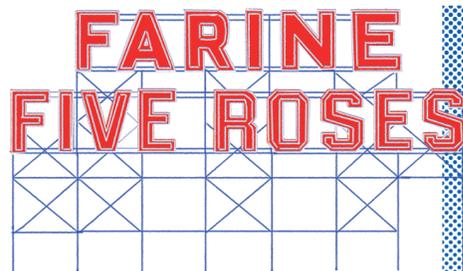
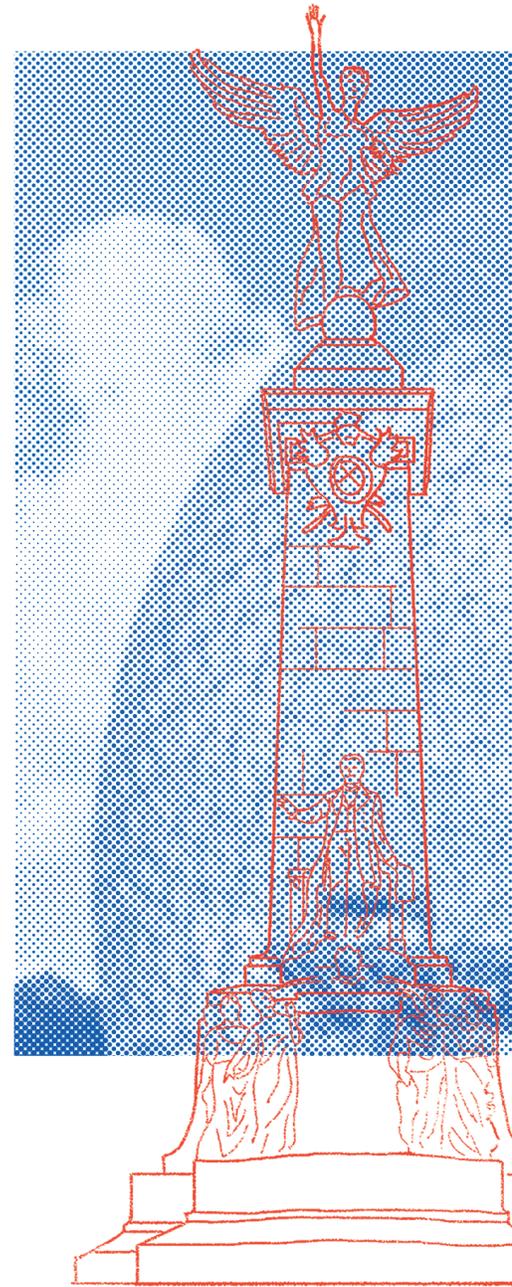
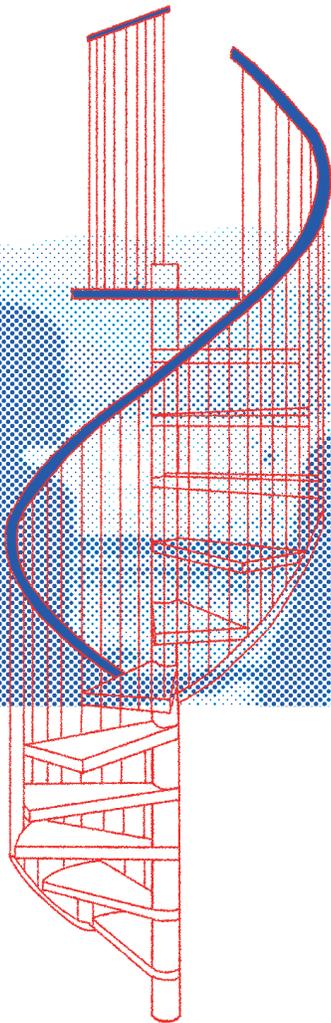


Colloque international

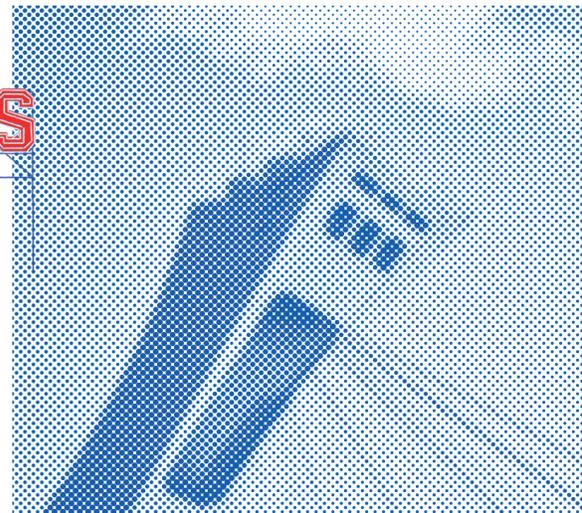
15-17 septembre 2022

1971-2021 50 ANS DE CORPUS MONTRÉALAIS

Archives nationales du Québec à Montréal (BAnQ Vieux-Montréal)
535, avenue Viger Est, Montréal



Organisé par
HÉLÈNE BLONDEAU
MARTY LAFOREST
WIM REMYSEN



Les corpus du Québec dans la Grande Grammaire du français

Anne Abeillé (Université de Paris-Cité) et Danièle Godard (Université de Paris-Cité)

Un des objectifs de la *Grande Grammaire du français* (GGF, Abeillé et Godard, 2021) a été de tenir compte des différents usages du français contemporain (écrit et oral, Internet, et différentes régions) en les intégrant dans une approche cohérente. On sait que les langues ne sont pas homogènes, ni géographiquement, ni socialement (Labov, 2006). Les grammaires du français, dans l'ensemble, se sont concentrées sur le français de France, et plus particulièrement sur l'usage standard ou normé. Par ailleurs, des ouvrages ont décrit le français parlé au Québec, lexique (e.g. Forest, 1996) et morphologie et syntaxe (e.g. Léard, 1995). La GGF s'appuie sur 34 grands corpus, ainsi que sur les études qui en sont issues. Pour le Québec, il s'agit des corpus de Montréal (Montréal 1971, 1984, voir Sankoff *et al.*, 1976 ; Thibault et Vincent, 1990), Ottawa-Hull (Poplack, 1996), le *Corpus du français parlé au Québec* (CFPQ) (Dostie, 2016). La GGF, qui a rassemblé 59 linguistes de 10 pays différents, dont 3 du Québec, participe donc à la valorisation des corpus québécois ainsi qu'à une meilleure connaissance de la variété des usages.

La GGF est descriptive et s'adresse à un large public. Elle prend en compte tous les usages (écrit, oral, Internet), et s'appuie sur 30 000 exemples écrits et oraux, postérieurs à 1950, non sourcés quand ils sont très courants (*Ça va ? C'est magnifique !*), annotés selon plusieurs critères : le jugement des locuteurs (acceptable 1a, inapproprié 1b, inacceptable 1c, variable 1d et non standard 1e) l'origine géographique, la source éventuelle.

- 1a. *J'ai un vélo.*
- 1b. *# J'ai un vélo et Marie non plus.*
- 1c. ** J'ai de vélo.*
- 1d. *% J'ai personne vu. (Franche-Comté, Jura)*
- 1e. *! C'est le livre que je t'ai parlé.*

Les exemples sourcés (plus de 5000) sont issus de grands corpus, par exemple les corpus d'Orléans, le CEFC, PFC et ceux du Québec.

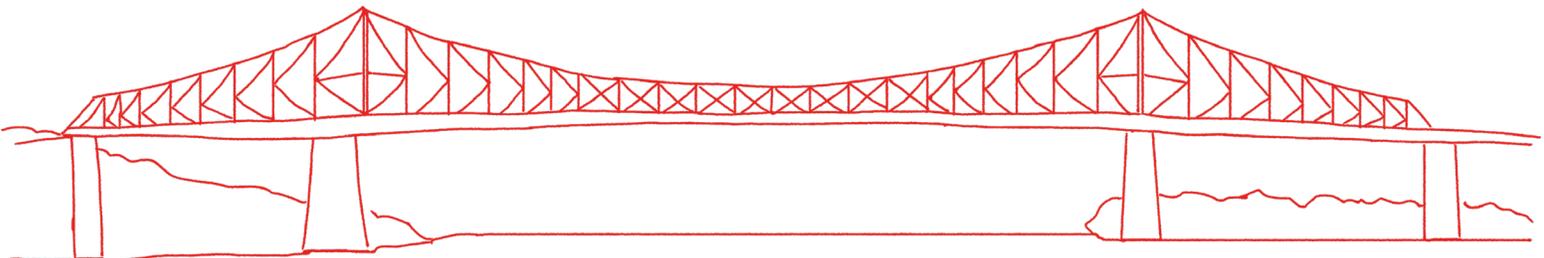
Les spécificités du français du Québec se laissent intégrer de différentes manières. D'une part la variation lexicale : « sacres » comme noms affectifs (*% ma crise de propriétaire*) ou comme particules de discours *% Câlisse, qui c'est qui paie ?* (Montréal 1984, loc. 2) ; création de connecteurs *% fait que, % fak, (% C'est dans le golfe du Saint-Laurent, fait que c'est pas la mer, Montréal 1984, loc. 4) ; pis* comme conjonction de coordination (*% Il va y avoir dérèglement des pôles, genre le Nord pis le Sud, ils vont changer, CFPQ 17, segment 3) ; la particule -tu* interrogative et exclamative (*% Le gouvernement te donne-tu beaucoup d'argent ?, CFPQ, 19, segment 1, p. 5).*

D'autre part, variation syntaxique, avec création d'une construction, comme l'infinitif à interprétation conditionnelle (*% Rester au bord du lac, je pêcherais plus souvent, Villiard et Vinet 1983).*

Nous avons travaillé sur les exemples extraits avec des collègues québécois pour savoir si le tour était stigmatisé (noté !) ou non (noté %), s'il était en progression (particule -tu) ou en régression.

Nous avons comparé l'évolution à celle de l'hexagone : omission de *ne* plus avancée, ou extension de l'auxiliaire *avoir* aux dépens d'*être* (*% ça était ouvert quand j'ai arrivé ici moi, Montréal 1971, 101 : 306) ; ou au contraire survivance : cooccurrence de pas* avec les autres négations (*% Ils n'iaisaient pas avec personne eux-autres, Montréal 1971, loc. 96), qui indique un fonctionnement différent de pas.*

Nous avons également noté si la même variation se trouvait ailleurs, par exemple le causataire avant l'infinitif (*% j'essayais de faire ma voix changer, Ottawa-Hull, 025/3247) se trouve aussi aux Antilles et à la Réunion. L'infinitif comme substitut (% tu te promènerais pas dans la rue pis dire salut au monde, îles de la Madeleine, Falkert 2007), existe aussi en Louisiane, en Wallonie, dans le Nord de la France, en Lorraine et en Suisse.*



Références bibliographiques

Abeillé A. & Godard D. (éds.), 2021, *La Grande Grammaire du français*, Actes Sud/Imprimerie Nationale Éditions, Arles.

Dostie, G., 2016, Le Corpus de français parlé au Québec (CFPQ) et la langue des conversations familiales : exemple de mise à profit des données à partir d'un examen lexico-sémantique de la séquence *je sais pas*, *Corpus*, 15.

Falkert, A., 2007. Valeur sémantique et comportement syntaxique de l'infinitif substitut dans quelques corpus oraux, *LINX* 57, p. 69-78.

Léard, J.-M., 1995. *Grammaire québécoise d'aujourd'hui*, Guérin universitaire, Montréal.

Poplack, S., 1996. Le corpus de français parlé à Ottawa-Hull, *Revue française de linguistique appliquée*, 1/2, p. 95-97.

Martineau, F. et Déprez, V., 2004, *Pas rien/Pas aucun* en français classique : variation dialectale et historique, *Langue française*, 143, p. 33-47.

Sankoff, D., Sankoff, G., Laberge, S. et Topham, M., 1976, Méthodes d'échantillonnage et utilisation de l'ordinateur dans l'étude de la variation grammaticale, *Cahier de linguistique*, 6.

Sankoff, G. et Thibault, P., 1977. L'alternance entre les auxiliaires *avoir* et *être* en français parlé à Montréal, *Langue française*, 34, p. 81-108.

Thibault, P. et Vincent, D., 1990, *Un corpus de français parlé : Montréal 84*, Presses de l'Université de Laval.

Villiard P. et M.-T. Vinet, 1983. Remarques sur l'expression de l'hypothèse en québécois, dans J.-M. Léard (éd), *Travaux de linguistique québécoise*, Presses de l'Université Laval, p. 209-221.



Dire l'approximation en micro-diachronie : de *disons* à *on va dire*

Lotfi Abouda (Université d'Orléans) et Marie Skrovec (Université d'Orléans)

En plus de son sens premier, le verbe *dire* peut fonctionner comme marqueur discursif, modalisant la forme ou le contenu d'un dire en train de se faire. Dans cet usage particulier, en plus de la forme *je dirais*, nous rencontrons le futur périphrastique avec *on (on va dire)* et l'impératif à la première personne du pluriel (*disons*), comme dans les deux exemples suivants :

1. [RF211] euh il m'arrive souvent de me servir d'un dictionnaire même très souvent de me servir d'un dictionnaire beaucoup plus souvent maintenant qu'il y a *disons* dix ans (ESL01_ENT_121).
2. bon euh tu galérais à finir ce ce jeu parce que il était tout aussi long que le jeu euh normal *on va dire* (ESL02_REPAS 0602)

Sur le plan syntaxique, faisant partie des modalisateurs parenthétiques (Perrin, 2012), *dire* se caractérise par une rection faible, pouvant tout aussi bien régir une subordonnée complétive, figurer en incise ou être postposé à une proposition (Blanche-Benveniste 1989, Urmson 1952, Andersen 1996). Sur le plan sémantico-pragmatique, malgré une grande diversité d'effets de sens pragmatiques (qu'il sera question d'identifier), *dire* a pour caractéristique commune de fonctionner dans ces emplois comme un marqueur méta-énonciatif marquant une distance du locuteur vis-à-vis de son dire.

Nous nous proposons dans cette recherche d'examiner les deux formes *disons/on va dire* par une exploration exhaustive de leurs occurrences dans un corpus oral d'environ 1,3 million de mots, extrait des ESLO (Enquêtes Sociolinguistiques à Orléans). Baptisé ESLO_MD_S (Abouda & Rendulic 2017), ce corpus, dont on présentera l'architecture, est composé de deux sous-corpus distincts, ESLO_MD (Abouda & Skrovec 2018), corpus micro-diachronique conçu pour des études en temps réel (avec combinaison d'études de tendance et par panel) et ESLO_Jeun-Inf, composé uniquement d'enregistrements d'ESLO2 mettant en scène des locuteurs jeunes, ce qui permet de mettre en perspective les résultats avec ceux du temps apparent. L'ensemble permet ainsi de cerner la dynamique du changement en combinant différentes approches.

Les résultats bruts montrent un affaissement très net de *disons* en l'espace de 40 ans : sur les 359 occurrences de *disons* identifiées dans ESLO_MD_S, 318 appartiennent à ESL01, 41 à ESL02, dont 5 seulement ont été identifiées dans Jeun-Inf.

Parallèlement, *on va dire*, marginal dans ESL01, connaîtra un essor remarquable dans ESL02 (Abouda & Skrovec 2016) : le nombre de ses occurrences passe de 5 à 183 (dont 61 dans la seule partie Jeun-Inf).

Si la parenté, mesurable par la possibilité de substitution entre elles, de ces deux formes morphologiquement distinctes rend légitime leur regroupement au sein d'un même paradigme fonctionnel, l'examen des facteurs internes (environnements syntaxiques, valeurs sémantiques, effets pragmatiques) et/ou externes (variables diaphasiques, diastratiques...) susceptibles de favoriser l'apparition d'une forme plutôt que d'une autre nous semble nécessaire.

Cette étude permettra, grâce à l'annotation syntaxique et sémantico-pragmatique de l'ensemble de leurs occurrences, opérée grâce au logiciel de textométrie TXM¹, de retracer l'évolution de ces deux marqueurs au cours des quarante dernières années. En plus de procéder à des relevés quantitatifs précis permettant de mesurer leur émergence ou régression, elle permettra, grâce à l'usage des métadonnées, de vérifier si la fréquence de telle ou telle forme est tributaire du genre interactionnel et/ou sensible aux variables sociodémographiques des locuteurs.

¹ Heiden & al. (2010) et Heiden (2010). Voir également <http://textometrie.ens-lyon.fr/>

Références bibliographiques

- Abouda, L. & Rendulic, N. (2017). « Séquence d'introduction de discours représenté : *faire* ou *dire* ? », *Discours* [En ligne], 21, mis en ligne le 22 décembre 2017. URL : <http://journals.openedition.org/discours/9353>.
- Abouda, L. & Skrovec, M. (2016). « Du mouvement au figement : pragmatization de la forme *on va dire*. Étude micro-diachronique sur un corpus oral », *Language Design*, Special Issue 2016, 121-145.
- Abouda, L. & Skrovec, M. (2018). « Pour une micro-diachronie de l'oral : le corpus ESL0-MD », SHS Web of Conferences 46, 11004, Congrès Mondial de Linguistique Française. <https://doi.org/10.1051/shsconf/20184611004>
- Andersen, H. L. (1996). « Verbes parenthétiques comme marqueurs discursifs », in C. Muller (éd.), *Dépendance et intégration syntaxique : subordination, coordination*, Tübingen : M. Niemeyer, 307-315.
- Anscombe, J.-C. (1985). « De l'énonciation au lexique : mention, citativité, délocutivité », *Langages*, 80, 9-34.
- Apothéloz, D. (2003). « La rection dite "faible" : grammaticalisation ou différentiel de grammaticité ? », *Verbum XXV* : 3, 241-262.
- Benveniste, E. (1966). « De la subjectivité dans le langage », *Problèmes de linguistique générale*, 1, Paris, Gallimard, 258-266.
- Blanche-Benveniste, C. (1989). « Constructions verbales "en incise" et rection faible des verbes », *Recherches sur le français parlé*, 9, 53-74.
- Delahaie, J. (2015). « *dis, dis donc, disons* : du verbe au(x) marqueur(s) discursif(s) », *Langue française* 186, 31-48.
- Dostie, G. (2004), *Pragmaticalisation et marqueurs discursifs. Analyse sémantique et traitement lexicographique*, Bruxelles : De Boeck/Duculot.
- Dostie, G., Pusch, C. D. (2007). « Présentation. Les marqueurs discursifs. Sens et variation », *Langue française*, 154, 3-12.
- Gachet, F. (2009). « Les verbes parenthétiques : un statut syntaxique atypique ? », *Linx* 61, 13-29.
- Gómez-Jordana Ferary, S. & Anscombe, J.-C. (2015). « Introduction : *dire* et ses marqueurs », *Langue française* 186, 5-12.
- Heiden, S. & al. (2010). « TXM, une plateforme logicielle open-source pour la textométrie : conception et développement », in Actes du 10th International Conference on the Statistical Analysis of Textual Data, Rome, Edizioni Universitarie di Lettere Economia Diritto 2, 1021-1032.
- Heiden, S. (2010). « The TXM Platform : Building Open-Source Textual Analysis Software Compatible with the TEI Encoding Scheme », in 24th Pacific Asia Conference on Language, Information and Computation, Sendai, Japan, Waseda University, 389-398.
- Khatchatourian, E. (2008). « Les marqueurs de reformulation formés à partir du verbe *dire* », in M.-C. Le Bot, M. Schuwer & É. Richard (éds), *La Reformulation. Marqueurs linguistiques. Stratégies énonciatives*, Rennes : Presses Universitaires de Rennes, 19-33.
- Oppermann-Marsaux, E. (2012). « *Di* de l'ancien français au français classique », in C. Guillot et al. (éds), *Le changement en français, études de linguistique diachronique*, Berne : Peter Lang, 265-280.
- Oppermann-Marsaux, E. (2014). « Les emplois du marqueur discursif *de* du moyen français jusqu'au français classique », in J.-C. Anscombe, E. Oppermann-Marsaux & A. Rodríguez Somolinos (éds), *Médiativité, polyphonie et modalité en français : études synchroniques et diachroniques*, Paris : Presses Sorbonne Nouvelle, 179-196.
- Péroz, P. (2013). « "C'est juste pour dire". Variation sémantique et régularité des opérations linguistiques dans le cas du verbe *dire* », *Pratiques*, 159, 257-273.
- Perrin, L. (2012). « Modalisateurs, connecteurs, et autres formules énonciatives », *Arts et Savoirs* [En ligne], 2. URL : <http://aes.revues.org/500> ; DOI : 10.4000/aes.500
- Saunier, E. (2012), « *Disons*, un impératif de dire ? », *L'Information grammaticale*, 131, 25-34.
- Steuckardt, A. (2005a). « Les marqueurs formés sur *dire* », in A. Niklas-Salminen & A. Steuckardt (dir.) (2005). *Les marqueurs de glose*. Presses de l'Université de Provence, 51-65.
- Steuckardt, A. (2005b). « Histoire de quelques correctifs formés sur *dire* », *Langue française* 186, 13-30.
- Steuckardt, A. (2014). « Polyphonie et médiativité dans un marqueur émergent : *on va dire* », in J.-C. Anscombe, E. Oppermann-Marsaux, A. Rodríguez Somolinos (éds.), *Médiativité, polyphonie et modalité en français. Études synchroniques et diachroniques*. Paris : Presses Sorbonne Nouvelle, 67-84.
- Urmson, J. O. (1952). Parenthetical verbs. *Mind*, 61, 480-496.

Moncton comme pôle urbain francophone : quels corpus pour décrire le français en usage ?

Laurence Arrighi (Université de Moncton), Basile Roussel (Université de Moncton) et Isabelle Violette (Université de Moncton)

Si Montréal est indéniablement « métropole et capitale culturelle du Canada français, ainsi que ville internationale et cosmopolite attrayante » (Breton-Carbonneau et Heller, 2021), il n'en demeure pas moins que d'autres centres urbains au Canada s'avèrent des pôles de francophonie dynamiques, attractifs et singulièrement intéressants pour la sociolinguistique. Moncton, principal centre urbain du Nouveau-Brunswick, présente une situation remarquable pour l'Acadie dans la mesure où la ville n'a cessé d'accroître son pouvoir d'attraction et son rayonnement comme pôle francophone au niveau local, régional, national et international depuis les années 1970 (Boudreau, 2014). L'une des manifestations de cette importance croissante de Moncton comme symbole contemporain de l'Acadie se manifeste dans le fait que le vernaculaire local, le chiac, est parfois présenté comme la langue des tout.e.s. les Acadien.ne.s. Toutefois comme en attestent déjà les études de corpus menées sur le français acadien, les pratiques linguistiques en Acadie, y compris au sein même de la ville de Moncton, sont largement hétérogènes (Arrighi, 2005 ; Balcom *et al.*, 2008 ; King, 2013 ; Neumann-Holzschuh et Mitko, 2018 ; Perrot, 2014 ; Roussel, 2020 ; Wiesmath, 2006, parmi d'autres). De plus, il convient de souligner que la plupart des corpus recueillis en Acadie datent d'il y a au moins 20 ans, ce qui atteste de la nécessité d'en impulser de nouveaux.

Notre objectif dans cette communication est tout d'abord de présenter une rétrospective des corpus disponibles au *Centre de recherche en linguistique appliquée* (CRLA) de l'Université de Moncton, duquel nous sommes tous les trois membres chercheurs. Nous mettrons en particulier l'accent sur les paramètres qui ont guidé les principes de représentativité et d'authenticité mobilisés pour recueillir ces corpus. Nous interrogerons également leurs éventuels écueils autour de la question : qu'est-ce qui a compté jusqu'à présent comme représentant.e légitime de « l'acadianité linguistique » ?

De plus, étant donné la reconfiguration actuelle de la francophonie acadienne, nous proposons une évaluation prospective afin d'identifier ce qui manque à une appréhension à la fois plus large et plus fine de la situation linguistique de Moncton et sa région. Outre les mobilités géographiques régionales, nationales et internationales mentionnées plus haut, il faut noter aussi la mobilité sociale ascendante des francophones de la ville. Dans un autre ordre d'idée, nous nous demanderons aussi quelles peuvent être de nouvelles avenues pour la récolte de corpus plus écologiques (Gadet et Guérin, 2016).

Ce tour d'horizon se veut une contribution principalement à l'axe 1 (La recherche sur le français de Montréal et des grands centres urbains de la francophonie) du colloque tout en rejoignant l'axe 2 autour du développement de corpus. En lien direct avec le contexte du colloque (les 50 ans des grands corpus montréalais), nous entendons aussi souligner le rôle des corpus dans la contribution à une connaissance plus informée de toute situation linguistique.

Références bibliographiques

- Arrighi, L. (2005). *Étude morphosyntaxique du français parlé en Acadie. Une approche de la variation et du changement linguistique en français. Tome I*. Thèse de doctorat, Université d'Avignon.
- Balcom, P., Beaulieu, L., Butler, G. R., Cichocki, W. et King, R. (2008). « The linguistic study of Acadian French ». Numéro thématique de la *Revue canadienne de linguistique/Canadian Journal of Linguistics*, n° 53(1).
- Boudreau, A. (2014) « Des voix qui se répondent : analyse discursive et historique des idéologies linguistiques en Acadie : l'exemple de Moncton. » *Minorités linguistiques et société/Linguistic Minorities and Society*, n° 4, p. 175-199. <https://doi.org/10.7202/1024697ar>
- Breton-Carbonneau, G. et Heller, M. (2021) « Le rapport centre-périphérie et les mobilités structurées : les jeunes Franco-Manitobains et Montréal. » *Francophonies d'Amérique*, n° 52, p. 85-103. <https://doi.org/10.7202/1082863ar>
- Gadet, F. et Guérin, E. (2016) « Construire un corpus pour des façons de parler non standard : "Multicultural Paris French" », *Corpus* [En ligne], n° 15, <http://journals.openedition.org/corpus/3049>
- King, R. (2013). *Acadian French in time and space*. Durham, N. C. : Duke University Press.
- Neumann-Holzschuh, I. et Mitko, J. (2018). *Grammaire comparée des français d'Acadie et de Louisiane (GraCoFAL). Avec un aperçu sur Terre-Neuve*. Boston : De Gruyter.
- Perrot, M.-È. (2014). « Le trajet linguistique des emprunts dans le chiac de Moncton : quelques observations », *Minorités linguistiques et société/Linguistic Minorities and Society*, n° 4, p. 200-218.
- Roussel, B. (2020). *À la recherche du temps (et des modes) perdu(s) : une étude variationniste en temps réel du français acadien parlé dans le nord-est du Nouveau-Brunswick*. Thèse de doctorat, Université d'Ottawa.
- Wiesmath, R. (2006). *Le français acadien : analyse syntaxique d'un corpus oral recueilli au Nouveau-Brunswick / Canada*. Paris : L'Harmattan.

Un nouveau corpus montréalais : les Français à Montréal

Julie Auger (Université de Montréal) et Nadège Fournier (Université de Montréal)

À l'instar des grandes villes nord-américaines, Montréal présente une importante diversité ethnique, culturelle et linguistique. Si le français demeure à ce jour la langue la plus parlée sur son territoire (Montréal en statistiques, 2017), le type de français parlé varie considérablement d'un foyer à l'autre. Pour bien comprendre la diversité et la richesse du français montréalais, il est nécessaire de s'intéresser de près à l'ensemble des variétés qui le constituent.

Dans le cadre de ce colloque qui souligne et célèbre les grands corpus grâce auxquels le français montréalais est très bien décrit, nous proposons une communication qui décrit un corpus en gestation qui vise à combler une lacune considérable : la documentation et l'analyse du parler des Français de Montréal. Depuis 10 ans, cette communauté connaît une forte croissance dans les grandes villes québécoises (Consulat général de France, 2020), particulièrement à Montréal. Notre projet se penche sur les pratiques langagières de ces Néo-Montréalais afin de mieux comprendre, d'une part, leur contribution au français de la métropole et, d'autre part, de déterminer dans quelle mesure ces locuteurs d'une variété de français considérée comme étant standard et prestigieuse incorporent des éléments montréalais dans leur parler. Pour ce faire, nous avons mis sur pied un projet en 2019 afin de constituer le corpus *Les Français à Montréal*. Nous procédons présentement aux premières analyses quantitatives.

Dans le but de faciliter les comparaisons avec les corpus existants, nous réalisons des entrevues semi-dirigées auprès de deux groupes de Français : des locuteurs établis à Montréal depuis plus de 5 ans et des locuteurs nouvellement arrivés. Afin de déterminer dans quelle mesure les Néo-Montréalais d'origine française modifient leur parler en fonction de leur interlocuteur, chaque participant est interviewé par deux enquêtrices : une Montréalaise et une Française. Le protocole de chaque entrevue est inspiré de ceux utilisés dans les corpus Sankoff/Cedergren, Montréal 1984 et FRAN-HOMA, mais les questions sont adaptées pour prendre en compte le consultant et l'enquêteur en présence afin de minimiser les chevauchements et favoriser des conversations naturelles. De plus, afin de mieux comprendre l'aspect longitudinal de l'acquisition d'un autre dialecte (par exemple, traits acquis, ordre d'acquisition et parcours individuels), nous interviewons les participants du second groupe tous les six mois pendant leurs trois premières années dans la métropole. Finalement, une innovation importante de notre corpus concerne l'inclusion d'une tâche expérimentale, inspirée du domaine émergent de la sociolinguistique cognitive et du travail de Chevrot *et al.* (2018), qui teste la production et la perception des traits montréalais de nos participants. Les résultats de cette tâche permettront de déterminer dans quelle mesure l'acquisition des traits montréalais est guidée par leur perception. La combinaison de ces approches complémentaires permet de recueillir des données riches provenant de locuteurs appartenant à une population montréalaise jamais étudiée. Nos analyses préliminaires permettent de déterminer que des mots du discours comme *fait que* et la prononciation d'un [a] en fin de mot font partie des traits acquis par certains Néo-Montréalais d'origine française.

Références bibliographiques

Blondeau, Hélène. 2020. La sociolinguistique au Québec des années 1960 à 2000. In Gaétane Dostie (dir.) *Combats pour la linguistique au Québec (1960-2000). Courants, théories, domaines* (Gramm-R), Peter Lang, 251-279.

Blondeau, Hélène, Mireille Tremblay, Anne Bertrand & Elizabeth Michel. 2021. A new milestone for the study of variation in Montréal French : The Hochelaga-Maisonneuve sociolinguistic survey. *Corpus 22* : 1-16.

Sankoff, David, Gillian Sankoff, Suzanne Laberge & Marjorie Topham. 1976. Méthodes d'échantillonnage et utilisation de l'ordinateur dans l'étude de la variation grammaticale. *Cahier de linguistique 6* : 85-125.

Thibault, Pierrette & Michelle Daveluy. 1990. *Un corpus de français parlé : recherches sociolinguistiques*. Québec : Université Laval.

/tUt/ en français montréalais : 44 ans plus tard...

Davy Bigot (Université Concordia), Christophe Beaulieu (Université Concordia) et Emily Leavitt (Université Concordia)

En français québécois (FQ), les formes masculines *tout* ([tu]) et *tous* ([tu]-[tus]) du français standard (FS) peuvent être prononcées [tUt], comme dans les exemples suivants :

1. (FQ) j'ai bu [tUt] mon vin vs j'ai bu [tu] mon vin (FS);
2. (FQ) j'ai nettoyé [tUt] mes chars vs j'ai nettoyé [tu] mes chars (FS);
3. (FQ) mes chums sont [tUt] venus vs mes chums sont [tus] venus (FS).

Burnett (2013 : 211) souligne que /tUt/ est « l'un des éléments les plus étudiés du système de quantification du FQ ». En effet, plus d'une quinzaine d'études portant sur le phénomène ont été publiées. Toutefois, très peu d'entre elles se sont penchées sur l'emploi de /tUt/ du point de vue de la linguistique variationniste (Lemieux et al., 1985 ; Lemieux et al., 1986 ; Daveluy, 2005 ; Bigot, 2010 ; Labelle-Hogue, 2012). En outre, les données sont peu comparables. Certains corpus sont distants de plus de 30 ans. Certaines recherches sont basées sur des entrevues semi-dirigées informelles, alors que d'autres sont construites autour d'entrevues formelles ou d'épisodes d'une série télévisée. Il ressort néanmoins de leur mise en commun que l'emploi de /tUt/ n'est ni systématique, ni aléatoire. Au contraire, on constate que son utilisation est variable, et que celle-ci obéit à des facteurs linguistiques et extralinguistiques précis.

L'essentiel des données sociolinguistiques sur /tUt/ en français montréalais date maintenant de plusieurs décennies. Depuis les études de Lemieux et al. (1985 ; 1986), l'emploi de /tUt/ n'a fait, à notre connaissance, l'objet d'aucune autre analyse variationniste basée sur un corpus de français parlé à Montréal. Le problème de la représentativité contemporaine des données se pose donc naturellement. Dans le cadre de cette présentation, nous nous proposons de revisiter le phénomène à la lumière d'un corpus récent d'entrevues semi-dirigées recueilli en 2015, à Montréal, dans le quartier d'Ahuntsic-Cartierville.

Tout d'abord, nous brosserons un portrait sociolinguistique de /tUt/. Puis, nous présenterons le corpus de Montréal Ahuntsic-Cartierville, et les principaux points méthodologiques de notre étude. Enfin, nous dégagerons les tendances lourdes de notre analyse, et nous détaillerons les points de convergence et de divergence entre nos résultats et ceux obtenus dans les études antérieures. Notre recherche révèle essentiellement que 1) l'emploi de /tUt/ n'est plus la variante principale en français montréalais, 2) la charge sociostylistique de /tUt/ a évolué entre 1971 et 2015, et 3) les facteurs internes conditionnant l'usage de /tUt/ en 2015 divergent de ceux du corpus Sankoff-Cedergren de 1971.

Références bibliographiques

- Bigot, Davy. 2010. « La norme grammaticale du français québécois oral : des questions, une réponse », In Martineau F., LeBlanc C., et Y. Frenette (dir.), *Vues sur les français d'ici*, Québec : Presses de l'Université Laval, pp. 9-30.
- Burnett, Heather. 2013. « Structure événementielle et modification pragmatique : on connaît-tu tout sur /tUt/ ? », In Bigot D., Friesner M., et M. Tremblay (dir.), *Les français d'ici et d'aujourd'hui. Description, représentation et théorisation*. Québec : Presses de l'Université Laval, pp. 211-230.
- Daveluy, Michèle. 2005. *Les langues étendards. Allégeances langagières en français parlé à Montréal*. Québec : Nota bene.
- Labelle-Hogue, Simon-Pier. 2013. « État du vernaculaire dans la téléserie québécoise : l'exemple de la Petite vie », In Bigot D., Friesner M., et M. Tremblay (dir.), *Les français d'ici et d'aujourd'hui. Description, représentation et théorisation*. Québec : Presses de l'Université Laval, pp. 131-172.
- Lemieux, Monique, Anne St-Amour et David Sankoff. 1985. « tUt en français de Montréal », In Cedergren H., et M. Lemieux (dir.), *Les Tendances Dynamiques du Français Parlé de Montréal*. vol. 2, Montréal : Office de la langue française, pp. 7-90.
- Lemieux, Monique, Carmen Fontaine et David Sankoff. 1986. « Quantificateur et marqueur du discours », In Sankoff D. (dir.), *Diversity and Diachrony*. Amsterdam/Philadelphia : JBPC, pp. 381-390.

Convergence ou divergence en français laurentien : 40 ans de changement à Montréal et à Welland

Hélène Blondeau (University of Florida), Claire Bourély (Université de Montréal), Raymond Mougéon (Collège Glendon, Université York) et Mireille Tremblay (Université de Montréal)

Cette communication examine le changement linguistique dans les français de Montréal et de Welland (Ontario), deux variétés laurentiennes génétiquement reliées, mais évoluant dans des contextes différents. Nous évaluons dans quelle mesure ces deux variétés montrent une tendance à la convergence ou à la divergence.

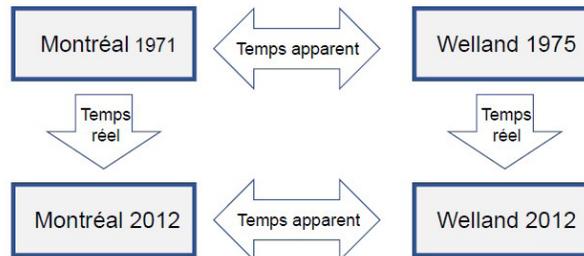


Figure 1: Représentation de l'étude sur les tendances communautaires

Notre étude s'appuie sur des corpus recueillis une première fois dans les années 1970 (Montréal 1971 et Welland 1975) et une seconde fois, 40 ans plus tard (Montréal 2012 et Welland 2012). Notre cadre conceptuel (Figure 1) permet de tester la validité du modèle de temps apparent en comparant les changements en temps réel aux temps 1 (années 1970) et 2 (années 2010) dans des contextes majoritaire (Montréal) et minoritaire (Welland).

Nous croisons les résultats pour deux variables :

L'usage des marqueurs de conséquence : deux variantes standard (*alors* et *donc*) et deux variantes vernaculaires (*ça fait que* et l'emprunt *so*)

l'utilisation des pronoms non clitiques du pluriel à la forme simple: *nous*, *vous*, *elles*, *eux* et leurs équivalents composés, *nous autres*, *vous autres*, *eux autres*, typiques du vernaculaire.

Divergence. Dans le cas des marqueurs de conséquence, la comparaison des corpus des années 1970 avec ceux des années 2010 révèle que les deux communautés avaient déjà commencé à diverger dans les années 1970 (absence de *so* à Montréal, différences dans la fréquence des variantes communes, certaines contraintes internes et externes ne sont pas partagées). Depuis, la divergence intercommunautaire s'est intensifiée : 1- Welland : croissance du connecteur *so* aux dépens de la variante vernaculaire traditionnelle (*ça fait (que)* – reflétant l'avancement du bilinguisme ; 2- Montréal : augmentation marquée de (*ça fait (que)*) et diffusion dans tous les groupes sociaux. Notre analyse de l'évolution de la variante standard *alors* révèle une autre facette de la divergence : 1- Welland : *alors* fait preuve de stabilité ; 2- Montréal : ce connecteur a subi un fort déclin.

Convergence ? Les résultats sur les pronoms indiquent un changement par le haut en faveur des formes simples dans les deux communautés, mais le changement est vigoureux à Montréal (8 % à 48 %), alors qu'il est plutôt modeste à Welland (10 % à 22 %). À Montréal, l'évolution rapide du changement découle du statut majoritaire et officiel du français au Québec, d'un accès accru à l'éducation et aux médias, ainsi qu'au contact avec des variétés non locales de français. À Welland, le caractère modeste du changement reflète la rareté de l'exposition aux variétés externes de français (l'immigration francophone étant quasi nulle à Welland) et la faiblesse relative de la pression normative, propre au statut minoritaire du français dans cette ville. L'adoption des formes simples est aussi associée au parler des bilingues anglo-dominants, locuteurs moins intégrés à la communauté francophone vernaculaire que les locuteurs franco-dominants.

En conclusion, notre étude de deux variables montre que si les français parlés à Montréal et à Welland partagent *grosso modo* les mêmes variantes, des facteurs propres aux contextes minoritaire/majoritaire entraînent des différences plus ou moins marquées dans l'évolution de la fréquence des variantes et des contraintes qui pèsent sur leur emploi.

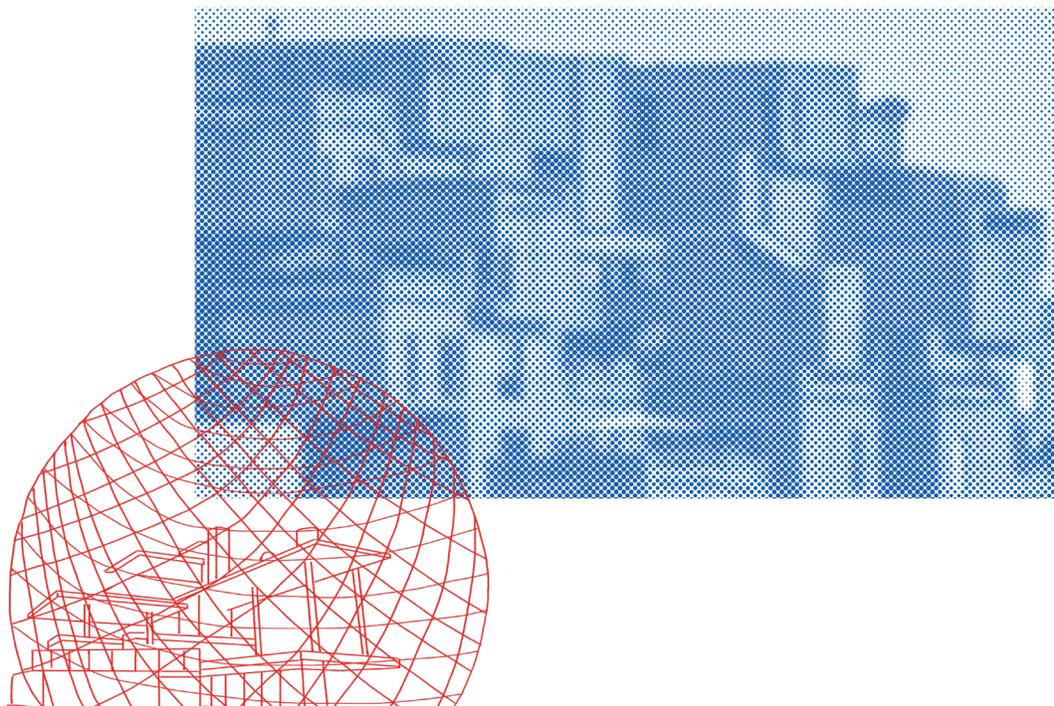
Références bibliographiques

Beniak, Édouard, Mougeon, Raymond et Valois, Daniel. (1985). *Contact des langues et changement linguistique : étude sociolinguistique du français parlé à Welland*. Québec : Centre International de Recherche sur le Bilinguisme.

Blondeau, Hélène, Tremblay, Mireille, Bertrand, Anne et Michel, Elizabeth. (2021). A new milestone for the study of variation in Montréal French : The Hochelaga-Maisonneuve sociolinguistic survey. *Corpus* 22 : 1-16.

Martineau, France et Séguin, Marie-Claude. (2016). Le corpus FRAN : Réseaux et maillages en Amérique française. *Corpus* 15.

Sankoff, David, Sankoff, Gillian, Laberge, Suzanne et Topham, Marjorie. (1976). Méthodes d'échantillonnage et utilisation de l'ordinateur dans l'étude de la variation grammaticale. *Cahiers de Linguistique de l'Université du Québec* 6 : 85-125.



La référence temporelle au futur en Belgique et au Québec : corpus oraux et corpus de textos

Hélène Blondeau (University of Florida), Anne Dister (Université Saint-Louis), Emmanuelle Labeau (Aston University) et Mireille Tremblay (Université de Montréal)

Cette communication propose un examen comparatif de la variation dans la référence temporelle au futur (RTF) dans le français du Québec et de Belgique, en faisant appel à des corpus de français parlé recueillis à Montréal et à Bruxelles de même qu'au corpus de textos sms4science (<http://www.sms4science.org/>). Depuis le travail fondateur d'Emirkanian et Sankoff (1985), les travaux consacrés à cette variable se sont principalement penchés sur l'oral et ont surtout examiné la situation une communauté à la fois. Dans la plupart des cas, le questionnement sur la variation dialectale reposait sur une discussion des autres études sans en contrôler les paramètres méthodologiques. Notre étude comparative définit le contexte variable de la même façon et adopte la même grille analytique.

Nous prenons appui sur une première analyse comparative de corpus de textos québécois et belge (Tremblay, Blondeau et Labeau, 2020) portant sur les trois variantes de la RTF, soit le présent à valeur de futur (PF), le futur synthétique (FS) et le futur analytique (FA). Alors que le PF présente des taux d'utilisation (Belgique (B) : 46,2 % ; Québec (Q) : 48,9 %) et des contraintes linguistiques semblables dans les deux communautés, il existe des différences entre les deux communautés en ce qui a trait à la fréquence des variantes synthétique (B : 38,5 % ; Q : 19,1 %) et analytique (B : 15,3 % ; Q : 32,0 %), le français de Belgique étant plus favorable au futur synthétique (1) que le français du Québec, qui lui préfère le futur analytique (2).

1. *Ouai en bref on fera jms ke s'engueulé kan on se parlera!* (SMS4Science, Belgique)
Oui en bref on fera jamais que s'engueuler quand on se parlera !
2. *Non jva le caller demain matin* (Texto4Science, Québec)
Non je vais le caller demain matin

Par ailleurs, le poids des contraintes linguistiques sur l'alternance FS/FA ne s'exerce pas de la même façon. La polarité qui exerce un effet important dans les textos québécois n'a pas d'effet significatif dans les textos belges et la spécification adverbiale n'affecte que les données belges. Cependant, les deux corpus convergent en ce qui a trait au rôle de la distance temporelle, bien que son effet soit plus fort dans les textos belges. Enfin, l'effet du type de verbe, un aspect peu étudié jusqu'ici (Poplack & Turpin 1999 ; Tristram 2021), joue un rôle dans les deux communautés. Notre examen comparatif de la référence temporelle au futur (RTF) dans le français du Québec et de Belgique, fait également appel à des corpus de français parlé recueillis à Montréal et à Bruxelles. L'analyse se concentre sur la variation entre le FS et le FA pour voir si la différence dialectale observée dans les corpus textos se retrouve à l'oral, tant en ce qui a trait au choix des variantes qu'aux contraintes linguistiques.

Cette étude permet non seulement de lever le voile sur la variation dialectale selon deux modalités, mais aussi de vérifier si ce cas de variation est impliqué dans un processus de changement en cours dans les deux communautés.

Références bibliographiques

Emirkanian, L. et Sankoff, D. (1985). « Le futur simple et le futur proche », dans M. Lemieux et H. Cedergren (dir.), *Les tendances dynamiques du français parlé à Montréal. Vol. 1*. Québec : Office de la langue française, p. 189-204.

Poplack, S. et Turpin, D. (1999). « Does the *Futur* have a future in (Canadian) French ? », *Probus*, 11(1), 133-164.

Tremblay, M., Blondeau, H. et Labeau, E. (2020). « Texting the future in Belgium and Québec : Present matters », *Journal of French Language Studies*, 30(1), 73-98. doi:10.1017/S0959269519000188

Tristram, A. (2021). « Variation and change in future temporal reference with *avoir* and *être* », *Journal of French Language Studies*, 31(1), 25-49. doi:10.1017/S0959269520000174

Corpus

Blondeau, H., Tremblay, M., Bertrand, A. et Michel, E. (2021), « A new milestone for the study of variation in Montréal French : The Hochelaga-Maisonneuve sociolinguistic survey », *Corpus* 22.

Dister, A. et Labeau, E. (2017). « *Le Corpus de français parlé à Bruxelles* : origine, hypothèses et développements », *Cahiers de l'AFLS* (Association for French Language Studies). 21.1.

Fairon, C., Klein, J.-R. et Paumier, S. (2006). « SMS pour la science. Corpus de 30.000 SMS et logiciel de Consultation », *Cahiers du Cental* 3.2., Presses Universitaires de Louvain.

Langlais, P. and Drouin, P. (2012). « Texto4Science : A Quebec French database of annotated text messages », *Linguisticae investigationes*, 35, 237-259.

Martineau, F. et Séguin, M.-C. (2016), « Le corpus FRAN : Réseaux et maillages en Amérique française », *Corpus* 15.

La gentrification du quartier Hochelaga dans les corpus oraux : de 1971 à 2012

Claire Bourély (Université de Montréal)

Cette communication se penche sur l'évolution des perceptions de la gentrification dans les discours de locuteurs résidents du quartier Hochelaga. Mon analyse compare les discours de deux locuteurs (18 et 25 ans) en 1971 et ceux de 38 locuteurs (entre 18 et 65 ans) en 2012, tous vivant dans le quartier Hochelaga. Il s'agira de décrire et analyser l'évolution des représentations sociales de la gentrification du quartier Hochelaga de 1971 à 2012.

À partir du grand corpus de 1971 (Sankoff-Cedergren), de nombreux travaux ont donné lieu à des descriptions du français parlé à Montréal (D. Sankoff et coll., 1976) dans une perspective sociolinguistique ciblant la variation linguistique à l'œuvre. D'autres se sont intéressés à la progression de la variation dans les corpus comparables de 1971, 1984 et 2012. À l'aide de méthodologies quantitatives appuyées sur l'analyse de facteurs, les études ont porté essentiellement sur la langue. Je me propose de travailler sur le contenu des entretiens, avec une visée qualitative.

Mon étude se fonde sur l'étude des discours des locuteurs entre 1971 et 2012. Je m'appuierai sur l'analyse thématique des différentes descriptions des conséquences de la désindustrialisation du quartier Hochelaga dans la moitié du XX^e siècle et de sa revitalisation à la fin du XX^e siècle. Mes travaux montrent que la gentrification déjà à l'œuvre en 1971 s'est accélérée, donnant lieu à des perceptions positives et négatives de la part des locuteurs.

Dans les deux exemples ci-dessous, les locuteurs évoquent le même lieu géographique, la même installation, le « mini putt », mais offrent deux représentations de l'espace distinctes à 40 ans d'écart.

1. C'est pas des changements brusques. Les constructions sont très rares. On construit... Les premiers débuts ils ont fait un gros coup Steinberg des choses de même. Puis après ça ça a arrêté. Puis là **ils vont peut-être construire je crois... j'ai entendu... sur l'autre bord un mini putt** pour jouer au bowling... pas au bowling au... au golf. (06, 1971)
2. Ben oui parce que avant c'était un Steinberg à place de de ce/ chez Métro. Avant c'était pas comme ça c'est/ C'était petit pis poussiéreux mais là là depuis cinq à dix ans là/ ah/ parce qu'avant là à à place d'Arôme là il y avait une une vieille roulotte à patates frites. Toute graisseuse. Toute dégueulasse. **Avec un un un mini putt avec euh des... un terrain vacant touT sale dégueulasse.** (020F65, 2012)

Il était alors pertinent de s'interroger sur la manière dont les discours sur la gentrification indexent les locuteurs (Silverstein, 2003) et leur confèrent une identité sociale située, entre discours du contre et aspects positifs du phénomène.

Cette étude permet, d'une part, de réfléchir aux enjeux de la gentrification et, d'autre part, de travailler sur les perceptions (Di Méo, 1994) des acteurs du quartier Hochelaga, locuteurs victimes de la gentrification ou eux-mêmes gentrificateurs (Clerval, 2008).

Devenir féministes à Paris, Marseille et Montréal : récits de vie dans le corpus CaFé

Heather Burnett (Université Paris Cité) et Julie Abbou (Université Paris Cité)

Le but de cette communication est de présenter le corpus CaFé, Cartographie linguistique des Féminismes. Il s'agit d'un corpus comparatif d'entretiens semi-dirigés à Paris, Marseille et Montréal avec des personnes qui sont engagées dans une activité en lien avec la cause des femmes, le féminisme ou les luttes pour les sexualités. Le sous-corpus parisien est en cours de transcription (42 entretiens réalisés majoritairement en 2021), tandis que le sous-corpus marseillais est en cours de constitution (10/40 entretiens en hiver 2022). Les 40 entretiens pour le corpus montréalais seront réalisés au printemps 2022. Les entretiens incluent une présentation de la trajectoire biographique des locutrices, une discussion sur leur positionnement politique et idéologique, et enfin une discussion sur leur rapport au langage.

L'objectif du corpus est de permettre des analyses intégrant la part sociale du langage dans les modélisations du langage. Cependant, le corpus se laisse mal saisir par les catégories sociodémographiques traditionnellement utilisées en sociolinguistique variationniste, car le féminisme n'est pas une catégorie homogène : elle concerne les femmes, mais pas seulement les femmes et pas toutes les femmes ; touche à leurs conditions matérielles d'existence en termes de rapports sociaux ; elle est un engagement idéologique mais dont l'échiquier conceptuel échappe aux découpages traditionnels de la politique partisane et à leurs acteurs ; enfin, c'est une catégorie qui a des impacts socio-économiques et symboliques sur les mobilités (carrière professionnelle vs carrière militante). Par exemple, le critère traditionnel "formation universitaire" ne répartit pas les personnes entretenues de manière pertinente, tandis que le critère "grandes écoles" est pertinent pour le volet parisien. De même, l'âge est relativement mauvais prédicteur de découpage, mais les générations (i.e. le contexte de politisation) pèsent sur les trajectoires féministes. Ces nouvelles catégories – qui articulent des positionnements idéologiques, des catégories socialement structurées, matériellement et historiquement, un contexte socio-historique et des mouvements d'identification – feront émerger de nouvelles analyses reliant variables sociolinguistiques et trajectoires biographiques.

Pour illustrer l'utilité de cette approche pour l'étude de la variation linguistique, nous présentons une étude de la façon de faire référence à une personne féminine (*femme* vs *meuf* vs *nana* vs *fille*) dans le sous-corpus parisien. Dans le corpus, il existe beaucoup de variation intra-locutrice, comme montrent les exemples en (1), tous dits par la locutrice 1.

- 1a. puis ma mère en plus c'est une **nana** qui a une grande gueule donc elle c'est pas la **femme** qui ferme sa gueule etc.
- 1b. j'ai relancé une **fille** que j'avais rencontrée l'année dernière qui lançait sa boîte
- 1c. il y avait pas mal de **nanas** et euh surtout des justement des **meufs** euh leaders

Nous comparerons les résultats obtenus selon les catégories socio-démographiques traditionnelles avec les résultats obtenus selon les nouvelles catégories proposées (formation élitaires ou non, génération et contexte de politisation, type d'engagement, positionnement dans le champ féministe).

Approche linguistique et épilinguistique de la mobilité discursive des étudiant/es arabophones à Montréal : autour des expressions figées

Anne-Sophie Calinon (Université Bourgogne Franche-Comté) et Katja Ploog (Université d'Orléans)

L'adoption du français par les néo-Montréalais est marquée, entre autres choses, par l'appropriation des expressions figées (ou : polylexicales, non littérales, préfabriquées). Nous définirons les expressions figées (désormais EF) par (a) leur non-compositionnalité sémantique, (b) la non-actualisation des référents, (c) la fixité syntaxique, (d) la restriction combinatoire (Legallois & Tuttin, 2013 : 4). Ces constructions sont des ressources linguistiques à teneur socio-indexicale marquée, dont le caractère emblématique les fait envisager comme véritable patrimoine culturel de la langue que l'on retrouve dans le « territoire du locuteur ».

Notre ancrage est celui de la (socio)linguistique de la mobilité (Ploog, Calinon & Thamin, 2020 ; Blommaert, 2010 ; Mondada, 1994), où la notion de *territoire* (Raffestin, 1980 ; Baud *et al.*, 2003 ; Berthomière & Hily, 2006) est appréhendée comme l'espace investi par le sujet, à la fois physiquement et discursivement. Dans une expérience de mobilité, l'élargissement du territoire spatial par l'investissement d'une terre "d'accueil", se trouve relayé par celui de son territoire langagier, construit et balisé par les différentes ressources linguistiques mises en œuvre dans l'élaboration discursive singulière. Par ailleurs, la réflexivité du sujet-locuteur le conduit à verbaliser son expérience (de la mobilité en l'occurrence) par des discours épilinguistiques.

Nous montrons comment les EF contribuent à la présentation de soi comme membre compétent de la communauté discursive. Leur émergence est relayée en effet par une indexicalité complexe qui élabore des ancrages multiples sur le territoire. La fluidité du discours en effet est facilitée par le recours à des formats les plus canoniques/routiniers possibles ; or, le nombre accru de contraintes morphosyntaxiques rend risquée l'élaboration de ces constructions. Leur utilisation effective signale que le risque a été mesuré, jusqu'à être revendiqué ici par un commentaire métadiscursif dans l'exemple suivant, issu de l'entretien avec Amir (Québec) :

et et en plus de cela euh en tant que x- en tant que moDÈLE euh je pense que que les études supérieures en France a un peu un peu euh **perdu de son de de comment de son ÉCLAT** (hum hum) si si vous permettez l'expression euh même au Maroc maintenant euh c'est plus vers le Canada (E15, 38 : 55)

Ce segment donne à voir une mobilité discursive complexe, en coarticulant trois types d'espaces, géographique (par la référence à la France et l'affirmation du locuteur de sa compétence socio-culturelle vis-à-vis de cet espace) ; l'espace discursif ensuite, par les commentaires (a priori) métadiscursifs, comportant, qui plus est, une allocution explicite (*si vous permettez*) ; espace langagier enfin, avec une EF (*perdre de son éclat*). Sur le fond, le locuteur assume le non-alignement potentiel vis-à-vis de son interlocutrice française, par sa désapprobation des conditions d'étude en France. Le commentaire métalinguistique *si vous permettez l'expression* indique que le locuteur a conscience d'employer une expression (en substance, *perdre de son éclat*). Puis, les disfluences nombreuses (répétitions, hésitations, reprises), jusqu'à la présence d'une balise de formulation (*comment*) laissent penser que l'élaboration de la construction est laborieuse, idée corroborée par la défaillance de l'accord (*les études / a perdu*). Il est remarquable que les décrochages énonciatifs marquent, de façon accentuée, les EF non littérales et métaphoriques, qui sont aussi les plus fréquemment touchées par une réalisation non standard...

À travers l'étude d'un corpus de 9 entretiens sociolinguistiques recueillis auprès d'étudiant/es maghrébin/es arabophones et inscrit/es dans une expérience de mobilité académique dans l'espace francophone (France et/ou Canada) dans le cadre du projet CEM (Dynamiques spatiales, langagières, identitaires de la circulation migratoire étudiante [Maghreb, France et Canada]), nous discuterons les constellations les plus typiques des élaborations constructionnelles relatives aux EF, et le travail de formulation qui en marque l'émergence par les disfluences et les commentaires métalinguistiques ou métadiscursifs. L'actualisation de ce répertoire à l'hétérogénéité manifeste cartographie la tension du territoire en mutation où évolue le locuteur.

Références bibliographiques

Canut, C. (2000) « Subjectivité, imaginaires et fantasmes des langues : la mise en discours "épilinguistique" », *Langage et société* 2000/3 (n° 93), p. 71-97.

Legallois, D. & Tutin, A. (2013) « Présentation : Vers une extension du domaine de la phraséologie ». *Langages*, 1(1), p. 3-25.

Ploog, K., Calinon, A.-S. & Thamin, N. (dirs.) (2020) *Mobilité : Histoire et émergence d'un concept en sociolinguistique*. Paris : L'Harmattan. Collection Espaces discursifs.

Blommaert, J. (2010) *The sociolinguistics of globalization*. Cambridge : Cambridge University Press.

Mondada, L. (1994) *Verbalisation de l'espace et fabrication du savoir : approche linguistique de la construction des objets de discours*, Université de Lausanne, Faculté des lettres.

Raffestin, C. (1980) *Pour une géographie du pouvoir*. Paris : Libraires techniques.

Baud et al., (2003). *Dictionnaire de géographie*. Paris : Hatier.

Berthomière, W. & Hily, M.-A. (2006) « Décrire les migrations internationales », *Revue européenne des migrations internationales*, vol. 22, n° 2, p. 67-82.



Récit de vie et souffrance : impact identitaire de la (non) transmission familiale à partir d'un corpus grenoblois de descendant·e-s italien·ne·s

Roxanne Comotti (Université Grenoble Alpes)

Cette communication s'inscrit dans l'axe *récit de vie* et se penche sur les représentations de locuteur·trices interrogé·es sur leurs (non)pratiques linguistiques à partir de la thématique de la souffrance.

Dans la littérature américaine, Fishman (1999) conceptualise l'intégration linguistique de migrant·es espagnol·es aux États-Unis sur un continuum de trois générations. Cette intégration résulte du processus par lequel les locuteur·trices de la première génération (G1) minimisent l'usage et la transmission de leur langue d'origine (LO) à leurs enfants (G2) au profit de la langue du pays de migration (PM). Dans ce processus, les locuteur·trices G2 né·es dans le PM – ou arrivant jeunes – supportent parfois un rôle linguistique charnière entre la LO et la langue du PM (Fibbi & Matthey, 2010). Les travaux en sociolinguistique mettent en avant les difficultés vécues des G1, à travers la « langue comme épreuve » (Deprez, 2002), ou encore la construction identitaire dans la migration (Giacomi, 1994). Fort est de remarquer que les recherches se concentrent principalement sur les récits de vie des G1. Lorsque l'on s'intéresse aux G2, ces dernier·ères sont représenté·es comme une génération creuse, qui a « volontairement mis entre parenthèses, notamment la langue d'origine [de leurs parents] » (Hansen, 1938 dans Fibbi & Matthey, 2010).

La communication se propose de dépasser la représentation de Hansen et se focalise sur la position *entre deux* des G2. Plus précisément, à partir d'entretiens de récits de vie (Bertaux, 1997) je questionnerai le concept de changement de langue (*language shift*) à travers le prisme de la souffrance dans le discours des G2. Je proposerai un élément de réponse à la question suivante : *Comment le discours de la souffrance dans les récits de vie des locuteur·trices de deuxièmes générations peut-il être mis en relation avec leur construction identitaire et la (non)transmission intergénérationnelle ?*

Le corpus analysé présente 17 récits de vie de G2 descendant·es d'une migration italienne à Grenoble arrivée entre 1950 et 1960. Les entretiens semi-dirigés ont été réalisés en 2017. Le logiciel AntConc est utilisé pour aider à l'exploration du corpus.

L'analyse se concentre sur les faits de parole thématissant la souffrance. Les marqueurs langagiers tels que les choix lexicaux et certaines manifestations non-verbales (silence, soupir, larmes...) sont privilégiés dans la mesure où je considère qu'ils participent à la fois de la représentation de la souffrance dans la construction identitaire des G2 et à la fois de la rupture de la transmission familiale intergénérationnelle. On peut ainsi isoler des séquences « profératoires » qui ont la particularité de « dire en signe ce qui était déjà présent en sens » (De Robillard, 2020) et ce dans une énonciation empreinte d'émotion comme l'exemple suivant : « moi ma mère [...] elle parlait couramment l'italien elle l'écrivait couramment [...] elle nous a jamais dit un mot en italien même à table +voilà +d'accord ? » (F12f1E2, *nos italiennes*).

La communication avancera des éléments de réponse à la problématique présentée, à partir des données extraites du corpus en cours d'analyse.

Références bibliographiques

Bertaux, D. (1997). Les récits de vie. Perspective ethnosociologique. Nathan Université, Tours.

De Robillard, D. (2020), communication lors de la Journée d'étude Dynamique et enjeux éthico-épistémologique du récit de vie : réfléchir en termes d'expérience, 10 mars 2020, <https://calenda.org/751441>.

Deprez, C. (2002) La langue comme « épreuve » dans les récits de migration. Dans Bulletin VALS-ASLA (Association suisse de linguistique appliquée) 76, 39-52.

Fibbi, R., Matthey, M. (2010), « Petits-enfants de migrants italiens et espagnols en Suisse », *Hommes & migrations* [En ligne], 1288.

Fishman, J. A. (1999), « Language Maintenance and Language Shift as a Field of Inquiry. A Definition of the Field and Suggestions for its Further Development », in *Linguistics*, n° 9, 1964, pp. 32-70.

Giacomi, A. (1994). Récits de migration et construction d'images identitaires. Dans Bres, J. (éd.) *Le récit oral*. Université Paul Valéry-Montpellier III, Praxiling, 297-308.

Hansen, M. (1938), *The Problem of the Third Generation Immigrant*, Rock Island, Augustana Historical Society.

50 ans de liaison à Montréal

Marie-Hélène Côté (Université de Lausanne)

La liaison est un des phénomènes phares du système sonore du français, par la variation qu'elle manifeste et la multiplicité des facteurs qui en déterminent la réalisation (Côté 2011 ; Durand *et al.* 2011). Cette variation a été surtout reliée à la dimension stylistique – on réalise davantage de liaisons en contexte formel –, la dimension diatopique ayant été traditionnellement négligée (contrairement au schwa, pour lequel la variation diatopique est depuis longtemps bien établie). Le français montréalais constitue une exception, les spécificités de la liaison montréalaise ayant été décrites à partir du corpus Sankoff-Cedergren (Ameringen 1977 ; Ameringen & Cedergren 1981 ; de Jong 1993 ; cf. aussi Tousignant 1978). Plus récemment, la constitution du corpus Phonologie du français contemporain (PFC ; Durand, Laks & Lyche 2002 ; Detey *et al.* 2016 ; www.projet-pfc.net) a changé la donne. D'une part, ce corpus vise à documenter l'ensemble du monde francophone ; d'autre part il inclut le codage de la liaison dans tous les contextes de liaison potentielle. Il permet donc une comparaison systématique de la réalisation de la liaison à travers les variétés de français (cf. Côté 2017). Le Québec est par ailleurs particulièrement bien représenté dans le corpus PFC, grâce à 32 points d'enquêtes et 440 locuteurs répartis sur le territoire, incluant 34 locuteurs montréalais, enregistrés entre 2009 et 2017 (Côté 2014 ; Côté & Saint-Amant Lamy à paraître).

La combinaison des données de Montréal 1971, telles que décrites dans Ameringen (1977), Ameringen & Cedergren (1981) et De Jong (1993), et des codages effectués sur les conversations du corpus PFC permet d'explorer au moins trois perspectives : (1) Comment a évolué la liaison à Montréal depuis 1971 ? (2) La liaison montréalaise présente-t-elle des spécificités par rapport aux autres régions du Québec ? (3) Comment la liaison montréalaise se positionne-t-elle par rapport au reste de la francophonie ?

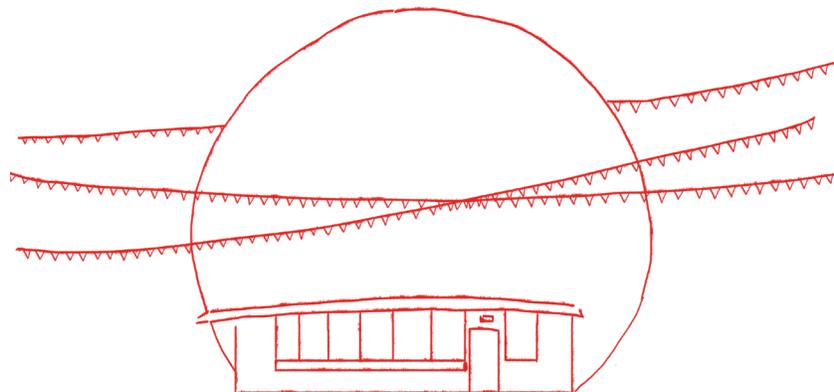
Pour les aborder, nous explorerons en particulier trois catégories de liaison :

- Les liaisons fréquentes au Québec, mais ne correspondant pas à la norme prescriptive : liaison en [t] après *suis* et *es* (Côté 2016), liaison en [l] après *ça* (Morin 1982) ;
- Les liaisons catégoriques dans la norme prescriptive (et d'usage en Europe), mais variables au Québec : liaison en [z] après *ils*, liaison en [n] après *on* (de Jong 1993) ;
- Les liaisons variables tant en Europe qu'au Québec, mais avec des fréquences différentes : liaison en [t] après *est* et *sont* nettement plus fréquente au Québec, liaison en [z] après les prépositions monosyllabiques (*chez*, *dans*) plus fréquentes en Europe (Côté 2017).

Sur la base de plus de 30 000 contextes de liaison codés pour l'instant dans 16 points d'enquête de PFC-Québec, dont 13 712 à Montréal, on observe un rapprochement entre le Québec et l'Europe, sans différence marquée entre Montréal et le reste du Québec. Seule l'absence de [z] après *ils* se maintient fermement, alors que la liaison en [l] après *ça* paraît moins fréquente à Montréal que dans l'ensemble du Québec. Ces liaisons mettent en lumière plusieurs facteurs potentiellement conflictuels et dont il s'agit de distinguer les effets : la norme prescriptive, la liaison comme marqueur stylistique positif, la norme d'usage européenne et l'analogie pouvant opérer entre différents contextes de liaison.

Références bibliographiques

- Ameringen, Arie van. 1977. La liaison en français de Montréal. Mémoire de maîtrise, Université du Québec à Montréal.
- Ameringen, Arie van & Henrietta J. Cedergren. 1981. Observations sur la liaison en français de Montréal. In David Sankoff & Henrietta J. Cedergren (éds) *Variation omnibus*. Edmonton : Linguistic research, 141-149.
- Côté, Marie-Hélène. 2011. French liaison. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (éds) *Companion to phonology*. Malden : Wiley-Blackwell, 2685-2710.
- Côté, Marie-Hélène. 2014. Le projet PFC et la géophonologie du français laurentien. In J. Durand, G. Kristoffersen & B. Laks, avec la collaboration de J. Peuvergne (éds) *La phonologie du français : normes, périphéries, modélisation*. Nanterre : Presses Universitaires de Paris Ouest, 175-198.
- Côté, Marie-Hélène. 2016. Variation in Canada : Trois-Rivières in Quebec. In Sylvain Detey, Jacques Durand, Bernard Laks & Chantal Lyche (éds) *Varieties of spoken French*. Oxford : Oxford University Press, 449-462.
- Côté, Marie-Hélène. 2017. La liaison en diatopie : esquisse d'une typologie. *Journal of French Language Studies* 27 : 13-25.
- Côté, Marie-Hélène & Hugo Saint-Amant Lamy. À paraître. The « Phonologie du français contemporain » project in Quebec : methodological and dialectometric considerations. In Elissa Pustka, Carmen Quijada Van den Berghe & Verena Weiland (éds) *Corpus dialectology : from methods to theory (French, Italian, Spanish)*. Amsterdam/Philadelphie : John Benjamins.
- De Jong, Daan. 1993. Sociophonological aspects of Montreal French liaison. In William J. Ashby, Marianne Mithun, Giorgio Peressinotto & Eduardo Raposo (éds) *Linguistic perspectives on the Romance languages. Selected papers from LSRL XXI*. Amsterdam/Philadelphie : John Benjamins, 127-137.
- Detey, Sylvain, Jacques Durand, Bernard Laks & Chantal Lyche (éds) 2016. *Varieties of spoken French*. Oxford : Oxford University Press.
- Durand, Jacques, Bernard Laks, Basilio Calderone & Atanas Tchobanov. 2011. Que savons-nous de la liaison aujourd'hui ? *Langue française* 169 : 103-135.
- Durand, Jacques, Bernard Laks & Chantal Lyche. 2002. La phonologie du français contemporain. Usages, variétés et structure. In Claus D. Pusch & Wolfgang Raible (éds) *Romanistische Korpuslinguistik. Korpora und gesprochene Sprache / Romance Corpus Linguistics. Corpora and Spoken Language*. Tübingen : Narr, 93-106.
- Morin, Yves-Charles. 1982. De quelques [l] non étymologiques dans le français du Québec : notes sur les clitiques et la liaison. *Revue québécoise de linguistique* 11 : 2. 9-47.
- Tousignant, Claude. 1978. La liaison consonantique à Montréal. Mémoire de maîtrise, Université de Montréal.



Trajectoire de l'omission de *il* dans les constructions impersonnelles à Montréal

Claire Djuikui Dountsop (Université de Montréal)

L'étude du *il* impersonnel a permis de classer le français oral comme une langue *semi-prodrop* (Miller and Monachesi 2003, Zimmermann and Kaiser 2014, Rizzi 1986). Ce statut a été confirmé dans des études sur différentes variétés de français parlé (Auteure et al. 2020, Widera 2017), mais l'alternance entre la réalisation et l'omission de *il* reste variable, tel qu'exemplifié en (1) et (2). Dans le cas de Montréal, nos travaux antérieurs ont montré un changement en cours vers l'omission de *il*, avec un renversement de cette tendance dans les années 80. Dans le présent travail, nous confirmons ces tendances, mais démontrons que le renversement observé avait déjà commencé dans les années 70.

Les données proviennent de trois corpus stratifiés, faits d'entrevues semi-dirigées. Celles-ci ont été recueillies en 1971 (N=120) (Sankoff et al. 1976), en 1984 (N=72) (Thibault et Vincent 1990) et en 2012 (N=50) (Blondeau et al. 2012). Le corpus de 1984 comprend 60 des locuteurs enregistrés une première fois en 1971. Les analyses, faites avec Goldvarb X, sont basées sur 5197 occurrences impliquant les verbes *falloir* et *sembler*. Chaque occurrence a été codifiée pour la présence vs l'absence de *il*, ainsi que pour des facteurs extralinguistiques (le sexe, l'âge, la classe socioéconomique) et linguistiques (type de proposition, présence de l'auxiliaire, négation avec *pas*).

Les tendances générales suggèrent un changement en cours dans la communauté en temps réel, en faveur de l'omission de *il*, le taux d'omission étant de 47 % en 1971, de 69 % en 1984 et de 66 % en 2012. La séparation des locuteurs de 1971 enregistrés une fois (1971a) de ceux qui l'ont été deux fois (1971b) montre que le changement est bidirectionnel, d'abord en faveur de l'omission (1971a) et ensuite en faveur de la réalisation de *il* (1971b, 1984 et 2012). Notre analyse met en évidence le rôle du genre et de la classe socioéconomique dans le changement. Les jeunes hommes de la classe haute mènent le changement en faveur de l'omission alors que le changement vers la réalisation est mené par les jeunes femmes des classes intermédiaire (1971b et 2012) et basse (1984).

Le verbe *falloir*, plus utilisé, favorise l'omission de *il*, confirmant ainsi l'effet de fréquence observé dans les études antérieures (Auteure et al. 2020, Fonseca-Greber 2004, Zimmerman & Kaiser 2014, Culbertson & Legendre 2014). De plus, notre analyse confirme que les propositions principales (3) et la présence de la particule négative *pas* (4) favorisent également l'omission de *il*, alors que la présence de l'auxiliaire *avoir* (5) la défavorise. Dans les trois périodes étudiées, les mêmes tendances sont observées, ce qui témoigne de la stabilité de la grammaire dans la communauté.

Notre étude confirme qu'il y a un renversement dans la trajectoire de l'omission de *il* dans la communauté montréalaise. On observe un changement en cours bidirectionnel, d'abord vers l'omission et ensuite vers la réalisation. En plus de définir la trajectoire de l'omission, ce travail constitue un argument de plus en faveur du statut *semi-prodrop* du français oral.

Exemples

- (1) a. **Ø** Faut je me rende à Tofino demain matin. (Montréal 2012)
b. **I** faut vraiment employer le terme exact. (Montréal 1971)
- (2) a. J'ai dit "**Ø** me semble ' l'année passée c'était dix." (Montréal 2012)
b. **I** me semble que c'était gai. (Montréal 1971)
- (3) a. Des fois **Ø** fallait arrêter parce qu'on allait au chalet. (Montréal 1984)
b. Mes frères s'en plaignaient beaucoup quand **i** fallait le bêcher deux fois par année. (Montréal 2012)
- (4) a. Oh **Ø** faut pas je me trompe parce que je vas être dans la mardo là. (Montréal 2012)
b. S'il veut réussir le moins **i** faut qu'il en fasse de la médecine. (Montréal 1984)
- (5) a. Pour commencer **Ø** faudrait je sache qu'est-ce-que c'est en gros. (Montréal 1971)
b. **I** a fallu qu'elle s'habitue elle aussi. (Montréal 1971)

Références bibliographiques

Auger, J. (1994), *Pronominal Clitics in Québec Colloquial French : A Morphological Analysis* (thèse de doctorat), University of Pennsylvania (Philadelphia).

Auteure et al. (2020), « A Real-time Analysis of the Variable Use of Expletive “il” in Montréal French », *University of Pennsylvania Working Papers in Linguistics*, Vol. 26, Iss. 2, Article 8.

Blondeau, H. ; F. Martineau ; M. Tremblay ; Frenette, Y. (2012), *Sous-corpus variationniste de Hochelaga-Maisonneuve 2012, du Corpus FRAN* (dir. F. Martineau).

Culbertson, J. (2010), *Convergent evidence for categorical change in French : from subject clitic to agreement marker*. *Language* 86 : 1, p. 85-132.

Culbertson, J. et G. Legendre (2014), « Prefixal Agreement and Impersonal “il” in Spoken French : Experimental Evidence », *Journal of French Language Studies*, vol. 24, n° 1, p. 83-105.

Fonseca-Greber, Bonnie B. (2004), « Zero Marking in French Impersonal Verbs : A Counter Trend to Clitic Morphologization ? », dans Marc Ettliger, Nicholas Fleisher et Mischa Park-Doob (dir.), *Proceedings of the 30th Annual Meeting of the Berkeley Linguistics Society*, Berkeley, Berkeley Linguistics Society, p. 81-92.

Miller, P. et P. Monachesi (2003), « Les pronoms clitiques dans les langues romanes », dans Danièle Godard (dir.), *Les langues romanes. Problèmes de la phrase simple*, Paris, CNRS éditions, p. 67-123.

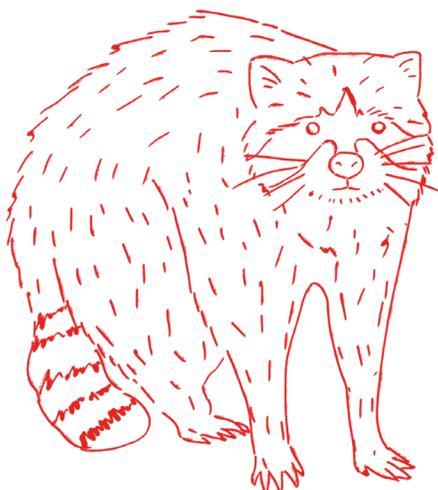
Rizzi, L. (1986), *Null Objects in Italian and the Theory of pro*, *Linguistic Inquiry* 17, p. 501-557.

Sankoff, D., G. Sankoff, S. Laberge et M. Topham (1976), *Méthode d'échantillonnage et utilisation de l'ordinateur dans l'étude de la variation grammaticale, La sociolinguistique au Québec, Cahiers de linguistique de l'Université du Québec*, 6 : 85-125.

Torres Cacoullos, R. & C. E. Travis (to appear), *Variationist typology : Shared probabilistic constraints across (non-)null subject languages*, *Linguistics*.

Widera, Carmen (2017), « L'emploi du pronom explétif *il* en français oral contemporain », Paper presented at Association for French Language Studies, Toronto.

Zimmermann, M. & G. Kaiser (2014), *On expletive subject pronoun drop in Colloquial French*. *French Language Studies* 24. Cambridge University Press, p. 107-126.



Le corpus « Traits d'union » : les représentations du français des intervenants auprès des Néo-Montréalais

Mireille Elchacar (Université TÉLUQ), Angéline Martel (Université TÉLUQ) et Nancy Gagné (Université TÉLUQ)

Notre proposition de communication se situe dans l'axe 1 de l'appel à communications (« La recherche sur le français de Montréal et des grands centres urbains de la francophonie »). Nous souhaitons étudier les représentations du français des intervenants travaillant auprès des collectivités ethnoculturelles de Montréal et ses environs. Cette recherche s'inscrit dans la suite des travaux du « Projet Traits d'union : compétences interculturelles en action », au carrefour des enjeux linguistiques et culturels.

Mené conjointement par l'Institut Jacques-Couture² et le Groupe d'expertise pour le développement des cités interculturelles au Québec, le projet *Traits d'union* a pour objectif de répertorier les besoins en compétences et communication interculturelles des intervenants de divers services publics. Pour évaluer la situation, onze groupes de discussion ont été organisés avec 85 personnes travaillant en proximité avec les collectivités ethnoculturelles, surtout à Montréal mais également dans les banlieues de Longueuil et Laval. Cinq secteurs ont été retenus : les bibliothèques municipales, le Service de police de Montréal, Incendie Montréal, Services Québec et Urgences-Santé. Les transcriptions de ces entretiens composent le *corpus Traits d'union*, à partir duquel nous travaillerons. La thématique de la langue a émergé spontanément comme fondamentale lors des échanges : « Les participants des différents groupes mettent l'accent sur l'enjeu que représente la langue. » (Gagné et al., 2020 : 42)

Depuis les années 2000, plusieurs études ont porté sur la représentation du français (Poll, 2005), de ses variétés géographiques (Moreau et al., 2007), du français québécois (Kircher, 2012 ; Laur, 2002), ou, plus récemment, de différents accents au sein du Québec (Remysen et al., 2020). Certaines recherches abordent la représentation du français des populations immigrantes (Harvey, 2016 ; Amireault et Lussier, 2008 ; Kircher, 2009 ; Maurais, 2008 ; Benzakour, 2004). D'autres études se penchent sur les représentations culturelles dans un contexte d'apprentissage du français (Amireault, 2007).

L'originalité de notre recherche s'avère de mettre en lumière les représentations du français dans un contexte de relations interculturelles. Le corpus *Traits d'union* permet d'accéder aux représentations des intervenants, qui eux-mêmes entretiennent déjà des rapports aux langues françaises, anglaises et étrangères en sol québécois. Comment les intervenants perçoivent le rôle du français dans leurs interactions avec les populations immigrantes ? Les enjeux de la variation géographique du français et du français québécois sont-ils présents dans le corpus ? Comment est perçu l'usage de l'anglais dans de tels contextes ? Dans quelles circonstances la connaissance des lois linguistiques joue-t-elle parfois sur les représentations des intervenants : « [s]elon les participants, la langue représente parfois un mur qu'il est difficile d'abattre dès lors que ses fondements reposent sur les législations » (Gagné et al., 2020 : 42).

En conclusion, nous tenterons de dégager les situations interculturelles dans lesquelles la langue est problématique et d'indiquer les solutions apportées par les intervenants.

² Affilié à l'Université TÉLUQ, l'Institut Jacques-Couture a pour principal objectif de soutenir les innovations en enseignement, en recherche et dans les services aux collectivités en lien avec la thématique « Accueils, Échanges, Sociétés ».

Références bibliographiques

Amireault, V. (2007). Représentations culturelles et identité d'immigrants adultes de Montréal apprenant le français. (Thèse de doctorat). Université McGill.

Amireault, V. et Lussier, D. (2008). *Représentations culturelles, expériences d'apprentissage du français et motivations des immigrants adultes en lien avec leur intégration à la société québécoise : étude exploratoire*. Québec : Office québécois de la langue française.

Benzakour, F. (2004). « Les stéréotypes associés aux constructions sur la langue dans le contexte d'immigration récente au Québec », *Québec français*, 132, 65-67.

Gagné, N., Martel, A., Bédard, J.-L., Bleton, P., Massana Macia, M., Rioux, G., Wilson, K. et al. (2020). *PROJET TRAITS D'UNION : compétences interculturelles en action. Rapport de recherche portant sur cinq secteurs d'intervention à Montréal, Longueuil et Laval* (Québec).

Harvey, Marie-Hélène (2016). « Enseignement du français québécois et exposition à ses différents usages : représentations d'apprenants immigrants adultes à Montréal » Mémoire. Montréal (Québec, Canada), Université du Québec à Montréal, Maîtrise en didactique des langues.

Poll, B. (2005) *Le français, langue pluricentrique ? Études sur la variation diatopique d'une langue standard*. Frankfurt am Main : Peter Lang.

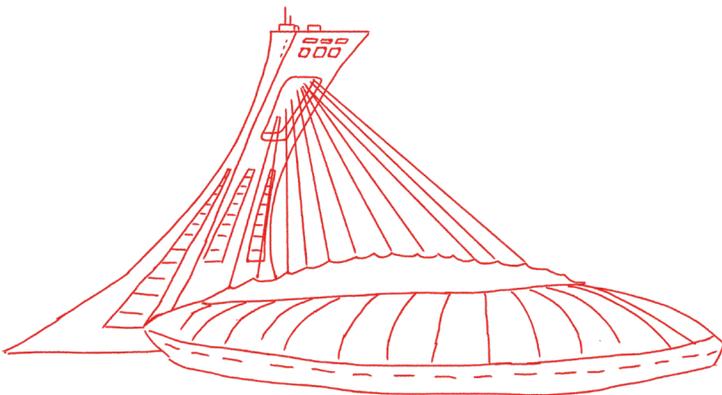
Kircher, R. (2009). Language attitudes in Quebec : A contemporary perspective. (Thèse de doctorat). Université Queen Mary.

Kircher, R. (2012). « How pluricentric is the French language ? An investigation of attitudes towards Quebec French compared to European French », *Journal of French Language Studies*, 22(03), 345-370.

Laur, E. (2002). « La qualité, le statut et la perception du français au Québec » dans B. R. Y. Bourhis (dir.), *L'Aménagement linguistique au Québec : 25 ans d'application de la Charte de la langue française*, Québec : Office québécois de la langue française.

Maurais, J. (2008). *Les Québécois et la norme*. Québec, Office québécois de la langue française. Récupéré à https://www.oqlf.gouv.qc.ca/etudes/etude_07.pdf

Moreau, M.-L., Bouchard, P., Demartin, S., Gadet, F., Guerin, E., Harmegnies, B., Huet, K., Laroussi, F., Prikhodkine, A., Singy, P., Thiam, N. et Tyne, H. (2007) *Les accents dans la francophonie : une enquête internationale*. Fernelmont : Éditions modulaires européennes.



Annotation et analyse de données sociophonologiques sur grands corpus : présentation de la plateforme Phonometrica

Julien Eychenne (Université de Sherbrooke) et Léa Courdès-Murphy (Université Toulouse – Jean Jaurès)

Le tournant du XXI^e siècle a révélé un intérêt croissant pour l'analyse des corpus oraux, comme l'attestent par exemple le développement de la socio-phonologie et de la phonologie de corpus (Durand *et al.* 2014, Detey *et al.* 2016, Honey 2017, Przewozny *et al.* 2020). Ce phénomène a considérablement élargi la communauté des spécialistes de la parole au-delà de son noyau d'origine, ce qui a naturellement abouti à l'émergence de nouveaux besoins et de nouvelles attentes en termes de gestion et d'analyse de données orales. Dans cette communication, nous offrons une présentation de la plateforme logicielle Phonometrica³, un outil pour l'annotation et l'analyse de corpus oraux alignés sur le signal sonore disponible sous licence libre, qui fonctionne sur les trois principaux systèmes d'exploitation (Windows, macOS, Linux) et qui offre une interface conviviale et facile à utiliser, aussi bien pour des chercheurs et des chercheuses confirmé.e.s que pour des étudiant.e.s.

Nous présenterons tout d'abord l'architecture générale du logiciel, en montrant les fonctionnalités de gestion de projet, de création d'annotations et d'ajout de métadonnées. Nous discuterons notamment sur le modèle de données utilisé pour représenter les annotations, à savoir les *graphes d'annotation* (Bird & Liberman 2001), un formalisme qui a été développé pour faciliter le traitement et l'échange d'annotations multi-niveaux alignées sur le signal. Nous montrerons également les possibilités d'interopérabilité avec plusieurs logiciels existants, dont Praat (Boersma & Weenink 2021).

Dans la deuxième partie, nous ferons une démonstration du moteur de recherche qui permet d'extraire des concordances sur des chaînes textuelles ou des variables sociophonologiques codées symboliquement, telles que la liaison en français. Au-delà des requêtes simples permettant l'extraction de concordances en contexte sur un seul niveau d'annotation, le logiciel offre la possibilité d'effectuer des requêtes complexes sur plusieurs niveaux. Pour ce faire, il distingue trois types de relations : *dominance* (ex. : des syllabes qui dominent des segments), *précédence* (ex. : un déterminant précède un nom dans un niveau d'annotation en parties du discours) et *alignement* (ex. : un mot est aligné avec sa partie du discours). Les métadonnées du projet peuvent par ailleurs être exploitées pour filtrer les requêtes et peuvent être exportées avec les résultats vers un format tabulé afin d'effectuer une analyse statistique.

Enfin, nous présenterons les fonctionnalités d'extraction de mesures acoustiques, qui s'appuient sur le moteur de recherche. Elles permettent non seulement d'effectuer différents types de mesures pour l'ensemble des concordances d'une requête mais aussi, pour chaque concordance, de retourner vers le signal afin de vérifier et au besoin corriger les mesures. Ces fonctionnalités simplifient et rendent plus accessible le travail d'analyse acoustique, qui demande en général l'écriture de scripts dédiés pour extraire les données, ainsi qu'un fastidieux travail de vérification manuelle.

Références bibliographiques

- Bird, Steven & Mark Liberman. 2001. A Formal Framework for Linguistic Annotation. *Speech Communication* 33(1-2). 23-60. [https://doi.org/10.1016/S0167-6393\(00\)00068-6](https://doi.org/10.1016/S0167-6393(00)00068-6).
- Boersma, Paul & David Weenink. 2021. Praat : doing phonetics by computer [Computer program]. Version 6.1.38, retrieved 2 January 2021 from <http://www.praat.org/>.
- Detey, Sylvain, Jacques Durand, Bernard Laks & Chantal Lyche (éds). 2016. *Varieties of Spoken French*. Oxford : Oxford University Press.
- Durand, Jacques, Ulrike Gut & Gjert Kristoffersen (dir.). 2014. *The Oxford Handbook of Corpus Phonology*. Oxford : Oxford University Press.
- Durand, Jacques, Bernard Laks & Chantal Lyche. 2002. La phonologie du français contemporain : usages, variétés et structure. In C. Pusch & W. Raible (dir.), *Romanistische Korpuslinguistik-Korpora und gesprochene Sprache/Romance Corpus Linguistics : Corpora and Spoken Language*. Tübingen : Gunter Narr Verlag, 93-106.
- Honey, John. 2017. Sociophonology. In Coulmas, Florian (éd.). *The Handbook of Sociolinguistics*, 92-106. Oxford : Wiley-Blackwell.
- Przewozny, Anne, Cécile Viollain & Sylvain Navarro (éds). 2020. *The Corpus Phonology of English : Multifocal Analyses of Variation*. Edinburgh : Edinburgh University Press.

³ <http://www.phonometrica-ling.org>

LangAge au cours de la vie : les péripéties linguistiques de l'âge avancé

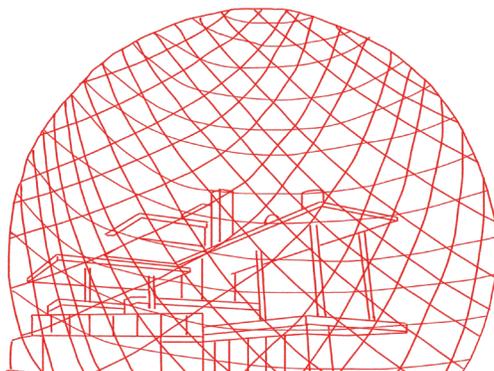
Annette Gerstenberg (Universität Potsdam) et Marta Lupica Spagnolo (Universität Potsdam)

Le corpus de Montréal a permis d'observer la variable *âge* dans de nombreuses perspectives : la stratification des âges représentée dans le corpus rend possible l'analyse détaillée de la distribution des variantes, laquelle dans une perspective longitudinale a mis en évidence des éléments substantiels pour la recherche sur le changement linguistique (Sankoff 2017 : 32-43). De plus, les auteurs du corpus ont accordé une attention particulière à l'importance de la notion de génération dans cette dynamique. Ce concept n'a pas été, comme c'est souvent le cas, lié à un groupe d'âge de manière purement quantitative, mais a été soigneusement situé dans son contexte historique et social (Clermont & Cedergren 1979 : 27).

Le corpus LangAge présenté ici met en avant cette perspective longitudinale, en tant que corpus de spécialité, et se consacre exclusivement à un seul et unique groupe d'âge et aux traits générationnels qui le caractérisent (Gerstenberg 2011). Cette spécialisation qui s'attache en premier lieu à décrire la catégorie démographique de 70 ans et plus est possible par la proximité thématique avec les ESLO (Bergounioux et al. 1992 ; Baude & Dugua 2011), et de leur conception sociologiquement large. L'idée de représenter l'âge avancé « normal » (Brouillet & Sysseau 2000) au moyen d'un corpus sociolinguistique dans une perspective longitudinale s'est développée en continu à travers plusieurs prises de contact (2005, 2012, 2015/16 ; El Sherbiny Ismail et al. 2022). De plus, différents genres communicatifs composent l'ensemble des enregistrements disponibles : plus de 110 entretiens biographiques avec plus de 740 000 tokens sont au cœur du projet, complété par des ateliers d'écriture organisés dans le cadre d'une université populaire, des conversations à bâtons rompus, ainsi qu'une étude individuelle longitudinale, avec des enregistrements quasi annuels répartis sur une période de 9 ans (Gerstenberg & Hamilton 2022).

Quelles solutions ont été trouvées face aux défis de la construction longitudinale d'un corpus de données orales ? Les interviews du corpus LangAge ont été rendues disponibles, dans la mesure de l'obtention des autorisations des participants, sur un serveur accessible via l'application web LaBB-CAT (2012-2021 ; Fromont & Hay 2012). Différentes options de LaBB-CAT comme la composition de sous-corpus et le choix ciblé de séries d'entretiens permettent de profiter de la dimension longitudinale de LangAge. Les questions éthiques liées au traitement des données sensibles impliquent non seulement la propre biographie des participants, mais des informations ayant également trait aux familles et réseaux de connaissances. Les principes sur lesquels s'est appuyée la réalisation de l'anonymisation en tiennent compte, ainsi que la création de plusieurs types d'utilisateurs, en fonction de la nature des données autorisées ou non pour leur diffusion publique. Les standards employés et les autorisations disponibles permettent, dans la version finale du corpus, de respecter les principes FAIR (Findable, Accessible, Interoperable, Reusable ; GO FAIR 2021) récemment avancés au *gold standard* du traitement des données aux humanités numériques. L'encodage XML assure notamment l'interopérabilité des données et leur usage dans d'autres contextes, rendu possible par un modèle de licence *creative commons*. En outre, l'outil LaBB-CAT permet l'extraction des transcriptions alignées, en format *.textGrid, des recherches KWIC avec ses fichiers audios correspondants (Fuchs et al. 2021).

La phase la plus récente de l'utilisation du corpus concerne l'implémentation de l'annotation semi-automatique des traits morphosyntaxiques (parts of speech, lemmatisation) et les perspectives de recherche qui en résultent.



Références bibliographiques

LangAge = Gerstenberg, Annette & équipe LangAge. 2017-. *LangAge corpora*. University of Potsdam : Department of Romance Studies. www.langage-corpora.org.

Baude, Olivier & Céline Dugua. 2011. (Re)faire le corpus d'Orléans quarante ans après : "quoi de neuf, linguiste ?" . *Corpus* 10. 99-118. <https://doi.org/10.4000/corpus.2036>.

Bergounioux, Gabriel, Jean Baraduc & Céline Dumont. 1992. L'Étude Socio-Linguistique sur Orléans (1966-1991): 25 ans d'histoire d'un corpus. *Langue française* (1). 74-93. 10.3406/lfr.1992.5812.

Brouillet, Denis & Arielle Syssau (eds.). 2000. *Le vieillissement cognitif normal. Vers un modèle explicatif du vieillissement*. Bruxelles : De Boeck ; Larcier.

Clermont, Jean & Henrietta J. Cedergren. 1979. Les r de ma mère sont perdus dans l'air. In Pierrette Thibault (ed.), *Le français parlé : Études sociolinguistiques*, 13-28. Amsterdam : Linguistic Research.

El Sherbiny Ismail, Eman, Annette Gerstenberg, Marta Lupica Spagnolo, Friederike Schulz & Anne Vandenbroucke. 2022. L'âge avancé en perspective longitudinale et ses outils : LangAge, un corpus au pluriel. *SHS Web Conf. (SHS Web of Conferences) – 8^e Congrès Mondial de Linguistique Française* 138. 10003 [1-14]. DOI [10.1051/shsconf/202213810003](https://doi.org/10.1051/shsconf/202213810003).

Fromont, Robert & Jennifer Hay. 2012. LaBB-CAT : an Annotation Store. In Paul Cook & Scott Nowson (eds.), *Proceedings of Australasian Language Technology Association Workshop 2012*, 113-117. Dunedin (NZ) : Otago University. <https://aclanthology.org/U12-1015.pdf>.

Fuchs, Susanne, Laura L. Koenig & Annette Gerstenberg. 2021. A Longitudinal Study of Speech Acoustics in Older French Females : Analysis of the Filler Particle *eah* across Utterance Positions. *Languages* 6(4). 211. DOI [10.3390/languages6040211](https://doi.org/10.3390/languages6040211).

Gerstenberg, Annette. 2011. *Generation und Sprachprofile im höheren Lebensalter : Untersuchungen zum Französischen auf der Basis eines Korpus biographischer Interviews* (Analecta Romanica 76). Frankfurt am Main : Vittorio Klostermann.

Gerstenberg, Annette & Heidi E. Hamilton. 2022. Older adults' conversations and the emergence of "narrative crystals" : a new approach to frequently told stories. *Narrative Inquiry*. 1-34. DOI [10.1075/ni.21075.ger](https://doi.org/10.1075/ni.21075.ger).

GO FAIR International Support and Coordination Office. 2021. *FAIR Principles*. Leiden (NL), Paris (F), Hamburg (D) : GO FAIR. <https://www.go-fair.org/fair-principles/> (2022-06-01).

LaBB-CAT = Fromont, Robert. 2012-2021. *LaBB-CAT : Language, Brain & Behaviour Corpus Analysis Tool*. University of Canterbury (NZ) : New Zealand Institute of Language, Brain and Behaviour. Online : <https://labbcats.canterbury.ac.nz/system> (2022-06-01).

Sankoff, Gillian. 2017. Before There Were Corpora. In Suzanne Wagner & Isabelle Buchstaller (eds.), *Panel Studies of Variation and Change* (Routledge studies in language and intercultural communication), 21-51. New York : Routledge.

Étudier le français parlé à Montréal : l'apport du corpus Ébullition

Anna Giaufret (Université de Gênes), Wim Remysen (Université de Sherbrooke) et Philippe Rioux (Université de Sherbrooke/ Université Concordia)

Parmi les nombreux outils qui permettent d'analyser le français parlé à Montréal, nous présenterons ici le corpus Ébullition, une base de données contenant les transcriptions de l'élément textuel des bandes dessinées québécoises, dans laquelle il sera possible de faire des recherches sur la base de plusieurs paramètres. Au sein du corpus, qui a été constitué afin qu'il soit représentatif de la bande dessinée québécoise au fil de son histoire, il sera possible de sélectionner les œuvres sur la base du lieu d'édition ou de l'auteur, afin de délimiter un corpus montréalais.

Ce corpus, ainsi que l'ont démontré les études précédentes (notamment Giaufret 2021), constitue un précieux matériau pour l'analyse du français parlé, car il en fournit une représentation graphique qui permet de travailler sur les stéréotypes, les marqueurs et les indicateurs (selon la terminologie de Labov 2001). Il permet également de se concentrer sur les formes néographiques qui constituent la frontière entre oral et écrit. Par la nature même du genre littéraire à laquelle elle appartient, la bande dessinée est aussi un corpus riche pour l'étude des pratiques lexicales, notamment les usages les plus marqués socialement et stylistiquement parlant, souvent plus difficiles à saisir. Et pour ce qui est de la bande dessinée montréalaise plus particulièrement, elle fournit aussi souvent l'occasion d'observer de plus près les questions liées à la cohabitation du français et de l'anglais dans la ville, ces deux langues se côtoyant parfois dans les échanges entre les protagonistes.

Dans le cadre de notre communication, nous nous limiterons à présenter les trois points suivants. Premièrement, la sélection d'œuvres à l'origine de la constitution du corpus et la place qu'y occupe la bd montréalaise. Deuxièmement, la délicate tâche de transcription du corpus, travail qui a été réalisé avec le protocole CBML du langage XML TEI. Ce langage permet un traitement très fin des bandes dessinées retranscrites, grâce notamment à la prise en compte des différents éléments constitutifs des bd, permettant ainsi des requêtes textuelles ciblées (dans les bulles, dans les récitatifs ou encore dans les éléments textuels qui apparaissent en toile de fond). Troisièmement, quelques études de cas permettant d'illustrer l'exploitation des contenus du corpus, par exemple le recours à des onomatopées et à des idéophones ou encore l'utilisation de formes néographiques à la frontière entre représentation de la prononciation et tentatives de codification.

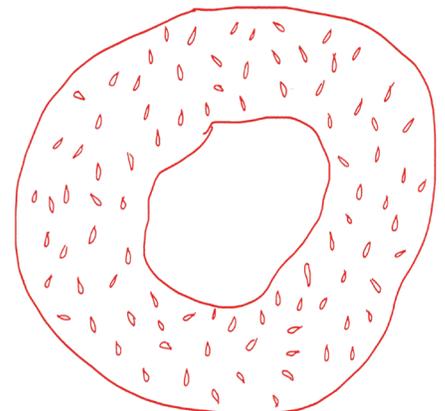
Références bibliographiques

Dargnat, Mathilde (2007), « L'oral au pied de la lettre : raisons et déraisons graphiques », *Études françaises*, vol. 1, n° 43, p. 83-100.

Giaufret, Anna (2021), *La représentation de l'espace urbain et du français parlé montréalais dans la bande dessinée*, Québec, PUL.

Koch, Peter, et Wulf Österreicher (2001), « Langage parlé et langage écrit », dans Günter Holtus, Michael Metzeltin et Christian Schmitt (dir.), *Lexikon der Romanistischen Linguistik*, vol. 1. Tübingen, Niemeyer, p. 584-627.

Labov, William (2001), *Principles of Language Change I : Social Factors*, Oxford, Blackwell.



Changements consonantiques en français parisien de 1989 à 2015 : une étude en temps réel socialement différenciée

Anita Berit Hansen (Université de Copenhague)

Le français de France a souffert d'un manque d'études phonétiques en temps réel, comparé au québécois (Hornsby & Pooley 2001 ; Gadet 2004). Les travaux basés sur la reproduction de l'enquête d'Orléans présentent à cet égard une exception (Eshkol-Taravella et al. 2011 ; Dugua & Baude 2017 ; Abouda & Skrovec à paraître), ainsi que le font certaines études longitudinales sur les professionnels de la parole publique (Smith 2000 ; Candea et al. 2013 ; Laks & Peuvergne 2017). Le grand centre urbain de Paris, par contre, reste peu étudié sous cet angle, mis à part les études vocaliques de Hansen (1994, 1997, 2001, 2021) et de Hansen & Juillard (2011).

Cet état des choses nous laisse avec une lacune empirique pour la Capitale à propos de deux phénomènes consonantiques pourtant suggérés en progrès par certains, vu leur fréquence élevée parmi les jeunes et leur présence parmi les présentateurs ou hommes politiques, i.e. la chute post-obstruante des liquides /l/ et /r/ en fin de mot (*tab(le)*, *quat(re)*) (cf. Wachs 1997 ; Pustka 2011), et la palatalisation ou l'affrication (pal/aff) des occlusives dentales et vélaires (*voiture* [vwaʦʏr], *placard* [plakʁaʁ]) (cf. Jamin 2005 ; Trimaille 2010 ; Paternostro 2017). Les deux phénomènes ont été associés au français populaire parisien, le dernier étant parfois lié également aux milieux multiculturels. Quelles ont été leurs trajectoires respectives auprès de Parisiens de classe moyenne pendant les dernières décennies ?

Nous opposons ici des interviews recueillies à Paris en 1989-1993 (Hansen 1998) et en 2011-2015 (Hansen ms), en focalisant sur des jeunes de 15 à 25 ans qui viennent des couches moyennes inférieures (CMI) et supérieures (CMS) respectivement (23 personnes ayant tous un lien solide avec la région parisienne). À partir de 420 occurrences d'obstruante +l/r+ə# et de 3500 occurrences de /t/ et /k/, nous constatons d'abord que les taux généraux n'indiquent pas de progrès de la chute des liquides, ni de la pal/aff sur /t/ ou /k/. Des analyses poussées témoignent cependant à la fois d'une propagation sociale du bas vers le haut et d'une généralisation contextuelle de ces phénomènes. Ainsi, contrairement à 25 ans plus tôt, les chutes des liquides sont vers 2015 significativement plus fréquentes parmi les locuteurs CMS que parmi les locuteurs CMI (51,4 % vs 35,1 %, tous contextes de droite confondus), et présentent des progrès devant pause (de 24,3 % à 41 %) surtout dus aux locuteurs favorisés. Pour les occlusives /t/ et /k/, une propagation sociale de la pal/aff ne peut pas être démontrée, l'ampleur de celle-ci restant trois à quatre fois moins élevée parmi les locuteurs CMS que parmi les locuteurs CMI (3-4 % vs 11-12 %, dans les deux corpus, tous contextes de droite confondus), mais le groupe CMS copie désormais les réalisations phonétiques plus bruyantes des CMI en contexte prévocalique (en plus de [tʃ] et [kʃ], aussi [tʃ], [ts], [kç]), et les pal/aff de 2011-2015 touchent des occurrences plus diversifiées que vers 1990 (des /t/ pas seulement devant voyelle antérieure fermée, mais aussi moyenne ; des /k/ pas seulement devant voyelle antérieure, mais aussi postérieure).

Références bibliographiques

- Abouda, Lotfi & Marie Skrovec (éds). (à paraître). *Micro-diachronie de l'oral*, Langages, numéro thématique.
- Candea, Maria, Martine Adda-Decker & Lori Lamel. (2013). « Recent evolution of non-standard consonantal variants in French broadcast news ». *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. 412-416.
- Dugua, Céline & Olivier Baude. (2017). « La liaison à Orléans, corpus et changement linguistique : une première étude exploratoire ». *Journal of French Language Studies*. 27(1) : 41-54.
- Eshkol-Taravella, Iris, Olivier Baude, Denis Maurel, Linda Hriba, Céline Dugua & Isabelle Tellier. (2011). « Un grand corpus oral "disponible" : le corpus d'Orléans 1968-2012 ». *Traitement Automatique des Langues, ATALA*, 2011, *Ressources Linguistiques Libres*. 53(2) : 17-46.
- Gadet, Françoise. (2004). « Mais que font les sociolinguistes ? ». *Langage et Société*. 107(1) : 85-94.
- Hansen, Anita Berit. (1994). « Étude du E caduc : stabilisation en cours et variations lexicales ». *Journal of French Language Studies*. 4(1) : 25-54.
- Hansen, Anita Berit. (1997). « Le nouveau [ə] prépausal dans le français parlé à Paris ». In Perrot, Jean (coord.). *Polyphonie pour Iván Fónagy. Mélanges offerts en hommage à Iván Fónagy par un groupe de disciples, collègues et admirateurs*. Paris, L'Harmattan. 173-198.
- Hansen, Anita Berit. (1998). *Les voyelles nasales du français parisien moderne. Aspects linguistiques, sociolinguistiques et perceptuels des changements en cours*. Copenhagen, Museum Tusulanum Press.

- Hansen, Anita Berit. (2001). « Les changements actuels des voyelles nasales du français parisien : confusions ou changement en chaîne ? ». *La Linguistique*. 37 (2) : 33-47.
- Hansen, Anita Berit. (2021). « Parisian French “unstable e” in word-initial syllables : engaged in a lexically diffused process of stabilization ? Evidence from repeated real-time studies ». In Recasens, Daniel & Fernando Sánchez-Miret (éds). *Sound Change in Romance : Phonetic and Phonological Issues*. München, Lincom Europa. 25-44.
- Hansen, Anita Berit. (ms). *Changements phonétiques à Paris autour du millénaire (1989-2015) : une approche multidimensionnelle*.
- Hansen, Anita Berit & Caroline Juillard. (2011). « La phonologie parisienne à trente ans d'intervalle : les voyelles à double timbre ». *Journal of French Language Studies*. 21 (3) : 313-359.
- Hornsby, David & Tim Pooley. (2001). « La sociolinguistique et les accents français d'Europe ». In Hintze, Marie-Anne, Tim Pooley & Anne Judge (éds). *French accents : Phonological and sociolinguistic perspectives*. London, CILT. 305-343.
- Jamin, Mikaël. (2005). *Sociolinguistic Variation in the Paris Suburbs*, thèse de doctorat non-publiée. University of Kent at Canterbury.
- Laks, Bernard & Julie Peuvergne. (2017). « La liaison en français contemporain dans la parole publique (1999-2015) ». *Journal of French Language Studies*. 27 (1) : 55-72.
- Paternostro, Roberto. (2017). « Les jeunes ont-ils un accent ? ». In Gadet, Françoise (coord.). *Les parlers jeunes dans l'Île-de-France multiculturelle*. Paris, Ophrys. 55-71.
- Pustka, Elissa. (2011). « Le conditionnement lexical de l'élision des liquides en contexte post-consonantique final ». *Langue Française*. 169 : 19-38.
- Smith, Alan. (2000). « Linguistic Change on British and French Public Service Radio », thèse PhD non publiée. University of Newcastle upon Tyne.
- Trimaille, Cyril. (2010). « Consonnes dentales palatalisées/affriquées en français contemporain : indicateurs, marqueurs et/ou variantes en développement ? ». In Abecassis, Michaël & Gudrun Ledegen (éds). *Les Voix des Français*. Oxford/New York, Peter Lang. 89-100.
- Wachs, Sandrine. (1997). « Le relâchement de la prononciation en français parlé en Île de France : analyses linguistique et sociolinguistique par générations », thèse de doctorat non publiée. Université de Paris X, Nanterre.

De la variation diaphasique en français montréalais vernaculaire à travers le multilinguisme rap

Emily B. Leavitt (Université Concordia)

Les études marquantes qui nous ont fourni les corpus de français parlé montréalais de 1971, 1984 et 1995 (Sankoff, Sankoff, Laberge et Topham, 1976 ; Thibault et Vincent, 1990 ; Vincent, Laforest et Martel, 1995) ont suscité un essor des recherches sociolinguistiques québécoises et ont rendu possible la comparaison longitudinale de données contemporaines. L'un des apports de cette recherche est la mise en évidence d'une structure hiérarchique des variétés multiples et distinctes du français montréalais (Bigot, 2011).

L'étude récente qu'ont menée Blondeau et Tremblay (2016) apporte des preuves qualitatives à l'appui d'un processus de diversification diaphasique en cours qui provoque l'émergence d'une nouvelle variété vernaculaire du français montréalais. Il s'agit d'un « vernaculaire populaire émergent » qui se différencie du « vernaculaire traditionnel » par une grande fréquence d'éléments lexicaux empruntés de langues autres que l'anglais, telles que le créole haïtien et l'arabe, ainsi que de segments d'alternance codique intraphrastiques (Blondeau et Tremblay, 2016, p. 32-35).

À notre connaissance, aucune étude n'a encore corroboré ce résultat. Pour cette raison, nous proposons une étude sociolinguistique variationniste quantitative portant sur la variation diaphasique en français montréalais vernaculaire. Nous cibons une communauté de pratique liée à la culture populaire des jeunes montréalais – celle du mouvement *rap québécois* à Montréal – en raison, d'une part, des qualités informelle et plurilingue de ses paroles constatées par Sarkar *et coll.* (Low et Sarkar, 2012, 2014 ; Sarkar et Allen, 2007 ; Sarkar et Winer, 2006) et, d'autre part, de sa diversité culturelle.

Notre objectif principal est d'évaluer la validité du constat de Blondeau et Tremblay (2016) pour qui la diversification fait apparaître une nouvelle variété vernaculaire du français montréalais. Afin d'y parvenir, nous proposons une analyse des propriétés différentielles de cette variété, à savoir, la langue source et la catégorie typologique d'où proviennent les occurrences d'emprunt lexical et d'alternance codique dans les paroles de chansons *rap*.

Nous effectuerons une analyse quantitative multivariée pour en dégager les modèles de variation par rapport à ces deux variables ainsi que le conditionnement éventuel de certains facteurs sociaux (le lieu de naissance, l'origine ethnique, la génération immigrante, l'âge et le genre de l'artiste et l'année de sortie de ses œuvres) et linguistiques (l'intégration et la catégorie grammaticale des emprunts ou la structure et la longueur des segments d'alternance codique). Notre analyse se base sur l'examen d'un corpus audio textuel composé de 2558 chansons recueillies auprès d'un échantillon de 50 artistes. Nous utiliserons le codage Python pour effectuer de manière automatique l'identification de la langue et l'étiquetage des parties du discours des occurrences attestées, ce qui facilitera le traitement de notre corpus à grande échelle.

Dans un premier temps, nous présenterons les recherches antérieures traitant du français montréalais vernaculaire. Nous décrirons ensuite notre corpus, notre typologie des phénomènes multilingues et les grandes lignes de notre méthodologie. Enfin, nous terminerons par la discussion de nos résultats dont les préliminaires tendent à démontrer l'émergence d'une nouvelle variété vernaculaire à partir des années 2000 dans le lyrisme *rap québécois*.

Références bibliographiques

- Bigot, D. (2011). De la norme grammaticale du français parlé au Québec. *Arborescences*, 1(1). <https://doi.org/10.7202/1001939ar>
- Blondeau, H. et Tremblay, M. (2016). Le traditionnel et l'émergent. L'apport de jeunes montréalais issus de l'immigration au français vernaculaire. Dans H. Blondeau et W. Remysen (dirs), *Du local au global : pratiques et idéologies linguistiques en contexte montréalais*, *Cahiers Internationaux de Sociolinguistique*, 10(1), 19-45. <https://doi.org/10.3917/cisl.1602.0019>
- Low, B. et Sarkar, M. (2012). Le plurilinguisme dans les arts populaires, un terrain inexploré? L'étude du langage mixte du rap montréalais en guise d'exemple. *Kinéphanos*, 3(1), 2047. Récupéré de <https://www.kinephanos.ca/2012/rap-montrealais/>
- Low, B. et Sarkar, M. (2014). Translanguaging in the Multilingual Montreal Hip-Hop Community : Everyday Poetics as Counter to the Myths of the Monolingual Classroom. Dans A. Blackledge et A. Creese (dirs), *Heteroglossia as Practice and Pedagogy* (p. 99-118). London : Springer-Verlag.
- Sankoff, D., Sankoff, G., Laberge, S. et Topham, M. (1976). Méthodes d'échantillonnage et utilisation de l'ordinateur dans l'étude de la variation grammaticale. *Cahiers de linguistique de l'Université du Québec*, 6(1), 85-125. <https://doi.org/10.7202/800043ar>
- Sarkar, M. et Allen, D. (2007). Hybrid Identities in Quebec Hip-Hop : Language, Territory and Ethnicity in the mix. *Journal of Language, Identity and Education*, 6(2), 117-130. <https://doi.org/10.1080/15348450701341253>
- Sarkar, M. et Winer, L. (2006). Multilingual Codeswitching in Quebec Rap : Poetry, Pragmatics and Performativity. *International Journal of Multilingualism*, 3(3), 173-192. <https://doi.org/10.2167/ijm030.0>
- Thibault, P. et Vincent, D. (1990). *Un corpus de français parlé. Montréal 84 : Historique, méthodes et perspectives de recherche. Recherches sociolinguistiques*. Québec : Université Laval.
- Vincent, D., Laforest, M. et Martel, G. (1995). Le corpus de Montréal 1995 : adaptation de la méthode d'enquête sociolinguistique pour l'analyse conversationnelle. *Dialangue*, 6(1), 29-46.



La transition linguistique vers une identité non-binaire en français : les récits de personnes queers en Acadie

Isabelle LeBlanc (Université de Moncton)

Dans le cadre d'une ethnographie queer en Acadie, trente personnes interviewées se posent toutes la même question formulée de cette façon par Renaud (pseudonyme) : « comment est-ce que je parle en français de personnes non-binaires ? ». Renaud poursuit sa réflexion en ajoutant : « Pour longtemps, je n'avais pas les outils et j'effectuais des recherches en ligne. Je trouvais des petits morceaux ici et là, mais aussitôt que je trouvais des informations sur les pronoms je ne trouvais rien sur les accords. Je trouve que c'est difficile avec la langue française pis c'est pour ça que je ne pense pas qu'on devrait avoir un standard, mais je pense qu'on devrait avoir un guide pour les options en tout cas ».

En adoptant un cadre théorique appartenant aux recherches sociolinguistiques sur le genre, nous chercherons à comprendre les dynamiques entre normes, réseaux et agentivité dans les discours sur la langue. En effet, les locutaires construisent et déconstruisent leurs rapports aux normes linguistiques dans leurs mises en récit de leurs trajectoires sociolangagières et l'objectif sera de réfléchir au témoignage queer comme une forme de contre-discours ou discours alternatif (Moïse et Hugonnier 2019) – afin de penser le processus langagier queer comme un processus langagier de « déshétérosexualisation » de l'espace public (Chetcuti-Osorovitz 2010).

En mobilisant les récits issus d'un corpus de trente entretiens conversationnels (45 heures) effectués avec des personnes de différentes générations, nous allons sélectionner quelques extraits à partir de critères qui permettent d'examiner comment *se dire* – se raconter est un processus sociolangagier empreint de rapports de pouvoir et de différentes formes de violence linguistique (Kibbey 2019) au sein d'une même communauté francophone.

Cette communication s'inscrit dans une volonté de reconnaître l'importance d'inclure les groupes marginalisés dans les analyses de discours sur l'identité et la langue afin de ne pas se limiter à penser le genre en termes de spectre (féminin-neutre-masculin), mais en termes « d'espace multidimensionnel (sans linéarité) » (Beaubatie 2021) ce qui permet de *désorienter* les points de repère du genre *typique tout en racontant des trajectoires individuelles*.

Réfléchir au « gender history » (Zimman 2009) plutôt que seulement au « gender identity », dans la mesure où la narration de soi permet de comprendre la pluralité des expériences identitaires caractérisée par des dimensions d'actes discursifs qui i) s'inscrivent dans différents espaces discursifs ; ii) évoluent dans le temps ; iii) varient selon le contexte social d'interaction interindividuelle. Cette proposition s'inscrit dans l'axe 3 du colloque (l'étude des récits de vie), notamment en ce qui concerne l'idée que « les corpus sociolinguistiques peuvent ainsi devenir des lieux de mémoire » favorisant une meilleure compréhension des enjeux et motivations identitaires.

Références bibliographiques

- BEAUBATIE, E. 2021, « L'espace social du genre », dans Emmanuel Beaubatie (dir.), *Transfuges de sexe. Passer les frontières du genre*, La Découverte, 135-166.
- CHETCUTI, N. et GRECO, L. (dir.), 2012, *La face cachée du genre. Langage et pouvoir des normes*, Presses Sorbonne Nouvelle.
- CHETCUTI-OSOROVITZ, N. 2010, *Se dire lesbienne : vie de couple, sexualité, représentation de soi*, Payot.
- COUTANT, A. « Langage », dans Juliette Rennes (dir.), *Encyclopédie critique du genre*, La Découverte, 359-368.
- DUCHÊNE A. et MOÏSE C. (dir.), 2011, *Langage, genre et sexualité*, Montréal, Nota Bene.
- ECKERT, P. 2016, « Linguistics, Language and Cognition, Sociolinguistics », *Third Wave Variationism*, Consultation en ligne le 1^{er} mars 2022, [<https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199935345.001.0001/oxfordhb-9780199935345-e-27?print=pdf>]
- GADET, F. 2021, « Changement linguistique », *Langage et société*, HS1, 41-46.
- GRECO, L. 2014, « Les recherches linguistiques sur le genre : un état de l'art », *Langage et société*. Recherches linguistiques sur le genre. Bilan et perspectives, n° 148, 11-29.
- KIBBEY, T. 2019, « Transcriptivism : An ethical framework for modern linguistics », *Proceedings Linguistic Society of America*, vol. 4, n° 45, 1-13.
- MOÏSE, C. et HUGONNIER, C. 2019, « Discours de haine dissimulée, discours alternatifs et contre-discours », *Semen : revue de sémio-linguistique des textes et du discours*, n° 47.
- WITTIG, M. 2018 (réédition du texte de 1978), *La pensée straight*, Éditions Amsterdam.
- ZIMMAN, L. 2009, « "The other kind of coming out" : Transgender people and the coming out narrative genre », *Gender and Language*, vol. 3, n° 1, 54-80.

Documenter les formes jugées inappropriées dans des discours parlementaires : développer un outil d'analyse adapté

Geneviève Lemieux Lefebvre (Université du Québec à Montréal)

Dans le cadre de nos recherches doctorales, nous avons analysé des formes jugées inappropriées produites à l'Assemblée nationale du Québec, lors de la période de questions et réponses orales pour l'ensemble des séances de la 40^e législature (octobre 2012 à mars 2014). Notre corpus s'est construit grâce aux transcriptions détaillées faites après chaque séance parlementaire et accessibles en ligne.

Nous souhaitons rendre compte des pratiques discursives en vigueur à l'Assemblée nationale, que nous identifions comme une communauté de pratique qui fixe ses propres règles de fonctionnement et pour laquelle les échanges sont régis par des procédures bien spécifiques. Dans cette optique, nous avons tenté d'agir en tant qu'observatrice et nous avons sélectionné uniquement les formes pour lesquelles il y a intervention d'un tiers, qui signale explicitement qu'il y a eu production d'une forme qui s'écarte du règlement parlementaire ou qui va au-delà de ce qui est admis dans ces joutes politiques.

Pour parvenir à nos objectifs, nous avons élaboré une grille développée spécifiquement pour le corpus à l'étude et en fonction des critères d'évaluation mis sur pied. Basée à la fois sur des méthodologies de recherches existantes (Harris, 2001 ; Culpeper, 2010, 2011, 2016 ; Bull et Wells, 2012 ; Sivenkova, 2013) et sur les observations préliminaires du corpus parlementaire, cette grille prend en compte l'interaction puisqu'elle regroupe des critères d'analyse utiles tant pour la forme jugée inappropriée (notre objet d'étude principal) que pour l'intervention qui signale cette intervention et le tour de parole qui y fait suite. Cet ensemble de critères offre une meilleure compréhension du contexte d'utilisation des formes jugées inappropriées, puisque l'analyse des interventions facilite l'identification de ce qui est pris en faute, tandis que le tour de parole qui suit mesure l'impact de ce signalement sur le déroulement des séances. Cette analyse fait donc ressortir des éléments propres à l'interaction, ce que la transcription d'échanges suivis comme ceux disponibles sur le site de l'Assemblée nationale nous permet d'exploiter.

Nous avons analysé les formes jugées inappropriées selon deux approches complémentaires, basées sur des méthodologies existantes. Suivant Harris (2001) et Bull et Wells (2012), nous avons identifié quatre structures discursives particulières, parmi lesquelles l'attaque dirigée contre un individu prédomine. En adaptant les formules d'impolitesse de Culpeper (2010, 2011, 2016), nous avons aussi déterminé que les critiques ciblées sont nettement dominantes. L'analyse des interventions a mis en lumière le travail de contrôle en place pour maintenir le décorum parlementaire. La suite des échanges a permis de faire des parallèles avec les travaux réalisés sur les réponses produites au Parlement (Sivenkova, 2013), mais aussi de constater que les interventions mènent majoritairement à des reprises, pour lesquelles les personnes prises en faute ont la possibilité de poursuivre après avoir été interrompues.

Les préoccupations de recherche précédemment évoquées s'inscrivent en complémentarité avec l'axe 2 du présent colloque puisqu'elles touchent à la documentation linguistique et à l'accessibilité à des corpus peu étudiés pour lesquels les outils d'analyse restent à développer.

Références bibliographiques

- Bull, P. et Wells, P. (2012). Adversarial Discourse in Prime Minister's Questions. *Journal of Language and Social Psychology*, 31(1), 30-48.
- Culpeper, J. (2010). Conventionalised impoliteness formulae. *Journal of Pragmatics*, 42(12), 3232-3245.
- Culpeper, J. (2011). *Impoliteness : using language to cause offence*. Cambridge ; New York : Cambridge University Press.
- Culpeper, J. (2016). Impoliteness Strategies. Dans Capone, A. et Mey, J. L. (dir.), *Interdisciplinary Studies in Pragmatics, Culture and Society* (p. 421-445). Cham : Springer International Publishing.
- Harris, S. (2001). Being Politically Impolite : Extending Politeness Theory to Adversarial Political Discourse. *Discourse & Society*, 12(4), 451-472.
- Sivenkova, M. (2013). On the metapragmatics of British, German and Russian political questions and answers. Dans Fetzer, A. (éd.) *The Pragmatics of Political Discourse : Explorations across cultures*. John Benjamin Publishing Compagny. Philadelphie, Pragmatics & Beyond New Series(228), 21-46.

Le rôle de facteurs cognitifs dans l'omission du complémenteur *que* à Montréal

Yiming Liang (Université Paris Cité), Pascal Amsili (Université Sorbonne Nouvelle) et Heather Burnett (Université Paris Cité)

L'omission du complémenteur *que* (1) est une variable sociolinguistique classique en français canadien (Sankoff *et al.*, 1971 ; Sankoff, 1980 ; Warren, 1994, parmi d'autres). Malgré l'identification d'un grand nombre des facteurs explicatifs sociaux et linguistiques, un débat subsiste sur l'indépendance des effets phonologiques et syntaxiques (Connors, 1975 ; Martineau, 1985 ; Dion, 2003). Par ailleurs, si les facteurs cognitifs de l'omission du *that* en anglais ont fait l'objet de nombreux travaux, peu d'études les ont pris en compte pour le français. En particulier, Jaeger (2010) montre que la densité informationnelle (opérationnalisée comme la probabilité qu'un verbe enchaîne une complétive) a un effet prédominant sur l'omission du *that*, devant une dizaine d'autres facteurs. Cela soutient donc l'hypothèse de l'*Uniformité de la densité informationnelle*, qui prédit que, face à deux variantes, un locuteur préférera celle qui distribue l'information de façon plus uniforme sur une phrase (Levy & Jaeger, 2007 ; Jaeger, 2010). Cette hypothèse n'a pas encore été testée dans d'autres langues comme le français.

1. Bien je pense (que) c'est : c'est important ... (locuteur 128, Corpus Montréal 84)

L'objectif de cette étude est donc double : a) utiliser l'ensemble du corpus *Montréal 84* (Thibault & Vincent, 1990) pour vérifier les résultats d'études précédentes, en particulier ceux concernant l'effet du contexte phonologique à droite et du sujet de la complétive ; b) tester deux facteurs cognitifs (la fréquence du verbe matrice et sa probabilité d'enchaîner une complétive), pour voir si l'hypothèse de l'*Uniformité de la densité informationnelle* est aussi valable en français.

Grâce à une annotation semi-automatique du corpus, nous avons extrait 5818 complétives sur dix-sept verbes fréquents, et annoté dix facteurs sociaux, linguistiques et cognitifs pour chaque observation. À l'aide d'un modèle à effets mixtes (Johnson, 2009) contrôlant la variation intrinsèque des locuteurs, notre étude montre que, conformément à la majorité des études antérieures, les effets de l'éducation, de la profession, du contexte phonologique à droite et du sujet de la complétive sont significatifs. Notamment, la hiérarchie de la sonorité dans le contexte suivant *que* a l'effet le plus important sur le choix de prononcer *que*. Conformément à l'hypothèse de l'accessibilité (Ferreira & Dell, 2000), un sujet plus accessible comme *je* et *tu* est corrélé avec un taux d'omission plus élevé qu'un autre pronom, et un sujet nominal défavorise clairement l'omission. La régression séquentielle montre également que ces deux effets ne peuvent pas se réduire l'un à l'autre. En outre, il s'avère que les effets des deux facteurs cognitifs sont tous très significatifs ($p < 0,001$) : un verbe plus fréquent favorise l'omission du *que*, tandis qu'une complétive plus prédictible tend à apparaître sans *que*, ce second effet confirme que le lien entre la réduction syntaxique et la prédictibilité d'une structure existe aussi en français.

Notre étude fournit ainsi une mise à jour sur notre compréhension de l'omission de *que*, et montre comment les facteurs grammaticaux, sociaux et cognitifs interagissent pour affecter la variation dans une langue. Plus généralement, notre étude montre que les grands corpus sociolinguistiques des années 1970 et 1980 peuvent continuer à nourrir de nouvelles découvertes pour une compréhension plus fine de la variation, notamment sur ses aspects cognitifs, 50 ans après leur construction.

Références bibliographiques

Connors, K. (1975). L'effacement de *que* : règle syntaxique. *Recherches linguistiques à Montréal*, 4, 17-33.

Dion, N. (2003). L'effacement du *que* en français canadien : Une étude en temps réel. Mémoire de maîtrise, Université d'Ottawa.

Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive psychology*, 40(4), 296-340.

Jaeger, T. F. (2010). Redundancy and reduction : Speakers manage syntactic information density. *Cognitive psychology*, 61(1), 23-62.

Johnson, D. E. (2009). Getting off the goldvarb standard : Introducing rbrul for mixed-effects variable rule analysis. *Language and linguistics compass*, 3(1), 359-383.

Levy, R. P., & Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. In *Advances in neural information processing systems* (pp. 849-856). B. Schölkopf, J. Platt, T. Hoffman (eds.), Cambridge, MA : MIT Press.

Martineau, F. (1985). *Elision variable de (que) dans le parler d'Ottawa-Hull*. Thèse de doctorat non publiée, Université d'Ottawa.

Sankoff, G. (1980). *The social life of language*. University of Pennsylvania Press.

Sankoff, G., Sarrasin, R., & Cedergren, H. (1971). Quelques considérations sur la distribution de la variable *que* dans le français de Montréal. In *Congrès de l'association-française pour l'avancement des sciences*. Sherbrooke.

Thibault, P., & Vincent, D. (1990). Un corpus de français parlé : Montréal 84. *Université Laval, Québec*.

Warren, J. (1994). Plus ça change, plus c'est pareil : The case of "que" in Montreal French. *Culture*, 14(2), 39-49.

Construire un corpus éthique à l'ère numérique : le cas exceptionnel de l'« Instagrammable »

Erin McInerney (Université de Strasbourg et University of Glasgow)

L'étude et la construction de corpus multimodaux sont récemment devenues un centre d'intérêt croissant pour les chercheurs en sociolinguistique. Définir ce terrain de recherche et en comprendre l'intérêt demande au chercheur d'accepter de naviguer en eaux troubles. Comment recueillir de nouvelles données multimodales et respecter les lois rigoureuses de la protection des données des individus ? Comment comprendre le fonctionnement spécifique des plateformes de dissémination des nouveaux médias ? Chacun de ces éléments est en évolution permanente. Par conséquent, ils viennent brouiller davantage un domaine qui suscite énormément de questionnements.

Certains chercheurs ont tenté de démêler les nombreuses questions qui se posent sur les nouveaux médias. Un volume coordonné par Sofia Rüdiger et Daria Dayter (2020) met en avant la perspective corporelle dans l'étude des réseaux sociaux en ligne, tout en remarquant « l'incertitude entourant l'éthique et la légalité de l'usage des données provenant des médias sociaux » (p. 3). Sloan et Quan-Haase (2016) ont rassemblé un grand éventail de commentaires méthodologiques sur les recherches sociologiques qui s'appuient sur les données de masse. La dimension éthique y est également évoquée. Des appels à remettre en question cette notion essentielle d'éthique se trouvent dans les travaux individuels et collaboratifs de chercheurs (tels que Boyd, 2010 ; Auray, 2012 ; Bonneau, 2012 ; Proulx et Rueff, 2018 ; Bahary-Dionne, 2021), ainsi que dans les publications de groupes de recherche (voir l'Association of Internet Researchers, 2012 ; 2019), dans des guides informels pour les jeunes chercheurs (Townsend et Wallace, 2016), ou tout simplement, dans des éditoriaux en ligne (Casilli et Tubaro, 2017).

Il reste cependant des réseaux sociaux qui n'ont pas été beaucoup étudiés. À ce jour, les sociolinguistes ont préféré, en grande partie, exposer des données recueillies sur des réseaux sociaux en ligne tels que Twitter et Facebook. Par conséquent, les discussions sur la notion d'éthique tournent souvent autour de ces deux plateformes. La présente communication propose d'attirer l'attention sur une autre plateforme, Instagram. Figurant parmi les cinq sites web les plus visités dans le monde, avec plus de 1,6 milliard d'utilisateurs, Instagram représente un immense potentiel pour les travaux et corpus futurs des chercheurs, quelle que soit leur discipline. Ces chiffres font naître une question qui est au cœur de cette discussion : si un aussi grand nombre de personnes se servent de ce site pour échanger, pourquoi les chercheurs s'y sont si peu intéressés ? Cette communication vise à identifier certains des obstacles qui se présentent lors de la constitution d'un corpus à partir de données recueillies sur Instagram. Trois questions guideront progressivement la discussion, au cours de laquelle nous nous pencherons sur les possibles raisons pour lesquelles Instagram serait moins souvent choisie que d'autres plateformes dans nos recherches. Parmi celles-ci, nous citons l'aspect multimodal propre aux données recueillies sur (et créées pour) Instagram, que nous proposons de repenser davantage comme une ouverture que comme un obstacle. Pour terminer, une approche de collecte et d'analyse de données sera présentée, afin d'illustrer comment les corpus sociolinguistiques doivent évoluer en parallèle avec les enjeux éthiques associés à l'ensemble de données qui les forment.

Références bibliographiques

- Bayard, C. (2018). Les mères célèbres sur Instagram : ce que nous révèlent leurs mises en scène de l'allaitement. *Enfances, Familles, Générations*, 31, 187-206.
- Blackwood, R. (2019). Language, images, and Paris Orly airport on Instagram : Multilingual approaches to identity and self-representation on social media. *International Journal of Multilingualism*, 16(1), 7-24. <https://doi.org/10.1080/14790718.2018.1500257>
- Boyd, D. (2010). *Privacy and Publicity in the Context of Big Data*. Raleigh, North Carolina.
- Caldeira, S. P., Van Bauwel, S., & Ridder, S. D. (2020). 'Everybody needs to post a selfie every once in a while' : Exploring the politics of Instagram curation in young women's self-representational practices. *Information, Communication & Society*, 1-18. <https://doi.org/10.1080/1369118X.2020.1776371>
- Labov, W. (1972). *Sociolinguistic patterns* (11. print). Univ. of Pennsylvania Press.
- Piazzesi, C., & Lavoie Mongrain, C. (2020). Selfies de femmes, négociation normative et production de culture visuelle sur Instagram et Facebook. *Recherches féministes*, 33(1), 135. <https://doi.org/10.7202/1071246ar>
- Tiidenberg, K. (2015). Selfies| Odes to Heteronormativity : Presentations of Femininity in Russian-Speaking Pregnant Women's Instagram Accounts. *International Journal of Communication*, 9(13), 1746-1758.

Les mathématiques douloureuses : récit pour des vies professionnelles à venir

Claudine Moïse (Université Grenoble Alpes) et Martine Pons (Université Grenoble Alpes)

Les mathématiques occupent en France une place de distinction dans le cursus scolaire. Elles ouvrent non seulement la porte à des métiers à forte sélection mais renvoient une image d'excellence, d'intelligence, de logique et de créativité. Face à cette situation et à ces représentations, certain-es élèves expriment un fort rejet voire une détestation de cette matière, qui peut avoir des conséquences dans l'avenir, notamment pour les filles (Vouillot, 2014).

Pour saisir ce rejet des mathématiques et voir comment il se met en récit différemment chez les garçons et chez les filles (Moïse et Pons 2021), nous avons mené une enquête ethnographique entre sciences de l'éducation, sociolinguistique et analyse du discours. Nos données ont été collectées pendant trois ans, de 2016 à 2019, et sont constituées d'observations de classes, de 13 focus groupes d'élèves de seconde (trois ans avant la fin du cursus scolaire), d'un questionnaire auprès de 152 étudiants, en 2^e et 3^e années de licences (baccalauréat québécois) de Lettres, de Langues et de sciences du langage, d'analyses de deux films documentaires sur les l'enseignement des mathématiques, d'entretiens auprès d'enseignant-es et de 22 entretiens auprès de 11 filles et 11 garçons âgé.es de 18 à 25 ans.

Pour cette communication, nous nous intéresserons aux récits de vie des jeunes gens, ancien-nés élèves, en lien avec les mathématiques et, plus particulièrement ici, avec leurs choix professionnels. Face aux échecs en mathématiques, l'insouciance de l'adolescence et les choix professionnels s'accompagnent aussi de rêves d'avenir brisés, de frustrations au long cours et donc de changements d'itinéraires. Entre (dés)espoir et principe de réalité, cette période de vie acte des ambitions impossibles. Renoncements et réajustements, batailles et trouvailles, évitements et contournements construisent pas à pas un à-venir. Par l'activité narrative et la mise en récit comme transfiguration des événements (Ricœur 1990, Oraofiamma 2002, 2008), les histoires individuelles s'organisent, s'agencent et s'imbriquent pour faire sens. Elles permettent d'affirmer des décisions et des projections vers un à-venir où des vies s'orientent, rebondissent, et se construisent pour désavouer ou pas le sentiment d'échec et de relégation. Ces jeunes femmes et ces jeunes hommes nous ont confié leurs difficultés, leurs émotions et leurs espoirs. L'approche compréhensive dans les échanges, entre conversation, relances, développement, recadrage, reformulations et précisions, a permis d'analyser faits et émotions mis en récit. Ils et elles ont balayé leur parcours scolaire, se sont raconté-es en mettant en lumière les points saillants voire douloureux d'une expérience de vie en lien avec les mathématiques et donc notamment les paliers d'orientation subis ou choisis, les concordances et inadéquations avec leurs projets d'à-venir, les évidences et les ruptures intimes. Dans une perspective genrée, les mathématiques donnent alors à comprendre leur rôle dans les lignes droites, les bifurcations ou la souffrance ressenties.

Pour cette communication, il s'agira d'analyser la mise en discours de ces moments de choix, de renoncements et d'accommodation face à un avenir professionnel à construire. Comment les difficultés en mathématiques, voire le traumatisme qui en résulte, mettent à mal les choix professionnels des élèves ? En quoi différent-ils entre les filles et les garçons, quand il semble que la façon de dire l'échec n'augure pas des mêmes conséquences professionnelles pour les unes et pour les autres ? Comment racontent-elles et ils ces moments de vie qui engagent l'avenir, et quelles en sont les caractéristiques discursives de mise en narration ?

Références bibliographiques

Moïse Claudine et Martine Pons, 2021, *Ce que les mathématiques font aux filles*, Paris, Bréal.

Oraofiamma Roselyne, 2002, « Le travail de la narration dans le récit de vie », Niewiadomski Christophe et de Villers Guy (dir.), *Souci et soin de soi, liens et frontières entre histoire de vie, psychothérapie et psychanalyse*, Paris, L'Harmattan, p. 163-191.

Oraofiamma Roselyne, 2008, « Les figures du sujet dans le récit de vie », numéro 145, *L'individu et ses appartenances*, partie 2, *Informations sociales*, p. 68-81.

Ricœur Paul, 1990, *Soi-même comme un autre*, Paris, Seuil, Points Essais.

Vouillot Françoise, 2014, *Les métiers ont-ils un sexe ? Pour sortir des sentiers battus de l'orientation des filles et des garçons*, Paris, Belin.

Un nouveau corpus sur l'espagnol parlé à Montréal : le COLEM

Enrique Pato (Université de Montréal)

Du point de vue social, politique et économique, Montréal est devenue la ville « Latine » par excellence de tout le Canada. En effet, elle concentre un pourcentage (14,2 %) de distribution du groupe minoritaire latino-américain équivalent à celui d'autres grandes villes canadiennes, telles que Vancouver, Toronto et Ottawa-Gatineau (15 %). Le groupe latino-américain qui caractérise la Région métropolitaine de Montréal (RMM) compte plus de 75 000 personnes (Bastien et Bélanger 2010, Pato 2017, Pato 2020).

Le COLEM (esp. *Corpus oral de la lengua española en Montreal*, fr. *Corpus oral de la langue espagnole à Montréal*), c'est un corpus de langue orale conçu pour connaître la situation linguistique de la communauté hispanophone et pour approfondir notre compréhension de l'espagnol à Montréal (RMM). Il permet aussi de mieux connaître l'adoption du français par les néo-Montréalais. Ce double champ de recherche est tout nouveau au Québec.

Le but précis du COLEM, inspiré du corpus Sankoff-Cedergren (Sankoff et al. 1976 : 88-89), est d'obtenir des données sur la nature, l'étendue et la fonction de la diversité linguistique à l'intérieur de l'espagnol parlé par les immigrants à Montréal, afin d'éclaircir la situation d'une population souffrant d'une considérable aliénation linguistique. Le protocole a été créé pour « provoquer chez les répondants des commentaires spontanés et prolongés » (Sankoff et al. 1976 : 113). Il s'agit donc des histoires de vie et des récits autobiographiques qui mettent de l'avant l'individu, sa trajectoire, ses opinions et ses goûts.

Le COLEM est un corpus de la langue parlée, résultant d'une série de conversations (153 entrevues semi-dirigées jusqu'à présent) enregistrées dans un milieu naturel, d'une durée approximative d'une heure et structurées en fonction d'un protocole commun. Le style de collecte est la conversation familiale (discours spontané). La sélection des participants est réalisée à partir de critères sociolinguistiques, des situations de contact linguistique et de la migration.

Cette présentation offrira une description du COLEM et un échantillon de l'espagnol parlé à Montréal en considérant ses aspects propres (grammatical et lexical) en tant qu'éléments d'un système cohérent (« espagnol montréalais » ?) partagé par tous les membres de la communauté, mais aussi ses différences. En effet, comme c'est le cas dans d'autres contextes géographiques, par exemple la ville de New York (Otheguy et Zentella 2012), la langue des hispanophones à Montréal peut être caractérisée par : 1) la continuité structurelle de l'espagnol d'origine ; 2) le contact linguistique avec le français et l'anglais ; et 3) le nivellement dialectal, résultant du contact avec d'autres variétés de l'espagnol.

Les données du COLEM nous permettent de mieux cerner plusieurs phénomènes du contact linguistique et servent à l'étude de l'espagnol, non seulement au Québec mais également au Canada. Le corpus peut intéresser aussi d'autres disciplines, comme les études portant sur la migration et les enjeux sociaux et culturels.

Références bibliographiques

Bastien, Y./Bélanger, A. 2010. *Un portrait comparatif de la situation de l'emploi chez les immigrants et les minorités visibles dans les RMR de Montréal, Ottawa-Gatineau, Toronto et Vancouver*. Montréal : INRS-UCS.

Otheguy, R./Zentella, A. C. 2012. *Spanish in New York : Language contact, dialect leveling, and structural continuity*. Nueva York : Oxford University Press.

Pato, E. 2017. « La realidad lingüística en Canadá y la situación del español en Montreal », *Oltreoceano*, 13 : 29-39.

Pato, E. 2020. « El español en contacto con el francés en Quebec y su estudio gracias al *Corpus oral de la lengua española en Montreal* (COLEM) », *Boletín Hispánico Helvético*, 35-36 : 263-287.

Sankoff, D./Sankoff, G./Laberge, S./Topham, M. 1976. « Méthodes d'échantillonnage et utilisation de l'ordinateur dans l'étude de la variation grammaticale », *Cahiers de linguistique*, 6 : 85-125.

1971-2016 : 45 ans de variation sociolinguistique observable dans un nouveau corpus de français montréalais

Béatrice Rea (University of Oxford)

Ma communication prend sa source dans une étude que j'ai menée en 2015-2020 sur le français montréalais et qui visait plus précisément à examiner l'alternance entre les auxiliaires *avoir* et *être* (ex. *J'ai tombé*) dans la vingtaine de verbes intransitifs et des pronominaux (ex. *Je m'ai fait mal*) requérant conventionnellement *être* (Grevisse 2011 : §782). Étant donné qu'en 45 ans le paysage linguistique montréalais a connu de grands changements et que le niveau de scolarisation moyen a augmenté, ce projet se voulait une mise à jour de l'article phare de Sankoff & Thibault (1977)⁴ sur l'alternance entre *avoir* et *être* dans le corpus Sankoff-Sankoff-Cedergren (1971).

J'ai donc tenté de déterminer par le biais d'entrevues sociolinguistiques (que j'ai réalisées en 2016) avec 52 locuteurs natifs (± 55 heures au total) l'état de cette variable morphosyntaxique aujourd'hui. Mes résultats démontrent que le taux d'*avoir* est passé d'environ 34 % à 10 %, entre 1971 et 2016.

À des fins de comparaison, mes entrevues consistaient en une série de questions similaires à celles du corpus Sankoff-Sankoff-Cedergren (1971). L'échantillon des locuteurs s'inspire aussi de ce corpus : les 52 participants, dont l'année de naissance se situe entre 1954 et 1998, sont répartis de façon égale entre les différentes macro-catégories sociales (âge, sexe, profil socio-professionnel), à raison d'environ quatre locuteurs par cellule.

Bien que l'étiquette « francophone(s) et Montréalais d'origine » (Sankoff et al. 1976) ait été plutôt évidente en 1971, divers changements démographiques (fusions municipales, immigration, exode rural, gentrification, diversification linguistique des quartiers traditionnellement francophones, etc.) ont compliqué la tâche de trouver des locuteurs « de souche ». Cette difficulté avait aussi été rencontrée par Labov et al. (2005 : 27) pour les villes américaines d'Atlanta, Dallas et New York lors de la création de leur *Atlas of North American English*. C'est pourquoi mon corpus inclut aussi des locuteurs du Grand Montréal.⁵

Ma communication a donc pour but de présenter les détails de ce corpus, puisque ce nouvel outil sera accessible en ligne et interrogeable sur la plateforme FDLQ. Plusieurs métadonnées seront disponibles (niveau d'éducation, lieu de naissance, lieu d'habitation actuel, occupation, occupation des parents et/ou du partenaire de vie, niveau de maîtrise/fréquence d'utilisation de l'anglais, contact avec d'autres langues et d'autres variétés de français, lien avec l'intervieweuse).

⁴ Sankoff a publié récemment une version de cet article (voir Sankoff 2019) incluant les données des corpus de 1971, 1984 (Thibault & Vincent) et 1995 (Vincent, Laforest & Martel) et s'attarde au changement linguistique au cours de la vie.

⁵ À l'époque, les corpus de quartiers de Hochelaga-Maisonneuve de Blondeau et al. (2012) et de Saint-Michel-Montréal-Nord de Blondeau et al. (2013), ainsi que celui d'Ahuntsic-Cartierville de Bigot & Papen (2018), n'étaient pas encore accessibles.

Références bibliographiques

Bigot, D. & R. Papen (2018), « Deux nouveaux corpus de français parlé au Québec : Sainte-Anne-des-Lacs et Montréal (Ahuntsic-Cartierville) », *Les français d'ici 2018*, Université Concordia, Montréal, mai 23-25.

Sous-corpus variationniste Hélène Blondeau-France Martineau-Mireille Tremblay-Yves Frenette de Hochelaga-Maisonneuve 2012, du Corpus FRAN (dir. F. Martineau).

Sous-corpus variationniste Hélène Blondeau-France Martineau-Mireille Tremblay de Saint-Michel-Montréal-Nord 2013, du Corpus FRAN (dir. F. Martineau).

Grevisse, M. (2011), *Le bon usage. Grammaire française refondue par André Goosse*, 15^e éd., Gembloux : Duculot.

Labov, W., S. Ash & C. Boberg (2005), *The Atlas of North American English : Phonetics, Phonology and Sound Change Volume 1*, Berlin : De Gruyter.

Sankoff, D., G. Sankoff, S. Laberge & M. Topham (1976), « Méthodes d'échantillonnage et utilisation de l'ordinateur dans l'étude de la variation grammaticale », *Cahiers de linguistique*, n° 6, p. 85-125.

Sankoff, G. & P. Thibault (1977), « L'alternance entre les auxiliaires *avoir* et *être* en français parlé à Montréal », *Langue française* n° 34, p. 81-108.

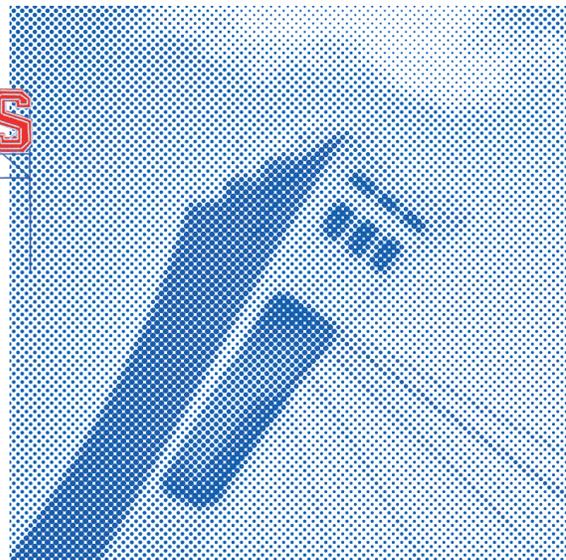
Sankoff, G. & P. Thibault (1980), « The alternation between the auxiliaries *avoir* and *être* in Montréal French », dans G. Sankoff (éd.) *The Social Life of Language*, Philadelphie : University of Pennsylvania Press, p. 311-345.

Sankoff, G. (2019), « Language change across the lifespan : Three trajectory types », *Language* 95(2), p. 197-229. <https://doi.org/10.1353/lan.2019.0029>

Thibault, P. & D. Vincent (1990), *Un corpus de français parlé : Montréal 84. Historique, méthodes et perspectives de recherche*, Québec, Département de langues et linguistique.

Vincent, D., M. Laforest & G. Martel (1995), « Le corpus de Montréal 1995 : adaptation de la méthode d'enquête sociolinguistique pour l'analyse conversationnelle », *Dialangue*, n° 6, p. 29-45.

FARINE
FIVE ROSES



Les corpus oraux montréalais à l'ère numérique : des bobines magnétiques à la plateforme du Fonds de données linguistiques du Québec (FDLQ)

Wim Remysen (Université de Sherbrooke) et Hugo Saint-Amant Lamy (Université de Sherbrooke)

Cinquante ans après leur genèse, force est de constater que les corpus sociolinguistiques montréalais de 1971, 1984 et 1995 se sont révélés de puissants catalyseurs pour l'analyse et la description du français parlé au Québec, notamment dans la perspective variationniste. Il n'est pas exagéré d'affirmer que les auteures et auteur qui ont pris l'initiative de constituer le premier de ces corpus – Gillian Sankoff, David Sankoff et Henrietta Cedergren – font figure de pionniers. Non seulement leur entreprise a-t-elle inspiré quantité d'autres linguistes par la suite, les nombreux travaux réalisés à partir des données recueillies ont aussi grandement contribué à une meilleure connaissance de plusieurs phénomènes caractéristiques du français québécois⁶. Dans un contexte encore fortement marqué par des préjugés négatifs envers cette variété de français, de tels travaux avaient une pertinence certaine. L'élaboration de corpus subséquents en 1984 et 1995, où certains locuteurs de 1971 sont interviewés à nouveau, a permis d'ajouter une dimension diachronique précieuse au riche portrait sociolinguistique du projet original. Cette richesse a suscité l'envie et l'intérêt de sociolinguistes partout dans le monde.

Aujourd'hui, Montréal 1971 et ses suites n'ont aucunement perdu de leur intérêt, bien au contraire. Bien plus qu'une simple pièce d'anthologie, ils ouvrent la possibilité d'étudier le changement qui affecte le français parlé à Montréal, notamment par des comparaisons avec des corpus constitués à date plus récente, et permettent de prendre le pouls des représentations de la langue qui existaient à l'époque. La granularité sociale et la profondeur longitudinale inégalées des trois corpus en font en outre un étalon de choix à l'aune duquel interpréter globalement la variation en français québécois. Cela dit, la préservation des données et leur accessibilité sont des enjeux majeurs. Ainsi, la numérisation de ces corpus, que souhaitaient déjà les auteures des corpus⁷, est relativement récente et leur mise en commun et leur dépôt n'ont jamais été réalisés.

Dans le cadre de cette communication, nous présenterons un projet de recherche récemment entrepris par le Centre de recherche interuniversitaire sur le français en usage au Québec en vue de pérenniser et de mutualiser les corpus montréalais, tant patrimoniaux (parmi lesquels la trilogie de 1971, 1984 et 1995) que contemporains (comme le sous-corpus montréalais du corpus PFC-Québec ou d'autres corpus récemment constitués). Ce projet consiste à créer une plateforme, baptisée Fonds de données linguistiques du Québec, qui rend possible la consultation de différents corpus et l'interrogation des contenus. D'outils utilisés ponctuellement pour des recherches spécifiques, les corpus intégrés au Fonds sont appelés à devenir des ressources d'usage courant, facilement accessibles et partagées par la communauté des chercheurs.

Nous présenterons les principales orientations théoriques, méthodologiques et techniques qui sous-tendent la réalisation de cette plateforme à partir de l'exemple des corpus montréalais qu'elle réunira. Les défis et enjeux sont en effet nombreux et concernent toutes les étapes de préparation des données, allant de la numérisation des entrevues à leur affichage, en passant par les questions de transcription et d'anonymisation des données.

Références bibliographiques

Blondeau, Hélène et Marty Laforest (2015), « Bibliographie des travaux fondés sur les données des grands corpus sociolinguistiques de Montréal », document manuscrit, 24 p.

Daveluy, Michelle et Marty Laforest (1994), « Entretiens : *Tout ça a commencé en 1970...* », *Culture*, vol. 14, n° 2, p. 11-26.

⁶ En fait foi l'impressionnante bibliographie faisant l'inventaire des recherches menées à partir des données de ce corpus (ainsi que des deux autres qui ont suivi). Cette bibliographie a été composée par Hélène Blondeau et Marty Laforest (2015).

⁷ Voir à ce sujet les propos recueillis par Michelle Daveluy et Marty Laforest auprès des instigatrices des trois corpus (et publiés dans Daveluy et Laforest 1994).

À l'écoute du passé : la pertinence des corpus patrimoniaux dans l'étude du français acadien en temps réel

Basile Roussel (Université de Moncton, campus de Shippagan), Wladyslaw Cichocki (University of New Brunswick) et Louise Beaulieu (Université de Moncton, campus de Shippagan)

S'il est un élément, parmi d'autres, qui témoigne de l'impact qu'a eu la création du corpus Sankoff-Cedergren (Sankoff et al., 1976), c'est bien celui d'avoir inspiré de nombreux travaux sur l'étude de la langue parlée vernaculaire d'un point de vue diachronique (Martineau, 2005 ; Poplack et St-Amand, 2009 ; Sankoff, 2006 ; Thibault et Daveluy, 1989 ; parmi d'autres). L'emploi de données écrites et orales qui ont survécu à l'épreuve du temps implique cependant une réflexion sur leur représentativité. Ces données, souvent éparses, ont parfois des styles différents et fournissent généralement peu d'informations sociales relativement aux locuteurs, ce qui soulève certaines questions. Comment déterminer si ces données présentent une image précise de la complexité du vernaculaire parlé dans le passé ? Est-il pertinent de comparer ces données à celles issues de corpus contemporains ?

Dans l'étude de ces questions, nous mettons à profit deux corpus patrimoniaux de français acadien parlé dans le nord-est du Nouveau-Brunswick : le *Corpus Collection Dominique Gauthier, M.D.* (recueilli dans les années 1950) et le *Corpus Collection Société Historique Nicolas-Denys* (recueilli en 1975). Ces corpus incluent un bassin de locuteurs nés entre 1870 et 1909. Pour les besoins de la présente étude, ces deux corpus sont comparés au *Corpus FANENB* recueilli dans la même région en 1990-91 (Beaulieu, 1995). Cet autre corpus a été créé en suivant le même standard méthodologique que celui de Montréal.

Afin d'étudier la représentativité des données patrimoniales en tant que manifestation de la variété vernaculaire, nous examinons trois éléments qui peuvent influencer le degré de formalité dans la performance linguistique et l'emploi du vernaculaire. Le premier élément porte sur les conditions d'enregistrement des corpus : les considérations physiques et techniques, le profil social de l'interviewer (contexte situationnel) et les consignes données aux locuteurs. Le deuxième s'intéresse au genre de discours employé : conte et histoire de vie. Le dernier porte sur la nature et l'étendue du registre des locuteurs à partir de leurs caractéristiques sociales et de la structure de leur communauté. De plus, nous regardons la présence dans les données de traits linguistiques qui sont attestés dans la forme vernaculaire des variétés acadiennes.

Dans le but de comprendre si une comparaison entre les corpus patrimoniaux et les corpus contemporains est pertinente, nous examinons le *Corpus FANENB* à l'aide des mêmes éléments. Cette comparaison permettra de déterminer le degré de compatibilité entre les données ainsi que la nature des analyses linguistiques possibles entre les données patrimoniales et les divers regroupements de locuteurs d'un corpus contemporain. Finalement, nous présentons brièvement les résultats d'analyses diachroniques se rapportant à deux variables morphosyntaxiques faites à partir des trois corpus discutés : les séquences en tête des propositions tensées en *si* (Ex. : *si/si que tu veux*), et les désinences postverbales à la troisième personne du pluriel (Ex. : *ils pêchent/i pêchont*).

En somme, cette présentation démontre la pertinence, dans l'étude du changement linguistique, des corpus représentant un état de langue parlée antérieurement.

Références bibliographiques

Beaulieu, L. (1995). *The social function of linguistic variation : A sociolinguistic study in a fishing community of the north-eastern coast of New Brunswick*. Thèse de doctorat, University of South Carolina.

Martineau, F. (2005). Perspectives sur le changement linguistique : aux sources du français canadien. *Revue canadienne de linguistique/Canadian Journal of Linguistics* 50 : 173-213.

Poplack, S. et A. St-Amand (2009). Les "Récits du français québécois d'autrefois" : reflets du parler vernaculaire du 19^e siècle. *Revue canadienne de linguistique/Canadian Journal of Linguistics* 54(3) : 511-546.

Sankoff, G. (2006). Age : Apparent time and real time. *Elsevier Encyclopedia of Language and Linguistics* (2^{ième} édition). Oxford : Elsevier.

Sankoff, D., G. Sankoff, S. Laberge et M. Topham (1976). Méthodes d'échantillonnage et utilisation de l'ordinateur dans l'étude de la variation grammaticale. *Cahiers de linguistique* 6 : 85-125.

Thibault, P. et M. Daveluy (1989). Quelques traces du passage du temps dans le parler des Montréalais, 1971-1984. *Language Variation and Change* 1(1) : 19-45.

Optimiser la transcription des corpus linguistiques à l'aide de la technologie intelligente : l'exemple du COLEM

Marimar Rufino Morales (Université de Montréal)

Au fil des dernières décennies, les progrès technologiques ont propulsé la méthodologie de préparation des corpus linguistiques et ont allégé les étapes de collecte et de stockage de matériel, d'alignement, d'annotation. En revanche, lorsqu'il s'agit de transcrire le matériel, la charge interprétative et la nature même des conversations font obstacle à l'automatisation du processus. L'intelligence artificielle et l'apprentissage profond ont assuré la parité entre la reconnaissance automatique de la parole et l'être humain (Xiong *et al.* 2016). Or, lorsque cette technologie est confrontée à la conversation spontanée ou à la variation, on obtient des résultats encore trop instables (Ravanelli *et al.* 2018). De ce fait, les interviews destinées à l'étude de la langue orale sont toujours transcrites à l'aide d'un lecteur et d'un clavier (Durand 2017).

Et pourtant, d'autres environnements professionnels ont recours à des logiciels de dictée pour un passage de l'oral à l'écrit. On y parvient, en mettant l'accent sur une collaboration humain-ordinateur là où la technologie de reconnaissance automatique de la parole fait défaut, à augmenter la précision des transcriptions. Les techniques et les stratégies employées diffèrent, tout comme les noms qui leur sont attribués, mais toutes celles-ci visent à stabiliser les fluctuations des outils informatiques et à gagner du temps.

Nous nous sommes focalisée sur la redite offline (Rufino Morales 2020), inspirée d'une méthode de sous-titrage à la télévision en temps réel, en raison du parallélisme entre l'oralité présente dans la plupart des scénarios des médias audiovisuels et celle qui caractérise les données collectées pour créer un corpus de la langue parlée. Nous avons adopté la redite offline pour transcrire les entrevues du *Corpus oral de langue espagnole à Montréal (COLEM)* (Pato dir.) et avons réussi, tout comme à la télévision, à optimiser le processus. Les transcriptions obtenues sont plus riches, tiennent compte du protocole du COLEM, et ce, en moins de temps que les autres méthodes.

Nous ferons une démonstration de redite offline et présenterons les résultats des données recueillies lors de la transcription du COLEM : un corpus composé d'interviews semi-dirigées, enregistrées dans un environnement naturel, qui reflètent l'espagnol en contact avec le français et l'anglais dans la Région Métropolitaine de Montréal (Pato 2017).

Nous étayerons l'exemple du COLEM en découvrant de nouveaux groupes de recherche qui ont adopté la redite offline pour transcrire des corpus oraux.

Références bibliographiques

Durand, Jacques. 2017. Corpus Phonology, *Oxford Research Encyclopedia of Linguistics*. doi : 10.1093/acrefore/9780199384655.013.145.

Pato, Enrique. 2017. La realidad lingüística en Canadá y la situación del español en Montreal. *Oltreoceano* 13, 27-37. doi : 10.1400/253045.

Ravanelli, Mirco, Brakel, Philemon, Omologo, Maurizio y Bengio, Yoshua. 2018. Light gated recurrent units for speech recognition. *IEEE Transactions on Emerging Topics in Computational Intelligence* 2(2), 92-102. doi : 10.1109/TETCI.2017.2762739.

Rufino Morales, Marimar. 2020. El reablado *off-line* para optimizar la transcripción de corpus orales en español. En Silvia Martínez Martínez (Ed.), "Nuevas tendencias en Traducción e Interpretación" (pp. 17-27). Granada : Comares.

Xiong, Wayne, Droppo, Jasha, Huang, Xuedong, Seide, Frank, Seltzer, Mike, Stolcke, Andreas, Yu, Dong y Zweig, Geoffrey. 2016. Achieving human parity in conversational speech recognition. *arXiv:1610.05256*.

La vie montréalaise racontée au sein de la Petite-Italie et du Petit-Portugal

Fabio Scetti (Université Paul-Valéry Montpellier 3)

Les communautés italienne et portugaise de Montréal sont des véritables pièces du puzzle urbain de la ville. La « Petite-Italie » et le « Petit-Portugal » ont été fondés respectivement dans les années 1910 et dans les années 1950, et depuis, ils se sont étendus sur le territoire urbain au-delà de leur dite « enclave ethnique ». La *Piccola Italia*, qui chevauche le marché Jean-Talon au Nord du Mile End et la *Comunidade* portugaise, qui se trouve sur le Plateau-Mont-Royal en descendant le boulevard Saint-Laurent vers le centre-ville, restent des lieux de référence du passage de ces deux flux migratoires à Montréal.

La fondation des deux communautés a marqué l'histoire de la ville et de la province du Québec. Elles ont joué un rôle important pendant la Révolution tranquille en ce qui concerne l'intégration linguistique, notamment suite à la Loi 101 et à la scolarisation des enfants. Nos recherches doctorales et postdoctorales, de 2011 à 2021, et le recueil d'un corpus sociolinguistique – fait d'observations ethnographiques de la vie de quartier et d'entretiens – nous permettent de raconter l'histoire de Montréal et de vie des Montréalais durant plus d'un siècle. Grâce aux récits, nous pouvons ainsi redessiner l'histoire politique, sociale et linguistique. Les histoires individuelles peuvent se regrouper dans une seule et unique histoire collective, faite de discours circulant au sein de ces communautés, de pratiques langagières diverses et de représentations.

Ces entrevues nous montrent le rôle des langues et des dialectes dans les familles et communautés ; ces dernières devenues de véritables lieux d'ancrage de l'identité des groupes. Après une description historique et socio-politique des contextes étudiés, cette contribution fournit des exemples des langues parlées et de vie. Ces récits se révèlent importants lors de la description de Montréal aujourd'hui, en se tournant en éléments complémentaires dans la construction de l'identité montréalaise. Ces communautés et leurs quartiers respectifs deviennent pour cette raison des véritables îlots de mémoire, communautaire mais aussi urbaine.

Références bibliographiques

Gabaccia, Donna R. (2006). « Global Geography of "Little Italy" : Italian Neighborhoods in Comparative Perspective », *Modern Italy*, n° 1, p. 9-24.

Ramirez, Bruno, and Michael Del Balso (1980). *The Italians of Montreal : from sojourning to settlement, 1900-1921*. Montréal : Éditions du Courant.

Scetti, Fabio (2019). *La communauté portugaise de Montréal : langue et identité*. Québec : Presses de l'Université Laval.

Simão Andrade, Miguel (2007). « La Commission des écoles catholiques de Montréal et l'intégration des immigrants et des minorités ethniques à l'école française de 1947 à 1977 », *Revue d'histoire de l'Amérique française*, vol. 60, n° 4, p. 455-486.

Zucchi, John (2007). *A history of ethnic enclaves in Canada* (vol. 9), Ottawa : Canadian Historical Association.

Nouvelles voies, nouvelles voix en sociolinguistique : ce que nous disent les corpus

Table ronde organisée par Hélène Blondeau (Université de Floride) et Anne-José Villeneuve (Université de l'Alberta)

Intervenantes et intervenants : Guillaume Beauchamp (Université de Montréal), Alexandra Dupuy (Université de Montréal), Isabelle LeBlanc (Université de Moncton) et Anne-José Villeneuve (Université de l'Alberta)

Que l'on se rattache à la sociolinguistique ou à l'anthropologie linguistique, nos pratiques de recherches et la constitution de corpus s'ancrent généralement sur des terrains d'études au sein de groupes de sujets ou d'acteurs sociaux. Comme le faisait remarquer Eckert (2013), quelle que soit la manière dont ces groupes sont définis, et quelle que soit leur taille, ils sont toujours divers et les différences observées en leur sein se présentent sous de nombreuses formes. En effet, que l'on définisse ces groupes comme des communautés linguistiques, des communautés de paroles, des réseaux sociaux ou des communautés de pratique, ils sont tous marqués par la diversité.

Profitant de cet événement qui célèbre les 5(1) ans de la constitution du corpus montréalais Sankoff-Cedergren, cette table ronde propose une discussion rétrospective et prospective sur la façon d'aborder et de mieux représenter la diversité en sociolinguistique. Cette question peut se poser d'une part en réfléchissant à nos façons de faire dans la constitution des corpus sociolinguistiques et à la manière de les interroger et de les interpréter. D'autre part, la diversité en sociolinguistique peut s'aborder en réfléchissant à la pluralité présente dans les cercles de sociolinguistes qui constituent nos communautés de pratique universitaires.

Ce travail réflexif nous amène donc à proposer une table ronde qui nous l'espérons rejoindra tant les chercheuses et les chercheurs de la génération émergente que celles et ceux qui circulent dans les cercles sociolinguistiques depuis plus longtemps.

