

« *So* » ... *fait que*... le français québécois vernaculaire se diversifie-t-il?
Une étude sociolinguistique de l'alternance entre *fait que* et « *so* » dans le
corpus « RapKeb21 »

Emily Brooke Leavitt
Université de Sherbrooke / Canada
emily.leavitt2@usherbrooke.ca

Reçu : 02/10/2023,

Accepté: 31/12/2023,

Publié: 31/12/2023

"So" ... *Fait Que*... Is Quebec French Vernacular Undergoing Diversification? A
Sociolinguistic Study of the Alternation Between "Fait Que" And "So" in the
"Rapkeb21" Corpus

Abstract:

The aim of this study was to test the hypothesis of an ongoing change within the contemporary urban vernacular in Quebec. We observed variation between (ça) *fait (que)* and "so" as markers of consequence. The social conditioning of this variable reveals a possible case of change in progress, because female Quebec rap artists show increased use of "so", and because the youngest artists produce more than their counterparts. Our results suggest that both variants have become indexicalized, with (ça) *fait (que)* indexing an identity rooted in Quebec origins and "so" indexing a local yet also global rap identity.

Keywords: sociolinguistics, language contact, Quebec rap, Quebec French, contemporary urban vernacular

Résumé :

L'objectif de cette étude a été de vérifier l'hypothèse d'un changement en cours au sein du vernaculaire urbain contemporain au Québec. Nous avons attesté de la variation entre (ça) *fait (que)* et « *so* » en fonction de marqueur de conséquence. Le conditionnement social de cette variable révèle un cas possible de changement en cours parce que les artistes rap québécois de genre féminin font un usage élevé de « *so* » et ce sont les artistes les plus jeunes qui le produisent davantage. Nos résultats suggèrent que les deux variantes sont indexicalisées, (ça) *fait (que)* indexant une identité liée aux origines québécoises et « *so* » renvoyant plutôt à une identité à la fois locale et mondialisée.

Mots-clés : sociolinguistique, contact des langues, rap québécois, français québécois, vernaculaire urbain contemporain

Introduction

Le Québec représente une situation de contact des langues et un site prometteur à l'étude sociolinguistique de la variation au sein du français en Amérique du Nord. Les études antérieures ont amplement montré le caractère distinct de la variété de français québécoise tant au niveau de son évolution historique et linguistique que des normes et contraintes sociolinguistiques en son sein. La présente étude aborde le problème de la diversification du français vernaculaire au Québec et à Montréal en particulier. Cette étude s'inscrit à la fois dans la tradition des études sociolinguistiques montréalaises et dans la veine de recherches traitant des divers vernaculaires urbains contemporains, comme celles portant sur les corpus « Multicultural Paris French » (Gadet et Guerin, 2016) et « Multicultural London English » (Cheshire et al, 2013).

1. Cadre conceptuel

L'objectif principal de la présente étude est de vérifier la théorie exposée par Blondeau et Tremblay (Blondeau, 2018, 2020; Blondeau et Tremblay, 2016, 2022a) de la diversification en cours du français vernaculaire au Québec et, plus précisément, en contexte urbain à Montréal. Nous nous intéressons à la production langagière d'artistes du genre rap québécois dont l'un des caractéristiques les plus saillants est sa nature vernaculaire. Nous proposons une analyse de la variation et des contraintes sociolinguistiques opérant sur elle quant à l'alternance entre deux variantes vernaculaires, (*ça fait (que)*), du français, et « *so* », de l'anglais, au sein d'un corpus de paroles de chansons rap québécoises, le corpus « RapKeb21 » (Leavitt, E. B. 2022).

1.1 Question de recherche

Notre question de recherche est si l'alternance entre variantes vernaculaires (*ça fait (que)* et « *so* » s'avère être contrainte par des facteurs sociaux renvoyant à l'existence de deux pôles normatifs au Québec, l'un à Montréal et l'autre délocalisé au sein de la communauté linguistique québécoise hors de Montréal, prévue par la théorie de diversification. Cette variable a été choisie en raison de la récence relative et la variante « *so* », n'ayant jamais été attestée au Québec dans les usages informels spontanés produits lors d'entretiens sociolinguistiques (Blondeau et al, 2019; Blondeau et Tremblay, 2022b), mais ayant récemment fait surface au sein du sous-corpus québécois francophone de textos Texto4Science (Langlais et Drouin, 2012) comme le constatent Blondeau et Tremblay (2022a).

1.2 Hypothèses

Notre hypothèse globale est que les deux variantes vernaculaires (*ça fait (que)*) et « *so* » se révéleront exister dans la production lyrique des artistes de rap québécois, mais faute de données sur leur alternance, nous ne posons pas d’hypothèse au sujet de leur ampleur relative.

Plus précisément, nous nous attendons à ce que la variante (*ça fait (que)*) soit plus présente chez les artistes de Montréal en raison de la hausse constatée à Montréal depuis 2012 de sa forme phonologique /fɛk/ aux dépens des autres variantes (Blondeau et al, 2019; Blondeau et Tremblay, 2022b; Blondeau et al, 2021).

Selon les travaux récents, la variante « *so* » s’avère catégoriquement absente des données d’entretiens sociolinguistiques semi-dirigés menés à Montréal selon la tradition labovienne (Labov, 1984) (Blondeau et al, 2019; Blondeau et Tremblay, 2022b; Blondeau et al, 2021), or elle a récemment été attestée par Blondeau et Tremblay (2022a) au sein du sous-corpus québécois francophone de textos, Texto4Science (Langlais et Drouin, 2012). Pour reprendre les mots de ces chercheuses, « cette présence de *so*, dont l’emploi est principalement attesté à l’oral dans des communautés minoritaires bilingues comme en Ontario et en Acadie, tranche avec ce qui a été observé dans les études antérieures sur le français montréalais (Blondeau et al., 2019) » (Blondeau et Tremblay, 135 :2022).

Nous estimons que c’est la nature vernaculaire de la variante « *so* » qui explique sa présence dans les textos mais aussi son absence dans les entretiens. Pour cette raison, nous posons l’hypothèse que la nature vernaculaire de la modalité de communication que représente l’expression artistique rap s’avérera être propice à la production de « *so* ». En revanche, cette hypothèse étant basée sur un constat qualitatif de la présence de « *so* » dans le corpus de textos (Blondeau et Tremblay, 2022a) dont on ignore la fréquence, nous attendons à découvrir si elle se présente dans notre corpus, à quelle fréquence et quels facteurs influencent sa production.

1.3 Le français vernaculaire au Québec et à Montréal

De nombreuses études récentes contribuent à décrire le vernaculaire français émergent à Montréal de manière diachronique à travers la comparaison du sous-corpus Hochelaga-Maisonneuve 2012 (Martineau et Seguin, 2016) aux corpus Montréal 1971 (Sankoff et al, 1976), 1984 (Thibault et Vincent, 1990) et 1995 (Vincent et al, 1995). En plus de la hausse de (*ça fait (que)*) à Montréal (Blondeau et al, 2019; Blondeau et Tremblay, 2022b; Blondeau et al, 2021), plusieurs études démontrent effectivement de la divergence entre le vernaculaire montréalais émergent et le vernaculaire délocalisé au sein de la communauté linguistique

québécoise hors de Montréal. Certaines variantes typiquement associées au vernaculaire québécois traditionnel sont actuellement en régression au sein du français montréalais vernaculaire : la contraction de la préposition *dans* et un déterminant subséquent (Blondeau et al, 10-11 :2021), les formes pronominales complexes (*nous-autres, eux-autres, vous-autres*) (Blondeau et al, 2022) et de la variante *m'as* (1^e personne du singulier du verbe *aller* comme auxiliaire au futur périphrastique) en faveur des variantes *j'vas* et *j'vais* (toutes deux plus formelles, la seconde davantage) (Blondeau et al, 2016). À l'opposé, une variante en recul plus largement au Québec s'avère être à la hausse à Montréal, l'emploi du futur synthétique (Blondeau, 2006).

Le multilinguisme est devenu l'un des caractéristiques les plus saillants du français montréalais vernaculaire. C'est ce que révèle notamment l'étude descriptive de Blondeau et Tremblay (2022a) basée sur un sous-corpus québécois de messages textos *Texto4Science* (Langlais et Drouin, 2012). La situation de contact des langues particulièrement intenses à Montréal se manifeste dans les usages linguistiques par des phénomènes comme le *codeswitching* (ou l'alternance codique) (Lamarre, 2013, 2014; Low et Sarkar, 2012) entre le français, l'anglais et de diverses langues d'immigration telles l'arabe, le créole haïtien et l'espagnol, l'emprunt à l'anglais et son intégration au français (Poplack, 2018; Friesner, 2009), et le *crossing* (ou le franchissement) (Rampton, 1995) aux langues dont la personne n'est pas un locuteur (Sarkar, 2008). Les répertoires linguistiques de Montréalais(es) étant plurilingues, les identités linguistiques revendiquées au sein de la communauté linguistique montréalaises sont aussi plurielles, hybrides et fluides. À travers l'indexicalité, les pratiques multilingues sont porteuses de sens social et servent d'outils lors de l'expression d'identités nuancées et liées, éventuellement de manière simultanée, au réseau social, à la participation à une certaine communauté de pratique et au milieu de socialisation du locuteur.

2. Cadre théorique : le genre rap

Si le rap, également appelé par analogie hip-hop, est relativement nouveau, il jouit aujourd'hui d'une popularité considérable et fait partie intégrale de la culture populaire, surtout de la jeunesse. *Rap* est effectivement un hyperonyme qui désigne à la fois un genre musical et une tradition plus large qui dépasse le cadre d'un genre musical et regroupe de différentes modalités d'interprétation, ses hyponymes, – la danse break, l'art du DJ, le rap et la composition lyrique (Androutsopoulos, 43-44 :2009). C'est à ce premier, le genre *rap*, dont nous faisons référence.

Le rap est un genre musical qui a été inventé aux années 1970 dans les rues du Bronx, à New York, aux États-Unis, par des jeunes qui cherchaient un moyen d'échapper au sort trop souvent réservé aux jeunes marginaux, à savoir la participation à la vie des gangs de rue. À cette époque, de jeunes musiciens se sont

**« So » ... fait que... le français québécois vernaculaire se diversifie-t-il?
Une étude sociolinguistique de l'alternance entre fait que et « so » dans le corpus «
RapKeb21 »**

réunis autour de la composition musicale de manière collective, partageant du matériel ainsi que leurs compétences techniques, leurs répertoires musicaux et leurs innovations stylistiques. Leur trame de fond était le rythme, qu'ils créaient à travers le découpage et le mixage de segments de chansons axés sur la percussion, sur lequel ils transposaient des paroles rapides, rimées, rythmées et rhétoriquement créatives.

La composition lyrique rap est largement destinée à faire entendre des voix autrement marginalisées à travers la représentation de réalités interdites, négligées ou sous-représentées dans d'autres formes d'expression artistique. Certains thèmes sont mis en évidence, notamment : les injustices sociales, les relations de famille, les divers aspects de la vie de la rue, les divers aspects de la jeunesse, l'adolescence et l'âge adulte, les perspectives d'avenir, la construction et l'expression identitaire particulièrement par la mise en relation et la négociation de relations entre sociotypes et la vantardise des compétences en matière de rap.

Certaines règles formant un code de conduite opèrent au sein du rap et servent à préserver son caractère démocratique profondément enraciné. Ce n'est qu'en respectant ces règles qu'un artiste peut se permettre de participer au « rap game, » en termes rap. La première renvoie au respect d'authenticité, ou en termes rap, « keepin' it real » (Lee, 2010). Puisque le réalisme est au cœur du genre, la légitimité de l'artiste et donc sa viabilité d'être pris au sérieux par les autres sont soumises à l'évaluation par l'autrui. La seconde règle de conduite renvoie à l'acceptation d'appartenir à une collectivité, désignée la « nation mondiale du rap » (Alim, 3 :2009). De la même manière que les inventeurs de rap participaient dans une entreprise collective ouverte à toutes parties intéressées à faire entendre leur voix, tout artiste rap doit être prêt à coexister dans un espace culturel parmi d'autres artistes qui peuvent éventuellement représenter des alliés, des concurrents ou des rivaux, voire les trois à la fois. La troisième règle renvoie à la reconnaissance d'une mémoire collective d'efforts des autres artistes passés, présents et futurs tels, notamment, les fondateurs du genre. La dernière est d'adopter le rap comme « a whole philosophy of life, an ethos that involves clothes, a style of talk and walk, a political attitude, and often a philosophical posture of asking hard questions and critically challenging established views and values » (Shusterman, 61 :2005).

Dans cette optique, la représentation de la vie réelle et des réalités modernes supplante les conventions sociétales de comportements langagiers et stylistiques normatifs, ce qui signifie que toute infraction à ces normes sociétales est autorisée. Pour cette raison, le rap se caractérise par un langage rude et vernaculaire, fertile aux usages non standard tels l'argot, les jurons et la vulgarité, parmi d'autres. C'est

effectivement la richesse de ces pratiques linguistiques non standard qui explique l'intérêt des chercheurs en linguistique pour l'étude du rap.

2.1 Le rap québécois

Tout comme le genre rap jouit d'une popularité considérable, le rap québécois en jouit d'une aussi. Cette popularité se reflète sur une main par l'engouement des médias concernant l'ascension fulgurante vers la célébrité de certains artistes (tels Lost, Tizzo x Shreez et Yes)¹, l'incarcération d'autres (Lost, VT et White-B)², et le langage de certains groupes en particulier (Dead Obies, Maky Lavender et Muzion)³. Sur une autre, elle est palpable sur les ondes radios où sont diffusées des chansons de certains artistes (tels Loud, FouKi et Koriass)⁴, ainsi que sur les ondes télévisées, notamment, par la diffusion sur TéléQuébec de l'émission télévisée *La fin des faibles*⁵ (un concours de rappeurs en herbe dont les juges sont Koriass, Souldia et Sarahmée) et par la diffusion annuelle de la cérémonie de remise des palmarès ADISQ et l'annonce des nominations aux prix pour l'année 2023⁶. La popularité du rap québécois se manifeste également sur la scène par certains artistes qui font des tournées de spectacles (Loud, Monk.E et Sarahmée)⁷ et par ceux qui jouent aux festivals musicaux, et plus particulièrement le festival Les Francos de Montréal (Lost, Mike Shabb et Rowjay)⁸.

2.2 L'étude sociolinguistique du rap québécois

Les premières études sociolinguistiques menées sur le rap québécois sont issues du projet « Hip Hop Language and Youth Culture in Quebec: Multilingual Identity Practices In and Out of Classrooms » mené entre 2002 et 2014 par Mela Sarkar et ses collègues à McGill University (Low et Sarkar, 2012, 2014 ; Low et al, 2009 ; Sarkar, 2008, 2009 ; Sarkar et Allen, 2007 ; Sarkar et Low, 2012 ; Sarkar et Winer, 2006 ; Sarkar et al, 2005). Ces travaux représentent un ensemble d'analyses de discours de paroles de chansons rap québécoises abordées d'une perspective ethnographique qualitative. Somme toute, leurs analyses portent sur 12 albums parus de 1997 à 2004 ainsi que 10 entrevues d'artistes et six entrevues de jeunes amateurs de rap.

Tout d'abord, les résultats de leurs analyses mettent en avant « une grande aisance à mélanger plusieurs influences linguistiques, autant dans les paroles

¹ Voir (Couturier, 2022; Jbeili, 2023; Leblanc, 2023b).

² Voir (Bordeleau, 2022; Leblanc, 2023a; Perron, 2022).

³ Voir (Bilefsky, 2020; Leavitt, S., 2022; Renaud, 2021).

⁴ Voir (Blanchet, 2021; Côté, 2018; Lauzon, 2018).

⁵ Voir (Télé-Québec, 2022).

⁶ Voir (Houle et Bouchard, 2017; PalmarèsADISQ, 2022, 2023).

⁷ Voir (Agence QMI, 2022; Alexy, 2022; Étienne, 2020).

⁸ Voir (Leblanc, 2023a; Renaud, 2023).

**« So » ... fait que... le français québécois vernaculaire se diversifie-t-il?
Une étude sociolinguistique de l'alternance entre fait que et « so » dans le corpus «
RapKeb21 »**

composées (donc préalablement réfléchies et assidûment retravaillées avant de paraître sur disque) que lors de conversations spontanées » (Low et Sarkar, 31 :2012). Les variétés de langues concernées sont le français québécois, l'anglais de l'Amérique du Nord, l'anglais vernaculaire Africain-Américain (AAVE), l'anglais rap mondial (un champ lexical de mots en anglais associé au rap mondial), le créole haïtienne, le créole jamaïcain, le créole trinitadien, l'espagnol et l'arabe (Sarkar, 2009).

Des entrevues avec les jeunes amateurs, Low et Sarkar (2012) constatent dans la production d'adolescents de diverses origines du *crossing* (Rampton, 1995) en créole haïtien et en créole jamaïcain. Toutefois, chez certains artistes pour qui ces langues font véritablement partie du répertoire, il ne s'agit pas de *crossing*, plutôt d'autres types de mélanges. Voici un extrait cité par Low et Sarkar (2012)⁹ :

Show me respect, j'suis la true ill nana / Pourquoi t'es venu, si tu *front sou kote*? / Fais pas ton **mean**, j'vois ton *bounda* sauter / Si tu sais pas danser, qu'est-ce que t'as à te moquer? / Dis-le, t'aimes mes **moves**, pas vrai, t'es choqué? / Hey! Mais qu'est-ce ta main fait là? Go away! / *Neg pa lave, pafume...* No way! (J.Kyll de Muzion, « Lounge with us », *Mentalité mounne morne*, 1999) (p. 32).

Concernant ces pratiques plurilingues, les auteures concluent que, « la production artistique est ici un baromètre de transformation sociale [...] qui témoigne d'un changement d'attitude et du développement de nouvelles compétences qui concernent le plurilinguisme » (Low et Sarkar, 35 :2012).

Ce plurilinguisme remplit, selon elles, plusieurs fonctions socioculturelles. La principale est la promotion de la diversité d'identités revendiquées surtout par la nouvelle génération de jeunes multilingues au Québec issus de la migration, la mondialisation et l'urbanisation (Sarkar et Allen, 2007). Les autres comprennent la résistance, la revendication de droits linguistiques, la déstabilisation de la hiérarchie des langues dominantes et le questionnement de ce qui constitue aujourd'hui le fait d'être Québécois(e) (Sarkar et al, 2005, 2007, 2008, 2009).

Quant à nous, les contributions les plus significatives des travaux de Sarkar et ses collègues sont : 1) le constat que les pratiques langagières au sein du rap québécois sont représentatives de « la diversité linguistique et culturelle que l'on

⁹ Ici, la mise en forme indique la langue source : le français québécois standard (romain maigre), le français québécois non standard (romain gras), le français européen (romain gras et souligné), l'anglais nord-américain standard (romain souligné), l'anglais vernaculaire Africain-Américain (italique maigre) et le créole haïtien (italique gras).

retrouve dans les écoles et dans la communauté de jeunes adultes » (Low et Sarkar, 36 :2012); 2) la mise en place d'une veine de recherche dédiée à l'analyse sociolinguistique du multilinguisme dans le rap québécois; et, de leur dernière étude, 3) « la valeur du discours rap en tant que matériel pédagogique et didactique dans la mesure où il peut servir à sensibiliser les enseignants à la culture populaire des jeunes urbains, et à représenter le capital socioculturel des jeunes » (Leavitt, E.B., 35 :2022).

Le relais des recherches sociolinguistiques sur le rap québécois a été pris par l'auteure de la présente étude lors de son analyse de la prosodie dans deux chansons de rap québécoises (Leavitt, 2021). Les résultats révèlent d'une part que les schémas d'accent et d'intonation dans le rap suivent largement les mêmes tendances que l'on constate en français québécois oral. De l'autre, le rap s'avère original dans son utilisation de procédés rythmiques tels le trémolo et la symétrie mélodique entre groupes rythmiques ainsi que dans la manière intentionnelle dont s'effectue l'alternance entre le français et l'anglais pour aboutir à la perturbation, la satisfaction ou l'amplification de l'effet prosodique des accents tonique, initial et expressif (Leavitt, 2021).

La prochaine étude de Leavitt (Leavitt, E.B., 2022) fut son mémoire de maîtrise rédigé sous la direction du professeur Davy Bigot de l'Université Concordia. Il s'agit d'une analyse sociolinguistique variationniste de trois variables (l'alternance entre *alors, donc, (ça) fait (que)* et « *so* »; l'alternance entre *j'vais, j'vas* et *m'as*; la neutralisation de *tout* et *tous* en [tʷt]) au sein du mégacorpus « RapKeb21 » de paroles de chansons rap québécoises créé et publié en libre accès par l'auteure (Leavitt, E.B., 2022). Ses résultats montrent une divergence significative entre la production d'artistes « ayant pour modèle normatif une variété de français montréalaise distincte » et ceux « ayant pour modèle normatif le français québécois » (Leavitt, E.B., iii :2022). Ces premiers privilégient la variante empruntée de l'anglais « *so* », tandis que les derniers favorisent *(ça) fait (que)* et la variante *j'vas* pour la deuxième variable, celles-ci étant les variantes les plus typiquement associées au français québécois traditionnel. Par ailleurs, les résultats de Leavitt (Leavitt, E.B., 2022) révèlent une systématisme des contraintes linguistiques sur la production rap qui s'avère être semblable à ce que l'on atteste dans les études antérieures sur les pratiques langagières (hors du rap) au sein de la communauté québécoise plus généralement.

3. Méthodologie

Nous avons effectué une analyse variationniste de l'alternance entre *(ça) fait (que)* et « *so* », notre variable dépendante, et des facteurs sociaux et linguistiques (variables indépendantes) qui la conditionnent au sein du corpus « RapKeb21 » (Leavitt, E.B., 2022), un corpus de paroles de chansons rap québécoises. Nous avons d'abord

**« So » ... fait que... le français québécois vernaculaire se diversifie-t-il?
Une étude sociolinguistique de l'alternance entre fait que et « so » dans le corpus «
RapKeb21 »**

identifié 300 occurrences¹⁰ de *(ça) fait (que)* (sous toutes ses formes, *ça fait que*, *ça fait* et *fait que*) et 927 occurrences de « so » au sein du corpus (N=1227). Nous avons extrait ces occurrences et nous avons par la suite codé chacune en fonction de l'artiste l'ayant produite et de quatre facteurs sociaux (âge, genre, lieu de naissance et origines géographiques) et de deux facteurs sociopsychologiques (identité linguistique et identité culturelle) concernant l'artiste, ainsi qu'en fonction de l'année de parution de la chanson dans laquelle elle figure et de la fonction grammaticale qu'elle sert (marqueur de conséquence ou marqueur de discours). Nous avons ensuite soumis les 1227 occurrences codées à une analyse multivariée par régression logistique à effets mixtes afin d'évaluer le conditionnement de l'alternance entre *(ça) fait (que)* et « so ».

3.1 Données

Les données dont il est question viennent du corpus « RapKeb21 » (Leavitt, E.B., 2022) que nous avons créé entre 2020 et 2021. Il représente les œuvres complètes d'un échantillon de 80 artistes de rap québécois, dont 37 artistes solitaires et 43 artistes appartenant à un groupe composé de deux membres ou plus, qui en compte 13 au total. Les artistes de l'échantillon ont été sélectionnés en fonction de critères d'appartenance à la communauté rap québécoise et de critères basés sur la popularité et la visibilité de l'artiste, la prééminence de l'artiste au sein du mouvement rap local et de la culture populaire québécoise, et la représentativité de la communauté québécoise plus large qui sont expliquées plus en détail dans Leavitt (Leavitt, E.B., 71-72 :2022).

Les données du corpus correspondent à un total de 2965 chansons. Chacune est représentée par trois fichiers : un fichier audio *.wav*, un fichier texte *.txt* de ses paroles, et un fichier *.TextGrid* composé des horodatages qui signalent l'emplacement temporel de chacun des mots de paroles dans l'audio de la chanson. Ainsi, le corpus se compose de 8895 fichiers divisés entre trois séries selon le format. La création de chacune des séries est détaillée dans Leavitt (Leavitt, E.B., 74 :2022). La série textuelle qui représente la version orthographique des paroles compte un total de 1 454 425 mots.

Aux fins de la présente étude, nous avons effectué une recherche par logiciel concordancier de toute occurrence des formes de *(ça) fait (que)* et de « so » au sein de l'ensemble du corpus. Nous avons extrait 1227 occurrences au total avec le contexte les précédant et les suivant au sein des paroles. Chacun des occurrences a

¹⁰ Nous utilisons le terme « occurrence » dans la même manière que l'on utilise le terme « token » en anglais.

été codée en fonction de la variable dépendante, voire la variante à laquelle il renvoie, soit *(ça) fait (que)* (n=300), soit « *so* » (n=927). Ensuite, les occurrences ont tous été codés en fonction des variables indépendantes.

3.2 Variables indépendantes

Les facteurs que nous avons traités de variables indépendantes sont de quatre types, soit social, sociopsychologique, linguistique et aléatoire. Nous avons tenu compte de cinq facteurs sociaux : âge, genre, lieu de naissance et origines géographiques de l'artiste et année de parution de la chanson. Les catégories d'âge de l'artiste et d'année de parution de la chanson ont été codé de manière directe comme variables continues, leurs valeurs ayant de 16 à 59 pour la catégorie d'âge et de 1997 jusqu'à 2021 pour la catégorie d'année. La catégorie de genre tient compte d'artistes de genre masculin, de genre féminin et de genre non binaire. Le choix de traiter du genre au lieu du sexe représente une prise de distance par rapport à la tradition destinée à nous permettre de tenir compte des réalités modernes et de constituer une catégorie basée principalement sur des inférences sociales et, dans certains cas, sur des informations biographiques lorsque ces dernières étaient disponibles. Les catégories de lieu de naissance et d'origines géographiques représentent des variables continues, c'est-à-dire qu'elles ont été constituées en fonction de la distribution de l'échantillon. Les valeurs qu'elles regroupent correspondent soit à la ville et la province pour des lieux au Canada, soit au pays pour des lieux hors du Canada. Le lieu de naissance signifie l'endroit précis où l'artiste est né(e) et les origines géographiques signifient l'endroit précis d'où viennent ou où sont nés ses parents. Si l'un des parents vient ou est né au Canada et l'autre hors du Canada, ce sont les origines géographiques du dernier qui sont prises en compte.

Nous avons tenu compte de deux facteurs sociopsychologiques : identité linguistique et identité culturelle. La catégorie d'identité linguistique se divise entre « francophone », « anglophone » et « bilingue ». Pour chacun des artistes, une recherche par concordancier a été effectuée sur l'ensemble des paroles de leurs œuvres complètes afin d'identifier le nombre de chacune des trois désignations en français et en anglais (*francophone*, *anglophone*, *bilingue* et « *bilingual* ») d'identité linguistique revendiquée par l'artiste. Celle qui s'est avérée majoritaire a été attribuée à l'artiste et son code ajouté à toute occurrence produite par l'artiste. Si l'artiste n'a revendiqué aucune de désignations recherchées, nous lui avons attribué un code nul. La catégorie d'identité culturelle renvoie aux différentes combinaisons de quatre désignations d'identité civique locale revendiquées par l'artiste en français ou en anglais : *Montréalais(e)* ou « *Montrealer* » (M), *Québécois(e)* ou « *Quebecker* » (Q), *Canadien(ne)* ou « *Canadian* » (C) et *Américain(e)* ou « *American* » (A). Encore une fois, une recherche par concordancier a été effectuée pour chaque artiste au sein des paroles de ses œuvres complètes, mais cette fois, nous

**« So » ... fait que... le français québécois vernaculaire se diversifie-t-il?
Une étude sociolinguistique de l’alternance entre fait que et « so » dans le corpus «
RapKeb21 »**

lui avons attribué un code correspondant à la combinaison de différentes identités revendiquées à au moins une reprise ou, dans le cas d’absence totale de désignations, un code nul. Les différentes combinaisons attestées sont : ‘MQCA’, ‘MQC’, ‘MQA’, ‘MQ’, ‘MA’, ‘M’, ‘Q’ et nul.

Nous avons pris en compte un seul facteur interne à la langue, soit la fonction grammaticale de l’occurrence. Celle-ci pourrait être soit marqueur de conséquence, soit marqueur de discours. Nous avons consulté chacune des 1227 occurrences et son contexte linguistique et en avons déduit la fonction grammaticale. Voici un exemple des deux fonctions pour chacune des deux variantes.

- 1) *All the places I’ve been, fait que* [marqueur de conséquence] *mes shoes s’abîment.*¹¹
- 2) *Jamais d’mic check, ça fait que* [marqueur de discours] *yes, respect Bas-Canada.*¹²
- 3) *Jamais égalé, so* [marqueur de conséquence] *pourquoi tu imites, huh?*¹³
- 4) *Comme le silence pour faire durer l’suspense, so...* [marqueur de discours].¹⁴

Enfin, nous avons traité l’individu en tant que variable aléatoire afin de contrôler l’effet de la variation individuelle et dégager quels autres effets ont un poids supérieur au seuil de l’individu.

3.3 Analyses

À partir des dix codes attribués à toutes les occurrences, nous avons créé une trame de données (ou « *dataframe* ») de 1227 rangées sur 10 colonnes. Nous avons importé cette trame de données dans RStudio (RStudio Team, 2020), un environnement de programmation dans le langage R (R Core Team, 2021). Nous avons utilisé la trousse (ou « *library* ») *lme4* (Bates et al, 2015). Nous avons soumis les données à une analyse par régression logistique¹⁵ en traitant l’occurrence comme variable dépendante, les facteurs sociaux, sociopsychologiques et interne comme variables indépendantes à effet fixe et l’individu comme variable indépendante à effet aléatoire. Dans ce modèle, les variables catégorielles sont traitées à travers une

¹¹ (Yes McCan, 2016).

¹² (Alaclair Ensemble, 2013).

¹³ (Izzy-S, 2018).

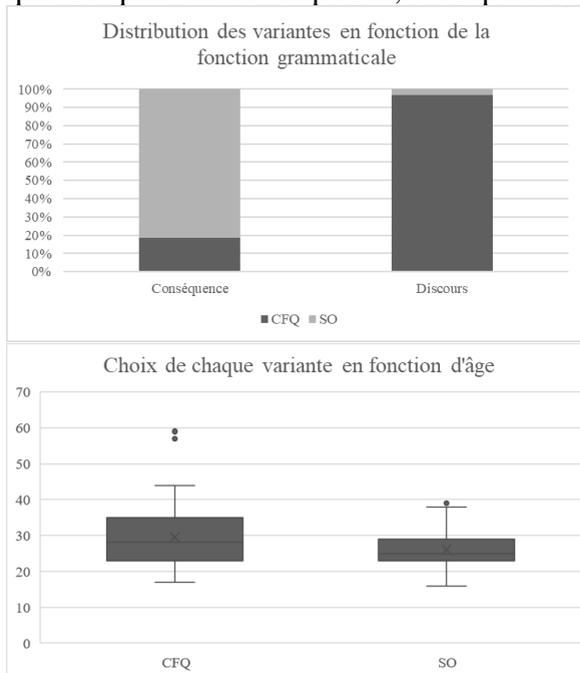
¹⁴ (Yes McCan, 2016).

¹⁵ La formule que nous avons utilisée pour effectuer l’analyse est la suivante :
rapkeb21.mrem=lmer(Variante ~ Âge + Genre + Nais + Orig + IdLing + IdCult + Année + Fonction + (1| Artiste), data=rapkeb21)

technique de contrastes binaires dite de « traitement » entre leur première sous-catégorie (ou « modalité ») qui se voit jouer le rôle de « modalité de référence » et chacune des autres sous-catégories (ou « modalités ») (Bringé et Golaz, 2017). La variable dépendante est traitée de manière binaire où la valeur '1' renvoie à la variante « so » et la valeur '0' renvoie à la variante (*ça*) fait (*que*). Les effets de facteurs sortis du modèle sont sur l'échelle logarithmique. Ainsi, la valeur de zéro sert de valeur pivot, c'est-à-dire que les valeurs positives indiquent une tendance vers « so » et les valeurs négatives indiquent une tendance vers (*ça*) fait (*que*).¹⁶

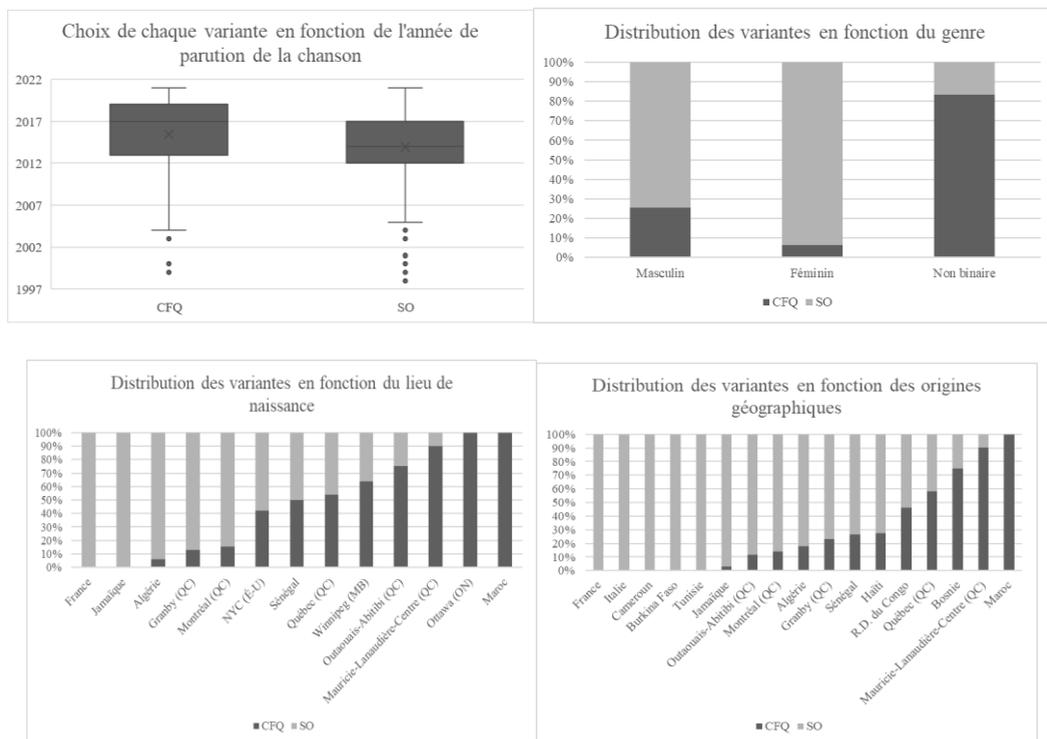
4. Résultats

La distribution globale des occurrences (N=1227) entre les deux variantes est de 24,45% (n=300) de (*ça*) fait (*que*) et de 75,55% (n=927) de « so ». Nous présentons ci-dessous la distribution des variantes (*ça*) fait (*que*) (CFQ) et « so » (SO) en fonction de chacune des variables indépendantes. Nous discutons brièvement de ces distributions afin de partir d'un portrait global des données avant de procéder aux résultats d'analyses statistiques. En fonction du seul facteur interne, la fonction grammaticale, les deux variantes sont distribuées de manière largement inverse. CFQ est majoritaire pour la fonction de marqueur de discours, et SO est majoritaire en tant que marqueur de conséquence, bien que ces tendances ne soient pas catégoriques.



¹⁶ Pour plus d'informations sur d'analyse multivariée par régression logistique, voir Tagliamonte et Baayen (2012).

**« So » ... fait que... le français québécois vernaculaire se diversifie-t-il?
Une étude sociolinguistique de l’alternance entre fait que et « so » dans le corpus «
RapKeb21 »**



Figures 1 à 6 Distribution des variantes en fonction des facteurs sociolinguistiques

Pour les deux facteurs sociaux de type continu, la distribution des variantes est présentée sous la forme de diagramme de quartiles. Si pour chacune des deux variables l'âge moyen des artistes l'ayant produite est dans la vingtaine, l'âge moyen des artistes ayant produit CFQ est légèrement plus élevé que celui correspondant à SO et l'éventail d'âges est plus important pour la production de CFQ que pour ce dernier. Aucun artiste d'âge supérieur à 40 ans n'a produit SO.

Quant à l'année de parution de la chanson, les deux variantes sont distribuées différemment. Le premier quartile des occurrences de chacune des deux variantes a été produit à travers une période relativement plus longue que celle du prochain quartile. Cette accélération est plus marquée pour la variante SO, le premier quartile ayant été produit sur une période de 7 ans et le second sur une période de 2 ans (2012-2014). À partir de 2014, l'accélération de SO s'est atténuée, pendant que la production de CFQ s'est accélérée, la moitié de ses occurrences ayant été produite entre 2017 et 2021.

Pour les trois facteurs sociaux de type catégoriel, la distribution des variantes est présentée sous la forme de diagrammes à barres empilés. Si la variante SO est majoritaire chez les artistes de genre féminin et masculin, elle l'est davantage pour ces premiers. En revanche, parmi ces artistes, c'est chez les artistes de genre masculin que CFQ est plus présent, mais il est exclusivement majoritaire chez les artistes de genre non binaire.

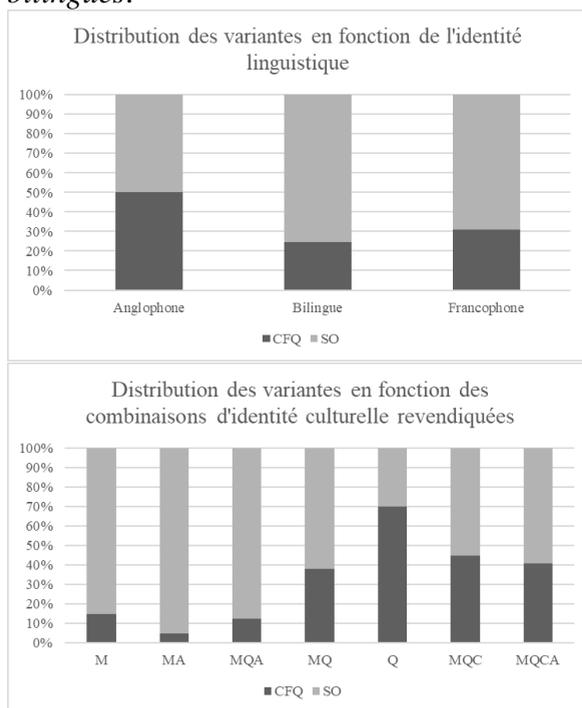
Les distributions des variantes en fonction du lieu de naissance et des origines géographiques présentent certaines similarités, mais aussi des différences subtiles et plus marquées. Les utilisateurs catégoriques de SO sont nés en France ou en Jamaïque, tandis que les utilisateurs catégoriques de CFQ sont nés à Ottawa en Ontario ou au Maroc. Il existe de la variation par rapport à tous les autres lieux de naissance. Les artistes ayant une production variable mais majoritaire de la variante SO sont nés en Algérie, à Granby, à Montréal ou à la ville de New York aux États-Unis. Les artistes nés au Sénégal n'ont tendance à produire ni SO, ni CFQ. Ceux faisant preuve d'une production variable mais majoritaire de CFQ sont nés à la ville de Québec, à Winnipeg au Manitoba, dans les régions d'Outaouais-Abitibi et de Mauricie-Lanaudière-Centre du Québec, ou à Ottawa en Ontario.

En fonction des origines géographiques, la distribution des variantes révèle plus d'utilisateurs catégoriques de « *so* » que dans le cas précédent. À l'instar des artistes nés en France, ceux d'origines françaises, mais également les artistes originaires de l'Italie, le Cameroun, le Burkina Faso et la Tunisie, produisent exclusivement SO. À l'opposé, les seuls utilisateurs catégoriques de CFQ ont des origines marocaines tout comme les artistes nés au même pays. Comme les artistes nés dans les mêmes lieux, les artistes ayant des origines jamaïcaines, montréalaises, algériennes, granbyennes et sénégalaises font un usage variable mais majoritaire de SO. Dans ce groupe figurent ceux ayant des origines d'Outaouais-Abitibi au Québec, haïtiennes ou congolaises. En revanche, les artistes produisant de manière variable mais quand même majoritaire de CFQ ont des origines de la ville de Québec, de la Bosnie ou de la région Mauricie-Lanaudière-Centre du Québec. Il existe une grande cohérence entre ces deux facteurs sociaux qui n'est altérée, si l'on exclut les groupes qui n'ont pas d'homologue dans l'une des deux variables, que par les artistes nés au Sénégal ou en région d'Outaouais-Abitibi et les artistes ayant des origines géographiques sénégalaises ou de région d'Outaouais-Abitibi. Ceux nés au Sénégal n'ont aucune préférence entre les deux variantes, tandis que ceux d'origines sénégalaises produisent plus de SO que de CFQ. Pendant que ceux nés en région d'Outaouais-Abitibi produisent plus de CFQ que de SO, ceux ayant que des origines de région Outaouais-Abitibi produisent davantage de SO.

Les distributions des variantes en fonction des deux facteurs sociopsychologiques, l'identité linguistique et l'identité culturelle, sont présentées ci-

**« So » ... fait que... le français québécois vernaculaire se diversifie-t-il?
 Une étude sociolinguistique de l'alternance entre fait que et « so » dans le corpus «
 RapKeb21 »**

dessous, aussi sous forme de diagrammes à barres empilés. Ces données ne renvoient pas à l'échantillon au complet (N=1084 et N=1179, respectivement), certains artistes n'ayant pas revendiqué d'identité à travers l'une des désignations recherchées. Il est très intéressant à constater que les artistes pour qui *anglophone* est la désignation d'identité linguistique dominante font des proportions égales des deux variantes, pendant que les artistes à identité dominante *francophone* ou *bilingue* produisent davantage de la variante SO empruntée à l'anglais que de la variante CFQ originaire du français. Cette dernière tendance est légèrement plus marquée pour les artistes *bilingues*.



Figures 7 et 8 Distribution des variantes en fonction des facteurs sociopsychologiques

La distribution des variantes en fonction de l'identité culturelle s'avère être plus complexes en raison du choix de tenir compte des différentes combinaisons de désignations *Montréalais(e)* (M), *Québécoise(s)* (Q), *Canadien(ne)* (C) et *Américain(e)* (A) (ainsi que leurs équivalents en anglais). Le seul groupe qui fait un usage majoritaire de CFQ comprend les artistes qui revendiquent exclusivement une identité québécoise. À l'opposé, ceux revendiquant une identité exclusivement montréalaise produisent SO comme variante majoritaire. Hormis ce groupe, les seuls

à ne revendiquer d'identité partiellement québécoise se disent *Montréalais(es)* et *Américain(es)* et sont les utilisateurs les plus exclusifs de SO. Les autres combinaisons comprennent toutes une dimension montréalaise et québécoise. Les artistes qui en rajoutent une dimension américaine produisent davantage de SO que ceux se disant *Montréalais(es)* et *Québécois(es)*. Ceux qui en rajoutent une dimension canadienne produisent relativement plus de CFQ que ces derniers, autant plus pour ceux qui n'en rajoutent pas aussi une dimension américaine.

Présentons maintenant les résultats statistiques de l'analyse des effets de variables indépendantes sur la production des variantes obtenus du modèle de régression logistique à effets mixtes. Rappelons que le modèle tient compte des effets fixes des variables d'âge, genre, lieu de naissance, origines géographiques, année de parution, identité linguistique, identité culturelle et fonction grammaticale, ainsi que de l'effet aléatoire de l'individu.¹⁷ La valeur des variantes SO et CFQ a été définie à '1' et à '-1', respectivement. La variance expliquée par le modèle est de 0,486. La statistique sommaire du modèle est de : $F(44, 1182) = 25.42, p < 2,2e-16$. Le tableau ci-dessous présente la valeur estimée des coefficients pour l'effet des catégories d'âge, année, fonction, genre, identité linguistique et identité culturelle.

	Valeur estimée	Écart-type	Z	p	
Ordonnée à l'origine (« intercept »)	-5.006	6.904	-0.725	0.469	
Âge	-0.002	0.003	-0.531	0.596	
Année	0.003	0.003	0.877	0.381	
Fonction - Marqueur de discours	-0.747	0.040	-18.719	<2E-16	***
Genre - Masculin	-0.169	0.070	-2.423	0.016	*
Genre - Non binaire	-0.939	0.223	-4.216	2.68E-5	***
Identité linguistique - Bilingue	-0.029	0.346	-0.083	0.934	
Identité linguistique - Francophone	-0.003	0.352	-0.008	0.993	
Identité culturelle - MA	0.142	0.093	1.528	0.127	
Identité culturelle - MQ	0.109	0.105	1.037	0.300	
Identité culturelle - MQA	0.156	0.063	2.475	0.013	*
Identité culturelle - MQC	0.010	0.069	0.148	0.882	
Identité culturelle - MQCA	0.151	0.078	1.945	0.052	.
Identité culturelle - Q	0.033	0.122	0.273	0.785	

Tableau 1 Résultats de l'analyse par régression logistique à effets mixtes

¹⁷ Voici le modèle tant qu'il est défini dans R : `rapkeb21.mrem=lmer(Variante ~ Âge + Genre + Nais + Orig + IdLing + IdCult + Année + Fonction + (1| Artiste), data=rapkeb21)`.

**« So » ... fait que... le français québécois vernaculaire se diversifie-t-il?
Une étude sociolinguistique de l'alternance entre fait que et « so » dans le corpus «
RapKeb21 »**

Nous constatons tout d'abord que les effets des variables continues d'âge et d'année ainsi que celui de la variable catégorielle d'identité linguistique ne se sont pas avérés statistiquement significatifs. Selon nos analyses, l'effet le plus important sur l'alternance entre SO et CFQ est celui de la fonction grammaticale. Le coefficient de la fonction de marqueur de discours tend plus vers le négatif que celle de l'ordonnée à l'origine qui englobe la fonction de marqueur de conséquence, ce qui signifie un effet favorable vers la variante CFQ en tant que marqueur de discours. Cette différence est hautement significative ($p < 2E^{-16}$). Le prochain effet le plus important est celui de la catégorie de genre. Le coefficient de la sous-catégorie non binaire tend le plus vers le négatif et celui de la catégorie masculine tend aussi plus vers le négatif, bien que de manière moins prononcée, par rapport au coefficient de la sous-catégorie féminine représentée à l'origine. Ces différences sont significatives, la première d'autant plus ($p = 2.68E^{-5}$) que la seconde ($p = 0.016$). Le dernier effet présenté dans ce tableau est celui de l'identité culturelle. Quant à cet effet, seul deux sous-catégories font preuve d'une différence significative par rapport aux autres. Ce sont la combinaison de dimensions identitaires montréalaise, québécoises et américaines qui produit le coefficient le plus élevé, ce qui indique un effet favorable vers la variante SO. Le prochain coefficient le plus élevé est celui de la combinaison des quatre dimensions montréalaise, québécoise, canadienne et américaine qui montre aussi un effet favorable vers la variante SO. Si l'on compare ces deux combinaisons (MQA et MQCA), nous constatons que la différence la plus remarquable entre elles et les autres combinaisons de désignations identitaires est la présence de la dimension américaine. Bien que ce ne surpasse pas le seuil de signification, l'autre combinaison comprenant la dimension américaine (MA), a le troisième coefficient le plus important.

Les derniers résultats dont nous discutons sont ceux correspondant aux facteurs sociaux, lieu de naissance et origines géographiques. Nous avons exclu du tableau suivant les sous-catégories dont la différence par rapport aux autres n'est pas significative, sauf celles de Montréal et de Québec, pour faciliter la comparaison. Par rapport au lieu de naissance, nous constatons cinq différences significatives au sein des données. Il s'agit dans chaque cas d'une tendance vers le négatif qui signalent une préférence pour la variante CFQ. Cet effet est le plus prononcé chez les artistes nés en région d'Outaouais-Abitibi, suivis de près par les artistes nés au Maroc, à Ottawa et en région de Mauricie-Lanaudière-Centre du Québec. Les artistes nés à Winnipeg au Manitoba font également preuve d'une préférence pour la variante CFQ, bien que moins prononcée.

Valeur estimée	Écart-type	Z	p
----------------	------------	---	---

Ordonnée à l'origine ("intercept")	-5.006	6.904	-0.725	0.469	
Naissance - Montréal (QC)	-0.278	0.185	-1.502	0.133	
Naissance - Québec (QC)	-0.232	0.228	-1.017	0.309	
Naissance - Mauricie-Lanaudière...	-0.750	0.128	-5.852	6.29E-09	***
Naissance - Winnipeg (MB)	-0.506	0.238	-2.127	0.034	*
Naissance - Maroc	-0.870	0.477	-1.822	0.069	.
Naissance - Ottawa (ON)	-0.787	0.391	-2.014	0.044	*
Naissance - Outaouais-Abitibi (QC)	-1.125	0.337	-3.336	0.001	***
Origines - Montréal (QC)	0.233	0.164	1.423	0.155	
Origines - Québec (QC)	-0.149	0.206	-0.721	0.471	
Origines - Liban	-0.600	0.358	-1.678	0.094	.
Origines - Cameroun	0.413	0.212	1.951	0.051	.
Origines - Bosnie	-0.434	0.214	-2.029	0.043	*

Tableau 2 Résultats d'analyse par régression logistique à effets mixtes (cont.)

Enfin, trois sous-catégories d'origines géographiques s'avèrent présenter une différence significative par rapport aux autres. Dans ce cas, les coefficients correspondant aux artistes d'origines bosniennes et libanaises tendent vers le négatif, soulignant ainsi un effet favorable pour la variante CFQ, tandis que celui des artistes d'origines camerounaises tend vers le positif et signale ainsi un effet favorable pour la variante SO.

5. Discussion

Tout d'abord, il faut apprécier la forte présence de « *so* » (75,5%) dans notre corpus. Cette variante n'a que très récemment été attestée au Québec et, ce, uniquement au sein d'un corpus de textos d'adolescents (Blondeau et al, 2022). Il est aussi intéressant à constater que « *so* » n'est guère en concurrence avec (*ça*) fait (*que*) en tant que marqueur de discours. Les deux variantes ne sont concurrentes que sous forme de marqueur de conséquence, leur distribution au sein de cette catégorie étant semblable à leur distribution globale au sein du corpus. Ce conditionnement interne l'emporte largement sur les effets de facteurs sociaux. Nous nous demandons s'il ne s'agit de cas de variation que pour la fonction grammaticale de marquer une relation de conséquence. En revenant sur les données brutes, nous n'observons que trois occurrences de « *so* » en tant que marqueur de discours, un nombre si faible qu'il pourrait s'agir de valeurs aberrantes. Si l'on considère les trois occurrences, reproduites ci-dessous, deux ont été produites dans la même chanson de Dead Obies et l'autre dans une chanson de One Nessa. Il s'avère fort probable dans ces deux

**« So » ... fait que... le français québécois vernaculaire se diversifie-t-il?
Une étude sociolinguistique de l'alternance entre fait que et « so » dans le corpus «
RapKeb21 »**

emplois différents que « *so* » apparaît en position finale d'énoncé afin de servir effectivement à la rime.

- 1) *Comme le silence pour faire durer l'suspense, so... / More shows, No-Doz.*¹⁸
- 2) *I know you fine but you ain't mine so / That's why we on then we off.*¹⁹

Ainsi, il semble juste d'affirmer que l'emploi de « *so* » comme marqueur de discours ne résulte pas de conditionnement social, plutôt de conditionnement de la rime. Alors, les effets sociaux que nous observons sur la production des variantes « *so* » et (*ça*) *fait (que)* ne se produisent que lorsqu'elles servent de marqueurs de conséquence. Cela implique que la distribution latente de ces deux variantes lorsqu'il est question de variation est encore plus asymétrique qu'il n'y paraît à première vue.

Si l'on exclut les occurrences de la sous-catégorie de marqueurs de discours, il s'avère que « *so* » prédomine davantage et représente 81,41% (n=924) des occurrences pendant que (*ça*) *fait (que)* n'en représente que de 18,59% (n=211). Car cette nouvelle répartition permet d'éclairer la situation de conditionnement social de la variable, nous avons effectué une nouvelle analyse par régression à effets mixtes en tenant uniquement en compte les 1135 occurrences produites comme marqueurs de conséquence. Nous avons vérifié que les effets sur la variation sont quasiment identiques à ce que nous avons présenté dans la section précédente, à la seule exception de l'effet des origines géographiques. Il est très intéressant à constater cette fois une différence significative chez les artistes ayant des origines de la ville de Québec. La différence est négative et signale que ces artistes tendent à utiliser (*ça*) *fait (que)* plus des autres artistes sauf ceux d'origines libanaises, bosniennes et congolaises qui tendent, eux aussi, à produire (*ça*) *fait (que)*. Si nous avons observé auparavant une différence significative, bien que faible, chez les artistes ayant des origines camerounaises, cet effet disparaît lors de cette nouvelle analyse. Tous les autres résultats étant les mêmes que ceux déjà rapportés, continuons notre discussion des autres variables indépendantes.

¹⁸ (Dead Obies, 2016, DO 2 Get, *Gesamtkunstwerk*).

¹⁹ (One Nessa, 2015, On and Off, *The Archer*).

	Valeur estimée	Écart-type	Z	p
Ordonnée à l'origine ("intercept")	-6.030	7.823	-0.771	0.441
Origines - Montréal (QC)	0.055	0.126	0.435	0.664
Origines - Québec (QC)	-0.454	0.201	-2.257	0.024 *
Origines - Liban	-0.741	0.352	-2.104	0.036 *
Origines - Bosnie	-0.753	0.220	-3.428	0.0006 ***
Origines - R.D. du Congo	-0.466	0.211	-2.206	0.028 *

Tableau 3 Résultats de la 2^e analyse par régression pour l'effet d'origines géographiques

Ce que nous observons en fonction du genre de l'artiste nous intrigue. Il semble bien que les artistes de genre féminin se démarquent des autres artistes par une forte préférence pour la variante « *so* », pendant que les artistes de genre non binaire se démarquent tout à fait dans l'autre sens en privilégiant (*ça*) *fait (que)*. Bien que l'effet d'âge ne surpasse pas le seuil de signification statistique, la distribution des variantes en fonction d'âge révèle que ce sont les artistes les plus jeunes qui produisent « *so* ». La combinaison de ces deux effets pourrait éventuellement signaler un changement en cours concernant la progression de « *so* » chez les artistes rap.

Il reste encore à considérer les effets significatifs du lieu de naissance de l'artiste et de son choix de désignations identitaires culturelles. Les artistes nés à l'ouest du Québec, en région Mauricie-Lanaudière-Centre du Québec et dans d'autres provinces canadiennes font tous preuve d'une production de (*ça*) *fait (que)* plus importante que celle de l'échantillon global d'artistes.

Concernant la revendication des désignations identitaires dont nous avons tenu compte, la seule tendance qui atteint le seuil de signification statistique semble surtout être liée à la dimension américaine lorsqu'il est question d'identités plurielles. Les artistes qui se disent être Américain(e)s sont plus enclins à produire la variante « *so* ». Faute d'informations supplémentaires, nous nous contentons de soumettre la possibilité, qui devra être explorée, que l'aspect multilingue de l'utilisation de cette variante soit motivé par le désir d'exprimer une identité dépassant le cadre local. Dans un contexte moderne de mondialisation, il semble possible que certains artistes fassent usage de certaines variantes linguistiques indexant des associations identitaires non seulement locales mais aussi plus larges. Cela pourrait être motivé par un désir latent de se connecter à la nation rap mondiale en favorisant, consciemment ou inconsciemment, une variante linguistique mettant l'accent sur les compétences multilingues.

Conclusion

En résumé, l'objectif de notre étude était de vérifier la théorie de Blondeau et Tremblay (Blondeau, 2018, 2020; Blondeau et Tremblay, 2016, 2022a) de la diversification du français vernaculaire contemporain au Québec et en contexte urbain à Montréal. Notre question de recherche concernant plus particulièrement l'alternance entre (*ça*) fait (*que*) et « so » était d'évaluer si cette dernière est contrainte par des facteurs sociaux d'une manière à suggérer l'existence de deux pôles normatifs quant au vernaculaire au Québec. Nous avons posé l'hypothèse qu'il existe un pôle centré dans la métropole montréalaise et un autre qui serait plutôt délocalisé sur l'ensemble de la province. Nous avons effectivement constaté des tendances significatives qui signalent la progression de la variante « so » en tant que marqueur de conséquence dans la production lyrique des 80 artistes de l'échantillon représenté dans notre corpus « RapKeb21 » (Leavitt, E.B., 2022). Ce sont particulièrement les artistes de genre féminin et de plus jeune âge qui produisent davantage cette variante. Il semble que malgré la hausse de « so » au sein de la communauté rap québécoise l'utilisation de la variante (*ça*) fait (*que*) indexe dans une certaine mesure une association aux origines québécoises. Notre analyse semble également mettre en évidence la possibilité que la variante « so » fait aussi l'objet d'une indexicalisation renvoyant, elle, à une identité plurielle et complexe liée à la fois aux pratiques langagières locales et mondiales. Ces constats sont généralisables pour leur part à la communauté rap québécoise et ne peuvent être considérés comme représentatifs des pratiques linguistiques de la population québécoise et de ses différentes communautés linguistiques. Il convient toutefois d'approfondir la question pour déterminer si les mêmes schémas peuvent être observés dans certaines communautés linguistiques, telles que celle des jeunes Montréalais(es), dont il a été démontré lors des études antérieures²⁰ qu'ils partagent certaines pratiques langagières avec les artistes rap québécois avec qui ils interagissent par le biais de leur participation à la culture populaire ainsi qu'au public du rap québécois.

²⁰ (Sarkar et al, 2005, 2006, 2007, 2008, 2009, 2012, 2014).

Références

- AGENCE QMI, (2022, 29 avril), « Loud reprend la tournée », *Le Journal de Montréal*.
- ALACLAIR ENSEMBLE, (2013), *Les maigres blancs de l'Amérique du Noir*, Disques 7ième Ciel.
- ALEX Y Kalam. 2022, 19 avril. « Après une tournée en Afrique, Monk.E revient au Québec avec du nouveau matériel », *Journal du Hip Hop*.
- ALIM H. Samy, (2009), « Straight outta Compton, straight aus Munchen : Global linguistic flows, identities, and the politics of language in a global hip hop nation », in, *Global linguistic flows: Hip hop cultures, youth identities, and the politics of language*, sous la direction de H. Samy Alim, Awad Ibrahim et Alastair Pennycook, pp. 1–22, New York, NY, Routledge.
- ANDROUTSOPOULOS Jannis, (2009), « Language and the three spheres of hip hop », in, *Global linguistic flows: Hip hop cultures, youth identities, and the politics of language*, sous la direction de H. Samy Alim, Awad Ibrahim et Alastair Pennycook, pp. 43–62, New York, NY, Routledge.
- BATES Douglas et al, (2015), « Fitting Linear Mixed-Effects Models Using lme4 », in *Journal of Statistical Software* 67, 1–48.
- BILEFSKY Dan, (2020, 7 avril), « ‘What Rhymes With Purell?’ Français Rappers Push Language Boundaries in Quebec », *The New York Times*.
- BLANCHET Jean-François, (2021, 24 décembre), « Roxane Bruneau et FouKi ont dominé les ondes radio de Québec en 2021 », *ICI Québec Radio-Canada*.
- BLONDEAU Hélène, (2006), « La trajectoire de l'emploi du futur chez une cohorte de Montréalais francophones entre 1971 et 1995 », in, *Revue de l'Université de Moncton* 37(2), 73–98.
- BLONDEAU Hélène, (2018), « Pratiques langagières à Montréal et à Bruxelles : diversité et dualité linguistique », in, *Revue française de linguistique appliquée* 23, 73-88.
- BLONDEAU Hélène, (2020), « Pratiques langagières et diversité culturelle chez de jeunes Montréalais : le français dans la métropole », in, *Attribuer un sens. La diversité des pratiques langagières et les représentations sociales*, sous la direction de Kristin Reinke, pp. 151-175, Presses de l'Université Laval.
- BLONDEAU Hélène, MARTINEAU France et FRENETTE Yves, (2016), « Francophonie montréalaise et globalisation : évolution des pratiques langagières en contexte à Hochelaga-Maisonneuve », in, *Cahiers internationaux de sociolinguistique* 10, 159-182.
- BLONDEAU Hélène, MOUGEON Raymond et TREMBLAY Mireille, (2019), « Analyse comparative de ça fait que, alors, donc et so à Montréal et à Welland: Mutations sociales, convergences, divergences en français laurentien », in, *Journal of French Language Studies* 29, 35-65.
- BLONDEAU Hélène et TREMBLAY Mireille, (2016), « Le traditionnel et l'émergent : l'apport de jeunes Montréalais issus de l'immigration au français vernaculaire », in, *Cahiers internationaux de sociolinguistique* 10, 19-45.

**« So » ... fait que... le français québécois vernaculaire se diversifie-t-il?
Une étude sociolinguistique de l’alternance entre fait que et « so » dans le corpus «
RapKeb21 »**

- BLONDEAU Hélène et TREMBLAY Mireille, (2022a), « Écrire son vernaculaire: variation et normes communautaires dans les messages textes en français québécois », in, *Journal of French Language Studies* 32(2), 120–144.
- BLONDEAU Hélène et TREMBLAY Mireille, (2022b), « The hidden dimensions of a change from below: Consequence markers in Montreal French », in, *Revue Canadienne de Linguistique* 67(1-2), 22-52.
- BLONDEAU Hélène et al. 2021. « A new milestone for the study of variation in Montréal French: The Hochelaga-Maisonneuve sociolinguistic survey », in, *Corpus* 22, 1-16.
- BLONDEAU Hélène, TREMBLAY Mireille et BOURÉLY Claire, (2022), « La dynamique de la variation pronominale : micro-diachronie des formes simples et complexes en français montréalais », in, *Langages* 226, 69-82.
- BORDELEAU Jean-Louis, (2022, 28 avril), « Le rappeur White-B écope d'une peine de prison de deux ans moins un jour », *Le Devoir*.
- BRINGÉ Arnaud et GOLAZ Valérie, (2017), « Modèles logistiques, modèles contextuels », in, *Manuel pratique d'analyse multiniveau*, 41-54, Ined Éditions [En ligne].
- CHESHIRE Jenny et al, (2013), « Language contact and language change in the multicultural metropolis », in, *Revue française de linguistique appliquée* 18, 63-76.
- CÔTÉ Émilie, (2018, 22 mars), « Après la France, la FM », *La Presse*.
- COUTURIER Jules, (2022, 7 octobre), « Shreez: un rap qui évolue tout en restant fidèle », *Métro Média*.
- DEAD OBIES, (2016), *Gesamtkunstwerk*. Bonsound.
- ÉTIENNE Anne-Lovely, (2020, 6 février), « Les coups de cœur de Sarahmée », *Le journal de Montréal*.
- FRIESNER Michael Lee, (2009), *The social and linguistic predictors of the outcomes of borrowing in the speech community of Montréal* (No de publication 3363297) [Thèse de doctorat], University of Pennsylvania.
- GADET Françoise et GUERIN Emmanuelle, (2016), « Construire un corpus pour des façons de parler non standard : « Multicultural Paris French » », in, *Corpus* 15 [En ligne].
- HOULE Nicolas et BOUCHARD Geneviève, (2017, 26 octobre), « Le hip-hop sort de la marge à l'ADISQ », *Le Droit*.
- IZZY-S. 2018. *Le sablier*. Explicit Nation.
- JBEILI Christophe, (2023, 12 juin), « New In Rap Queb: Lost is Quebec's Biggest Rapper Right Now », *Complex Canada*.
- LABOV William, (1984), « Field methods of the project on linguistic change and variation », in, *Language in Use: Readings in Sociolinguistics*, sous la direction de John Baugh et Joel Sherzer, pp. 28- 66, Englewood Cliffs, NJ, Prentice Hall.
- LAMARRE Patricia, (2013), « Catching “Montréal on the Move” and Challenging the Discourse of Unilingualism in Québec », in, *Anthropologica* 55, 41–56.

- LAMARRE Patricia, (2014), « Bilingual winks and bilingual wordplay in Montreal's linguistic landscape », in, *International Journal of the Sociology of Language* 228, 131-151.
- LANGLAIS Philippe et DROUIN Patrick, (2012), « Texto4Science: A Quebec French database of annotated text messages », in, *Linguisticae investigationes* 35, 237–259
- LAUZON Véronique, (2018, 15 septembre), « Koriass tue l'éléphant », *La Presse*.
- LEAVITT Emily Brooke, (2021), « A prosodic comparison of accentuation, intonation, and rhythm between Quebec French rap production and normative Quebec French discourse », in, *Midwest Journal of Undergraduate Research* 12, 118–136.
- LEAVITT Emily Brooke, (2022), *De la variation linguistique dans le rap québécois. Une étude sociolinguistique* (N° de publication 991396) [Mémoire de maîtrise], Université Concordia, Montréal.
- LEAVITT Sarah, (2022), 26 août. « Transit agency cuts Montreal rapper's promotional video over use of English », *CBC Radio-Canada*.
- LEBLANC Pascal, (2023a, 9 juin), « Lost en quatre temps », *La Presse*.
- LEBLANC Pascal, (2023b, 19 juin), « Yes Mccan change de nom et annonce un nouvel album », *La Presse*.
- LEE Jamie Shinhee, (2010), « Glocalizing keepin' it real: South Korean hip-hop playas », in, *The languages of global hip hop*, sous la direction de Marina Terkourafi, pp. 139–161, London, Continuum International Publishing Group.
- LOUD LARY AJUST, (2014), *Blue Volvo*. Audiogram.
- LOW Bronwen et SARKAR Mela, (2012), « Le plurilinguisme dans les arts populaires, un terrain inexploré ? L'étude du langage mixte du rap montréalais en guise d'exemple », in, *Kinéphanos. Revue d'études des médias et de culture populaire* 3, 20–47.
- LOW Bronwen et SARKAR Mela, (2014), « Translanguaging in the multilingual Montreal hip-hop community: Everyday poetics as counter to the myths of the monolingual classroom », in, *Heteroglossia as practice and pedagogy, Educational Linguistics* 20, sous la direction de Adrian Blackledge et Angela Creese, pp. 99–118, Springer-Verlag.
- LOW Bronwen, SARKAR Mela et WINER Lise, (2009), « “Chus mon propre Bescherelle”: Challenges from the hip-hop nation to the Quebec nation », in, *Journal of Sociolinguistics* 13, 59–82.
- MALSTROM Kasheem. 2001. *Ange ou demon*. Ilov Records.
- MARTINEAU France et SÉGUIN Marie-Claude, (2016), « Le Corpus FRAN : réseaux et maillages en Amérique française », in, *Corpus de français parlé et français parlé des corpus* 15, 1-24.
- ONE NESSA, (2015), *The Archer*. Melomayne; Ravalian; Vader NOGAMZ.
- PALMARÈS ADISQ, (2022, 2 novembre), « Dévoilement des gagnant.es du Gala de l'Industrie et du premier Gala de l'ADISQ 2022 », ADISQ.
- PALMARÈS ADISQ, (2023), « Les nominations : album de l'année rap », Galas ADISQ.
- PERRON Louis-Samuel, (2022, 9 juin), « Un rappeur condamné à trois ans de prison soupçonné d'avoir menacé une avocate », *La Presse*.

**« So » ... fait que... le français québécois vernaculaire se diversifie-t-il?
Une étude sociolinguistique de l’alternance entre fait que et « so » dans le corpus «
RapKeb21 »**

- POPLACK Shana, (2018), *Borrowing. Loanwords in the Speech Community and in the Grammar*, New York, NY, Oxford University Press.
- RAMPTON Ben, (1995), *Crossing: Language and ethnicity among adolescents*. (Real language series), London et New York, NY, Longman.
- R CORE TEAM, (2021), « R: A language and environment for statistical computing », *R Foundation for Statistical Computing*, Vienna, Austria.
- RENAUD Philippe, (2021, 29 mars), « Les leçons de la polémique du franglais dans le rap québécois », *Le Devoir*.
- RENAUD Philippe, (2023, 3 juin), « Le rap français reprend les scènes du Québec », *Le Devoir*.
- RSTUDIO TEAM, (2020), « RStudio: Integrated Development for R », *RStudio*, Boston, MA, PBC.
- SANKOFF David et al. 1976. « Méthodes d’échantillonnage et utilisation de l’ordinateur dans l’étude de la variation grammaticale », in, *Cahiers de Linguistique de l’Université du Québec* 6, 85–125.
- SARKAR Mela, (2008), « « Ousqu’on chill à soir ? » : pratiques multilingues comme stratégies identitaires dans la communauté hip-hop montréalaise », in, *Diversité Urbaine, Numéro thématique : Plurilinguisme et identités au Canada* 1, 27-44.
- SARKAR Mela, (2009), « “Still reppin’ pour mi gente”. The transformative power of language mixing in Quebec hip hop », in, *Global linguistic flows: Hip hop cultures, youth identities, and the politics of language*, sous la direction de H. Samy Alim, Awad Ibrahim et Alastair Pennycook, pp. 139–158, New York, NY, Routledge.
- SARKAR Mela et ALLEN Dawn, (2007), « Hybrid identities in Quebec hip-hop: Language, territory and ethnicity in the mix », in, *Journal of Language, Identity and Education* 6(2), 117–130.
- SARKAR Mela et LOW Bronwen, (2012), « Multilingualism and popular culture », in, *The Routledge handbook of multilingualism*, sous la direction de Marilyn Martin-Jones, Adrian Blackledge et Angela Creese, pp. 403–418, London, Routledge.
- SARKAR Mela et WINER Lise, (2006), « Multilingual codeswitching in Quebec rap: Poetry, pragmatics and performativity », in, *International Journal of Multilingualism* 3(3), 173–192.
- SARKAR Mela, WINER Lise et SARKAR Kobir, (2005), « Multilingual code-switching in Montreal hip-hop: Mayhem meets method or, “tout moune qui talk trash kiss mon black ass du nord” », in, *Proceedings of the 4th International Symposium on Bilingualism*, sous la direction de James Cohen, Kara T. McAlister, Kellie Rolstad et Jeff MacSwan, pp. 2057–2074, Somerville, MA, Cascadilla Press.
- SHUSTERMAN Richard, (2005), « Rap aesthetics: violence and the art of keeping it real », in, *Hip hop and philosophy*, sous la direction de Derrick Darby et Tommie Shelby, pp. 54-64, Open Court.

- TAGLIAMONTE Sali A. et BAAYEN Harald R, (2012), « Models, forests, and trees of York English: *Was/were* variation as a case study for statistical practice », in, *Language Variation and Change* 24, 135-178.
- TÉLÉ-QUÉBEC, (2022), *La fin des faibles*. Télé-Québec TV.
- THIBAUT Pierette et VINCENT Diane, (1990), « Un corpus de français parlé : recherches sociolinguistiques. Montréal 84: historique, méthodes et perspectives de recherche », in, *Recherches Sociolinguistiques* 1, Québec, Département de Langues et Linguistique, Université Laval.
- VINCENT Diane, LAFOREST Marty et MARTEL Guylaine, (1995), « Le corpus de Montréal 1995 : adaptation de la méthode d'enquête sociolinguistique pour l'analyse conversationnelle », in, *Dialangue* 6, 29-45.
- YES MCCAN, (2016), Yes McCan vs Hermano Salvatore. *WordUP! #TREIZE*. WordUP! Battles.

Biographie de l'auteur

Emily B. Leavitt est doctorante en sciences du langage à l'Université de Sherbrooke, chargée de cours à l'Université Laval et titulaire d'une maîtrise en sciences humaines spécialisée en sociolinguistique de l'Université Concordia. Ses recherches sont subventionnées par les Fonds de recherche du Québec – Société et culture.