

AIDAinformazioni

RIVISTA SEMESTRALE DI SCIENZE DELL'INFORMAZIONE

Fondata nel 1983 da Paolo BISOGNO

N. 3-4 — Anno 38 — luglio-dicembre 2020

Proprietario della rivista
UNIVERSITÀ DELLA CALABRIA

Direttore Scientifico
Roberto GUARASCI
Università della Calabria

Direttore Responsabile
Fabrizia Flavia SERNIA

Comitato scientifico

Anna Rovella, UNIVERSITÀ DELLA CALABRIA
Maria Guercio, SAPIENZA UNIVERSITÀ DI ROMA
Giovanni Adamo, CONSIGLIO NAZIONALE DELLE RICERCHE
Claudio Gnoli, UNIVERSITÀ DEGLI STUDI DI PAVIA
Ferruccio Diozzi, CENTRO ITALIANO RICERCHE AEROSPAZIALI
Gino Roncaglia, UNIVERSITÀ DELLA TUSCIA
Laurence Favier, UNIVERSITÉ CHARLES-DE-GAULLE LILLE 3
Madjid Ihadjadene, UNIVERSITÉ VINCENNES-SAINT-DÉNIS PARIS 8
Maria Mirabelli, UNIVERSITÀ DELLA CALABRIA
Agustín Vivas Moreno, UNIVERSIDAD DE EXTREMADURA
Douglas Tudhope, UNIVERSITY OF SOUTH WALES
Christian Galinski, INTERNATIONAL INFORMATION CENTRE FOR TERMINOLOGY
Béatrice Daille, UNIVERSITÉ DE NANTES
Alexander Murzaku, COLLEGE OF SAINT ELIZABETH, USA

Comitato di redazione

Antonietta Folino, UNIVERSITÀ DELLA CALABRIA
Erika Pasceri, UNIVERSITÀ DELLA CALABRIA
Maria Taverniti, CONSIGLIO NAZIONALE DELLE RICERCHE
Maria Teresa Chiaravalloti, CONSIGLIO NAZIONALE DELLE RICERCHE
Assunta Caruso, UNIVERSITÀ DELLA CALABRIA

Segreteria di Redazione

Valeria Rovella, UNIVERSITÀ DELLA CALABRIA

ARACNE

AIDAinformazioni

RIVISTA SEMESTRALE

«AIDAinformazioni» è una rivista scientifica che pubblica articoli inerenti le Scienze dell'Informazione, la Documentazione, la Gestione Documentale e l'Organizzazione della Conoscenza. È stata fondata nel 1983 quale rivista ufficiale dell'Associazione Italiana di Documentazione Avanzata e nel febbraio 2014 è stata acquisita dal Laboratorio di Documentazione dell'Università della Calabria.

La rivista si propone di promuovere studi interdisciplinari oltre che la cooperazione e il dialogo tra profili professionali aventi competenze diverse, ma interdipendenti. I contributi possono riguardare *topics* quali Documentazione, Scienze dell'informazione e della comunicazione, Scienze del testo e del documento, Organizzazione e Gestione della conoscenza, Terminologia, Statistica testuale e Linguistica computazionale e possono illustrare studi sperimentali in domini specialistici, casi di studio, aspetti e risultati metodologici conseguiti in attività di ricerca applicata, presentazioni dello stato dell'arte, ecc.

«AIDAinformazioni» è riconosciuta dall'ANVUR come rivista di Classe A per l'Area 11 – Settore 11/A4 e censita per le Aree 10 – Scienze dell'antichità, filologico-letterarie e storico-artistiche; 11 – Scienze storiche, filosofiche, pedagogiche e psicologiche; 12 – Scienze giuridiche; 14 – Scienze politiche e sociali, così come dall'AERES (Agence d'évaluation de la recherche et de l'enseignement supérieur) che la annovera tra le riviste scientifiche dell'ambito delle Scienze dell'Informazione e della Comunicazione. La rivista è, inoltre, indicizzata in: ACNP – Catalogo Italiano dei Periodici; BASE – Bielefeld Academic Search Engine; ERIH PLUS – European Reference Index for the Humanities and Social Sciences – EZB – Elektronische Zeitschriftenbibliothek – Universitätsbibliothek Regensburg; Gateway Bayern; KVK – Karlsruhe Virtual Catalog; Letteratura Professionale Italiana – Associazione Italiana Biblioteche; The Library Catalog of Georgetown University; SBN – Italian union catalogue; Summon™ – by SerialsSolutions; Ulrich's; UniCat – Union Catalogue of Belgian Libraries; Union Catalog of Canada; LIBRIS – Union Catalogue of Swedish Libraries; Worldcat.

I contributi sono valutati seguendo il sistema del *double blind peer review*: gli articoli ricevuti dal comitato scientifico sono inviati in forma anonima a due *referee*, selezionati sulla base della loro comprovata esperienza nei *topics* specifici del contributo in valutazione.

Condizioni di acquisto

(spese di spedizione in Italia incluse)

Abbonamento annuale euro 42,00

Fascicolo euro 26,00

Per ordini

telefono/fax: 06 4551463

e-mail: info@giocchinoonoratieditore.it

online: <http://www.aracneeditrice.it/>

Skype: aracneeditrice

Modalità di pagamento

Bonifico bancario intestato a:

Gioacchino Onorati editore S.r.l. unip.

IBAN: IT 28 B 03069 38860 10000003170

presso Banca Intesa Sanpaolo

(filiale di Ariccia)

Causale: *Abbonamento «AIDAinformazioni»*

Editore

Gioacchino Onorati editore S.r.l. – unip.

via Vittorio Veneto, 20

00020 Canterano

(06) 4551463

www.aracneeditrice.it

info@aracneeditrice.it

info@giocchinoonoratieditore.it

Copyright ©

Gioacchino Onorati editore S.r.l. – unip.

ISBN 978-88-255-3333-0

edizione a stampa ISSN 1121-0095

edizione elettronica ISSN 1594-2201

Indice

Editoriale

- 7 «La luce della scienza cerco e'l beneficio»: l'eredità di Leonardo Da Vinci illumini la rinascita dell'Italia post-pandemia
Fabrizia Flavia Sernia

Contributi

- 15 Digitalizzazione e Patrimonio Culturale Digitale. Ambiguità, ipotesi, definizioni
Nicola Barbuti, Mauro De Bari
- 33 Comunicazione on line sul COVID-19. Scelte terminologiche comuni per la statistica pubblica ufficiale
Patrizia Collesi, Michela Troia
- 55 “Il furto del secolo” nella Lione settecentesca dei banchieri luterani
Paolo De Gasperis
- 67 Vista sul *sentiment* dell'Italia di Instagram durante la pandemia da COVID-19. Il racconto di #andràtuttobene
Mariacristina Falco, Serena Pelosi, Pierluigi Vitale
- 93 Rappresentatività e copertura semantica dei kos
Antonietta Folino, Francesca Parisi

- 113 La néonymie à l'épreuve de l'ADT. Le cas du terme "francisation"
 au Québec
 Fernando Funari, Eleonora Marzi
- 129 Assisted Contracts Validation through LCS problem resolution
 Francesca Martinucci, Michele Stingo, Massimiliano Polito
- 143 Dal pavimento di pece al soffitto di cristallo. Ricostruzione di
 una carriera politica al femminile, fra testimonianze e fonti
 d'archivio
 Lucia Montesanti, Francesca Veltri
- 171 La raccolta dei dati nelle pandemie
 Erika Pasceri, Anna Perri
- 197 Dagli *Éléments de syntaxe structurale* di Lucien Tesnière alle
 Universal Dependencies. Percorsi terminologici
 Erica Quintavalle
- 213 Analyse du forum www.psychoactif.org. Identification des con-
 versations et pratiques concernant l'injection de drogues
 Marc Tanti, Hugo Moquet, Perrine Roux
- 241 Multidimensionnalité et variation terminologique dans le do-
 maine de l'édition de textes
 Silvia Domenica Zollo

Note e rubriche

- 259 Riflessioni in ambito terminologico (a distanza) durante l'au-
 tunno 2020
 Claudio Grimaldi

La néonymie à l'épreuve de l'ADT

Le cas du terme “francisation” au Québec

FERNANDO FUNARI*, ELEONORA MARZI**

ABSTRACT: Through a case study — the analysis of the discursive behaviors of the term “francization” — the article proposes to show that any neological phenomenon in terminology does not arise only from a demiurgic act having its source in the practices of standardization. On the contrary, an important role is played by the spontaneous behaviors taking place in the speech. In order to understand the cultural and intercultural data that underlies the discursive uses and behavioral habits of *francisation*, we take an approach inspired by textometry and corpus linguistics. We will question a corpus created from the website of the Ministère de l'Immigration, de la Francisation et de l'Intégration (MIFI) of the Government of Québec. After having retraced the fundamental stages of the development of *francisation* in European and American, French and Quebec terminology databases, we will analyze their semantic behavior by studying these co-occurrences in the corpus chosen on the TXM and Hyperbase softwares.

Keywords: *Francisation*, Language policies, Terminology, Neology, Textual data analysis.

1. Terminologie, néonymie, discours

Francisation est un néologisme terminologique qui s'est imposé depuis au moins quinze ans dans le discours officiel québécois et qui connaît une diffusion importante, d'abord dans le domaine du travail et de l'entreprise, puis dans d'autres domaines (économique, social, culturel, interculturel). L'Office québécois de la langue française entend par *francisation* le processus visant « la généralisation de l'utilisation du français comme langue de travail et des communications dans l'Administration et les

* Fernando Funari, Dipartimento di Formazione, Lingue, Intercultura, Letterature e Psicologia, Università degli Studi di Firenze.
fernando.funari@unifi.it.

** Eleonora Marzi, Dipartimento di Lingue, Letterature e Culture Moderne, Alma Mater Studiorum – Università di Bologna.
eleonora.marzi@unibo.it.

Cet article est le produit d'une collaboration stricte entre ses auteurs ; cependant, les paragraphes 1 et 2 doivent être attribués à Fernando Funari ; les paragraphes 3 et 4 à Eleonora Marzi.

entreprises »¹. Le terme figure comme le moteur de plusieurs initiatives et actions de politique linguistique du Québec et se prête, dès lors, à voyager entre de typologies et de genre de discours très différents.

Dire qu'il s'agit d'un néologisme² ce n'est pas dire toute la vérité, car l'innovation qui est à la base de sa création et de la diffusion de son usage concerne uniquement le changement de sa définition. Un terme *francisation* existe déjà en français comme pratique d'acclimatation linguistique des emprunts étrangers ou bien dans le domaine du droit marin français, pour indiquer la possibilité pour un bâtiment d'arborer un pavillon français. Ce type d'innovation terminologique, qui relève de l'attribution d'un signifié nouveau à un signifiant existant, implique non seulement des facteurs de changement géographiques et historiques (la transformation du sens correspond pour *francisation* au déplacement d'un côté à l'autre de l'Atlantique) ; mais aussi une dimension socio-culturelle fondamentale, où le sens fondamental de *francisation* — la gestion linguistique de l'altérité par rapport à une identité — change de référent selon le point de vue du locuteur qui à chaque fois s'en fait porteur. En particulier, ce cas de néonymie se contextualise dans le cadre d'une réappropriation de contenus linguistiques et symboliques hexagonaux au sein des politiques linguistiques au Québec. Ce cas d'étude se propose donc d'interroger le comportement lexical et textuel de ce terme dans le but de cerner les démarches culturelles à l'œuvre dans la reprise discursive des terminologies.

Les recherches en terminologie plus récentes montrent un gain d'intérêt important pour la dimension texto-discursive du terme. En abandonnant la vocation onomasiologique traditionnelle, la terminologie s'intéresse toujours plus à l'analyse de la textualisation des langues spéciales et à leurs comportements dans des contextes discursifs différents. Notre hypothèse est donc que c'est le comportement discursif d'un terme à déterminer sa force néonymique. En d'autres mots, nous ambitionnons de montrer que tout phénomène néologique en terminologie ne relève pas uniquement d'un acte démiurgique ayant sa source dans les pratiques de normalisation et d'aménagement linguistique ; au contraire, un rôle important est joué par les comportements spontanés ayant lieu dans le discours. Maintes études ont défendu cette hypothèse, en approchant différentes typologies textuelles et selon différentes démarches métrologiques. L'approche socioterminologique a par exemple adapté la boîte à

1. OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE, *Grand Dictionnaire Terminologique* (GDT), 2005, <<http://www.granddictionnaire.com/>> (dernière consultation : 31/10/2020).

2. Ou *néonyme*, comme Rondeau nous apprend à appeler l'invention linguistique en terminologie. Cf. G. RONDEAU, *Introduction à la terminologie*, Gaëtan Morin, Québec 1984.

outil sociolinguistique à l'étude des termes « sous l'angle de la diversité de leurs usages sociaux, ce qui englobe à la fois l'étude des conditions de circulation et d'appropriation des termes, envisagés comme des signes linguistiques, et non comme des étiquettes de concepts »³. La prise en compte de la traduction dans l'analyse de la textualisation des terminologies est proposée par Rachele Raus dans *La terminologie multilingue*⁴ alors que des considérations détaillées sur la porosité entre les approches terminologiques et lexicologiques ont été récemment répertoriées par John Humbley dans *La néologie terminologique*⁵.

Afin de comprendre la donnée culturelle et interculturelle qui sous-tend les usages discursifs et les habitudes comportementales d'un terme — dans notre cas : *francisation* — nous nous réclamons d'une approche inspirée par l'analyse des données textuelles (ADT). Nous interrogerons un corpus créé à partir du site web du Ministère de l'Immigration, de la Francisation et de l'Intégration (MIFI)⁶. Après avoir retracé les étapes fondamentales de l'aménagement de *francisation* dans de banques de données terminologiques européennes et américaines, françaises et québécoises, nous en analyserons le comportement sémantique en à travers les deux logiciel TXM⁷ et Hyperbase⁸, à travers en particulier les mesures statistiques de l'index de spécificité, de la concordance et de la poly-occurrence ou co-occurrence généralisée.

Une interprétation de ces résultats visera à comparer l'aménagement terminologique de *francisation* dans les sources terminographiques officielles et son usage en contexte discursif (en l'occurrence : le discours officiel du MIFI). En particulier, la dimension discursive du néonyme sera analysée en comparant les traits définitoires proposés par les ressources terminologiques (IATE, GDT etc.) et les traits ressortant, par contre, des cooccurrences dans le corpus choisi. L'analyse textométrique est donc

3. F. GAUDIN, *La socioterminologie*, in *La terminologie : nature et enjeux*, in « Langages », n. 157, 2005, pp. 80–92.

4. R. RAUS, *La terminologie multilingue. La traduction des termes de l'égalité H/F dans le discours international*, De Boeck, Bruxelles 2013.

5. J. HUMBLEY, *La néologie terminologique*, Lambert-Lucas, Limoges 2018.

6. Ministère de l'Immigration, de la Francisation et de l'Intégration du Gouvernement du Québec <<https://www.immigration-quebec.gouv.qc.ca/fr/accueil.html>>, (dernière consultation : 31/10/2020).

7. S. HEIDEN, J-P. MAGUÉ, B. PINCEMIN, TXM : *Une plateforme logicielle open-source pour la textométrie – conception et développement*, in “Proceedings of 10th International Conference on the Statistical Analysis of Textual Data – JADT 2010”, edited by S. Bolasco, vol. II, Edizioni Universitarie di Lettere Economia Diritto, Rome 2010, pp. 1021–1032, <<http://textometrie.ens-lyon.fr/spip.php?article60&lang=fr>> (dernière consultation : 31/10/2020).

8. Hyperbase 5.3 a été conçu par Étienne Brunet dans le cadre de l'UMR 6039 CNRS, Université de Nice, <<http://ancilla.unice.fr/default.html>> (dernière consultation : 31/10/2020).

l'instrument de questionnement de l'interaction des traits définitoires établis dans le cadre de l'aménagement terminologique officiel avec les traits des unités lexicales avec lesquels le terme entre en contact en contexte discursif.

2. De “francisation” à “francisation”, vers une définition de *défini-tion*

En terminologie, définition et *definiendum* entretiennent un rapport beaucoup plus étroit qu'en lexicologie : si le sens d'un mot est donné grâce à la connaissance du système linguistique, celui du terme est déterminé par son insertion dans un système de connaissance spécialisé. La définition, précise Humbley, « situe le terme dans son univers non pas linguistique, mais scientifique ou technique »⁹. En ce sens, le rôle fondamental de la définition en néologie terminologique a été souligné à maintes reprises. Si selon l'approche onomasiologique traditionnelle le terme n'est que l'expression linguistique d'un concept spécialisé, la création d'un nouveau terme est un processus cognitif subalterne par rapport à la création d'un nouveau concept. Des phénomènes de néonymie peuvent dès lors se produire seulement au niveau des modifications des définitions. Humbley précise à cet égard que « [s]i l'on se tient à la doctrine de l'indivisibilité du signe linguistique, on peut dire qu'un nouveau signifié attribué à un signifiant existant aboutit à l'émergence d'un nouveau signe. De ce point de vue, on peut affirmer que tout changement de définition relève de la néologie »¹⁰.

Tel est le cas de *francisation*, qui traverse une histoire complexe, liée de temps à autres aux manières et aux habitudes de la gestion culturelle de l'altérité et non dissemblable de l'histoire d'autres termes beaucoup plus embarrassants (appropriation, assimilation, etc.). Le verbe *franciser* est défini dans *Le Robert* comme acte de « donner une forme française à (une expression, un mot étranger) » (*ad vocem*). Il apparaît également comme terme du droit marin pour signifier « Faire passer un navire sous pavillon français » (*Le Robert, ad vocem*), où l'ancrage national de la signification saute aux yeux. Le verbe, attesté à partir du xvi^e siècle, serait une dérivation par jonction du suffixe causatif *-iser*, exprimant une transformation vers l'idée représentée par le radical, en l'occurrence *français*. De son côté, le TLFi, confirme la définition de *franciser* comme l'action de « Rendre français ; donner un caractère français à quelque

9. J. HUMBLEY, *op. cit.*, p. 53.

10. Ivi, p. 54.

chose ou à quelqu'un » (TLFI, *ad vocem*). Sous le domaine “linguistique” le terme *franciser* signifie : « a) Donner une forme française à un mot d'une langue étrangère » et « b) Introduire, répandre la langue française dans un lieu où elle n'est que peu ou pas parlée ; faire parler français » (TLFI, *ad vocem*).

Le substantif *francisation* est attesté dans le TLFI, où il apparaît encore comme terme du droit, avec ce signifié : « Action de reconnaître à un navire le droit de battre pavillon français et de jouir des privilèges qui y sont attachés ; résultat de cette action (TLFI, *ad vocem*) et, encore, sous le domaine “Linguistique” : « Action de franciser ; résultat de cette action » (*Ibidem*).

Dans les ressources terminologiques québécoises — et notamment dans le *Grand dictionnaire terminologique de la langue française* — le terme en question est présent dans trois fiches. Les deux premières relèvent encore du domaine “droit”, respectivement la fiche 17583068 (« Formalité conférant à un bâtiment de mer le droit de naviguer sous pavillon français », 1974), et la fiche 17572196 (« Droit d'arborer le pavillon national, avec les prérogatives et les servitudes qui en découlent », 1978). La troisième fiche présente une innovation intéressante du point de vue du domaine conceptuel de référence du terme, en l'occurrence “sociologie” (fiche 2068763, datée 2005). La définition qui est donnée est la suivante (v. Tab. 1) :

Tableau 1. Définition de Francisation.

<i>Terme</i>	Francisation
<i>Domaine</i>	Sociologie
<i>Auteur</i>	Office québécois de la langue française, 2005
<i>Définition</i>	Processus qui vise la généralisation de l'utilisation du français comme langue de travail et des communications dans l'Administration et les entreprises.
<i>Note</i>	Au Québec, la francisation vise une utilisation généralisée du français notamment dans les communications internes, dans les communications avec les clients et le public, dans l'affichage et la publicité, dans les documents de travail et les technologies de l'information.

Trois observations permettent de qualifier le terme représenté dans cette fiche comme néologisme terminologique : d'abord, d'un point de vue conceptuel, le changement de domaine. Le sens général de “rendre français”, commun aux définitions précédentes subit un processus de terminologisation en vertu duquel *francisation* entre à plein titre dans le cadre des éléments conceptuels constituant les sciences sociales. Deuxièmement, l'aspect du terme est affecté par une définition insistant sur une visée temporelle durative. Si la forme originale se présentait sous forme causative (rendre français quelqu'un ou quelque chose), la dimension durative est appuyée par l'introduction de “processus” dans la définition.

Une troisième observation concerne le radical faisant l’objet d’un processus de transformation : le radical *français* est en effet explicitement ancré dans l’espace–temps des politiques linguistiques québécoises, qui concernent dans ce cas certains secteurs du monde du travail et de la vie publique : l’entreprise, l’administration, la communication interne et externe des organisations, les technologies de l’information. Selon le *Sophiste* de Platon — ouvrage fondatrice d’une conception négative de l’altérité — toute image de l’Autre ressort d’une opération de négation linguistique et ontologique de l’objet du prédicat. L’altérité est en ce sens un concept flou qui dépend de l’identité qui se trouve à chaque fois coincé entre un “ne” et un “pas”. *Francisation* relève donc de la gestion de ce qui est “étranger” ; mais cette définition change à selon de l’espace–temps de l’énonciation, précisé, dans la fiche 2068763 du GDT, de manière très visible par le déictique “Au Québec” placé en position très saillante au début de la note.

On peut résumer ces observations dans le tableau suivant (v. Tab. 2), qui propose une synopse des traits définitoires présents dans la fiche analysée :

Tableau 2. Synopse des traits définitoires.

	TLFI	GDT
<i>domaine</i>	linguistique	sociologie
<i>aspect</i>	ponctuel	duratif
<i>identité</i>	France	Québec
<i>altérité</i>	non–France	?

La notion d’altérité évoquée — plus ou moins explicitement — dans les entrées des dictionnaires terminologiques est donc un élément relevant de l’énonciation et doit donc être étudiée en contexte. Dans les pages qui suivent, ces traits seront mis à l’épreuve d’une étude textométrique de la circulation discursive du terme en question, afin de démontrer l’hypothèse d’un apport fondamental des relations sémiques à l’œuvre dans le discours au processus de création terminologique. Dans une visée plus large, cette étude de cas souhaite montrer aussi que les phénomènes de néologie terminologique sont irréductibles aux opérations de normalisation linguistique institutionnelle ; au contraire ils s’identifient aux contextes d’usages et aux points de vue des différents énonciateurs impliqués.

3. L'analyse textuelle des données à l'intérieur du discours du MIFI

L'analyse des données textuelles — ADT (parfois appelée logométrie, textométrie ou lexicométrie selon qu'elle se concentre respectivement au discours, au texte et au lexique) se sert de mesures statistiques pour la recherche des informations textuelles à plusieurs niveaux : lexicale, syntaxique et sémantique. Nous avons construit notre corpus avec les mêmes principes de la linguistique de corpus (représentativité du phénomène à analyser ; balancement en termes de quantités des sous-corpus ; dimensions définies ; format numérique donc manipulable et consultable en manière automatique ; possibilité d'annotations) dans l'optique de considérer la communication en ligne Ministère de l'Immigration, de la Francisation et de l'Intégration (MIFI) comme un univers discursif autonome. Nous considérons le site web du MIFI tel un corpus spécifique (dédiée donc au discours politique de l'immigration), monolingue (même si une version en langue anglaise est également disponible) et synchronique (l'intervalle de temps considéré étant la durée du mandat du ministère). A ce propos nous avons observé que le discours du Ministère, et même son nom, change de manière profonde au rythme des différents mandats.

Ces réflexions justifient un approche *corpus-driven*, inductif, où les données sont étudiées non pas comme des expressions du système linguistique général mais en contexte, tels des éléments caractéristiques d'un discours spécifique.

3.1. Collecte et prétraitement

La collecte du corpus a été faite manuellement, ce qui nous a permis de créer un fichier de metadata pour pouvoir localiser l'information linguistique à l'intérieur de l'architecture du site. Nous observons, tout d'abord, que l'organisation hiérarchique du site (construite sur trois volets : “Immigrer au Québec”, “Favoriser l'intégration” et “Connaître le Ministère”) correspond à deux acteurs énonciatifs différents : le locuteur qui est le Ministère, et l'interlocuteur qui se divise à son tour en deux moments distincts : avant et après l'installation au Québec. À partir de cette intuition nous avons divisé le corpus en trois sous-corpora (respectivement indiqué comme partie I, partie II et partie III), dont chaque partie prend en considération un rôle discursif différent. L'intérêt de cette étude est donc aussi d'étudier les comportements de *francisation* non seulement en fonction d'un acteur discursif spécifique (le migrant) mais aussi, et surtout, à partir de la visée temporelle de son action concrète.

En général le principe qui nous a guidé pour la collecte du corpus a été celui de considérer le site web comme un ensemble de textes auto-référenciés. Cela signifie qu'on n'a pas pris en considération toute référence à des liens externes — comme par exemple une page d'actualités d'un quotidien qui relate une mesure politique et ses effets, ou le contenu des documents en pièce jointe. La raison réside dans le fait que ces contenus, quoique “portés” par le locuteur, ne relèvent pas directement de sa responsabilité énonciative. En concret, ce principe est réalisé en éliminant toutes références aux liens ou à l'action des usagers (*cliquez ici, haut de page*) ou à des informations à l'égard de la matérialité des documents en pièce jointe (les dimensions en *ko* des pdf ainsi que l'appellation “PDF”).

Une fois collecté, le corpus a été lemmatisé et annoté automatiquement selon les catégories du discours à travers le logiciel TreeTagger¹¹. Le corpus annoté automatiquement a été ensuite revu et corrigé manuellement. Une version non lemmatisée a été également gardée car, comme nous verrons plus tard, pour certains aspects le *case sensitive* du terme *francisation* est discriminant. Il en résulte un corpus ainsi formé (v. Tab. 3) :

Tableau 3. Composition du Corpus.

<i>token</i>	5.619
<i>occurrences</i>	263.555
<i>verbes</i>	67.748
<i>noms</i>	71.423

Le rapport entre les *tokens* et les occurrences nous renseigne sur le degré de richesse lexicale ou de répétitivité du vocabulaire utilisé dans le corpus en question. Dans ce cas, le niveau de diversité lexicale, attesté sur un taux de 0,021%, s'avère très peu élevé, ce qui peut être interprété comme un signe de technicité, la répétition étant une des caractéristiques du “style de spécialité” qui s'oppose aux usages synonymiques en terminologie¹². Quoique intuitive, cette observation peut nous confirmer que le cadre où s'insèrent les usages de *francisation* est celui d'un discours

11. À propos du logiciel TreeTagger : H. SCHMID, *Improvements in Part-of-Speech Tagging with an Application to German*, in “Proceedings of the ACL SIGDAT-Workshop”, Dublin, 1995; Id., *Probabilistic Part-of-Speech Tagging Using Decision Trees*, in “Proceedings of International Conference on New Methods in Language Processing”, Manchester 1994.

12. Voir à ce propos : H. NILSSON, *Être ennuyeux : nouvelles perspectives sur la responsabilité terminologique du traducteur dans le contexte européen*, in *Terminologie(s) et traduction. Les termes de l'environnement et l'environnement des termes*, sous la direction de A.-M. Velicu, S. Berninski, Peter Lang, Bern 2018, pp. 371–386.

technique, dont le domaine et les relations conceptuelles restent à être interrogées.

3.2. Les concordances

Une première analyse sera conduite à partir de l'étude de la concordance considérée comme la liste des occurrences d'un mot en contexte ; sa fonction principale est d'analyser le contexte immédiat d'un mot. Les premières étapes d'analyse du corpus visent à en définir des traits spécifiques mais aussi à faire sortir des indices susceptibles d'ouvrir des pistes intéressantes. Nous commençons donc avec une prime observation des occurrences du mot *francisation* où la présence de majuscule/minuscule nous indique une fréquence différenciée, selon les proportions suivantes (v. Tab. 4) :

Tableau 4. Index du terme « francisation / Francisation ».

Francisation	191
francisation	86
TOTAL	361
Pourcentage du corpus	0,13 %

Notre hypothèse est que les deux alternatives (*Francisation* vs *francisation*) ne se limitent pas à une différence typographique mais correspondent à un emploi différent du terme. *Francisation* à *f* majuscule se réfère à la formule institutionnelle désignant le Ministère de l'Immigration, Francisation et de l'Intégration. Cela est également démontré par l'écart de l'indice entre les premières trois cooccurrences avec toutes les autres (qui ont été supprimé de la table ci-dessous). Nous observons que même la proposition “de” — la plus présente de la langue française — apparaît avec un écart de 16 points en montrant que les termes *Francisation*, *Intégration*, *Immigration* et *Ministère* forment une expression polylexicale à visée dénomminative (v. Tab. 5).

Tableau 5. Cooccurrences du terme “Francisation”.

Occurrence	Pos	Fréquence	CoFréquence	Indice	DistMoy
<i>intégration</i>	NOM	495	159	136.9062	3.2767296
<i>immigration</i>	NOM	702	139	88.4071	3.4532375
<i>ministère</i>	NOM	605	123	79.7038	5.8455286
<i>de</i>	ART	10170	492	63	2,9
...

Les usages de *francisation* (*f* minuscule) sont résumés dans le tableau qui suit (v. Tab. 6). Le tableau est généré automatiquement en affichant

un contexte de 5 mots des deux côtés et a ensuite été filtré pour regrouper les occurrences du même type.

Tableau 6. Concordances du terme « francisation ».

Occ.	ContexteGauche	Pivot	ContexteDroite
1	améliorer la	francisation	
1	animateurs en	francisation	
2	séances de formation en	francisation	
3	cours en	francisation	
1	des organismes en	francisation	
1	employés œuvrant en	francisation	
7	en matière de	francisation	
5	Extranet	francisation	
1	faciliter l'insertion professionnelle et la	francisation	
1	Favoriser la	francisation	
3	intervenants de la	francisation	
1	l'offre de	francisation	
1	les activités de	francisation	
2	les démarches d'immigration, de	francisation	
1	leur parcours de	francisation	
1	Mérites en	francisation	
4	partenaires en	francisation	
2	personne responsable de la	francisation	
5	Plan d'action pour la	francisation	
3	Pour une	francisation	réussie
7	programme de	francisation	
1	Publications sur la	francisation	
1	Québec et à l'étranger, à la	francisation	
6	référentiel commun en	francisation	
2	réussite de votre	francisation	
13	service de	francisation	
3	soutien à la	francisation	
1	telles que l'hébergement, la	francisation	
1	parcours de	francisation	
3		francisation	en ligne
1		francisation	à temps complet
1		francisation	et des relations interculturelles

On peut remarquer que la plupart des concordances concerne des actions gouvernementales (*programmes de ; plan d'action ; service de*) réalisées à partir d'instruments informatiques (*en ligne, extranet*) ; des responsables ou du personnel qui travaille dans ce secteur (*animateurs ; employés ; partenaires*) et qui anime les différents aspects de la *francisation* ; celle-ci est donc envisagée comme processus construit sur différents étapes (*parcours de ; séances de formation en ; cours en ; démarches de*). Ces données nous confirment une vision complexe du concept du point de vue de la visée temporelle (à la fois durative, inchoative, répétitive) qui sous-tend sa configuration terminologique.

Cette visée durative (qui confirme la définition donnée en GDT) se construit autour d'une position centrale occupée par le simulacre de l'interlocuteur, à qui le déictique de seconde personne s'adresse dans l'occurrence : “réussite de votre francisation”. L'action de *franciser* se réfère donc à un être physique, doté de caractéristiques psychologiques et sociales. La dimension institutionnelle (dans le GDT la francisation concernait plutôt des organisations et des entreprises et leur communication) est donc réduite à l'échelle de l'individu, en l'occurrence le migrant. Le “processus” est donc visé dans l'optique d'une “réussite” personnelle qui doit être prise en charge par l'interlocuteur. En ce sens, la présence d'un “référentiel commun”, servant à décrire « sous forme de comportements observables les compétences en français des personnes immigrantes »¹³, indique aussi que le processus de *francisation* est un acte mesurable et quantifiable.

3.3. L'index de spécificité

La lecture des données issues de la concordance montrent une disproportion entre les trois éléments figurant dans la dénomination du ministère (*immigration* ; *francisation* ; *intégration*) : si l'action de *franciser* est focalisée sur la figure du migrant–interlocuteur, le rapport de *francisation* avec les deux étapes de l'installation du migrant au Québec (respectivement : immigration et intégration) restent à être interrogées. L'index de spécificité peut en ce sens montrer le poids relatif d'un terme à l'intérieur d'un sous–corpus. On entend par index de spécificité le calcul statistique des occurrences d'un mot, qui peut apparaître en surnombre ou en sous–effectif dans chacune des trois partitions de notre corpus (I. “Immigrer au Québec” ; II. “Favoriser l'intégration” ; III. “Connaître le Ministère”). Le tableau 7 résume donc les différences de poids du terme *francisation* relativement à sa présence dans chacun des trois volets : La section “Connaître le Ministère” montre donc une présence plus significative :

Tableau 7. Index de spécificité des la CQL “mots = ‘Francisation / francisation’”.

	freq. tot.	I.	index	II.	index	III.	index
francisation	86	56	-1,7	6	-1,4	24	4,9
Francisation	191	149	0,7	5	-6,9	37	3,4

13. MIFI – MINISTÈRE IMMIGRATION, FRANCISATION, INTÉGRATION, <<http://www.immigration-quebec.gouv.qc.ca/fr/partenaires/partenaires-francisation/referentiel-commun.html#rc01>> (dernière consultation : 27/11/2020).

Le même tableau peut être étudié selon une représentation en *barplot*. Le graphique (v. Fig. 1) nous permet d’observer l’absence et surtout la différence entre les deux variantes du terme en question (*Francisation* vs *francisation*) comme le signe d’usages sémantiques opposés : d’un côté sa présence dans la dénomination officielle du MIFI ; de l’autre, l’usage contextuel en discours.

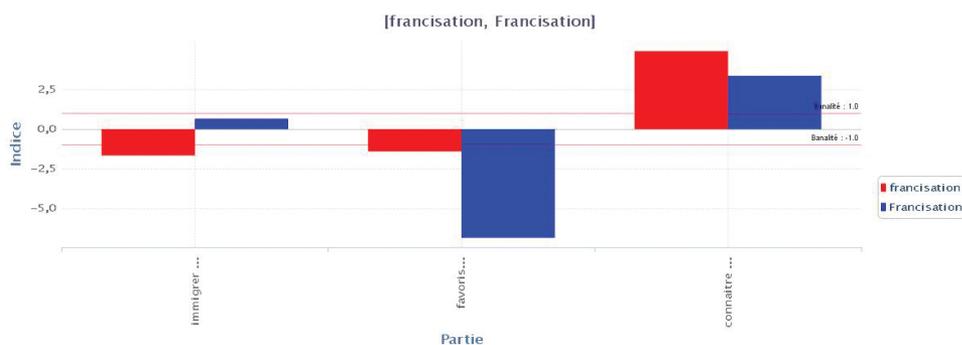


Figure 1. Index de spécificité des la cql “mots = ‘Francisation / francisation’”.

3.4. *La polyoccurrence ou occurrence généralisée*

L’instrument de la cooccurrence diffère de celui de la concordance dans la mesure où il nous permet de surpasser le concept d’immédiate proximité syntaxique pour rechercher des liens sémantiques. Dans le domaine de l’ADT, des nombreuses études ont été conduites sur la capacité des cooccurrences de définir des réseaux thématiques d’un ouvrage ou de caractériser un discours. Ils sont ici en jeu la polarisation d’un mot et sa capacité attractive dans le contexte. Toutefois comme ont observé Victorri & Fuchs « cette influence n’est pas à sens unique mais réciproque entre les différents éléments »¹⁴. Le contexte ici joue un rôle fondamental : Mayaffre définit la cooccurrence comme la forme minimale du contexte, mais c’est après Viprey que, grâce au concept de cooccurrence généralisée, le texte commence à être considéré (au niveau statistique) comme une entité réticulaire où tous les mots sont corrélés et où le sens de chacun est le résultat d’une négociation de sens mesurée par la distance, la fréquence et la cooccurrence. En fait la polyoccurrence (ou cooccurrence généralisée) considère les cooccurrences non pas comme

14. B. VICTORRI, C. FUCHS, *La polysémie : construction dynamique du sens*, Hermès, Paris 1996.

en couple deux par deux, mais comme une matrice où les mots sont en relation les uns avec les autres.

La cooccurrence est caractérisée par une valeur de fréquence du mot qui co-occure, puis par une valeur qui indique la co-fréquence des deux mots ensemble, l'indice de spécificité qui nous fournit une mesure relative à l'ensemble du corpus et nous permet de déchiffrer la "significativité" de l'association, et enfin la valeur de la distance entre les deux mots. Toutes ces mesures sont visibles grâce au graphique de la cooccurrence généralisée, qui nous permet d'avoir une vision générale du texte et, par la suite, de nous concentrer sur les relations du mots *francisation*. L'usage du graphique de cooccurrence répond donc à deux nécessités : d'un côté celle d'avoir une vision d'ensemble du tissu textuel du corpus (qui devient donc un texte unitaire) et de faire ainsi que cette représentation ne soit pas simplement graphique, mais qu'elle représente également des mesures statistiques. Nous avons cherché le mot *francisation* (en dehors de l'appellation du Ministère, donc en minuscule) tout en gardant la division sur les trois parties du corpus.

Cette dernière partie de notre analyse a été conduite à l'aide du logiciel Hyperbase, qui permet de calculer la poly-occurrence et de la représenter en graphique, à travers l'outil *thème*¹⁵. Nous observons les polarisations réticulaires du terme *francisation*, en filtrant les résultats dans le but d'obtenir comme champ d'application l'ensemble des parties du discours (substantif, adjectif, verbe, adverbe).

Le graphique 2 (v. Fig. 2), qui se réfère au sous-corpus 1 ("Immigrer au Québec"), montre que le terme *application* est associés à d'autres unités, nécessaires pour son identification. Le terme avec un impact majeur est *référentiel*, à son tour associé à *commun*. Comme précisé plus haut, les aspects graphiques représentent des mesures mathématiques : en ce cas l'épaisseur de la ligne se réfère à l'index de cooccurrence. Cela veut dire, dans le graphique suivant, que *référentiel commun* apparaît presque uniquement à coté de *francisation* :

15. Voir E. BRUNET, *HYPERBASE© Logiciel hypertexte pour le traitement documentaire et statistique des corpus textuels, Manuel de référence*, janvier 2011, <<http://hyperbase.unice.fr/hyperbase/doc/manuel.pdf>> (dernière consultation : 31/10/2020).

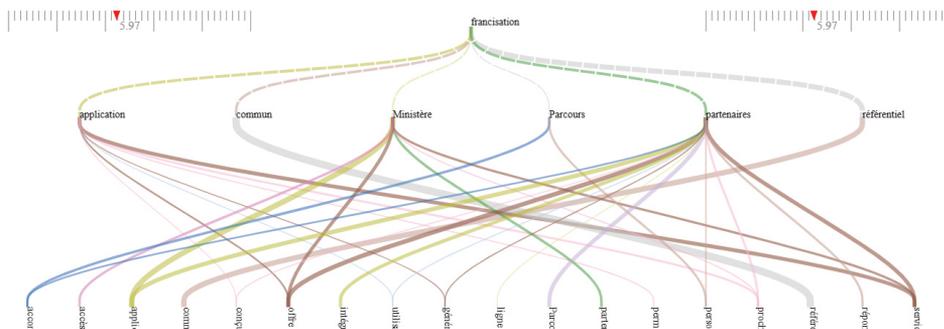


Figure 2. Poly-occurrence : Partie I : Immigrer au Québec.

Dans la partie II le terme *francisation* montre une cooccurrence élevée avec les expressions “personnes”, “immigrantes”, “contexte” et “autoformation”. Nous pouvons observer comment les relations qui relient ces termes sont bidirectionnelles (v. Fig. 3). Cela veut dire par exemple que *autoformation* et *contexte* paraissent ensemble de manière assez régulière, et cela nous indique un probable “ensemble sémantique” typique du discours. Les mots qui paraissent ensemble formant des expressions idiomatiques s’appellent “collocates” et font partie du système linguistique en général¹⁶. Les concordances bidirectionnelles qui présentent un haut taux indice comme dans notre cas, peuvent être considérées des “expressions idiomatiques” du discours particulier que nous sommes en train d’analyser. Notre hypothèse est que “autoformation en contexte” ou “contexte d’autoformation” représentent des concepts typiques du discours du MIFI accessibles à partir du terme *francisation*.

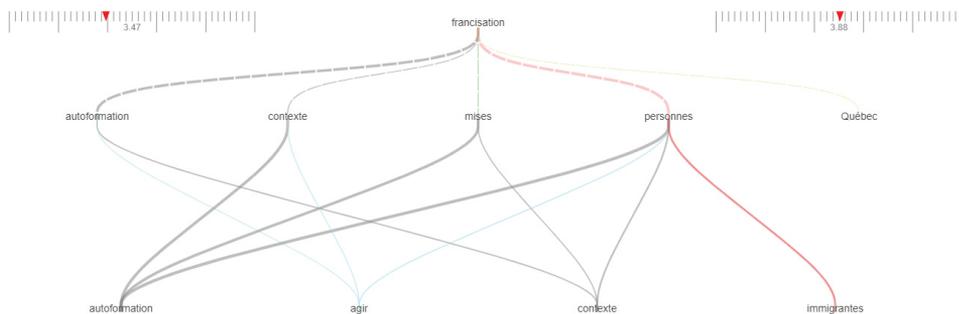


Figure 3. Poly-occurrence : Partie II : Favoriser l’intégration.

16. D. ANCTIL, O. TREMBLAY, *Les collocations : des combinaisons de mots privilégiées*, in « Correspondance » vol. XXI, n. 3, CDM, 2016.

En ce qui concerne la partie III, on remarque que *francisation*, *intégration* et *personnes immigrantes* composent une aire de polarité importante (Graphe 5). Ici aussi on découvre que “personnes” et “immigrantes” forment une expression quasi figée du discours, indiquant un acteur fondamental de la francisation. Également le mot “intégration” paraît comme un autre pôle du réseau sémantique que les poly-occurrences contribuent à mettre en évidence (v. Fig. 4) :

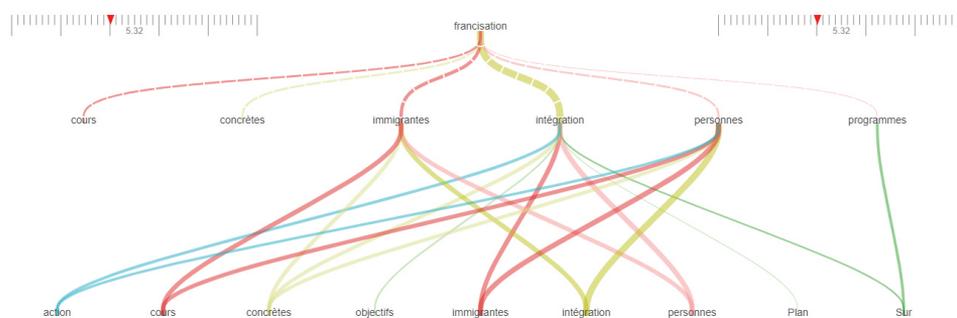


Figure 4. Poly-occurrence : Partie III : Connaître le Ministère.

4. Conclusion

À partir des analyses conduites, le terme *francisation* se montre comme une action communautaire, publique, réticulaire, visant la personne plutôt que l'organisation, l'individu plutôt que le groupe. Elle possède aussi une dimension aspectuelle qui la rend un élément de transition parmi les deux étapes de l'installation du migrant au Québec, soit l'immigration et l'intégration. Bien qu'elle ne corresponde précisément ni à l'un ni à l'autre de ces deux processus, la *francisation* semble se configurer comme un état intermédiaire, fonctionnel à la métamorphose de l'interlocuteur du statut d'*immigrant* à celui d'*intégré*. A travers l'analyse de la concordance, de l'index de spécificité et des polyoccurrences du terme en question, on a donc pu constater toute une série de comportements lexicaux qui échappent à la normalisation terminologique des néonymes et qui obligent à une vision plus large des phénomènes terminologiques : en ce sens il est impératif d'inclure des considérations d'ordre culturel, politique et social. Bien que partiels, les résultats invitent donc à un dialogue entre approches prescriptive et descriptives en terminologie, qui expriment, sous la forme d'un jeu de miroirs, l'apport indispensable de la visée textuelle et discursive dans les études en néologie.

Références bibliographiques

- ANCTIL D., TREMBLAY O., *Les collocations : des combinaisons de mots privilégiées*, in « Correspondance » vol. XXI, n. 3, CCDMD, 2016.
- GAUDIN F., *La socioterminologie*, in *La terminologie : nature et enjeux*, in « Langages », n. 157, 2005, pp. 80–92.
- HEIDEN S., MAGUÉ J-P., PINCEMIN B., *TXM : Une plateforme logicielle open-source pour la textométrie — conception et développement*, in “Proceedings of 10th International Conference on the Statistical Analysis of Textual Data – JADT 2010”, edited by S. Bolasco, vol. II, Edizioni Universitarie di Lettere Economia Diritto, Rome 2010, pp. 1021–1032, <<http://textometrie.ens-lyon.fr/spip.php?article60&lang=fr>>.
- HUMBLEY J., *La néologie terminologique*, Lambert–Lucas, Limoges 2018.
- NILSSON H., *Être ennuyeux : nouvelles perspectives sur la responsabilité terminologique du traducteur dans le contexte européen*, in *Terminologie(s) et traduction. Les termes de l’environnement et l’environnement des termes*, sous la direction de A.–M. Velicu, S. Berninski, Peter Lang, Bern 2018, pp. 371–386.
- OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE, *Grand Dictionnaire Terminologique (GDT)*, 2005, <<http://www.granddictionnaire.com/>>.
- RAUS R., *La terminologie multilingue. La traduction des termes de l’égalité H/F dans le discours international*, De Boeck, Bruxelles 2013.
- RONDEAU G., *Introduction à la terminologie*, Gaëtan Morin, Québec 1984.
- SCHMID H., *Improvements in Part-of-Speech Tagging with an Application to German*, in “Proceedings of the ACL SIGDAT–Workshop”, Dublin, 1995.
- , *Probabilistic Part-of-Speech Tagging Using Decision Trees*, in “Proceedings of International Conference on New Methods in Language Processing”, Manchester 1994.
- VICTORRI B., C. FUCHS, *La polysémie : construction dynamique du sens*, Hermès, Paris 1996.