writing and society : an introduction*, Cambridge University Press.

CROWLEY, D. J. et Paul HEYER (2011), *Communication in History : Technology, Culture, Society* (6e édition), Boston, MA, Allyn & Bacon/Pearson.

CRYSTAL, David (2000), *Language Death*, Cambridge, Cambridge University Press.

CRYSTAL, David (2001), *Language and the Internet*, Cambridge University Press.

CRYSTAL, David (2005), « The scope of Internet Linguistics » *Communication virtuelle, American Association for the Advancement of Science*, février. [En ligne]. URL : <https://davidcrystal.com> [Site consulté le 19 mars 2022].

CRYSTAL, David (2011), *Internet Linguistics: A Student Guide*, Routledge.

CUNLIFFE, Daniel (2007), « Minority languages in the Internet: New threats, new opportunities », dans CORMACK, M. et N. HOURIGAN [dir.], *Minority Language Media: Concepts, Critiques, and Case Studies*, Multilingual Matters, p. 133– 150.

CUNLIFFE, Daniel (2009), « The Welsh language on the Internet: Linguistic resistance in the age of the network society », dans GOGGIN, Gerard et Mark McLELLAND [dir.], *Internationalizing Internet Studies: Beyond Anglophone Paradigms*, Routledge, p. 96– 111.

CUNLIFFE, Daniel, Delyth MORRIS et Cynog PRYS (2013), « Investigating the differential use of Welsh in young speakers' social networks: a comparison of communication in face-to-face settings, in electronic texts and on social networking sites », dans HAF GRUFFYDD JONES, Elin et Enrique URIBE-JONGBLOED [dir.], *Social Media and Minority Languages: Convergence and the Creative Industries*, p. 74-86.

DAVID Jacques, et Harmony GONCALVES (2007), « L'écriture électronique, une menace pour la maîtrise de la langue ? », *Le français aujourd'hui*, 1(156), p.3 9-47. [En ligne]. URL : <https://www.cairn.info/revue-le-francais-aujourd-hui-2007-1-page-39.htmLa> [Site consulté le 5 décembre 2023].

DESHAIES, Denise (1991), « Contribution à l'analyse du français québécois: étude des pronoms personnels », *Revue québécoise de linguistique théorique et appliquée*, 10(3), p. 11-40.

DESHAIES, Denise et Eve LAFORGE (1981), « Le futur simple et le futur proche dans le français parlé dans la ville de Québec », *Langues et linguistique*, 7, p. 21-37.

DESSUREAULT-DOBER, Diane (1974), *Étude sociolinguistique de /Ca fait que/ : Coordonnant logique et marqueur d'interaction*, Mémoire de maîtrise, Université du Québec à Montréal.

DEUMERT, Ana et Kristin Vold LEXANDER (2013), « Texting Africa: Writing as performance », *Journal of Sociolinguistics*, 17(4), p. 522–546. [En ligne]. URL : <https://doi.org/10.1111/josl.12043> [Site consulté le 15 janvier 2024].

DEVELOTTE Christine, et Marie-Anne PAVEAU (2017), « Pratiques discursives et interactionnelles en contexte numérique. Questionnements linguistiques », *Langage et société*, 2-3(160-161), p. 199-215. [En ligne]. URL : <https://www.cairn.info/revue-langage-et-societe-2017-2-page-199.htm> [Site consulté le 21 décembre 2023].

DE VILLERS, Marie-Éva (2004), *Description et analyse synchroniques de faits lexicaux propres au français du Québec dans Le Devoir: illustration de la norme réelle du français québécois*, thèse de doctorat, Université de Montréal.

DE VILLERS, Marie-Éva (2005), *Le vif désir de durer. Illustration de la norme réelle du français québécois*, Montréal, Éditions Québec Amérique.

Di VITO, Sonia (2015), « Le traitement lexicographique de la synonymie à l'aide de la linguistique de Corpus : Le cas de profit et de ses synonymes », *Éla. Études de linguistique appliquée*, 178, p. 177-193. En ligne]. URL : <https://www.cairn.info/revue-ela-2015-2-page-177.htm> [Site consulté le 15 juin 2023].

DOSTIE, Gaétane (2006), « Régularité et spécificité dans le paradigme des marqueurs consécutifs. Fait que en français québécois », *Cahiers de lexicologie*, 89(2), p.75–96.

DOSTIE, Gaétane (2013), « Les associations de marqueurs discursifs—De la cooccurrence libre à la collocation », *Linguistik online*, 62(5). [En ligne]. URL :<https://bop.unibe.ch/linguistik-online/article/view/1304> [Site consulté le 24 novembre 2023].

DRAY, Susan et Rob DRUMMOND (2019), « Mixing methods ? Linguistic ethnography and language variation », dans TUSTIN, Karin [dir.], *The Routledge Handbook of Linguistic Ethnography*, Routledge, p.241-255.

DUBOIS, Jean, *et al.* (2002), *Dictionnaire de linguistique*, Larousse.

EISENSTEIN, Elizabeth L. (1980), *The Printing Press as an Agent of Change*, Cambridge University Press.

EISENSTEIN, Jacob (2015). « Systematic patterning in phonologically-motivated orthographic variation », *Journal of Sociolinguistics* 19(2), p. 161–188.

EMIRKANIAN, Louissette et David SANKOFF (1986), « Le futur simple et le futur périphrastique dans le français parlé », dans *Morphosyntaxe des langues romanes, Actes du XVIIe Colloque international de linguistique et de philologie romanes*, Aix-en-Provence, p. 397-407.

ENCREVÉ, Pierre (1976), « Présentation : Labov, linguistique, sociolinguistique », dans LABOV, William (1976), *Sociolinguistique*, Paris, Éditions de Minuit.

ERIKSON, Thomas (1999), « Persistent conversation : An Introduction », *Journal of Computer-Mediated Communication*, 4(4), [En ligne]. URL : <https://academic.oup.com/jcmc/article/4/4/JCMC447/4584406> [Site consulté le 18 janvier 2024].

ESRC (2015), *ESRC Framework for Research Ethics*. [En ligne]. URL : <http://www.esrc.ac.uk/files/funding/guidance-for-applicants/esrc-framework-for-researchethics-2015/.pdf> [Site consulté le 16 janvier 2023].

FAFA, Fatima Zohra (2017), *Spécificités du langage sms : du système linguistique et paralinguistique*, Mémoire de maîtrise, Université Mohamed Seddik Ben Yahia, Pôle universitaire de Tassoust Jijel, Faculté des lettres et des langues, Département des lettres et langue française. [En ligne]. URL : <http://dspace.univ-jijel.dz:8080/xmlui/bitstream/handle/123456789/4093/Sp%c3%a9cificit%c3%a9s%20du%20langage%20sms.pdf?sequence=1&isAllowed=y> [Site consulté le 15 décembre 2023].

FALKERT, Anika (2006), « La mutation achevée du connecteur *ça fait que* dans le français acadien des Îles-de-la-Madeleine », *Revue de l'Université de Moncton*, 37(2), p. 39-53. [En ligne]. URL : <https://id.erudit.org/iderudit/015838ar> [Site consulté le 20 décembre 2023].

FAUDAOUS, Jean-Claude (1990), *Dictionnaire des abréviations courantes de la langue française*, Paris, Maison du dictionnaire.

FERRARA, Kathleen, Hans BRUNNER et Greg WHITTEMORE (1991), « Interactive written discourse as an emergent register », *Written Communication*, 8(1), p. 8-34. [En ligne]. URL : <https://doi.org/10.1177/0741088391008001002> [Site consulté le 15 avril 2022].

FERRE, Gaelle (2011), « Multimodal analysis of discourse markers ‘donc’, ‘alors’ and ‘en fait’ in conversational French », *ICPhS VXII*, Hong Kong, Chine, p. 671-674. [En ligne]. URL : <https://hal.science/hal-00617521/document> [Site consulté le 20 décembre 2023].

FEUILLARD, Colette (2017), *Usage, norme et codification : de la diversité des situations à l'utilisation du numérique*, Louvain-la-Neuve, Éditions E.M.E.

FONSECA-GREBER, Bonnibeth Beale (2007), « The emergence of emphatic ‘ne’ in conversational Swiss French », *Journal of French Language Studies*, 17(3), p. 249–275. [En ligne]. URL : <https://doi.org/10.1017/S0959269507002992> Site consulté le 20 décembre 2023].

FOUSER, Robert, Narahiko INOUE et Chungmin LEE (2000), « The pragmatics of orality in English, Japanese and Korean computer-mediated communication », dans PEMBERTON, Lynn et Simon SHURVILLE [dir.], *Words on Web : Computer Mediated Communication*, Exeter-Portland, Intellect Books, p. 52-62.

FRANZKE, Aline Shakti, Anja BECHMANN, Michael ZIMMER, Charles ESS et l'Association of Internet Researchers (2020), *Internet Research: Ethical Guidelines 3.0*. [En ligne]. URL : <https://aoir.org/reports/ethics3.pdf> [Site consulté le 16 janvier 2023].

FRIGINAL, Eric (2013), « Twenty-five years of Biber's multi-dimensional analysis: Introduction to the special issue and an interview with Douglas Biber », *Corpora* 8 (2), p.137-152. [En ligne]. URL : <https://doi.org/10.3366/cor.2013.0038> [Site consulté le 20 décembre 2023].

FRIGINAL, Eric et Jack HARDY (2014), *Corpus-Based Sociolinguistics : A Guide for Students*, Routledge.

FRIGINAL, Eric, Oskana WAUGH et Ashley TITAK (2017), « Linguistic variation in Facebook and Twitter posts », dans FRIGINAL, Eric [dir.], *Studies in Corpus-Based Sociolinguistics*, Routledge, p. 342–362.

GADET, Françoise (1995), « Norme, variation, évaluation », *La genèse de la norme. Colloque de la SHESL, Histoire épistémologie langage*, 11(1), p.18-22. [En ligne]. URL : [https://www.persee.fr/doc/hel\\_0247-8897\\_1995\\_num\\_11\\_1\\_3399](https://www.persee.fr/doc/hel_0247-8897_1995_num_11_1_3399) [Site consulté le 11 décembre 2022].

GADET, Françoise (2003), *La variation sociale en français*, Paris, Orphrys.

GADET, Françoise (2007), « À la mémoire de Jacques Anis », *Glottopol, Revue de sociolinguistique en ligne : Regards sur l'internet, dans ses dimensions langagières. Penser les continuités et discontinuités. En hommage à Jacques Anis*, n° 10, juillet. [En ligne]. URL : [http://glottopol.univ-rouen.fr/telecharger/numero\\_10/gpl10\\_temoignage.pdf](http://glottopol.univ-rouen.fr/telecharger/numero_10/gpl10_temoignage.pdf) [Site consulté le 13 décembre 2022].

GARCIA, Ofelia et Ricardo OTHEGUY (2020), « Plurilingualism and translanguaging: commonalities and divergences », *International Journal of Bilingual Education and Bilingualism*, 23(1), p. 17–35. [En ligne]. URL : <https://doi.org/10.1080/13670050.2019.1598932> [Site consulté le 15 décembre 2023].

GARVIN, Paul (1983), « Le rôle des linguistes de l'École de Prague dans le développement de la norme linguistique tchèque », dans BÉDARD, Édith et Jacques MAURAI [dir.], *La norme linguistique*, Québec, Conseil supérieur de la langue française. [En ligne]. URL : [https://www.cslf.gouv.qc.ca/bibliotheque-virtuelle/publication-html/?tx\\_iggeplusplus\\_pi4%5Bfile%5D=publications/pubf101/f101p1.html](https://www.cslf.gouv.qc.ca/bibliotheque-virtuelle/publication-html/?tx_iggeplusplus_pi4%5Bfile%5D=publications/pubf101/f101p1.html) [Site consulté le 22 mars 2022].

GEORGES, Fanny, Antoine SEILLES, Guillaume ARTIGNAN, Bérenger ARNAUD, Nancy RODRIGUEZ, Jean SALLANTIN et Mountaz HASCOËT (2009), « Sémiotique et visualisation de l'identité numérique: une étude comparée de Facebook et Myspace », *Actes de la conférence H2PTM'09*, Paris, Hermès, p. 257-268. [En ligne]. URL : <https://hal.science/hal-00410952> [Site consulté le 10 janvier 2024].

GILLESPIE, Tarleton, Pablo J. BOCZKOWSKI et Kirsten A. FOOT (2014), *Media Technologies : Essays on Communication, Materiality, and Society*, Cambridge, Massachusetts, The MIT Press.

GOUVERNEMENT DU CANADA (2023), « Français et anglais : Vers une égalité réelle des langues officielles au Canada », *Patrimoine Canada*. [En ligne].

URL : <https://www.canada.ca/fr/patrimoine-canadien/organisation/publications/publications-generales/egalite-langues-officielles.html> [Site consulté le 18 janvier 2024].

GOUVERNEMENT DU QUÉBEC (2020), *Plan stratégique 2018-2023*, Office québécois de la langue française. [En ligne]. URL : <https://www.oqlf.gouv.qc.ca/office/2018-2023-plan-strategique.pdf> [Site consulté le 18 janvier 2024].

GRAFSTRÖM, Åke (1969), « On remplaçant nous en français », *Revue de linguistique romane*, 33(131-132), p. 270-298.

GREGORY, Michael et Susanne CARROLL (1978), *Language and Situation : Language Varieties and Their Social Contexts*, Routledge and Kegan Paul.

GROSSETTI, Michel (2014), « Que font les réseaux sociaux aux réseaux sociaux : Réseaux personnels et nouveaux moyens de communication », *Réseaux*, 184-185, p. 187-209. [En ligne]. URL : <https://www.cairn.info/revue-reseaux-2014-2-page-187.htm> [Site consulté le 20 juin 2022].

GUIGNARD, Jean-Baptiste (2011), « Linguistique cognitive et modèles catégoriaux : Quelques considérations épistémologiques », *Corela*, 9(2). [En ligne]. URL : <http://journals.openedition.org/corela/2127> [Site consulté le 18 janvier 2024].

HALTÉ, Pierre (2020), « Les émoticônes : de la signification des affects aux stratégies conversationnelles », *Communiquer*, 28. [En ligne]. URL : <https://www.erudit.org/en/journals/communiquer/1900-v1-n1-communiquer05334/1069900ar/abstract/> [Site consulté le 20 août 2021].

HALVORSEN, Stefan (2021), *T'as-tu déjà vu ça ? Les constructions interrogatives en français québécois*, Mémoire de maîtrise, Université d'Oslo. [En ligne]. URL :

<https://www.duo.uio.no/bitstream/handle/10852/88388/1/T-as-tu-d-j--vu--a--Stefan-Halvorsen--Masteroppgave-i-FRA4194.pdf> [Site consulté le 10 août 2023].

HANSEN, Anita Berit et Isabelle MALDEREZ (2000), « La négation en français parlé—une enquête en région parisienne », *Le Français parlé, Actes du colloque international Université de Copenhague du 29 au 30 octobre 1998*.

HANSEN, Marianne et Andreas STAEHR (2021), « Digital language practices and generational differences », *NyS, Nydanske Sprogstudier*, 1(59), p. 113-156.

HERRING, Elizabeth (2015), *Por qué piko ? Examining Paraguayan Spanish question markers through dissection of Paraguayan tweets and other media*, Thèse de doctorat, Indiana University Bloomington.

HERRING, Susan (1996), *Computer-Mediated Communication : Linguistic, Social and Cross-Cultural Perspectives*, John Benjamins.

HERRING, Susan (2002), « Computer-mediated communication on the internet », *Annual Review of Information Science and Technology*, 36(1), p. 109-168. [En ligne]. URL : <https://doi.org/10.1002/aris.1440360104> [Site consulté le 20 juin 2021].

HERRING, Susan (2004), « Computer-mediated discourse analysis: An approach to researching online behavior », dans BARAB, Sarah, Rob KLING et James GRAY [dir.], *Designing for Virtual Communities in the Service of Learning (Learning in Doing: Social, Cognitive and Computational Perspectives)*, Cambridge, Cambridge University Press, p. 338-376.

HERRING, Susan (2007), « A faceted classification scheme for computer-mediated discourse », *Language@Internet*, 4(1). [En ligne]. URL :

<https://www.languageatinternet.org/articles/2007/761> [Site consulté le 19 mars 2022].

HERRING, Susan (2013), « Discourse in web 2.0: familiar, reconfigured, and emergent », dans TANNEN, Deborah, et Anna Marie TRESTER [dir.] *Discourse 2.0 : Language and New Media*. Washington, DC, Georgetown University Press, p. 1-25.

HILLE, Sanne et Piet BAKKER (2014), « Engaging the Social News User », *Journalism Practice*, 8(5), p. 563-572. [En ligne]. URL : <https://doi.org/10.1080/17512786.2014.899758> [Site consulté le 15 avril 2023].

HILTE, Lisa, Reinhild, VANDEKERCKHOVE et Walter DAELEMANS (2018), « Social media writing and social class: A correlational analysis of adolescent CMC and social background », *International Journal of Society, Culture & Language*, 6(2), 73-89. [En ligne]. URL : [https://www.ijsc.com/article\\_31069.html](https://www.ijsc.com/article_31069.html) [Site consulté le 15 décembre 2023].

HUNDT, Marianne, Nadja NESSELHAUF et Carolin BIEWER (2007), *Corpus Linguistics and the Web*, Internationales Wissenschaftsforum, Fritz Thyssen-Stiftung, Universität Heidelberg.

HYMES, Dell (1972), « Models of the interaction of language and social life », dans GUMPERZ, John et Dell HYMES [dir.], *Directions in Sociolinguistics. The Ethnography of Communication*, Oxford, Basil Blackwell, p. 35-71.

HYMES, Dell (1974), *Foundations in Sociolinguistics: An Ethnographic Approach*, Philadelphia, University of Pennsylvania Press.

HYMES, Dell (2001), *Foundations in Sociolinguistics : An Ethnographic Approach*, Routledge.

IORIO, Josh (2015), « Vernacular literacy : Orthography and literacy practices », dans GEORGAKOPOULOU, Alexandra et Tereza SPILIOTI [dir.], *The Routledge Handbook of Language and Digital Communication* (1e éd.), Routledge, p. 164-176.

JEANJEAN, Colette (1988), « Le futur simple et le futur périphrastique en français parlé : Étude distributionnelle », dans Claire BLANCHE-BENVENISTE *et al.* [dir.], *Grammaire et histoire de la grammaire : Hommage à la mémoire de Jean Stéfanini*, PU de Provence, p. 235-257.

JEWITT, Casey (2016), « Multimodal analysis », dans GEORGAKOPOULOU, Alexandra et Tereza SPILIOTI [dir.] *The Routledge Handbook of Language and Digital Communication* (1e éd.), Routledge, p. 69-84.

JOHANN, Michael, Cornelia WOLF et Alexander GODULLA (2021), « Managing relationships on Facebook: A long-term analysis of leading companies in Germany », *Public Relations Review*, 47(3). [En ligne]. URL : <https://doi.org/10.1016/j.pubrev.2021.102044> [Site consulté le 20 août 2021].

JOHANSSON, Marjut (2017), « Everyday opinions in news discussion forums: Public vernacular discourse », *Discourse, Context & Media*, 19, p. 5-12. [En ligne]. URL : <https://www.sciencedirect.com/science/article/pii/S2211695816301088> [Site consulté le 13 décembre 2023].

JOHNSON, Ian (2013), « Audience design and communication accommodation theory: use of Twitter by Welsh-English biliterates » dans HAF, Elin, et URIBE-JONGBLOED, Enrique [dir.], *Social Media and Minority Languages: Convergence and the Creative Industries*, Multilingual Matters, p. 99-118

JONES, Elin (2013), « Minority language media, convergence culture and the indices of linguistic vitality », dans Elin JONES et Enrique URIBE-JONGBLOED [dir.], *Social Media and Minority Languages: Convergence and the Creative Industries* Multilingual Matters, p. 58-74.

JONGBLOED-FABER, Lysbeth, Jolie van LOO, et Leonie CORNIPS (2017), « Regional Languages on Twitter: A Comparative Study Between Frisian and Limburgish », *Dutch Journal of Applied Linguistics*, 6(2), p. 174–196. [En ligne]. URL : <https://doi.org/10.1075/dujal.16017.jon> [Site consulté le 20 septembre 2023].

JÜNGER, Jakob et Till KEYLING (2019), *Facepager : An application for automated data retrieval on the web*. [En ligne]. URL : <https://github.com/strohne/Facepager/> [Site consulté le 2 avril 2024].

KÄRKI, Katja (2017), *Alternance codique en anglais dans des groupes Facebook français et québécois : cas des grimpeurs*, Mémoire de maîtrise, Université de Helsinki.

KENNEDY TERRY, Kristen. (2017), « Contact, context, and collocation: The emergence of sociostylistic variation in L2 French learners during study abroad », *Studies in Second Language Acquisition*, 39(3), p. 553-578. [En ligne]. URL : <https://doi.org/10.1017/S0272263116000061> [Site consulté le 15 août 2023].

KLIMOVA, Blanka, et Marcel PIKHART (2020), « Cognitive and applied linguistics aspects of using social media: the impact of the use of Facebook on developing writing skills in learning English as a foreign language », *European Journal of Investigation in Health, Psychology and Education* 10(1), p. 110-118. [En ligne]. URL : <https://doi.org/10.3390/ejihpe10010010> [Site consulté le 20 décembre 2023].

KONAN, Koffi (2013), « Regard sur une forme langagière ivoirienne à partir de la structure “n’a qu’à+ infinitif” : valeur restrictive ou valeur modale ?=(Looking to an ivoirienne linguistic form from the structure “... n’a qu’à+ infinitif”: a restrictive or modal value) », *Estudios románicos*, n° 22, p. 43-52.

LABERGE, Suzanne (1977), *Étude de la variation des pronoms sujets définis et indéfinis dans le français parlé à Montréal*, Thèse de doctorat, Université de Montréal.

LABOV, William (1972), « The social motivation of a sound change », *Sociolinguistic Patterns*, Philadelphia, University of Pennsylvania University Press, p. 1-41.

LABOV, William (1973), *Sociolinguistic Patterns*, Philadelphia, University of Pennsylvania Press.

LABOV, William (1976), *Sociolinguistique*, Paris, Éditions de Minuit.

LABOV, William (2006), « The study of language in its social context », dans *The Social Stratification of English in New York City*, Cambridge, Cambridge University Press, p. 3-17.

LACAZE, Grégoire (2022), « Réflexions autour de la variation diamésique », *E-rea* 19(2). [En ligne]. URL : <http://journals.openedition.org.ezproxy.library.yorku.ca/erea/13428> [Site consulté le 27 septembre 2022].

LACERTE, Christine (2006), *Représentations sociales du code conversationnel du clavardage chez les jeunes et chez les experts québécois*, Mémoire de maîtrise, Université du Québec à Montréal. [En ligne]. URL : <https://archipel.uqam.ca/3237/1/M9617.pdf> [Site consulté le 7 septembre 2021].

LANGLAIS, Philippe et Patrick DROUIN (2012), « Text4Science: A Quebec French database of annotated text messages », *Linguisticae Investigationes*, 35(2), p. 237–259. [En ligne]. URL : <https://doi.org/10.1075/li.35.2.07lan> [Site consulté le 20 janvier 2022].

LARRIVÉE, Pierre et Alexander GURYEV (2021), « Variantes formelles de l'interrogation. Présentation », *Langue française*, 4, p. 9-24. [En ligne]. URL : <https://www.cairn.info/revue-langue-francaise-2021-4-page-9.htm> [Site consulté le 15 mars 2022].

LÉARD, Jean-Marcel (1983), « Le statut de *fak* en québécois: un simple équivalent de *alors* ? », *Travaux de Linguistique Québécoise*, 4, p. 59-100.

LEBRUN, Monique et Nathalie LACELLE (2012), « Les usages linguistiques des adolescents québécois sur les médias sociaux », *Alsic*, 15(1). [En ligne]. URL : <https://journals.openedition.org/alsic/2462> [Site consulté le 27 novembre 2020].

LEEMAN, Danielle (2004), « L'emploi de *juste* comme adverbe d'énonciation », *Langue française*, 142(2), p.17-30. [En ligne]. URL : <https://www.cairn.info/revue-langue-francaise-2004-2-page-17.htm> [Site consulté le 20 décembre 2024].

LOI SUR LES NOUVELLES EN LIGNE (2023), *Énoncé concernant la Charte, C-18*. [En ligne]. URL : [https://www.justice.gc.ca/fra/sjc-csj/pl/charte-charter/c18\\_1.html#shr-pg0](https://www.justice.gc.ca/fra/sjc-csj/pl/charte-charter/c18_1.html#shr-pg0) [Site consulté le 13 décembre 2023].

LU, Xuan, Wei AI, Xuanzhe LIU, Qian LI, Ning WANG, Gang HUANG, Qiaozhu MEI (2016), « Learning from the ubiquitous language: An empirical analysis of emoji usage of smartphone users », dans *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UBIComp*, septembre 12-16, Heidelberg, p. 770–780. [En ligne]. URL : <https://doi.org/10.1145/2971648.2971724> [Site consulté le 15 janvier 2023].

LUCKHARDT, Kristin (2009), *Stilanalysen zur Chat-Kommunikation : eine korpusgestützte Untersuchung am Beispiel eines medialen Chats*, Thèse de doctorat, Technische Universität Dortmund. [En ligne]. URL : <http://hdl.handle.net/2003/26055> [Site consulté le 16 janvier 2023].

LYSTER, Roy (1996), « Question forms, conditionals, and second-person pronouns used by adolescent native speakers across two levels of formality in written and spoken French », *The Modern Language Journal*, 80(2), p. 165-182. [En ligne]. URL : <https://doi.org/10.1111/j.1540-4781.1996.tb01157.x> [Site consulté le 20 août 2021].

MAGUÉ, Jean-Philippe *et al.* (2015-2020), *SoSweet : une sociolinguistique de Twitter*.

MAGUÉ, Jean-Philippe, ROSSI-GENSANE, Nathalie et Pierre HALTÉ (2020), « De la segmentation dans les tweets : signes de ponctuation, connecteurs, émoticônes et émojis », *Corpus*, 20. [En ligne]. URL : <http://journals.openedition.org/corpus/4619> [Site consulté le 20 août 2021].

MANSFIELD, John, Henry LESLIE-O'NEILL et Haoyi LI (2023), « Dialect differences and linguistic divergence: A crosslinguistic survey of grammatical variation », dans *Language Dynamics and Change* (à paraître), [En ligne]. URL : [https://brill.com/view/journals/ldc/13/2/article-p232\\_3.xml?language=en](https://brill.com/view/journals/ldc/13/2/article-p232_3.xml?language=en) [Site consulté le 20 décembre 2023].

MARCELLESI, Jean Baptiste et Bernard GARDIN (1974), *Introduction à la sociolinguistique : la linguistique sociale*, Paris, Larousse.

MARCHAL, Bruno (2014), « Réseaux sociaux et pratiques linguistiques : comment intégrer l'usage du français dans un environnement aussi concurrentiel que celui de l'anglais ? », *3ème Conférence internationale de Phuket - Le français face au plurilinguisme-pluriculturalisme : États de lieux et perspectives de la didactique de l'anglais et du français en contexte plurilingue* – English and French : Observations and teaching perspectives in multilingual and multicultural context, Université Prince de Songkla, Phuket, Thaïlande, Thaïlande. [En ligne]. URL : <https://halshs.archives-ouvertes.fr/halshs-02074179/document> [Site consulté le 20 décembre 2023].

MARCHESSAULT, Guy (2014), « Les impacts des Web 1.0 et Web 2.0 sur la religion : un effort de prospective », *Revue Lumen Vitae*, 69(1), p. 33-45. [En ligne]. URL : <https://www.cairn.info/revue-lumen-vitae-2014-1-page-33.htm> [Site consulté le 15 avril 2023].

MARCOCCIA, Michel (2004), « La communication écrite médiatisée par ordinateur : faire du face à face avec de l'écrit », *Journée d'études de l'ATALA Le traitement automatique des nouvelles formes de communication écrite*, juin 2004, Paris, France. [En ligne]. URL : <https://hal.archives-ouvertes.fr/hal-02952347/> [Site consulté le 20 août 2021].

MARCOCCIA, Michel (2012), « Définitions et négociations de la norme scripturale dans un forum de discussion d'adolescents », *Éla. Études de linguistique appliquée*, 166, p. 157-169. [En ligne]. URL : <https://www.cairn.info/revue-ela-2012-2-page-157.htm> [Site consulté le 21 août 2021].

MARCOCCIA, Michel (2016), *Analyser la communication numérique écrite*, Armand Colin.

MARCOCCIA, Michel et Nathalie GAUDUCHEAU (2007), « L'analyse du rôle des smileys en production et en réception : un retour sur la question de l'oralité des écrits numériques », *Glottopol : Revue de sociolinguistique en ligne*, 10, p. 39-55. [En ligne].

URL : [http://glottopol.univ-rouen.fr/telecharger/numero\\_10/gpl10\\_complet.pdf](http://glottopol.univ-rouen.fr/telecharger/numero_10/gpl10_complet.pdf) [Site consulté le 20 août 2021].

MARTEL, Pierre (2017), « Le français québécois », *Conférence prononcée devant les membres de l'APPRUS*. [En ligne]. URL :

<https://www.usherbrooke.ca/apprus/fileadmin/sites/apprus/documents/conferences/Conference-PierreMartel-17mai2017.pdf> [Site consulté le 11 janvier 2024].

MARTEL, Pierre et Hélène CAJOLET-LAGANIÈRE (1995), « Oui...au français québécois standard », *Interface*, 16 (5), p. 14-24. [En ligne]. URL :

[https://www.usherbrooke.ca/crifuq/fileadmin/sites/crifuq/contributions/CAJOLET\\_Interface\\_1995.pdf](https://www.usherbrooke.ca/crifuq/fileadmin/sites/crifuq/contributions/CAJOLET_Interface_1995.pdf) [Site consulté le 11 janvier 2024].

MARTEL, Pierre et Hélène CAJOLET-LAGANIÈRE (1996), *Le français québécois : usages, standard et aménagement*. Sainte-Foy, Québec: Institut québécois de recherche sur la culture.

MARTEL, Pierre et Hélène CAJOLET-LAGANIÈRE (2007), « Le français québécois : Une langue à plusieurs visages », *Visage*, 9 (2), p. 1-3. [En ligne]. URL :

[https://www.usherbrooke.ca/crifuq/fileadmin/sites/crifuq/contributions/CAJOLET\\_Sommets.pdf](https://www.usherbrooke.ca/crifuq/fileadmin/sites/crifuq/contributions/CAJOLET_Sommets.pdf) [Site consulté le 11 janvier 2024].

MARTINEAU, France et Raymond MOUGEON (2003), « Sociolinguistic research on the origins of *ne* deletion in European and Quebec French », *Language*, 79(1), p. 118-152. [En ligne]. URL : <https://www.jstor.org/stable/4489387> [Site consulté le 20 août 2021].

MARTINEAU, France et Raymond MOUGEON (2005), « *Vais, vas, m'as* in spoken Quebec and metropolitan French : A diachronic and dialectal perspective », communication, *Linguistics Symposium on Romance Languages*, Austin, University of Texas, février 2005.

MARTINEAU, France et Marie-Claude SÉGUIN (2016), « Le Corpus FRAN : réseaux et maillages en Amérique française », *Corpus*, 15. [En ligne]. URL : <http://journals.openedition.org/corpus/2925> [Site consulté le 14 janvier 2022].

MARTINET, André (1980), *Eléments de linguistique générale*, Paris, Armand Colin.

MASSICOTTE, France (1986), « Les expressions de la restriction en français de Montréal », dans SANKOFF, David [dir.], *Diversity and Diachrony*, Philadelphia, Benjamins, p. 325-332.

MASSICOTTE, France (1984), *Analyse sociolinguistique de RIEN QUE en français de Montréal*, Mémoire de maîtrise, Montréal, Université du Québec à Montréal.

MAURAS, Jacques (2003), *Analyse linguistique de 4000 courriels*, Conseil supérieur de la langue française. [En ligne]. URL : <https://numerique.banq.qc.ca/patrimoine/details/52327/46104 ?docref=EZttnwF7nVg7-4XKrUS9VQ> [Site consulté le 22 mars 2022].

MBEY MAKANG, Moïse (2014), « Étude transformationnelle de la syntaxe verbale en français du Cameroun : cas des locuteurs sous-scolarisés de Ngaoundéré », *Studii de gramatică contrastivă*, 21, 35–57. [En ligne]. URL : <https://www.ceeol.com/search/article-detail?id=673389> [Site consulté le 15 décembre 2023].

MCCULLOCH, Gretchen (2019), *Because Internet: Understanding the New Rules of Language*, New York, Riverhead Books.

MCCULLOCH, Gretchen et Jeffrey LAMONTAGNE (2020), « La phonologie du français sur Twitter », dans BIGOT, Davy *et al.* [dir.], *Les français d'ici en perspective*, Presses de l'Université Laval, p. 171-195.

McLUHAN, Marshall (1994), *Understanding Media: The Extensions of Man*, MIT Press.

MEISNER, Charlotte (2017), *La variation pluridimensionnelle: une analyse de la négation en français*, Peter Lang International Academic Publishers.

META (2023), « Changes to News Availability on Our Platforms in Canada », About, *Facebook*. [En ligne]. URL : <https://about.fb.com/news/2023/06/changes-to-news-availability-on-our-platforms-in-canada/#translation> [Site consulté le 13 décembre 2023].

META (2024), « Éviter les spams et les arnaques », Pages d'aide, *Facebook*. [En ligne]. URL : <https://fr-ca.facebook.com/help/1584206335211143> [Site consulté le 17 janvier 2024].

MILETTE, Mélanie (2012), *Médias sociaux: Enjeux pour la communication*, Québec, Les Presses de l'Université du Québec. [En ligne]. URL : <https://doi.org/10.2307/j.ctv18pgp4w> [Site consulté le 20 décembre 2023].

MILROY, Lesley (1987), *Language and Social Networks*, Blackwell.

MIONI, Alberto (1983), « Italiano tendenziale : Osservazioni su alcuni aspetti della standardizzazione », *Scritti linguistici in onore di Giovan Battista Pellegrini*, 2, p. 495-517.

MOREAU, Marie-Louise (1997), *Sociolinguistique : les concepts de base*, Mardaga.

MOSEGAARD-HANSEN, Maj-Britt (1997), « *Alors* and *donc* in spoken French: A reanalysis », *Journal of Pragmatics* 28(2), p. 153-187. [En ligne]. URL : [https://doi.org/10.1016/S0378-2166\(96\)00086-0](https://doi.org/10.1016/S0378-2166(96)00086-0) [Site consulté le 15 décembre 2023].

MOUGEON, Raymond (1976), « Bilingualism and language maintenance in the Gaspé Peninsula, Quebec, Canada », *Anthropological Linguistics*, 18(2), p. 53–69.

MOUGEON, Raymond (2013), « Recherche sociolinguistique historique sur l'emploi des verbes *aller* et *s'en aller* exprimant le futur en français parlé hexagonal », communication au Département d'études françaises de l'Université York, campus Keele, 1er mai.

MOUGEON, Raymond, Sandrine HALLION, Davy BIGOT et Robert PAPEN (2016), « Convergence et divergence sociolinguistique en français laurentien: l'alternance *rien que/juste/seulement/seulement que/ne. . . que* », *Journal of French Language Studies*, 26(2), p. 115–154. [En ligne]. URL : <https://doi.org/10.1017/S095926951600003X> [Site consulté le 20 mars 2022].

MOUGEON, Raymond, Terry NADASDI et Katherine REHNER (2002), « État de la recherche sur l'appropriation de la variation par les apprenants avancés du FL2 ou FLE », *Acquisition et interaction en langue étrangère*, 17, p. 7-50. [En ligne]. URL : <https://doi.org/10.4000/aile.847> [Site consulté le 20 mars 2022].

MOUGEON, Raymond et Katherine REHNER (2013), « Socio-stylistic dimensions of the (non) use of negative particle *ne* in French Canadian high schools ». [En ligne]. URL : [https://www.glendon.yorku.ca/rmougeon/website\(publish\)/documents/nedeletionoct2013.pdf](https://www.glendon.yorku.ca/rmougeon/website(publish)/documents/nedeletionoct2013.pdf) [Site consulté le 1 aout 2023].

MOUGEON, Raymond et Katherine REHNER (2015), « Stylistic and discursive functions of French negative particle *ne* in an educational context », *Journal of Sociolinguistics*, 19(5), p. 585-611. [En ligne]. URL : <https://doi.org/10.1111/josl.12164> [Site consulté le 2 aout 2023].

NADASDI, Terry et Christina KEPPIE (2004), « Le mot *juste* en français albertain », dans PAPEN, Robert *et al.* [dir.], *Les parlers français de l'Ouest*, Cahiers franco-canadiens de l'Ouest, Presses universitaires de Saint-Boniface, 16(1–2), p. 67–78.

NGUYEN, Thuy (2015), *Code Switching: A Sociolinguistic Perspective*, Anchor, Hamburg.

NUGRAHA, Danang Satria (2024), « Using Big Data Approach to Create a Treebank of Informal and Formal Indonesian », *South Asian Research Journal of Engineering and Technology*, 6(1), p. 1-8. [En ligne]. URL : [https://sarpublishation.com/media/articles/SARJET\\_61\\_1-8.pdf](https://sarpublishation.com/media/articles/SARJET_61_1-8.pdf) [Site consulté le 20 décembre 2023].

NUNBERG, Geoffrey (1999), « Les enjeux linguistiques d'Internet », *Critique internationale*, 4, p. 105-121. [En ligne]. URL : [https://www.persee.fr/doc/criti\\_1290-7839\\_1999\\_num\\_4\\_1\\_1525](https://www.persee.fr/doc/criti_1290-7839_1999_num_4_1_1525) [Site consulté le 20 août 2021].

OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE (2023), « Emploi des auxiliaires avec le verbe *passer* », *Banque de dépannage linguistique*. [En ligne]. URL : <https://vitrinelinguistique.oqlf.gouv.qc.ca/24106/la-grammaire/le-verbe/auxiliaires/emploi-des-auxiliaires-avec-le-verbe-passer> [Site consulté le 27 juillet 2023].

OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE (2023), « Le futur proche », *Banque de dépannage linguistique*. [En ligne]. URL : <https://vitrinelinguistique.oqlf.gouv.qc.ca/24122/la-grammaire/le-verbe/temps-grammaticaux/futur/le-futur-proche> [Site consulté le 1 août 2023].

OKSUZ, Dogus Can, Vaclav BŘEZINA et Patrick REBUSCHAT (2021), « Collocational processing in L1 and L2: The effects of word frequency, collocational frequency, and association », *Language Learning*, 71(1), p. 55-98. [En ligne]. URL : <https://doi.org/10.1111/lang.12427> [Site consulté le 20 avril 2023].

ONG, Walter J. (2002), *Orality and Literacy : The Technologizing of the Word* (2e éd.), Routledge.

ORGANISATION INTERNATIONALE DE LA FRANCOPHONIE (2015), *Veille internationale sur la culture et le commerce et numérique*, 10(3), Centre d'études sur l'intégration et la mondialisation, Université du Québec à Montréal. [En ligne]. URL : <https://www.ieim.uqam.ca/IMG/pdf/oif-volume10-numero3avril2015ceim.pdf> [Site consulté le 22 décembre 2023].

PAGE, Ruth, David BARTON, Johann W. UNGER, et Michele ZAPPAVIGNA (2014), *Researching Language and Social Media : A Student Guide*, Routledge, Taylor & Francis Group.

PALFREYMAN, David et Muhamed al KHALIL (2003), « “A funky language for teenzz to use:” Representing Gulf Arabic in instant messaging », *Journal of Computer-Mediated Communication*, 9(1). [En ligne]. URL : <https://academic.oup.com/jcmc/article/9/1/JCMC917/4614456> [Site consulté le 15 décembre 2023].

PANCKHURST, Rachel (1998), « Analyse linguistique du courrier électronique », dans GUÉGEN, Nicolas et Laurence TOBIN [dir.] *Communication, société et Internet*, Paris, L’Harmattan, p. 47-60.

PAOLILLO, John C. (2001), « Language variation on Internet Relay Chat: A social network approach », *Journal of Sociolinguistics*, 5(2), p. 180–213. [En ligne]. URL : <https://doi.org/10.1111/1467-9481.00147> [Site consulté le 12 janvier 2023].

PAOLILLO, John C. (2016), « Network Analysis », dans GEORGAKOPOULOU, Alexandra et Tereza SPILIOTI [dir.], *The Routledge Handbook of Language and Digital Communication*, London & New York, Taylor and Francis Group, p. 36-54.

PAVEAU, Marie-Anne (2015), « Ce qui s’écrit dans les univers numériques », *Itinéraires* 2014-1. [En ligne]. URL : <http://journals.openedition.org/itineraires/2313> [Site consulté le 22 novembre 2022].

PAVEAU, Marie-Anne (2019), « Introduction. Écrire, parler, communiquer en ligne : nos vies sociolangagières connectées », *Langage et société*, 167, p. 9-28. [En ligne]. URL : <https://www.cairn.info/revue-langage-et-societe-2019-2-page-9.htm> [Site consulté le 22 novembre 2022].

PEREA, François (2010), « L’identité numérique : de la cité à l’écran. Quelques aspects de la représentation de soi dans l’espace numérique », *Les Enjeux de l’information et de la communication*, p. 144-159. [En ligne]. URL : <https://doi.org/10.3917/enic.010.0800> [Site consulté le 20 août 2021].

PERRIAULT, Jacques (2012), « Réseaux de communication horizontale, un aperçu à travers le temps », *Le temps et les médias*, n° 18, p. 143-158. [En ligne]. URL : <https://www.cairn.info/revue-le-temps-des-medias-2012-1-page-148.htm> [Site consulté le 15 février 2024].

PIEROZAK, Isabelle (2007), « Prendre Internet pour terrain », *Glottopol, Revue de sociolinguistique en ligne : Regards sur l'internet, dans ses dimensions langagières. Penser les continuités et discontinuités. En hommage à Jacques Anis*, 10, p. 4-10. [En ligne]. URL : [http://glottopol.univ-rouen.fr/telecharger/numero\\_10/gpl10\\_00pierzak.pdf](http://glottopol.univ-rouen.fr/telecharger/numero_10/gpl10_00pierzak.pdf) [Site consulté le 13 décembre 2022].

POIRIER, Christian *et al.*, (2023), *Pratiques de découverte de contenus culturels et environnements numériques : Regards sur la découverte culturelle au Québec*, Institut national de la recherche scientifique, Centre Urbanisation Culture Société, Chaire Fernand-Dumont sur la culture. [En ligne]. URL : [https://espace.inrs.ca/id/eprint/13245/1/D%C3%A9couverteCulturelle\\_Poirier.pdf](https://espace.inrs.ca/id/eprint/13245/1/D%C3%A9couverteCulturelle_Poirier.pdf) [Site consulté le 20 décembre 2023].

POIRIER, Franck (2017), « Quelle modalité pour l'interaction avec les petits appareils mobiles et vestimentaires : texte ou vocal ? Comment choisir entre clavier et assistant personnel ? », *29ème conférence francophone sur l'interaction homme-machine*, Août 2017, Poitiers, France. [En ligne]. URL : <https://hal.archives-ouvertes.fr/hal-01577674/document> [Site consulté le 21 mars 2022].

POIRIER, Franck et Igor SCHADLE (2004), « État de l'art des méthodes de saisie de données sur dispositifs nomades : typologie des approches », dans *Proceedings of the 16th Conference on l'interaction homme-machine (IHM 04)*, Association for computing machinery, New York, NY, USA, p. 133–140. [En ligne]. URL : <https://doi.org/10.1145/1148613.1148632> [Site consulté le 27 mars 2022].

POPLACK, Shana et Anne ST-AMAND (2007), « A real-time window on 19th-century vernacular French: The *Récits du français québécois d'autrefois* », *Language in society*, 36(5), p. 707-734. [En ligne]. URL : <https://doi.org/10.1017/S0047404507070662> [Site consulté le 27 janvier 2024].

POPLACK, Shana et Anne ST-AMAND (2009), « Les Récits du français québécois d'autrefois : Reflet du parler vernaculaire du 19e siècle », *Canadian Journal of Linguistics*, 54(3), p. 511–546.

POUDAT, Céline, Ciara R. WIGHAM et Loïc LIÉGEOIS (2020), « Les corpus de la communication médiée par les réseaux : une introduction », *Corpus*, 20. [En ligne]. URL : <http://journals.openedition.org/corpus/4720> [Site consulté le 27 mars 2022].

PRADO, Daniel (2016), « Les langues romanes minoritaires et l'Internet », *Hermès, La Revue*, 75, p. 101-110. [En ligne]. URL : <https://www.cairn.info/revue-hermes-la-revue-2016-2-page-101.htm> [Site consulté le 15 octobre 2023].

QUONIAM Luc et Charles-Victor BOUTET (2008), « Web 2.0, la révolution connectique », *Document numérique*, 11(1-2), p. 133-143. [En ligne]. URL : <https://www.cairn.info/revue-document-numerique-2008-1-page-133.htm> [Site consulté le 10 avril 2023].

RAINIE Lee et Barry WELLMAN (2012), *Networked : The New Social Operating System*, Cambridge Massachusetts, MIT Press.

RAND, David et David SANKOFF (1990), *GoldVarb - Version 2.0 : User Manual*.

RÉA, Béatrice (2016), « “Je m’ ai fait mal quand j’ ai tombé” : Questions d’ auxiliarité et de réflexivité dans le français parlé à Montréal (1971-2013) », *ScriptUM : la revue du colloque VocUM, 1, Revue de l’ édition 2014 de VocUM : Colloque multidisciplinaire sur le langage*. [En ligne]. URL : <https://scriptum.vocum.ca/index.php/scriptum/article/view/26> [Site consulté le 15 août 2023].

RÉA, Béatrice (2020), *Je m’ ai fait mal quand j’ ai tombé: a real- and apparent-time study of auxiliary alternation in intransitive and pronominal verbs in spoken Montréal French (1971-2016)*, Thèse de doctorat, Université Oxford. [En ligne]. URL : <https://ora.ox.ac.uk/objects/uuid:0546b011-d2a2-41cf-80ba-e96613def21b> [Site consulté le 15 août 2023].

REHNER, Katherine et Raymond MOUGEON (1998), « Use of restrictive expressions *juste, seulement, and rien que* in Ontario French », *Revue de l’ ACLA*, 19(1-2), p. 89-110. [En ligne]. URL : <https://yorkspace.library.yorku.ca/server/api/core/bitstreams/14544361-078e-42c6-8218-c2e67f3e6157/content> [Site consulté le 12 juillet 2023].

REINKE, Kristin et Luc OSTIGUY (2016), *Le français québécois d’ aujourd’ hui*, De Gruyter.

ROSENBACH, Anette (2007), « Exploring constructions on the web: a case study », dans HUNDT, Marianne, Nadja NESSELHAUF et Carolin BIEWER [dir.], *Corpus Linguistics and the Web*, Internationales Wissenschaftsforum, Fritz Thyssen-Stiftung, Universität Heidelberg, Stiftung Universität Heidelberg.

ROULEAU, Héloïse (2019), *Nouvel essor du rap québécois : développement numérique d’ une culture en marge de l’ industrie*, Mémoire de maîtrise, Université de Montréal, Faculté de musique. [En ligne]. URL : <https://papyrus.bib.umontreal.ca/xmlui/handle/1866/24566> [Site consulté le 20 décembre 2023]

SANKOFF, David, Sali A. TAGLIAMONTE, et Eric SMITH (2018), *Goldvarb Z: A Variable Rule Application for Macintosh*, Department of Linguistics, University of Toronto.

SANKOFF, Gillian (1973/1980), *The Social Life of Language*, University of Pennsylvania Press.

SANKOFF, Gillian et Pierrette THIBAUT (1977), « L'alternance entre les auxiliaires "avoir" et "être" en français parlé à Montréal », *Langue Française*, 34, p. 81-108. [En ligne]. URL : [https://www.persee.fr/doc/lfr\\_0023-8368\\_1977\\_num\\_34\\_1\\_4818](https://www.persee.fr/doc/lfr_0023-8368_1977_num_34_1_4818) [Site consulté le 20 août 2021].

SANKOFF, Gillian et Diane VINCENT (1977), « L'emploi productif du *ne* dans le français parlé à Montréal », *Français moderne*, 45, p. 243–.

SCHWARTZ, Andrew H., Johannes C. EICHSTAEDT, Margaret L. KERN, Lukasz DZIURZYNSKI, Stephanie M. RAMONES, Megha AGRAWAL, Achal SHAH, Michal KOSINSKI, David STILLWELL, Martin E. P. SELIGMAN, Lyle H. UNGAR, (2013) « Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach », *PLoS ONE* 8(9). [En ligne]. URL : <https://doi.org/10.1371/journal.pone.0073791> [Site consulté le 9 juillet août 2024].

SCHWÖRER, Jakob et Xavier ROMERO-VIDAL (2020), « Radical right populism and religion: mapping parties' religious communication in Western Europe », *Religion, State and Society*, 48(1), p. 4–21. [En ligne]. URL : <https://doi.org/10.1080/09637494.2019.1704611> [Site consulté le 20 août 2021].

SEARGEANT, Philipp et Caroline TAGG (2014), *The Language of Social Media : Identity and Community on the Internet*, Palgrave Macmillan.

SHAHBAZNEZHAD, Hamidreza, Rebecca DOLAN, et Mona RASHIDIRAD (2021), « The role of social media content format and platform in users' engagement behavior », *Journal of Interactive Marketing*, 53(1), p. 47-65. [En ligne]. URL : <https://doi.org/10.1016/j.intmar.2020.05.001> [Site consulté le 15 décembre 2022].

SIEBENHAAR, Beat (2008), « Quantitative Approaches to Linguistic Variation in IRC: Implications for Qualitative Research », *Language@Internet*, 5. [En ligne]. URL : <https://www.languageatinternet.org/articles/2008/1615> [Site consulté le 20 août 2021].

SIEMUND, Peter (2008), « Linguistic Universals and Vernacular Data », dans FILIPPULA, Markku, Juhani KLEMOLA, et Heli PAULASTO [dir.], *Vernacular Universals and Language Contacts: Evidence from Varieties of English and Beyond*, London, Routledge.

SPILIOTI, Tereza (2017), « Media convergence and publicness: Towards a modular and iterative approach to online research ethics », *Applied Linguistics Review*, 8(2–3), p. 191–212. [En ligne]. URL : <https://doi.org/10.1515/applirev-2016-1035> [Site consulté le 17 juillet 2022].

STÆHR, Andreas (2014), *Social Media and Everyday Language Use Among Copenhagen Youth*, Thèse de doctorat, Department of Scandinavian Studies and Linguistics, University of Copenhagen. [En ligne]. URL : [https://nors.ku.dk/ansatte/?pure=files%2F122495013%2FPh.d.\\_2014\\_Staehr.pdf](https://nors.ku.dk/ansatte/?pure=files%2F122495013%2FPh.d._2014_Staehr.pdf) [Site consulté le 15 juin 2022].

STAEHR, Andreas (2015), « Reflexivity in Facebook interaction – Enregisterment across written and spoken language practices », *Discourse, Context & Media*, 8, p. 30-45. [En ligne]. URL : <https://doi.org/10.1016/j.dcm.2015.05.004> [Site consulté le 15 juin 2022].

STAEHR, Andreas (2016a), « Linguaging and normativity on Facebook », dans ARNAUT, Karel *et al.* [dir.], *Engaging Superdiversity: Recombining Spaces, Times and Language Practices*, Multilingual Matters, Bristol, Blue Ridge Summit, p. 170-195.

STAEHR, Andreas (2016b), « Normativity as a social resource in social media practices », dans MADSEN, Lian Malai *et al.* [dir.], *Everyday Linguaging: Collaborative Research on the Language Use of Children and Youth*, Berlin, München, Boston, De Gruyter Mouton, p. 71-94.

STANFORD, James (2016), « A call for more diverse sources of data: Variationist approaches in non-English contexts », *Journal of Sociolinguistics*, 20, p. 525-541. [En ligne]. URL : <https://doi.org/10.1111/josl.12190> [Site consulté le 12 décembre 2021].

STARK, Elisabeth (2012), « Negation marking in French text messages », *Linguisticae Investigationes*, 35(2), p. 341–366. [En ligne]. URL : <https://doi.org/10.1075/li.35.2.11sta> [Site consulté le 15 août 2022].

STARK, Elisabeth (2014), « Réanalyses dans la graphie : « L’écrit spontané » dans les SMS et le statut des pronoms clitiques du français contemporain », *Langages (Paris)*, 196, p. 131–148. [En ligne]. URL : <https://www.cairn.info/revue-langages-2014-4-page-131.htm> [Site consulté le 15 août 2022].

STARK, Elisabeth et Isabella RIEDEL (2012), « L’accord du participe passé dans les SMS francophones du corpus SMS suisse », *Romanistisches Jahrbuch*, 63(1), p. 116–138. [En ligne]. URL : [https://www.zora.uzh.ch/id/eprint/88692/1/Stark\\_L%27accord\\_du\\_participe\\_pass%C3%A9.pdf](https://www.zora.uzh.ch/id/eprint/88692/1/Stark_L%27accord_du_participe_pass%C3%A9.pdf) [Site consulté le 15 août 2022].

STARK, Elisabeth, Simon UEBERWASSER et Beni RUEF (2009-2015), *Swiss SMS Corpus*, University of Zurich.

STORRER, Angelika (2014), « Sprachverfall durch internetbasierte Kommunikation ? : Linguistische Erklärungsansätze – empirische Befunde », dans PLEWNIA, Albrecht et Andreas WITT [dir.], *Sprachverfall ? : Dynamik - Wandel – Variation*, Berlin, Boston, De Gruyter, p. 171-196.

SVIDRONOVA, Maria Murray, Alena KASCAKOVA et Gabriela BAMBUSEKOVA (2019), « Social media in the presidential election campaign: Slovakia 2019 », *Adm si Manag Public*, 33, p. 181-194. [En ligne]. URL : <https://www.proquest.com/openview/920b59d416341c82fdc692d834c09b51/1?pq-origsite=gscholar&cbl=986350> [Site consulté le 20 août 2021].

SZMRECSANYI, Benedikt et Alexandra ENGEL (2022), « A variationist perspective on the comparative complexity of four registers at the intersection of mode and formality », *Corpus Linguistics and Linguistic Theory*, 19(1), p. 79-113. [En ligne]. URL : <https://doi.org/10.1515/cllt-2022-0031> [Site consulté le 10 janvier 2023].

TAGG, Caroline (2015), *Exploring Digital Communication: Language in Action*, Taylor and Francis.

TAGLIAMONTE, Sali (2006), *Analysing Sociolinguistic Variation*, Cambridge University Press.

TAGLIAMONTE, Sali A. et Derek DENIS (2008), « Linguistic ruin ? Lol! Instant messaging and teen language », *American Speech*, 83(1), p. 3–34. [En ligne]. URL : <https://doi.org/10.1215/00031283-2008-001> [Site consulté le 20 décembre 2022].

TANNEN, Deborah (2013), « Medium is the meta-message: conversational style in new media interaction », dans TANNEN, Deborah et Anna Marie TRESTER [dir.], *Discourse 2.0 : Language and New Media*, Washington, DC, Georgetown University Press.

TANNEN, Deborah et Anna Marie TRESTER (2013), *Discourse 2.0 : Language and New Media*, Washington, DC, Georgetown University Press.

TATOSSIAN, Anaïs (2020), « Nouvelles pratiques orthographiques à l'ère du numérique », dans REINKE, Kristin [dir.], *Attribuer un sens : La diversité des pratiques langagières et les représentations sociales*, Les Presses de l'Université Laval, p. 101-123.

TÉTU, Martin (2012), « Des vertus culturelles du piratage à l'ère numérique : ou comment le *peer-to-peer* peut contribuer à la circulation du patrimoine québécois et à la diversité culturelle », *Éthique publique*, 14(2). [En ligne]. URL : <http://journals.openedition.org/ethiquepublique/1015> [Site consulté le 20 décembre 2023].

THIBAULT, Pierrette (1991), « La langue en mouvement : simplification, régularisation, restructuration », *Linx : Ces langues que l'on dit simple*, 25, p. 79-92. [En ligne]. URL : [https://www.persee.fr/doc/linx\\_0246-8743\\_1991\\_num\\_25\\_2\\_1228](https://www.persee.fr/doc/linx_0246-8743_1991_num_25_2_1228) [Site consulté le 20 août 2021].

THIBAULT, Pierrette et Michelle DAVELUY (1989), « Quelques traces du passage du temps dans le parler des Montréalais, 1971-1984 », *Language Variation and Change*, 1(1), p. 19-45. [En ligne]. URL : <https://doi.org/10.1017/S0954394500000107> [Site consulté le 20 août 2021].

THIELE, Daniel et Tjasa TURNSEK (2022), « How right-wing populist comments affect online deliberation on news media Facebook pages », *Media and Communication (Lisboa)*, 10(4), p. 141–154. [En ligne]. URL : <https://doi.org/10.17645/mac.v10i4.5690> [Site consulté le 15 janvier 2023].

THOMAS, Alain (2015), « *Nous/on* : De la réalité linguistique à la salle de classe », *Arborescences*, (5), p. 126–138. [En ligne]. URL : <https://doi.org/10.7202/1032665ar> [Site consulté le 20 août 2021].

THURLOW, Crispin (2003), « Generation txt ? The sociolinguistics of young people's text messaging », *Discourse Analysis Online*, 1(1). [En ligne]. URL : <https://extra.shu.ac.uk/daol/articles/v1/n1/a3/thurlow2002003.html> [Site consulté le 20 août 2021].

TOIVONEN, Jade (2023), *Ça fait un bon moment que j'ai pas posté on here-l'utilisation de l'anglais dans les vlogs français et québécois sur YouTube*, Thèse de doctorat, Université de Helsinki.

TREMBLAY, Mireille (2016), « Les constructions partitives pronominales en français: une analyse de corpus », *Actes du CMLF 2016 - 5e Congrès mondial de linguistique française*. [En ligne]. URL : <https://doi.org/10.1051/shsconf/20162714010> [Site consulté le 15 novembre 2023].

TREMBLAY, Mireille (2020), « Le texto : une pratique langagière distincte ? », dans REINKE, Kristin [dir.], *Attribuer un sens : La diversité des pratiques langagières et les représentations sociales*, Les Presses de l'Université Laval, p. 79-100.

TREMBLAY, Mireille, Hélène BLONDEAU et Emanuelle LABEAU (2019), « Texting the future in Belgium and Québec: Present matters », *Journal of French Language Studies*, 30(1), p. 73–98. [En ligne]. URL : <https://doi.org/10.1017/S0959269519000188> [Site consulté le 15 janvier 2024].

van COMPERNOLLE, Rémi A. (2007), « The variable use of *ne* in public French-language chat », dans GERBAULT, Jeannine [dir.], *La Langue du cyberspace: de la diversité aux normes*, L'Harmattan, Paris, p. 251–264.

van COMPERNOLLE, Rémi A. (2008a), « Morphosyntactic and phonological constraints on negative particle variation in French-language chat discourse », *Language Variation and Change*, 20(2), p. 317–339. [En ligne]. URL : <https://doi.org/10.1017/S0954394508000112> [Site consulté le 20 août 2021].

van COMPERNOLLE, Rémi A. (2008b), « Language variation in online personal ads from Quebec: The case of *ne* », *Language@Internet*, 5(1). [En ligne]. URL : <https://www.languageatinternet.org/articles/2008/1666> [Site consulté le 20 août 2021].

van COMPERNOLLE, Rémi A. (2008c), « *Nous* versus *on* : Pronouns with first person plural reference in synchronous French chat », *Canadian Journal of Applied Linguistics*, 11(2), p. 85–110. [En ligne]. URL : <https://journals.lib.unb.ca/index.php/CJAL/article/view/19917> [Site consulté le 20 août 2021].

van COMPERNOLLE, Rémi A. (2009), « Emphatic *ne* in informal spoken French and implications for foreign language pedagogy », *International Journal of Applied Linguistics*, 19(1), p. 47-65. [En ligne]. URL : <https://doi.org/10.1111/j.1473-4192.2009.00213.x> [Site consulté le 20 août 2021].

van COMPERNOLLE, Rémi A. (2010), « The (slightly more) productive use of *ne* in Montreal French chat », *Language Sciences (Oxford)* 32(4), p. 447–463. [En ligne]. URL : <https://doi.org/10.1016/j.langsci.2009.07.004> [Site consulté le 20 décembre 2023].

van COMPERNOLLE, Rémi et WILLIAMS, Lawrence (2007), « De l'oral à l'électronique : la variation orthographique comme ressource sociostylistique et pragmatique dans le français électronique », *Glottopol : Revue de sociolinguistique en ligne*, 10, p. 56-69. [En ligne]. URL : [http://glottopol.univ-rouen.fr/telecharger/numero\\_10/gp110\\_complet.pdf](http://glottopol.univ-rouen.fr/telecharger/numero_10/gp110_complet.pdf) [Site consulté le 20 août 2021].

van DIJCK, José (2013), *The Culture of Connectivity : A Critical History of Social Media*, Oxford University Press.

VARIS, Piia (2016), « Digital Ethnography », dans GEORGAKOPOULOU, Alexandra et Tereza SPILIOTI [dir.], *The Routledge Handbook of Language and Digital Communication*, London & New York, Taylor and Francis Group, p. 55-68.

VÉZINA, Robert (2009), *La question de la norme linguistique*, Québec, Conseil supérieur de la langue française, [En ligne]. URL : <https://www.cslf.gouv.qc.ca/publications/pubf302/cslfnormelinguistique.pdf> [Site consulté le 22 mars 2022].

VILLENEUVE, Anne-José (2017), « Normes objectives et variation socio-stylistique : le français québécois parlé en contexte d'entrevues télévisées », *Arborescences*, 7, p. 49–66. [En ligne]. URL : <https://doi.org/10.7202/1050968ar> [Site consulté le 20 août 2021].

VICENTINI, Alessandra (2003), « The economy principle in language: notes and observations from early modern English grammars », *MPW - Mots Palabras Words*, 3, p. 37-57. [En ligne]. URL : <https://www.ledonline.it/mpw/allegati/mpw0303vicentini.pdf> [Site consulté le 18 janvier 2024].

VITALI-ROSATI, Marcello (2014), *Égarements. Amour mort et identités numériques*, Hermann.

WALKER, James A. (2010), *Variation in Linguistic Systems*, Florence, Routledge.

WILLIAMS, Geoffrey (1998), « Collocational networks: Interlocking patterns of lexis in a corpus of plant biology research articles », *International Journal of Corpus Linguistics*, 3(1), p. 151–171. [En ligne]. URL : <https://doi.org/10.1075/ijcl.3.1.07wil> [Site consulté le 20 août 2023].

WILLIAMS, Geoffrey (2002), « In search of representativity in specialised corpora: Categorisation through collocation », *International Journal of Corpus Linguistics*, 7(1), p. 43–64. [En ligne]. URL : <https://doi.org/10.1075/ijcl.7.1.03wil> [Site consulté le 20 août 2023].

WILLIAMS, Geoffrey (2006), « La linguistique et le corpus: Une affaire prépositionnelle » *Texte, revue de linguistique en ligne*. [En ligne].  
URL : <http://www.revetexto.net/Parutions/Livres-E/Albi-2006/Williams.pdf> [Site consulté le 26 juin 2024].

WILLIAMS, Lawrence (2009), « Sociolinguistic variation in French computer-mediated communication: A variable rule analysis of the negative particle *ne* », *International Journal of Corpus Linguistics*, 14(4), p. 467–491. [En ligne]. URL : <https://doi.org/10.1075/ijcl.14.4.02wil> [Site consulté le 20 août 2021].

WOEHLING, José (2001), « Les anciens et les modernes : une conciliation difficile au Québec », *Raisons politiques*, 2, p. 195-206. [En ligne]. URL : <https://doi.org/10.3917/rai.002.0195> [Site consulté le 20 décembre 2023].

WONG GONZALES, Wilkinson Daniel (2024), « Sociolinguistic Analysis with Missing Metadata ? Leveraging Linguistic and Semiotic Resources Through Deep Learning to Investigate English Variation and Change on Twitter », *Applied Linguistics* (à paraître), Oxford University Press. [En ligne]. URL : <https://doi.org/10.1093/applin/amad086> [Site consulté le 5 février 2024].

WÜEST, Jakob (2009), « La notion de diamésie est-elle nécessaire ? », *Travaux de linguistique*, 59, p. 147-162. [En ligne]. URL : <https://www.cairn.info/revue-travaux-de-linguistique-2009-2-page-147.htm> [Site consulté le 15 décembre 2023].

XIAO, Robert (2015), « Collocation », dans BIBER, Douglas et Randi REPPEN [dir.], *The Cambridge Handbook of English Corpus Linguistics*, Cambridge University Press, p. 106-124.

YATES, Simeon (1996), « Oral and written linguistic aspects of computer conferencing : a corpus based study », dans HERRING, Susan [dir.], *Computer-Mediated Communication. Linguistic, Social and Cross-Cultural Perspectives*, Amsterdam / Philadelphia, John Benjamins, p. 29-46.

ZHANG, Yi (2020), « Adopting Japanese in a popular Chinese video-sharing website: heteroglossic and multilingual communication by online users of Bilibili.com », *International Multilingual Research Journal*, 14(1), p. 20–40. [En ligne]. URL : <https://doi.org/10.1080/19313152.2019.1627856> [Site consulté le 14 août 2023].

**Annexe : Bilan des résultats d'études antérieures**

Variantes				
Lexique				
Conséquence	Donc	Alors	(ça) fait que	So
Blondeau <i>et al.</i> (2019) oral	19 %	5 %	76 %	0 %
Blondeau et Tremblay (2022) oral	18,4 %	2,6 %	78,9 %	Exclu
Martineau et Séguin (2016) oral	16,47 %	15,68 %	67,72 %	0,12 %
Contrainte	Juste	Seulement	Rien que	Ne...que
Massicotte (1986) oral	39,7 %	19 %	41,3 %	Exclu
Thibault et Daveluy (1989) oral	1971 - 41 % 1984 - 58 % Moyenne 52 %	1971 - 25 % 1984 - 12 % Moyenne 17 %	1971 - 33 % 1984 - 30 % Moyenne 31 %	Exclu
Verbes				
Futur	Simple (analytique)		Périphrastique (synthétique)	
Blondeau (2006) oral	1971 - 14 % 1984 - 23 % 1995 - 20 %		1971 - 86 % 1984 - 77 % 1995 - 80 %	
Tremblay <i>et al.</i> (2019) SMS	63 %		37 %	
Syntaxe				
Négation	Ne		Ø	
Sankoff et Vincent (1977) oral	0,5 %		99,5 %	
van Compernelle (2010) (t)chat	5,1 %		94,9 %	

Interrogation	-tu	Inversion	Est-ce que	Sans inversion
Bally (2022) SMS	15,1 %	19,1 %	8,5 %	57,3 %
Bigot (2011) oral formel	4,7 %	9,3 %	79 %	7 %
Halvorsen (2021) oral	15,6 %	4,8 %	8,5 %	71,1 %
Lyster (1996) oral informel	16,7 %	44 %	37,4 %	1,9 %
Villeneuve (2017) oral formel	6 %	13,5 %	37,3 %	43,3 %
Pronoms				
Simplets et composés	Nous-nous autres	Vous-vous autres	Eux-eux-autres	
Blondeau (2000) oral	Nous 15 % N-autres 85 %	Vous 0 % V-autres 100 %	Eux 12 % Eux-autres 88 %	
Blondeau et al (2014) SMS	Sur 79 occurrences (1e, 2e et 3e personnes du pluriel) - 1 occurrence composée			
Nous et on	Nous		On	
Blondeau et al (2014) SMS	3,7 %		96,3 %	
Laberge (1977) oral	1,6 %		98,4 %	

Alternance des auxiliaires avoir et être au passé composé		
Verbes	Fréquences d'emploi de formes non standards	
	Sankoff et Vincent (1977) oral	Réa (2020) oral
Aller	0,7 %	0,4 %
Arriver	9 %	0,6 %
Décéder	N/A	N/A
Descendre	50 %	32,4 %
Devenir	N/A	N/A
Rendre	N/A	N/A
Rester	N/A	19,4 %
Revenir	6,5 %	1,3 %
Partir	64 %	2,1 %
Passer	10 %	59 %
Sortir	33 %	19 %
Venir	7 %	0 %