

La résilience du *ne* en français contemporain : Étude d'un marqueur stylistique à Paris, Montréal et Marseille

Marie Flesch¹, Julie Abbou², Mireille Tremblay³ et Heather Burnett¹

¹LLF (CNRS-Université Paris Cité) - France

²Università di Torino - Italie

³Université de Montréal - Canada

*marie.flesch@gmail.com

Résumé. Cet article présente une nouvelle contribution à l'abondante littérature sur l'analyse de l'utilisation et de la particule négative *ne* en français contemporain (*Je ne l'aime pas* vs *Je l'aime pas*). Il présente une étude du corpus CaFé (Cartographie linguistique des féminismes), composé de 102 entretiens sociolinguistiques semi-dirigés réalisés à Montréal, Paris et Marseille en 2021 et 2022 avec des militant.es féministes et queer (2 millions de mots). Il met en lien la variabilité dans la réalisation du *ne* avec un facteur linguistique (type de sujet) et plusieurs facteurs discursifs (forme d'adresse, thèmes abordés, personnes connues/non connues de la chercheuse) et sociodémographiques (âge, éducation, engagement, ville). L'analyse de 25 560 négations montre que le type de sujet, l'âge, les thèmes abordés et la relation entre la chercheuse et les personnes enregistrées ont un effet significatif sur la réalisation du *ne*, comme l'ont montré de nombreuses études. L'originalité de notre étude tient au fait que le corpus CaFé permet une comparaison directe entre les trois villes. Cela révèle qu'il y a une variation dialectale très limitée et que les taux d'emploi du *ne* sont très semblables dans les trois sous-corpus : 11.22 % à Paris, 10.75 % à Marseille et 8.87 % à Montréal. La variation dialectale est limitée à un contexte morphosyntaxique précis : les sujets lexicaux non redoublés par un pronom. Notre étude souligne l'importance de construire des corpus comparables pour étudier la variation sociale et dialectale.

1 Introduction

Cet article présente une nouvelle contribution à l'abondante littérature sur l'analyse de l'utilisation de la particule négative *ne* en français contemporain. L'alternance entre les phrases négatives où l'adverbe de la négation *pas* est variablement redoublé par la particule préverbale *ne* (eg. *Je ne l'aime pas* vs *Je l'aime pas*) est une des variables sociolinguistiques les plus étudiées du français (voir Gadet 2000 pour un résumé). Même si la littérature antérieure met en évidence des différences dans les taux et les conditions de l'emploi du *ne* selon les communautés étudiées (voir ci-dessous), elle présente néanmoins une vision du phénomène assez unifié: il s'agirait d'un changement en cours (Ashby 1976, 1981, Armstrong et Smith 2002, Martineau et Mougeon 2003) en lien avec le cycle de Jespersen (Jespersen 1917), où la négation obligatoirement bipartite du français classique serait en train de disparaître en faveur d'une négation simple avec *pas*. Cette analyse s'appuie sur le fait que, dans plusieurs études (Ashby 1981, 2001, Coveney 1996, Hansen et Malderez 2004, Avanzi et al. 2016, entre autres), un effet d'âge a été observé : les locuteurs plus jeunes emploieraient moins de *ne* que les plus âgés. De plus, des études diachroniques en temps réel, comme Martineau et Mougeon (2003) et Hansen et Malderez (2004), ont démontré que les taux de *ne* sont plus bas dans les corpus plus récents. En ce qui concerne les contextes linguistiques, sociaux et discursifs qui conditionneraient ce changement en cours, le *ne* serait préservé après des sujets lexicaux (Diller 1983, Armstrong et Smith 2002, Donaldson 2017, Hansen et Malderez 2004, Martineau et Mougeon 2003, Meisner 2010, Meisner et Pomino 2014, Villeneuve 2017, entre autres), dans le parler des locuteurs et locutrices des classes sociales favorisées (Ashby 1981, 2001, Buson et al. 2022, Coveney 1996, Hansen et Malderez 2004, entre autres), et dans des contextes formels (Ashby 1981, Armstrong 2002, Fonseca-Greber 2007, Meisner 2010, Villeneuve 2017, entre autres). À ces variations grammaticales, sociales et stylistiques s'ajoute l'observation de la variation géographique: malgré le fait que les études européennes présentent des taux d'emploi du *ne* très variés, allant jusqu'à 78 % dans des entretiens informels avec des locuteurs de classes « supérieures » en France (Diller 1983), les taux d'emploi du *ne* observés au Canada tendent à être

largement inférieurs (par exemple, 0.5 % dans des entretiens informels à Montréal (Sankoff et Vincent 1977)).

Il reste néanmoins plusieurs points à clarifier dans l'analyse générale décrite ci-dessus. Le premier point concerne la variation géographique. Il est vrai que les taux de *ne* observés dans des études sur le français parlé québécois sont, en moyenne, largement inférieurs à ceux des études sur les parlers européens (par exemple, à notre connaissance, le taux le plus élevé dans des entretiens au Québec (8%) a été observé par Villeneuve 2017 dans son étude d'entretiens semi-formels télévisés), mais il y a eu quelques études sur le territoire français (Garel 1997, Armstrong 2002) qui ont aussi observé des taux très bas (autour de 1%) rivalisant avec les études canadiennes. Ces résultats soulèvent la possibilité que la variation observée entre le Canada et la France n'est pas strictement d'ordre géographique, mais pourrait être due à d'autres différences entre les corpus utilisés dans les études sur ces deux continents. Malheureusement, il n'y a pas encore eu d'étude comparative franco-québécoise basée sur des corpus d'entretiens directement comparables. Notre article fournit une telle étude: nous présentons une étude quantitative sur l'emploi du *ne* dans le corpus *Cartographie linguistique des féminismes (CaFé)* (Abbou et Burnett 2023). Le corpus CaFé est un corpus d'entretiens sociolinguistiques avec 102 militant(e)s féministes et queer à Paris (42 entretiens), Montréal (40 entretiens) et Marseille (20 entretiens). Il s'agit d'entretiens informels basés sur un questionnaire commun et des critères de recrutement communs, ce qui rend les trois sous-corpus les plus comparables possibles. Dans notre étude, basée sur 25 560 phrases négatives, nous trouvons des taux d'emploi du *ne* qui ne varient que très peu à travers les trois corpus (11.22 % à Paris, 10.75 % à Marseille et 8.87 % à Montréal). De plus, nous observons que les mêmes facteurs linguistiques et sociaux conditionnent cette variable dans les trois corpus. Selon des analyses statistiques (régression logistique effectuée avec le logiciel R version 4.1.1., R Core Team 2021), nous trouvons les effets d'âge et du sujet grammatical attendus dans les trois corpus : les locutrices plus âgées utilisent plus de *ne* dans les trois corpus, et l'apparence du *ne* est favorisée après des sujets lexicaux plutôt que des pronoms. Ces résultats mettent en question s'il existe réellement une différence dialectale importante entre la France et le Canada.

Le fait que nous ayons trouvé un taux assez élevé de *ne* dans notre corpus nous permet d'investiguer d'autres questions ouvertes de façon plus détaillée, en particulier, la définition de la formalité. Comme mentionné ci-dessus, la formalité est souvent mentionnée comme une propriété importante pour le conditionnement du *ne*, mais les études antérieures l'ont souvent conceptualisée de façons différentes. Par exemple, Armstrong (2002), Meisner (2010) et Villeneuve (2017) ont montré que le registre (style d'entretien) a un effet sur l'emploi du *ne* : les entretiens formels favorisent l'emploi du *ne*, comparé avec les entretiens informels ou les conversations entre pairs. À l'intérieur d'un seul contexte social, la formalité et l'informalité peuvent être conceptualisées comme la connaissance ou non de l'intervieweur et de celui ou celle qui est interviewé(e) (Meisner 2010), la partie temporelle de l'entretien (Ashby 1981) ou même les sujets ou thématiques abordés dans les entretiens (Sankoff et Vincent 1977, Donaldson 2017). Le corpus CaFé est constitué uniquement d'entretiens semi-dirigés, donc nous ne pouvons pas tester les effets de registre. Par contre, le corpus varie selon les autres dimensions de formalité identifiées dans la littérature : un sous-ensemble des participantes sont connues de l'intervieweuse avant l'entretien, et l'entretien (90 minutes en moyenne) est structuré en trois parties successives, qui abordent chacune des thématiques distinctes. Nos résultats montrent que ces trois dimensions (la connaissance, la place temporelle et les thématiques) ont toutes un effet sur l'emploi du *ne*, et sont donc à investiguer (ou à contrôler) dans les études futures sur cette variable sociolinguistique.

L'article commence par décrire le corpus CaFé avant de détailler la méthodologie employée pour annoter les négations et codifier les facteurs linguistiques et sociaux, en revenant sur les études antérieures sur la négation en français. Il présente ensuite les résultats de notre étude, avec une section consacrée à des statistiques descriptives, puis un modèle de régression. Enfin, il propose une discussion de nos résultats.

2 Le corpus CaFé

Le corpus CaFé est un corpus sociolinguistique francophone de 102 entretiens semi-dirigés, menés auprès de personnes engagées dans le féminisme, la cause des femmes, le queer ou les luttes pour les sexualités, à Paris (42 entretiens), Montréal (40 entretiens) et Marseille (20 entretiens). Deux entretiens ont été réalisés avec plusieurs personnes : un à Montréal (deux personnes) et un à Paris (trois personnes). Comme on pourrait s'y attendre dans un corpus axé sur le féminisme, la très grande majorité des participant.e.s sont des femmes (cis ou trans), mais il y a aussi une petite minorité d'hommes (cis et trans) et de personnes non-binaires. Paris, Montréal et Marseille sont trois grandes villes multiculturelles, et la composition des milieux féministes le reflète. Beaucoup de nos participant.e.s sont bilingues (français-arabe, français-kabyle, français-portugais, français-anglais, entre autres), et nous avons fait le choix de ne pas exclure des participant.e.s dont la langue maternelle n'est pas le français, pourvu que leur engagement féministe se passe en français. En conséquence, nous avons trois locutrices qui ont appris le français après l'enfance : une dont la langue maternelle est l'espagnol mexicain, une dont la langue maternelle est l'anglais nigérian et une dont la langue maternelle est le suédois. Les 105 locutrices et locuteurs enregistré.es ont été recruté.es sur la base de leur activité dans le mouvement féministe ou queer, avec une attention à maintenir une représentation qualitative des différents rapports sociaux (notamment âge, éducation, etc.).

Les participant.e.s ont été recruté.e.s selon 5 catégories d'engagement féministe : 1) **Universitaire** : l'engagement professionnel pour la formation, la création et la production de connaissances féministes. Il s'agit ici principalement d'universitaires travaillent en études de genre, mais également de personnes s'employant à diffuser une lecture féministe de la science (universitaires, médecins) ou à promouvoir les femmes dans les sciences. 2) **Professionnelle** : l'activité professionnelle de fabrication, de négociation ou d'application des lois, règlements, procédures en lien avec le féminisme, la cause des femmes ou les droits sexuels. Il s'agit ici typiquement de professionnelles de l'égalité, des questions de genre et/ou de sexualité (chargée de mission ou référent.e égalité femmes/hommes, travailleuses de la diversité du secteur public ou privé, travailleuses communautaires, avocates, formatrices). 3) **Associative** : les activités associatives ou communautaires bénévoles au sein d'associations ou d'organismes. Ce travail peut être du lobbying sur secteur spécifique ou plus largement un travail d'élaboration de cadre d'action sociale. Ce sont ici des militantes, porte-paroles, etc. d'associations, de collectifs, ou de syndicats. 4) **Média** : le travail, salarié ou bénévole, de diffusion médiatique et numérique des idées féministes. Cette modalité concerne les médias, le monde du livre et de la presse (papier comme numérique) et rassemble des auteures, éditrices, bibliothécaires, rédactrices de revue, traductrices, personnalités publiques, journalistes, etc. 5) **Collectif** : les activités réticulaires de contestation ou de solidarité, activités bénévoles qui prennent place dans des espaces militants non-organisés (militance en ligne, collectifs non-organisés, regroupement autour de lieux, réseaux informels, etc.). Nous faisons l'hypothèse que la façon de s'investir dans des « causes » de façon scientifique, salariée, organisée ou numérique offre un curseur sur différents axes : degré d'institutionnalisation, axe mainstream/radical, axe collectif/individuel, etc. sans investir idéologiquement ces termes. Cela nous permet de rassembler dans les entretiens des personnes situées très différemment sur l'échiquier politique (sur la dichotomie gauche/droite, universaliste/intersectionnaliste, légaliste/anti-autoritaire, etc.) avec pour point commun une appropriation des questions féministes ou de sexualités. Pour une description détaillée de la constitution du corpus et ses caractéristiques, voir Abbou et Burnett (à paraître).

Le corpus, recueilli entre 2021 et 2022, comptabilise 168 heures de parole et de 2 millions de mots, intégralement transcrits en transcription orthographique enrichie. Chaque entretien, d'une durée moyenne d'1h30, aborde trois phases distinctes : d'abord un récit biographique, puis une discussion sur différents aspects idéologiques du féminisme, et enfin une séquence consacrée au langage. Le corpus a été balisé au format XML, avec deux unités de structures : les tours de parole et les trois phases de l'entretien.

3 Étude variationniste du *ne* : méthodologie

3.1 Extraction et annotation des négations

Nous avons utilisé le logiciel TXM (Heiden et al., 2010) pour extraire les occurrences de neuf marqueurs négatifs : *aucun/aucune, guère, jamais, ni, nulle part, pas, personne, plus* et *rien*. Ces requêtes ont généré 40 043 concordances que nous avons annotées manuellement avec le codage suivant : 0 pour absence de *ne* et 1 pour présence d'un *ne*. Nous avons écarté plusieurs cas : les cas où aucun *ne* ne pouvait être employé (« il faut dire non-sexiste pas inclusif », « faut pas chercher »), les cas où *plus* n'est pas un adverbe négatif (« un mot plus adapté au contexte francophone »), et les cas où la présence ou l'absence d'un *ne* est neutralisée par la liaison (« on était pas beaucoup »). Dans les cas de co-occurrences de deux marqueurs de négation (« il y a plus rien »), qui ont généré deux concordances distinctes, nous avons supprimé les doublons. Après cette étape, il reste 25 539 négations dans le jeu de données.

3.2 Facteurs sociaux

Les personnes enregistrées sont âgées de 19 à 83 ans. L'âge médian est 35 ans (*écart interquartile (EI) = 22*) dans l'ensemble du corpus ; il est de 36.5 ans à Marseille (*EI = 17.5*), de 31 ans à Montréal (*EI = 12*) et de 38 ans à Paris (*EI = 32*). Nous avons réparti les locutrices et locuteurs dans trois groupes d'âge : moins de 35 ans (*n = 50*), 35 à 59 ans (*n = 33*), et plus de 60 ans (*n = 22*) (tableau 1). Pour les niveaux d'études, nous avons créé quatre groupes : baccalauréat français ou diplôme d'études collégiales québécois (8.57 % des personnes), licence française ou baccalauréat québécois (31.43 %), master ou maîtrise (34.29 %), et doctorat (25.71 %). Nous avons également intégré à notre étude un facteur « grande école », que nous considérons comme un indicateur de la classe sociale, vu que son admission reste, en très grande partie, réservée à l'élite économique et intellectuelle française (voir Pasquali 2021). Dans le corpus CaFé, 16 personnes (15.24 %) sont passées par une grande école française (Sciences Po, ENS, HEC, etc.). La majorité des locutrices passées par une grande école (12) sont dans le corpus parisien ; aucune n'est dans le corpus montréalais, le système d'études supérieures canadien n'ayant pas d'équivalent des grandes écoles françaises. La majorité des locutrices passées par une grande école (12) sont dans le corpus parisien ; aucune n'est dans le corpus montréalais¹. Enfin, le facteur « engagement » décrit les différentes pratiques féministes des personnes composant le corpus. Il est composé des cinq catégories structurant le corpus (voir section 2) : universitaire, professionnel, associatif, média et collectif. Ces catégories ont clairement un lien avec la catégorie socioprofessionnelle, et pourraient donc influencer l'emploi variable du *ne*.

Tableau 1. Facteurs sociaux pris en compte dans l'étude

| | Marseille | Montréal | Paris | Total |
|----------------------|-----------|----------|-------|-------|
| Groupes d'âge | | | | |
| 1 (moins de 35 ans) | 9 | 23 | 18 | 50 |
| 2 (35 à 59 ans) | 7 | 12 | 14 | 33 |
| 3 (plus de 60 ans) | 4 | 6 | 12 | 22 |
| Éducation | | | | |
| Bac | 3 | 5 | 1 | 9 |
| Licence | 6 | 16 | 11 | 33 |
| Master | 5 | 13 | 18 | 36 |
| Doctorat | 6 | 7 | 14 | 27 |
| Grande école | | | | |
| Non | 16 | 41 | 32 | 89 |
| Oui | 4 | 0 | 12 | 16 |

| Engagement | | | | |
|---------------|---|----|----|----|
| Associatif | 2 | 16 | 10 | 28 |
| Collectif | 3 | 7 | 5 | 15 |
| Media | 5 | 6 | 11 | 22 |
| Professionnel | 4 | 3 | 9 | 16 |
| Universitaire | 6 | 9 | 9 | 24 |

3.3 Facteurs linguistiques et discursifs

3.3.1 Travaux antérieurs

Nous avons pris en compte quatre facteurs linguistiques et discursifs dans notre étude : la nature du sujet grammatical (pronom ou syntagme nominal), la phase de l'entretien (début, milieu ou fin), la modalité d'adresse utilisée par l'intervieweuse (*tu* ou *vous*), et le fait que celle-ci connaisse ou pas les interviewé·es avant le début de l'entretien. Ces facteurs ont déjà été explorés à divers degrés par la littérature sur le *ne*.

L'influence du type de sujet (syntagme nominal ou pronom) sur la productivité du *ne* est bien documentée. Dans un corpus d'entretiens informels réalisés en 1975 dans le sud de la France, Diller (1983) note un taux de réalisation de 61 % pour les pronoms sujets contre 97 % pour les sujets lexicaux. L'effet de la nature du sujet est encore plus prononcé dans le corpus d'entretiens informels d'Hansen et Malderez (2004), avec 56.4 % de maintien du *ne* suivant un sujet nominal contre 5.8 % suivant un pronom. Armstrong et Smith (2002), qui remarquent aussi un taux de *ne* plus faible pour les pronoms que pour les autres sujets dans leurs deux corpus radiophoniques, font par ailleurs état d'un effet de fréquence : plus un pronom est fréquent, plus le *ne* chute. Ainsi, chez Meisner (2010), le *ne* est retenu à 87 % avec les sujets lexicaux et seulement à 15 % avec *je* et 6 % avec *ce*, qui sont les deux pronoms sujets les plus fréquents de son corpus. Avec des pronoms moins fréquents, le taux de réalisation est plus élevé : 42 % pour *vous* et 71 % pour *nous*, par exemple.

Les trois autres facteurs discursifs que nous étudions ici ont été plus rarement analysés. L'influence de la temporalité de l'entretien a été mesurée par Ashby (1981), dans un petit échantillon de trois personnes, enregistrées à deux moments et dans deux cadres différents. Dans la seconde partie de l'entretien, le taux de réalisation du *ne* a été divisé par deux chez les trois locuteur·ices. Armstrong (2002) a enregistré des adolescent·es dans deux contextes différents (entretien avec le chercheur et conversations entre adolescent·es), notant un taux légèrement plus élevé de la réalisation du *ne* dans les entretiens, mais uniquement chez les filles.

Le lien entre le tutoiement/vouvoiement et la productivité du *ne* a été décrit par Villeneuve (2017), dans un corpus d'interviews télévisées québécoises. La modalité d'adresse était un meilleur prédicteur de la production du *ne* que l'âge ou la catégorie professionnelle : l'autrice note un taux de réalisation de 28.6 % quand les journalistes utilisaient le *vous*, et de 7.8 % quand ils utilisaient le *tu*. Enfin, dans un corpus composé d'enregistrements réalisés dans divers contextes, Meisner (2010) note que le fait qu'une personne connaisse ses interlocuteurs et interlocutrices semble avoir un impact sur le taux de réalisation du *ne*, avec davantage d'omissions dans les conversations entre personnes qui se connaissent.

3.3.2 Codification des facteurs linguistiques et discursifs

La nature des sujets des verbes sur lesquels portent les négations a été annotée manuellement. Nous avons créé deux catégories : les sujets lexicaux et les sujets pronominaux, une catégorie qui inclut les sujets redoublés (« mes copines elles comprennent pas »). Nous avons réuni dans une troisième catégorie les cas où la négation porte sur un infinitif, un gérondif ou un impératif. Le jeu de données compte 1440 sujets lexicaux, 23 633 pronoms et 487 cas sans sujet.

L'annotation en phases de l'entretien, réalisée lors de la création du corpus au format XML, s'appuie sur le questionnaire utilisé : 1 pour la trajectoire biographique, 2 pour la discussion idéologique, et 3 pour la discussion métadiscursive. La phase 1 est la plus courte (377 418 mots) ; la phase 2, qui est la plus longue, contient 1 183 888 mots, et la dernière phase 649 239 mots.

L'annotation de la modalité d'adresse a été réalisée manuellement en inspectant les entretiens. L'intervieweuse tutoie 74 personnes et vouvoie 27 personnes tout au long de l'entretien (Tableau 2). Dans quatre cas, elle passe du *vous* au *tu* au cours d'entretien ; pour ces cas, nous avons utilisé les repères temporels présents dans les transcriptions pour identifier le passage du *vous* au *tu*, puis annoté manuellement les occurrences de négations en fonction de leur position dans l'entretien. Enfin, l'annotation des personnes connues ou inconnues de l'intervieweuse a été réalisée grâce aux informations communiquées par celle-ci. Avant l'entretien, elle connaissait à divers degrés 16 personnes sur 105, dont 11 à Marseille, ville dont elle est originaire.

Tableau 2. Modalités d'adresse et personnes connues de la chercheuse qui a réalisé les entretiens, par ville

| | Marseille | Montréal | Paris | Total |
|---------------------------------------|-----------|----------|-------|-------|
| Tutoiement / vouvoiement | | | | |
| Usage exclusif du tutoiement | 19 | 36 | 19 | 74 |
| Usage exclusif du vouvoiement | 1 | 4 | 22 | 27 |
| Passage du <i>vous</i> au <i>tu</i> | 0 | 1 | 3 | 4 |
| Connaissances de la chercheuse | | | | |
| Personnes non connues | 9 | 39 | 41 | 89 |
| Personnes connues | 11 | 2 | 3 | 16 |
| Total | 20 | 41 | 44 | 105 |

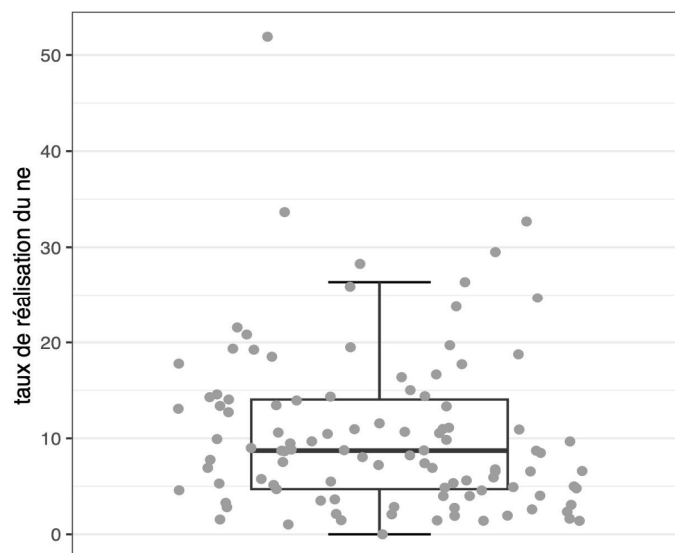
4 Résultats

4.1 Résultats descriptifs

4.1.1 Résultats généraux

Sur les 25 560 négations pour lesquelles la réalisation d'un *ne* était possible, nous avons relevé 2645 occurrences du *ne*, soit un taux de réalisation global de 10.35 %. Le taux de réalisation moyen par personne est de 10.49 % ($ET= 8.47$) ; la médiane est de 8.72 % ($EI = 9.33$). Les valeurs s'étendent de 0 % à 51.90 % (Graphique 1). Cinq personnes se démarquent par un taux élevé de réalisation du *ne* et peuvent être considérées comme des « outliers » sur le plan statistique. La personne qui a retenu le plus de *ne* (76 omissions pour 82 réalisations) est une femme parisienne de 67 ans. Les autres taux de réalisation importants sont ceux de deux Montréalaises dont la langue maternelle n'est pas le français (50 et 83 ans, taux : 33.66 % et 32.68 %), d'une Marseillaise (50 ans, taux : 29.50 %) et d'une Parisienne (70 ans, taux : 28.28 %). Les trois francophones natives sont titulaires d'un doctorat et les deux non-natives d'une licence et d'un diplôme d'école secondaire. À l'opposé, un quart des personnes enregistrées ont un taux de réalisation inférieur à 4.71 %. Une personne n'a pas utilisé de *ne* du tout (245 omissions) : il s'agit d'une

femme montréalaise de 40 ans (niveau d'éducation 1). Neuf autres personnes affichent un taux de réalisation très faible, inférieur à 2 %. La plus âgée a 40 ans, et la plus jeune 19 ans ; l'âge médian de ces personnes est de 26.5 ans. Trois sont dans le corpus parisien et sept dans le corpus montréalais.



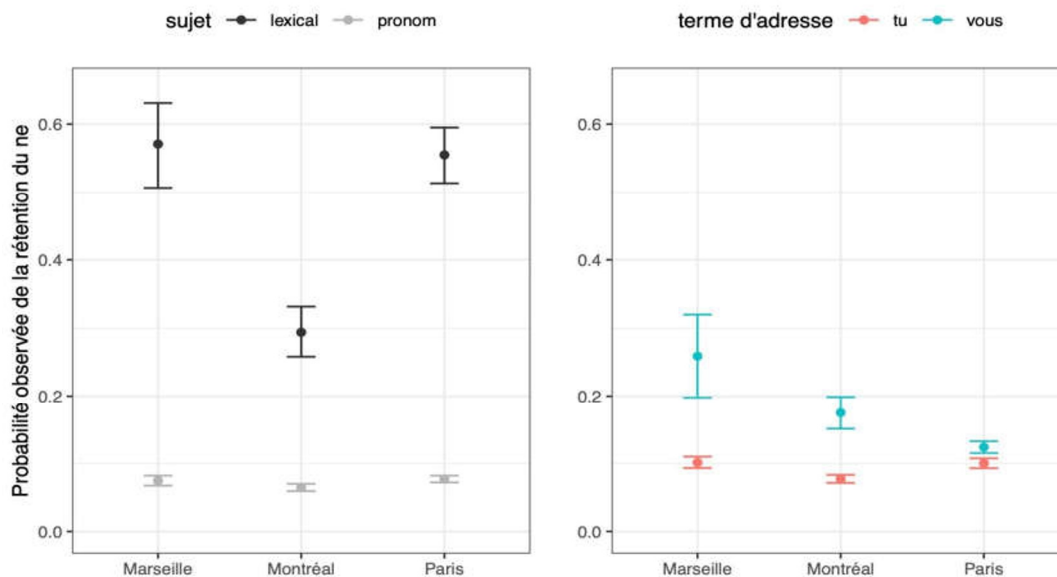
Graphique 1. Taux de réalisation du *ne*, exprimé en pourcentage, par personne

4.1.2 Facteurs linguistiques et discursifs

Tout d'abord, le taux de réalisation du *ne* varie en fonction du mot de négation employé. L'adverbe *pas* est de loin le mot de négation le plus fréquent dans le corpus (22 165 occurrences) ; le taux de réalisation du *ne* est de 7.14 % (2160 réalisations). Viennent ensuite *plus* (1185 occurrences, taux de *ne* : 14.26 %), *jamais* (1032 occurrences, taux de *ne* : 11.34 %), *rien* (678 occurrences, taux de *ne* : 13.86 %), *aucun* (241 occurrences, taux de *ne* : 20.33 %), *personne* (212 occurrences, taux de *ne* : 24.53 %), *ni* (42 occurrences, taux de *ne* : 7.14 %), *nulle part* (4 occurrences, aucun *ne* réalisé) et *guère* (1 occurrence seulement, avec *ne* réalisé).

Ensuite, le taux de réalisation du *ne* est fortement influencé par le type de sujet. Il y a 656 *ne* pour 784 omissions pour les sujets lexicaux (n = 1440), soit 45.56 % de réalisation, 1710 *ne* pour 21 923 omissions pour les sujets pronominaux (n = 23 633), soit 7.24 % de réalisation, et 279 réalisations pour 208 omissions pour les cas sans sujet (n = 487), soit un taux de réalisation de 42.71 %. Le graphique 2 (A), qui présente les probabilités observées de la réalisation du *ne* pour les sujets lexicaux et les pronoms par ville, met en lumière l'écart important entre sujets lexicaux et pronoms. Cet écart est plus faible dans le corpus montréalais que dans les deux autres corpus.

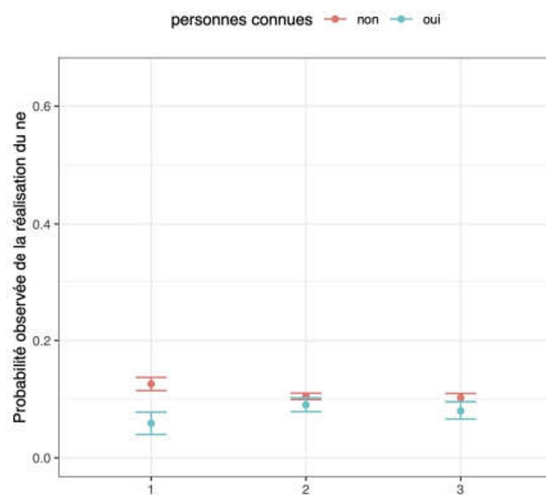
Sur les 18 410 négations dans le corpus qui ont été employées dans des entretiens avec des personnes que la chercheuse a tutoyées (ou des parties d'entretiens, lorsqu'il y a un passage du *vous* au *tu*), le *ne* a été produit 1682 fois et omis 16 728 fois, soit un taux de réalisation de 9.14 %. Ce taux est plus élevé dans les entretiens utilisant le vouvoiement : il est de 13.47 % (963 réalisations pour 6187 omissions). Le graphique 2 (B) suggère que cette tendance est présente dans les trois corpus.



Graphique 2. Probabilité observée de la réalisation du *ne* par type de sujet et par ville (A), et en fonction du terme d'adresse, par ville (B)

21 512 négations ont été produites par les 89 personnes non connues de la chercheuse, et 4048 négations par les 16 personnes qu'elle connaissait. Le taux de réalisation global est légèrement plus faible dans les entretiens réalisés avec des personnes connues : il est de 8.33 % (337 réalisations pour 3711 omissions). Il est de 10.73 % (2308 réalisations pour 19 204 omissions) dans les entretiens avec les personnes non connues de la chercheuse.

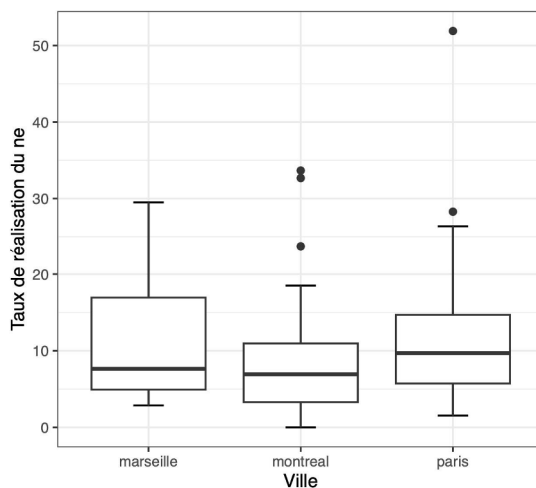
Au cours de l'entretien, le taux de réalisation global du *ne* baisse légèrement, passant de 11.67 % pendant la première phase (440 réalisations/3330 omissions) à 10.24 % dans la seconde phase (1420 réalisations /12443 omissions), puis à 9.90 % dans la dernière phase (785 réalisations/7142 omissions). En explorant l'évolution du taux de réalisation du *ne* au cours des trois phases, nous avons remarqué la présence d'une possible interaction entre ce facteur et le fait que la chercheuse connaisse ou non les personnes qu'elle a enregistrées (Graphique 3). Dans la première phase de l'entretien, consacrée à la trajectoire biographique, les personnes connues ont moins réalisé le *ne* que les personnes non connues. Dans la seconde phase, consacrée à des thématiques moins personnelles (le féminisme), le taux de réalisation augmente chez les personnes connues, mais il baisse chez les personnes non connues. Durant la troisième phase, la probabilité de retenir le *ne* reste plus bas chez les personnes connues de la chercheuse.



Graphique 3. Probabilité observée de la réalisation du *ne* pour les personnes connues et non connues de la chercheuse, par phase de l'entretien

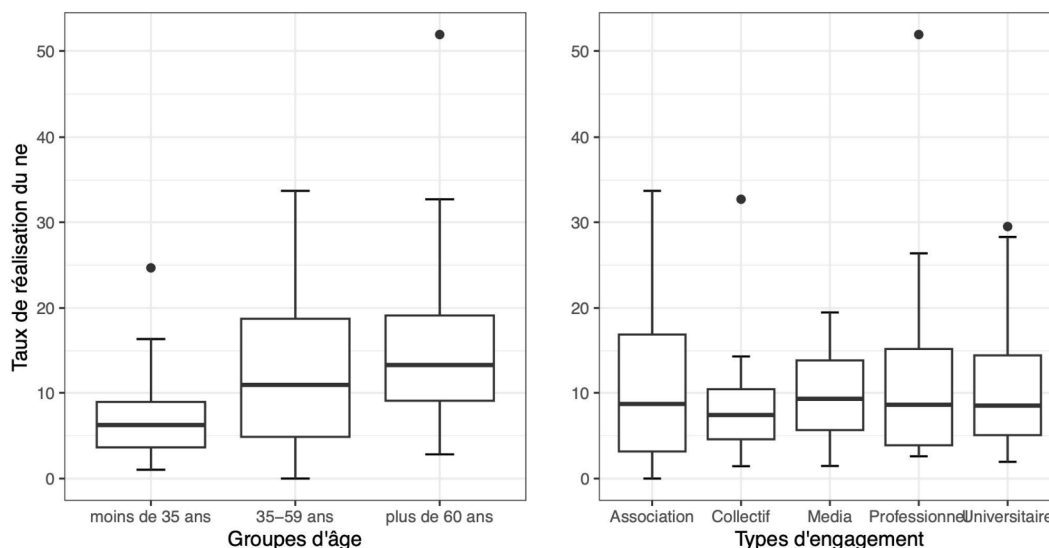
4.1.3 Facteurs sociaux

Les taux de réalisation globaux sont de 11.22 % pour Paris (1351 réalisations/10 693 omissions), de 10.75 % pour Marseille (543 réalisations/4508 omissions) et de 8.87 % pour Montréal (751 réalisations/7715 omissions). Les taux médians, par personne, sont de 7.64 % à Marseille ($EI = 12.01$), 6.92 % à Montréal ($EI = 7.66$) et 9.68 % à Paris ($EI = 8.95$) (Graphique 4).



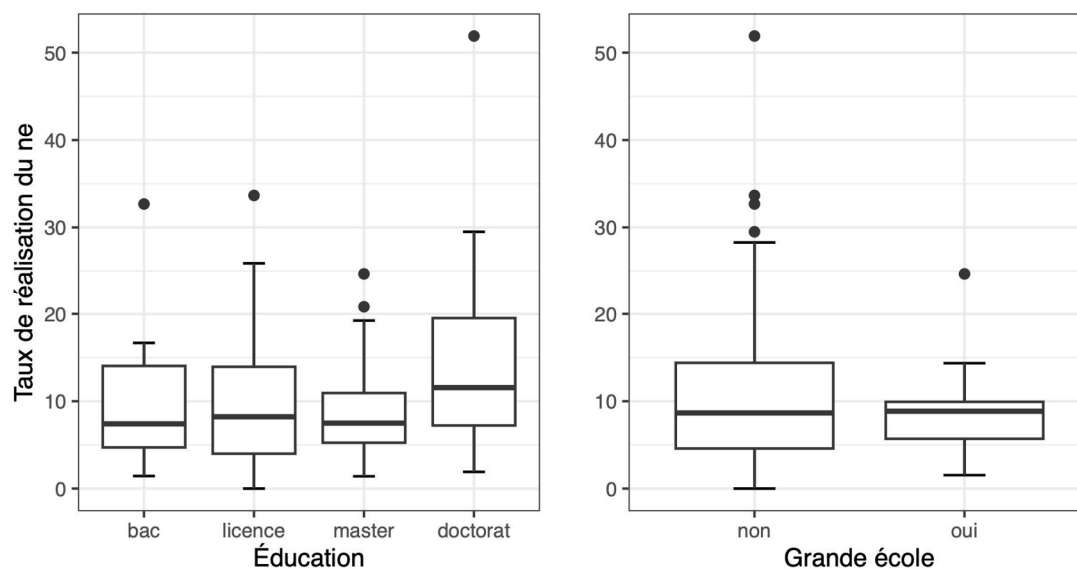
Graphique 4. Taux de réalisation du *ne* par ville, exprimé en pourcentage

L'âge semble être corrélé positivement à la production de *ne* dans les négations. Le taux de réalisation médian augmente d'un groupe d'âge à l'autre : il est de 6.24 % pour les moins de 35 ($EI = 5.29$), de 10.92 % pour les 35-59 ans ($EI = 13.89$), et de 13.36 % pour les plus de 60 ans ($EI = 10.05$). Il ne semble en revanche pas y avoir de lien entre le type d'engagement féministe et le taux de réalisation du *ne*, qui est proche dans les cinq groupes. Il est compris entre 7.41 % ($EI = 5.84$) pour les personnes ayant un engagement collectif et 9.30 % ($EI = 8.24$) pour les personnes ayant un engagement dans les médias (Graphique 5).



Graphique 5. Taux de réalisation du *ne* par groupes d'âge et par types d'engagement féministe

En ce qui concerne le niveau d'éducation, un seul groupe se détache : il s'agit des personnes titulaires d'un doctorat, qui ont un taux de réalisation médian de 11.56 % ($EI = 12.30$), alors que les médianes des trois autres groupes sont similaires (7.41 % pour le bac/DEC, 8.22 % pour la licence/bac, et 7.49 % pour le master/maitrise). Enfin, le fait d'être passé.e par une grande école ne semble pas être corrélé avec la réalisation du *ne*, avec des médianes très similaires dans les deux groupes : 8.66 % pour les personnes qui ne sont pas allées dans une grande école ($EI = 9.80$) et 8.86 % pour les personnes qui y sont allées ($EI = 4.23$) (Graphique 6).



Graphique 6. Taux de réalisation du *ne* par niveau d'éducation et pour le facteur « grande école »

4.2 Analyses statistiques

Pour savoir quels facteurs ont un effet significatif sur la réalisation du *ne*, nous avons construit un modèle de régression logistique à effets aléatoires (Tableau 3) avec le package R lme4 (Bates et al., 2015). L'intégration d'effets aléatoires correspondant à chaque personne a pour objectif de neutraliser la variation liée à des idiosyncrasies pour mettre en lumière la variation liée aux facteurs qui nous intéressent. Nous avons également intégré un effet aléatoire correspondant au type de mot négatif (*pas*, *plus*, *aucun*, etc.), car étudier la corrélation entre ce facteur et la réalisation du *ne* n'était pas le but de notre étude. Les prédicteurs du modèle sont les variables sociales, linguistiques et discursives présentées dans les sections précédentes, avec deux interactions : l'interaction entre la ville et le type de sujet, et l'interaction entre les personnes connues de la chercheuse et la phase de l'entretien. Les prédicteurs ayant un effet significatif ($p < 0.05$) sont indiqués en gras.

La taille d'effet est indiquée par les « odds ratios » (rapports des chances). Quand ceux-ci sont inférieurs à 1, la probabilité de réalisation du *ne* est plus faible pour le groupe indiqué que pour le groupe de référence. Les groupes de référence du modèle sont Montréal, les moins de 35 ans, le tutoiement, les pronoms, le niveau d'éducation « BAC/DEC », les personnes qui ne sont pas allées dans des grandes écoles, l'engagement associatif, la phase 1 de l'entretien et les personnes non connues de la chercheuse. Pour interpréter pleinement le modèle (c'est-à-dire, pour effectuer toutes les comparaisons possibles entre les groupes), nous avons changé ses niveaux de référence plusieurs fois.

Le modèle montre que le niveau d'éducation, le fait d'être passé.e par une grande école, le type d'engagement féministe et la modalité d'adresse (*tu* ou *vous*) n'ont pas d'effet significatif sur le taux de réalisation du *ne*. L'âge a un effet partiel : les 35-59 ans et les plus de 60 ans réalisent plus le *ne* que les moins de 35 ans ($OR = 1.64$ et 2.16), mais il n'y a pas de différence significative entre les deux groupes les plus âgés.

Le facteur qui a la taille d'effet la plus importante est le type de sujet. À Montréal, par exemple, la probabilité qu'un *ne* soit réalisé est bien plus élevée après un syntagme nominal qu'après un pronom ($OR = 7.50$), ou quand la négation porte sur un infinitif, gérondif ou impératif ($OR = 25.16$). La ville a également un effet significatif, qu'il faut interpréter en prenant en compte son interaction avec le type de sujet. Pour les pronoms, la probabilité de réalisation du *ne* est plus élevée à Marseille qu'à Montréal ($OR = 1.95$), mais il n'y a pas de différence significative entre Paris et Montréal, ni entre Paris et Marseille). Pour les sujets lexicaux, Montréal se démarque de Paris et Marseille, avec une probabilité plus élevée de la réalisation du *ne* pour ces deux villes (OR respectifs = 3.06 et 5.21 , $p < 0.001$).

Les personnes qui connaissaient déjà la chercheuse avant l'entretien réalisent moins le *ne* que les personnes qui ne la connaissaient pas ($OR = 0.25$), et ce pendant toutes les phases de l'entretien. Enfin, l'interaction entre la phase de l'entretien et le fait que les personnes enregistrées connaissent la chercheuse ou non est significative. Pour les personnes non connues de la chercheuse, la phase de l'entretien n'a pas d'effet : le taux de réalisation du *ne* reste constant tout au long de l'entretien. En revanche, chez les personnes que la chercheuse connaissait, le taux de réalisation change : il est significativement plus élevé dans la seconde et la troisième partie de l'entretien que durant la première.

Tableau 3. Modèle de régression logistique avec effets aléatoires

| <i>Prédicteurs</i> | <i>Odds Ratios</i> | <i>Intervalles de confiance</i> | <i>Valeur p</i> |
|--|--------------------|---------------------------------|-----------------|
| (Intercept) | 0.06 | 0.03 – 0.13 | < 0.001 |
| Ville [Marseille] | 1.95 | 1.14 – 3.35 | 0.015 |
| Ville [Paris] | 1.20 | 0.78 – 1.86 | 0.405 |
| Type de sujet [lexical] | 7.50 | 5.99 – 9.41 | < 0.001 |
| Type de sujet [sans] | 25.16 | 17.78 – 35.61 | < 0.001 |
| Groupe d'âge [35-59 ans] | 1.64 | 1.09 – 2.47 | 0.019 |
| Groupe d'âge [plus de 60 ans] | 2.16 | 1.30 – 3.58 | 0.003 |
| Forme d'adresse [vous] | 1.01 | 0.70 – 1.46 | 0.949 |
| Grande école [oui] | 1.07 | 0.63 – 1.84 | 0.793 |
| Niveau d'éducation [licence] | 0.67 | 0.35 – 1.30 | 0.238 |
| Niveau d'éducation [master] | 0.77 | 0.39 – 1.50 | 0.440 |
| Niveau d'éducation [doctorat] | 1.28 | 0.58 – 2.80 | 0.540 |
| Engagement [collectif] | 0.77 | 0.44 – 1.36 | 0.371 |
| Engagement [média] | 1.09 | 0.67 – 1.79 | 0.720 |
| Engagement [professionnel] | 1.02 | 0.60 – 1.73 | 0.950 |
| Engagement [universitaire] | 0.77 | 0.43 – 1.37 | 0.378 |
| Phase [2] | 0.93 | 0.80 – 1.08 | 0.315 |
| Phase [3] | 0.91 | 0.77 – 1.06 | 0.214 |
| Personnes connues [oui] | 0.25 | 0.12 – 0.51 | < 0.001 |
| Ville [Marseille] × Type de sujet [lexical] | 2.67 | 1.82 – 3.92 | < 0.001 |
| Ville [Paris] × Type de sujet [lexical] | 2.54 | 1.90 – 3.42 | < 0.001 |
| Ville [Marseille] × Type de sujet [sans] | 0.77 | 0.43 – 1.38 | 0.381 |
| Ville [Paris] × Type de sujet [sans] | 0.85 | 0.54 – 1.34 | 0.494 |
| Phase [2] × Personnes connues [oui] | 1.81 | 1.14 – 2.86 | 0.012 |
| Phase [3] × Personnes connues [oui] | 1.89 | 1.16 – 3.09 | 0.011 |

4.3 Discussion

Comme l'ont également montré de nombreuses études (Armstrong et Smith 2002, Diller 1983, Hansen et Malderez 2004, Meisner 2010, Villeneuve 2017), nos analyses indiquent que le type de sujet est le facteur qui a l'effet le plus fort : la probabilité qu'un *ne* soit réalisé est plus élevée pour les sujets lexicaux que pour les sujets pronominaux. Pour comprendre cet effet, il faut savoir qu'en français parlé, les sujets lexicaux sont très souvent redoublés par des pronoms clitiques, à tel point que plusieurs chercheur.e.s considèrent qu'il s'agit de marqueurs de l'accord verbal. Cette analyse des sujets pronominaux a été proposée pour le français parisien (Culbertson 2010, Liang 2023) et québécois (Roberge 1990, Auger 1994). Si les pronoms clitiques dans le corpus CaFé sont (très souvent) des marqueurs morphologiques d'accord, on s'attend à ce que le *ne*, un clitique syntaxique, ne puisse pas (ou peu) apparaître à l'intérieur de cette morphologie verbale. Par contre, dans les phrases que nous avons codées comme ayant un sujet lexical, il n'y a pas de marqueur d'accord avec le sujet, et le *ne* ne serait donc pas défavorisé dans ce contexte.

Ensuite, l'intégration de l'interaction entre ville et type de sujet dans le modèle de régression met en lumière une variation dialectale plus importante pour les sujets lexicaux que pour les pronoms. Si les sujets lexicaux restent des lieux privilégiés de la réalisation du *ne* dans toutes les villes, c'est davantage le cas à Paris et à Marseille qu'à Montréal. Cela suggère que, si l'omission du *ne* est un changement en cours, celui-ci a davantage progressé au Québec qu'en France.

Nos résultats révèlent toutefois une variation géographique moins prononcée que ce que l'on aurait pu attendre, au vu du très faible taux de réalisation du *ne* (0.5 %) trouvé par Sankoff et Vincent (1977) dans le corpus montréalais Sankoff-Cedegren et les taux plus importants (chez les adultes) rapportés par plusieurs études françaises de corpus informels : 15.8% en 1972-74 et 8.2% dans les années 1990 chez Hansen et Malderez (2004), 11 % chez Donaldson (2017) ou encore 18 % chez Ashby (2001). Grâce à la façon dont il a été conçu, le corpus CaFé permet une comparaison directe qui, dans le cas du *ne*, fournit une image nuancée des différences dialectales entre France et Québec, tout du moins pour les personnes très scolarisées.

Notre étude a deux résultats qui peuvent paraître surprenants : l'absence d'effet de l'éducation ou du facteur « grande école », qui contraste avec les études précédentes ayant remarqué une corrélation entre niveau d'étude, classe sociale et réalisation du *ne* (Ashby 2001, Buson et al., 2022, Coveney 1996, Diller 1983, Sankoff et Vincent 1977), et le taux de réalisation du *ne* à Montréal (8.87 %), bien plus élevé que celui trouvé par Sankoff et Vincent (1977). Ces résultats s'expliquent sans doute par le fait que les personnes qui composent le corpus CaFé sont, pour la plupart, très éduquées, ce qui est lié à la thématique du corpus (l'engagement féministe, qui a souvent pour cadre des sphères universitaires et médiatiques) : seules neuf personnes sur 105 n'ont pas obtenu de diplôme universitaire. Le taux de réalisation élevé du *ne* à Montréal dans CaFé peut par ailleurs s'expliquer par les changements intervenus dans la société québécoise : les Québécois.es sont aujourd'hui beaucoup plus scolarisés.es qu'en 1971, date de réalisation du corpus Sankoff-Cedegren. Il est possible que cette évolution ait un impact sur les pratiques langagières québécoises, qui sont peut-être plus « normées » aujourd'hui qu'il y a 50 ans, ce dont témoignerait le taux relativement élevé de la réalisation du *ne* dans le corpus CaFé. Il est à noter que d'autres travaux sur le français montréalais ont observé une progression de variantes de prestige dans des corpus plus récents, comparé au corpus Sankoff-Cedegren (voir Blondeau et al. 2019). Il est donc possible que la démocratisation de l'éducation supérieure en français au Québec ait entamé une convergence sur une norme plus internationale, au moins pour les locuteurs et locutrices qui ont des diplômes supérieurs.

Le fait que le *ne* soit moins réalisé lorsque les personnes enregistrées connaissent la chercheuse appuie les résultats de Meisner (2010). Cet effet interagit avec la phase de l'entretien, ce qui suggère un effet des thématiques abordées : le *ne* est davantage réalisé quand les personnes abordent des thématiques impersonnelles (l'engagement féministe, le langage). Cela fait écho à Sankoff et Vincent (1977), qui remarquent également une réalisation plus élevée du *ne* avec des sujets sérieux (la langue, la religion, ou l'éducation des enfants).

5 Conclusion

Grâce à sa taille, à sa composition et à la richesse des informations sociodémographiques qui l'accompagnent, le corpus CaFé fournit plusieurs apports à notre connaissance de la variation de la négation en français parlé contemporain. Tout d'abord, notre étude met en évidence la complexité du phénomène, liée de façon significative à des facteurs linguistiques (type de sujet, et bien entendu d'autres, non explorés ici), à des facteurs discursifs (connaissance de l'interlocutrice, sujets abordés) et à des facteurs sociaux (âge, ville). Elle déconstruit par ailleurs en partie la différence dialectale France/Québec dans la négation, en tout cas chez les personnes très scolarisées. Sur le plan méthodologique, elle montre qu'il est important d'explorer les effets éventuels d'interactions entre plusieurs variables, qui peuvent révéler d'intéressantes nuances : la différence dialectale France/Québec, limitée à un contraste entre Marseille et Montréal pour les pronoms, est plus prononcée pour les sujets lexicaux. Notre étude déconstruit aussi l'aspect « formalité » de la variation en mettant en lumière l'interaction entre le statut « connaissance de l'interlocutrice » et les thématiques abordées dans l'entretien : la formalité de la thématique abordée a un effet plus important si la locutrice et l'interlocutrice se connaissent. Plus généralement, notre étude permet de constater que le *ne* de la négation est toujours bien vivant, y compris au Québec, et qu'il reste toujours beaucoup d'aspects sociaux et stylistiques à étudier en lien avec cette variable sociolinguistique classique.

Financement

Ce travail de recherche a reçu un financement de l'ERC dans le cadre du programme de financement de la recherche et de l'innovation Horizon 2020 de l'Union européenne (accord de subvention n°850539).

Références bibliographiques

- Abbou, J., & H. Burnett. (2024). Devenir féministe et queer à Paris et Montréal: Récits de vie dans le corpus CaFé. À paraître dans H. Blondeau, M. Laforest & W. Remysen (dir). *Actes du Colloque 50 ans de corpus montréalais*. Montréal : Presses Universitaires de Montréal.
- Armstrong, N. (2002). Variable deletion of French *ne* : A cross-stylistic perspective. *Language Sciences*, 24(2), 153-173. [https://doi.org/10.1016/S0388-0001\(01\)00015-8](https://doi.org/10.1016/S0388-0001(01)00015-8)
- Armstrong, N., & Smith, A. (2002). The influence of linguistic and social factors on the recent decline of French *ne*. *Journal of French Language Studies*, 12. <https://doi.org/10.1017/S0959269502000121>
- Ashby, W. J. (1976). The loss of the negative morpheme, *ne*, in Parisian French. *Lingua*, 39(1-2), 119-137. [https://doi.org/10.1016/0024-3841\(76\)90060-7](https://doi.org/10.1016/0024-3841(76)90060-7)
- Ashby, W. J. (1981). The loss of the negative particle *ne* in French : A syntactic change in progress. *Language*, 57(3), 674-687. <https://doi.org/10.2307/414345>
- Ashby, W. J. (2001). Un nouveau regard sur la chute du *ne* en français parlé tourangeau : S'agit-il d'un changement en cours? *Journal of French Language Studies*, 11(1), 1-22. <https://doi.org/10.1017/S0959269501000114>
- Auger, J. (1994). *Pronominal clitics in Quebec colloquial French: A morphological analysis*. Thèse de doctorat. University of Pennsylvania.
- Avanzi, M., Béguelin, M.-J., & Diémoz, F. (2016). De l'archive de parole au corpus de référence : La base de données orales du français de Suisse romande (OFROM). *Corpus*, 15. <https://doi.org/10.4000/corpus.3060>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Blondeau, H., Mougeon, R., & Tremblay, M. (2019). Analyse comparative de *ça fait que, alors, donc* et *so* à Montréal et à Welland: mutations sociales, convergences, divergences en français laurentien. *Journal of French Language Studies*, 29(1), 35-65.

- Buson, L., Nardy, A., Rousset, I., & Zhang, C. (2022). Pratiques stylistiques en contexte scolaire : Usages contrastés du « ne » de négation chez les adultes et les enfants d'une école maternelle française. *SHS Web Conf. 8e Congrès Mondial de Linguistique Française*, 138, 06018. <https://doi.org/10.1051/shsconf/202213806018>
- Coveney, A. (1996). *Variability in spoken French : A sociolinguistic study of interrogation and negation*. Exeter : Elm Bank.
- Culbertson, J. (2010). Convergent evidence for categorial change in French: From subject clitic to agreement marker. *Language*, 85-132.
- Diller, A.-M. (1983). Subject NP structure and variable constraints : The case of *ne* deletion. In R. W. Fasold (Éd.), *Variation in the form and use of language. A sociolinguistics reader* (p. 167-175). Georgetown : Georgetown University Press.
- Donaldson, B. (2017). Negation in near-native French : Variation and sociolinguistic competence. *Language Learning*, 67(1), 141-170. <https://doi.org/10.1111/lang.12201>
- Fonseca-Greber, B. B. (2007). The emergence of emphatic 'ne' in conversational Swiss French. *Journal of French Language Studies*, 17(3), 249-275. <https://doi.org/10.1017/S0959269507002992>
- Gadet, F. (2000). Des corpus pour (ne)... pas. In M. Bilger (Éd.), *Corpus. Méthodologie et applications linguistiques* (p. 156-167). Paris : Honoré Champion.
- Garel, C. (1997). Étude de la négation dans la langue française à travers le langage des jeunes de 17-19 ans [mémoire de maîtrise].. Université de Paris X Nanterre.
- Hansen, A., & Malderez, I. (2004). Le ne de négation en région parisienne : Une étude en temps réel. *Langage et société*, 1(107), 5-30. <https://doi.org/10.3917/lis.107.0005>
- Heiden, S., Magué, J.-P., & Pincemin, B. (2010). TXM : Une plateforme logicielle open-source pour la textométrie—Conception et développement. *10th International Conference on the Statistical Analysis of Textual Data. JADT 2010*, Rome.
- Jespersen, O. (1917). *Negation in English and other languages*. København : Høst.
- Liang, Y. (2023). *Quantitative syntax, formal syntax and information theory: Bridging gaps by studying French variation*. Thèse de doctorat. Université Paris Cité.
- Martineau, F., & Mougeon, R. (2003). A sociolinguistic study of the origins of *ne* deletion in European and Quebec French. *Language*, 79(1), 118-152. <https://doi.org/10.1353/lan.2003.0090>
- Meisner, C. (2010). A corpus analysis of intra-and extralinguistic factors triggering *ne*-deletion in phonic French. In F. Neveu, V. Muni Toke, J. Durand, T. Klingler, L. Mondada, & S. Prévost (Éds.), *Congrès Mondial de Linguistique Française—CMLF 2010*. Paris: Institut de Linguistique Française. <https://doi.org/10.1051/cmlf/2010091>
- Meisner, C., & Pomino, N. (2014). Synchronic variation in the expression of French negation : A Distributed Morphology approach. *Journal of French Language Studies*, 24(1), 9-28. <https://doi.org/10.1017/S0959269513000355>
- Pasquali, P. (2021). *Héritocratie: Les élites, les grandes écoles et les mésaventures du mérite (1870-2020)*. La Découverte.
- R Core Team. (2021). *R: A language and environment for statistical computing*. <https://www.R-project.org/>
- Roberge, Y. (1990). *Syntactic recoverability of null arguments*. McGill-Queen's Press-MQUP.
- Sankoff, G., & Vincent, D. (1977). L'emploi productif du *ne* dans le français parlé à Montréal. *Français (Le) Moderne Paris*, 45(3), 243-256.
- Villeneuve, A. J. (2017). Normes objectives et variation socio-stylistique: le français québécois parlé en contexte d'entrevues télévisées. *Arborescences*, (7), 49-66.

ⁱ Même si le système d'études supérieures canadien ne possède pas l'équivalent des grandes écoles françaises, l'intégration de ce facteur reste pertinente : en effet, certaines personnes du corpus montréalais ont été scolarisées en

France, et il est courant pour des Québécois.e.s de passer du temps en France pendant leurs études supérieures. Il aurait donc été tout à fait possible que certaines personnes du corpus montréalais soient passées par une grande école.