



LANGUES SOCIÉTÉS



La variation dans la langue standard

Actes du colloque tenu les 13 et 14 mai 2002
à l'Université Laval
dans le cadre du 70^e Congrès de l'ACFAS

Sous la direction de Pierre Bouchard

Office québécois de la langue française

**NOMENCLATURE D'UN PROCHAIN DICTIONNAIRE
DU FRANÇAIS STANDARD EN USAGE AU QUÉBEC :
PROBLÉMATIQUE ET BILAN DES TRAVAUX**

Hélène Cajolet-Laganière

Université de Sherbrooke

Pierre Martel

Université de Sherbrooke

Michel Théoret

Université de Sherbrooke

Jean-Claude Boulanger

Université Laval

Louis Mercier

Université de Sherbrooke

Introduction

Toute langue porte l'empreinte de la communauté qui la parle et contribue à façonner son identité. Elle sert notamment de support à la représentation du monde et de l'univers culturel que celle-ci a élaborée. Elle apparaît comme un facteur essentiel au maintien et au développement de son identité culturelle. Dans ce contexte, la langue d'usage public d'une communauté linguistique donnée devient le véhicule de sa culture (littéraire, sociale, politique, scientifique, technique, administrative, etc.) et le moyen de promotion par excellence de son identité. Cette production langagière publique, associée par consensus à l'usage valorisé constitue pour cette communauté un bien collectif qu'il lui faut protéger, exploiter, diffuser et inscrire dans la réalité contemporaine et son développement technologique. Aussi importe-t-il de recenser les mots, sens, expressions et autres éléments qui caractérisent la langue standard en usage au Québec. Les textes sélectionnés comme lieux de ce recensement doivent rendre compte des particularités de l'espace géographique, physique, social, économique du Québec afin de bien refléter la langue réellement en usage. De fait, l'environnement naturel québécois présente une faune et une flore qui diffèrent de celles de l'Europe; le Québec possède des institutions politiques, sociales, culturelles, scolaires et autres conformes à ses besoins et à ses aspirations. Il a en outre développé une expertise dans des domaines diversifiés. Des milliers de

mots, de sens et d'expressions traduisent ces spécificités, qui ne se logent pas uniquement dans la langue orale ou familière, loin de là : ils constituent même une composante importante de la langue standard contemporaine, dont quiconque a besoin dans le cadre de sa vie sociale et professionnelle.

Comme vous le savez, un dictionnaire du français standard en usage au Québec est en cours de réalisation à l'Université de Sherbrooke, projet auquel sont associés plusieurs chercheurs du Centre d'analyse et de traitement informatique du français québécois (CATIFQ). Le présent article a pour but de décrire les étapes visant à constituer la nomenclature de ce futur dictionnaire, c'est-à-dire la liste des mots à retenir pour refléter la langue générale ainsi que les usages associés aux divers domaines d'activités socioéconomiques propres aux contextes québécois et nord-américain. Un mot d'abord sur les textes choisis comme lieu de recensement. Le mot *standard* employé dans la perspective adoptée renvoie donc à une catégorie de textes dont le niveau de langue est généralement soutenu (textes littéraires reconnus, documents administratifs et législatifs soignés, manuels scolaires dont plusieurs ont été primés, productions scientifiques, livres et essais de tous genres publiés par des intellectuels, textes journalistiques, etc.). Ces textes ont été soigneusement sélectionnés, stockés dans la Banque de données textuelles de Sherbrooke (BDTS) et indexés aux fins d'élaboration de cette nomenclature. Quant au concept de « variation », lequel constitue le thème propre du colloque où vient s'inscrire cette contribution, il faut retenir qu'il recouvre différents types de variation. Selon Moreau (1997), par exemple, il faut distinguer au moins quatre sortes de variation : la variation diachronique (dans le temps), la variation diatopique (dans l'espace), la variation diastratique (selon les classes ou groupes sociaux) et la variation diaphasique (selon les registres de langue). Les exemples que nous donnons dans la deuxième partie de ce texte relèvent principalement des variations diatopiques et diaphasiques.

De fait, la première tâche pratique inhérente à la rédaction d'un dictionnaire usuel est sans contredit relative à la nomenclature, c'est-à-dire à la sélection des mots à retenir, ou au contraire, à l'exclusion de l'ouvrage projeté. Cette sélection est de première importance et doit être effectuée avec la plus grande rigueur, en tenant compte de l'orientation de l'ouvrage, de l'objectif visé et du public ciblé. Pour les utilisateurs, trouver un mot à la nomenclature d'un dictionnaire signifie que le mot « existe », qu'il a un statut dans la langue. Aussi certains d'entre eux peuvent-ils être choqués par la présence de mots qui, pour diverses raisons, ne leur

semblent pas devoir s'y trouver (au Québec, par exemple, la présence de sacres, d'anglicismes, de mots trop familiers ou vulgaires). À l'opposé, l'absence du mot cherché peut être aussi interprétée négativement par les lecteurs : il peut penser que ce mot n'existe pas, qu'il est sans doute incorrect, critiqué, ou encore, qu'il est trop rare ou trop spécialisé pour être enregistré. Par conséquent, l'élaboration de la nomenclature pour un lexicographe représente une première opération à effectuer, parmi les plus délicates et les plus importantes. En outre, quelle que soit sa taille, un dictionnaire ne contient qu'une infime partie de la totalité du lexique d'une langue. Un choix judicieux des mots à inclure dans la nomenclature s'impose donc. Parmi la masse quasi infinie des vocables d'une langue, divers types ou catégories de mots peuvent être distingués. Cela concerne les grands types d'éléments lexicaux, comme les mots simples, les mots composés et les groupes complexes plus ou moins figés constituant les mots courants de la langue générale, les éléments de formation, etc. Sans oublier d'autres catégories de mots, comme les archaïsmes et les néologismes, les mots limités à certaines grandes aires linguistiques (belgicismes, francismes, québécoïsmes), les mots régionaux à l'intérieur d'une aire, les emprunts aux langues étrangères, dont les emprunts critiqués à l'anglais, les termes recommandés par des organismes officiels ou largement banalisés.

Devant cette tâche à accomplir, le lexicographe ne peut confier sa sélection au hasard. Ses choix seront nécessairement guidés par le profil du public auquel il destine son ouvrage. Selon qu'il rédige un dictionnaire scolaire pour jeunes élèves, un dictionnaire usuel pour grand public ou un dictionnaire plus spécialisé, ses critères de choix seront sans aucun doute différents. Ses choix seront aussi établis en fonction de l'orientation et des objectifs de l'ouvrage (dictionnaire normatif, descriptif, de difficultés, ou autres).

Dans le présent cas, le but visé par l'ouvrage est de décrire le français en usage au Québec en donnant la priorité aux emplois relevant du registre standard et en partant de ce standard pour hiérarchiser les autres usages retenus (oraux, familiers, critiqués, etc.), de sorte que la description proposée puisse servir de guide linguistique à la communauté québécoise. Cette description doit évidemment tenir compte des particularismes lexicaux des principales communautés francophones de manière à assurer l'intercommunication dans la francophonie. Cette hiérarchisation des usages autour d'un standard, légitimé et valorisé, suppose en même temps que soient reconnues et acceptées les spécificités linguistiques du Québec. À ce sujet, on peut rappeler le témoignage de Daniel Raunet, journaliste à

Radio-Canada, qui, lors d'une communication présentée dans le cadre des États généraux sur la situation et l'avenir de la langue française au Québec, a donné quantité d'exemples de situations langagières où les journalistes sont aux prises avec des questions de norme, des cas non résolus, des questions qui restent sans réponse. À partir de ce constat, il affirmait que : « Le vrai problème, c'est le non-dit, le silence, le tabou qui entoure nos choix dans un registre particulier : la langue soutenue. »

On peut rappeler également que, dès 1983, Josette Rey-Debove, co-directrice des dictionnaires *Le Robert*, affirmait clairement l'importance de décrire le français québécois comme un tout cohérent :

« [...] la seule façon de réaliser un bon dictionnaire québécois consisterait à partir de zéro et à décrire le français québécois comme s'il s'agissait de la seule langue au monde [...]. »

Allant dans le sens de ce commentaire de Rey-Debove, nous avons opté pour l'élaboration d'une nomenclature originale suivant trois étapes. Dans un premier temps, nous avons extrait une première liste de mots de la Banque de données textuelles de Sherbrooke (BDTS). La BDTS, représentative du français en usage au Québec (voir le schéma présenté ci-après), permet l'accès à des textes qui reflètent différents registres de langue, notamment le niveau standard, qui est actuellement en usage au Québec dans les principaux domaines d'activités. De fait, les documents stockés fournissent les mots de la langue générale et le vocabulaire de base de très nombreux domaines spécialisés. En outre, on y retrouve certains textes de nature didactique qui sont utilisés dans la formation des jeunes et qui, à ce titre, sont valorisés et jouent un rôle important d'exemplarité. Dans un deuxième temps, cette première liste de mots a été validée à la lumière des nomenclatures proposées dans les principaux dictionnaires usuels français et québécois. Enfin, le dépouillement d'ouvrages plus spécialisés a permis une dernière validation de la nomenclature. La première partie de l'exposé qui suit fait état des résultats du travail effectué quant au choix de la nomenclature. La deuxième partie présentera un certain nombre de cas de variation lexicale dans la langue standard et leur traitement au regard de la nomenclature.

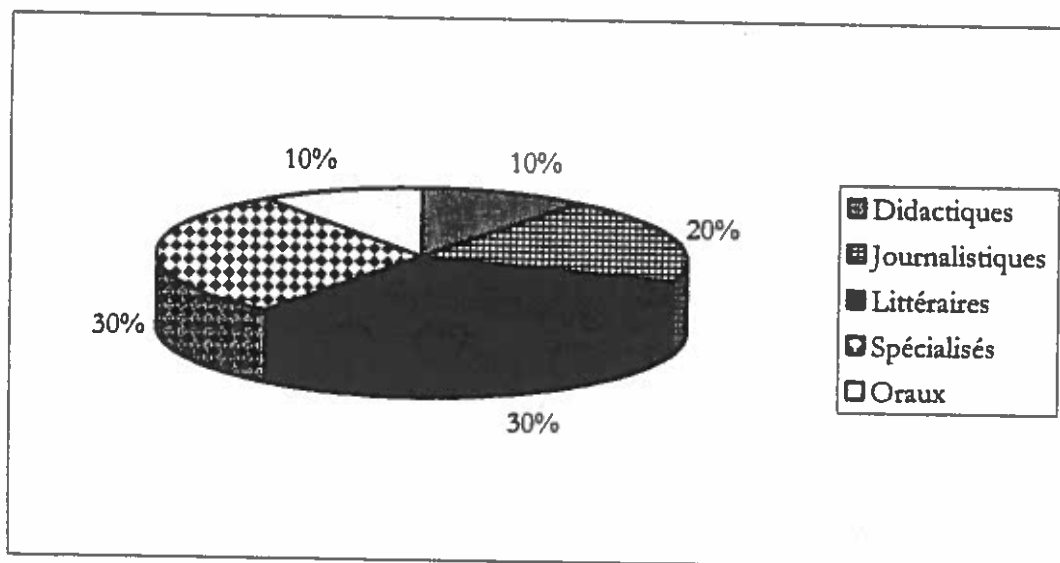
1. Étapes visant à la sélection de la nomenclature

Dans sa version actuelle, la BDTS est un corpus informatisé et indexé de plus de 37 millions de mots; elle est constituée de textes représentatifs des différents usages du français québécois. Il ne s'agit pas d'une représen-

tativité au sens statistique du terme : la notion de « représentativité » de la BDTS doit plutôt être associée à la variété des textes qui la composent et qui reflètent la langue générale (orale et écrite, dans différentes situations de communication) de même que les langues littéraire, journalistique et relativement plus spécialisée. La ventilation des pourcentages associés aux diverses catégories de textes correspond à l'objectif visé : donner accès aux différents registres de langue, notamment le niveau standard, actuellement en usage au Québec. Pour la notion de « français québécois », nous faisons référence aux textes que nous stockons aux fins de diverses analyses et qui ont été rédigés dans l'espace géographique du Québec. Nous ne faisons nullement allusion à quelque caractère différentiel de ces textes par rapport à des textes rédigés en France, par exemple. Nous considérons ces textes pour eux-mêmes et dans leur totalité.

La BDTS est constamment gardée à jour grâce à l'entrée de textes reflétant l'actualité socioculturelle, économique, politique et autres. À cette banque s'ajoutent enfin les cédéroms Actualités/Québec, Repère, Internet, etc., ainsi que le Réseau des corpus lexicaux québécois (dont une version réduite de la BDTS fait partie), géré par le Secrétariat à la politique linguistique du Québec; c'est la partie disponible en périphérie du corpus.

Figure 1



Description de la BDTS

Corpus informatisé de plus de 37 M de mots.

Plus de 10 000 textes et ouvrages représentatifs du français en usage au Québec

Domaines enrichis : faune et flore, politique, relations du travail, arts et culture, nord et hiver, acériculture, éducation, etc.

Corpus en périphérie : Actualités/Québec, Repère et Internet; Réseau des corpus lexicaux québécois; autres banques de données textuelles.

– Extraction de la nomenclature de la BDTS

L'élaboration d'une nomenclature originale extraite de la BDTS concerne toute une série d'éléments : les noms communs de la langue générale, les termes spécifiques à divers domaines, les emprunts à diverses langues, notamment à l'anglais, les abréviations, sigles et acronymes, les noms propres de personnes, de lieux, de marques et autres dénominations diverses. L'indexation des mots de la BDTS a fourni quelque 200 000 formes. Après lemmatisation de ces formes, nous avons obtenu quelque 57 000 vocables correspondant à des noms communs différents et quelque 40 000 vocables relevant de la catégorie des noms propres. De façon générale, les noms propres ne sont pas intégrés à la nomenclature des dictionnaires de langue, à l'exception d'une sélection de gentils, de marques ou noms déposés, de sigles et d'acronymes d'usage courant et de quelques termes du type Noël, Pâques, Œdipe, etc. Aussi, les données mentionnées dans la suite de ce texte ne concerneront-elles que les noms communs attestés dans la BDTS.

– Validation de la nomenclature extraite de la BDTS à la lumière d'autres nomenclatures

Dans le but d'élaborer la nomenclature la plus représentative possible du français en usage au Québec, nous avons validé la liste des vocables extraits de la BDTS en la comparant avec la nomenclature d'une série de douze dictionnaires relativement récents, usuels et scolaires, tant français que québécois. Il s'agit, d'une part, de sept dictionnaires scolaires pour les jeunes (de 6000 à 30 000 entrées), dont les éléments ont été réunis dans une seule nomenclature cumulative : *Larousse des débutants*, *Robert Benjamin*, *Mon premier dictionnaire français illustré*, *Dictionnaire CEC Jeunesse*, *Robert Junior illustré* (version nord-américaine), *Dictionnaire CEC Intermédiaire* (version électronique), *Dictionnaire HRW*. Et d'autre part, de quatre dictionnaires usuels et d'un dictionnaire de difficultés (de 28 000 à environ 60 000 entrées) : *Dictionnaire québécois d'aujourd'hui*, *Dictionnaire du français plus*, *Petit Robert électronique*, *Petit Larousse illustré*, *Multidictionnaire*.

Le total des vocables différents traités dans ces douze dictionnaires s'élève à environ 72 000; la BDTS, pour sa part, a fourni quelque 57 000 noms communs, dont quelque 41 000 sont enregistrés dans la nomenclature des dictionnaires comparés.

- Première observation quant à la variation dans le traitement lexicographique

De l'analyse des nomenclatures comparées il ressort une variation dans la représentation de la langue à décrire, notamment dans le choix des mots et l'étendue de ces nomenclatures. À titre indicatif, quelque 20 000 mots seulement sont communs aux six nomenclatures examinées. En outre, quelque 12 000 mots n'apparaissent que dans un seul dictionnaire. Le PLI contient quelque 3800 mots spécifiques; le PRÉ, environ 3700; le DFP, environ 2300; le DQA, environ 900; le MULTI, environ 500; la nomenclature fusionnée des sept dictionnaires scolaires, environ, 450. La BDTS, quant à elle, renferme quelque 16 000 noms communs qui ne sont enregistrés dans aucun de ces dictionnaires.

Cette étape de validation par comparaison avec la nomenclature des principaux dictionnaires disponibles sur le marché québécois a conduit à l'enrichissement de la nomenclature initialement extraite de la BDTS. Elle a été suivie par une seconde étape de validation. En effet, nous avons, dans un troisième temps, validé la nomenclature de la BDTS avec celles de répertoires spécialisés, notamment dans les domaines susceptibles de faire référence à bon nombre de spécificités québécoises (faune et flore, relations du travail, éducation, hiver, nord et acériculture, politique, arts et culture, avis de recommandation de l'OLF, emprunts critiqués à l'anglais, etc.) ainsi que dans certains domaines de pointe, comme l'informatique.

Nous avons donc examiné ces listes au regard de certains critères d'inclusion et d'exclusion. Ces critères sont de trois ordres :

- fréquence des mots dans la BDTS et dispersion dans les différents types de discours stockés dans la BDTS (littéraire, journalistique, didactique, spécialisé et oral);
- présence et sanction des mots dans les dictionnaires sélectionnés, avec prise en compte de leur diffusion dans la francophonie, des avis de recommandation ou d'officialisation dont ils ont fait l'objet, des critiques formulées à leur égard dans les ouvrages correctifs qui jouissent d'une certaine autorité;

- critères plus larges enfin, touchant la politique éditoriale : orientation et objectifs de l'ouvrage, public ciblé, etc.

– **Autre variation dans la représentation de la langue :**

la diversité dans l'adressage des entrées des dictionnaires

Lors de la comparaison de la nomenclature de la BDTS avec celles des différents dictionnaires sélectionnés, nous avons observé une grande diversité dans l'adressage des entrées. Ces cas de variation sont de divers ordres :

- **entrées se terminant ou non par un signe de ponctuation (! ?) :**

heu!, hop!, ohé!, ouste!, ah! ou *heu, hop, ohé, ouste, ah;*
n'est-ce pas? ou *n'est-ce pas*

- **mots avec ou sans majuscules accentuées :**

CÉE, ÉNAP ou CEE, ENAP

- **présence ou non d'espaces entre les lettres des sigles ou abréviations**

Ph. D ou *Ph.D;*

V. H. S. ou *V.H.S.*

B. A. ou *B.A.*

R. S. V. P. ou *R.S.V.P.*

- **emploi ou non de majuscules et/ou de points pour les sigles et abréviations :**

am ou *A.M.*

no ou *No*

Adn ou *A.D.N.* ; *ADN* ou *A. D. N.*

- **emploi ou non de lettres majuscules initiales pour les noms de fêtes ou d'autres séries de mots :**

Pentecôte ou *pentecôte*

Assomption ou *assomption*

Halloween ou *halloween*

Noël ou *noël*

Bureautique, Martini, État ou *bureautique, martini, état*, etc.

Bible ou *bible*

Coran ou *coran*

Évangile ou *évangile*

– variantes orthographiques :

alcootest ou *alcotest*
maffia ou *mafia*
marijuana ou *marihuana*
acuponcture ou *acupuncture*
cédrière ou *cèdrière*
malentendant ou *mal entendant*
minicassette ou *mini-cassette*
base-ball ou *baseball*
millefeuille ou *mille-feuille*
rhinopharynx ou *rhino-pharynx*
microorganisme ou *micro-organisme*

– variantes en genre :

haillons ou *haillon*
broussailles ou *broussaille*

– majuscules ou minuscules initiales pour les marques ou noms déposés :

Caméscope ou *caméscope*
Interphone ou *interphone*
Minitel ou *minitel*
Scotch ou *scotch*

Pour ces cas de variation dans l'adressage des entrées (quelques milliers ont été notés), il importe également de prendre des décisions au regard de certains critères, notamment un alignement sur le code typographique reconnu, diffusé et bien implanté au Québec (*baseball* et non *base-ball*, par exemple); un alignement sur certains éléments de modifications orthographiques proposés en 1990 (par exemple, la soudure de bon nombre de mots, notamment ceux avec préfixes);

2. Exemples de cas de variation lexicale

Dans cette deuxième partie, de manière à répondre aux objectifs visés par le thème du colloque, soit la variation dans la langue standard et les ouvrages de référence, nous présentons un certain nombre d'exemples de cas de variation lexicale relevés lors de l'établissement de notre nomenclature. Nous les avons traités au regard des critères mentionnés ci-dessus, soit fréquence et dispersion dans la BDTS, sanction donnée à ces mots par les différents dictionnaires et enfin, traitement lexicographique de ces mots pour ce qui est de la nomenclature, objet du présent exposé. Les cas de variation ont été groupés selon la typologie suivante.

Il va de soi que ces exemples sont très limités et qu'ils sont donnés à titre d'illustration : ils sont en effet représentatifs d'un grand nombre de cas similaires. Nous les avons groupés comme suit :

- Concurrence entre vocables employés surtout au Québec et vocables employés surtout ailleurs dans la francophonie;
- Concurrence entre vocables français et emprunts critiqués à l'anglais;
- Concurrence interne au Québec entre vocables français et emprunts à l'anglais;
- Concurrence non réelle entre emprunts à l'anglais et vocables recommandés, mais non utilisés (du moins au Québec);
- Concurrence entre vocables de la langue générale et vocables des langues de spécialité.

Type 1 : Concurrence entre vocables employés surtout au Québec et vocables employés surtout ailleurs dans la francophonie

Cas type	Employé surtout au Québec	Employé surtout ailleurs dans la francophonie
« disque de caoutchouc utilisé au hockey »	<i>rondelle</i>	<i>palet</i>
Fréquences dans la BDTS (37 millions)	(f = 145)	(f = 0)
Dispersion : catégories de textes	Tous les types	
Répertorié	Multi (Q), DQA	Multi (sens général), PRÉ, DQA (F), DFP
Critiqué	PRÉ	
Non répertorié		

Comme le montre le tableau ci-dessus, le mot *rondelle* a une fréquence de 145 dans la BDTS, alors que *palet* affiche la fréquence 0. *Rondelle* est en outre très dispersé et employé dans tous les genres de textes. *Rondelle* est de plus répertorié dans les dictionnaires québécois et n'est critiqué par aucun d'eux. Aussi, compte tenu de nos critères, nous retenons *rondelle* et *palet* dans la nomenclature du dictionnaire. *Rondelle* aura une entrée avec article, alors que *palet* aura droit à une entrée associée à un renvoi à *rondelle*, du moins pour le sens relatif au hockey. C'est dans l'article *rondelle* que sera présenté le maximum d'informations fonctionnelles associées au concept de « rondelle », notamment les renseignements d'ordre sémantique.

Type 1 : Concurrence entre vocables employés surtout au Québec et vocables employés surtout ailleurs dans la francophonie

Cas type	Employé surtout au Québec	Employé surtout ailleurs dans la francophonie
« spécialiste des interventions d'urgence »	<i>urgentologue</i>	<i>urgentiste</i>
Fréquences dans la BDTS (37 millions)	(f = 10)	(f = 0)
Dispersion : catégories de textes	Textes surtout reliés au domaine médical	
Répertorié	PLI (Q), PLI, DQA	PLI, PRÉ, ENF
Critiqué	DFP, Multi, PRÉ	Multi, DQA, DFP
Non répertorié		

Il s'agit d'un cas relativement similaire à *rondelle/palet*. *Urgentologue* est employé au Québec (fréquence 10 dans la BDTS), alors qu'*urgentiste* est absent de la Banque. *Urgentologue* est également répertorié dans les dictionnaires français (PLI) et québécois. Il n'est critiqué dans aucun dictionnaire. Pour sa part, même s'il est absent de la BDTS, *urgentiste* est répertorié dans les dictionnaires français et québécois. Les deux mots seront intégrés à la nomenclature du dictionnaire. Conformément à nos critères, *urgentologue* aura le statut d'entrée principale, alors que le mot *urgentiste* l'accompagnera comme variante morphologique. Chaque mot fera l'objet d'une remarque de nature diatopique.

Type 2 : Concurrence entre vocables français et emprunts critiqués à l'anglais

Cas type	Vocabulaire français	Emprunt critiqué à l'anglais
« personne qui travaille à la pige »	<i>pigiste</i>	<i>free lance</i>
Fréquences dans la BDTS (37 millions)	(f = 97)	(f = 3)
Dispersion : catégories de textes	Tous les types	
Répertorié	PRÉ, Multi, PLI, ENF, DQA, DFP	PLI
Critiqué		Multi (anglic.), PRÉ (anglic.)
Non répertorié		DQA, DFP

Cet exemple illustre un cas différent. Le mot largement en usage au Québec, et ce, dans tous les types de discours, est *pigiste* (fréquence 97 dans la BDTS, comparativement à la fréquence 3 pour *free lance*). En outre, les dictionnaires français et québécois recensent *pigiste* sans jamais le critiquer. À l'opposé, *free lance* est peu utilisé au Québec; il est de plus critiqué dans des ouvrages français et québécois. Compte tenu de son emploi courant en France, il fera l'objet d'une entrée renvoi à *pigiste*, où son statut sera précisé.

Type 2 : Concurrence entre vocables français et emprunts critiqués à l'anglais

Cas type	Vocabulaire français	Emprunt critiqué à l'anglais
« jeu de patience... »	<i>casse-tête</i>	<i>puzzle</i>
Fréquences dans la BDTS (37 millions)	(f = 115)	(f = 42)
Dispersion : catégories de textes	Tous les types	Surtout journalistiques et littéraires
Répertorié	PRÉ (C), PLI (Q), Multi (Q), ENF, DQA, DFP	PRÉ, PLI, ENF, DFP
Critiqué		Multi (anglic.)
Non répertorié		

Cet exemple illustre le cas de deux mots en usage dans les textes québécois. *Casse-tête* est nettement plus fréquent (fréquence 115 dans la BDTS) et plus largement dispersé (tous les types de textes) que *puzzle* (fréquence 42 dans la BDTS) et limité à certains types de discours (surtout journalistiques et littéraires). Les deux mots sont répertoriés dans les dictionnaires français et québécois; *casse-tête* fait l'objet de marques topolectales dans plusieurs ouvrages, mais n'est jamais critiqué. *Puzzle*, quant à lui, est critiqué. Aussi, conformément à nos critères, *casse-tête* fera l'objet d'un article; *puzzle* aura une entrée renvoi, et nous précisons dans l'article *casse-tête* le statut de *puzzle*.

Type 3 : Concurrence interne au Québec entre vocables français et emprunts à l'anglais

Cas type	Vocable français	Emprunt à l'anglais
« tranche de pain grillé »	<i>rôtie</i>	<i>toast</i>
Fréquences dans la BDTS (37 millions)	(f = 47)	(f = 94)
Dispersion : catégories de textes	Surtout littéraires	Surtout oraux
Répertorié	Multi (Q et vieilli dans la francophonie), PRÉ (vieilli ou région.), PLI, ENF, DQA, DFP	Multi (masc.), PRÉ, PLI, DFP, ENF
Critiqué		DQA (anglic.)
Non répertorié		

Cet exemple illustre le cas de deux mots largement employés au Québec; la concurrence se situe entre un vocable de facture française, *rôtie*, et un emprunt, parfois critiqué, à l'anglais, *toast*. Ce dernier a une fréquence plus élevée (fréquence 94 dans la BDTS) que son concurrent et apparaît surtout dans la langue orale, alors que *rôtie*, moins fréquent (fréquence 47 dans la BDTS), apparaît surtout dans les textes littéraires. Seul, un dictionnaire québécois critique le terme emprunté à l'anglais. Conformément à nos critères, c'est dans l'article *rôtie* que sera présenté le maximum d'informations fonctionnelles associées à la notion de « rôtie », notamment les renseignements d'ordre sémantique.

Type 4 : Concurrence non réelle entre emprunts à l'anglais et vocables recommandés mais non utilisés (situation au Québec)

Cas type	Emprunt à l'anglais	Vocable recommandé
« stratégie d'entreprise en vue d'attirer des consommateurs »	<i>Marketing</i>	<i>mercatique</i>
Fréquences dans la BDTS (37 millions)	(f = 625)	(f = 0)
Dispersion : catégories de textes	Tous les types	
Répertorié	Multi, PLI, ENF DQA (anglic.), PRÉ (anglic.), DFP (anglic.)	Multi (note), PRÉ (recommandation officielle)
Critiqué		PLI, DQA, DFP (remarque)
Non répertorié		

Dans l'exemple ci-dessus, la concurrence entre les deux termes n'est que virtuelle, puisque le terme recommandé, *mercatique* n'est pas implanté dans l'usage au Québec; seul le terme emprunté à l'anglais, *marketing*, est utilisé. En effet, *mercatique* a la fréquence 0 et *marketing*, la fréquence 625 dans la BDTS. En outre, *marketing* est employé dans tous les types de discours, malgré le fait qu'il soit critiqué par certains dictionnaires français et québécois; il est par ailleurs accepté par d'autres (Multi, GDT). Selon nos critères, le terme usuel *marketing* fera l'objet d'un article complet, précisant le statut recommandé du terme *mercatique*, qui lui, aurait un article renvoi.

Type 5 : Concurrence entre vocables de langue générale et vocables des langues de spécialité)

Cas type	Langue générale	Langue plus spécialisée
« matière plastique »	<i>polythène</i>	<i>polyéthylène</i>
Fréquences dans la BDTS (37 millions)	(f = 25)	(f = 43)
Dispersion : catégories de textes	Tous les types	Textes spécialisés
Répertorié	Multi, PRÉ (marque déposée), DQA, DFP	PRÉ, Multi, PLI, ENF, DQA, DFP
Critiqué	PLI	
Non répertorié		

L'exemple ci-dessus illustre la concurrence entre un mot de la langue générale, *polythène*, et un terme spécialisé, *polyéthylène*. Le terme spécialisé est plus fréquent (fréquence 43 dans la BDTS) que le mot de langue générale (fréquence 25). Toutefois, *polyéthylène* est employé surtout dans les textes spécialisés, alors que *polythène* est utilisé dans tous les types de textes. Aucun des deux mots n'est critiqué. Selon les critères retenus, l'article *polythène* sera le support principal des données sémantiques et nous ferons un article renvoi pour le terme spécialisé *polyéthylène*.

Type 5 : Concurrence entre vocables de langue générale et vocables des langues de spécialité)

	En usage au Québec	En usage hors Québec
Langue générale	<i>huard</i> (f = 17)	<i>plongeon</i> (f = 2)
Langue spécialisée (nom relativement banalisé de l'espèce la mieux connue)		
Dans la terminologie francophone unifiée, qui sert de référence depuis 1993, terme généralement adopté au Québec.	<i>plongeon huard</i>	
Dans la terminologie antérieure à 1993, terme encore très largement diffusé.	<i>huart</i> (à collier) (f = 20)	
Dans la terminologie européenne antérieure à 1993, terme dont l'usage se maintient en Europe.		<i>Plongeon imbrin</i>

L'exemple ci-dessus illustre un cas plus complexe de concurrence entre un mot de langue générale et un terme spécialisé, cette concurrence étant de plus affectée par la variation chronologique et la variation diatopique. De fait, dans la langue générale, on relève *huard* (fréquence 17 dans la BDTS) au Québec, alors que *plongeon* (fréquence 2 dans la BDTS) est utilisé dans le reste de la francophonie. Dans la langue spécialisée, les ornithologues québécois utilisent généralement *plongeon huard* pour désigner l'espèce la plus connue, conformément à la terminologie francophone unifiée depuis 1993, alors qu'en Europe, on continue plutôt à employer *plongeon imbrin*. Au Québec enfin, la même espèce est encore régulièrement désignée par le terme spécialisé qui avait cours avant 1993 et qui demeure très largement diffusé, soit *huart à collier*, avec la variation orthographique *huart* (fréquence 20 dans la BDTS). Conformément à nos critères, nous ferons un article complet sous *huard*, et un article renvoi pour *plongeon*. L'ensemble des informations fonctionnelles associées au concept de « huard » sera donné sous *huard*.

Conclusion

En conclusion, nous rappelons toute l'importance que revêt l'établissement de la nomenclature d'un dictionnaire. Notre choix de procéder à l'élaboration d'une nomenclature originale nous a contraints à la conception d'une banque de données textuelles représentative du français en usage au Québec (plus de 10 000 textes soigneusement sélectionnés et reflétant les usages dans les domaines littéraires, journalistiques, spécialisés, didactiques et oraux). Outre son apport essentiel quant à l'élaboration d'une nomenclature originale du français standard en usage actuellement au Québec, l'intérêt d'une telle banque concerne également la description des mots de cette nomenclature : analyse sémantique des mots en contextes, choix d'exemples et de citations, découpage des sens et élaboration des définitions, repérage et sélection de cooccurrents, de collocations et de phraséologismes de toutes sortes, hiérarchisation des usages en fonction de divers registres de langue (orale/écrite, générale/spécialisée, familière/soignée).

Nous avons extrait de la BDTS une nomenclature de quelque 57 000 noms communs, que nous avons enrichie et validée à l'aide de douze dictionnaires usuels existants et plus d'une centaine de répertoires spécialisés de domaines diversifiés. La variation constatée touche plusieurs phénomènes, principalement la variation graphique et la variation lexicale; cette dernière fait varier notamment l'étendue de la nomenclature des dictionnaires. En ce qui a trait à la sélection de la nomenclature du futur dictionnaire du français standard en usage au Québec, les critères suivants ont guidé notre choix : fréquence et dispersion des mots dans la BDTS, présence et sanction des mots dans les dictionnaires usuels, français et québécois, en tenant compte des mots largement diffusés dans la francophonie, présence et sanction des mots dans certaines nomenclatures particulières (répertoires d'avis de recommandation ou de normalisation, répertoires de mots critiqués, dictionnaires et vocabulaires spécialisés de pointe (informatique, par exemple), ou décrivant certains domaines spécifiques au Québec (relations du travail, faune et flore, éducation, hiver et nord, acériculture, politique, etc.).

Une fois les unités lexicales sélectionnées, nous avons procédé à une certaine hiérarchisation des usages, comme nous l'avons démontré dans les exemples précédents : les mots faisant partie de l'usage standard formeront des articles complets dans lesquels sera précisé le statut des formes en concurrence. Ces dernières auront aussi leur place dans la

nomenclature et elles feront l'objet d'une entrée renvoi, ou, le cas échéant, d'un renvoi pour le sens concerné.

L'établissement de cette nomenclature hiérarchisée est en outre essentielle pour le développement de logiciels de traitement de l'information textuelle, des logiciels de correction orthographique, de dictée et de reconnaissance vocales qui doivent s'adapter à l'usage du français au Québec et même en Amérique. À titre d'exemples, les milliers de mots, de sigles, d'appellations d'entreprises, de commerces, d'associations, d'organismes, d'institutions, de toponymes, de gentilés, etc., qui ne seront jamais pris en compte par les logiciels sans un repérage et une description de ce vocabulaire.