Document generated on 07/01/2025 2:16 a.m.

Cahiers franco-canadiens de l'Ouest

Le lexique nordique des jeunes Manitobains : constantes et variations sociolinguistiques dans le vocabulaire fondamental désignant les animaux

Liliane Rodriguez

Volume 29, Number 2, 2017

Territoire, langue et identité : présences nordiques dans l'Ouest canadien

URI: https://id.erudit.org/iderudit/1042268ar
DOI: https://doi.org/10.7202/1042268ar

See table of contents

Publisher(s)

Presses universitaires de Saint-Boniface (PUSB)

ISSN

0843-9559 (print)
1916-7792 (digital)

Explore this journal

Cite this article

Rodriguez, L. (2017). Le lexique nordique des jeunes Manitobains : constantes et variations sociolinguistiques dans le vocabulaire fondamental désignant les animaux. *Cahiers franco-canadiens de l'Ouest*, *29*(2), 429–455. https://doi.org/10.7202/1042268ar

Article abstract

What is called the *basic vocabulary* (or *fundamental*, or *available lexicon*) of young speakers constitutes a strong semantic base for their future lexical development, as teenagers and adults, even as future writers. Such vocabulary is indeed *fundamental* because it is acquired early, before thirteen years of age, to designate daily life objects representative of the world experience of young speakers: clothing, games, school, animals, etc. Our field work in twelve Manitoba communities, between 1990 and 2006, reveals that the semantic notion of *the North* is recurrent in the available lexicon of Franco-Manitobans. Based on examples from the lexicometric corpora established with the collected data using lexico-statistical methodologies, our examples in this article focus on the lexico-semantic field of Animals (the Fauna field). The *lexicometric indices* of these available words (which account for their actual use, and not passive understanding) were calculated for two sociolinguistic parameters: urban/rural and girl/boy. These indices converge, showing the lexicometric salience linked to Northern daily experience (e.g. bear, wolferine), as opposed to the low statistical representation of foreign referents (e.g. giraffe, elephant), in the development of the lexicon, the semantic memory, and the creative imagination of young Manitobans.

Tous droits réservés © Centre d'études franco-canadiennes de l'Ouest (CEFCO) et Presses universitaires de Saint-Boniface (PUSB), 2017

This document is protected by copyright law. Use of the services of Érudit (including reproduction) is subject to its terms and conditions, which can be viewed online.

https://apropos.erudit.org/en/users/policy-on-use/

érudit

This article is disseminated and preserved by Érudit.

Érudit is a non-profit inter-university consortium of the Université de Montréal, Université Laval, and the Université du Québec à Montréal. Its mission is to promote and disseminate research.

https://www.erudit.org/en/

CAHIERS FRANCO-CANADIENS DE L'OUEST
Vol. 29, nº 2, 2017, p. 429-455

Le lexique nordique des jeunes Manitobains: constantes et variations sociolinguistiques dans le vocabulaire fondamental désignant les animaux

Liliane RODRIGUEZ
University of Winnipeg

RÉSUMÉ

Ce que l'on appelle le *vocabulaire fondamental* (ou *lexique disponible*) des jeunes forme le socle sémantique sur lequel se construit, au fil des ans, le lexique du locuteur adolescent puis adulte, voire l'imaginaire de l'écrivain. Ce vocabulaire est *fondamental* car il s'acquiert jusqu'à l'âge de treize ans, pour désigner les objets de notre expérience immédiate du monde: les vêtements, les jeux, l'école, les animaux, etc. Nos enquêtes, réalisées en douze communautés du Manitoba entre 1990 et 2006, montrent que la notion sémantique du *Nord* occupe une place récurrente dans le vocabulaire fondamental des jeunes Franco-Manitobains. Extraits des corpus lexicométriques établis par des méthodes de statistique lexicale sur les données de ces enquêtes, nos exemples, ici, proviennent du champ lexico-sémantique «Les animaux» (le champ de la faune). Les *indices de disponibilité* de ces mots (leur usage réel, et non une simple compréhension passive) ont été calculés selon deux paramètres sociolinguistiques: urbain/rural et fille/garçon. Ces indices convergent, mettant en évidence la forte représentation lexicométrique d'une expérience nordique vécue (ours, carcajou, etc.) par rapport à la faible représentation de référents lointains (girafe, éléphant, etc.) dans le développement du lexique, de la mémoire sémantique et de l'imaginaire des jeunes Manitobains.

ABSTRACT

What is called the *basic vocabulary* (or *fundamental*, or *available lexicon*) of young speakers constitutes a strong semantic base for their future lexical development, as teenagers and adults, even as future writers. Such vocabulary is indeed *fundamental* because it is acquired early, before thirteen years of age, to designate daily life objects representative of the world experience of young speakers: clothing, games, school, animals, etc. Our field work in twelve Manitoba communities, between 1990 and 2006, reveals that the semantic notion of *the North* is recurrent in the available lexicon of Franco-Manitobans. Based on examples from the lexicometric corpora established with the collected data using lexico-statistical methodologies, our examples in this article focus on the lexico-semantic field of Animals (the Fauna field). The *lexicometric indices* of these available words (which account for their actual use, and not passive understanding) were calculated for two sociolinguistic parameters: urban/rural and girl/boy. These indices converge, showing the lexicometric salience linked to Northern daily experience (e.g. bear, wolferine), as opposed to the low statistical representation of foreign referents (e.g. giraffe, elephant), in the development of the lexicon, the semantic memory, and the creative imagination of young Manitobans.

1. Le «Nord» des jeunes Manitobains: du territoire au vocabulaire

L'entité complexe qu'est le Nord, avec ses multiples facettes géographique, sociale, politique, économique et environnementale est aussi, tout simplement, un cadre de vie quotidienne, par exemple celui des Manitobains et d'autres Canadiens. Nous suivrons le fil linguistique qui lie locuteurs et lieu de vie, pour mesurer le rôle du lieu de vie sur le lexique en usage chez les jeunes. Toute population entretient avec son territoire des relations étroites qui se répercutent sur sa langue, notamment sur le lexique acquis pendant l'enfance et l'adolescence. Nous avons interrogé la langue des jeunes Manitobains francophones (souvent bilingues), à la recherche de présences nordiques dans leur vocabulaire. Le Manitoba, en effet, se situe dans l'Ouest du Canada, mais il fit longtemps

partie de ce vaste espace que les habitants de la Nouvelle-France nommaient «le Nord-Ouest». Cette désignation courante au XIXe siècle survit aujourd'hui dans des noms tels que «Police montée du Nord-Ouest» (Brown, 1990, p. 419) et «Rébellion du Nord-Ouest» (soulèvement de 1885 contre le gouvernement canadien des Métis et de leurs alliés autochtones en Saskatchewan, *Historica Canada*, 2017). De fait, le Manitoba s'apparente indéniablement au Nord, par ses longs hivers aux températures négatives extrêmes, et la faune et la flore septentrionales qui constituent son «ADN climatique».

S'appuyant sur nos recherches sur le parler des jeunes (notamment Rodriguez, 2006), cette étude consacrée uniquement au vocabulaire de la faune montre que les références lexicales au Nord et à ses attributs familiers diffèrent des références au Sud et à ses attributs exotiques. Il est évident que notre lexique personnel est plus à même de s'enrichir de mots en lien direct avec notre vie journalière climatique, géographique et sociale (qui, pour nos jeunes témoins, inclut l'expérience scolaire, dont on ne saurait négliger l'importance en milieu linguistique minoritaire) qu'avec des réalités éloignées de notre quotidien. Mais, par quelles voies cette influence passe-t-elle dans l'acquisition lexicale? Cette étude ciblée sur la faune montrera comment ce que nous appellerons le *vocabulaire de proximité* prend racine dans la langue des jeunes, permettant à leur lexique de s'étendre et de se nuancer, en développant leurs moyens de communication et d'expression. Au-delà de ces aspects d'ordre linguistique, l'empreinte du vocabulaire de proximité chez les enfants va de pair avec le développement de leurs facultés métaphoriques et de leur sentiment identitaire. Le premier contact avec le milieu naturel géographique et climatique au seuil de l'acquisition lexicale devient non seulement vecteur de compétence lexicale mais aussi de créativité et d'identité. Cette empreinte du premier milieu naturel sur la parole et l'imaginaire a été analysée de façon magistrale par Bachelard (Bachelard, 1942) dans sa phénoménologie de l'imagination, à la source de la critique psychanalytique, herméneutique et postmoderne, depuis Ricœur (Ricœur, 1975) jusqu'à Sloterdijk (Sloterdijk, 2014). C'est dire l'importance du milieu climatique et géographique dans le développement du lexique des jeunes. C'est la description de la compétence lexicale au contact des milieux climatique et géographique qui restera au centre de notre

propos, toutefois son corrélat métaphorique (voire artistique) et identitaire sera évoqué à nouveau en conclusion. Mais d'abord, comment s'établit le lien entre le lieu de vie d'un locuteur et son vocabulaire?

2. Du *référent* à la *disponibilité*: méthodologie et corpus

Pour répondre à cette question, quelques définitions préliminaires sont nécessaires, notamment celle du *signe verbal*. Il se définit intrinsèquement comme l'ensemble du *signifiant*, la face phonique ou graphique du signe, et du *signifié*, sa face conceptuelle (Saussure, 1972 [1916], p. 99-100). Il faut ajouter un troisième élément constitutif du signe verbal, le *référent*, devenu l'objet de la sémiotique (Eco, 1992, p. 30). Le référent est le lien entre le signe verbal et le réel extérieur au signe, les *realia*, ces réalités extérieures à la langue. Ce renvoi d'un mot au réel représenté par ce mot a lieu qu'il s'agisse d'un réel tangible (exprimé par un mot concret comme neige) ou d'un réel intangible (exprimé par un mot abstrait comme liberté), car le processus mental de *catégorisation cognitive* est le même dans les deux cas (Rozencwajg et Corroyer, 2007, p. 25-40). Au Canada, la notion «Nord» fait référence aux deux types de réalité: une réalité concrète (un territoire palpable, avec son climat et ses habitants) ainsi qu'une réalité abstraite (l'idée que l'on s'en fait, la mémoire qu'on en a, et les mots qui y sont associés).

Pour les jeunes Manitobains, les référents des mots associés au Nord se trouvent concrètement dans les réalités auxquelles ces locuteurs sont exposés de façon récurrente. Leur vocabulaire de proximité en porte la marque, car le développement lexical et sémantique se fait dans l'expérience immédiate d'un milieu naturel où la neige n'est jamais bien loin. Comme les mots s'acquièrent toujours en contexte, l'un après l'autre, s'insérant un à un dans une organisation mentale dynamique où ils se renforcent et s'opposent les uns aux autres, le réseau lexical et sémantique en référence au Nord naît et se développe aussi de cette manière.

Le vocabulaire nordique analysé ci-dessous provient des corpus issus de nos trois enquêtes de terrain, réalisées en 1990–1991, en 1993 et en 2004–2006, dans le but de mesurer la *disponibilité lexicale* des jeunes, en douze points du Manitoba francophone par des études séquentielles et longitudinales.

Ces enquêtes totalisent une participation de 425 témoins. Les données des trois enquêtes ont permis d'établir trois corpus, comptant plus de 200 000 mots (occurrence totale des attestations), soit plus de 25 000 mots différents (dits *lemmes*). Celui de 1990–1991 est publié (Rodriguez, 2006, p. 205-519), les deux suivants sont en voie de publication. Les trois enquêtes ont ciblé seize champs lexico-sémantiques (Centres d'intérêt, notés *C. I.* ci-dessous), dont celui de la faune, le C.I. 14, Les animaux, étudié ici. Du fait du nombre élevé de témoins et d'attestations, les mesures statisitiques obtenues sont stables, et fiables quant à leur représentativité du groupe.

Le vocabulaire recueilli et élaboré en corpus constitue le *vocabulaire disponible* des jeunes. C'est le vocabulaire utile, stable et durable, nommé «français fondamental» (Gougenheim et al., 1964), qui est celui des jeunes, à l'âge-clé de huit à treize ans où il se manifeste. C'est celui auquel ils font appel spontanément pour s'exprimer. La disponibilité lexicale est donc d'abord un concept de psycholinguistique, mis au point par André Michéa: «Un mot disponible est un mot qui, sans être particulièrement fréquent, est cependant toujours prêt à être employé et se présente immédiatement et naturellement à l'esprit au moment où l'on en a besoin.» (Michéa, 1953, p. 340). En effet, lors de la collecte des mots disponibles, les témoins donnent les mots qui leur viennent spontanément à l'esprit, en libre association mentale, ce qui abolit la distinction entre langue parlée et langue écrite (Picoche, 1993, p. 59-63). L'évaluation de la disponibilité allie la psycholinguistique à la statistique car le vocabulaire disponible se mesure par des méthodes statistiques, présentées ci-dessous.

Mais, d'abord, comment recueille-t-on le vocabulaire disponible? Cela se fait par un type d'enquête de terrain portant sur un groupe de témoins (et non sur un seul individu) en chaque point géographique enquêté, ce qui assure une fiabilité statistique aux résultats. La recherche des mots disponibles se fait en trois étapes. Elle débute par une enquête de terrain où les témoins produisent des mots selon un protocole défini (travail individuel, salle de classe totalement dépourvue d'aides visuelles et verbales, et de tout adjuvant pédagogique, temps limité pour réponse, etc., Rodriguez, 2006, p. 83-85). Ensuite, les mots relevés lors de l'enquête sont lemmatisés (pour éliminer les

formes hétérographes des données brutes) et balisés de critères linguistiques (mot/notion, emprunt à l'anglais, mot régional, etc.) et sociolinguistiques (âge, sexe, milieu, etc.) à leur saisie dans le corpus de données informatisé. Enfin, ce corpus de données est transformé, par un programme d'analyse statistique élaboré spécifiquement pour cette recherche (Rodriguez 2006, p. 85-88), en «corpus de disponibilité» où *l'indice de disponibilité* de chaque mot est calculé. La disponibilité lexicométrique globale des mots est ainsi établie statistiquement, ainsi que pour chacun des critères balisés dans le corpus de saisie des données.

La disponibilité s'évalue donc selon une méthodologie et avec des outils appartenant à la *lexicométrie*, ou étude statistique du langage (Muller, 1993). La disponibilité est l'un des cinq domaines lexicométriques, les quatre autres étant la *répartition*, la *valence*, le *rang associatif* (à ne pas confondre avec le *rang de disponibilité*, indiqué dans nos tableaux) et la *fréquence* (Rodriguez, 2006, p. 74-76). Ces cinq domaines sont complémentaires, chacun ayant sa méthodologie spécifique pour recueillir des données linguistiques destinées à constituer un corpus, et tous ayant en commun le recours à des outils informatiques. La lexicométrie connaît aujourd'hui un essor considérable grâce à des logiciels performants pour étudier la fréquence en analyse textuelle (Salem et al, 2013), et grâce aux techniques empruntées à la géométique et au SIG (système d'information géographique) appliquées en dialectologie. Une présentation détaillée des relations entre domaines lexicométriques sortirait du champ de notre sujet.

La disponibilité est une mesure précise du vocabulaire spontanément utilisé par des locuteurs donnés, dans un contexte socio-géographique donné. Statistiquement, elle se définit comme la *fréquence relative*, calculée par la formule (100:T) *n*, ou (*n*:T) 100, où T représente le nombre de témoins et *n* la fréquence absolue (occurrence) de l'attestation. Elle permet notamment de connaître le nombre et la proportion de témoins utilisant chaque mot attesté. Par opposition, la *fréquence absolue* du lexique d'un ensemble de documents (calculée selon d'autres formules) sert à construire les banques de données lexicales électroniques. Il est important de comprendre que les mots de tête (les plus fréquents) obtenus par la méthode de la fréquence

absolue sont des mots grammaticaux (ne, pas, il, etc.) alors que les mots disponibles sont tous des mots thématiques, des noms concrets, au référent immédiatement identifiable. Sur les plans lexicométrique et sémantique, un mot disponible est donc le contraire d'un mot fréquent (Michéa, 1950, p. 188-189). C'est pourquoi la disponibilité (fréquence relative) et la fréquence (absolue) sont complémentaires dans la mesure du lexique d'une langue. De plus, précisément parce qu'elle fait apparaître les mots concrets, la méthode de la disponibilité est la mieux adaptée à la recherche d'un vocabulaire concret comme celui à référence nordique, qui est notre sujet.

3. Le corpus des mots disponibles du C.I. Les animaux

Comme nos corpus du vocabulaire des jeunes incluent de nombreux autres champs lexico-sémantiques que celui des animaux, nous avons dû circonscrire le vocabulaire de la faune dans la masse du lexique disponible de l'ensemble des trois synchronies (1990-1991, 1993 et 2006), qui est celle de dix-sept ans d'usage lexical fondamental. Le but de ces calculs est d'évaluer avec précision la présence lexicale nordique sur le plan quantitatif (statistique) et qualitatif (sémantique, aussi mesuré statistiquement) les mots du C. I. Les animaux, et de trois de ses sous-ensembles (animaux domestiques, nordiques et exotiques).

Pour le C.I. Les animaux, l'ensemble du triple corpus se dénombre ainsi: 895 mots différents (pour une occurrence totale de 11 968 mots) en 1991 (Rodriguez 2006, p. 445); 309 mots différents (pour une occurrence totale de 1302 mots) en 1993; et 450 mots différents (pour une occurrence totale de 2755 mots) en 2006. Pour ce C.I., le nombre de témoins est de 425, dont 206 filles et 219 garçons, 323 de milieu rural et 102 de milieu urbain. Nous avons calculé la disponibilité de ce vocabulaire de la faune globalement (pour la totalité des témoins), et en fonction des critères croisés balisés: filles, garçons, milieu rural et milieu urbain (notés F, G, R et U ci-après). Les calculs croisés permettent une vérification accrue par la comparaison des quatre groupes de résultats.

Dans cette évaluation du Nord lexical et des mots qui le dénotent, le connotent et l'expriment, le calcul de disponibilité pour la totalité des témoins se double donc de celui de quatre

groupes sociolinguistiques F, G, U et R. Notons que les groupes U et R sont à la fois géolinguistiques (liés au lieu géographique) et sociolinguistiques (liés au milieu sociétal). Le nombre des mots les plus disponibles rattachés à la notion Nord et de ceux du vocabulaire domestique et exotique ont été calculés et comparés. Cette comparaison a fait apparaître un nombre de mots très élevé et à forte disponibilité exprimant la notion Nord (tableaux 1a, 1b, 2a et 2b), ainsi que des particularités sémantiques et morphologiques du vocabulaire nordique (tableaux 3 et 4) – autant de traits absents du vocabulaire de la faune exotique.

Nous ne donnerons que le résultat quantitatif global (tous groupes confondus) de l'étude longitudinale: en dix-sept ans, la moyenne de mots disponibles par témoin est passée de 37 mots pour 32 témoins, à 41 mots pour 66 témoins. Cette différence dénote une dispersion lexicale, dont le calcul est utile à l'étude longitudinale, mais qui sort du cadre de notre propos. (À titre d'information, nous dirons simplement que la dispersion est le rapport entre l'occurrence et le nombre de lemmes, et la variance «un indice de dispersion autour de la moyenne» (Dugast, 1980, p. 35). La mesure lexicale de la variance peut être affinée statistiquement par la dispersion absolue (par exemple, l'étendue, X=Xmax-Xmin), les quantiles (indispensables en cartographie), les interquartiles, l'écart absolu moyen et l'écart type, ou de la dispersion relative (coefficient de variation, écart moyen relatif ou coefficient interquartile relatif). Le principe commun à tous ces critères de dispersion est l'établissement d'un rapport de deux nombres (ayant la même unité de mesure) montrant l'écart entre chaque valeur (mot) et la valeur médiane (la dispersion se note en %, à ne pas confondre avec le taux de disponibilité, qui s'indique aussi en %.).

4. Le vocabulaire disponible en référence au Nord chez les filles et les garçons

Les tableaux et leurs commentaires utilisent le même code de référence: F pour le groupe des filles, G pour le groupe des garçons, et A, N et E pour chaque sous-ensemble lexico-sémantique: types lexicaux A (autres, divers) se référant aux animaux domestiques ou petits animaux sauvages vivant près des fermes, et non spécifiques à une faune nordique; types lexicaux N se référant à la faune sauvage canadienne, donc

septentrionale; et types lexicaux E se référant à la faune exotique, étrangère au Canada.

Les tableaux 1a et 1b (en annexe) montrent les résultats quantitatifs du C.I. Les animaux, pour les 50 mots de tête (ceux de la plus forte disponibilité) dans les groupes F (206 filles) et G (219 garçons), présentés en parallèle dans les tableaux. Les mots y sont classés par rang décroissant, ce rang étant déterminé par l'indice de disponibilité, qui descend de 98% pour le rang 1 à 19% pour le rang 50 (seuls les rangs sont indiqués dans ces tableaux).

Sur le plan quantitatif, indépendamment des types lexicaux et de leur sens, le groupe F (filles) totalise 721 mots disponibles différents sur une occurrence de 8020 mots, avec une moyenne de 39 mots par témoin. Le groupe G (garçons) totalise 836 mots disponibles différents sur une occurrence de 8003 mots, avec une moyenne de 37 mots par témoin. Le groupe F a donc produit une plus grande occurrence que le groupe G, moins de mots différents, et une moyenne par témoin supérieure. Quantitativement, les deux groupes ont donc une production lexicale relativement similaire.

Sur le plan qualitatif, celui du sens des lexies, les tableaux 1a et 1b présentent, en trois colonnes distinctes pour F et G, les trois sous-groupes sémantiques des 50 mots de tête, A, N et E. Les types lexicaux A, N et E sont présents chez les deux groupes F et G, dans des proportions similaires, mais non identiques sémantiquement. Il y a chez les filles une prédominance, en nombre et en indice, des mots se référant à des animaux domestiques ou à la petite faune de la ferme, avec 25 mots du type A, contre 22 chez les garçons. Absents du groupe G, des mots disponibles comme poussin, hamster et abeille sont probablement dus, dans le groupe F, à une expérience plus fréquente des réalités de la ferme, constatée lors de notre enquête de terrain.

Contrairement aux mots de type A, ceux du type N sont plus nombreux chez les garçons (15 mots) que chez les filles (13 mots). Absents du groupe des filles, les mots disponibles lynx, cougar et coyote, désignant de grands mammifères, sont probablement dus à une exposition plus grande aux réalités référentielles de la chasse dans le groupe des garçons, une

explication qui s'appuie, comme dans le cas du groupe F, sur l'importance du contact avec le référent dans l'acquisition lexicale. En effet, comme nous avons pu le constater sur le terrain, à la campagne, les rôles traditionnels masculins et féminins sont généralement répartis: par exemple, la chasse aux mammifères et la moisson sont plutôt pratiquées par les hommes, et le travail au jardin et à la basse-cour plutôt par les femmes, même si, parfois, les deux domaines d'activité sont partagés.

Enfin, les mots du type lexical E, aux référents exotiques, présentent moins d'écart quantitatif entre les deux groupes F et G. Les mots de type E sont aussi les moins disponibles des trois types, en nombre comme en indice: 13 chez les garçons, 12 chez les filles. La prédominance des mots N aux référents nordiques sur les mots E aux références exotiques se confirme dans l'ensemble des résultats lexicométriques de ce C. I., au-delà des 50 premiers mots. Ce C. I. inclut en effet plus de 115 mots aux références nordiques, contre 55 mots aux référents exotiques. Du 51e rang au 895e et dernier rang, rangs de moyenne et faible disponibilité, les types N continuent néanmoins à être plus nombreux et plus disponibles que les types E.

5. Le vocabulaire disponible en référence au Nord en milieu rural et en milieu urbain

Le calcul lexicométrique pour les groupes rural et urbain porte sur 323 locuteurs de milieu rural (R) et 102 de milieu urbain (U). Les mots des tableaux 2a et 2b (en annexe) montrent les résultats quantitatifs du C. I. Les animaux pour les 50 mots de tête (ceux de la plus forte disponibilité) pour ces deux groupes, présentés en parallèle dans les tableaux. Les mots y sont classés par rang descendant, le rang étant déterminé par l'indice de disponibilité (seuls les rangs sont indiqués dans les tableaux). Ces indices vont de 98% pour le rang 1 à 19% pour le rang 50.

Sur le plan quantitatif, indépendamment des types lexicaux et de leur sens, le groupe R totalise 920 mots disponibles différents (sur une occurrence de 12 319 mots), avec une moyenne de 38 mots par témoin. Le groupe U totalise 527 mots disponibles différents, avec une moyenne de 36 mots par témoin. Le groupe R a donc produit plus de mots différents que

le groupe U (920 contre 527), avec une moyenne par témoin plus élevée (38 contre 36).

Sur le plan qualitatif, celui du sens des mots, les tableaux 2a et 2b présentent, en colonnes distinctes pour R et U, les trois sous-groupes sémantiques des 50 mots de tête: types lexicaux A N et E. Le groupe R a produit un grand nombre de types A (24) et de types N (15), et un petit nombre de types E (11). À l'opposé, le groupe U a produit moins de types A (21), moins de N (13) et beaucoup plus de E (16). Ces écarts résultent du vécu des jeunes témoins: contact répété avec les réalités nordiques pour les témoins de milieu rural, et contact *de visu* avec des réalités exotiques pour les témoins de milieu urbain (présence d'un zoo en ville, par exemple). Quant au rôle de la littérature enfantine, qui met régulièrement en scène une faune exotique, et qui est pratiquée de manière similaire dans les deux groupes R et U, elle semble moins affecter le groupe R que le groupe U.

Les mots de type N aux référents nordiques sont donc plus disponibles que ceux de type E aux référents exotiques. Ceci est également vrai chez les filles comme chez les garçons, mais dans le groupe rural plus que dans le groupe urbain. Visible dans les tableaux 1a, 1b, 2a et 2b des 50 mots de tête, cette prédominance se confirme sur l'ensemble des mots disponibles de ce C.I. des animaux, qui contient plus du double de mots à référent nordique que de mots à référent exotique.

6. Génériques et spécifiques dans le vocabulaire disponible en référence au Nord

Non seulement les mots disponibles de type lexical N sont plus nombreux que ceux de type E, mais ils sont aussi distincts des points de vue sémantique et morphologique. Les tableaux 3 et 4 (en annexe) présentent la précision sémantique et la diversité morphologique du vocabulaire nordique. Les calculs lexicométriques pour les deux mots-notions «baleine» et «renard», ainsi que pour les autres mots disponibles de ces deux tableaux, ont été réalisés sur cinq critères croisés: mot, F, G, U et R.

Le tableau 3, consacré à la notion «baleine», montre la variété lexicale, sémantique et morphologique des types lexicaux N. La notion Nord s'exprime par une diversité de

noms, dont la précision sémantique ne recule pas devant des mots plus rares, tels que orque et épaulard, également attestés – probablement renforcés par un usage scolaire vite assimilé dès qu'il porte sur des objets géographiquement proches des jeunes apprenants. En outre, les programmes scolaires font une large place à l'étude du milieu naturel septentrional dans les écoles canadiennes (donc nordiques). Le souci de précision des locuteurs donne lieu aussi à une diversité morphologique, avec de nombreux mots composés où au mot principal s'adjoint un modificateur nominal, adjectival ou prépositionnel, c'est-à-dire un générique suivi d'un spécifique, selon la nomenclature binominale de Linné. (Nous éviterons ici les termes lexicographiques hyperonyme/hyponyme, utiles aussi en sémantique pour analyser le glissement hyperonymique/hyponymique d'une même lexie, selon l'ensemble et le sous-ensemble sémique considéré). Le générique (mot-notion) obtient toujours un taux de disponibilité très élevé (34% pour baleine chez les filles et 35,2% chez les garçons), mais plusieurs spécifiques sont présents, de moyenne ou faible disponibilité. C'est le cas pour les spécifiques attestés du générique baleine, tels que baleine bleue, baleine grise, baleine blanche, baleine requin, etc. Même en anglais, des spécifiques N sont attestés (tel *humpback whale*). Contrairement aux types lexicaux N, les types E sont de la langue courante et seulement génériques, tels gorille et singe. Les mots moins courants, comme macaque ou marmoset, sont absents. De même, les mots E sont dénués de spécifiques: les génériques tigre et singe sont attestés, mais ni tigre du Bengale ni singe du Népal ne le sont.

Le tableau 4, consacré à la notion «renard», confirme la précision lexicale et sémantique des mots de type N et de leur diversité morphologique accrue: présence de composés (renard argenté, renard roux, etc.) et de dérivés, avec le nom du mâle, de la femelle et celui du jeune. C'est le cas de la série sémantique disponible renard/renarde/renardeau, qui n'a aucun équivalent parmi les mots de type E attestés, à l'exception des couples très courants tigre/tigresse et lion/lionne, que ce soit en français ou en anglais. Les mots anglais sont d'ailleurs absents des 50 mots de tête, et peu nombreux dans l'ensemble des mots disponibles de ce champ lexico-sémantique.

Comme les exemples des tableaux 3 et 4 le montrent, le nombre et la diversité des mots de type N se constate dans les quatre groupes, filles et garçons, locuteurs ruraux et locuteurs urbains. Cette diversité lexicale, sémantique et morphologique des types lexicaux N, renvoyant à un référent nordique, apparaît très nettement dans l'ensemble des corpus pour ce C.I. de la faune. Nous avons obtenu des résultats dans notre mesure du vocabulaire à référence nordique dans d'autres champs lexico-sémantiques (vêtements, métiers, transports, etc.), ce qui témoigne de la profondeur lexicale quantitative et qualitative des types N du vocabulaire disponible des jeunes.

7. Conclusions

Notre exploration du vocabulaire disponible des jeunes Manitobains visant la notion des animaux a montré que, statistiquement, les types lexicaux N à référent nordique sont très présents dans le vocabulaire disponible de ces jeunes, tant chez les filles que chez les garçons, tant en lieu/milieu rural qu'en lieu/milieu urbain. Les calculs lexicométriques effectués sur les trois corpus révèlent que les types N sont supérieurs en nombre et en diversité lexicale, morphologique et sémantique aux types lexicaux E à référent exotique, et que cette même précision se retrouve dans d'autres notions, par exemple celles des aliments et des transports, où les mots à contenu nordique sont toujours plus fréquents et variés que ceux évoquant le Sud. C'est une *constante* lexicométrique dans nos corpus.

Cette constante se présente avec une fréquence similaire dans les quatre groupes F/G et R/U, même s'il existe quelques nuances entre eux: les groupes G et R, par exemple, possèdent plus de mots nordiques disponibles; le groupe F possède plus de mots disponibles à référent domestique (maison ou ferme), tout en ayant aussi des mots nordiques. Cette constante quantitative et qualitative est caractéristique des types N, alors que la constante pour les types E est leur limite en nombre comme en diversité morphologique et sémantique. Pour les jeunes Manitobains, les mots E, à référent exotique ou lointain (animaux entrevus, au zoo, sur un écran de télévision ou d'ordinateur, ou dans un film ou un livre) ne sont que génériques, et restent dépourvus de ce qui caractérise les mots N: renvoi à la nature, à la couleur, à l'odeur et au son des réalités de l'expérience immédiate, d'où leur rôle essentiel dans l'imagination et l'imaginaire, ainsi que

mentionné plus haut. L'analyse lexicométrique des données de terrain permet ainsi de mesurer et de comprendre l'ampleur et la spécificité du vocabulaire disponible acquis dans l'enfance, et de procurer de nouvelles pistes pédagogiques encourageant la créativité verbale.

Ces conclusions sur l'analyse lexicométrique du vocabulaire de la faune nous ramènent à notre définition initiale du signe et de son référent. Le vocabulaire disponible des jeunes se construit en référence aux réalités géographiques et climatiques vécues par eux, et constitue le socle sur lequel se construira et s'élaborera le lexique de l'adulte. En plus du développement des réseaux lexical et sémantique initiés dès le jeune âge de nos témoins, ce vocabulaire de proximité du milieu nordique nourrit déjà leur faculté métaphorique, comme le prouvent certaines attestations (Sasquatch; homme abominable). Les résultats lexicométriques du vocabulaire disponible de nos témoins rejoignent l'analyse de Bachelard sur le rôle du milieu géoclimatique de l'enfance, véritable berceau de l'imagination, avec le potentiel métaphorique et identitaire du lexique qu'il suscite. Cela est vrai pour tout le monde, où que soit le lieu de l'enfance. Au sortir de l'enfance, le *vocabulaire fondamental* accompagne toute la vie le locuteur devenu adulte. La présence nordique chez l'adulte, avec la force de son pouvoir métaphorique et identitaire, se constate dans la langue des adultes, à fortiori dans celle des créateurs, non seulement de l'Ouest, mais ailleurs au Canada, où le Nord est toujours présent, à portée de route. Nous pensons à ces artistes forgés par le Nord, comme la musicienne inuite Tanya Tagaq, dont les chants de gorge (comme *Uja*) reflètent la langue acquise pendant sa petite enfance, dans le Grand Nord. Elle a reçu le Prix Polaris 2014 pour son album *Animism*. Il en est de même pour le cinéaste québécois Denis Côté et son modeste film vidéo *Les États nordiques*, qui lui a valu la surprise du prestigieux Léopard d'or 2005 de Laucarno (Suisse) pour avoir capté l'essence nordique de la vie et de la langue des habitants de Radisson, minuscule village de la Baie James, et pour avoir exprimé la quête du Nord, comme lieu de refuge et de réflexion sur soi. Avec son pouvoir métaphorique, voire artistique, le lexique de l'enfance, dont cette étude sur le vocabulaire nordique en témoigne, apporte une dimension inattendue à un domaine tel que la lexicométrie, et fournit aux

éducateurs et aux chercheurs une raison supplémentaire de se pencher sur le vocabulaire disponible des jeunes.

BIBLIOGRAPHIE

BROWN, Craig (1990) *Histoire générale du Canada*, Québec, Les Éditions du Boréal, 695 p.

ECO, Umberto (1992) *Le Signe*, Paris, Le Livre de Poche, Biblio Essais, 280 p.

DUGAST, Daniel, (1980) *La Statistique lexicale*, Genève, Slatkine, 107 pages.

GOUGENHEIM Georges, RIVENC, Paul, MICHÉA, René, SAUVAGEOT, Aurélien (1956) *L'élaboration du français fondamental: Étude sur l'établissement d'un vocabulaire et d'une grammaire de base,* Paris, Didier, 257 p.

GOUVERNEMENT DU CANADA, *Historica Canada*, (2017) Anthony Wilson-Smith, Publisher. encyclopédiecanadienne.ca

MICHÉA, René (1953) Michéa R. (1953) «Mots fréquents et mots disponibles, un nouvel aspect de la statistique lexicale », dans *Langues Modernes*, volume 47, n° 4, 338-344.

MULLER Charles (1993), *Initiation aux méthodes de la statistique linguistique*, Paris, Champion,

OROZENCWAJG, Paulette et CORROYER, Denis (2007) «L'analyse des processus cognitifs dans une version adaptée du test des similitudes de Wechsler», dans *Psychologie et Éducation*, 4, p. 25-40.

PICOCHE, Jacqueline (1993) *Didactique du vocabulaire français*, Paris, Nathan, 206 p.

PÉLISSIER, Daniel (2006) «Pourquoi et comment utiliser la lexicométrie pour l'analyse de focus groups?», dans *Présence numérique des organisations*, http://presnumorg.hypotheses.org/214.

REINERT, Max Reinert (1979), *Alceste, Analyse Lexicale par Contexte d'un Ensemble de Segments de Texte*, Paris, Laboratoire de Jean-Paul Benzécri, CNRS. Diffusion Société Image (Choeb Zafar, dir.). contact@image-zafar.com

RICOEUR, Paul (1975) *La Métaphore vive*, Paris, Seuil, 416 p.

RODRIGUEZ, Liliane (2006) *La langue française au Manitoba (Canada). Histoire et évolution lexicométrique*, Tübingen, Niemeyer, 519 p.

RODRIGUEZ, Liliane (2012) «Le rôle du marquage lexicométrique des régionalismes dans un corpus franco-canadien», *Jadt2012, Statistical Analysis of Textual Data*, Proceedings of the 11th International Conference Journées d'analyse statistique des donnés textuelles, p. 871-882.

SALEM, André, LAMALLE, Cédric, MARTINEZ, William, FLEURY, Serge (2003) *Outils de statistique textuelle*, Paris, Universite´ de la Sorbonne nouvelle - Paris 3, SYLED - CLA2T Version 3.41.

SAUSSURE, Ferdinand de, (1972 [1916]) *Cours de linguistique générale*, Paris, Payot, 510 p.

SLOTERDIJK, Peter (2014) *Les Lignes et les Jours. Notes (2008-2011)*, Paris, Éditeur Libella-Maren Sell, 624 p.

ANNEXES

Tableau 1a

Le vocabulaire disponible des jeunes Manitobains - C.I. 14 Les animaux

Les 50 mots les plus disponibles: rangs 1-25 (de 98,1% à 33,8% de disponiblité)
Catégorie sémantique: mots A (Autres) / mots N (Nord) / mots E (Exotiques)
Catégorie sociolinguistique: filles (F) / garçons (G)

FILLES (F) - 206 témoins **721 mots différents - moyenne par témoin: 39**			**GARÇONS (G) - 219 témoins** **836 mots différents - moyenne par témoin: 37**		
mots A	**mots N**	**mots E**	**mots A**	**mots N**	**mots E**
1. cheval			1. chat		
2. chat			2. chien		
3. chien			3. vache		
4. vache			4. cheval		
5. poisson			5. poisson		
		6. éléphant			6. lion
7. cochon			7. cochon		
8. oiseau					8. éléphant
9. lapin					9. tigre
	10. ours			10. ours	
		11. lion	11. souris		

			12. singe	12. oiseau		
13. poule				13. lapin		
			14. tigre			14. singe
15. souris					15. loup	
16. canard				16. serpent		
		17. loup		17. rat		
18. serpent				18. poule		
			19. girafe	19. tortue		
20. rat					20. chevreuil	
21. tortue						21. girafe
		22. renard		22. canard		
			23.kangourou		23. baleine	
			24. zébre			24. kangourou
25. mouton					25. renard	
total 1	15	3	7	14	5	6

Tableau 1b

Le vocabulaire disponible des jeunes Manitobains - C.I. 14 Les animaux

Les 50 mots les plus disponibles: rangs 15-50 (de 32,4% à 16,9% de disponiblité)
Catégorie sémantique: mots A (Autres) / mots N (Nord) / mots E (Exotiques)
Catégorie sociolinguistique: filles (F) / garçons (G)

FILLES (F) - 206 témoins **721 mots différents* - moyenne par témoin: 39**			**GARÇONS (G) - 219 témoins** **836 mots différents* - moyenne par témoin: 37**		
mots A	**mots N**	**mots E**	**mots A**	**mots N**	**mots E**
26. hamster				26. bison	
27. grenouille			27. mouton		
	28. chevreuil			28. castor	
	29. baleine				29. zébre
30. coq			30. écureuil		
31. écureuil			31. hamster		
	32. dinosaure[1]			32. lynx	
	33. bison				33. crocodile
		34. chameau	34. chèvre		
35. hibou				35. requin[2]	
		36. panda	36. boeuf		
	37. phoque		37. grenouille		

38. chèvre						38. hippopotame
		39. castor				39. chameau
			40. koala	40. hibou		
			41. hippopotame			41. panda
		42. ours polaire			42. dinosaure[1]	
			43. crocodile		43. phoque	
		44. mouffette			44. raton laveur	
		45. requin[2]				45. alligator
46. mouche					46. cougar	
47. abeille				47. mouche		
48. moustique					48. orignal	
		49. orignal			49. coyote	
50. poussin						50. gorille
total 2	10	10	5	8	10	7
total 1+2	25	13	12	22	15	13

* Nombre de types lexicaux N: 115+ et de types lexicaux E: 55+ dans l'ensemble du C. I.

1 Pour nos témoins, ce mot désigne les dinosaures de l'Alberta.

2 Il s'agit de nombreuses espèces de requins nordiques (aiguillat du Saint-Laurent, requin maraîche du Nord-Ouest, requin du Groënland, requin pèlerin, etc.)

Tableau 2a

Le vocabulaire disponible des jeunes Manitobains - C.I. 14 Les animaux

Les 50 mots les plus disponibles: rangs 1-25 (de 97,1% à 34,7% de disponiblité)
Catégorie sémantique: mots A (Autres) / mots N (Nord) / mots E (Exotiques)
Catégorie géolinguistique et sociolinguistique: milieu rural (R) / milieu urbain (U)

MILIEU RURAL (R): 323 témoins 920 mots différents* - moyenne par témoin: 38			**MILIEU URBAIN (U): 102 témoins 527 mots différents* - moyenne par témoin: 36**		
mots A	**mots N**	**mots E**	**mots A**	**mots N**	**mots E**
1. chat			1. chien		
2. chien			2. chat		
3. cheval			3. cheval		
4. vache					4. éléphant
5. poisson					5. lion
6. cochon			6. vache		
7. lapin					7. singe
		8. éléphant	8. cochon		
		9. lion	9. poisson		
	10. ours		10. oiseau		
11. oiseau					11. tigre

12. souris					12. ours	
			13. tigre	13. souris		
		14. loup		14. lapin		
15. poule				15. rat		
			16. singe	16. poule		
17. serpent						17. kangourou
18. canard				18. serpent		
19. rat				19. hamster		
20. tortue						20. girafe
			21. girafe		21. baleine	
		22. chevreuil		22. tortue		
		23. renard		23. écureuil		
24. mouton					24. loup	
			25. zèbre		25. requin[1]	
total 1	15	4	6	15	4	6

* Nombre de types lexicaux N: 115+ et de types lexicaux E: 55+ dans l'ensemble du C. I.

1 Il s'agit de nombreuses espèces de requins nordiques (aiguillat du Saint-Laurent, requin maraîche du Nord-Ouest, requin du Groënland, requin pèlerin, etc.).

Tableau 2b

Le vocabulaire disponible des jeunes Manitobains - C.I. 14 Les animaux

Les 50 mots les plus disponibles: rangs 15-50 (de 34,3% à 18,9% de disponiblité)
Catégorie sémantique: mots A (Autres) / mots N (Nord) /mots E (Exotiques)
Catégorie géolinguistique et sociolinguistique: milieu rural (R) / milieu urbain (U)

MILIEU RURAL (R): 323 témoins **920 mots différents* - moyenne par témoin: 38**			**MILIEU URBAIN (U): 102 témoins** **527 mots différents* - moyenne par témoin: 36**		
mots A	**mots N**	**mots E**	**mots A**	**mots N**	**mots E**
		26. kangourou			26. zèbre
	27. bison		27. canard		
	28. baleine				28. panda
29. grenouille				29. renard	
30. hibou			30. coq		
		31. chameau		31. dinosaure[1]	
32. chèvre			32. grenouille		
	33. phoque				33. koala
	34. castor		34. mouton		
35. écureuil					35. alligator
36. hamster				36. castor	
	37. dinosaure[1]			37. chevreuil	

			38. crocodile		38. lynx	
			39. hippopotame			39. crocodile
			40. panda			40. pingouin[3]
41. mouche					41. bison	
		42. lynx		42. boeuf		
43. coq						43. guépard
		44. ours polaire				44. hippopotame
		45. cougar			45. ours polaire	
		46. requin[2]		46. dauphin		
47. boeuf						47. gorille
		48.raton laveur			48. orignal	
49. abeille					49. raton laveur	
		50. orignal				50. chameau
total 2	9	11	5	6	9	10
total 1+2	24	15	11	21	13	16

* Nombre de types lexicaux N: 115+ et de types lexicaux E: 55+ dans l'ensemble du C. I.

1 Pour nos témoins, ce mot désigne les dinosaures de l'Alberta.

2 Il s'agit de nombreuses espèces de requins nordiques (aiguillat du Saint-Laurent, requin maraîche du Nord-Ouest, requin du Groënland, requin pèlerin, etc.).

3 Pour nos témoins, ce mot désigne le manchot (de la Terre-Adélie, dans l'Antarctique), sous l'influence de l'anglais et d'autres langues, le grand pingouin du Canada ayant d'ailleurs disparu au XIXe siècle.

Tableau 3

Le vocabulaire disponible des jeunes Manitobains -
C.I. 14 Les animaux

Catégorie sémantique: notion N «baleine»
Catégories sociolinguistiques: filles (F)/garçons (G) et milieu urbain (U)/rural (R)

notion «baleine»	**F**	**G**	**U**	**R**
baleine	√ 34%	√ 35,2%	√	√
baleine bleue	√	√	√	√
baleine grise	√	√	√	√
baleine blanche	√	√	√	√
baleine rose	√			√
bélouga	√	√		√
baleine bélouga	√			√
baleine requin		√		√
orque		√	√	
épaulard		√	√	
whale	√	√	√	√
beluga whale		√	√	
killer whale	√	√	√	√
sperm whale		√		√
humpback whale		√		√
total de mots différents	**9**	**13**	**9**	**12**

Tableau 4

Le vocabulaire disponible des jeunes Manitobains -
C.I. 14 Les animaux

Catégories sémantique et morphologique: notion N «renard»
Catégories sociolinguistiques: filles (F)/garçons (G) et milieu urbain (U)/rural (R)

notion «renard»	**F**	**G**	**U**	**R**
renard	√ 41,3%	√ 33,8%	√	√
renarde	√	√	√	√
renardeau	√			√
renard polaire	√		√	
renard blanc	√	√	√	
renard arctique	√	√		√
renard de l'Arctique		√	√	
renard gris		√	√	
renard argenté		√	√	√
renard roux	√	√	√	√
renard rouge		√	√	
total de mots différents	**7**	**9**	**9**	**6**