

Bachelorarbeit

zur Erlangung des Grades eines Bachelor of Arts
an der Universität Trier im Fachbereich II - Romanistik

**Les Gros Mots en français : quand la linguistique de corpus
rencontre la sociolinguistique**

Sommersemester 2023

Vorgelegt von:

Name: Quoc Hoang Anh, To
Matrikelnummer: 1554803
E-Mail: s2antooo@uni-trier.de
Adresse: 124 Bis Avenue de Flandre
75019, Paris, Frankreich
Abgabedatum: Paris, 10. September 2023

Erstgutachterin: Dr. Hanna, Merk
Zweitgutachter: Dr. Paul, Strohmaier

ERKLÄRUNG ZUR BACHELORARBEIT / MASTERARBEIT

Hiermit erkläre ich, dass ich die Bachelorarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und die aus fremden Quellen direkt oder indirekt übernommenen Gedanken als solche kenntlich gemacht habe. Die Arbeit habe ich bisher keinem anderen Prüfungsamt in gleicher oder vergleichbarer Form vorgelegt. Sie wurde bisher nicht veröffentlicht.

Datum

Unterschrift

Remerciements

Tout d'abord, je souhaite remercier Mme la Professeur Hanna Merk, qui a supervisé mon mémoire de licence. Je tiens à lui exprimer mes sincères remerciements pour son soutien, sa disponibilité et ses suggestions utiles durant la préparation de ce travail.

Je voudrais également remercier David Lefebvre pour des anecdotes intéressantes sur le Québec, des débats et des idées qui ont contribué de manière significative à la réalisation de ce travail. Un grand merci à lui, qui m'a soutenu avec beaucoup de patience, qui a rendu mes études possibles et qui reste toujours à ma disposition.

« You can know the name of a bird in all the languages of the world, but when you're finished, you'll know absolutely nothing whatever about the bird... So let's look at the bird and see what it's doing — that's what counts. »

R. Feynman

Résumé

Au travers d'une analyse de corpus, cette recherche sociolinguistique apporte des éléments de compréhension des tendances de jurons au sein du français québécois. Elle s'appuie sur le Corpus de français parlé au Québec (CFPQ) pour analyser méticuleusement la fréquence de dix jurons sélectionnés, en tenant compte de leurs variations morphologiques et de leur usage euphémique. Les variations dans le choix des jurons sont liées à des facteurs sociaux tels que le sexe et l'âge. Les résultats indiquent qu'un nombre légèrement plus élevé de femmes que d'hommes utilisent des jurons dans le CFPQ, bien que la fréquence d'utilisation soit significativement plus élevée chez les hommes que chez les femmes. Lorsque les femmes jurent, elles utilisent davantage d'euphémismes. Une tendance similaire est également observée chez les personnes âgées. Les expressions grossières atteignent un pic d'utilisation lorsque le locuteur a un âge dans la vingtaine pour diminuer par la suite. De manière intéressante, les individus utilisent plus souvent des gros mots lorsqu'ils conversent avec des personnes du même sexe que lorsqu'ils sont avec des personnes du sexe opposé. Aucune différence significative ne fut observée dans la fréquence des jurons entre les locuteurs de différents niveaux d'éducation. Les résultats suggèrent que les normes sociales et la perception du caractère offensant sont des facteurs sous-jacents qui affectent la façon de jurer à travers les différents groupes sociaux. Cette étude évalue également l'application de la linguistique de corpus dans l'analyse sociolinguistique, en soulignant ses avantages tels que la réfutation des stéréotypes et l'observation des changements linguistiques, mais aussi ses inconvénients tels que le caractère incomplet des données et les distorsions potentielles.

Mots-clés : Gros mots, jurons, sacres, linguistique de corpus, changement de langue, analyse sociolinguistique, français québécois

***Disclaimer : En raison de la nature de ce sujet de recherche, les termes grossiers sont cités à titre d'exemple.

Table des matières

| | | |
|--------|---|----|
| 1. | Introduction | 1 |
| 2. | Cadre théorique..... | 3 |
| 2.1. | Jurer | 3 |
| 2.1.1. | Gros mots, juron, injure..... | 3 |
| 2.1.2. | Sociolinguistique des jurons..... | 4 |
| 2.1.3. | Jurons et euphémisme..... | 6 |
| 2.1.4. | Étude de corpus sur les jurons | 6 |
| 2.2. | Situation linguistique au Québec..... | 8 |
| 2.2.1. | Politique linguistique..... | 8 |
| 2.2.2. | Le bilinguisme – L’alternance des codes | 9 |
| 2.2.3. | Sacres – Identité Québécoise..... | 11 |
| 3. | Méthode..... | 12 |
| 3.1. | Source des données..... | 12 |
| 3.2. | Identification des variables linguistiques | 13 |
| 3.3. | Analyse..... | 14 |
| 4. | Résultats..... | 16 |
| 4.1. | Résultats généraux..... | 16 |
| 4.2. | Sexe | 18 |
| 4.3. | Sexe des groupes de conversation | 19 |
| 4.4. | Âge..... | 19 |
| 4.5. | Niveau d’éducation..... | 20 |
| 4.6. | Euphémisme | 21 |
| 5. | Discussion..... | 23 |
| 5.1. | Niveau d’offensivité | 23 |
| 5.2. | Quand la linguistique de corpus rencontre la sociolinguistique..... | 27 |
| 6. | Conclusion..... | 31 |
| | Annexes | 34 |
| | Référence Bibliographique | 39 |

1. Introduction

Les mots grossiers, également connus sous le nom d'explétifs, de blasphèmes ou de langage vulgaire, sont des mots ou des phrases qui sont considérés comme socialement inappropriés ou offensants. Ils peuvent être utilisés pour exprimer des émotions fortes, telles que la colère ou la frustration, mais également pour choquer ou offenser les autres. Même si ces mots sont généralement tabous dans la plupart des contextes sociaux, ils constituent une caractéristique commune du langage parlé et se retrouvent dans presque toutes les langues et cultures (McEnery 2006, 02). De manière intéressante, le fait de jurer est une preuve évidente de caractéristiques sociolinguistiques puisque ces mots ne sont généralement pas enseignés dans les écoles ou dans des contextes formels d'acquisition d'une langue. Les enfants ont tendance à acquérir des expressions grossières via les interactions avec leurs amis, leurs parents ou les médias, et qu'ils continuent à les utiliser à des fréquences diverses tout au long de leur vie (Suganob-Nicolau 2016, 117). Par conséquent, les jurons sont un phénomène sociolinguistique qui mérite d'être étudié en raison de leur rôle dans la régulation du comportement social.

Durant les dernières années, la linguistique de corpus s'est imposée comme une méthodologie de recherche rigoureuse et polyvalente pour étudier divers aspects de la langue, notamment la grammaire, le discours, les facteurs et processus sociaux, ainsi que l'apprentissage et l'enseignement des langues (Brezina et Gablasova 2018; Ngula 2018). De nombreuses recherches utilisant cette méthode ont été menées pour donner un aperçu des phénomènes et des contextes culturels et historiques dans de nombreuses langues et différents contextes. Les jurons dans les variantes françaises n'ont pas été oubliés dans ce domaine. Plusieurs recherches se sont concentrées sur différentes langues régionales telles que le français acadien (Kasparian et Gérin 2012), le français mauricien (Auckle 2017) ou le *nouchi* – une variété de français d'Afrique de l'Ouest (Derive et Derive 2004; Fodjo 2021). Certaines recherches se concentrent également sur l'effet psychologique de l'insulte utilisée pour rabaisser quelqu'un – et prennent également en compte la gravité de l'insulte et sensibilisé aux différents niveaux d'offense résultant des jurons (Larchet 2017; Lévy 2017; Stapleton et al. 2022).

Au Québec, de toutes les caractéristiques de la langue, les sacres sont certainement les plus connus, voire des marqueurs de l'identité québécoise (Charette 1999; Blanchet 2017). En raison de la place prépondérante qu'ils occupent dans la culture, les sacres sont l'objet de nombreuses recherches scientifiques dans différents domaines de la linguistique, en

particulier dans celui de la pragmatique. Après une étude comparative entre les jurons en français hexagonal et en français québécois, Drescher constate que les sacres québécois connaissent une pragmatisation progressive pour devenir des marqueurs discursifs (Drescher 2007, 183). De même, Bässler souligne que les sacres apparaissent rarement de manière isolée, mais en relation avec des signaux non verbaux et sont utilisés pour intensifier les énoncés (Bässler 2009, 171). Alors que de nombreuses études contribuèrent à élucider la pertinence des sacres en tant que marqueurs de l'identité québécoise, les jurons non liés aux sacres dans le français du Québec reçurent comparativement moins d'attention. Compte tenu de la nature bilingue du Canada et du contact linguistique inévitable, il est raisonnable de s'interroger sur l'influence potentielle de l'anglais et d'autres langues sur le répertoire des grossièretés en français québécois.

Cette étude présente une analyse de corpus basée sur 10 jurons en comparant les fréquences et en analysant les divergences sociolinguistiques dans le *Corpus de français parlé au Québec* (CFPQ) développé par l'Université de Sherbrooke au Canada. L'objectif de cette investigation n'est pas de fournir un guide prescriptif sur la façon de jurer comme un Québécois, mais d'examiner la variation dans le choix des jurons sur la base de différents facteurs sociaux, et de chercher à déterminer les éléments sous-jacents qui affectent le comportement lié au fait de jurer. Cette étude vise à évaluer les avantages et les limites de l'utilisation de la méthode du corpus pour comprendre la nature des jurons dans le contexte québécois, ainsi que son application dans l'analyse sociolinguistique. Pour réaliser les objectifs, plusieurs questions de recherche (QR) sont proposées :

- QR1 : Quels sont les jurons les plus courants en français québécois ?
- QR2 : Dans quels schémas les jurons au Québec se retrouvent-ils le plus souvent ?
- QR3 : Dans quelle mesure les jurons révèlent-ils les facteurs sociaux et culturels au Québec qui façonnent l'utilisation des insultes ?

Cette analyse vise à examiner les modèles sociolinguistiques qui sous-tendent l'utilisation des jurons au Québec. Le reste de cette étude est organisé comme suit : la deuxième partie présente le contexte théorique de cette étude, en donnant un aperçu des études antérieures sur le langage grossier, la situation linguistique au Québec et la méthode de la linguistique de corpus. La troisième partie décrit la méthodologie utilisée, y compris la sélection et l'analyse des données. Il y est expliqué les critères de sélection de l'ensemble des données et la manière dont les données furent organisées. La section des résultats sera présentée dans la quatrième partie, avec un aperçu général, suivi d'une exploration des modèles

sociolinguistiques basés sur des facteurs tels que l'âge, le sexe, et l'éducation. Les résultats sont ensuite discutés et la question de recherche sera traitée dans la cinquième partie. Cette partie se termine par des remarques finales et des réflexions sur de possibles pistes d'approfondissement de ce sujet. La dernière section synthétise les résultats de cette étude, en répondant à la question de recherche et récapitule les implications des résultats.

2. Cadre théorique

2.1. Jurer

2.1.1. Gros mots, juron, injure...

La définition des *gros mots* représente un grand défi en raison de l'ambiguïté substantielle et du chevauchement des usages qui caractérisent la clarification conceptuelle de la terminologie des jurons. Cette ambiguïté inhérente conduit souvent à une classification interchangeable de termes tels que *gros mot*, *injure*, *insulte*, *juron*, *sacre* et autres. La définition de *gros mots* donnée par les dictionnaires semble moins restrictive que celle de mot grossier. Dans un sens plus large, elle est synonyme de *grossièreté*, qui fait référence à des attitudes, des comportements ou des commentaires inappropriés ou insensibles (« Dico en ligne Le Robert », s. d.; « Dictionnaire Usito », s. d.). Cela correspond à la définition de Rouayrenc (1998, 05) des *gros mots* comme des mots interdits par la société. Les dictionnaires de français de l'Hexagone et de français du Québec, tels que Le Robert et Usito, semblent établir une distinction entre *jurons* et *insultes*, en fonction de l'intention derrière l'action – qu'il s'agisse d'exprimer une forte émotion ou d'offenser quelqu'un (« Dico en ligne Le Robert », s. d.; « Dictionnaire Usito », s. d.).

Au sens large, les *gros mots* sont fréquemment considérés comme un terme générique qui englobe les *insultes* et les *jurons*. Une telle supposition est probablement trop simpliste. De nombreuses *injures* ne font pas partie de la catégorie des *gros mots* (par exemple, *blaireau*) et que certains *gros mots*, comme *branlette*, n'appartiennent ni aux *jurons* ni aux *injures* (Rouayrenc 1998, 05). Une définition basée sur le récepteur est présentée par Perea, qui établit une distinction entre les *jurons*, les *insultes personnelles* adressées à quelqu'un et les *insultes rituelles* utilisées dans le cadre de l'interaction sociale. Au niveau sociolinguistique, le *juron* est un signe de proximité émotionnelle et d'identité qui dépasse les règles de l'interaction sociale, souvent même en opposition directe avec elles (Perea 2011, 53-54). Huston propose une classification basée sur le diagramme du processus de communication de Bühler, qui différencie les *gros mots* selon leur fonction principale dans le langage : expressive, impressionnante et référentielle. Selon ce modèle, les *jurons* sont associés à la

fonction expressive, tandis que les *injures* sont associées à la fonction impressive. Autrement dit, les *jurons* sont produits pour exprimer des sentiments ou des émotions, tandis que les *injures* sont dirigées vers le destinataire (Huston 1980, 16-17).

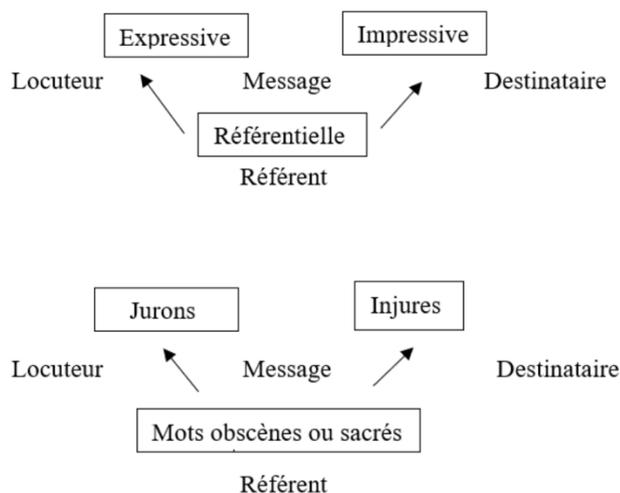


Figure 1 Diagrammes du processus de communication de Bühler et du diagramme du juron proposé par (Huston 1980, 16-17)

Après avoir établi les définitions, il est important de faire la distinction entre les *jurons* et les *insultes* en ce qui concerne la nature émotionnelle de l'énoncé. Les *jurons* diffèrent de l'*insulte* en ce sens qu'ils ne sont pas dirigés vers autrui et qu'ils ressemblent davantage à une interjection destinée à soutenir l'expression du locuteur dans l'euphorie, la jubilation, la colère ou la frustration, alors que les insultes et les injures visent à porter atteinte à la réputation des auditeurs. Cette étude ne se penche pas sur les *insultes* et les *injures*. De tels cas sortent du cadre de notre recherche.

2.1.2. Sociolinguistique des jurons

La façon dont un individu parle peut donner un aperçu de son identité sociale et peut influencer la perception qu'en ont les autres. Certaines études suggèrent que des facteurs sociaux tels que l'âge et le sexe jouent un rôle important dans le choix des jurons et des expressions grossières. Une étude sur les jurons en anglais a montré que les hommes ont tendance à avoir un vocabulaire plus riche de jurons, ainsi que les jurons associés à l'âge et au statut socio-économique (McEnery 2006, 30). Des résultats similaires sont également observés dans une étude récente sur le français informel sur une page Reddit (Flesch 2022). La norme sociale explique ce phénomène : les jurons des femmes sont souvent jugés plus sévèrement que ceux des hommes (Yaguello 1987, 44). Cependant, le niveau socio-économique et le niveau d'éducation semblent être des facteurs controversés. McEnery &

Xiao ont constaté une fréquence significative de l'utilisation de *fuck* chez les locuteurs anglophones ayant quitté l'école à 15 ans (2004, 246). En revanche, Stapleton affirme que la relation entre les jurons et le faible niveau d'éducation est simplement un stéréotype supposé (Stapleton 2010, 291).

Le choix des jurons peut varier considérablement en fonction de divers facteurs, notamment le contexte social dans lequel ils sont utilisés, la relation entre les locuteurs et les normes culturelles et linguistiques de la communauté. Selon Aitchison, les gens sont susceptibles de cesser d'utiliser l'argot et les jurons en vieillissant, car ils apprennent à adapter leur langage à différentes situations. Néanmoins, certaines personnes âgées ont tendance à utiliser des jurons pour paraître amicales (Aitchison 2006, 23). De même, leur emploi de jurons peut changer lorsque les locuteurs fréquentent des personnes qui ont des enfants en bas âge (Holmes 1992 dans McEnery et Xiao 2004, 241). Il a également été observé que ceux qui apprenaient le français en classe étaient moins susceptibles de jurer en français que les apprenants qui combinaient l'enseignement en classe avec une utilisation authentique en dehors de la classe ou qui apprenaient le français de manière naturaliste. Une forte socialisation en français prédit également une fréquence plus élevée de jurons en français que dans leur langue maternelle (Dewaele 2011, 111).

Des recherches antérieures sur le langage grossier permettent d'obtenir des informations sur sa nature offensante et son impact émotionnel. Les gens ont des perceptions différentes du niveau d'offense des jurons, et cela s'applique même à des mots qui sont généralement considérés comme moins offensants. Les femmes, par exemple, qui ne sont généralement pas associées aux jurons, sont plus sensibles à la nature offensante du langage (Beers Fägersten 2007, 32). Le degré d'offense varie selon que le locuteur est natif ou non natif. Les locuteurs non natifs ont tendance à surestimer le caractère offensant de la plupart des expressions grossières. Cela s'explique par les mises en garde contre le langage offensant dans les établissements d'enseignement (Dewaele 2016, 22). Malgré l'association courante des jurons avec l'agression verbale et leur caractère tabou, cela donne lieu à un phénomène baptisé *paradoxe des jurons*. Proposé par Beers-Fägersten (2012) dans son étude, ce terme souligne le fait que les jurons, bien qu'ils soient socialement inappropriés, sont largement utilisés dans le parler de tous les jours. Une fréquence notable de jurons a été constatée dans son analyse menée sur des locuteurs natifs américains, les jurons les plus utilisés n'étant pas nécessairement ceux considérés comme les plus offensants.

2.1.3. Jurons et euphémisme

Les expressions euphémistiques sont souvent employées de manière tactique pour remplacer les mots offensants par d'autres formes qui véhiculent des idées similaires afin d'éviter un langage inapproprié. Elles sont conçues pour être indirectes et imprécises afin d'aborder des sujets sensibles, tabous, embarrassants ou désagréables, tout en conservant un ton socialement acceptable (Hughes 2006 cité dans Jing-Schmidt 2021, 125). Le tabou et l'euphémisme, en tant que dichotomie, coexistent et collaborent harmonieusement. Le tabou établit ce qui est interdit, tandis que l'euphémisme implique l'utilisation d'expressions alternatives pour remplacer les formes taboues du langage (Widlak 1965, 934). Contrairement aux mots tabous, les euphémismes sont présents à différents niveaux et registres linguistiques, véhiculant un sens relativement neutre par rapport aux expressions taboues (Jamet 2010, 31; Guiraud 1975, 23).

Dans le but d'exprimer des émotions fortes tout en évitant la confrontation directe, la formation d'euphémismes consiste à remplacer ou à modifier des mots pour en atténuer l'impact tout en conservant le sens sous-jacent. Par exemple, les sacres au Québec tels que *tabarnaque* deviennent *tabarouette*, *câlize* devient *câlino*, *ciboire* devient *cibole* (Drescher 2007, 177). En effet, il s'agit souvent de mots qui sont réinterprétés pour s'adapter à une situation qui ne semble pas pertinente. En prononçant *punaise*, par exemple, le locuteur ne parle pas d'un insecte, mais d'une alternative à *putain*. Pour créer un euphémisme, les locuteurs choisissent souvent des mots qui présentent des similitudes phonétiques ou sémantiques avec les mots cibles, par exemple, *mince*, *merdre* ou *mercredi* au lieu de *merde* (Hirsch 1985, 55). Des études indiquent que les individus subissent des niveaux de stress plus élevés lorsqu'ils prononcent vocalement des mots obscènes que lorsqu'ils emploient des euphémismes, ce qui met en évidence l'influence psychologique du choix de mots spécifiques (Bowers et Pleydell-Pearce 2011). Les résultats de la recherche sont sans ambiguïté et ne sont pas inattendus, puisque les gens ont tendance à éprouver un plus grand stress lorsqu'ils prononcent un gros mot que lorsqu'ils emploient un euphémisme.

2.1.4. Étude de corpus sur les jurons

Des études montrent que la linguistique de corpus permet de révéler le rôle des jurons dans la société en examinant la manière dont ils sont utilisés et qui les utilise. Une étude sociolinguistique sur l'homophobie a été menée par Larchet (2017), qui a identifié trois catégories de propos homophobes : ceux ciblant les hommes gays en les féminisant, les propos genrés et sexualisés non explicitement homophobes, et les propos personnalisés

ciblant les lesbiennes. En outre, des études comparatives sur les jurons des adolescents ont également montré que certains thèmes de jurons, tels que les insultes à la mère, sont beaucoup plus fréquents dans les pays catholiques comme l'Espagne qu'en Norvège (Drange, Hasund, et Stenström 2014). Les corpus français sont également utilisés dans le traitement du langage naturel à diverses fins, comme la construction de modèles pour détecter le langage abusif et les discours de haine en ligne (Battistelli, Bruneau, et Dragos 2020). Il est certain que la linguistique de corpus contribue à de nombreuses interdisciplines linguistiques différentes telles que la sociolinguistique et la linguistique informatique.

La méthode linguistique du corpus fait néanmoins l'objet de critiques importantes, notamment de la part de Noam Chomsky, qui a identifié certains aspects problématiques inhérents à cette approche. Premièrement, l'analyse rétrospective des données de corpus est insuffisante pour la linguistique théorique. Selon Chomsky, c'est sur ce langage interne que le linguiste doit se concentrer en priorité. En revanche, le corpus linguistique met l'accent sur les aspects observables du langage externe, qui reflètent insuffisamment le langage interne (cf. la conversation de Chomsky cité dans Aarts 2001, 05-07). Les arguments de Chomsky sont bien ancrés dans le débat entre le rationalisme et l'empirisme. Alors que les théories rationalistes visent la plausibilité cognitive en représentant la manière dont le traitement du langage se produit, les approches empiristes s'appuient sur l'observation de données naturelles, dans ce cas le corpus (McEnery et Wilson 2005, 05). Le langage est intrinsèquement indénombrable, alors qu'un corpus, quelle que soit sa taille, reste une collection finie de langues, ce qui entraîne des biais potentiellement considérables dans les résultats (Ibid., 10).

Cette incomplétude est due aux multiples variations entre les différents contextes d'une langue et à l'impossibilité de collecter tous les exemples de phénomènes linguistiques (Stefanowitsch 2009, 09). Ceci n'est pas propre à la linguistique de corpus, mais s'applique également aux statistiques, où l'absence de données dans les résultats n'implique pas leur inexistence. Par conséquent, il est suggéré aux linguistes d'accepter et de reconnaître l'incomplétude des corpus et de construire des modèles généraux basés sur des données incomplètes et de les affiner itérativement avec l'inclusion de données supplémentaires (Stefanowitsch 2009, 08). Bien que l'argument de Chomsky reste valable¹, il est essentiel de

¹ McEnery et Wilson, qui soutiennent l'utilisation de l'approche linguistique de corpus, sont d'accord avec certains arguments de Chomsky, ceux que j'ai mentionnés ci-dessus (McEnery et Wilson 2005, 10). Chomsky a également mentionné les erreurs potentielles dans l'analyse de corpus, mais cet argument concernait avant tout les limites de la technologie dans les années 1960 et n'est donc plus valable aujourd'hui.

prendre en compte la contribution de Lakoff et Bucholtz, qui souligne que la linguistique de corpus ne propose que des énoncés ou des séquences de symboles, nécessitant une interprétation pour leur attribuer un sens. Les données du corpus elles-mêmes ne possèdent pas de sens intrinsèque, ce qui montre le rôle essentiel de l'introspection dans l'analyse cognitive du langage (Lakoff et Bucholtz 2004).

La critique de Chomsky sur l'incomplétude des données de corpus a considérablement influencé la linguistique de corpus moderne, incitant à améliorer la représentativité et l'équilibre dans la construction des corpus. Le corpus idéal devrait être constitué d'échantillons linguistiques reflétant fidèlement la langue étudiée et garantissant une représentation équilibrée grâce à une sélection méticuleuse (McEnery, Xiao, et Tono 2006, 05). La collecte de données n'est pas entièrement aléatoire, mais elle doit éviter les choix arbitraires. Ainsi, la représentativité implique de capturer l'utilisation quotidienne de la langue dans divers contextes à une échelle plus petite qui reflète l'échelle plus grande (Brezina et Gablasova 2018, 01). La position de l'approche linguistique de corpus fut renforcée depuis 1980 et fournit des analyses plus rapides et plus objectives des caractéristiques linguistiques que les méthodes basées sur l'introspection, car elles offrent également des informations fiables sur la fréquence et permettent aux chercheurs de vérifier et de reproduire les études (Ngula 2018, 207).

2.2. Situation linguistique au Québec

2.2.1. Politique linguistique

Le Québec, province canadienne dont les caractéristiques linguistiques sont distinctes de celles du français hexagonal, connut un développement linguistique unique. Alors que le Canada est un pays officiellement bilingue, l'anglais et le français étant tous deux reconnus comme langues officielles, la majorité de la population (80 %) parle l'anglais comme première langue dans tout le pays. À l'inverse, au Québec, 80 % des personnes parlent le français comme première langue (Houle 2018, 04). L'utilisation du français au Québec était historiquement limitée aux Canadiens français de souche jusqu'à l'adoption de politiques linguistiques à la fin des années 1960 (Oakes 2004, 539; Piriou 1992, 27). La loi 63, adoptée en 1969, permettait aux parents de choisir la langue d'enseignement applicable à leurs enfants, puis la loi 22 reconnaissait officiellement le français comme langue du Québec et mettait en œuvre des mesures visant à promouvoir son utilisation tout en restreignant l'accès à l'école anglaise (Gosseline s. d.). La Charte de la langue française de 1977, qui définit le français comme la langue officielle du gouvernement provincial, est l'une des réalisations

qui permirent de faire du français la langue officielle (Piriou 1992, 27-28). À partir de cette date, le français occupe une position dominante au Québec, servant de langue principale pour l'administration, la législation, la justice, l'éducation, le lieu de travail et les activités commerciales (Gosseline s. d.)

Malgré les efforts de protection et de promotion de la langue française, l'équilibre linguistique entre le français et l'anglais au Québec reste fragile. L'anglais occupe une position dominante non seulement au Canada mais aussi en Amérique du Nord en raison de la mondialisation et des exigences du marché (Ibid.). De plus, le faible taux de natalité chez les Québécois francophones et le nombre croissant d'immigrants non francophones posent des problèmes au gouvernement pour maintenir la prééminence de la langue française (Oakes 2004, 547).

2.2.2. Le bilinguisme – L'alternance des codes

La province de Québec témoigne d'une augmentation notable du bilinguisme entre les langues française et anglaise au cours des dernières années. Selon les données de l'Office québécois de la langue française, entre 1971 et 2016, une augmentation significative du bilinguisme a été observée, avec une croissance du pourcentage de 1,5 fois, englobant près de la moitié de la population québécoise en tant que bilingue (45 %). En 2016, la proportion de Canadiens anglophones du Québec qui affirment maîtriser le français et l'anglais est impressionnante (69 %). Notamment, parmi ceux avec une langue maternelle différente, 51 % avaient également adopté le français comme langue secondaire (*Rapport sur l'évolution de la situation linguistique au Québec* 2019). En outre, dans la ville de Montréal, environ les deux tiers des résidents canadiens anglophones maîtrisent l'anglais et le français (Boberg 2012, 496). Toutefois, il ne s'agit pas d'un phénomène national, c'est-à-dire que cela ne se produit pas au sein du Canada anglophone, mais est observé avec un changement radical uniquement au Québec, comme le montre la figure 2 ci-dessous (Corbeil et Houle 2016).

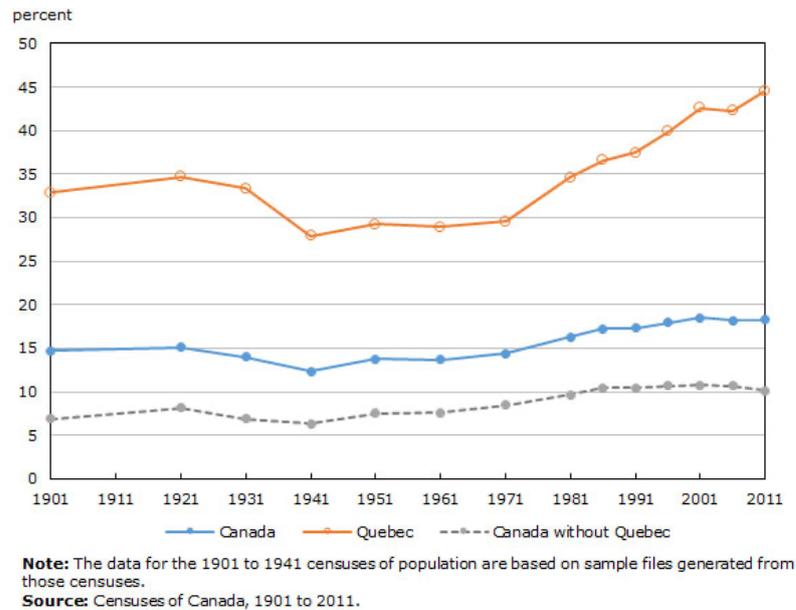


Figure 2 Répartition de la population au Canada, au Québec et au Canada sans le Québec, de 1901 à 2011, maîtrisant l'anglais et le français

Dans un contexte bilingue comme celui du Canada, l'alternance codique (ou code-switching en anglais) est un phénomène inévitable. Elle désigne la pratique consistant à alterner entre deux ou plusieurs langues au cours d'une même conversation ou dans un même contexte (Deliana et Rohani 2019, 202). Ceci a été d'ailleurs observé dans différentes minorités françaises, comme dans le cas du français acadien. De nombreux emprunts directs et dynamiques sous différentes formes simples comme des verbes, des noms, des adjectifs, des interjections et des phraséologismes furent trouvés dans le français acadien, voir dans les exemples suivants :

- (1) J'étais comme **fucking Christ** I anyways so là c'est ça j'i
- (2) Ça me **pisse** assez [= tellement] **off**/

(Kasparian et Gérin 2012, 130, 133)

La fréquence et l'intensité de l'alternance codique, ainsi que l'utilisation de l'anglais par les personnes bilingues au Québec sont influencées par des facteurs tels que la taille des populations anglophones et francophones. Sur l'île de Montréal et dans les villes autour de l'Ontario et de Toronto, le nombre de l'alternance codique en anglais dans les tweets français s'est avéré beaucoup plus élevé que dans le reste du Québec, ce qui est en corrélation positive avec la population anglophone (Kellert 2023, 06, 10-11). L'étude comparative de Carpuat sur ce phénomène d'alternance codique entre les Canadiens francophones et anglophones suggère que les Canadiens francophones font preuve d'un niveau élevé de maîtrise de l'anglais, d'après leur utilisation de l'alternance codique impliquant des citations, des

expressions idiomatiques et des expressions à plusieurs mots (2014, 112). En outre, Dewaele souligne que les personnes multilingues ont tendance à préférer jurer dans leur langue dominante, bien qu'elles puissent stratégiquement jurer dans une langue plus faible pour en atténuer l'impact ou contourner les conventions sociales (2004, 102).

2.2.3. Sacres – Identité Québécoise

Parmi les différentes caractéristiques du français québécois, les sacres se distinguent du français hexagonal². En français québécois, les sacres désignent des jurons religieux associés à des objets liturgiques tels que le calvaire, le tabernacle, l'hostie ou le calice (Bässler 2009, 160). Les origines des sacres sont ancrées dans l'histoire du Québec, avec une forte influence catholique sur les individus, suivie d'une période de transformation qui fit basculer le Québec vers la sécularisation dans les années 1960 (Bougaïeff 1980, 840). À cette époque, les sacres sont passés d'une transgression religieuse à une transgression plus sociale. La présence du sacre sous forme de jurons blasphématoires n'est pas un hasard dans une telle société. Il est observé que les *sacres* transmettent une plus grande intensité d'émotion chez les Québécois, quel que soit leur statut social, par rapport à ce qu'une expression en français standard pourrait atteindre (Ibid.). Donc, les sacres servent de véritables marqueurs de l'identité québécoise, facilitant la reconnaissance et la connexion entre les Québécois francophones.

De nombreuses recherches linguistiques mettent en lumière la remarquable complexité de la construction grammaticale et de la fonction des sacres. Il a été constaté que la notion de péché associée aux sacres avait presque disparu depuis les années 80 parmi la jeune génération (Vincent 1984, 59). Par conséquent, les sacres ne sont plus employés principalement comme des actes de blasphème ou de malédiction à l'égard du divin, mais servent plutôt de forme de serment (Bougaïeff 1980, 840). Outre leur utilisation courante en tant qu'interjection, les sacres peuvent s'intégrer dans la syntaxe de la langue parlée sous quatre catégories grammaticales : nominale, adjectivale, adverbiale et verbale (Bässler 2009, 160). Blanchet identifie trois caractéristiques distinctes qui font la différence entre les sacres québécois et le français hexagonal : la synonymie, la récursivité et la substitution (2017, 12-14). Leur sémantique conceptuelle étant perdue, cela leur permet d'être interchangeables sans altérer leur sens, voir l'exemple (3). Leurs propriétés récursives leur permettent d'être

² Les jurons à caractère religieux tels que *mon dieu*, *parbleu* et *bon sang* existent certainement en français de l'Hexagone. L'usage des jurons religieux serait relativement moins intense en France, et les jurons relèveraient principalement du domaine de la sexualité et de la scatologie (Bässler 2009, 160).

composés à l'infini à la suite sans ordre fixe, ce qui est impossible dans les jurons du français hexagonal (4a,b,c,d). De plus, les sacres peuvent être remplacés par d'autres formes de jurons sans en changer le sens (5a,b).

(3) Crisse de calisse de tabarnak

(4a) Crisse de tabarnak

(4b) Tabarnak de crisse

(5a) Ostie de crisse

(4c) Bordel de merde

(4d) *Merde de bordel

(5b) Ostie de tabarnak

3. Méthode

3.1. Source des données

Les données utilisées dans cette étude proviennent du *Corpus de français parlé au Québec* (CFPQ) – un corpus parlé collecté dans le cadre de la recherche à l'Université de Sherbrooke. Le CFPQ est composé de 31 sous-corpus comprenant des conversations occasionnelles entre amis proches ou membres de la famille. Chaque sous-corpus individuel a une durée d'environ 1 heure et 30 minutes. Au total, environ 46 heures d'enregistrement furent réalisées entre 2006 et 2019. Les transcriptions de ces sous-corpus comprennent 712,300 tokens et 28,638 types. Les conventions de transcription décrites sur le site web du corpus suivent des normes largement adoptées pour la représentation de la langue parlée, ce qui facilite la recherche par mots-clés (Dostie 2015, 03).

Le CFPQ est constitué de riches données orales et informelles du français québécois non influencées par des objectifs de recherche. Les conversations sont enregistrées dans des lieux privés, où les locuteurs tiennent souvent des conversations sans contrainte de sujet spécifique. Ainsi, cette approche permet de créer une atmosphère confortable entre les participants pour mener la conversation, mais aussi de minimiser le paradoxe de l'observateur, dans lequel l'observateur peut modifier l'utilisation de la langue des participants et, dans ce cas, l'utilisation de jurons (Labov 1972, 209). La sélection du corpus du CFPQ présente une opportunité prometteuse pour mieux comprendre le phénomène des jurons au Québec.

Comme indiqué dans le Tableau 1, les données suggèrent une distribution relativement équilibrée des sexes au sein de l'échantillon, sans presque aucun biais significatif en faveur de l'un ou l'autre sexe, avec 51 hommes (46 %) et 61 femmes (54 %). Les participants se répartissent de manière équilibrée entre plusieurs groupes d'âge, allant de moins de 20 ans à plus de 70 ans. Le niveau d'éducation de la population suggère une proportion relativement faible de personnes ayant un niveau d'éducation primaire, tandis que le niveau d'éducation supérieur représente une proportion substantielle de l'échantillon avec 40 %, suivi par le

niveau d'éducation intermédiaire. Ceux qui suivirent le Cégep, qui est spécifique au système éducatif québécois, représentent 25 % des participants, ce qui suggère une représentation notable des personnes ayant suivi ce type d'enseignement.

Tableau 1 Distribution des participants en fonction de l'âge, du sexe et du niveau d'éducation

| Variables | | Nombre | Pourcentage | Totale |
|------------------------|-----------------|---------------|--------------------|---------------|
| Sexe | Hommes | 51 | 46 % | 112 |
| | Femmes | 61 | 54 % | |
| Age | Moins de 20 ans | 11 | 10 % | 112 |
| | 20 – 29 ans | 27 | 24 % | |
| | 30 – 39 ans | 18 | 16 % | |
| | 40 – 49 ans | 16 | 14 % | |
| | 50 – 59 ans | 17 | 15 % | |
| | 60 – 69 ans | 11 | 10 % | |
| | Plus de 70 ans | 12 | 11 % | |
| Niveau d'études | Primaire | 7 | 6 % | 112 |
| | Secondaire | 33 | 29 % | |
| | Cégep | 28 | 25 % | |
| | Universitaire | 44 | 40 % | |

Lorsqu'il s'agit de corpus parlés, le défi de l'échantillonnage doit être pris en compte. Malgré les efforts déployés pour maintenir une distribution relativement équilibrée des participants, une taille d'échantillon inadéquate peut limiter la capacité à saisir toute la gamme des variations ou à fournir des estimations fiables des paramètres de la population. Malheureusement, ce problème reste inévitable dans le cas des corpus oraux, vu que leur taille est généralement beaucoup plus petite en volume que celle des corpus écrits (Gadet 2021, 74). Le traitement des corpus oraux et multimodaux implique de multiples étapes, telles que la conception du protocole, la collecte des données, la transcription, les révisions et l'anonymisation, qui nécessitent un temps et des ressources considérables (Ibid.). Il est donc crucial de reconnaître que l'absence de certaines caractéristiques linguistiques dans le corpus n'implique pas nécessairement leur absence dans l'usage quotidien de la langue.

3.2. Identification des variables linguistiques

Un ensemble de dix jurons et leurs variantes ont été utilisés pour analyser l'utilisation des jurons au Québec :

CÂLICE, OSTIE, TABARNAQUE, CRISSE, CIBOIRE, MARDE, FUCK,
SHIT, BULLSHIT, BITCH

La sélection de ces termes s’est inspirée de plusieurs critères, fondés sur des recherches antérieures (cf. Vincent 1984, 215; Charette 1999, 84), qui mettaient l’accent sur les sacres reconnus par les locuteurs natifs et sur la fréquence de leur utilisation. D’autres types de jurons ont été sélectionnés selon la classification intuitive de Charette (1999, 84), voir les annexes A et B. Les termes empruntés à l’anglais sont basés sur une étude des jurons les plus courants en Amérique du Nord (cf. Ten et Simatupang 2020). Les termes empruntés à l’anglais furent ajoutés afin d’examiner l’alternance codique effectué par les Québécois à la suite de l’augmentation du nombre de personnes bilingues français – anglais. Compte tenu de nos intérêts spécifiques, les recherches comprennent également des variations de formes euphémiques de jurons basées sur l’ensemble de dix jurons. Les mots-clés des jurons euphémiques furent choisis parmi les répertoires établis dans des études antérieures (cf. Charest 1974, 43; Thibault & Vincent 1981 cité dans Charette 1999).

Les requêtes de recherche pour les jurons du corpus furent générées en utilisant des symboles joker en conjonction avec des outils optionnels pour prendre en compte les formes variables des mots. Le tableau 2 représente le répertoire complet des requêtes de recherche pour les jurons, leurs variantes morphologiques et leurs formes euphémiques.

Tableau 2 Requêtes de recherche

| Les jurons | Mots-clé |
|-------------------|---------------------------|
| Ostie | *OSTI*, *ESTI*, *ASTI* |
| Crisse | *CRISS*, *CRIST*, *CRIF* |
| Tabarnaque | *TABAR* |
| Ciboire | *CIBO* |
| Merde/Marde | *MARD*, *MERD* |
| Fuck | *FUCK*, *FRIDGE*, *FRICK* |
| Shit | *SHIT*, *SHOOT*, *SNAP* |
| Damn | *DAMN*, DANG |
| Asshole | *ASSHOLE* |

3.3. Analyse

Avant d’importer les sous-corpus dans le logiciel d’analyse de corpus ANTCONC, les conventions seront supprimées afin d’obtenir des résultats plus transparents et plus précis. Le corpus fut conçu à l’origine pour capturer un large éventail de caractéristiques linguistiques, y compris des éléments pragmatiques tels que les gestes, les émotions, les mouvements et les variations d’intonation. Dans le cadre de cette étude, la diversité des caractéristiques présente toutefois plus d’inconvénients que d’avantages, comme l’illustrent les exemples suivants. Afin d’éliminer les collocations non pertinentes et d’obtenir un

résultat plus précis, la convention sera d'abord supprimée à l'aide de Python3, comme la montre l'annexe C.

| File | Left Context | Hit | Right Context |
|--------------|--|-----------------------------------|---|
| 21_sous-... | S : ben ouais <P12,L16> O : pis il est agronome [1 | ostie (dit en riant) <P12,L17> | Y : [1eh crise ok (|
| 21_sous-... | personne dont il parle <P24,L9> O : elle est [1danseuse | ostie (dit en riant) <P24,L10> | [1ma voisine je l' |
| 21_sous-... | L17> [2<f<tu veux jouer au Y : les mêmes chances [5 | ostie (dit en riant) <P39,L18> | [3ouin ben là les |
| 21_sous-... | O : <P42,L8> S : [3•ayoye [4tabarnaque* <P42,L9> Y : longtemps | ostie>> (dit en riant) <P42,L10> | S : (inaud.) <P42,L11> [4< |
| 26_sous_c... | vite parce que ça sent fort* (0:02:13.0) <P154,L13> MÉ : [6(RIRE) [7 | ostie (dit en riant) <P154,L8> | [1ouin <P154,L14> [5ok < |
| 27_sous_c... | je disai:s (dit avec moquerie) (0:02:02.0) <P19,L17> G : (RIRE) [1 | ostie (dit en riant) <P19,L18> | A : M : [1(RIRE) <P19, |
| 30_sous_c... | à <P7,L21> SOUS-CORPUS 30 : segment 1. (Durée 10 minutes) MA : [1 | ostie:: (dit en riant) <P8,L1> | 8 V : V : [2(elle hoche |
| 6_sous_co... | oui mais là c'est pas de la miSÈRE [2ça | Ostie (dit en riant) <P115,L22> | [1ok <P115,L20> [1(inaud.) < |
| File | Left Context | Hit | Right Context |
| 25_sous_c... | épargnée cette fois -là en tout cas j'ai un | ostie | de speech de marde que je lui avais donné |
| 25_sous_c... | M : TAbarouette <P92,L9> C : mais j'ai fait un | ostie | de speech là j'ai dit écoute je suis |
| 21_sous-... | regarde ostie <P15,L8> O : ah non <P15,L9> Y : | ostie | de ah non <P15,L10> O : pis euh il |
| 30_sous_c... | on peut sacrer à la caméra ostie de con euh : | ostie | de bande de CONS bande de d'épais vous |
| 22_sous_c... | là <P142,L11> M : <P142,L12> MC on a des | ostie | de bonne vie <P142,L13> J : oui <P142,L14> |
| 27_sous_c... | là je te jure que tu euh tu fais un | ostie | de bout de chemin <P151,L24> G : oui <P151, |
| 30_sous_c... | P46,L17> M : si je mangeais juste pas comme un | ostie | de cochon là je serais gros de même t' |
| 30_sous_c... | non NON mais euh on peut sacrer à la caméra | ostie | de con euh : ostie de bande de CONS bande |

Figure 3 Illustration des concordances avant et après la suppression des conventions

Une fois que les requêtes de recherche appropriées furent définies et que la convention des 31 sous-corpus fut retirée, les questions de recherche purent être abordées en suivant la procédure suivante :

Question de recherche 1 : Une recherche dans CFPQ fut effectuée pour chacun des jurons à tour de rôle. Toutes les occurrences contenant les mots cibles et répondant aux exigences furent collectées pour former un ensemble de données à analyser. Un tri manuel fut appliqué pour filtrer les entrées non pertinentes telles que les énoncés inachevés ou les mots ayant une signification autre que le juron. Les euphémismes ont fait l'objet d'une attention particulière dans leur contexte, vu que le même mot dans son sens original peut être confondu avec un euphémisme. Par exemple, *câlino* qui peut être un sacre mais qui décrit aussi une personne qui aime les câlins, *mince* qui est souvent prononcé comme un lapsus mais qui a aussi une signification comme *maigre*. Les fréquences ont ensuite été collectées et normalisées pour la comparaison.

Question de recherche 2 : Une ségrégation a été effectuée entre les formes flexionnelles et dérivationnelles. La catégorisation ultérieure a consisté à classer les variantes flexionnelles dans des classes distinctes – à savoir les verbes, les noms et les adjectifs – en fonction des suffixes distinctifs qu'elles présentent.

Question de recherche 3 : La distribution sociale derrière les jurons fut étudiée en fonction du sexe du locuteur, de son âge, de son niveau d'éducation et des différents sexes composant les groupes. Pour ce faire, une liste des énoncés dans lesquels apparaissent les jurons fut générée, puis les locuteurs et leur milieu social furent identifiés. Cette analyse

sociolinguistique s'est également concentrée sur les jurons, les euphémismes et les jurons empruntés à l'anglais. Malheureusement, le nombre de tokens pour chaque sous-corpus n'étant pas disponible, la fréquence relative de l'analyse du genre du groupe a été estimée et normalisée. De plus amples détails sont donnés dans la section de résultat et à l'annexe D.

4. Résultats

4.1. Résultats généraux

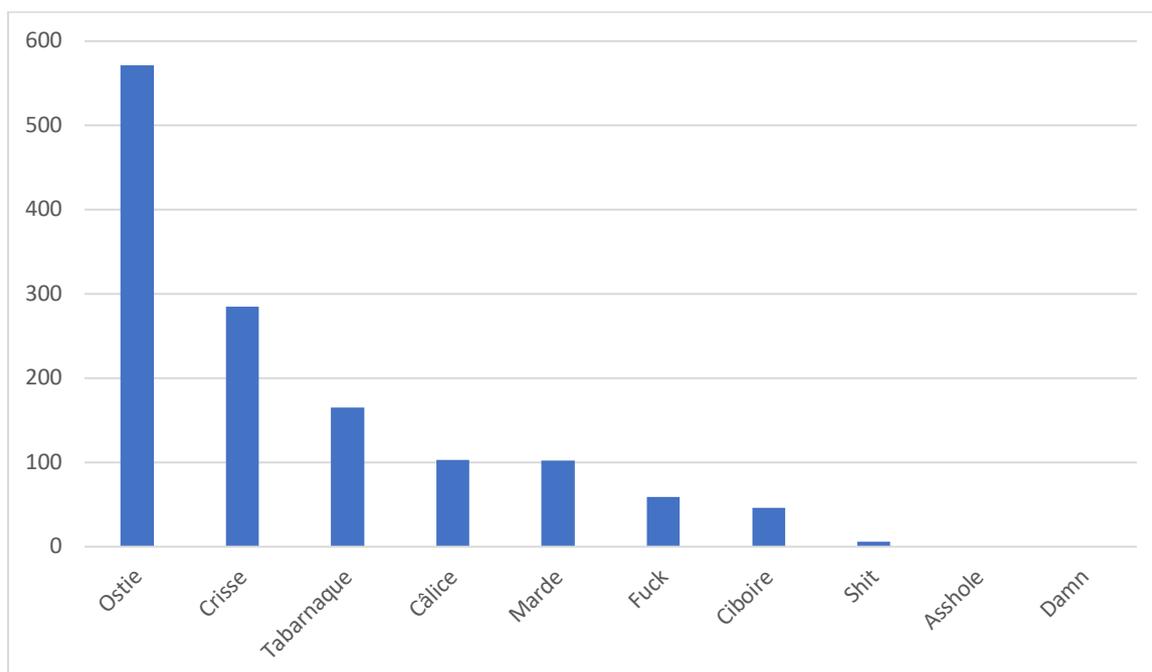


Figure 4 Fréquence des jurons dans le CFPQ

Le résultat montre que 1337 occurrences furent trouvées, dans lesquelles les dix jurons furent utilisés. La figure 4 illustre la fréquence de mots de ces jurons dans le corpus du CFPQ. Il en ressort que les sacres sont majoritairement utilisés comme jurons en français québécois, OSTIE étant le juron le plus fréquent. Les jurons semblent également être centrés autour de trois termes principaux : OSTIE, CRISSE et TABARNAQUE. Une tendance remarquable est également observée : la fréquence de ces trois mots diminue de moitié par rapport à celle du mot précédent. Par exemple, la fréquence du mot OSTIE, qui compte 571 occurrences, diminue de moitié à 285 occurrences pour le mot CRISSE, puis tombe à 165 occurrences pour le mot TABARNAQUE. De plus, CIBOIRE a une fréquence relativement plus faible que les autres sacres, ce qui en fait le deuxième juron le moins utilisé dans le corpus. MARDE est aussi fréquent que CÂLICE. Parmi les jurons dérivés de l'anglais, seuls FUCK et SHIT sont présents dans le corpus, bien qu'ils soient relativement peu courants,

avec respectivement 59 et 6 occurrences. Les résultats suggèrent que les sacres sont souvent utilisés pour jurer, mais que les locuteurs québécois utilisent également d'autres types de jurons tels que les jurons scatologiques et les jurons empruntés à l'anglais.

Tableau 3 Les jurons et leurs variantes morphologiques trouvés dans le CFPQ

| Originel | Euphémisme | Verbe | Adjectif | Adverbe |
|--------------------------------|---|--|-----------------|----------------|
| câlce | câlce, câlique, câlif | (se) câliser, se décâliser | décâlifé | |
| ciboire | cibole | | | |
| crisse | crif, cristie | (se) crisser, décrisser, recrisser | | crissement |
| tabarnaque | tabarnache, tabarouette, tabarnique, tabarnouche | | | |
| ostie/estie | ostifie, ostine, ostique | | | |
| marde/merde³ | | s'emmerder | | |
| fuck | | fucker | fucké | fucking |
| shit | | | | |

Le tableau 3, basé sur la catégorisation de Vincent et Thibault (1981 cité dans Charette 1999, 150), fournit les variantes morphologiques trouvées dans le corpus, les classant comme euphémismes et dans trois catégories lexicales distinctes : Verbe, Adjectif et Adverbe. Alors que CÂLICE et CRISSE ont les formes flexionnelles les plus diverses, TABARNAQUE a plus de formes dérivationnelles. En effet, si OSTIE est le juron le plus répandu au Québec, aucune autre forme flexionnelle n'a été rencontrée, son usage se limitant principalement à sa forme originale et à ses formes dérivées. Aucune forme euphémique n'a été observée parmi les jurons empruntés à l'anglais. Les six occurrences de SHIT sont insuffisantes pour tirer de véritables conclusions. Le mot FUCK connaît une évolution dynamique en français québécois, donnant lieu à de nouvelles formes flexionnelles telles que Fucker et Fucké(E). Le suffixe -er est souvent utilisé en anglais pour construire des noms (par

³ À la suite d'une discussion avec des locuteurs québécois, il semble que le français québécois ait une façon distincte d'utiliser MARDE et MERDE. Merde est probablement utilisé comme une forme atténuée de MARDE. Un autre point soulevé est que le français du Québec n'utilise pas souvent MARDE seul. Dans ce cas, MERDE est le choix préféré, de la même manière que le français hexagonal utilise *Merde!* Néanmoins, aucune recherche n'a été portée à ma connaissance qui mentionne cette observation.

exemple, teacher, reader), alors qu'en français québécois, ce suffixe sert à construire des verbes, comme dans l'exemple suivant :

(6) pourquoi c'est juste pour fucker pour nous fucker (P28, L5 Sous-corpus 30).

Il est important, cependant, de rappeler que les variantes morphologiques trouvées dans la CFPQ présentent une diversité beaucoup moins large que l'inventaire des sacres établi par Charette (cf. Charette 1999, 141). Cela suggère l'existence de nombreuses autres variantes morphologiques des jurons analysés que le corpus n'a pas permis de saisir. Plus de détails seront abordés dans la section de discussion

4.2. Sexe

Passant à la catégorie sociale, le tableau 4 illustre la distribution de la fréquence des jurons selon le sexe du locuteur dans le CFPQ. Les résultats présentent notamment une observation inattendue : dans le contexte de données de la CFPQ, une plus grande proportion de femmes que d'hommes sont impliquées dans les jurons. Plus précisément, le nombre de participantes dépasse celui des participants par un facteur de 1,5, représentant respectivement 52 et 38 participants. L'analyse proportionnelle des données, avec un total de participants au CFPQ de 51 hommes et 61 femmes, a révélé que 70 % des hommes et 80 % des femmes jurent. Cependant, la taille de l'échantillon de cette étude n'était que de 112 participants. Compte tenu de la taille limitée de l'échantillon, les conclusions définitives doivent être tirées avec prudence. Un échantillon plus large et plus diversifié est nécessaire pour déterminer si les femmes sont plus nombreuses à jurer que les hommes.

Tableau 4 Distribution des jurons selon le sexe

| | Femmes (n = 52) | Hommes (n = 38) |
|-------------------|---------------------------|---------------------------|
| Câlce | 50 | 53 |
| Crisse | 66 | 219 |
| Ciboire | 20 | 26 |
| Ostie | 131 | 440 |
| Tabarnaque | 55 | 110 |
| Marde | 50 | 52 |
| Fuck | 33 | 26 |
| Shit | 2 | 4 |
| Total | 407 | 930 |

Cependant, en ce qui concerne la fréquence des jurons, les données révèlent le contraire. Les hommes ont tendance à utiliser ces jurons plus fréquemment que les femmes, comme le montre le taux d'occurrence plus élevé chez les hommes (69%) que chez les femmes (31%).

Notamment, cette tendance est significative pour les jurons courants CRISSE, OSTIE et TABARNAQUE, les occurrences masculines étant trois fois plus fréquentes dans le cas d’OSTIE et de CRISSE et deux fois plus élevées dans le cas de TABARNAQUE. Il est intéressant de noter que le mot FUCK est la seule exception, pour laquelle une fréquence plus élevée parmi les locuteurs féminins fut trouvée dans le corpus.

4.3. Sexe des groupes de conversation

Le tableau 5 présente la distribution des jurons parmi les groupes de conversation classés selon leur composition (mixte ou de même sexe), y compris la fréquence et la fréquence relative estimée par million de mots (FRE). La fréquence indique une légère divergence dans les schémas de jurons entre les groupes mixtes et les groupes de même sexe, avec une légère tendance à l’utilisation plus fréquente chez les locuteurs de même sexe. Toutefois, la présence d’un plus grand nombre de groupes mixtes dans l’échantillon soulève des questions, en raison du fait que la fréquence devrait être plus élevée dans le groupe ayant le plus de sous-corpora. Il est donc nécessaire de procéder à une estimation de la fréquence.

Tableau 5 Répartition des jurons en fonction du sexe du groupe

| | Fréquence | FRE (per million mots) |
|---------------------|-----------|------------------------|
| Groupe mixte | 623 | 1291 |
| Groupe de même sexe | 714 | 3107 |

La fréquence relative des jurons fut estimée sur la base du nombre total de tokens dans le corpus, étant donné que les groupes mixtes étaient plus nombreux que les groupes du même sexe et que la taille du sous-corpus était inconnue. Cette estimation est réalisable, car les sous-corpus furent enregistrés pendant presque la même durée d’une heure et demie. La formule d’estimation se trouve à l’annexe D. Une fois les données estimées et normalisées, la présence de jurons au sein de groupes distincts basés sur le genre devient évidente. Notamment, les groupes de même sexe présentent une fréquence de jurons nettement plus élevée que les groupes mixtes, avec plus du double du nombre de jurons par million de mots observés. Néanmoins, il est important de reconnaître que la formule fournit une estimation plutôt qu’une mesure exacte de la fréquence des jurons, et que les fréquences réelles peuvent s’écarter des estimations.

4.4. Âge

L’impact de l’âge sur les préférences de jurer devient évident à travers les données fournies dans le tableau 6. Les jurons ont tendance à s’intensifier et à atteindre leur apogée au début de l’âge adulte, mais ils diminuent progressivement au fil du temps. Fait

remarquable, la tranche d'âge des 20 à 29 ans est la plus répandue, puisqu'elle représente 44 % des cas de jurons observés. De manière inattendue, une baisse notable est observée dans la tranche d'âge des 30 – 39 ans, qui représente 13,8 % des occurrences. La fréquence des jurons connaît une résurgence dans la tranche d'âge des 40 – 49 ans, marquant un second pic. Après cela, la fréquence des jurons diminue progressivement. Les proportions pour les groupes d'âge 50 – 59, 60 – 69 et plus de 70 ans sont plus faibles, avec respectivement 2,2 %, 3 % et 1,1 %.

Tableau 6 Répartition des jurons par âge

| Âge | < 20 | 20 – 29 | 30 – 39 | 40 – 49 | 50 – 59 | 60 – 69 | > 70 |
|--------------------|------|---------|---------|---------|---------|---------|------|
| Câllice | 12 | 53 | 11 | 11 | 6 | 8 | 2 |
| Crisse | 5 | 178 | 18 | 70 | 3 | 6 | 5 |
| Ciboire | 0 | 17 | 13 | 9 | 1 | 3 | 3 |
| Ostie | 18 | 277 | 41 | 197 | 33 | 5 | 0 |
| Tabarnaque | 3 | 59 | 28 | 45 | 12 | 13 | 5 |
| Marde | 13 | 37 | 29 | 15 | 3 | 5 | 0 |
| Fuck | 14 | 23 | 12 | 8 | 2 | 0 | 0 |
| Shit | 1 | 2 | 3 | 0 | 0 | 0 | 0 |
| Total | 66 | 646 | 155 | 355 | 60 | 40 | 15 |
| Pourcentage | 5 % | 48 % | 12 % | 27 % | 4 % | 3 % | 1 % |

En examinant plus en détail, les résultats suggèrent une différence dans l'utilisation des jurons entre les différentes générations. OSTIE est le juron le plus utilisé dans la plupart des groupes, des moins de 20 ans aux 50 – 59 ans. Ce n'est toutefois pas le cas chez les personnes de plus de 60 ans, où les occurrences de TABARNAQUE et de CRISSE sont plus nombreuses. En outre, l'utilisation de FUCK est plus fréquente dans la population jeune, avec un pic de 23 occurrences dans le groupe des 20 – 29 ans, suivi par le groupe des moins de 20 ans. Aucune occurrence n'a été trouvée dans la génération plus âgée. La même tendance est observée dans le cas de SHIT, mais le nombre extrêmement faible d'occurrences trouvées dans le corpus n'est pas suffisant pour tirer des conclusions sur cet usage particulier.

4.5. Niveau d'éducation

En se penchant sur une autre catégorie sociale pertinente, celle du niveau d'éducation, la distribution de la fréquence d'utilisation des jurons en fonction du niveau d'éducation du locuteur est présentée dans le tableau 7. Sur l'ensemble de l'échantillon, 0,7 % des participants qui emploient des jurons n'ont suivi qu'une formation au niveau primaire. Il n'y a qu'une légère différence dans la fréquence des jurons entre les trois autres groupes, les

occurrences les plus élevées étant observées chez les personnes dont le niveau d'études le plus élevé est le secondaire, suivies par les locuteurs titulaires d'un Cégep et d'un diplôme universitaire. Plus précisément, les jurons observés au sein de chaque niveau d'études – secondaire, cégep et université – représentent environ un tiers de l'échantillon. Néanmoins, en y regardant de plus près, de nombreux jurons sont plus fréquents dans un groupe que dans l'autre. Par exemple, MARDE, SHIT et CIBOIRE apparaissent plus souvent dans le groupe des diplômés universitaires, OSTIE et TABARNAQUE dans le groupe des diplômés du Cégep, et CÂLICE, CRISSE et FUCK dans le groupe des diplômés du secondaire. Dans l'ensemble, cela montre que les personnes issues de différents groupes éducatifs jurent de la même manière, mais que la distinction réside principalement dans le choix des mots.

Tableau 7 Distribution du niveau d'éducation

| | Primaire | Secondaire | Cégep | Universitaire |
|--------------------|-----------------|-------------------|--------------|----------------------|
| Câllice | 1 | 46 | 22 | 34 |
| Crisse | 0 | 119 | 90 | 76 |
| Ciboire | 3 | 10 | 10 | 23 |
| Ostie | 0 | 195 | 239 | 137 |
| Tabarneau | 5 | 52 | 58 | 50 |
| Marde | 0 | 29 | 22 | 51 |
| Fuck | 0 | 26 | 8 | 25 |
| Shit | 0 | 1 | 1 | 4 |
| Total | 9 | 478 | 450 | 400 |
| Pourcentage | 0,7 % | 35,8 % | 33,7 % | 29,9 % |

4.6. Euphémisme

Les résultats montrent que les jurons euphémiques ne représentent qu'un quart des occurrences totales dans le CFPQ, comme le montre la figure 5. Sur les cinq sacres examinés, CÂLICE présente la fréquence la plus élevée de la forme atténuée, qui représente plus des deux tiers de ses occurrences, dépassant ainsi sa forme originale. Les jurons euphémiques sont également plus nombreux que leur forme originale dans le cas de CIBOIRE et TABARNAQUE, où ils sont utilisés deux fois plus souvent. Les deux jurons les plus fréquemment rencontrés dans le corpus, OSTIE et CRISSE, ont des fréquences d'utilisation euphémique relativement faibles par rapport aux autres sacres. Plus précisément, OSTIE a la fréquence d'euphémisation la plus faible, représentant environ 4 % de ses occurrences totales.

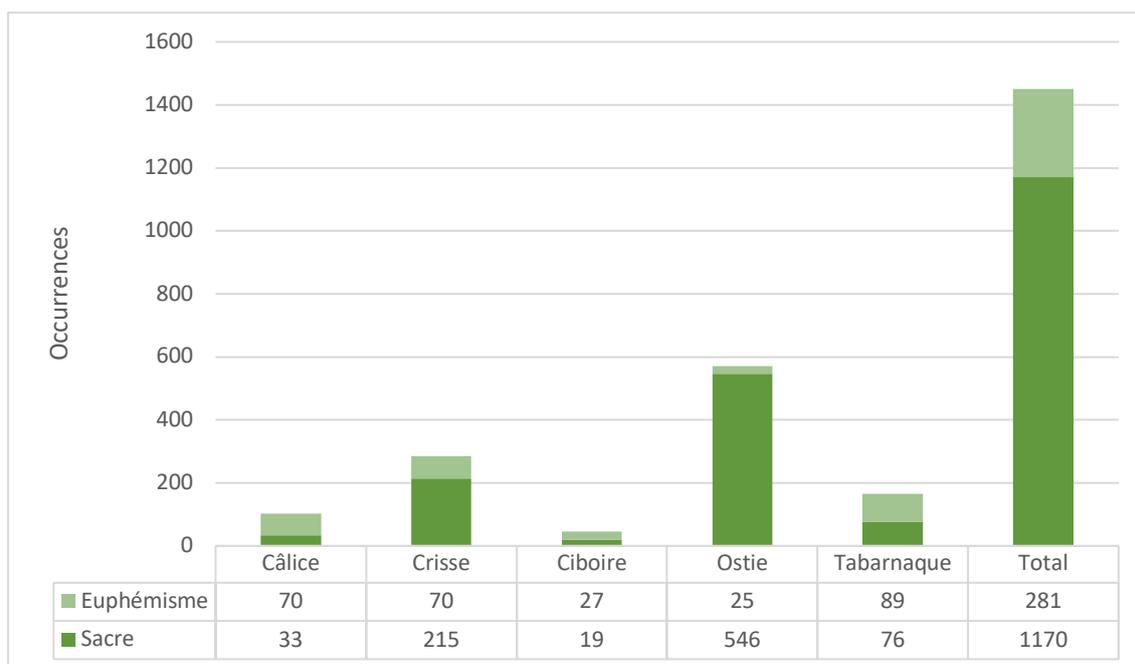


Figure 5 Répartition des euphémismes comparée au total des occurrences

En ce qui concerne le sexe des locuteurs, comme le montre le tableau 8, les données attestent que, dans l'ensemble, les locuteurs masculins ont légèrement plus d'occurrences d'euphémismes que les femmes, avec 161 et 120 respectivement. Cependant, lorsque la fréquence est normalisée et que la proportion d'euphémismes dans l'inventaire des jurons pour les hommes et les femmes est comparée, la situation change. Les locuteurs masculins ont des occurrences de jurons nettement plus élevées, mais l'utilisation d'euphémismes ne représente qu'un cinquième des jurons trouvés dans les occurrences masculines (19 %). En revanche, les femmes, dont les résultats ci-dessus montrent une fréquence relativement faible de jurons, ont en réalité une proportion plus élevée de jurons euphémiques, avec près d'un quart des occurrences prononcées par des locutrices (37 %). Les résultats suggèrent une corrélation négative entre le sexe en ce qui concerne les jurons et les euphémismes.

Tableau 8 Proportion d'euphémismes dans le nombre total d'occurrences de jurons par sexe

| | Homme | Femme |
|------------------------------|-------|-------|
| Fréquence | 161 | 120 |
| Occurrence proportion | 19 % | 37 % |

Une tendance similaire peut être observée en ce qui concerne l'âge des **orateurs** dans le tableau 9. Le groupe des 20 – 29 ans, qui compte le plus grand nombre de jurons, est celui qui se sert le moins de l'euphémisme, puisqu'il ne représente que 12 % du total des jurons prononcés dans ce groupe d'âge. Le groupe de locuteurs âgés de 60 – 69 ans et de plus de 70 ans a la plus basse fréquence de jurons, mais plus de la moitié de leurs énoncés sont des

formes d'euphémisation, en particulier dans le cas des plus de 70 ans. Ces résultats indiquent une corrélation potentielle entre l'âge et l'utilisation d'expressions euphémiques dans les jurons. Néanmoins, il est important de reconnaître que les données sont limitées.

Tableau 9 Proportion d'euphémismes dans le nombre total d'occurrences de jurons par âge

| Âge | < 20 | 20 – 29 | 30 – 39 | 40 – 49 | 50 – 59 | 60 – 69 | > 70 |
|-------------------|--------|---------|---------|---------|---------|---------|--------|
| Fréquence | 19 | 78 | 42 | 95 | 16 | 21 | 10 |
| Proportion | 28,8 % | 12,1 % | 27,1 % | 26,8 % | 26,7 % | 52,5 % | 66,7 % |

5. Discussion

Dans cette partie, une évaluation du caractère offensant, de son association avec les normes sociales et de son influence sur le choix des mots sera présentée, sur la base des résultats de l'étude. Le rapport entre le niveau d'offensivité et la conscience des individus à cet égard sera exploré. Ensuite, la convergence de la sociolinguistique et de la linguistique de corpus sera examinée afin de mieux comprendre les modèles et les changements linguistiques. Les avantages de la linguistique de corpus, notamment la réfutation des idées préconçues et la possibilité d'observer les changements linguistiques, seront soulignés, parallèlement à ses inconvénients, tels que le caractère incomplet des données et les distorsions potentielles.

5.1. Niveau d'offensivité

Le fait de jurer semble être lié à la façon dont le locuteur perçoit l'offense. Les résultats révèlent qu'un nombre considérable de femmes jurent, mais qu'elles le font moins fréquemment que les hommes. En particulier, les jurons euphémiques sont fréquents chez les femmes et les personnes âgées. La perception de caractère offensant découle potentiellement des normes sociales, qui sont des règles tacites servant comme lignes directrices implicites qui dictent le comportement, les valeurs et les attentes acceptables au sein d'une communauté en particulier. Certains comportements sont socialement plus acceptables, tandis que d'autres le sont moins. Les individus modifient fréquemment leur langage et leurs expressions afin d'éviter tout langage inapproprié ; lorsqu'un langage vulgaire est utilisé, il est souvent accompagné d'excuses (Widlak 1965, 941; Huston 1980, 74). Les normes sociales jouent donc un rôle essentiel dans le discernement du langage approprié, influençant ainsi la façon dont les locuteurs perçoivent les mots qu'il est acceptable de prononcer et ceux qu'il convient d'éviter. De nombreuses recherches précédentes ont mis en évidence des caractéristiques sociales de l'utilisation du langage qui

peuvent être liées au jugement de l'offensivité. Étant donné que les jurons prononcés par les femmes ont tendance à susciter une plus grande désapprobation sociale que ceux prononcés par les hommes (Stapleton 2010, 293), les locutrices de différentes cultures ont tendance à jurer moins, mais à éprouver des émotions négatives plus fortes à l'égard du langage offensant que les locuteurs masculins (Sapolsky, Shafer, et Kaye 2011, 13; Berowa, Ella, et Lucas 2019, 21).

Les jurons prononcés par les femmes semblent être un sujet de recherche pertinent. Il a été observé que les jurons à caractère très offensif sont plus souvent associés à des locuteurs masculins, tandis que les mots moins offensifs sont plutôt utilisés par des locutrices (McEnery 2006, 30). En introduisant le concept de « langage des femmes », Lakoff souligne que les attentes de la société encouragent les femmes à éviter les expressions émotionnelles fortes et assertives telles que les jurons, tout en favorisant l'articulation de l'atténuation et de l'incertitude (1973, 45). Par conséquent, les femmes emploient plus souvent des euphémismes dans leur discours par rapport aux hommes, comme le montrent les résultats du corps de cette étude. Les tabous, l'étiquette et les normes sociales restent présents dans la société, mais leur caractère et leurs principes changent constamment (Widlak 1965, 933). Sous l'influence de l'émancipation des femmes, les rôles des hommes et des femmes évoluent. Bougaïeff (1980) observe une modification notable des jurons chez les femmes françaises du Québec. Cette modification pourrait expliquer la forte proportion d'hommes et de femmes qui jurent dans le CFPQ, bien que leur fréquence soit moins élevée que celle des hommes.

L'âge du locuteur joue un rôle important dans le résultat comportemental des jurons, en particulier dans le contexte social et la conscience de l'environnement. En raison du caractère inapproprié associé aux jurons, les locuteurs adultes ont tendance à s'abstenir d'utiliser des jurons, quel que soit leur sexe, en présence de jeunes enfants (Jay 1992 cité dans Beers-Fägersten 2012, 11). Il existe également une corrélation positive entre la fréquence des jurons chez les parents et l'âge de leurs enfants, les parents étant plus susceptibles de jurer lorsque leurs enfants sont plus âgés (Berger 2002 cité dans Beers-Fägersten 2012, 13). Cela correspond au deuxième pic de fréquence des jurons observé dans le groupe des 40 – 49 ans selon nos résultats. Autrement dit, les personnes qui ont de jeunes enfants à la maison sont susceptibles d'utiliser des jurons moins souvent que celles qui n'en ont pas, et celles qui ne vivent pas avec leurs parents peuvent les utiliser plus fréquemment (McEnery et Xiao 2004, 241). Avec l'âge, une meilleure conscience des normes sociales et de la perception du langage peut conduire à une plus grande maîtrise de soi et à l'utilisation d'un langage plus

acceptable socialement. Ainsi, l'effet des normes sociales sur la perception de l'offense devient évident, ce qui contribue à la diversité des préférences en matière de jurons au sein des groupes sociaux.

La conscience qu'ont les locuteurs des implications offensantes et des jugements sociaux potentiels que leurs interlocuteurs peuvent porter sur leurs choix linguistiques est un facteur crucial à prendre en compte. Les jurons peuvent générer un sentiment de vulgarité, qui se traduit par la perception d'une compétence, d'une fiabilité, d'une sociabilité, d'une sympathie et d'une politesse moindres (DeFrank et Kahlbaugh 2019, 136). Cependant, les jugements sur les jurons dépendent de nombreux facteurs, tels que le lien entre le locuteur et l'auditeur, le contexte, le sexe du locuteur et le mot qu'il utilise (Jay et Janschewitz 2008, 272). Nos observations ont révélé que la conscience d'un jugement potentiel entraînait une fréquence de jurons plus élevée dans les conversations entre personnes de même sexe que dans les conversations entre personnes de sexe différent, même si les participants se connaissaient bien et que les conversations étaient privées. La différence entre les sexes pourrait découler d'un préjugé sexiste sous-jacent, comme l'indiquent les recherches de DeFrank et Kahlbaugh sur l'impression de sexe en présence de profanations. Les participants masculins percevaient les jurons des femmes comme plus offensants que ceux des hommes, tandis que les participantes féminines trouvaient les jurons des hommes plus offensants que ceux des femmes (2019, 137). Pour éviter de paraître insolents, les locuteurs hésitent parfois à utiliser un langage offensant en présence de personnes du sexe opposé. Cette sensibilisation influence leur comportement par rapport aux jurons, et conduit à une réduction spontanée des jurons afin de maintenir un certain niveau d'adéquation sociale.

Les degrés variables du caractère offensant des jurons témoignent de la diversité des perceptions, qui se reflète dans l'adoption d'euphémismes. Les jurons, autrefois considérés comme une transgression gravement punissable, ont ensuite donné naissance aux euphémismes, destinés à atténuer les connotations négatives et à exprimer les sentiments sans enfreindre les normes divines ou les règles de la communauté (Vincent 1984, 57). Il convient de rappeler que, même à la fin du XIX^e siècle, la célèbre « Merdre » d'Ubu, bien que moins brutale et dissimulée sous forme d'euphémisme, a suscité un niveau significatif d'indignation publique et critique (Kor 2017, 137). Nos résultats révèlent que le juron au Québec est principalement dérivé des sacres, malgré l'existence d'autres catégories de jurons ; de plus, les variations euphémiques ne se retrouvent que dans les inventaires de sacres. Ces résultats mettent en lumière les différents degrés d'offensivité : les sacres sont susceptibles d'avoir un niveau d'offensivité relativement élevé. L'une des explications de l'absence

d'euphémisme dans d'autres types de jurons est que les jurons autres que les sacres au Québec peuvent être considérés comme moins offensants, peut-être à un point tel que la création de formes alternatives est jugée inutile. En particulier, une autre explication plausible de cette distinction réside dans la différence de perception des jurons en termes d'offensivité entre la première langue (L1) et la deuxième langue (L2) des locuteurs.

Les variations entre L1 et L2 se manifestent également dans le domaine du processus émotionnel, ce qui souligne que la manière dont les émotions sont perçues et exprimées peut également influencer le choix des mots. De nombreuses études portent sur les différentes perceptions des locuteurs multilingues. Harris, Ayçiçeği et Gleason (2003) ont réalisé une étude physiologique et affirment que les participants présentent des réactions physiologiques nettement plus fortes lorsqu'ils sont exposés à des jurons dans leur L1, le turc, que dans leur L2, l'anglais. Malgré les niveaux de compétence et la fréquence d'utilisation des deux langues, la L1 est souvent préférée pour jurer, mais les jurons dans la L2 sont perçus comme moins intenses au niveau émotionnel (Dewaele 2010). Dans certaines cultures comme la culture chinoise, l'utilisation de la L2 pour jurer est souvent privilégiée pour transmettre des émotions fortes, dans la mesure où les contraintes sociales au sein de la culture chinoise encouragent la suppression de l'expression émotionnelle forte (Caldwell-Harris et al. 2011). Toutefois, il est essentiel de reconnaître que les mots qui auraient pu être considérés comme offensants dans des études antérieures peuvent aujourd'hui ne plus être profanes en raison du processus graduel de désensibilisation au fil du temps (DeFrank et Kahlbaugh 2019). D'autres recherches diachroniques sont nécessaires pour comprendre le niveau d'offensivité du français québécois au fil du temps.

Après avoir examiné le niveau d'offensivité, il est évident que les termes jurés comme sacres sont plus offensants que les autres types de jurons et sont utilisés exclusivement plus fréquemment dans le contexte du Québec. Les résultats de cette étude sont conformes au paradoxe des jurons proposé par Beers-Fägersten (2012), qui montre une corrélation positive entre la fréquence des occurrences de jurons et des évaluations plus fortes de l'offensivité. Ceci débouche sur la question suivante : Qu'est-ce qui motive les gens à dire ces mots ? Si la première motivation de l'utilisation de jurons est souvent de transmettre des émotions intenses telles que la colère, la frustration, la joie ou la surprise (Jay 2009, 155), ce n'est pas le seul objectif. Les jurons au quotidien visent d'autres objectifs que l'expression d'émotions fortes ou d'une simple colère. Les linguistes conviennent largement que la variation contextuelle des jurons est étroitement liée à la relation entre les interlocuteurs et au contexte social (Jay et Janschewitz 2008, 285; Jay 2009, 157; Kapoor 2016, 04). Un seul mot a le

potentiel de susciter des émotions variées en fonction des intentions des locuteurs et du contexte de leur conversation. Lorsqu'il est utilisé judicieusement, le juron permet d'exprimer l'appartenance à une communauté spécifique, de renforcer les liens interpersonnels, de consolider la crédibilité du message et l'autorité du locuteur, et d'affirmer les identités de groupe (Stapleton et al. 2022, 07-08). Il est essentiel de répliquer ces résultats et d'étudier plusieurs corpus afin de découvrir les facteurs sous-jacents qui motivent les choix de jurons dans différents contextes.

5.2. Quand la linguistique de corpus rencontre la sociolinguistique

Nos résultats démontrent clairement un intérêt partagé entre la linguistique de corpus et la sociolinguistique. L'incorporation de corpus dans la sociolinguistique n'est pas un développement récent ; au contraire, elle gagne régulièrement du terrain dans la recherche sociolinguistique. Le laboratoire de sociolinguistique de l'Université d'Ottawa sert de plaque tournante pour le stockage de nombreuses heures d'enregistrements de discours spontanés. Cette collection comprend d'importants dépôts de données sur les langues parlées dans différentes communautés d'Amérique du Nord, spécialement conçus pour l'analyse sociolinguistique (Poplack 2021, 01-02). La linguistique de corpus s'est avérée utile pour détecter la corrélation entre les facteurs sociaux et l'utilisation de la langue, pour soutenir ou rejeter certaines hypothèses sociales sur la base d'une méthode scientifique empirique. Les corpus ainsi que les données démographiques se révèlent être des sources précieuses pour l'analyse synchronique et diachronique de l'utilisation de la langue (Beeching 2006, 52). Les analyses de corpus facilitent également l'examen des phénomènes linguistiques liés aux contextes sociaux. Des phénomènes tels que la grammaticalisation, la généralisation sémantique et la pragmatization nécessitent une exploration dans des cadres temporels et sociaux (Ibid., 53 – 54). Mair souligne une convergence notable entre ces deux sous-domaines linguistiques : tous deux privilégient l'étude des énoncés dans des contextes discursifs plutôt que des règles abstraites, et soulignent l'importance des données de fréquence et des statistiques (Mair 2004, 121).

Il est courant de penser que les mots tabous sont employés lorsque les locuteurs ont du mal à trouver une expression appropriée, signe d'un manque de vocabulaire (Anderson et Trudgill 1992, 65). Le jurement a également une connotation négative et est souvent associé à des individus dont la réputation est moins favorable, tels que les criminels ou les personnes perçues comme étant sans éducation (Patrick 1901, 117; Stapleton 2020, 29). Nos données indiquaient néanmoins qu'il n'y avait pas de corrélation entre le niveau d'éducation et le fait

de jurer. Certaines études réfutent pareillement cette supposition. Giordano & College (2016) affirment qu'aucune relation statistique n'a pas été trouvée entre les jurons et l'intelligence après avoir effectué une analyse sur les scores de QI et l'enquête. Jay & Jay (2015) réfutent la conception erronée selon laquelle les personnes qui utilisent des jurons ont un vocabulaire limité, affirmant que la capacité à produire un langage tabou n'indique pas un manque de richesse linguistique. Sans qu'une corrélation entre le niveau d'éducation et le comportement lié aux jurons puisse être observée, un lien intrigant apparaît en ce qui concerne la perception du langage offensant. Un sondage révèle que les Québécois de bas niveau d'éducation considèrent qu'il n'est pas nécessaire de marquer les gros mots dans les dictionnaires, alors que les personnes de niveau d'éducation plus élevé soutiennent généralement l'étiquetage réfléchi du langage vulgaire (Maurais 2008, 35). Ainsi, la linguistique de corpus est manifestement une méthode vigoureuse pour étudier la façon dont la langue est liée à la société, comme le montre cette étude sur les jurons au Québec. Avec des exemples linguistiques réels, des preuves concrètes de la sociolinguistique furent fournies à la fois dans l'analyse qualitative et quantitative.

Le changement linguistique est l'un des facteurs qui peut être observé à travers la linguistique de corpus. Même si les linguistes se concentrent souvent sur l'environnement linguistique à l'origine du changement, il semble que des facteurs sociaux aient été négligés (Wagner 2012, 371). Labov a été le précurseur d'une approche visant à examiner les changements de caractéristiques linguistiques en analysant les différents âges au sein d'une communauté donnée au lieu de s'appuyer uniquement sur les archives historiques. Son étude de la stratification sociale, qui porte sur la centralisation des diphtongues sur l'île de Martha's Vineyard et l'utilisation distincte de /r/ dans les grands magasins de New York, a permis d'établir une compréhension fondamentale du lien entre les facteurs sociaux et l'évolution du langage (Labov 1962; 1972). Avec des données adéquates sur la classe sociale, le sexe, l'origine ethnique et les attitudes des locuteurs à l'égard du parler local, il est possible de déduire la motivation sociale qui sous-tend les changements linguistiques. Cela conduit à l'interprétation par Labov des modèles de changement au sein de l'individu et de la communauté, comme le montre la Figure 6 ci-dessous (Labov 1994, cité dans Sankoff et Blondeau 2007, 562)

| SYNCHRONIC PATTERN | INTERPRETATION | INDIVIDUAL | COMMUNITY |
|------------------------|--|------------|-----------|
| flat | 1. Stability | stability | stability |
| regular slope with age | 2. Age grading | change | stability |
| regular slope with age | 3. Generational change (= apparent-time interpretation) | stability | change |
| flat | 4. Communal change | change | change |

Figure 6 Modèle de changement chez l'individu et dans la communauté

La gradation en fonction de l'âge et le changement générationnel sont les deux types de changement qui suscitent le plus d'intérêt de la part des sociolinguistes. L'usage de la langue change avec l'âge, avec un impact au niveau individuel ; dans le cas du français québécois, les jurons sont plus fréquents dans la vingtaine et diminuent progressivement – un changement au niveau individuel sans impact à l'échelle de la communauté. Ce type de changement correspond davantage à une échelle d'âge qu'à un changement de langue (Sankoff 2019, 199). En revanche, le changement générationnel suggère une stabilité linguistique chez les individus jusqu'à l'âge adulte, tandis que les changements linguistiques commencent par la jeune génération et s'étendent progressivement à la génération plus âgée, pour finalement façonner la langue utilisée par la communauté (Sankoff et Blondeau 2007, 562). Compte tenu de l'augmentation de la population bilingue français – anglais au Québec, la hausse des jurons anglais parmi la jeune génération laisse entrevoir une tendance potentielle à l'intégration de ces mots dans le français québécois. Le troisième scénario implique des changements à la fois à l'échelle de la communauté et au niveau individuel, lorsque l'ensemble de la communauté adopte un trait linguistique, conduisant à une distribution uniforme à travers les groupes d'âge et les générations. Pour les recherches futures, il sera essentiel d'effectuer des analyses diachroniques supplémentaires pour différencier le changement générationnel du changement générationnel accompagné d'une variation de l'âge lors de l'observation de cette transformation communautaire.

Il est certain que le corpus présente de nombreuses avancées qui contribuent grandement à examiner les phénomènes sociaux dans le cadre de la sociolinguistique. Nos résultats soulignent non seulement les avantages de la linguistique de corpus dans l'analyse sociolinguistique, mais ils mettent également en évidence des limites spécifiques qu'il convient de reconnaître. Certains problèmes secondaires, tels que la faible fréquence des données, empêchent de tirer des conclusions précises, comme le montrent les résultats concernant le juron anglais SHIT, ou la combinaison d'une faible fréquence de jurons et d'un nombre élevé d'euphémismes dans le groupe d'âge des plus de 70 ans. Ces facteurs empêchent de déterminer si les jurons utilisés sont principalement des euphémismes.

L'occurrence d'une faible fréquence pourrait potentiellement entraîner du *cherry picking*, où seules les données les plus favorables à une déclaration sont intentionnellement choisies (Morse 2010), ce qui conduit à une mauvaise interprétation et, de manière plus critique, à une généralisation excessive et à des conclusions inexactes. De plus, l'incomplétude des données reste problématique dans le processus de la linguistique de corpus, ce qui fit l'objet de nombreuses études dans différentes disciplines. Ce caractère incomplet des données met en évidence le défi de l'absence, qui donne lieu à deux préoccupations principales. Le premier concerne l'absence d'éléments qui ne peuvent être trouvés dans le corpus, tandis que le second concerne l'incapacité d'expliquer certains choix linguistiques (Baker et al. 2008, 296; Partington 2014, 120). De même, Mair déclare que tous les processus de grammaticalisation « ne laissent pas une empreinte statistique » dans les corpus (2004:133).

Ce sujet a été abordé par Chomsky qui s'opposa à l'utilisation de corpus linguistiques dans une interview :

Chomsky: The verb *perform* cannot be used with mass word objects: one can *perform* a task but one cannot *perform* labour.

Hatcher: How do you know, if you don't use a corpus and have not studied the verb *perform*?

Chomsky: How do I know? Because I am a native speaker of the English language.

(Hill 1962 cité dans McEnery et Wilson 2005, 11)

La perspective de Chomsky souligne en effet l'importance de l'intuition du locuteur natif lorsqu'il s'agit d'émettre des jugements grammaticaux. L'intuition du locuteur natif sert de protection contre l'absence d'éléments ou d'ensembles linguistiques dans le corpus. Cette préoccupation est aggravée pour les locuteurs non natifs, qui n'ont pas cette capacité intuitive à évaluer la grammaire et qui s'appuient sur les corpus pour obtenir des informations précises sur la fréquence linguistique. En comparaison avec des études antérieures menées par des locuteurs natifs du français québécois telles que Vincent & Thibaut (1981) et Charette (1999), comme détaillé dans l'annexe A et B, nos résultats révèlent un inventaire de *sacres* relativement plus restreint. Cet écart soulève une préoccupation critique : de nombreux *sacres* et leurs variantes morphologiques semblent être tombés hors du corpus. Cela met en évidence le dilemme de la fiabilité de l'analyse empirique par le corpus linguistique par rapport à l'intuition d'un locuteur natif. Pourtant, une étude sur la façon dont les locuteurs natifs et les apprenants de la langue italienne jugent la collocation des mots montre qu'il y a une solide correspondance entre leurs intuitions et les résultats du corpus pour les mots qui sont utilisés très fréquemment ou très rarement (Siyanova-Chanturia et Spina 2015). Ainsi, McEnery et Wilson suggèrent de demander

conseil à des locuteurs natifs de la langue en question afin de traiter les structures syntaxiques peu familières mais grammaticalement correctes (2005, 12). Un lien potentiel pourrait également exister entre la méthode de la linguistique de corpus et l'intuition du locuteur natif. Des recherches supplémentaires sont nécessaires pour explorer comment l'intuition et l'analyse de corpus peuvent être utilisées plus efficacement pour répondre aux limites de la méthodologie de la linguistique de corpus.

Pour résumer, la linguistique de corpus sert efficacement les objectifs sociolinguistiques, notamment en étudiant la corrélation entre les variables linguistiques et sociales, qui est au cœur de la recherche sociolinguistique depuis plus de trois décennies (Mair 2004, 01). Les corpus sont efficaces pour contribuer aux phénomènes sociolinguistiques et au changement linguistique, en ce sens qu'ils fournissent de nombreuses données authentiques, qui peuvent être analysées statistiquement tout en réduisant les biais. Le caractère incomplet des données est l'un des inconvénients à prendre en compte. Dans le cadre de la recherche, il est essentiel de ne pas traiter les données de corpus comme équivalentes à la langue réelle et d'évaluer soigneusement les conclusions tirées des données de corpus avant de les généraliser à une population plus large (McEnery et Hardie 2012, 26). Il est suggéré de poser les bonnes questions, d'effectuer des recherches précises et de tirer des conclusions valables basées à la fois sur ce qui est présent et ce qui est absent dans les données relatives à un usage spécifique (Solan et Gales 2017, 1357). Ces étapes nécessitent un examen méticuleux, faute de quoi les données, comme l'a suggéré Mair (2004, 22), peuvent se transformer en un « Zahlenfriedhof » – un enchevêtrement de statistiques sans explication.

6. Conclusion

En conclusion, cette recherche montre que les jurons font partie intégrante du langage quotidien et qu'ils possèdent des caractéristiques linguistiques distinctes qui sont susceptibles de faire l'objet d'une analyse linguistique. L'objectif de cette étude n'était pas de fournir un manuel faisant autorité pour jurer comme un locuteur natif du français québécois. Le premier objectif était d'examiner les choix de jurons au Québec d'un point de vue sociolinguistique, tout en explorant la contribution des corpus à l'analyse sociolinguistique. Essentiellement, cette étude représente la convergence de la sociolinguistique et de la linguistique de corpus. L'enquête s'est penchée sur la prévalence des jurons dans le français informel parlé au Québec, en se concentrant sur les dix formes de jurons les plus fréquentes. Il est important de noter que cette étude ne prétend pas que les

résultats s'appliquent à tous les locuteurs du Québec, mais ce sont les résultats qui furent trouvés dans le CFPQ.

En revenant aux questions de recherche, cette étude révèle que les jurons au Québec vont au-delà des sacres, englobant d'autres jurons comme MARDE, FUCK et SHIT. Cependant, les jurons les plus courants se concentrent sur les trois sacres principaux, OSTIE étant le plus courant, suivi de CRISSE et de TABARNAQUE. SHIT est moins fréquemment utilisé, suivi de CIBOIRE (QR 1). En outre, cette étude a classé les variations morphologiques en groupes distincts, révélant que les euphémismes sont exclusivement présents dans le groupe de sacres. En comparaison avec des études menées par des locuteurs natifs du français québécois, de nombreuses variantes morphologiques des mots cibles sont absentes du corpus (QR 2). Le sexe joue un rôle dans la fréquence des jurons, les hommes jurant plus souvent que les femmes et utilisant moins d'euphémismes. Les jurons sont plus fréquents chez les jeunes adultes âgés de 20 à 29 ans, avec un deuxième pic dans la tranche d'âge des 40 – 49 ans. Il est intéressant de noter que l'utilisation d'euphémismes augmente avec l'âge. L'étude réfute également l'hypothèse selon laquelle le niveau d'éducation influe de manière significative sur les habitudes en matière de jurons, vu que la distribution est la même pour les différents niveaux d'éducation dans les données. Les emprunts à l'anglais comme FUCK et SHIT sont liés à la jeune génération. Néanmoins, les faibles occurrences conduisent à interpréter ces résultats avec prudence (QR 3).

Ces résultats conduisent à considérer les implications sociolinguistiques. Le CFPQ a indéniablement révélé les tendances concernant les jurons à travers différents facteurs sociaux. En adoptant un point de vue plus large, une attention particulière est portée aux implications inhérentes et aux explications potentielles des variations dans les choix de jurer. Les jurons sont beaucoup plus complexes qu'une simple expression de colère ou qu'une manifestation d'impolitesse. Dans le domaine du français québécois, le système de jurons reflète une tapisserie historique d'influence catholique et de contact linguistique avec l'anglais. Au-delà de ce contexte, les jurons englobent la production de mots ou d'expressions qui sont considérés comme tabous dans les normes sociales, mais qui persistent comme un défi aux conventions de la société. Cette interaction complexe conduit à un effet en cascade, où les normes de la société influencent la façon dont les mots sont perçus comme offensants ou (in)appropriés, ce qui se traduit par le choix de jurer. Les résultats de nos recherches soulignent l'impact significatif de ces normes sociales sur le choix des jurons. Par exemple, les femmes sont souvent jugées plus sévèrement lorsqu'elles utilisent des jurons, et il est généralement admis que les parents devraient éviter d'utiliser un

langage vulgaire lorsqu'ils sont en présence de jeunes enfants. Notre étude révèle également que les sacres sont généralement perçus comme plus offensants que les autres formes de jurons au Québec. Cette conclusion découle de l'observation d'une fréquence importante d'expressions euphémiques dans les résultats. La présence d'euphémismes souligne la prise de conscience du caractère offensant de certains mots tabous. Il est intéressant de noter que certains mots qui apparaissent plus fréquemment font également l'objet d'une utilisation plus importante d'euphémismes. Cela indique que les gens reconnaissent le caractère tabou du mot et s'efforcent de réduire leur impact en utilisant un langage plus atténué. Néanmoins, ces tentatives n'empêchent pas complètement le fait de jurer. Les preuves recueillies dans le cadre de cette étude soutiennent fortement l'existence du **paradoxe des jurons**.

Ces résultats soulignent sans aucun doute l'intérêt partagé pour l'analyse des variations communes à la linguistique de corpus et à la sociolinguistique. Ils indiquent de manière décisive que la linguistique de corpus remplit efficacement certains objectifs de la sociolinguistique, à savoir l'étude de l'interaction entre des variables linguistiques dépendantes et des variables sociales indépendantes. Ils confirment également la pertinence des corpus informatisés et des pratiques de corpus standard pour améliorer notre compréhension des changements linguistiques et des facteurs sociolinguistiques qui façonnent les phénomènes linguistiques. En appliquant des méthodes scientifiques empiriques, les chercheurs peuvent justifier ou contester des présomptions sociales. Cependant, l'étude actuelle dévoile aussi une série de difficultés qui exigent l'attention de la linguistique de corpus, notamment en ce qui concerne les questions de représentativité, de biais statistiques et, plus particulièrement, le problème persistant de l'incomplétude des données. Le caractère incomplet suscite des critiques, telles que l'exclusion d'éléments du corpus qui sont néanmoins présents dans la langue naturelle. Ce dilemme pose une question : la solution consiste-t-elle à privilégier la linguistique de corpus ou à se fier à l'intuition du locuteur natif ? Les corpus, aussi neutres soient-ils, peuvent-ils vraiment présenter des données sans un certain degré de distorsion ? Cette question souligne la nécessité de poursuivre les recherches sur la précision de l'intuition du locuteur natif et sur la manière dont elle peut être appliquée à la linguistique de corpus.

Annexes

ANNEXE A : LES SACRES ET LEURS VARIANTES MORPHOLOGIQUES POTENTIELLES

Les sacres les plus utilisés en 1981

| Lexème religieux | Orthographe | Dérivés | Verbe | Adjectif | Adverbe |
|------------------|--------------------|--|--------------------------------------|--------------------------|-------------|
| Baptême | batême | batince batêche | ----- | ----- | ----- |
| Calice | câlice | câlasse câlique câline câiboire | câlisser décâlisser | câlissant décâlissant | câlissement |
| Calvaire | calvaire | calvince calvette calvasse | ----- | ----- | ----- |
| Christ | criss | clis, crif cristi cristophe cristal crousse christmas | crisser décrisser déconcrisser | crissant décrissant | crissement |
| Ciboire | ciboère ciboire | Cibole Cibolac saint-cibole saint-cicroche | désainci- boiriser* | ----- | ----- |
| Hostie | osti | ostic hostin hospice asti astic | Ostier* Ostifier* | ----- | ----- |
| Tabernacle | tabarnac | tabarnouche tabaslac tabarouette tabamanne tabarnic | Tabarnaquer* | ----- | ----- |

Source : THIBAULT et VINCENT, 1981, p. 215.

(Charette 1999, 150)

Les sacres et leurs dérivés¹²

| Sacre | Notes | Dérivés |
|------------|--|--|
| Calice | Il faut prononcer <i>câlisse</i> . | Câlisseur (euse), décâlisseur (euse), câlissage, décâlissage, câlissable, décâlissable, recâlissable, câlisser, décâlisser, recâlisser, contre-câlisser. |
| Calvaire | Il est moins employé que les autres sacres et ne semble pas se diversifier. Il est plus employé dans la langue parlée que dans la langue écrite. | |
| Christ | C'est certainement le plus employé de tous les sacres mais aussi le moins effectif. | Crisseur (euse), décrisseur (euse), déconcrisseur (euse), crissailleur (euse), décrissailleur (euse), crissage, décrissage, déconcrissage, crissage, décrissage, crissable, décrissable, déconcrissable, crisser, décrisser, déconcrisser, reconcrisser, crissailler, décrissailler, contre-crisser. |
| Ciboire | On entend le plus souvent <i>cibouère</i> ou <i>cibouère</i> . | Déciboire, déciboirable, déciboiriser, désaintciboiriser. |
| Hostie | On entend presque toujours <i>esti</i> ou <i>sti</i> . Ce sacre est un des plus employés mais ne semble pas se diversifier ni à l'oral, ni à l'écrit. | |
| Sacrement | On entend presque toujours <i>sacramant</i> . Nous le retrouvons très rarement dans la langue écrite. | Désacramentiser. |
| Tabernacle | On prononce toujours <i>tabarnak</i> . C'est ce sacre qui semble produire le plus d'effet que ce soit dans la langue parlée ou dans la langue écrite. | |

(Charette 1999, 141)

ANNEXE B : CLASSEMENT DES JURONS BASE SUR LA FONCTION
D'INTERJECTION EN FRANÇAIS QUÉBÉCOIS

Tableau 3.1

Les interjections en français québécois : appellations intuitives

| Interjection | | | | | |
|----------------------|------------|-----------|-----------|----------------------|----------------|
| Juron au sens large | | | | Gros mot | |
| Juron au sens étroit | Cas limite | Sacre | Blasphème | Termes scatologiques | Termes sexuels |
| Bâtard | Baptême | Câlice | Maudit + | Bullshit | Fuck |
| Bateau | Maudit | Calvaire | (sacre) | Fuck | [...] |
| Câline | Viarge | Ciboire | | Shit | |
| Sacrifice | [...] | Crisse | | Merde | |
| Taboire | | Hostie | | [...] | |
| Torpinouche | | Sacrament | | | |
| Torrieu | | Tabarnak | | | |
| [...] | | [...] | | | |

(Charette 1999, 84)

ANNEXE C : CODE POUR LA SUPPRESSION DE LA CONVENTION

```
1 import re
2 import os
3 import PyPDF2
4 from natsort import natsorted
5
6 # Take the directory where the files are stored
7 directory = '/Users/anhlefevre/Desktop/CleanCorpus/'
8
9 print('-----')
10 print('Liste des fichiers PDF à traiter : ')
11 for filename in natsorted(os.listdir(directory)):
12     f = os.path.join(directory, filename)
13     if os.path.isfile(f) and f.endswith('.pdf'):
14         print(f)
15 print('-----')
16
17 for filename in natsorted(os.listdir(directory)):
18     f = os.path.join(directory, filename)
19     if os.path.isfile(f) and f.endswith('.pdf'):
20         pdf = PyPDF2.PdfReader(open(os.path.expanduser(f), 'rb'))
21         number_of_pages = len(pdf.pages)
22         print("number_of_pages of file :", f, " : ", number_of_pages)
23         outputfilename = f+'.txt'
24         print(outputfilename)
25         if os.path.exists(os.path.expanduser(outputfilename)):
26             os.remove(os.path.expanduser(outputfilename))
27
28         for x in range(0, number_of_pages):
29             page = pdf.pages[x]
30             text = page.extract_text()
31             #text = re.sub("([\[]).*?([\]])", "", text)
32             text = re.sub("([\(\)].)*?([\)])", "", text)
33             #text = re.sub("([\<]).*?([\>])", "", text)
34
35             # Retirer ce qui est après une parenthèse ouverte
36             text = re.sub("(?=\(.*", "", text)
37             # Retirer ce qui est avant une parenthèse fermée
38             text = re.sub(".*(?:\})", "", text)
39             file = open(os.path.expanduser(outputfilename), 'a')
40             file.writelines(text)
41         print('Fichier ', outputfilename, ' traité')
```

ANNEX D : FORMULAIRE DE CALCUL DE LA FRÉQUENCE RELATIVE ESTIMÉE

Total des mots dans les 31 groupes : 712300

Total des jurons dans 10 groupes mono sexe : 714

Total des jurons dans 21 groupes mixtes : 623

En supposant que tous les textes aient à peu près la même longueur, un texte contient en moyenne $712300/31 = 22977$ mots.

Pour une conversation entre membres d'un même sexe

- Estimation du nombre de mots dans 10 textes : $22977*10 = 229770$ mots.
- Estimation de la fréquence relative des gros mots dans 10 textes : $714/229770*1000000 = 3107$ mots par million de mots.

Pour les conversations entre personnes de sexe différent :

- Estimation du nombre de mots dans 21 textes : $22977*21 = 482517$ mots.
- Estimation de la fréquence relative des gros mots dans 21 textes : $623/482517*1000000 = 1291$ mots par million de mots

Référence Bibliographie

Support papier

- Aarts, Bas. 2001. « Corpus linguistics, Chomsky and Fuzzy Tree Framents ». Dans *Corpus linguistics and linguistic theory.* , 5-13. Amsterdam: Rodopi.
- Aitchison, Jean. 2006. « Whassup? Slang and swearing among schoolchildren ». *Education review* 19 (2) : 18-24.
- Anderson, Lars-Gunnar, et Peter Trudgill. 1992. *Bad Language*. Penguin Books.
- Auckle, Tejshree. 2017. « Youth discourse in multilingual Mauritius: The pragmatic significance of swearing in multiple languages 1 ». *Stellenbosch Papers in Linguistics Plus* 51 : 61-86. <http://spilplus.journals.ac.za>.
- Baker, Paul, Costas Gabrielatos, Majid Khosravini, Michał Krzyzanowski, Tony McEnery, et Ruth Wodak. 2008. « A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press ». *Discourse and Society* 19 (3) : 273-306. <https://doi.org/10.1177/0957926508088962>.
- Bässler, Valerie. 2009. « Les sacres en français québécois - beaucoup plus que de simples décharges émotives ». Dans *Français du Canada - Français de France VIII: Actes du huitième Colloque international Trèves, du 12 au 15 avril 2007*. Sous la direction de Beatrice Bagola et Hans-J. Niederehe, 159-75. Tübingen : Max Niemeyer.
- Battistelli, Delphine, Cyril Bruneau, et Valentina Dragos. 2020. « Building a formal model for hate detection in French corpora ». Dans *Procedia Computer Science.* , 176 : 2358-65. Elsevier B.V. <https://doi.org/10.1016/j.procs.2020.09.299>.
- Beeching, Kate. 2006. « Synchronic and diachronic variation: the how and why of sociolinguistic corpora ». Dans *Corpus Linguistic around the world.* , 56 : 49-61.
- Beers Fägersten, Kristy. 2007. « A sociolinguistic analysis of swearword offensiveness ». *Saarland Working Papers in Linguistics (SWPL)* 1 : 14-37.
- Beers-Fägersten, Kristy. 2012. *Who's Swearing Now? The Social Aspects of Conversational Swearing*.
- Berowa, Annie Mae C., Jennibelle R. Ella, et Rochelle Irene G. Lucas. 2019. « Perceived Offensiveness of Swear Words Across Genders ». *The Asian EFL Journal* 25 (5.2) : 164-87. <https://www.researchgate.net/publication/341576891>.
- Blanchet, Diane. 2017. « Sacre-verbos et structure d'arguments universelle ». Montréal : Université du Québec.

- Boberg, Charles. 2012. « English as a minority language in Quebec ». *World Englishes* 31 : 493-502. <https://doi.org/10.1111/j.1467-971X.2012.01776.x>.
- Bougaïeff, André. 1980. « Un Trait du français populaire et familier au Québec: le système des “sacres” ». *The French Review* 53 (6) : 839-47.
- Bowers, Jeffrey S., et Christopher W. Pleydell-Pearce. 2011. « Swearing, euphemisms, and linguistic relativity ». *PLoS ONE* 6 (7). <https://doi.org/10.1371/journal.pone.0022341>.
- Brezina, Vaclav, et Dana Gablasova. 2018. « The corpus method ». Dans *English Language*. Sous la direction de Jonathan Culpeper, 595-609. London : Macmillan Education UK.
- Caldwell-Harris, Catherine L., Jimmy Tong, Winvy Lung, et Sinlan Poo. 2011. « Physiological reactivity to emotional phrases in Mandarin-English bilinguals ». *International Journal of Bilingualism* 15 (3) : 329-52. <https://doi.org/10.1177/1367006910379262>.
- Carpuat, Marine. 2014. « Mixed-Language and Code-Switching in the Canadian Hansard ». *Proceedings of the First Workshop on Computational Approaches to Code Switching*, 107-15.
- Charest, G. 1974. *Le livre des sacres et blasphèmes québécois*. Collection Connaissance des pays québécois. Montréal : L’Aurore. <https://doi.org/978-0885320004>.
- Charette, Eric. 1999. « Du sacre au nom de qualité: Approche théorique et description grammaticale ». Sherbrook : Université de Sherbrook.
- Corbeil, Jean-Pierre, et René Houle. 2016. « The evolution of English–French bilingualism in Canada from 1901 to 2011 ». Québec.
- DeFrank, Melanie, et Patricia Kahlbaugh. 2019. « Language Choice Matters: When Profanity Affects How People Are Judged ». *Journal of Language and Social Psychology* 38 (1) : 126-41. <https://doi.org/10.1177/0261927X18758143>.
- Derive, Jean, et Marie-Jo Derive. 2004. « Processus de création et valeur d’emploi des insultes en français populaire de Côte-d’Ivoire ». *Langue française* 144 (1) : 13-34. <https://doi.org/10.3406/lfr.2004.6805>.
- Dewaele, Jean Marc. 2016. « Thirty shades of offensiveness: L1 and LX English users’ understanding, perception and self-reported use of negative emotion-laden words ». *Journal of Pragmatics* 94 (mars) : 112-27. <https://doi.org/10.1016/j.pragma.2016.01.009>.
- Dewaele, Jean-Marc. 2004. « Blistering barnacles! What language do multilinguals swear in?! » *Estudios de Sociolinguística* 5 (1) : 83-105.

- Dewaele, Marc. 2011. « Variation in self-reported frequency of use of french swearwords among L2 and L3 users of French ». Dans *Le français en contact: Hommages à Raymond Mougeon*. Sous la direction de F Martineau et T Nadasdi, 87-115. Presses de l'Université Laval.
- Dostie, Gaétane. 2015. « Le Corpus de français parlé au Québec (CFPQ) et la langue des conversations familières : Exemple de mise à profit des données à partir d'un examen lexico-sémantique de la séquence je sais pas ». *Corpus*, janvier.
- Drange, Eli-Marie Danbolt, Ingrid Kristine Hasund, et Anna-Brita Stenström. 2014. « "Your mum! "Teenagers' swearing by mother in English, Spanish and Norwegian ». *International Journal of Corpus Linguistics* 19 (1) : 29-59. <https://doi.org/10.1075/ijcl.19.1.02dra>.
- Drescher, Martina. 2007. « Sacres québécois et jurons français: Vers une pragmaticalisation des fonctions communicatives? » *Français du Canada - Français de France VIII*. <https://doi.org/10.1515/9783110231045.177>.
- Flesch, Marie. 2022. « Pratiques langagières informelles des femmes et des hommes en ligne : étude quantitative d'un corpus de commentaires publiés sur le site internet Reddit ». *SHS Web of Conferences* 138 : 12007. <https://doi.org/10.1051/shsconf/202213812007>.
- Fodjo, Kouakou Kouman. 2021. « Les insultes “mots-doux” dans les échanges conversationnels dans l'espace public en Côte d'Ivoire ». *Multilinguales*, n° 15 (juin). <https://doi.org/10.4000/multilinguales.6309>.
- Gadet, Françoise. 2021. « Corpus ». *Langage et société* Hors série (HS1) : 73-79. <https://doi.org/10.3917/lhs.hs01.0074>.
- Giordano, Frank, et Manhattan College. 2016. « The Relationship between Profanity and Intelligence ». *Yale Review of Undergraduate Research in Psychology*.
- Guiraud, Pierre. 1975. *Les gros mots*. Presses universitaires de France.
- Harris, Catherine L., Ayşe Ayçiçeği, et Jean Berko Gleason. 2003. « Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language ». *Applied Psycholinguistics* 24 (4) : 561-79. <https://doi.org/10.1017/S0142716403000286>.
- Hirsch, Robert. 1985. « Taxonomies of Swearing ». Dans *Perspectives on Swearing*. , 37-59.
- Houle, René. 2018. « L'évolution des populations de langue maternelle au Canada, de 1901 à 2016 ». Québec .
- Huston, Nancy. 1980. *Dire et interdire: éléments de jurologie*. Paris : Payot.

- Jamet, Denis. 2010. « Historique et procédés linguistiques de l'euphémisme ». Dans *Empreintes de l'euphémisme: Tours et détours*. Sous la direction de Denis Jamet et Manuel Jobert, 31-49. Paris : L'Harmattan. <https://hal.science/hal-01395558>.
- Jay, Kristin L, et Timothy B Jay. 2015. « Taboo word fluency and knowledge of slurs and general pejoratives: deconstructing the poverty-of-vocabulary myth ». *Language Sciences* 52 : 251-59. <https://doi.org/https://doi.org/10.1016/j.langsci.2014.12.003>.
- Jay, Timothy. 2009. « The Utility and Ubiquity of Taboo Words ». *Perspectives on psychological science* 4 (2) : 153-61.
- Jay, Timothy, et Kristin Janschewitz. 2008. « The pragmatics of swearing ». *Journal of Politeness Research* 4 (2) : 267-88. <https://doi.org/10.1515/jplr.2008.013>.
- Jing-Schmidt, Zhuo. 2021. « Euphemism ». Dans *Handbook of Pragmatics Online*. Sous la direction de Jan-Ola Östman et Jef Verschueren, 24 : 125-45. John Benjamins Publishing Company. <https://doi.org/10.1075/hop.24.eup1>.
- Kapoor, Hansika. 2016. « Swears in Context: The Difference Between Casual and Abusive Swearing ». *Journal of Psycholinguistic Research* 45 (2) : 259-74. <https://doi.org/10.1007/s10936-014-9345-z>.
- Kasparian, Sylvia, et Pierre M. Gérin. 2012. « Une forme de purification de la langue : étude des jurons et des gros mots chez des minoritaires francophones, le cas des Acadiens ». *Francophonies d'Amérique*, n° 19 (octobre) : 125-38. <https://doi.org/10.7202/1005314ar>.
- Kellert, Olga. 2023. « Probing sociodemographic influence on code-switching and language choice in Quebec with geolocation of tweets ». *Frontiers in Psychology* 14. <https://doi.org/10.3389/fpsyg.2023.1137038>.
- Kor, Di Yanna. 2017. « « ... la merdre n'était pas mauvaise » : Ubu roi-une pièce pour cinq sens ». *Itinera* 13 : 137-48. <http://gallica.bnf.fr>.
- Labov, William. 1962. « The social history of a sound change on the island of Martha's Vineyard ». Massachusetts : Columbia University.
- . 1972. « The Social Stratification of (r) in New York City Department Stores ». Dans *Sociolinguistic Patterns*. Sous la direction de William Labov, 168-77. Philadelphia : University of Pennsylvania Press.
- Lakoff, R T, et M Bucholtz. 2004. *Language and Woman's Place: Text and Commentaries*. Studies in Language and Gender. Oxford University Press.
- Lakoff, Robin. 1973. « Language and woman's place ». *Language in Society* 2 (1) : 45-80.

- Larchet, Keltoume. 2017. « Les injures à caractère homophobe: Exploitation des enquêtes, Cadre de vie et sécurité ». *Grand Angles*. Vol. 42.
- Lévy, Alexandre. 2017. « Insulte et traitement du féminin. Étude psychanalytique ». *Bulletin de psychologie* Numéro 550 (4) : 291-99. <https://doi.org/10.3917/bupsy.550.0291>.
- Maurais, Jacques. 2008. « Les Québécois et la norme : l'évaluation par les Québécois de leurs usages linguistiques ». Québec : Office québécois de la langue française.
- McEnery, Anthony, Richard Xiao, et Yukio Tono. 2006. *Corpus-Based Language Studies An Advanced Resource Book*. New York : Routledge.
- McEnery, Anthony, et Zhonghua Xiao. 2004. « Swearing in modern British English: The case of fuck in the BNC ». *Language and Literature* 13 (3) : 235-68. <https://doi.org/10.1177/0963947004044873>.
- McEnery, T, et A Hardie. 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge Textbooks in Linguistics. Cambridge University Press.
- McEnery, Tony. 2006. *Swearing in English Bad Language, Purity and Power from 1586 to the Present*. Routledge.
- McEnery, Tony, et Andrew Wilson. 2005. « Early corpus linguistics and the Chomskyan revolution ». Dans *Corpus Linguistics: An Introduction*. , 2nd éd., 1-27. Edinburgh : Edinburgh University Press.
- Ngula, Richmond Sadick. 2018. « Corpus Methods in Language Studies ». Dans *Perspectives on Conducting and Reporting Research in the Humanities*. , 205-23. University of Cape Coast Press.
- Oakes, Leigh. 2004. « French: a language for everyone in Québec? » *Nations and Nationalism* 10 (4) : 539-58.
- Partington, Alan. 2014. « Mind the gaps ». *International Journal of Corpus Linguistics* 19 (1) : 118-46. <https://doi.org/10.1075/ijcl.19.1.05par>.
- Patrick, G. T. W. 1901. « The psychology of profanity ». *Psychological review* 8 (2) : 113-27.
- Perea, François. 2011. « Les gros mots, paradoxes entre subversion et intégration ». *La lettre de l'enfance et de l'adolescence* 83-84 (1) : 53. <https://doi.org/10.3917/lett.083.0053>.
- Pirou, Jean-Pierre. 1992. « La Situation de la langue française au Québec ». *Cahiers de l'Association internationale des études françaises* 44 (1) : 25-38. <https://doi.org/10.3406/caief.1992.1776>.

- Poplack, Shana. 2021. « Le corpus comme portail pour l'étude de la variation (socio)linguistique* ». *Corpus*, n° 22 (janvier) : 1-16. <https://doi.org/10.4000/corpus.5422>.
- Rapport sur l'évolution de la situation linguistique au Québec*. 2019. Québec : Office québécois de la langue française.
- Rouayrenc, Catherine. 1998. *Les gros mots. Que sais-je ?* Vol. 3. Paris cedex 14 : Presses Universitaires de France.
- Sankoff, Gillian. 2019. « Language change across the lifespan: Three trajectory types ». *Language* 95 (2) : 197-230.
- Sankoff, Gillian, et Hélène Blondeau. 2007. « Language Change across the Lifespan: /r/ in Montreal French ». *Language* 83 (3) : 560-88.
- Sapolsky, Barry S., Daniel M. Shafer, et Barbara K. Kaye. 2011. « Rating offensive words in three television program contexts ». *Mass Communication and Society* 14 (1) : 45-70. <https://doi.org/10.1080/15205430903359693>.
- Siyanova-Chanturia, Anna, et Stefania Spina. 2015. « Investigation of Native Speaker and Second Language Learner Intuition of Collocation Frequency ». *Language Learning* 65 (3) : 533-62. <https://doi.org/10.1111/lang.12125>.
- Solan, Lawrence M, et Tammy Gales. 2017. « Corpus Linguistics as a Tool in Legal Interpretation ». *BYU Law Review* 6 (5) : 1311-59.
- Stapleton, Karyn. 2010. « Swearing ». Dans *Interpersonal Pragmatics*. , 6 : 289-307.
- . 2020. « Swearing and perceptions of the speaker: A discursive approach ». *Journal of Pragmatics* 170 (décembre) : 381-95. <https://doi.org/10.1016/j.pragma.2020.09.001>.
- Stapleton, Karyn, Kristy Beers Fägersten, Richard Stephens, et Catherine Loveday. 2022. « The power of swearing: What we know and what we don't ». *Lingua* 277 (octobre). <https://doi.org/10.1016/j.lingua.2022.103406>.
- Stefanowitsch, Anatol. 2009. *Corpus Linguistics: A Guide to the Methodology*.
- Suganob-Nicolau, Maria Fe. 2016. « Swear words among young learners: A case study of the elementary students ». *Indonesian Journal of English Language Teaching* 11 (2) : 117-32.
- Ten, Angela, et Ervina CM Simatupang. 2020. « The Use of Swear Words in The Corpus of Contemporary American English (COCA): Sociolinguistics Study ». *International Journal of Psychosocial Rehabilitation* 24 (2) : 7678-87.
- Vincent, Diane. 1984. « Le sacre au Québec: transgression d'un ordre religieux ou social? » *Culture* 4 (2) : 55-61.

- Wagner, Suzanne Evans. 2012. « Age Grading in Sociolinguistic Theory ». *Linguistics and Language Compass* 6 (6) : 371-82. <https://doi.org/10.1002/lnc3.343>.
- Widlak, Stanislas. 1965. « L'interdiction linguistique en français d'aujourd'hui ». *Revue belge de philologie et d'histoire* 43 (3) : 932-45. <https://doi.org/10.3406/rbph.1965.2587>.
- Yaguello, Marina. 1987. *Les mots et les femmes : essai d'approche socio-linguistique de la condition féminine / Marina Yaguello*. Prismes. Paris : Payot.

Support numérique

- « Dico en ligne Le Robert ». s. d. Consulté le 24 août 2023.
- « Dictionnaire Usito ». s. d. Université de Sherbrooke. Consulté le 24 août 2023.
- Gosseline, Jacques. s. d. « La législation linguistique du Québec ». Le dictionnaire en ligne Usito. Consulté le 24 août 2023.
https://usito.usherbrooke.ca/articles/th%C3%A9matiques/gosselin_maillet_1.
- Mair, Christian. 2004. « Corpus linguistics and grammaticalisation theory: Statistics, frequencies, and beyond ». <https://www.researchgate.net/publication/287724976>. Consulté le 24 août 2023.

Liste des illustrations

| | |
|--|----|
| Figure 1 Diagrammes du processus de communication de Bühler et du diagramme du juron proposé par (Huston 1980, 16-17)..... | 4 |
| Figure 2 Répartition de la population au Canada, au Québec et au Canada sans le Québec, de 1901 à 2011, maîtrisant l'anglais et le français | 10 |
| Figure 3 Illustration des concordances avant et après la suppression des conventions..... | 15 |
| Figure 4 Fréquence des jurons dans le CFPQ..... | 16 |
| Figure 5 Répartition des euphémismes comparée au total des occurrences..... | 22 |
| Figure 6 Modèle de changement chez l'individu et dans la communauté..... | 29 |
| | |
| Tableau 1 Distribution des participants en fonction de l'âge, du sexe et du niveau d'éducation | 13 |
| Tableau 2 Requêtes de recherche | 14 |
| Tableau 3 Les jurons et leurs variantes morphologiques trouvés dans le CFPQ | 17 |
| Tableau 4 Distribution des jurons selon le sexe | 18 |
| Tableau 5 Répartition des jurons en fonction du sexe du groupe..... | 19 |
| Tableau 6 Répartition des jurons par âge | 20 |
| Tableau 7 Distribution du niveau d'éducation | 21 |
| Tableau 8 Proportion d'euphémismes dans le nombre total d'occurrences de jurons par sexe | 22 |
| Tableau 9 Proportion d'euphémismes dans le nombre total d'occurrences de jurons par âge | 23 |