



Attribution et variation du genre d'emprunts à l'anglais, à l'italien, au japonais et à l'arabe dans le lexique du français

Mémoire

Rémi Belleau

Maîtrise en linguistique

Maître ès arts (M.A.)

Québec, Canada

© Rémi Belleau, 2016

Résumé du mémoire

Le genre des emprunts lexicaux référant à des entités non-sexuées en français est parfois considéré comme arbitraire, alors qu'il est parfois vu comme motivé par sa forme physique et/ou ses significations. Puisque les avis diffèrent à ce sujet, nous nous sommes intéressés à analyser de nombreux critères pouvant contribuer à l'attribution du genre d'un emprunt. Nous avons constitué quatre corpus, chacun composé de textes issus d'une communauté linguistique emprunteuse (Québec et Europe) et d'un niveau de formalité (formel ou informel). Nous avons observé que le genre des emprunts varie considérablement dans de nombreux cas. Nous constatons que les emprunts des langues à genre (italien, arabe) conservent généralement leur genre originel. Les critères sémantiques et de forme physique peuvent autant justifier le genre d'un emprunt l'un que l'autre. Le critère sémantique le plus opératoire intègre chaque emprunt dans un paradigme conceptuel regroupant plusieurs unités lexicales sous une même conceptualisation et, généralement, un genre commun au paradigme.

Abstract

Gender assignment to lexical non-human loanwords in French is not easily predictable. Some authors will consider it as arbitrary, others as motivated from physical or semantic properties of a given word. Because points of view on this matter are highly divided, we investigated and analysed many criteria that may contribute to gender assignment to loanwords in French. We compiled four corpora, each one composed of texts originating from one of two different linguistic communities (the province of Québec or Europe) and either a formal or informal context. We noticed that a loanword's assigned gender varies considerably between usages in many cases. However, in general, loanwords from languages having a binary gender feature, such as Italian or Arabic, usually keep their original gender. Semantic criteria may justify loanword gender as much as physical criteria. The most productive semantic criterion is the conceptual paradigm into which the loanword is integrated, which groups together many lexical units under a similar conceptualisation, and usually a common gender.

Table des matières

Résumé du mémoire.....	iii
Abstract.....	v
Table des matières.....	vii
Liste des tableaux.....	xi
Liste des figures.....	xiii
Liste des sigles.....	xv
Remerciements.....	xvii
1. Introduction.....	1
1.1 Objectifs de recherche.....	1
1.2 Hypothèses de recherche.....	3
1.2.1 Qu'est-ce qui détermine le genre des noms morphologiquement simples?.....	3
1.2.2 Les emprunts des langues à genre conservent-ils leur genre?.....	4
1.2.3 Dans quelle mesure un genre à valeur générique peut-il être attribué?.....	4
2. État de la question.....	7
2.1 Sur l'emprunt.....	7
2.2 Sur le genre en français.....	9
2.2.1 L'aspect sémantique des genres sémantiquement spécifiques.....	11
2.2.2 L'aspect sémantique de l'archigenre.....	16
2.2.3 Le genre et l'animéité.....	20
2.3 Sur l'attribution du genre en français.....	20
2.3.1 Critères de forme physique.....	21
2.3.2 Critères sémantiques.....	24
2.3.3 Conservation du genre.....	28
2.4 Le concept et la conceptualisation.....	28
3. Cadre théorique.....	31
3.1 Position sur l'aspect sémantique du genre.....	31
3.2 Position sur l'attribution du genre et sur le choix des critères retenus.....	33
4. Méthodologie.....	37
4.1 Objet d'étude et sélection des unités à examiner.....	37
4.1.1 Choix des langues donneuses.....	37
4.1.2 Sélection des unités.....	37
4.2 Recherche préliminaire et comparaisons lexicographiques.....	39
4.3 Constitution des corpus.....	40
4.3.1 Corpus pilotes.....	40
4.3.2 Corpus définitifs.....	42
4.3.3 Volume réel et relatif des corpus définitifs.....	46
4.4 Identification du genre.....	47
4.4.1 Genre des occurrences des emprunts en corpus.....	47
4.4.2 Genre majoritaire pour analyse par critère.....	48
4.5 Remplissage de la grille d'analyse.....	49
4.5.1 Critères de forme physique.....	49
4.5.2 Critères sémantiques.....	53

4.5.3 Conservation du genre	55
5. Résultats et analyse	57
5.1 Résultats généraux	57
5.1.1 Nombre d'occurrences d'emploi pour chaque genre	57
5.1.2 Degré de présence des unités par langue par corpus	60
5.2 Résultats spécifiques	62
5.2.1 Genre majoritaire	62
5.2.2 Retour sur la recherche préliminaire	63
5.3 Résultats par critères	65
5.3.1 Réanalyse morphologique (phonétique et graphique)	66
5.3.2 Syllababilité finale et entière	70
5.3.3 Finale graphique et similitude graphique	73
5.3.4 AP par phonème final et pénultième	75
5.3.5 AP par rime riche	79
5.3.6 Paronymie	80
5.3.7 IP par similitude conceptuelle	81
5.3.8 IP par hyperonymie	83
5.3.9 IP par ellipse d'un signe entier	83
5.3.10 Conservation du genre	85
5.3.11 Initiale vocalique	87
5.3.12 Opérationnalité simultanée de plusieurs critères	88
5.4 Paradigmes d'unités lexicales	90
5.4.1 Types de voiles et couvre-chefs	90
5.4.2 Types de pâtes	91
5.4.3 <i>Party</i> et les événements festifs	92
5.4.4 Types de saucissons	95
5.4.5 Substituts protéinés	96
5.5 Bilan	96
6. Discussion et retour sur les hypothèses	97
6.1 Retour sur les critères	97
6.1.1 Taux de justification des critères	97
6.1.2 Première classification : forme physique <i>versus</i> signification(s)	98
6.1.3 Deuxième classification : critères phonétiques <i>versus</i> critères graphiques	100
6.1.4 Autres constats	101
6.1.5 Discussion	102
6.2 Retour sur les hypothèses	105
6.2.1 Hypothèse 1 : forme physique et signification(s)	105
6.2.2 Hypothèse 2 : conservation du genre et emprunts aux LAG	105
6.2.3 Hypothèse 3 : archigenre et emprunts aux LSG	106
6.3 Ce qu'on peut retirer de cette étude	106
6.3.1 Neuf constats	106
6.3.2 Possibles interactions des critères	108
6.4 Limites et causes d'erreurs	110
7. Conclusion	115

8. Bibliographie	119
8.1 Ouvrages linguistiques et métalinguistiques	119
8.2 Ressources textuelles exploitées	125
9. Annexes	127
9.1 Nombre d'occurrences et pourcentage de masculin en corpus	127
9.1.1 Emprunts à l'anglais	127
9.1.2 Emprunts au japonais	128
9.1.3 Emprunts à l'arabe	129
9.1.4 Emprunts à l'italien	130
9.2 Résultats obtenus à partir des critères de réanalyse morphologique	131
9.2.1 Réanalyse morphologique phonétique	131
9.2.2 Réanalyse morphologique graphique	135
9.3 Résultats obtenus à partir des critères de syllabité finale graphique, de syllabité entière, de finale graphique et de similitude graphique	139
9.3.1 Emprunts à l'anglais	139
9.3.2 Emprunts au japonais	140
9.3.3 Emprunts à l'arabe	141
9.3.4 Emprunts à l'italien	142
9.4 Résultats obtenus à partir des critères d'AP par phonème final, d'AP par phonème pénultième et d'AP par rime riche	143
9.4.1 Emprunts à l'anglais	143
9.4.2 Emprunts au japonais	144
9.4.3 Emprunts à l'arabe	145
9.4.4 Emprunts à l'italien	146
9.5 Résultats obtenus à partir du critère de paronymie	147
9.5.1 Emprunts à l'anglais	147
9.5.2 Emprunts au japonais	148
9.5.3 Emprunts à l'arabe	149
9.5.4 Emprunts à l'italien	150
9.6 Résultats obtenus à partir du critère d'IP par similitude conceptuelle	151
9.7 Résultats obtenus à partir des critères d'IP par hyperonymie et d'IP par ellipse d'un signe entier	156
9.7.1 Emprunts à l'anglais	156
9.7.2 Emprunts au japonais	157
9.7.3 Emprunts à l'arabe	158
9.7.4 Emprunts à l'italien	159
9.8 Résultats obtenus à partir des critères de conservation du genre et de l'initiale vocalique	160
9.8.1 Emprunts à l'anglais	160
9.8.2 Emprunts au japonais	160
9.8.3 Emprunts à l'arabe	161
9.8.4 Emprunts à l'italien	161

Liste des tableaux

Tableau 1 : présentation des corpus pilotes

Tableau 2 : Volume réel et relatif des corpus

Tableau 3 : Composition et taille réelle des corpus

Tableau 4 : Nombre d'occurrences d'emploi de chaque genre sur toutes les unités à l'étude par corpus

Tableau 5 : Nombre d'occurrences d'emploi de chaque genre pour chaque langue dans chaque corpus

Tableau 6 : Proportion d'utilisation des unités par langue par corpus sur le total de chaque corpus

Liste des figures

Figure 1 : représentation bidimensionnelle du genre

Figure 2 : représentation du genre par les marques

Figure 3 : niveau de généralité des lexèmes *homme* et *femme*

Liste des sigles

En ordre alphabétique, voici les sigles employés dans ce travail de recherche:

AP : *analogie phonétique*, processus d'attribution du genre où l'unité reçoit le genre d'une unité à forme physique similaire.

CV : *consonne-voyelle*, représente la structure syllabique finale d'une unité se terminant par une consonne puis une voyelle.

CVC : *consonne-voyelle-consonne*, représente la structure syllabique finale d'une unité se terminant par une voyelle puis une consonne.

DHLF : *Dictionnaire historique de la langue française*, ouvrage de 1992, réédité en 2006, et rédigé par Alain Rey et collaborateurs.

F : *féminin*, genre grammatical.

FO : *faiblement opératoire*, se dit des critères dont le taux de justification ne dépasse pas le seuil de 60%.

GLLF : *Grand Larousse de la langue française*, ouvrage de 1971 rédigé par Guilbert et collaborateurs.

GM : *genre majoritaire*, le genre le plus représentatif en corpus des unités à l'étude.

IP : *intégration paradigmaticque*, processus d'attribution du genre où l'unité intègre un paradigme lexico-conceptuel.

LSG : *langue sans genre*, définit les langues donneuses ne possédant pas la catégorie du genre dans le substantif, comme c'est le cas de l'anglais ou du japonais.

LAG : *langue avec genre* ou *langue à genre*, définit les langues donneuses possédant la catégorie du genre dans le substantif, comme c'est le cas de l'italien ou de l'arabe.

M : *masculin*, genre grammatical.

MAF : *marques associées au féminin*, l'ensemble des formes physiques (graphiques et phonétiques) associées au genre féminin.

MAM : *marques associées au masculin*, l'ensemble des formes physiques (graphiques et phonétiques) associées au genre masculin.

PO : *potentiellement opératoire*, se dit des critères dont le taux de justification dépasse le seuil de 60%.

V : *variable*, décrit le genre majoritaire en corpus comme étant majoritairement à la fois masculin et féminin.

Remerciements

Avant de commencer ce mémoire, je tiens à remercier plusieurs personnes pour l'aide qu'ils m'ont apportée, aide sans laquelle je ne pourrais présenter ici le fruit d'un long labeur.

Je tiens à remercier tout d'abord mon directeur de maîtrise Bruno Courbon pour sa patience et pour les nombreux commentaires constructifs dont il m'a fait part. Sans son aide essentielle et ses remises en question nécessaires, je n'aurais pu compléter un travail d'une telle ampleur. Je voudrais également le remercier pour m'avoir donné accès au logiciel du projet de recherche qu'il dirige : *Appropriation du vocabulaire spécialisé par les francophones : canaux de diffusion et modes d'intégration*. Grâce à cela, j'ai pu constituer mon corpus français informel beaucoup plus rapidement que les autres corpus.

En second lieu, je tiens à remercier les membres du comité d'évaluation de ce projet. Le comité est composé de messieurs Bruno Courbon, Jacques Ouellet et Patrick Duffley. Je voudrais les remercier pour la rétroaction qu'ils m'ont fournie. Leurs commentaires m'ont permis de perfectionner ce travail.

En troisième lieu, je tiens à remercier mesdames Claudia Borgonovo et Johanna-Pascale Roy pour leur aide méthodologique qui a contribué à la rigueur nécessaire à ma formation.

En quatrième lieu, je tiens à remercier Kendall Vogh, Myriam Paquet-Gauthier, Hugo Mailhot et Caroline Sigouin pour avoir suscité ma réflexion et pour avoir remis en question avec leur esprit critique ce que je croyais acquis.

En cinquième lieu, j'aimerais remercier Kendall Vogh pour son aide à la traduction en anglais de l'abstract. Je voudrais également remercier Amal Zgati pour sa gentillesse et sa connaissance de l'arabe.

En sixième lieu, je voudrais remercier mes assistants à la constitution des corpus, Frédérique Lamirande, Sébastien Roy, Rose Déry et Marie-Claude Boutin. Sans leur aide, la constitution des corpus formels aurait été bien plus longue.

En dernier lieu, je tiens à remercier du fond du cœur ma famille et mes amis pour leur support et plus particulièrement Marie-Hélène Gosselin pour son aide à la révision du texte et pour son constant soutien moral dans ce long projet.

Merci également à ceux que j'oublie, mais dont la contribution ne peut être ignorée.

1. Introduction

1.1 Objectifs de recherche

Le genre grammatical en français est une catégorie grammaticale traditionnellement conçue comme binaire dont les deux constituants sont couramment dénommés *masculin* et *féminin*. Ces appellations causent fréquemment des confusions entre les multiples usages du concept de *genre* (Courbon, à paraître), qu'on pourrait appeler *sexe* (ou *sexe biologique*), *genre* (grammatical) et *gender* (*idem*, ou *genre social*). Les usages du concept de genre nommés *sexe* et *gender* ne seront pas abordés dans ce travail de recherche. Le terme *genre*, au minimum pluricentenaire (voire plurimillénaire), est souvent associé avec le concept de *sexe* dans l'esprit des locuteurs. Cette association cause fréquemment des confusions entre ces deux concepts, car, en langue, elle ne reflète qu'une partie du lexique, celle des unités désignant des entités animées humaines (bien qu'il existe de nombreuses unités dont le genre grammatical utilisé ne correspond pas avec le sexe, telles que *victime*, *sentinelle* ou *star*). Tel que Milner (1989:191) le mentionne, les genres, afin d'être désambiguïsées de leur valeur « sexosocioaffective », pourraient être dénommés *groupe* « *le* » et « *groupe* » *la*.

Selon Audet et Ouellet (1984:41), l'alternance du morphème du genre (par exemple : *le mode / la mode*, *le garde / la garde*, *le rond / la ronde*) a deux fonctions. Elle sert « soit pour marquer le renvoi à un référent différent, soit pour exprimer la distinction des sexes ». Or, des observations personnelles régulières nous permettent de montrer des différences dans l'attribution du genre qui ne sont pas corrélées au sexe ou qui ne renvoient pas à des référents différents. Par exemple, *un radis / une radis* « tubercule rouge comestible », *un prank / une prank* « canular, blague », *un purse / une purse* « sac à main », et *un advi^{MD} / une advi^{MD}* « comprimé d'ibuprofène »¹. Considérant de tels exemples (et plusieurs autres), nous sommes venu à la conclusion que les différentes explications examinées (voir section 2.2.1) de la motivation du genre en français ne pouvaient pas fournir une explication complète du problème. Nous avons donc cherché, dans notre mémoire, à remettre en question cette conception binaire que sont les groupes *le* et *la* d'un point de vue sémantique en l'appliquant aux emprunts, sans avoir l'ambition de pouvoir répondre entièrement au problème. Nous n'avons aucune prétention quant à l'exhaustivité de notre analyse du problème. Nous avons simplement cherché à offrir une piste supplémentaire à la question de la variation du genre.

La question de l'arbitraire et de la motivation du genre a fait couler beaucoup d'encre. Cependant, l'attribution du genre aux emprunts a été bien moins étudiée et elle suppose un minimum de motivation, sans quoi les

¹ À cela nous désirons ajouter les unités *streptocoque* et *deal* que nous avons récemment entendues utilisées au masculin et au féminin. De même, le genre des unités *moustiquaire*, *trampoline*, *tentacule*, *pétale*, *pétoncle*, *nacre*, *termite*, *rail*, *après-midi* et *neurone* (ainsi que nombre d'unités commençant par une voyelle phonétique ou des noms de styles musicaux tels que *le pop/la pop*) varie sensiblement. On peut parler de différences interindividuelles du genre (Courbon 2015b).

locuteurs ne s'entendraient pas sur le genre à utiliser. Les études sur l'attribution du genre aux emprunts ont dévoilé nombre de critères formels ou sémantiques pouvant justifier le genre d'emprunts. Or, parmi ces études, seul un petit nombre ont considéré plus d'un critère (Poplack *et al.* 1982, Guilford 1999). De plus, toutes n'analysaient que des emprunts à l'anglais. Avec de nombreux critères pouvant tous (supposément) justifier au même titre le genre des emprunts lexicaux, nous avons décidé de faire le point sur le sujet en comparant individuellement une quinzaine de critères et en regardant lesquels pouvaient réellement justifier le genre d'emprunts de plusieurs langues.

Notre travail de recherche a un double objectif. Il vise d'abord à redéfinir la notion de genre pour prendre en compte la variation du genre (ci-dessus) que nous observons à travers les usages et les communautés. Ensuite, notre mémoire cherche à valider ou infirmer les critères d'attribution du genre que nous avons trouvés dans la littérature. Le premier objectif de recherche est de nature qualitative, cherchant à compléter et à nuancer les connaissances sur l'aspect sémantique de la motivation du genre en français, et ce, particulièrement dans le cas des emprunts. La sémantique ayant un objet d'étude plus abstrait que l'étude de la forme physique, notre étude a dû se limiter à quelques facettes de la sémantique². C'est donc dire que quelques aspects ont été laissés de côté, notamment ceux de la comparaison de significations des unités entre les individus et les communautés. Le second objectif est davantage de nature quantitative, empirique, prenant pour base des corpus composés d'usages produits par des locuteurs. Nous utilisons des notions statistiques basiques (pourcentages et moyennes). La méthode statistique est employée dans son plus simple appareil afin de mettre en évidence les phénomènes que nous étudions sans devoir utiliser des concepts mathématiques plus avancés qui compliqueraient inutilement l'accès aux conclusions que nous tirons. Nous comparerons des critères d'attribution du genre de nature sémantique et de forme physique, et nous sommes conscient qu'une telle comparaison, sur un plan quantitatif, comporte ses risques d'erreurs de méthode et d'interprétation. Néanmoins, nous avons cherché à mettre sur un pied d'égalité tous les types de facteurs d'attribution du genre afin de leur donner une base commune sur lesquels nous pouvions tirer des conclusions sur leur potentiel d'efficacité. Nous ne cherchons pas à confronter le physique au sémantique, car nous considérons qu'il est probable que le genre puisse trouver source dans le physique et le sémantique combinés. Notre approche multifactorielle peut être considérée comme « artificielle », car elle compare des éléments d'ordres différents (physique et sémantique). Elle peut être considérée comme réductrice, voire volontairement naïve, car elle n'a pas pour objectif de décider définitivement quelle serait la raison principale de l'attribution d'un genre à un emprunt. Nous ne cherchons pas à prouver l'efficacité d'un groupe de critères

² Montrer une corrélation entre des formes physiques et un genre peut être assez facile à établir, car la forme est matérielle, tangible et concrète; à l'inverse, montrer une corrélation entre la sémantique et un genre est une entreprise beaucoup plus ardue, car le sens est immatériel, intangible, abstraite, mais perçu et (ré)interprété. Nous tenterons donc d'approximer la signification des unités et des genres, car il s'agit d'un objet abstrait dont le sens saisi doit être formulé pour être communiqué dans notre étude.

sur un autre groupe, mais à vérifier la validité potentielle de chacun des facteurs et de suggérer des pistes de recherches ultérieures avec d'autres unités.

Si le genre des emprunts à l'anglais a largement été étudié, le genre d'emprunts à d'autres langues en français est resté très peu étudié. Toutes les études que nous avons recensées ont trouvé une majorité d'emprunts portant la marque du masculin pour une minorité d'emprunts portant celle du féminin (voir note 36). Nous avons cherché, dans la seconde partie de notre mémoire (davantage quantitative), à voir dans quelle mesure cette prédominance du masculin serait applicable aux emprunts à d'autres langues. Pour cela, nous avons également choisi des emprunts à l'italien, au japonais et à l'arabe. Nous avons également cherché à mettre en évidence différentes raisons possibles d'attribution du genre aux emprunts selon les auteurs et à tester dans quelle mesure leurs conclusions peuvent être applicables ou non aux emprunts de notre échantillon.

1.2 Hypothèses de recherche

1.2.1 Qu'est-ce qui détermine le genre des noms morphologiquement simples³?

Nous avons fait le choix dans ce travail de recherche de ne pas nous placer sur la question de la prééminence d'un ordre de facteurs (sémantiques *versus* de forme physique) par rapport à un autre⁴. Ce choix est motivé par la volonté d'être aussi impartial que possible quant à l'évaluation des facteurs d'attribution du genre. Pour cela, nous considérons que la forme physique et l'aspect sémantique d'une unité peuvent contribuer à motiver le genre d'un emprunt. Dans une perspective sémasiologique où la forme physique d'une unité donne le genre, un morphème associé à un genre pourrait motiver son emploi. Pour éviter de voir le genre justifié par ce critère de manière systématique, nous avons choisi de sélectionner uniquement des unités « morphologiquement simples » (*i.e.* base seule, ex : *table*, *mur*), plutôt que des unités « morphologiquement complexes » (*i.e.* base + suffixe, ex : *banan+ier*, *marqu+eur*).

Comme nous le verrons dans la section 2.1 sur l'emprunt, les emprunts ne sont pas généralement décomposables en morphèmes une fois empruntés, ce qui les rend morphologiquement opaques, donc simples aux yeux d'un locuteur ne parlant pas la langue donneuse. Suivant ces postulats, nous sommes

³ Nous entendons par cette appellation des noms composés d'une base seule sans suffixe ajouté. On qualifie ces noms de *monomorphématiques* ou *monomorphémique* (Kertész & Rákosi 2012:119). Cela n'inclut pas toutes les unités lexicales simples, car nombre d'entre elles sont des « mots morphologiquement complexes », ou *polymorphématiques*.

⁴ On oppose le point de vue *sémasiologique* au point de vue *onomasiologique*. Prenons l'exemple de *bananier*. La vision sémasiologique veut que le suffixe *-ier* donne le genre masculin (citons quelques auteurs qui suivent ce point de vue : Lyster (2006), Surridge (1984, 1985, 1986, 1994), Grevisse (1964, 1980, 1986, 2008), Tucker *et al.* (1970), Guilford (1999)). À l'inverse, le point de vue onomasiologique suppose que le masculin est ancré dans le concept d'ARBRE et que le genre motive le choix du suffixe *-ier* plutôt que *-ière*. Pour Audet & Ouellet (1984:37), « le choix du morphème [de genre] approprié se fait en corrélation avec la valeur lexicale du substantif » et ce choix du genre est défini comme une « référence interne » au concept (p.41).

amené à formuler notre **première hypothèse** supposant que le genre d'un emprunt morphologiquement simple soit autant déterminé par sa forme physique que par sa (ou ses) signification(s). Nous croyons possible d'observer une influence combinée par les critères de forme physique et les critères sémantiques avec les unités de notre échantillon, selon les corpus que nous avons constitués.

1.2.2 Les emprunts des langues à genre conservent-ils leur genre?

Notre **deuxième hypothèse** se base sur une étude que Stolz (2009) a menée sur des emprunts italiens en maltais. Elle est venue à la conclusion que, dans la majorité des cas, le genre de l'emprunt dans la langue donneuse était identique dans la langue cible. À partir de cela, nous formulons une seconde hypothèse, supposant qu'un emprunt d'une langue à genre (abrégé LAG) conservera son genre s'il offre suffisamment de ressemblances avec les genres du français. Il va sans dire que les genres de l'italien sont très similaires aux genres du français, au nombre de deux, possédant la même origine et environ les mêmes fonctions. Or, les genres sémitiques de l'arabe sont différents en ce qu'ils permettent différentes fonctions⁵ que le français doit exprimer à travers la dérivation ou la composition lexicale. Comme Stolz (2009) l'a suggéré, il est possible d'apercevoir une correspondance⁶ entre les genres des langues romanes et sémitiques (italien et maltais en l'occurrence, tel que l'exemple de la page 332 le montre : « malt. *abbiltà* « habilité » (<it. *abilità*) »). Cette constatation nous permet de supposer qu'il pourrait y avoir une correspondance similaire entre les genres du français et ceux de l'arabe, et qu'on pourrait observer une conservation du genre dans la majorité des cas.

1.2.3 Dans quelle mesure un genre à valeur générique peut-il être attribué?

Notre **troisième hypothèse** se base sur la redéfinition du genre que nous effectuerons dans la section 2.2. Nous observons que les genres masculins et féminins correspondent à des conceptualisations spécifiques des réalités et que les emprunts ont fréquemment, dans l'esprit de locuteurs peu familiers avec l'unité lexicale empruntée, une conceptualisation vague ou abstraite de la réalité. Par conséquent, les locuteurs utilisent de manière générale un genre à valeur générique (marqué par la sémiologie du masculin) pour les unités dont ils ont une conceptualisation vague ou abstraite. Considérant que les emprunts aux langues à genre (abrégé LAG) conserveront majoritairement leur genre, on devrait voir un autre mécanisme d'attribution général ne se basant ni sur la forme physique ni sur le sens pour les emprunts aux langues sans genre (abrégé LSG). Nous supposons que ce processus consiste en l'attribution d'un genre à valeur générique et qu'on devrait voir s'actualiser dans les emprunts aux LSG une plus grande fréquence de marques associées au masculin. Nous pensons donc que le genre des emprunts de ces langues se verra motivé d'abord par un genre à valeur

⁵ Il permet de former des noms abstraits (*Hasan* « bon », *Hasanat* « bonté »), des noms collectifs (*kāhen* « prêtre », *kāhenāt* « prêtres »), des singulatifs (*nomina unitatis*; *baqar* « troupeau de vaches », *baqarat* « vache ») et des diminutifs (hébreu **yāniq* « jeune arbre »; *yāniqat* « brindille, petite branche »). (Ibrahim 1973:44)

⁶ Par correspondance, on entend une partie des fonctions (distinction des sexes) et une conservation potentielle du genre.

générique, car les emprunts ne sont généralement pas clairement définis dans l'esprit de la plupart des locuteurs qui les utilisent. Dans la mesure où l'attribution d'un genre à valeur générique n'est pas possible, nous supposons que le genre sera motivé par sa représentation spécifique, sa ou ses significations, ou par sa forme physique.

2. État de la question

Nous présenterons, à travers l'état de la question, le cadre théorique dans lequel nous nous situons.

2.1 Sur l'emprunt

Puisque nous nous intéressons aux emprunts du lexique du français, il nous faut d'abord définir la notion d'emprunt. Pour nous, l'emprunt est un signe, qui peut être intégré au lexique d'une langue, et qui est composé d'une forme physique et d'une ou plusieurs significations manifestées à travers de multiples emplois contextuels. Les emprunts sont intrinsèquement des créations externes au français (Rey, Siouffi & Duval 2007:1216) et, comme tout signe, ils sont en général réinterprétés et adaptés sémiologiquement et sémantiquement. Le nouveau signe peut être identique à un signe déjà présent dans le lexique de la langue et il peut référer à une réalité familière ou étrangère aux locuteurs. Cette définition exclut donc la distinction *emprunt lexical* et *emprunt sémantique* (Humbley 1974, DHLF 2006:1228-1230), distinction classique qui oppose la forme physique au sens. Nous ne pensons pas qu'il soit possible d'emprunter uniquement une forme physique – car tout signe, emprunté ou indigène, est la combinaison d'une forme physique et d'une ou plusieurs significations⁷. De plus, nous ne croyons pas non plus qu'il soit possible d'emprunter uniquement une signification – car, pour toute innovation de signification, Nyckees (2006:45) explique qu'une signification nouvelle apparaît comme « une instanciation de la signification ancienne » et s'autonomise pour se détacher de l'ancienne, étendant le sens d'une unité à de nouveaux usages⁸. Un dernier commentaire par rapport à notre définition de l'emprunt serait qu'ils ne sont pas à priori décomposables en morphèmes, sans une connaissance minimale de la langue donneuse, donc *kamikaze* (*kami* « dieu », *kaze* « vent ») et *pizza* (*pizz* = base lexématique, *-a* = morphème flexionnel) ne se décomposeraient pas aisément sans des connaissances de base du japonais ou de l'italien. On peut cependant déduire quelques unités morphémiques de la langue donneuse avec des comparaisons telles que *karaté-karatéka*, *judo-judoka* et, de là, on peut féminiser en *karatékaté* et *judokate*. La sémiologie des emprunts n'est généralement pas aussi transparente que celle des unités non empruntées.

Nous excluons de notre étude le phénomène appelé *réemprunt*, dans lequel une langue A donne un mot à une langue B qui en redonnera une forme similaire ou dérivée à la langue A. Nous excluons les réemprunts,

⁷ Pour nous, le sens est discursif, énonciatif et contextuel, et produit ponctuellement à travers l'emploi de l'unité (Nyckees 1998:248). La signification est propre à la langue et à l'unité, indépendamment du contexte. Elle est construite socioculturellement par les usagers pour leur communauté (p.258). Nous étudions la signification à travers un échantillon d'emplois.

⁸ Nous distinguons ici *usage* et *emploi*, en ce que « l'usage sémantique » relève « de l'expérience pratiquée », de « l'habitude linguistique 'routinisée' », du collectif et du partagé, et que l'emploi relève du sens contextuel (construit en discours), du ponctuel et de l'individuel (Courbon 2012b:118). « L'usage sémantique remplit le rôle d'intermédiaire entre la diversité sémantique des réalisations singulières et le sémantisme en puissance » (*ibid.*). Dans le cas de notre étude, nous nous intéressons à l'usage de certains emprunts à travers les emplois contextuels.

car nous voulons éviter une potentielle interférence du genre de la forme dans la langue donneuse. Les autres conceptions de l'emprunt, en termes de *nécessité* et de *luxé* (DHLF 2006:1229, Lafourcade 2009:28), en termes de *xénisme* et de *pérégrinisme* (Guilbert 1975:92, Gaudin & Guespin 2000:296), en terme de *calque* (Dubois *et al.* 2007:73-74), et en termes d'emprunt *interne* et *externe* (Dubois *et al.* 2007:178, Boulanger & Malkowska 2008:117) sont également laissées de côté, car elles ne satisfont pas nos critères d'analyse ou soulèvent des distinctions non pertinentes pour notre analyse⁹.

De manière générale, les conceptions de l'emprunt représentent le phénomène comme l'adoption (volontaire ou non) d'une unité (lexicale, morphologique, etc.) à une autre langue (par les usagers ou par la langue elle-même)¹⁰ et/ou comme le résultat de ce processus. Seule la définition relevée par Louis Deroy (1956:18), traduisant Vittore Pasani (1946), a retenu notre attention. Pasani définit l'emprunt comme suit : « une forme d'expression qu'une communauté linguistique reçoit d'une autre communauté ». Cette définition apporte une nuance intéressante, en ce qu'elle suggère la transmission d'une unité à une langue receveuse, et non l'inverse. Cette perspective inverse suggère qu'une langue à genre (LAG) donnant un substantif peut avoir tendance à imposer son genre à la langue receveuse. Cet aspect nous intéresse et nous conserverons ce point de vue pour le critère d'analyse de conservation du genre.

Une notion à définir qui affecte notre définition de l'emprunt est celle d'héritage. L'héritage s'oppose au phénomène d'emprunt en ce que les mots hérités nous proviennent directement du latin et sont passés d'une génération à l'autre, en évoluant sur le plan physique et sémantique (GLLF 1971:1579, DHLF 2006:1228). Cette notion est problématique, car elle implique qu'un emprunt, aussi intégré et ancien soit-il¹¹ ne peut être considéré comme un mot hérité, malgré ses évolutions (phonétiques, sémantiques, etc.) et son transit intergénérationnel. L'héritage permet également le phénomène de réemprunt, car un mot d'origine latine peut avoir voyagé par une autre langue et revenir en français, comme c'est le cas de *standard*¹². Ce phénomène sera de moindre importance dans notre étude principalement synchronique, car, passé un certain niveau

⁹ L'attribution d'un genre n'est pas une question de nécessité ou de luxe, car le genre est une notion obligatoire en français. Le genre répond à un besoin d'expression ou d'expressivité (Malkowska 2004:31), qu'il s'agisse d'exprimer une nouvelle réalité ou de transmettre un effet stylistique. Un emprunt est toujours pris par nécessité d'exprimer une nouvelle réalité ou par nécessité de transmettre un effet stylistique (prestige, moquerie, etc.). Il s'agit d'un faux dilemme. Pour ce qui est de *xénisme* et *pérégrinisme*, le caractère d'« étrangéité » d'un emprunt n'affecte pas le choix du genre, car nous étudions les tendances d'usages et un emprunt, intégré ou non à la langue, aura un genre dans chacun des emplois qui en sont faits. Que l'emprunt soit fait à partir d'un signe similaire ou dissimilaire, partagé ou non, emprunté pour la première ou la millième fois, cela ne devrait pas affecter le genre donné à l'emprunt, car celui-ci est, postulons-nous, déterminé par sa forme physique ou par ses significations. La notion de calque linguistique n'est pas intéressante à étudier dans le cadre du genre, car la nouvelle signification est attachée à un signe faisant déjà partie du lexique du français, et donc, le genre ne change pas. La notion d'emprunt interne ou externe fait référence au passage d'un mot d'un domaine à celui d'un autre domaine, comme les termes économiques *prêt*, *intérêt* et *débit* qui proviennent du domaine général de la langue; dans ces cas, le genre ne se modifie généralement pas.

¹⁰ Cf. DHLF (2006:1228), Dubois *et al.* (2007:177), Loubier (2011:10), *Le Petit Robert* (2013:857).

¹¹ Comme *personne* (latin *persona*) provenant de l'étrusque, ou *physique* (latin *physica*) du grec (*Petit Robert* 2015).

¹² De l'anglais *standard*, anciennement en français *estandard* (*Petit Robert* 2015).

d'intégration¹³, un emprunt peut ne plus être considéré comme tel, mais comme héritage. Selon nous, le genre des substantifs (hérités et empruntés) peut être considéré de deux manières: le genre est *hérité*¹⁴ des locuteurs qui l'ont produit auparavant; il est également continuellement réévalué et réassigné par les locuteurs qui tentent – volontairement ou non – de régulariser (par analogie ou autres procédés) ou d'uniformiser le lexique, ce qui implique que même des substantifs intégrés depuis des siècles au lexique du français peuvent changer de genre¹⁵.

2.2 Sur le genre en français

S'il va de soi que nous étudierons le genre, la notion même de genre n'est pas totalement partagée dans la littérature scientifique sur le sujet. Afin d'éviter toute confusion, définissons à l'instant ce que nous appellerons plus tard *genre*. La notion de genre¹⁶ s'oppose d'abord et avant tout à celle d'espèce. Le générique est au genre ce que le spécifique est à l'espèce. Le genre étant d'abord « une représentation notionnelle d'ordre conceptuel » (Malkowska 2004:40), nous considérons que le genre n'existe pas seulement dans une dimension (dite horizontale) opposant traditionnellement (en français) le masculin au féminin. Au contraire, nous considérons que le genre existe dans deux dimensions basées sur le degré de la conceptualisation du concept, opposant d'abord les deux genres précédents à un genre à valeur générique subsumant, dominant, supplantant ceux-ci. C'est ce que nous appelons la verticalité du genre. Nous pouvons donc illustrer notre modèle ainsi :

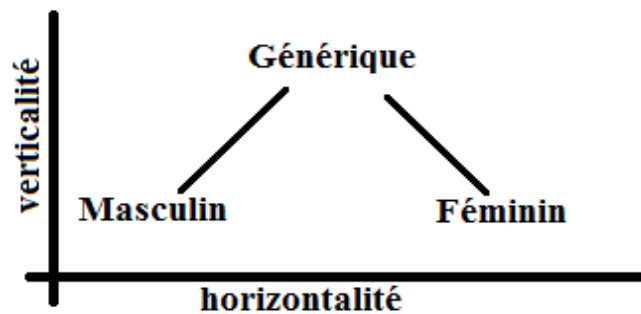


Figure 1 : représentation bidimensionnelle du genre

¹³ Différents points de vue s'opposent sur la notion d'héritage. La définition la plus stricte veut que seul un mot d'origine latine conservé dans le lexique du français soit considéré comme hérité. À l'inverse, un point de vue opposé voudrait, dès que l'unité est partagée par deux locuteurs, qu'elle soit intégrée dans leur lexique individuel et devienne héritée dès qu'elle est partagée par un nouveau locuteur qui apprend (hérite) ce lexème des deux premiers locuteurs.

¹⁴ Notre emphase.

¹⁵ Tel qu'*aigle* [masculin] (latin *aquila* [féminin]). On retrouve encore *une aigle* en héraldique et dans le domaine militaire (*l'aigle romaine*). *Navire* était également féminin pendant longtemps (latin *navis* [féminin]) (cf. doublon *la nef*) (Arnauld & Lancelot 2010:57 [1660]).

¹⁶ On considérera le genre dans notre étude uniquement sous son aspect grammatical. On l'oppose ici à *genre naturel* ou *biologique*, une classification purement référentielle où le masculin correspond au sexe mâle et le féminin au sexe femelle.

Cependant, une telle représentation a deux problèmes. Premièrement, elle superpose le genre à valeur générique aux genres masculin et féminin. Un regard rapide à cette représentation, sans autres explications, pourrait laisser entendre, à ceux qui associent (à tort dans ce travail de recherche) le mot *masculin* aux hommes et le mot *féminin* aux femmes, que les sexes sont dominés, ou du moins qu'un des sexes est dominé, ce qui est loin d'être notre intention. Précisons que « l'emploi des expressions *forme féminine* ou *forme masculine* sera motivé par un souci d'économie : [...], puisqu'aucune forme n'est en soi « féminine » ni « masculine ». » (Courbon, à paraître). Il en va de même pour les mots *masculin* et *féminin* lorsqu'utilisés seuls. Le deuxième problème est que cette représentation illustre uniquement l'aspect sémantique du genre, mais qu'au niveau sémiotique (et à l'esprit de la plupart des locuteurs), la situation est différente. Nous illustrerons donc notre vision du genre ainsi :

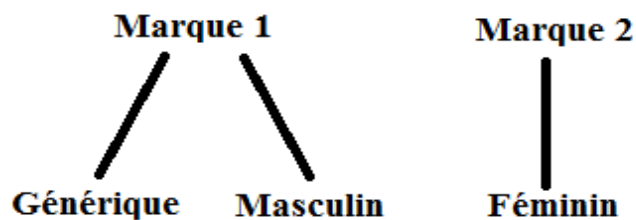


Figure 2 : représentation du genre par les marques

Considérant que le genre est une propriété intrinsèque à (au minimum) tous les substantifs et qu'il relève de la manière dont le locuteur se représente une réalité, nous postulons qu'il existe sémantiquement trois genres, deux à conceptualisation spécifique (le masculin et le féminin) et un à conceptualisation générique¹⁷, qui sont actualisés sémiotiquement sous deux formes (typiquement appelés *masculin* et *féminin* que nous avons appelés ci-haut *marque 1* et *marque 2*)¹⁸. C'est donc dire que ce « genre générique » (que nous appellerons

¹⁷ Notons que « conceptualisation générique » ne renvoie pas à « référence générique ». La référence générique renvoie à un large groupe d'individus, alors que la conceptualisation générique a trait à la manière dont la réalité est représentée, et non la quantité d'individus impliqués dans la référence. Donc, loin de la conceptualisation générique, la majorité des noms collectifs indéfinis portent la marque associée au féminin : *pléthore, multitude, myriade, panoplie, quantité, somme, kyrielle, plupart, foule, majorité, moitié, légion, multiplicité, pluralité, ribambelle, série, suite, succession*, etc. Cette série n'inclut pas non plus les noms collectifs (dont le genre varie souvent, prônant souvent la marque du féminin, mais non exclusif) pour référer à des humains (*foule, horde, famille, équipe, gang, clan, troupe, peuple*), à des animaux (*troupeau, banc, meute, essaim, nuage, nuée, colonie, horde, harde, volée, faune, race, portée*) et à des objets (*tas, ramassis, foison, agglomérat, amas, attirail, rame, rangée, batterie, arsenal, flotte, trousseau, paquet*). La prédominance du féminin pour les noms collectifs indéfinis serait due selon Malkowska (2004:73) à l'intervention du genre féminin inanimé « dans la conceptualisation de la mesure relative d'une contenance approximative, de ce qui se conçoit comme une série indéfinie ou comme une partie d'une série indéfinie ».

¹⁸ Notre postulat. Nous n'utiliserons pas les distinctions *générique/particulière* ou *spécifique/non-spécifique/attributive* de Riegel, Pellat et Rioul (2009:962). Nous employons ici les termes *générique* et *spécifique* comme antagonistes, désignant deux modes d'existence référentielle extensive opposés, *générique* renvoyant à « l'ensemble d'une classe ou d'une sous-classe d'individus » et *spécifique* renvoyant à « un ou des individus particuliers dans l'univers de discours » (*ibid.* p.283).

désormais **archigenre**¹⁹) n'a pas de forme propre et doit s'exprimer sémiotiquement à travers une marque, généralement celle du masculin. À travers notre étude, nous appellerons ces marques²⁰ *marques associées au masculin* (MAM) et *marques associées au féminin* (MAF) afin de ne conserver les appellations de *masculin* et de *féminin* uniquement aux genres sémantiques.

Afin de justifier le choix de l'appellation *archigenre*, nous avons tenu à lister les notions ou les appellations que nous avons exclues. Les voici en ordre alphabétique :

Genre commun : laisse entendre que le masculin et le féminin sont unis en un genre commun (Labrosse 1996:36), comme c'est le cas en suédois et en danois. Certains usages de *genre commun* le confondent avec le genre par défaut (Labrosse 2001:35) : « la forme dont on se sert dès qu'on ne spécifie pas qu'il s'agit d'un être de sexe féminin ».

Genre indifférencié (Grevisse 1986:755) : utilisé dans les contextes où le masculin et le féminin doivent être réunis par un seul pronom.

Genre neutre (ou *neutre* utilisé seul, Labrosse 2001:34) : a deux effets. Le premier fait le parallèle avec le genre grammaticalement neutre présent dans de nombreuses langues (allemand, russe, latin, etc.). Le deuxième a pour effet de neutraliser la distinction masculin-féminin, comme le *genre indifférencié*.

Genre non marqué (Roché 1992, Ndiaye 1994, Kilarski 1997, Carroll 2005) : utilise le concept morphologique de *marque*, un concept binaire biaisé par une vision sémiotique et non sémantique qui ne cherche pas à expliquer l'ajout de cette marque.

Genre par défaut (Kilarski 1997) : appellation trop restrictive, en ce sens qu'elle neutralise la verticalité du genre sémantiquement générique en attribuant automatiquement le masculin aux emprunts, sans même laisser le choix du genre le plus représentatif (de la réalité de l'emprunt) aux locuteurs.

Genre premier (ou *premier genre* – Labrosse 1996:28, 31, cité dans Elmiger 2013:117) : laisse entendre qu'un genre viendrait avant un autre, ce qui n'est vrai que dans la conceptualisation du genre à valeur générique qui précède les genres spécifiques. Or, puisque l'archigenre est généralement exprimé à travers la marque du masculin, certaines personnes pourraient attribuer l'appellation *genre premier* au masculin et étendre cette conception au *gender* social associé aux hommes, ce qui n'est pas le cas.

2.2.1 L'aspect sémantique des genres sémantiquement spécifiques

Lorsqu'on observe des paires minimales sémantiques et morphologiques où le genre diffère, telles que le sont par exemple les paires *plant-plante* et *point-pointe*²¹, il apparaît évident que le masculin et le féminin dénotent des valeurs conceptuelles différentes, n'ayant aucun lien avec la représentation des sexes qu'on peut retrouver dans les noms animés. Regardons, dans cette section, comment les auteurs de la littérature ont abordé la question de l'aspect sémantique des genres (que nous considérons comme spécifiques). Nous nous

¹⁹ Nous avons décidé d'appeler le genre générique « archigenre », basé sur le même modèle qu'*archiphonème* et *archiséme*, de la racine grecque *ἀρχι-* signifiant la prééminence, le règne du concept comme suprême.

²⁰ Nous utiliserons le terme *marque* comme l'ensemble des traces physiques des genres (actualisé à travers le phénomène d'accord). Nous séparons *marque associée au masculin* (MAM) et *marque associée au féminin* (MAF). Nous ne faisons donc pas référence à l'opposition morphologique *marqué/non marqué*.

²¹ Voir la suite de phrases (1) à (9), dans cette section. Ces phrases mettent en évidence le « genre sous sa forme pure » (Togebly 1965:143).

situerons par rapport à ces théories davantage dans le troisième chapitre (cadre théorique) de notre travail de recherche.

Les premiers linguistes à avoir assigné aux genres une valeur sémantique distincte et « non arbitraire » sont Damourette et Pichon (1968, première édition en 1911). Pour eux, à travers un processus mental qu'ils nomment *sexuiseemblance*, l'esprit humain affecterait à chaque substantif un genre en fonction de caractéristiques stéréotypiques associées aux conceptions des sexes (pour le locuteur), sexes qui seraient représentés par les genres grammaticaux. Ainsi, pour eux, la mer serait du genre féminin, car elle est « d'aspect changeant comme une femme, journalière, d'humeur mobile comme une jolie capricieuse, attirante et dangereuse comme une beauté perfide » (p. 371-2). De même, « l'*arc* a pour femme l'*arche*; ils ont pour enfant un petit arc, un *arceau*. L'*arc* est géométrique et militaire; l'*arche* est plus matérielle, est d'une substance lourde; l'*archet* d'ailleurs, leur autre fils, ressemble à sa mère par sa tendre robustesse » (p. 379, cité dans Malkowska 2004:10). Claire Michard (2002:52) classe ce type de raisonnement, et à juste titre lorsqu'appliqué au lexique entier, de « délire sexualisant ». Si nous reconnaissons que certaines caractéristiques prototypiques de la perception sociale des sexes peuvent être assignées aux genres correspondants, notamment dans le cas d'allégories ou de métaphores, nous sommes d'avis que le genre des unités lexicales référant à des entités non sexuées ne se base pas sur les sexes pour classer les noms.

Hjelmslev (1938:140-151) a été le premier linguiste à attribuer aux genres une signification qui ne se basait pas sur la distinction sexuée homme vs femme. Pour lui, le masculin et le féminin reflètent, dans une dimension dynamique, la concentration et l'expansion, dans une dimension statique un état discret ou un état compact et, dans une dimension subjective, l'idée de massif vs ponctuel (sans spécifier à quel genre correspond quelle propriété). Togeby (1965:143) précise que « le masculin désignerait la concentration, le féminin l'expansion ». Togeby ajoute également des valeurs aux genres qu'aucun autre auteur n'a mis en évidence jusqu'à présent : le masculin mettrait en avant l'idée de passivité alors que le féminin serait lié à « l'idée de production ou d'activité » (*ibid.*). À l'inverse, Martinet (1956:94, cité dans Violi 1987:24) voit, dans son étude du féminin en indoeuropéen, « le vague sentiment d'une passivité et d'une réceptivité de la terre » où il est possible qu'on puisse bâtir une mythologie autour de la fertilité paradoxale²² de la terre. Si on se fie à Lafont (2001:140), cette opposition passivité/activité au sein du genre remonterait au latin où le masculin était « plutôt du côté de l'existant concret » alors que « le féminin [était] du côté de l'acte ou du concept. » Finalement, pour Arrivé (1997:89), la notion d'expansion associée au féminin amène une valeur augmentative (non mentionnée par Hjelmslev et Togeby), valeur à laquelle il amène des contre-exemples à valeur diminutive. Nous nous opposons à ce point de vue, en ce que la valeur augmentative et diminutive (par le biais

²² La terre est à la fois passive et productive. Elle reçoit la semence et la fait croître.

de suffixes) n'est généralement pas un effet du genre²³, mais le résultat de processus de dérivation métaphorique et métonymique, processus qui ne sont pas motivés de prime abord par le genre.

Zubin et Köpcke (1984:51) ont observé en allemand que les genres masculins et féminins, dans le suffixe *-mut* (*der Hochmut* « l'arrogance », *die Demut* « l'humilité », *die/der Kleinmut* « la petitesse d'esprit »), ont des valeurs respectives d'extroversion et d'introversio. Le masculin (ou l'extroversion) connote l'agressivité, le rejet, la distance et l'égoïsme, alors que le féminin (ou l'introversio) connote la soumission, l'acceptation, l'accessibilité, la vulnérabilité et l'altruisme. Nous ne croyons pas que ces valeurs puissent être transmissibles au système du français, parce qu'elles sont probablement propres au suffixe *-mut*.

Audet et Ouellet (1984:47-51; et Malkowska (2004) dans leurs traces) ont étudié la valeur sémantique du genre et ils sont venus à la conclusion que chaque genre déterminait un mode d'existence différent, objectif pour le masculin inanimé et relatif pour le féminin inanimé (p. 49). Pour eux, « [c]onçu féminin, le désigné apparaît comme le lieu virtuel de la désignation, un lieu externe indéfini, non limité ou non délimité, abstrait ou général » (p. 47), alors que « conçu masculin, ce même désigné apparaît comme le lieu actuel de cette désignation, un lieu interne, défini ou délimité, particulier ou concret » (*ibid.*). Le genre masculin inanimé et le genre féminin inanimé mettent en évidence différentes oppositions telles que l'effectif et le virtuel, l'objectif et le relatif, le particulier et le général, le défini et l'indéfini (p. 51). Malkowska ajoute à ces oppositions celle de l'authentique et du semblable (p. 76). Si ces valeurs peuvent justifier le genre de paires minimales sémantiques et morphologiques (telles que *plant/plante* et *point/pointe*), nous ne voyons cependant pas comment elles peuvent expliquer la variation du genre que nous observons dans de nombreuses unités qui réfèrent aux mêmes réalités avec les mêmes concepts. Par exemple, *le minestrone* et *la minestrone* réfèrent les deux au même type de soupe, mais l'appellation porte surtout les MAM en Europe et les MAF au Québec.

Nous sommes d'avis qu'Audet et Ouellet, tout comme Hjelmlev et Togeby, réussissent assez bien à justifier le genre d'unité dans la variété standard du français, mais que leurs explications sont défectives ou incomplètes lorsqu'on étudie la variation du genre de la même unité lexicale. Pour cette raison, nous avons décidé de reprendre partiellement les valeurs sémantiques qui nous semblaient correspondre à la différence de conceptualisation que nous observions préalablement. Nous listons ci-dessous ce qui nous semble généralement applicables selon les auteurs mentionnés plus haut.

²³ Car il existe de nombreux suffixes diminutifs masculins (*-et* dans *archet*, *-ot* dans *îlot*) auxquels correspondent des suffixes diminutifs féminins (*-ette* dans *maisonnette*, *-ote/otte* dans *pelote* ou *jugeotte*) ainsi que de nombreux autres suffixes asymétriques des deux genres (*-ule* dans *globule*, *-elle* dans *ruelle*). Dans un grand nombre de cas, le genre est identique à celui du lexème simple (ici *une maison*, *un globe*, *une rue*), mais on peut apporter à cela des contre-exemples, tels qu'*un îlot* dérivé d'*une île*, *une charrette* dérivé d'*un char* et *un bottillon* dérivé d'*une botte*.

- Masculin spécifique : passivité (Togebly et Hjelmslev), existant concret (Lafont), actuel ou effectif, défini, délimité, particulier, concret, objectif (termes d'Audet et Ouellet);
- Féminin spécifique : expansion (Togebly et Hjelmslev), activité (Togebly), acte (Lafont), virtuel, externe, indéfini, non limité, non délimité, abstrait, général, relatif (termes d'Audet et Ouellet).

Nous percevons dans les genres spécifiques un dynamisme qui ne semble pas avoir été traité pour les entités inanimées²⁴. Par exemple, l'opposition *le voile* « morceau de tissu porté sur le corps ou sur la tête » et *la voile* « morceau de tissu qui fait avancer un navire » nous semble manifester un dynamisme dans le deuxième cas. Ainsi, le féminin est conçu comme actif, productif, dynamique (car la voile contribue au mouvement), alors que le masculin est conçu comme statique, délimité, car il n'engendre pas le mouvement²⁵. Nous croyons que ce dynamisme peut être interne ou externe pour le féminin (mouvement vers l'intérieur ou l'extérieur) alors que le masculin aura tendance à marquer des entités passives, statiques. Donnons l'exemple de *plant-plante*. On conçoit *plant* comme statique, atomique (indivisible) ou terminal, comme une simple unité comptable sans autre fonction, tels que les exemples suivants peuvent le montrer :

1. Les autorités ont saisi trente-cinq plants de cannabis.
2. Il y a environ mille plants dans ce champ.

À l'inverse, *plante* manifeste une productivité, une activité que sa contrepartie masculine ne manifeste pas, tel que les combinaisons suivantes le démontre : *une plante grimpante, une plante carnivore, une plante toxique, une plante médicinale*. *Machin* et *machine* montrent le même genre de dynamisme : *machin* renvoie à un simple objet indéfini, statique et terminal (sans usage particulier), mais *machine* comporte un dynamisme ou une productivité, car l'unité renvoie à une série de mécanismes ayant un ou des usages particuliers. *Bas* et *base* suivent cette opposition : la *base* renvoie à la partie soutenant une structure, fournissant (produisant) la stabilité et la force de celle-ci, alors que *bas* renvoie simplement à une partie statique, passive et atomique d'une structure.

Ajoutons maintenant les nouvelles valeurs que nous attribuons aux genres masculin et féminin spécifiques à celles que nous avons précédemment retenues.

²⁴ Audet et Ouellet (1984:52) parlent de mode d'activité (ou d'agir) défini ou indéfini, effectif ou virtuel, particulier ou général, mais uniquement dans le cas des entités conçues comme animées. Nous étudions ici les signes à référent inanimés. Togebly a donné une très courte liste d'unités avec exemples, la majorité étant morphologiquement complexe. Lafont s'est intéressé au latin et non pas aux unités du français.

²⁵ On retrouve les mêmes conceptions lorsqu'on observe les combinaisons avec *voile*. *Mettre le voile sur quelque chose*, c'est cacher une réalité, l'empêcher d'accomplir son potentiel (activité), alors que *mettre les voiles* énonce une activité, un départ vers l'extérieur. Cet exemple avec *voile* nous a été suggéré par Bruno Courbon.

Ajoutons que tout objet qui crée un mouvement n'est pas nécessairement féminin. Ainsi, *foe* et *spinnaker* portent la MAM, malgré qu'ils soient des types (très spécifiques) de voiles de bateau. Leur genre pourrait s'expliquer notamment par un paradigme sémantique presque entièrement masculin de type de voiles de bateau, auquel on ajouterait également *yankee, génois, tourmentin, clinfoc, trinquette* (féminin), *solent, cacatois, perroquet, tapecul, artimon* et *gennaker*.

- Masculin spécifique : statique, passif, atomique (indivisible), terminal, actuel/effectif, défini, délimité;
- Féminin spécifique : dynamique (intérieurité ou extériorité, concentration ou expansion), actif, virtuel, indéfini, non limité ou non délimité.

Observons quelques autres paires afin de voir comment le masculin représente la conception atomique, alors que le féminin illustre cette activité ou productivité dont nous faisons état :

1. Le chlore et le sodium sont les **composants** du sel.²⁶ / Es-tu bien sûre d'avoir analysé toutes les **composantes** du problème?²⁷
2. Il a fait un **point** avec la **pointe** de son crayon.²⁸
3. Il a planté ce nouveau **fraisier** dans sa **fraisière**.
4. La **garde** nationale défend le palais; un **garde** est posté à cette entrée.
5. Ce **marin** s'est engagé dans la **marine** marchande.
6. Faire une **passe**. / Faire un **pas**.
7. Le **cours** d'une vie. / La **course** contre la montre.²⁹
8. Le **jour** et la nuit. / J'ai travaillé toute la **journée**.
9. Il a douze **ans**. / Je vous souhaite une bonne **année**.

Commentons l'exemple (7). *Cours* est davantage statique dans les combinaisons *cours d'eau*, *reprendre son cours*, *au cours de sa vie*, *un cours de français*; *course* est inversement plus actif : *la course à pied*, *faire ses courses*, *regarder la course de chevaux*, *arrêter quelqu'un dans sa course*. Si les deux unités comportent un sème de mouvement, celui de *course* semble plus puissant, pressant ou exacerbé, alors que celui de *cours* semble plus lent ou constant. La durée de l'action semble plus délimitée (incluant une fin de l'action) au masculin qu'au féminin. N'oublions pas que nous tentons de faire une approximation du sens avec des traits plus ou moins actualisés ou présents dans chaque opposition. Nous ne pensons pas que seul le dynamisme (voir son aspect sémantique entier) est en mesure de distinguer ces unités (*cf.* notamment Audet et Ouellet 1984:50).

Les exemples (8) et (9) nous montrent que la valeur dynamique du féminin n'est pas contradictoire avec l'usage de suffixes. Le suffixe *-ée* est un tel exemple. L'opposition du masculin et du féminin dans le morphème *-ier/ière* semble marquer la même opposition. Pour cette paire, le choix du morphème de genre dépend donc de la valeur sémantique désirée. À partir d'autres suffixes, tels que *-et/ette* (*briquet/briquette*), *-ant/lante* (*collant*, *volant*, *contenant* et *imprimante*, *variante*) et *-teur/trice* (*distributeur/distributrice*), plusieurs substantifs pourraient également trouver leur motivation dans le choix de la valeur sémantique associé au morphème de genre.

²⁶ http://bdl.oqlf.gouv.qc.ca/bdl/gabarit_bdl.asp?id=4416, consulté en mars 2015.

²⁷ *Ibid.*

²⁸ Ouellet (2014:252). Les exemples suivants proviennent tous du même ouvrage, soit les pages 251, 252, 261 et 262.

²⁹ Les exemples qui suivent sont nôtres.

L'aspect du dynamisme que nous avons souligné ci-haut n'est pas contradictoire avec la théorie d'Audet et Ouellet (1984) et Ouellet (2014). Au contraire, il semble il y avoir plus d'affinités entre le défini, le délimité et le statique qu'avec le dynamique. Le féminin dynamique engendre un mouvement qui ne saurait être défini ou délimité. La virtualité du féminin ne fait qu'accentuer le mouvement sans limite du féminin dynamique, alors que l'effectivité du masculin circonscrit davantage l'unité comme terminale, statique et concrète.

Un dernier commentaire sur l'aspect sémantique des genres spécifiques est qu'il est normal que le féminin porte plus souvent des connotations d'affectivité que le masculin : cela n'a rien à voir avec le sexe [+femme] ou le *gender* correspondant, mais plutôt avec le fait que seul au féminin est associé une valeur de productivité, et donc, seul ce genre spécifique peut produire ces effets. C'est à cause de son activité, de sa générativité, de sa matricialité s'il produit des effets affectifs. Un exemple concret serait le cas des insultes : s'il est possible de traiter un homme de *chien* et une femme de *chienne*, il est possible de traiter un homme de *chienne*, mais pas une femme de *chien*. Les noms d'insultes sont plus souvent féminins que masculins : *pute, putain, grue, oie, poule, chienne, bitch, vache*. On ne verra pas aussi fréquemment *gigolo, jars, coq, bœuf*. Ce constat n'a pas été relevé dans *Les gros mots* de Guiraud (1991)³⁰.

2.2.2 L'aspect sémantique de l'archigenre

Notre postulat théorique séparant l'archigenre des deux genres spécifiques se base sur le type de conceptualisation associé à chaque concept. Si les deux genres spécifiques avaient des valeurs sémantiques propres à un type de conceptualisation, nous sommes d'avis que l'archigenre n'a pas de valeur sémantique autre qu'une idée de vague ou d'imprécision quant à l'identité de la réalité. C'est ce manque de conceptualisation spécifique qui distingue l'archigenre des genres spécifiques. L'expression de l'archigenre est caractérisée par une visée non-spécifique du locuteur, par une conceptualisation vague (Zubin et Köpcke 1986:144, 146)³¹. Les unités ainsi conçues sont peu nombreuses dans le lexique indigène du français. La plupart des unités sont clairement définies dans l'esprit des locuteurs, à l'exception des concepts nouveaux pour le locuteur (néologismes individuels), des concepts génériques (aussi appelés *superordonnés*, voir section 2.3.2 (notamment la note 55), ne pas confondre avec référence générique comme l'indique la note 17) ou des emprunts. Ces concepts ont généralement une conceptualisation plus vague ou abstraite dans l'esprit

³⁰ Quelques exemples animaliers donnés sont *moule, vache, chameau et cochon*, mais l'auteur ne regarde pas comment le genre ajoute une valeur à l'énoncé. On note tout de même plusieurs significations de noms d'animaux renvoyant à une « femme méchante, bête, laide », « grossière, malpropre », « dévergondée, prostituée » (Guiraud 1991:60, 86-89).

³¹ Ajoutons que cette conceptualisation vague et générale d'une référence archigénérique va à l'encontre d'Audet et Ouellet (1984) en ce que le masculin devrait refléter un ensemble défini ou concret. C'est pour cette raison que nous séparons cette conception générique du masculin spécifique.

des locuteurs. Nous considérons que la différence entre l'archigenre et le masculin spécifique s'opère sur une base conceptuelle et non sémantique (significations) ou sémiologique (forme physique)³².

Dans la majorité des cas, l'archigenre est, spontanément et dans la quasi-totalité des cas, exprimé par la MAM. Cependant, puisqu'il est impossible de regarder directement la conceptualisation d'un concept pour un locuteur, il nous faut considérer que chaque MAM peut représenter l'archigenre ou le masculin spécifique.

Considérons :

10. L'**homme** est mortel.
11. Une pomme par jour éloigne le **docteur** pour toujours.
12. Le **chien** est le meilleur ami de l'homme.

En utilisant les lexèmes *homme*, *docteur* et *chien* des phrases (10), (11) et (12), le locuteur n'a pas en tête un homme, un docteur ou un chien particulier, mais le concept même, dans son interprétation la plus générale, d'un humain, d'un docteur ou d'un chien parmi tant d'autre (Hirtle 2007:108). Le lexème *homme* peut également être utilisé dans une conceptualisation davantage spécifique, pour référer à l'ensemble des individus de sexe masculin que le concept d'*homme* représente, comme en (13), ou pour renvoyer à un individu en particulier, comme en (14) :

13. L'**homme** et la femme³³.
14. Le **docteur** m'a dit de faire plus d'exercice.

Il est possible d'observer des cas où la MAF a une valeur sémantique générique, mais ceux-ci renvoient généralement à la référence sexuée³⁴, comme en (15) :

15. La journée internationale de la **femme**.

Cependant, la référence au sexe constitue un niveau de spécificité supérieur à celui de l'archigenre (sans référence au sexe), dont la visée est non-spécifique. Nous pouvons illustrer en exemples les lexèmes *homme* et *femme* ainsi :

³² Le concept d'archigenre n'est pas une tentative de réinventer la roue ou de faire une révolution dans le domaine du genre. Il s'agit plutôt de fournir une explication sur la prédominance générale (et généralisée) de la MAM dans le cas des emprunts. Nous pensons que le concept d'archigenre peut être applicable aux emprunts comme aux termes génériques (donc ayant une conceptualisation plus vague ou générale). Pour les termes plus génériques, voir les sections 2.3.2 (et la note 55) et 3.1 (et la note 64).

³³ On peut voir cet exemple comme représentant l'ensemble des hommes et des femmes, ou un seul homme et une seule femme. Nous appelons ces différents niveaux *infragénériques* et *spécifique*. Pour plus d'informations, sur les niveaux de généralité, voir la figure 3 plus bas.

³⁴ Ce qui pourrait être expliqué par le postulat de Michard (2003:45^e paragraphe), pour qui les noms désignant des hommes ont d'abord le trait [animé] avant d'avoir le trait sexué [homme], tandis que les noms désignant des femmes ont d'abord le trait sexué [femme] avant d'avoir le trait [animé].

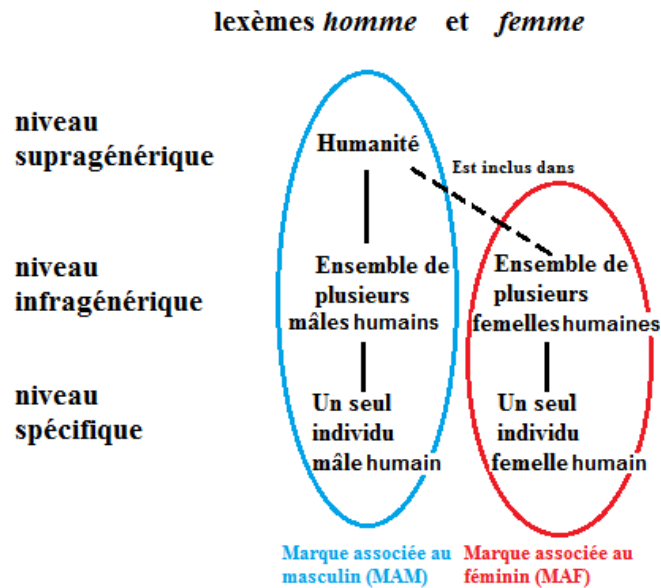


Figure 3 : niveau de généricité des lexèmes *homme* et *femme*

L'archigénre pourrait être considéré comme le représentant morphosémantique du niveau supragénérique. Il ne faut pas confondre les niveaux de généricité d'un même signe (d'un même signifié) avec les sèmes *microgénériques*, *mésogénériques* et *macrogénériques* de Rastier (1987:49-50). Ces termes s'appliquent respectivement à un *taxème* (*grosso modo* un paradigme lexical envisagé du point de vue sémantique), à un *domaine* (groupe de taxèmes) et à une *dimension* (« classe de généralité supérieure », p. 50), donc au minimum un regroupement de signes et non pas à travers les emplois d'un seul signifié. De plus, selon Rastier (1987:52), « aucun sème n'est donc par nature spécifique ou générique ». Les sèmes génériques ont pour fonction de rapprocher deux sémèmes (« contenu d'un morphème », p. 275), alors que les sèmes spécifiques ont comme fonction d'opposer deux sémèmes. Afin de nous dissocier de cette vision globale de la généricité, et pour n'étudier que la généricité des emplois d'un signe, nous conserverons la terminologie de niveaux *infra-* et *supragénériques*.

Si on peut rapprocher les niveaux infragénérique et spécifique à la notion d'extension (« qui opère la représentation des notions qui sont à la base de tous les concepts, ce qui met en cause l'évocation d'un ensemble de valeur partielle ou intégrale » (Ouellet 2005:169)), il ne faut pas confondre l'extension avec l'archigénre. Les exemples (10), (11) et (12) examinés ci-haut emploient les concepts dans leur plus grande extension, comme englobant tous les individus des groupes *homme*, *docteur* et *chien*. Les concepts d'*homme*, de *docteur* et de *chien* sont généralement clairement définis à l'esprit des locuteurs du français et ne sont probablement pas sémantiquement génériques au sens archigénérique.

Si l'on considère que le lexème *homme* est sémantiquement symétrique à celui de *femme*, il faut considérer son niveau supragénérique comme l'archigenre. Or, *homme* « humanité » ne peut être considéré comme l'extension maximale de l'unité lexicale *homme* qu'à condition de considérer la paire *homme-femme* comme sémantiquement (référentiellement) asymétrique. Advenant cela, *homme* « humanité » apparaît comme une unité lexicale à plus grande extension que *femme* dans son extension la plus grande. Là où l'exemple (13) supposait une extension similaire (restreinte ou large), l'exemple (10) permet l'interprétation d'*homme* sur toute son extension possible (de « cet homme-ci est mortel » à « l'humanité est mortelle »). L'unité *humanité* « ensemble des individus de la race humaine », bien que féminin et référant au même groupe d'individus que l'exemple (10), ne peut être considéré comme archigenre, non seulement puisqu'il porte la MAF, mais également parce qu'il est probablement défini clairement à l'esprit de tous les locuteurs. Si *homme* peut (selon les analyses) être associé à l'archigenre, *humanité* ou *femme* ne le peuvent pas en raison de leur conceptualisation plus spécifique que celle de l'exemple (10). Les exemples suivants démontrent la conceptualisation plus spécifique de ces deux unités :

16. Il aura marqué l'**humanité** toute entière.
17. Il a perdu son **humanité**.
18. Il a fait preuve de beaucoup d'**humanité**.
19. Le lycée a favorisé les sciences au détriment des **humanités** classiques³⁵.
20. Elle ne s'était jamais sentie aussi **femme**.
21. Il a pris **femme**, elle a pris mari.
22. Un professeur **femme**.

Puisque l'extension est seulement actualisée en discours, le terme même d'extension ne nous permet pas d'évaluer les différentes significations sans un contexte particulier. Afin de palier cela, nous avons donc choisi les termes *supragénérique*, *infragénérique* et *spécifique* pour référer à la visée « archigénérale » (maximale, de puissance), générale ou spécifique d'un signe qu'a le locuteur.

Nous considérons donc que l'archigenre s'actualise le plus souvent par la marque associée au masculin. Les exemples donnés entre (10) et (14) sont tous des exemples d'animés humains. Le cas de l'archigenre hors des noms animés n'a pas été vraiment étudié. Cependant, nos observations nous permettent de proposer que l'archigenre devrait pouvoir s'appliquer également aux inanimés, puisqu'il a la fonction de marquer la conceptualisation vague ou générale que les locuteurs ont des réalités, qu'elles soient animées ou inanimées.

On voit désormais l'intérêt de cette conception du genre en lien avec le phénomène de l'emprunt. Lorsqu'on emprunte un mot et qu'on a une représentation générale (parfois presque non existante) de la réalité de

³⁵ <https://fr.wiktionary.org/wiki/humanité> (Consulté en mai 2016), Marie Duru-Bellat, *Le bac, un standard bicentenaire*, Propos recueillis par Géraldine Véron, Le journal du CNRS, juin 2008

l'emprunt, il serait logique de penser qu'on puisse attribuer majoritairement l'archigendre, à moins que le locuteur n'en ait eu l'expérience suffisante pour en avoir une conceptualisation spécifique, auquel cas le locuteur décidera inconsciemment si la référence correspond plus au masculin spécifique ou au féminin spécifique. Ce phénomène est attesté dans la littérature, en ce que la marque du masculin (comme générique ou spécifique, les auteurs ne distinguent pas) est toujours la marque majoritaire dans le cas des emprunts à l'anglais³⁶.

2.2.3 Le genre et l'animéité

Il reste primordial de dissocier le genre, marqué sémiologiquement par le masculin et le féminin, de l'animéité, non marquée sémiologiquement³⁷, mais distinguant par référence logique l'animé et l'inanimé (Audet & Ouellet 1984). L'appellation *animé* fera référence dans notre travail à la notion d'*animé humain*. Si, selon Hirtle (2007:107-111), le genre grammatical anglais est un sous-système du système d'animéité anglais, le genre et l'animéité en français sont consubstantiels. Voici quatre exemples illustrant cette existence commune, tous provenant de Ouellet (2014:249) et démontrant qu'un seul signe peut être utilisé comme animé ou comme inanimé, peu importe son genre :

- Masculin animé (humain) : Cet enfant obéit à son **tuteur**.
- Masculin inanimé : Un **tuteur** soutient cette plante.
- Féminin animé (humain) : Elle affronte la **partie** adverse.
- Féminin inanimé : Il joue une **partie** d'échecs.

Nous concentrerons nos recherches sur les genres masculins et féminins inanimés, afin d'éviter que les substantifs désignant des entités animées puissent donner leur genre par le biais d'une référence sexuelle, malgré leur conceptualisation différente des noms inanimés³⁸.

2.3 Sur l'attribution du genre en français

Avant notre travail de recherche, quelques auteurs ont cherché la motivation du genre des emprunts, le plus souvent à l'anglais. Soulignons ici ceux qui nous ont inspiré à développer sur le sujet : Arndt (1970) sur les emprunts à l'anglais en allemand; Correa-Zoli (1973) sur les emprunts à l'anglais en italien américain; Poplack, Pousada et Sankoff (1982) sur les emprunts à l'anglais en espagnol de Puerto Rico et en français

³⁶ Poplack *et al.* (1982:21) 78% de masculin; Guilford (1999:67, 78, 80) 81%, 84% et 81.9%, 76.73% pour tout son corpus, Raymond (2011:74) 87%, Nymansson (1995:107) 90.6%; Humbley (1974:67) près de 90% (pas de chiffre exact).

³⁷ Sauf dans quelques paires, notamment *qui/quoi*.

³⁸ Pour les inanimés, on dira que la référence correspond à un lieu de désignation soit virtuel, externe indéfini, non limité ou indélimité, abstrait ou général (pour le féminin), soit actuel, interne, défini ou délimité, particulier ou concret (pour le masculin) (Audet & Ouellet 1984:47). Pour les animés, le genre correspondra au masculin si « la désignation implique un mode d'activité bien défini, une fonction d'agent spécifique et particulière », ou au féminin si la désignation ne « définit pas une fonction d'agent particulière et implique un agir indéfini » (*ibid.*).

montréalais (première étude multifactorielle); Surridge (1982 et 1984), Nymansson (1995) et Guilford (1999) sur les emprunts à l'anglais en français; Lafourcade (2009, mémoire de maîtrise) sur les emprunts à l'espagnol d'Amérique du Sud en français; Raymond (2011, mémoire de maîtrise) sur les emprunts à l'anglais en français franco-ontarien de Sudbury. Poursuivons maintenant à l'attribution du genre au lexique du français.

Il existe trois points de vue sur l'attribution du genre : celui des auteurs qui soutiennent la thèse de l'arbitraire³⁹, qui sont d'avis qu'aucune règle ne régit l'attribution du genre des noms inanimés, celui de linguistes qui sont d'avis que c'est la forme physique du mot qui détermine le genre⁴⁰, et celui d'autres linguistes qui sont d'avis que l'attribution se base surtout sur des facteurs sémantiques⁴¹.

Afin de respecter cette séparation théorique, et pour suivre la distinction (d'inspiration hjelmsléviennne) faite par Ibrahim (1973:52-53) entre les processus de « forme » [physique : phonétique, phonologique, graphique] ou de « contenu » [sémantique, référentiel, morphologique], nous avons examiné l'attribution du genre à travers des critères d'attribution du genre de ces deux ordres, tout en observant, dans quelle mesure l'archigenre est attribué ou le genre d'origine est conservé.

Selon Surridge (1985-1986), l'**analogie morphologique**, qui consiste à donner à un emprunt le genre d'un mot qui a le même morphème final (*kharidjisme* serait masculin à cause de sa finale en *-isme*, toujours masculine), est le critère d'attribution le plus efficace, et c'est pour cela que nous n'avons examiné que des mots morphologiquement simples. Les seuls critères morphologiques considérés sont donc la réanalyse morphologique, dont le résultat se nomme parfois « faux morphème », et, lorsqu'il s'agit d'emprunts d'une langue à genre, la conservation du genre de la langue donneuse.

2.3.1 Critères de forme physique

Les critères tenant de la forme physique⁴² du signe seront décrits dans cette section. Le premier est la **syllabicit  finale** et se base sur la terminaison syllabique, dictant qu'un certain type de finale sera associ    un genre particulier. Pour Harley (1979) et Carroll (2005), si un signe se termine par une syllabe ouverte (voyelle finale, CV), il a davantage tendance    tre masculin; s'il se termine par une syllabe ferm e (consonne finale CVC), il sera davantage f minin (ex : *fran ais-fran aise*, *allemand-allemande*, *pas-passe*, *machin-*

³⁹ Grevisse (1964, 1980, 1986, 2008), Corbett (1991), Andriamamonjy (2000:423), Lowe (2007), De Villers (2009), Riegel, Pellat & Rioul (2011:170).

⁴⁰ Tucker *et al.* (1970, 1977), Surridge (1985, 1986), Carroll (1989, 2005), Lyster (2006).

⁴¹ Damourette & Pichon (1968 [1911]), Surridge (1982, 1984), Audet & Ouellet (1984), Ouellet (2014).

⁴² Nous avons choisi l'appellation *forme physique*, car elle nous semblait plus juste que *forme* utilis e seule. En effet, la notion de *forme* peut  tre utilis e dans une vari t  de contextes, tels que *forme s mantique* et *s mantique formelle*. Nous n'utiliserons pas l'opposition *s mantique-s miotique* de Benveniste (1974:21), car, selon notre terminologie, *s miotique* r f re   *signe*, *signe*  tant la combinaison d'une forme physique et d'une ou plusieurs significations. Par *forme physique*, nous entendons   la fois une s rie de phon mes ou de graph mes ainsi que la repr sentation mentale de la forme sonore ou graphique.

machine). Pour leur part, Tucker, Lambert et Rigaud (1977:36-37) ont mené une étude analysant la finale phonétique d'un mot et la corrélation de celle-ci avec le genre du mot et ils y ont suggéré que les mots se terminant par une voyelle sont plus souvent masculins, sauf s'ils se terminent par /i/ ou /ĩ/, auxquels cas ils seraient davantage féminins. La finale /e/ n'a pas permis d'identifier un genre plus productif. Les résultats de Tucker *et al.* (1977), de Harley (1979) et Carroll (2005) tiennent de la forme phonétique. Or, puisque nous étudions des signes graphiques comme unité d'étude, il faut inverser cette constatation pour que la finale vocalique (graphique) soit féminine (*française, allemande*) et non pas masculine. Cela suit également la « règle » générale dictant qu'il faut ajouter un –e à un mot français pour le rendre masculin⁴³.

Le deuxième critère est la **syllabicité entière** et se base sur le nombre de syllabes phonétiques dans le signe. Selon Surridge (1982:37), les emprunts anglais monosyllabiques sont plus souvent féminins en français québécois qu'en français européen. Au Québec, les emprunts anglais monosyllabiques porteraient environ autant la marque du masculin que celle du féminin, alors qu'en Europe, la marque du masculin prédominerait à 95%.

Le troisième critère est **similitude graphique** d'un signe avec la graphie d'un autre signe (Arndt 1970). Cette similitude peut être entière (comme *chip*₁ « puce électronique » et *chip*₂ « croustille ») ou partielle (*gig* « spectacle musical » et *gigue* « danse »). Pour des raisons méthodologiques, les homographes complets ont été retirés de l'étude afin d'éviter un risque de contamination du genre en français de la forme homonymique (ou identique sémiologiquement).

Le quatrième critère est la **finale graphique** du mot (Lyster 2006:85). Lyster a montré que, pour les noms indigènes au lexique du français, la finale graphique permet de prédire le genre avec 98% d'efficacité. Les emprunts ont été comparés à ce modèle en tenant compte des diacritiques en finale de mot.

Nous avons séparé en plusieurs critères ce que plusieurs auteurs appellent l'*analogie phonétique* (Surridge 1984:59, Nymansson 1995:99, Guilford 1999:68), car elle représentait différents procédés. Le cinquième critère est l'analogie phonétique (abrégée AP) **par phonème final**. Malgré que les mots morphologiquement simples ne furent pas séparés des morphologiquement complexes de leur corpus, nous nous sommes basé sur les chiffres de Tucker *et al.* (1970, 1977), et nous avons comparé la finale phonétique de chaque unité avec le genre correspondant.

⁴³ Fréquemment enseignée aux apprenants du français langue seconde. Nous ne commenterons pas l'hypergénéralisation que cette règle apporte.

Le sixième critère est l'analogie phonétique **par phonème pénultième**. Tucker *et al.* (1977) ont démontré⁴⁴ que le phonème pénultième (voire antépénultième) joue un rôle dans l'attribution du genre, notamment dans les finales en [ɔ̃] qui sont presque exclusivement féminines lorsque précédés de [j] et masculines dans les autres contextes.

Le septième critère est l'analogie phonétique **par rime riche**. Haugen (1953:444-445) a démontré qu'on donne un genre à des emprunts parce qu'ils ont des finales phonétiques identiques à d'autres mots du lexique du français. Basé sur ce modèle, ce critère a été utilisé conjointement avec un dictionnaire de rimes (Warnant 2007) et le genre des mots avec le plus de phonèmes similaires (au minimum trois) sera comparé.

Le huitième et dernier critère est la **paronymie**. Ce critère compare le côté sonore (physique ou mental) d'un signe à un autre lui ressemblant et assigne au premier signe le genre du second. Pour la paronymie, nous avons établi des critères de similitude n'exigeant pas une voyelle identique, mais une forme syllabique similaire. Il fallait également que tous les phonèmes soient identiques, sauf un.

Il existe un dernier critère de forme physique (et dans une certaine mesure, sémantique) et il s'agit de ce que Surridge (1985-1986) appelle l'*analogie suffixale* ou *morphologique* (où l'emprunt prend le genre de mots qui ont la même finale suffixale). Cependant, puisque nous n'incluons pas dans notre échantillon à l'étude des noms comportant un suffixe, nous avons analysé un critère différent. Reichler-Béguelin (1993:2.2.2) a démontré que les locuteurs peuvent remotiver un signe selon les connaissances métalinguistiques qu'ils possèdent et assigner des propriétés ou une étymologie que l'unité ne possède pas a priori. Nous postulons que ce qu'on peut appeler « faux morphème » ou « faux suffixe » est en fait le résultat d'une réanalyse d'un signe (que nous appelons **réanalyse morphologique**) où le locuteur assigne un sens à une finale de forme identique (phonétique et/ou graphique) à un morphème existant dans la langue. Un exemple attesté dans la littérature serait le fait que les emprunts (à l'arabe) se terminant en /a/ soient féminins, puisqu'il existe un morphème similaire dans la langue cible (Badarneh 2007:63). De plus, la finale graphique *-um* de l'emprunt *drum* pourrait être réinterprétée et réanalysée en la finale *-um* de mots tels que *consortium* ou *référendum*. De même (pour un préfixe), la forme *bikini* a donné *monokini*, car le *bi-* a été réinterprété comme le préfixe *bi-* dans *bisexuel* ou *bidimensionnel*. On retrouve d'ailleurs, au Québec, la forme *burkini* (ou *burqini*) formé de *burqa* et *bikini*, forme montrant avec quelle facilité les locuteurs jouent avec la morphologie de leur langue. Notons que la « base » (ce qui n'est pas réinterprété, *dr-* dans *drum* et *-kini* dans *bikini*) n'a pas besoin d'être a priori signifiante; si le locuteur décide que la finale est signifiante, c'est suffisant pour assigner un genre⁴⁵. Si

⁴⁴ Avec leurs tableaux en annexe I et II, p.67 à 125.

⁴⁵ Il y a plusieurs morphèmes en français, notamment *-oune* en français québécois, qui ne se collent pas à des bases transparentes, par exemple, *ti-coune*, *toutoune*, *nounoune* (et *noune*), *poupoune*, *foufoune*, *guidoune*, *pitoune*, *floune* (et *sloune*), *poune*, *gougoune*, *moumoune*, *minoune*, *balloune*, etc., qui ont tous une valeur de proximité ou de péjorativité

la base était immédiatement signifiante, l'emprunt n'aurait pas été retenu pour l'étude, parce que nous l'aurions considéré comme un nom morphologiquement complexe. La transparence (ou sémanticité manifeste) d'une base n'est pas nécessaire pour qu'un nom (réanalysé comme morphologiquement complexe) puisse intégrer le paradigme d'un morphème, tels que le sont les bases opaques suivantes : *-stit-* dans *constitution* et *céc-* dans *cécité*. Ces bases sont décrites comme *non autonome* par Mortureux (2008:24). Pour conclure, la base n'a pas à être signifiante seule, car elle devient signifiante avec la finale. Puisque les morphèmes (vrais ou faux) ont une sémiologie à la fois orale et écrite, nous avons considéré le faux morphème sous ces deux formes physiques.

2.3.2 Critères sémantiques

Les critères basés sur la ou les significations des emprunts seront examinés dans cette section. De nombreux auteurs parlent d'un processus d'attribution du genre nommé « analogie sémantique » (Humbley 1974; Poplack *et al.* 1982; Surridge 1982-1984; Nymansson 1995). Les auteurs ne sont pas tous d'accord sur le fonctionnement de ce processus. Le plus souvent, *analogie sémantique* réfère au fait qu'un emprunt prend le genre d'un mot indigène signifiant approximativement la même chose. Ainsi, *bike* serait masculin à cause de *vélo*, et non de *bicyclette*, malgré la plus grande ressemblance de la forme physique phonétique. Cette approche d'un mot de référence, appelé *homologue* (Humbley 1974, Surridge 1984), *parasynonyme* (Nymansson 1995) ou *next lexical equivalent* (Stolz 2009), est à notre sens trop restrictive. Il nous semble plus probable que l'emprunt intègre un paradigme conceptuel que nous appellerons *paradigme latent*, paradigme auquel le locuteur fait correspondre une nouvelle unité et attribue à ce néologisme le genre majoritaire ou le plus représentatif (donc *bike* est intégré dans le paradigme avec entre autres *vélo*). Nous postulons que le paradigme latent⁴⁶ est un réseau virtuel de concepts comportant une quantité indéfinie de signes à conceptualisation similaire et comportant plusieurs sèmes génériques et spécifiques communs, supposant un certain degré de substituabilité des éléments qui le composent⁴⁷. Il se situe à l'intersection du niveau conceptuel (notionnel, donc abstrait, socialisé et peu formalisé) et du niveau lexical (linguistique à finalité référentielle, donc plus concret, mais quand même mental, et formalisé sous forme de signes). Nous postulons que le paradigme latent est individuel, donc non partagé par la communauté, car il est construit par l'expérience du locuteur. Pour un individu ayant plus de connaissances dans un domaine donné, son

(Laforest et Vincent 2004:65, note 5). Nombre de noms indigènes au lexique français qui se terminent avec une finale identique à un suffixe, mais dont la « base » n'est pas porteuse de signification, ont le même genre que la finale commune : *savon*, *poisson*, *email* (Huot 2012:50). On peut facilement reconnaître *-on* et *-ail*, et pourtant la « base » seule n'est pas sémantisable.

⁴⁶ Nous devons cette notion de latence lexicale à Courbon (2015a:190-191) pour qui les formes réduites (par exemple *télé* plutôt que *télévision*) sont bien souvent utilisées avec les formes longues « présente[s] en filigrane [...] dans un paradigme dénominateur ».

⁴⁷ Le paradigme lexical est généralement une série ouverte de concepts, mais on compte quelques séries fermées pour des raisons socioculturelles, tels que les noms de saisons ou de points cardinaux, ces derniers limités au nombre de quatre éléments en français.

paradigme latent sera plus développé que celui d'un non-initié, d'où une potentielle variation du genre chez l'individu. Nous appellerons le processus d'intégration d'un néologisme dans un paradigme *intégration paradigmatique*.

Au sein d'un paradigme, certains signes ou concepts peuvent être plus saillants que d'autres et ainsi imposer leur genre sur le reste du paradigme. Dans certains cas, l'unité la plus influente est un hypéronyme ou un nom générique qui détermine le genre du nouveau concept intégré, comme dans le cas du paradigme des arbres presque tous masculins, dominés par le concept ARBRE. Dans d'autres cas, il s'agit d'un nom déterminant l'extension de l'emprunt (un spécificateur, comme le cas de *café latte*, où *latte* vient spécifier *café*) où le spécifié (*café*; qui tombe et qui restera latent au moins temporairement) donne le genre à l'emprunt spécificateur (*latte*). Parfois, aucun concept ne prime sur l'ensemble des autres par sa saillance. Puisqu'il n'est pas toujours possible de prévoir l'influence et la saillance de chaque nom d'un paradigme, nous avons tenu compte de chaque possibilité dans notre analyse.

Souvent, il y a emprunt parce qu'une langue doit dénommer une réalité nouvelle. C'est le cas, parmi tant d'autres, de *sushi*. Ne possédant pas de correspondant référentiel identique (à l'instar de *bike* qui correspond entre autres à *vélo*), le paradigme latent devient un critère sémantique indispensable d'attribution du genre. Sans le paradigme latent, ce serait dire que les emprunts sans correspondant ne peuvent pas trouver la motivation de leur genre à travers des similitudes conceptuelles, ce qui nous ne permettrait pas de mettre à l'épreuve ce critère pour de nombreux emprunts. Le paradigme latent de tels néologismes est alors constitué de noms à conceptualisation similaire, comme *roulé*, *rouleau*, ou *wrap* dans le cas de *sushi*⁴⁸. Le choix de tels correspondants s'est fait avec la comparaison de signes comportant un potentiel combinatoire similaire⁴⁹. Cette méthode (basée sur Courbon 2009 et 2012a) permet d'accéder à un choix de signes de façon relativement objective, car les correspondants sont fournis par des locuteurs et non par des chercheurs, et ainsi d'éviter la justification d'un genre par des correspondants rares et hors d'usage⁵⁰.

Nous avons analysé l'intégration paradigmatique (abrégée *IP*) à travers trois critères. D'abord, l'**IP par similitude conceptuelle** réfère au processus d'intégration d'un emprunt dans un paradigme latent où les

⁴⁸ On pourrait donc énoncer cette conceptualisation (ou conception latente) comme : [+COMESTIBLE], [PETITE TAILLE] [DONT LE CONTENU EST ENROULÉ SUR SOI-MÊME DANS UNE COUCHE EXTÉRIEURE].

⁴⁹ Par exemple, *tatami* est trouvé entre autres dans la combinaison [agenouillé/assis] sur le *tatami*. On peut comparer avec des signes d'une grande proximité sémantique, tels que *tapis*, *plancher* ou *matelas*, à condition que ceux-ci puissent entrer un minimum dans les mêmes combinaisons.

⁵⁰ Par exemple, Surridge (1982:31) utilise *galoche* comme homologue féminin à *shouclaque* (féminin), *galoche* étant un terme totalement inusité au Québec, d'autant plus que *shouclaque* présente une grande variabilité de genre selon nos observations. Un autre exemple (Surridge, *ibid.*) : *slappe* aurait en français canadien, aux dires de Surridge, l'homologue *dosse*, un homologue qui n'est plus actuel pour un mot dont on pourrait plus facilement trouver les unités latentes *gifle*, *claque* ou *frappe*.

concepts sont similaires. La connexion entre le paradigme latent et l'emprunt se fait par le biais de l'expérientiel et de la similitude entre les deux référents ou concepts, et donc l'emprunt adopte le genre majoritaire du paradigme. Ainsi, à supposer les emprunts *highway*, *freeway*, *parkway*, *trail* ou *track*, ceux-ci risquent davantage de prendre le genre féminin, commun au paradigme des noms de voies⁵¹ (Rice 2006:1403, SurrIDGE & Lessard 2008:94). L'IP d'emprunt contribue à unifier et ordonner la « répartition des genres selon la structuration sémantique du lexique » (Roché 1992:119).

Le deuxième critère d'ordre sémantique, l'IP **par hyperonymie** (Ibrahim 1973:23, Grevisse 1986:770, Guilford 1999:68), réfère au processus où le signe intègre le paradigme latent composé entre autres d'un ou plusieurs hyperonymes de ce signe. Ce processus expliquerait pourquoi les noms d'arbres sont le plus souvent masculins, car leur hyperonyme *arbre* porte la MAM, relevant probablement de l'archigenre. En regard à la classification cognitive des concepts, Rosch *et al.* (1976) ont établi une tripartition hiérarchique représentant l'organisation des catégories sur trois niveaux conceptuels : le niveau superordonné (exemples de Kleiber (1990:83, 88) : *animal*, *fruit*, *meuble*), le niveau de base (aussi *niveau basique*; *chien*, *pomme*, *chaise*) et le niveau subordonné (*épagneul*, *golden*, *chaise pliante*). Selon Rosch *et al.* (1976; également cité dans Tversky 1986:64-65), le niveau de base est le niveau de catégorisation fournissant le plus d'informations à propos d'une réalité pour un moindre effort cognitif. Les auteurs indiquent que le niveau de connaissances d'un domaine spécifique d'un individu peut modifier le statut d'un élément subordonné et en faire un élément de niveau de base. Il est donc faux d'affirmer que la distribution d'éléments de niveau de base est la même pour tout le monde, mais il est raisonnable de croire qu'un grand nombre de locuteurs partagent les mêmes niveaux de conceptualisation.

En se basant sur cette hiérarchie, certains auteurs (Zubin & Köpcke 1986, Thornton 2009) ont étudié comment le genre était attribué selon le niveau de catégorisation d'un concept. Zubin et Köpcke ont remarqué que la plupart des superordonnés en allemand sont de genre neutre⁵², alors que les termes de niveau basique sont généralement masculins ou féminins, car ils portent plus de traits sémantiques distinctifs (Kleiber 1990:84-85, 88)⁵³. Si on part du constat que le neutre allemand agit comme l'archigenre du français (le genre à la

⁵¹ *Route*, *autoroute*, *rue*, *ruelle* s'opposent à *sentier* et *chemin* dans leur conceptualisation, en ce que les noms masculins possèdent un trait sémantique de destination, de finalité que les autres termes n'ont pas nécessairement (*cheminer vers quelque part*, *chemin faisant*, *sortir des sentiers battus*).

⁵² Comme *das Obst* « le fruit », *das Gemüse* « le légume », *das Metall* « le métal », *das Fahrzeug* « le véhicule », *das Instrument* « l'instrument »; et les « hypergénériques » *das Ding* « la chose », *das Objekt* « l'objet », *das Wesen* « l'être ».

⁵³ Deux notes à ce sujet: les concepts de niveau basique « fournissent un maximum d'information avec un minimum d'efforts cognitifs » (Rosch 1978:28). De plus, "increasing specificity from the superordinate to the basic level leads to a large gain in informativeness, but further increases in specificity do not increase informativeness, but do increase the mental burden of categories and distinctions" (Tversky 1986:65, cités dans Thornton 2009:28).

conceptualisation plus vague⁵⁴), lorsqu'on regarde les superordonnés en français, on se rend compte que la plupart de ceux-ci portent la marque du masculin⁵⁵, et sont conceptuellement vagues, donc génériques. Cela nous semble donc être l'influence de l'archigenre.

Thornton (2009) a pour sa part étudié l'hyponymie en italien et a remarqué que les unités lexicales exprimant des concepts subordonnés étaient le plus souvent du même genre que le terme basique. Puisque les termes basiques en français sont autant de genre masculin que de féminin (car leur conceptualisation est plus définie et précise que celle des concepts superordonnés), il n'est pas possible de prévoir le genre d'une unité lexicale subordonnée sans connaître le genre de son supérieur basique. Pour Thornton (2009:31, *notre traduction*), « la plupart des emprunts dénotent des entités particulières et très spécifiques, qui ont plutôt un statut de subordonné ». À partir de cette considération, on peut déduire le genre d'un emprunt en connaissant le genre de son hyperonyme, qu'il s'agisse d'un concept de niveau de base (masculin ou féminin) ou d'un concept superordonné (genre à valeur générique, donc marque du masculin en français). Afin de tester ce critère, nous avons trouvé les hyperonymes des unités examinées à partir d'une formule « X est une sorte de Y », reconnue pour sa validité dans l'étude des réseaux sémantiques (Levrat & Sabah 1990:89, Murphy 2010:114-115).

On pourrait reprocher à qui veut utiliser ce critère que l'hyperonyme *fruit* ne saurait justifier le genre de nombre de noms de fruits (*pomme, fraise, framboise, poire, mûre, citrouille, tomate*, etc.), ce qui est vrai dans le cas de ces fruits, malgré que le critère précédent d'IP par similitude conceptuelle puisse expliquer leur genre. À l'inverse, tous les noms de fruits ne portent pas la MAF : on trouvera de nombreux fruits avec la marque du masculin, tels que *kiwi, ananas, poivron, pomélo, pamplemousse, potiron, citron, bleuet, coing, raisin, avocat, pois, maïs, concombre, melon*, etc. Si certains noms de fruits cités (portant chacun des genres) ne sont pas prototypiques ou conceptualisés comme fruits (la tomate, la citrouille et le poivron pourraient être catégorisés comme légumes), il n'en reste pas moins qu'ils sont biologiquement le fruit d'une fleur pollinisée portant au minimum une graine et qu'ils ne correspondent pas à la tige, à la feuille ou au tubercule du plant.

Finalement, avant de passer au troisième critère sémantique, nous sommes d'avis que le critère d'IP par hyperonymie ne devrait pas être considéré comme une sous-catégorie de l'IP par similitude conceptuelle, car

⁵⁴ "Superordinate concepts are conceptually vague, undifferentiated, while basic level concepts are richly specified both perceptually and functionally" (Zubin & Köpcke 1986:146).

⁵⁵ Comme *fruit, légume, arbre, animal, oiseau, poisson, véhicule, outil, lieu, événement, meuble, aliment, mets*, etc.; quelques exceptions (*plante, roche, machine*) qui pourraient être expliquées à travers des oppositions morphologiques (*plant/plante, roche/rocher, machin/machine*). Davantage d'exemples et de contre-exemples peuvent être trouvés dans Nelson (2005:25) et SurrIDGE et Lessard (2008:201-202). Pour des exemples d'unités encore plus génériques, appelées *hypergénériques* (telles que *bidule, truc, machin*), voir la section 3.1 et la note 64.

si les paradigmes latents qu'ils analysent peuvent être formés des mêmes concepts ou de concepts différents, les concepts sont regroupés en paradigme selon des critères d'ordre différent (ressemblance et taxonomie).

Le troisième et dernier critère sémantique, l'IP **par ellipse d'un signe entier**, réfère au type de réduction sémiotique (Courbon 2015a) par lequel un syntagme nominal polylexical (par exemple, *café latte*), en français, l'un (*latte*) déterminant l'extension de l'autre (*café*), s'élide du substantif déterminé (*café*); le substantif qui le déterminait (*latte*) prend le genre du substantif qui était déterminé (*café*), celui-ci devenant latent (Guilford 1999:70, Boulanger & Malkowska 2008:129)⁵⁶. Un autre exemple serait *la bento*, à l'origine *la boîte bento*⁵⁷.

2.3.3 Conservation du genre

Puisque nous étudions des emprunts à l'italien et à l'arabe, il nous faut considérer la possibilité que l'existence du genre dans la langue d'origine puisse avoir une influence sur le genre attribué en français. Stolz (2009) a démontré qu'une interaction entre les genres de l'arabe et des langues romanes est possible. Dans notre grille d'analyse, nous avons comparé le genre de l'emprunt dans sa langue d'origine (fourni par des dictionnaires de la langue d'origine) et le genre majoritaire observé en corpus.

2.4 Le concept et la conceptualisation

En dehors du genre, la question de l'organisation de la pensée et du langage conjoints est une question qui va au-delà de la visée de ce mémoire. Néanmoins, nous employons à plusieurs reprises les termes de *concept* et de *conceptualisation* qui peuvent être employés avec des significations différentes. Il convient donc de définir les termes que nous employons.

L'humain vit continuellement de nouvelles expériences qui sont perçues par ses sens et analysées par son cerveau. L'esprit humain doit généraliser les informations qu'il reçoit pour mieux les classer, les interpréter et raisonner à partir de ces informations. Chaque expérience vécue est singulière, particulière, ponctuelle et l'humain regroupe, par le processus de conceptualisation, chaque expérience sous une catégorie mentale qu'on appelle le concept. La conceptualisation est une opération cognitive complexe qui consiste à unifier la multiplicité d'expériences singulières (directes ou indirectes) en une catégorie mentale et qui participe à d'autres procédés mentaux que nous n'aborderons pas ici (raisonnement, jugement, décision; Meunier 2006:19). La conceptualisation est également le processus qui, au moment de l'acte de référence, structure les concepts et les associe pour constituer des phrases et formuler un discours d'une portée logique déterminée (Ouellet 2005:149). Le résultat des généralisations des expériences, le concept, est d'abord et avant tout une construction mentale qui est applicable à une multiplicité d'expériences et qui peut être

⁵⁶ Il est très important que l'ellipse ait lieu en français et non dans la langue donneuse, puisque le déterminant élide doit être français, pour pouvoir donner son genre.

⁵⁷ Aussi fréquent : *le bento (box)*.

combinée ou transformée avec n'importe quel autre concept afin de créer de nouveaux concepts ou signifiés (*ibid.*, Meunier 2006:9-11). S'il est vrai que l'expérience généralisée devient concept (ou est classée sous un concept), on n'a cependant pas besoin de vivre physiquement ou directement chaque expérience pour en constituer le concept; ainsi, le concept de MORT est accessible à tout locuteur sans que celui-ci n'ait à mourir. Dire que le concept est basé sur l'expérience, c'est avouer qu'il puisse varier d'individu en individu, puisque les expériences sont singulières et varient. Le niveau d'expertise d'un sujet donné peut varier considérablement entre les individus, et même au niveau des connaissances considérées comme partagées (*cf.* « l'expertise », Nyckees 2007, par. 143) on peut observer des variations de signification, donc de conceptualisation⁵⁸.

La relation entre concept et sémiotique est forte, mais n'est nécessaire que dans un but d'expression ou de raisonnement (*ibid.* p. 24) : pour exprimer un contenu (conceptuel), il faut avoir recours à des signes. Ceux-ci ne sont donc pas nécessaires lorsque les concepts servent d'autres fonctions, notamment au jugement ou à la décision. Dans le cas de la mise en discours du concept par un signe, « chaque fois qu'il s'exprime, un locuteur produit et associe les notions qui composent les concepts ainsi que les signes qui les marquent » (Ouellet 2014:37).

Maintenant que nous avons vu dans la section deux sur quelles bases se fondait notre étude, poursuivons avec notre cadre théorique, puis par la méthodologie.

⁵⁸ Cela nous intéresse encore plus, en ce que la variation de genre peut ne pas s'appliquer qu'aux emprunts, mais également aux mots « communément partagés », tels que *moustiquaire*, *trampoline*, *tentacule*, *pétale*, etc.

3. Cadre théorique

3.1 Position sur l'aspect sémantique du genre

Nous avons mentionné plus haut (section 2.2.1) différents points de vue quant à l'aspect sémantique du genre. Nous allons expliciter dans cette section notre position quant aux théories vues ci-haut.

Damourette et Pichon (1968[1911]) ont été les premiers à attribuer une valeur aux genres. Bien que nous ne trouvions pas raisonnable l'idée d'appliquer au lexique entier des caractéristiques prototypiques de la perception sociale des sexes, leurs efforts, misant à appliquer une valeur « sémantique » quelconque à toutes les unités lexicales de la langue, ont servi de jalon au domaine. Nous reconnaissons que, dans certains cas (notamment ceux des métaphores allégoriques), le genre peut être attribué en fonction de caractéristiques sexuelles ou sociales, mais l'argument contraire pourrait être avancé. Par exemple, si la justice ou la liberté sont représentées par des femmes (*cf. Iustitia*, déesse romaine de la justice, et *La Liberté guidant le peuple* d'Eugène Delacroix), on pourrait argumenter que le genre grammatical (souvent associé aux sexes ou aux *genders*) a dicté le choix du sexe de l'individu symbolique. De même, selon Henriette Gezundhajt⁵⁹, la mort est généralement représentée par une femme faucheuse en français et un homme faucheur en allemand (*der Tod*). De plus, si le soleil et la lune sont généralement représentés par un homme et une femme en français⁶⁰ (et dans les langues latines), on pourrait argumenter que le genre grammatical de ces mots en allemand (respectivement féminin, *die Sonne*, et masculin, *der Mond*) influence la perception de cette allégorie⁶¹. En ce qui a trait à notre étude, nous n'attribuons pas de valeurs sexuelles ou sociales aux genres spécifiques ou générique, parce la perception des sexes ne nous semble pas pouvoir définir le genre d'emprunts au-delà des entités animées (et encore⁶²).

Bien que Hjelmslev fût le premier linguiste à attribuer aux genres spécifiques une valeur sémantique n'étant pas basée sur la distinction sexuée homme vs femme, les distinctions qu'il établit dans son article de 1938 ne sont pas directement corrélées à un genre en spécifique. Nous croyons pouvoir établir à quel genre correspond quelle valeur (notamment avec Togeby 1965), mais sans confirmation d'une autre source de Hjelmslev, il nous est impossible d'être certain de ce que l'auteur avait à l'esprit. En revanche, Togeby perçoit dans le genre féminin une productivité ou une activité (sans doute « l'expansion » que Hjelmslev avait à l'esprit) et dans le masculin une passivité, deux valeurs que nous observons dans nombre de paires (voir

⁵⁹ Département d'études françaises de l'Université de Toronto, notamment ici : <http://www.linguistes.com/mots/genre.html>.

⁶⁰ Pensons à Charles Trenet qui chantait « *Le Soleil a rendez-vous avec la lune* ». La paire *Le Soleil/la lune* est également discuté par Courbon (à paraître).

⁶¹ On pourrait ajouter encore quelques exemples : la mer, personnifiée comme une femme par Michel Rivard (« Je voudrais voir la mer et danser avec elle »), masculin en espagnol (*el mar*), mais utilisé parfois au féminin (*la mar*), notamment au Pérou (Courbon à paraître).

⁶² Nous renvoyons à Ouellet (2014) pour une étude plus poussée du genre en ce qui a trait aux unités animées.

2.2.1). Lafont (2001) semble être du même avis. Évidemment, ces valeurs (ainsi que toutes les autres sur lesquelles nous nous basons) ne sauraient être applicables à toutes les unités. Par exemple, *une chaise* est-elle vraiment plus active ou productive qu'*un banc*? Nous ne pensons pas que les valeurs que nous avons soulignées (et que nous soulignerons dans cette section) puissent être appliquées à toutes les unités portant une marque de genre ou l'autre, sans tenir compte d'autres valeurs sémantiques, contextuelles ou combinatoires.

Martinet voyait dans le féminin indo-européen à l'inverse une passivité ou une réceptivité associée à la terre. Si nous sommes d'accord avec la notion de réceptivité (car recevoir n'empêche pas de produire), Martinet ne fournit pas d'exemple en français permettant de soutenir son propos, propos de prime abord contradictoire⁶³.

Arrivé (1997) associait au féminin une valeur augmentative à travers la dérivation, valeur à laquelle il amenait des contre-exemples à valeur diminutive. À notre avis, distinguer les genres à travers des suffixes ne peut permettre une comparaison sur un même pied d'égalité, car les suffixes portent de nombreuses valeurs sémantiques qui leur sont propres. Il faut donc comparer, afin de faire sortir de manière plus évidente les valeurs sémantiques des genres, des paires minimales sémiologiques, qu'elles soient morphologiquement simples (*pas-passe*) ou complexes (*distributeur-distributrice*).

En ce qui a trait à Zubin et Köpcke (1984), leur étude sur le genre en allemand ne nous permet de tirer aucune conclusion en regard des genres en français. Les valeurs sémantiques relevées pourrait être le produit du suffixe *-mut*, suffixe qui ne trouve pas d'équivalent en français. Nous avons décidé de conserver à l'écart leur théorie pour notre étude.

Audet et Ouellet (1984) et Malkowska (2004) ont relevé de nombreuses valeurs sémantiques qui nous semblent justifiées selon un bon nombre de paires minimales. Si les oppositions (en termes d'effectif et de virtuel, d'objectif et de relatif, de particulier et de général, de défini et d'indéfini) peuvent justifier le genre de bon nombre d'unités, il nous semble que la question de productivité vs passivité ne soit pas abordée en ce qui a trait aux entités inanimées. Nous sommes d'avis que cette distinction permettrait également la justification de plusieurs exemples.

La question de l'archigenre a été peu traitée dans la littérature en dehors de la valeur générique du masculin, généralement des animés humains, comme une catégorisation plus générale que celles des sexes. Zubin et Köpcke (1986) ont remarqué que de nombreux termes superordonnés (« noms de catégories », voir 2.3.2) et hypergénériques (exemples en français : *chose, truc, objet, machin*) en allemand portaient le genre neutre.

⁶³ Rappelons-le, la terre reçoit les semences (réceptivité) et fait croître le plant (productivité). La notion de productivité va contre celle de passivité.

Une étude rapide des hypergénériques en français canadien auprès de notre entourage nous permet de suggérer que les hypergénériques porteraient un peu plus souvent la MAM que celle du féminin (15 à MAM, 11 à MAF)⁶⁴. Les catégories « de base », superordonnées aux concepts de base, semblent porter également la MAM plus souvent que la MAF (voir notes 52 et 55). Bien qu'on observe des cas contraires, la majorité des superordonnés semblent porter la marque du masculin, ce qui est un argument supplémentaire en faveur de l'existence d'un genre à conceptualisation vague ou générale.

En conclusion, en ce qui a trait à l'aspect sémantique du genre, nous sommes d'avis que les genres spécifiques marquent des distinctions sémantiques que l'archigenre ne relève pas en raison de sa définition et conceptualisation vague qui n'a pas besoin de souligner de telles distinctions. L'archigenre n'est pas systématique dans tous les cas, mais il semble étendre son influence depuis les termes hypergénériques et superordonnés jusqu'aux emprunts, comme nous le constaterons un peu plus bas.

3.2 Position sur l'attribution du genre et sur le choix des critères retenus

Comme nous l'avons mentionné à de nombreuses reprises, nous sommes d'avis que le genre peut se trouver motivé par la signification et par la forme physique. Cette position semble contraire à tous les auteurs dont nous avons retenu le point de vue (qu'ils priment l'aspect sémantique ou formel), et peut sembler de prime abord contradictoire. Elle est cependant nécessaire afin de vérifier sur un même pied d'égalité dans quelle mesure les nombreux critères que nous avons retenus peuvent justifier le genre d'un emprunt. Notre position ne rend pas nos résultats moins valides que ceux de nos prédécesseurs; à l'inverse, on pourrait argumenter que la visée large conservée tout au long de l'étude nous permet plus clairement d'étudier l'attribution du genre des emprunts sans tenir de jugement quant au caractère justificateur des critères retenus. Rappelons que nous n'avons pas la prétention d'être capable d'expliquer la variation du genre de tous les emprunts, sinon que de suggérer des pistes de recherche en faveur d'un aspect ou de l'autre de chaque unité.

Dire que les statistiques basées sur les études du français que nous utilisons dans notre analyse ne sont pas susceptibles d'être exploitables pour l'étude du genre des emprunts reviendrait à prendre parti pris de l'aspect onomasiologique de l'attribution du genre prônant le sens sur la forme. Poussé à l'extrême, ce point de vue reviendrait à dire qu'aucun emprunt ne pourrait intégrer la langue par sa sémiologie, parce que l'unité proviendrait nécessairement d'un système à sémiologie différente du français, donc différent en tout point et

⁶⁴ À noter que la forme sémiologique de ces unités varie ostensiblement, surtout à l'écrit.

Sont de genre masculin : *bidule, truc, cossin, machin, machintruc, objet, gugus* (ou *gugusse*), *gossage, trucmuche, grément* (ou *gréement*), *chmu* (ou *schmu, chmou*), *stuff, guédi* (ou *guédis*), *kit, bézouigne*.

Sont de genre féminin : *chose, patente* (à *gosse*; aussi *patante*), *pinouche, affaire, gogosse, be Belle* (ou *bébelle*), *greluches* (généralement pluriel), *babiole, bidouille, réguine, guiznoutte*.

non compatible. Ce point de vue fortement réductif ne saurait nous permettre d'accomplir une analyse transcendant tout préjugé sur l'attribution du genre aux emprunts.

En ce qui a trait aux critères que nous avons retenus, tous ont été mentionnés par au moins un auteur, excepté ceux de l'analogie phonétique (AP) par phonème pénultième et de la paronymie. Pour ce premier, nous étions simplement curieux de voir dans quelle mesure les résultats de Tucker *et al.* (1977), et donc le critère d'AP par phonème final, pouvaient être étendus à la finale d'un mot. De plus, la paronymie avait déjà avancé comme une possibilité de motivation par quelques auteurs; par exemple, Surridge (1994) a suggéré à l'inverse que la paronymie (ou dans son cas, l'homophonie des noms non homonymes) différencierait systématiquement le genre des unités (ex : *cours-cour*, *foie-foi-fois*, *col-colle*, *lieu-lieue*, *sel-selle*, ainsi que *mémoire-mémoire*, *rose-rose*, *orange-orange*⁶⁵). Dans le cas de tels avis contraires, nous avons choisi d'inclure ce critère afin de vérifier quel point de vue était le plus près de la réalité des emprunts.

Rappelons que les études de Tucker et de ses collaborateurs ont pris pour objet d'étude tous les substantifs d'un dictionnaire sans distinguer les types de formation ou dérivation des unités. La non-distinction entre la finale phonique et un morphème ajouté à une base (Nelson 2005:19) ignore complètement les différents types de formation de mots qui ont eu ou pourraient avoir une influence sur le genre. Ainsi donc, *thé* sera classé dans la même catégorie que *vulgarité*, ou *peur* dans la même catégorie qu'*instructeur*, choix méthodologique basé uniquement sur leur terminaison phonétique. Une simple comparaison d'unités morphologiquement simples aurait pu donner des résultats beaucoup plus représentatifs du lexique dans lequel les emprunts doivent s'intégrer. La sémiologie du français n'est pas non plus toujours transparente quant au genre; si elle distingue aisément *homme-femme* ou *artisan-artisane*, elle ne distinguera pas *rose-rose*, ou *mémoire-mémoire*. Cependant, aussi contradictoire que cela puisse être, la sémiologie du français ne distinguant pas le genre systématiquement peut rendre les résultats des études de Tucker *et al.* d'autant plus intéressants, car ils reflètent l'ensemble de la sémiologie du français et non une partie d'elle. Outre cela, considérant que les études de Tucker et ses collaborateurs bénéficient d'une grande notoriété dans le domaine de l'attribution du genre (aux unités indigènes et aux emprunts), et considérant que nous nous devons de rester impartial dans le choix de nos critères et dans notre analyse, nous avons tenu à mettre au grand jour le pouvoir justificateur potentiel de la terminaison phonétique dans l'attribution du genre aux emprunts.

Le critère de syllabité finale s'inspirant de Harley (1979) et de Carroll (2005) visait d'abord la forme phonétique du mot. Hors, puisque nous étudions des unités graphiques, nous avons décidé d'inverser la tendance en prenant compte des exemples donnés sous leur forme graphique.

⁶⁵ Nous ne sommes pas d'accord avec ces derniers exemples, car *le rose* et *la rose* (ainsi que d'autres exemples, p.165--169) semblent clairement partager le même signifié, ce qui les rendrait non pas homonymes, mais deux actualisations du même potentiel sémantique.

Les critères de réanalyse morphologique ont également été inspirés par SurrIDGE (1985-1986). Puisque nous avons choisi uniquement des unités à morphologie simple (donc sans suffixe), nous avons eu besoin de voir si les locuteurs voyaient ou pouvaient voir dans les unités empruntés des suffixes significatifs. Nous avons donc adapté *l'analogie suffixale* en une réanalyse morphologique.

Dans la littérature, les critères sémantiques se limitaient souvent à la correspondance d'une seule unité à un emprunt ou à des « classes sémantiques » restreignantes, tels que les noms d'éléments chimiques ou de voiture⁶⁶. Le choix souvent douteux des correspondants nous a fait réfléchir sur le sujet et nous avons plutôt développé le concept de paradigme latent (grâce à Courbon 2015a), préférant attribuer le genre à un groupe d'unités ou de concepts plutôt qu'à une seule unité. L'avantage de prendre en considération un genre collectif permet d'amenuiser le genre d'une unité qui pourrait contredire ce qu'un paradigme entier pouvait mettre en évidence. Afin de trouver un terrain d'entente entre des classes sémantiques et des correspondants simples, nous avons choisi d'analyser l'aspect sémantique principal des emprunts en les intégrant à des paradigmes d'unités à conceptualisation similaire.

Nous avons choisi d'inclure un maximum de critères dans notre analyse, malgré le fait que certains critères ne pourraient s'appliquer qu'à une petite minorité d'emprunts. Nous savons que plus le nombre d'unités analysables est petit, moins nos conclusions peuvent être représentatives. Néanmoins, elles peuvent servir à titre indicatif et de comparaison afin de suggérer l'influence ou la non-influence d'un critère.

Maintenant que nous avons pris position, voyons maintenant comme nous avons procédé à notre analyse.

⁶⁶ Bien que Nelson (2005) ait fait ressortir des groupes sémantiques plus larges et plus variés, notamment les types de carapace ou de coquille (*carapace, coquille, écaille, enveloppe, couenne*), les types d'appendice (*branche, main, nageoire, queue, palme*), les types d'espace vide (*espace, abîme, abysse, gouffre*) et les types de « flat blades » (*spatule, houe, pelle, scie, rame, truelle*).

4. Méthodologie

4.1 Objet d'étude et sélection des unités à examiner

4.1.1 Choix des langues donneuses

Dans le cadre de cette étude, nous avons observé des emprunts qui nous parviennent de l'anglais, de l'italien, du japonais et de l'arabe⁶⁷. La sélection de ces langues se base sur deux critères : la présence ou l'absence d'un genre grammatical dans le substantif et le degré de proximité linguistique par rapport au français. La proximité d'une langue avec une autre peut être considérée tant sur le plan génétique (parenté) que géographique. L'anglais et le japonais sont des LSG substantival alors que l'italien et l'arabe ont deux genres, le masculin et le féminin. L'italien et l'anglais sont des langues relativement proches au français, comparées à l'arabe ou au japonais (deux langues indo-européennes, la première germanique⁶⁸ et la seconde romane; les deux situées en Europe ou en Amérique du Nord, donc à proximité géographique du français); l'arabe et le japonais sont des langues lointaines par rapport au français (une sémitique, l'autre japonique; une de l'autre côté de la mer Méditerranée et l'autre sur des îles d'un autre océan).

4.1.2 Sélection des unités

Pour être sélectionné comme unité à l'étude, un emprunt devait répondre à certaines exigences. Aucun nom ne devait, dans l'une de ses significations, référer à une entité animée humaine⁶⁹, car il y a un risque d'interférence avec le sexe associé au nom inanimé (comme *kamikaze* « soldat qui se suicide » et « avion qui percute un bateau ou un immeuble »). Ensuite, aucun emprunt avec un suffixe français (ou assimilé en français⁷⁰) n'a été sélectionné. Aucun emprunt ne devait être composé de manière à être morphologiquement transparent pour un locuteur francophone unilingue⁷¹. De plus, puisque le genre peut⁷², mais pas dans tous les

⁶⁷ Nous considérons par « arabe » à la fois l'arabe classique et les arabes dialectaux.

⁶⁸ Une grande quantité de lexèmes de l'anglais sont d'origine française ou latine. Si les estimations varient d'un auteur à l'autre, on trouve toujours dans un texte anglais typique une estimation d'au moins 50% des vocables d'origine latine (et française). Joseph M. Williams (1975:67) totalise même 60% pour ces deux langues.

⁶⁹ Nous avons accepté une seule unité dénommant un animal, un chien, par l'emprunt *clebs*.

⁷⁰ Le suffixe emprunté à l'anglais *-ing* est toujours masculin lorsqu'utilisé en français (Guilford 1999:74-75), avec des mots tels que *camping*, *shopping*, *parking*. La corrélation du masculin et de la finale *-ing* démontre une assimilation en français, car nous faisons preuve de créativité avec des composés en *-ing* qui ne sont pas des emprunts à l'anglais (ex : *pressing*). C'est pourquoi ce morphème sera également laissé de côté. Pour ce qui est de la dérivation en arabe, Ibrahim (1973:45) explique qu'elle s'accomplit en manipulant la racine d'un mot de différentes manières (cf. morphologie templatique des langues sémitiques avec Shimron (2003) et Stump (1996)). Seul *-at* est un véritable suffixe dérivationnel, apparu plus tard dans les langues sémitiques.

⁷¹ Nous avons fait exception à cette règle pour deux unités (*jam* et *jam-session*), car il nous a semblé percevoir une différence du genre utilisé, notamment à cause de *session*. Des composés tels que *pan+cake*, *pipe+line*, *inter+view* n'ont pas été exclus, car il nous a semblé qu'ils ne formaient qu'une seule unité dont la morphologie n'est pas transparente pour un non-initié à l'anglais. Il en a été de même pour *spagh+etti*, *pane+ttone*, etc.

⁷² Ibrahim (1973:27-28).

cas⁷³, distinguer entre deux significations d'un même signe, nous avons exclu les emprunts qui avaient une forme graphique identique ou similaire à un substantif ou un sigle indigène au lexique du français⁷⁴, comme l'est *tag* « signature graffiti » qui est similaire à *TAG* « trouble d'anxiété généralisé » ou *full* « maison pleine [poker] » qui est phonétiquement similaire à *foule*. Finalement, chaque emprunt a dû être intégré au lexique du français après l'an 1800. Cette date n'est pas aléatoire : elle exclut les emprunts à l'arabe médiéval, les emprunts à l'italien de la période classique française, les emprunts à l'anglais d'avant la révolution industrielle. Cette date inclut les emprunts à l'anglais de la période « post-révolution industrielle », tous les emprunts au japonais, les emprunts à l'italien post-Renaissance et les emprunts à l'arabe de la période coloniale française en Afrique, et finalement les emprunts à l'arabe dans l'argot français postcolonial.

Dans le but de circonscrire nos corpus et d'observer la variation du genre, nous avons adopté une approche thématique tout en sélectionnant des emprunts de divers domaines. Les emprunts relatifs au domaine de l'alimentation (Boulanger & Malkowska 2008) forment environ la moitié thématique des emprunts à l'étude. L'autre moitié est composée entre autres d'emprunts relatifs au domaine du divertissement (Guilford 1999).

Avant la sélection finale, nous avons constitué une « pré-liste » qui regroupait tous les emprunts correspondants à nos critères que nous avons trouvés notamment dans le *Petit Robert 2015*, à partir d'une recherche selon la langue donneuse des mots. Cette pré-liste contenait environ 400 unités. Nous avons procédé à une recherche de chacune des unités dans des corpus pilotes (voir la section 4.3.1) pour quantifier le nombre moyen d'occurrences de chaque emprunt. Nous avons exclu les unités trop peu fréquentes (sous le 25% de moyenne ajustée de fréquence) et les termes techniques relevant d'un domaine de spécialité (habituellement accompagnés d'une marque technique dans le *Petit Robert 2015*, comme *rebab*, *sakieh*, *simoun*, *erg*, *makimono*, *kakémono*, *biwa*). De là, nous avons constitué une seconde liste d'environ 160 mots à partir de l'épuration des 400 premiers. Nous avons ensuite cherché les 200 premières occurrences de chaque emprunt dans les mêmes corpus pilotes afin d'observer la variation en genre possible des emprunts. La majorité des emprunts qui présentaient une variation, même d'une seule occurrence, ont été retenus (19 en anglais, 6 en japonais, 15 en arabe, 10 en italien). Les moins fréquents (en deçà de 25 occurrences sur 200,

⁷³ Une même unité peut avoir plusieurs significations différentes et un même genre. Pour preuve, en France, les deux sens de *clip* (« pince » et « segment vidéo ») sont utilisés au masculin. On trouve également *pitch* « résumé, synopsis », « argumentaire », « fréquence d'une onde sonore perçue comme la hauteur du son », « goudron », toutes les significations étant masculines. Nous ne croyons pas qu'il soit possible d'affirmer que les significations de chaque signe correspondent à deux signes différents sans avoir recours à une étude diachronique du mot.

⁷⁴ Cette décision nous assure une certaine indépendance de l'unité et évite ce qu'Ullmann (1959:224) appelle la *collision homonymique*, c'est-à-dire l'interaction entre deux formes homonymiques. Évidemment, notre choix travaille à l'encontre des critères de similitude graphique et de paronymie, mais nous avons sélectionné nos unités et composé les corpus avant la sélection du critère de paronymie. Cela explique donc pourquoi un si petit nombre de paronymes peut être trouvé aux unités que nous avons sélectionnées.

50 occurrences par corpus pilote) ont été exclus, à moins de présenter une variation substantielle de genre, ou de faire partie d'une thématique.

Au final, 100 unités ont été sélectionnées, 25 par langue. Les voici, regroupés par langue, puis par thématique :

1. **anglais (autres sujets)** : *gig, jam, jam-session, party, rave, slam, hype, interview, break, bug* (et *bogue*), *clip, scoop, trip, drum, groove, pipeline, dope, shit, speed, van* (et *vanne*);
 anglais (alimentation) : *pancake, sandwich, mug, stout, cake*;
2. **japonais (autres sujets)** : *bonsaï, dojo, haiku, karaoké, shiatsu, sudoku, tatami, tsunami, manga, nunchaku, kimono, obi*;
 japonais (alimentation) : *bento, tempura, wasabi, miso, seitan, surimi, matcha, saké, soja* (et *soya*), *tofu, maki, sashimi, sushi*;
3. **arabe (autres sujets)** : *baraka, chicha, keffieh, chèche, niqab, burqa, hidjab* (et *hijab*), *intifada, fatwa, djihad* (et *jihad*), *charia, barda, bled, clebs, sarouel*;
 arabe (alimentation) : *brick, chorba, harissa, kefta* (et *kofta, kofte, kefte*⁷⁵), *merguez, tajine, taboulé, couscous, loukoum, méchoui*;
4. **italien (autres sujets)** : *omertà, vendetta* (et *vandetta*), *sépia*;
 italien (alimentation) : *cannelloni, farfalle, tagliatelle, spaghetti, pancetta, carpaccio, prosciutto, expresso* (et *espresso*), *cappuccino* (et *capuccino, cappucino, capucino*), *panini, ciabatta, panettone, calzone, pizza, bruschetta, antipasti* (et *antipasto*), *minestrone, ricotta, risotto, sabayon, salami, tiramisu*.

Pour des raisons méthodologiques et d'économie de temps, nous n'avons pas cherché toutes les formes graphiques de chaque emprunt. Seuls les emprunts avec des diacritiques ont été cherchés avec et sans accents ou trémas (donc *karaoké* et *karaoke*, *bonsai* et *bonsaï*). Nous sommes conscient que, dans l'optique où l'on considère la graphie d'un emprunt comme déterminante dans la motivation du genre, il serait possible d'imaginer que le genre puisse varier. Or, nous ne croyons pas que ce soit le but de la présente étude de vérifier d'autres critères graphiques que la finale. Des vérifications au sujet d'autres critères graphiques pourraient être entreprises, mais nous ne les ferons pas dans cette étude.

4.2 Recherche préliminaire et comparaisons lexicographiques

Une fois les unités décisives sélectionnées, nous avons cherché dans six dictionnaires la signification et le genre de chacune afin de vérifier s'il y a eu évolution soit du genre donné, soit de la formulation employée pour décrire les unités. Nous avons sélectionné trois éditions différentes du *Petit Larousse illustré*, celle de 1908, celle de 1957 et celle de 2012. Nous avons également sélectionné le *Grand Robert en ligne* (2014) qui nous a semblé bien compléter la sélection d'ouvrages européens. Nous avons donc deux éditeurs français

⁷⁵ En ce qui a trait aux formes *kofte/kofta*, notre informatrice arabophone nous a indiqué qu'elles étaient plus turques qu'arabes. Cependant, puisque notre liste d'unités et nos corpus étaient déjà complétés, nous avons conservé toutes ses formes probables en français.

différents et nous voulions compléter avec deux ouvrages québécois. Nous avons donc choisi le dictionnaire *Usito* (2014) et le *Dictionnaire Bélisle de la langue française au Canada*, édité vers 1958⁷⁶, d'abord pour leur année de parution (équivalente aux *Petit Larousse*), puis pour leur structure et leur mode de définition lexicographique semblables. Ces ouvrages nous semblaient relativement indépendants de la norme lexicographique parisienne et ont parfois été les seuls à attester une signification québécoise.

À ce stade, nous avons obtenu quelques résultats intéressants pour une recherche préliminaire. Premièrement, selon les ouvrages, le genre variait pour 13 emprunts (*ciabatta, calzone, harissa, interview, minestrone, party, pipeline, rave, sandwich, stout, tagliatelle, tempura, van/vanne*), ce qui semblait prometteur pour nos résultats. Deuxièmement, les emprunts faisant partie de la thématique alimentaire référaient souvent à la fois au plat entier, à la fois aux composantes du mets. Pour cela, nous croyons possible d'observer de la variation dans le genre de ces emprunts (*manger un spaghetti* (mets), *fait de spaghettis* (composantes)), car la conceptualisation de la réalité, comme entier ou comme ensemble de composantes individuelles, pourrait possiblement modifier le genre employé. Finalement, la formulation lexicographique de la définition de plusieurs emprunts était parfois très différente de l'usage, car la description lexicographique se veut scientifique, alors que l'usage utilise l'emprunt sous une autre signification. Par exemple, *soya/soja* est généralement décrit comme une plante légumineuse d'origine asiatique dont les graines contiennent des acides gras bénéfiques, alors qu'il suffit de regarder comment les locuteurs emploient ce terme pour se rendre compte qu'il parlent de la substance dont est fait le tofu ou autre aliment dans une interprétation générale de la matière.

4.3 Constitution des corpus

4.3.1 Corpus pilotes

Les corpus pilotes ont servi à la constitution de la liste finale d'unités à l'étude, afin de permettre la vérification de la fréquence de celles-ci et d'attester leur genre dans l'usage. Nous avons constitué quatre corpus pilotes sur le même modèle que les corpus finaux : deux québécois, deux européens, deux formels (ou semi-formels), deux informels, chacun incluant une approche thématique et non thématique. Dans le but d'estimer un nombre de mots de chaque corpus pilote, nous avons compté rapidement (à l'aide des mots clés *est* et *était*) le nombre d'articles (dans le cas de corpus formels, à partir de la base textuelle *Eurêka*) ou commentaires (pour le forum *Domaine Bleu*) et nous l'avons multiplié par la moyenne de mots par article/commentaire (établie selon les 100 premiers articles/commentaires de chaque quotidien/forum) afin d'obtenir un total estimé. Le nombre de commentaires et de mots pour les forums européens informels nous a été fourni par le logiciel du

⁷⁶ L'édition dont nous disposons ne comporte pas de date de parution. Nous avons donc déduit l'année de la photographie située avant l'introduction qui célèbre les 350 ans de la ville de Québec (1608-1958).

projet de recherche « Appropriation du vocabulaire spécialisé par les francophones : canaux de diffusion et modes d'intégration », dirigé par Bruno Courbon. Les sous-forums sélectionnés dans le corpus pilote informel européen pour l'approche thématique sont *nutrition, régime, cuisine, alimentation et santé*; ceux pour l'approche non-thématique sont *famille, loisir, people et vie pratique*. Le tableau suivant exprime la composition et la taille des corpus pilotes.

Corpus Pilote	Composition	Nombre d'articles ou de commentaires approximatif	Mots par article/commentaire	Total estimé de vocables
Formel québécois	<i>La Presse, Le Soleil, Le Journal de Québec, Le Nouvelliste, Le Droit, Le Devoir</i>	703 000 (2010-2013)	436	306 000 000
Formel européen	<i>Le Monde, Libération, Le Figaro, Sud Ouest, Le Temps, Le Soir</i>	1 326 000 (2010-2013)	539	715 000 000
Informel québécois	<i>Domaine Bleu</i> ⁷⁷ (non thématique)	8 255 554 (8 août 2014) 750 000	56,46	462 000 000
	<i>Recettes du Québec</i> ⁷⁸ (thématique)		43,08	32 000 000 ⁷⁹
Informel européen	<i>Doctissimo</i> ⁸⁰ (thématique)	6.7 millions	-	453 000 000 ⁸¹
	<i>Doctissimo</i> (non thématique)	17.2 millions		594 000 000 ⁸²

Tableau 1 : présentation des corpus pilotes

⁷⁷ <http://www.domainebleu.ca/>.

⁷⁸ <http://www.recettes.qc.ca/forum/>.

⁷⁹ Le nombre de commentaires étant impossible à obtenir, la moyenne de mots par commentaire (43.08) a été multipliée par le nombre de pages approximatif du mot *est* (parfois 751000 ou 748000) pour un total d'environ 32 millions de mots au minimum.

⁸⁰ <http://forum.doctissimo.fr/>.

⁸¹ Arrondi avec les forums *cuisine, nutrition-régime et alimentation et santé* (sous-forum de *famille*).

⁸² Arrondi avec les forums *famille, loisir, people et vie pratique*.

4.3.2 Corpus définitifs

Les corpus finaux ayant servis à l'analyse des 100 emprunts ont été constitués manuellement à partir des mêmes sources que les corpus pilotes. Puisque chaque forme a été vérifiée avant d'être incluse dans le corpus, les homographes de formes à l'étude ont été retirés afin de ne conserver que les signes à signification connexe à l'emprunt. De cette manière, *gig* « spectacle de musique » a été conservé, et *gig* « gigaoctet » a été retiré; *rave* « fête de musique techno » a été conservé, le légume *rave* a été exclus. Voici sur quelles bases les corpus finaux ont été constitués.

4.3.2.1 Corpus formels québécois et européen

Ces deux corpus ont été formés à partir de la base de données textuelles *Eurêka* en se basant sur les mêmes quotidiens que ceux des corpus pilotes. Pour chaque emprunt, seule la forme jugée « standard⁸³ » a été cherchée dans le moteur de recherche, sauf pour les emprunts suivants, puisqu'il existe des variantes utilisées dans un contexte formel : *antipasto-antipasti*, *expresso-espresso*, *kefta-kefte-kofta-kofte*, *djihad-jihad*, *hidjab-hijab*, *van-vanne*. Puisqu'*Eurêka* ne différencie pas les caractères diacritiques des lettres sans diacritique, les requêtes ont été faite sans diacritique (donc *bonsai* plutôt que *bonsai*).

Pour chaque emprunt, la période de recherche s'étendait du 1^{er} janvier 2008 au 31 décembre 2013. Nous avons choisi 2013 comme année de fin, car l'année 2014 n'était pas encore terminée lors de nos recherches et nous voulions une année où l'entièreté des résultats (publications, articles, etc.) serait archivée. Pour notre période de recherche, les 100 premières occurrences à genre explicite⁸⁴ pour chaque forme ont été retenues et les articles entiers ont été enregistrés en format *html* (et non *pdf*). Les résultats *html* ont été copiés intégralement dans un fichier bloc-note (*.txt*), le seul format à pouvoir être traité par le concordancier *Antconc*. Dans le cas où nous n'obtenions pas 25 occurrences à genre explicite, nous avons étendu la période de recherche sur 10 ans, donc du 1^{er} janvier 2003 au 31 décembre 2013. Certains emprunts ont été difficiles à attester dans un corpus. Par exemple, *shit* « drogue » a été attesté uniquement 5 fois sur 10 ans dans le corpus québécois, et *obi* « ceinture japonaise » uniquement 15 fois sur 10 ans dans le corpus européen. Les articles doublons ont volontairement été exclus du le corpus. Il reste cependant possible qu'un même article soit présent plusieurs fois s'il incluait plusieurs des unités que nous avons sélectionnées. Voici un exemple d'occurrences d'*obi* en corpus européen et de *shit* en corpus québécois.

- « Tous deux se retrouvaient chaque soir dans cet étrange tête-à-tête. Partageaient de mêmes rituels : la serviette dépliée puis posée sur l'abdomen comme un tablier, une **obi** japonaise; la lenteur à porter le vin au

⁸³ En nous basant sur les mêmes ouvrages de référence qui ont été utilisées pour constituer la liste d'emprunts.

⁸⁴ Afin de déterminer si le genre était explicite, l'emprunt devait apparaître avec un adjectif, un participe passé ou un déterminant discriminant le genre (*gros-grosse*, *mis-mise*, *le-la* etc.), sauf dans le cas d'un emprunt commençant par une voyelle et associé aux possessifs *mon*, *ton* ou *son*.

verre. Un soir, il lui adressa la parole de sa voix de violoncelle. Elle lui répondit en une langue qu'il ne reconnaissait pas (l'allemand). Il alla alors se coucher. » - François Simon (*Le Figaro*), 2010-08-06⁸⁵

- « les taches calligraphiques sur le podium, le premier mannequin en ballerines avance d'un pas feutré, enveloppé dans une robe-courtepointe en soie peinte d'un paysage abstrait. Puis on assiste à une variation plutôt réussie autour du **obi** (imprimé, en soie métallisée, à noeud plat géant...) sur des robes, en maille ou en jersey de coton mat, aux détails féminins : poches plaquées, maxi-col Claudine relevé, bavolet brodé d'une colombe. » - Hélène Guillaume (*Le Figaro*), 2009-03-07
- « Un habitant du quartier, José Vieira, explique que la police a dû intervenir plusieurs fois dans son immeuble. "Les gens viennent pisser dans la cour de l'immeuble, forcent la porte, veulent s'introduire, viennent fumer du **shit** dans le hall", raconte-t-il. » - Anne-Laure Mondesert (*La Presse*), 2008-01-05
- « Les banlieues sont un sujet qui touche de près Calogero, lui-même issu d'un quartier chaud de Grenoble. " C'est un hommage aux jeunes de banlieues qui ont envie de s'en sortir. En banlieue, il n'y a pas que des mômes qui brûlent des bagnoles et ont envie de vendre du **shit** ", estime le chanteur avant de confesser quelques vols à l'étalage. » - Anabelle Nicoud (*La Presse*), 2007-04-14

Nous avons choisi pour le corpus européen de prendre des quotidiens de plusieurs régions de l'Europe francophone : régions de France (*Sud Ouest*, *Le Figaro*, *Le Monde*, *Libération*), Suisse (*Le Temps*) et Belgique (*Le Soir*). Cette sélection nous permet d'avoir un aperçu du français écrit en Europe et permet d'aborder la question du genre des emprunts du point de vue de plusieurs communautés francophones.

Pour ce qui est des quotidiens québécois, nous avons sélectionné *Le Journal de Québec* (Québec), *Le Soleil* (Québec), *Le Devoir* (Montréal), *Le Nouvelliste* (Trois-Rivières), *La Presse* (Montréal) et *Le Droit* (Gatineau-Ottawa). Nous croyons que cette sélection est suffisamment représentative, car elle regroupe quatre centres urbains du Québec, dont deux sont en contact constant avec la langue anglaise, une des langues donneuses à l'étude.

Le nombre d'occurrences à genre explicite se chiffre à 7514 pour le corpus québécois, et 7772 pour le corpus européen.

4.3.2.2 Corpus informel québécois

Le corpus informel québécois a deux sources. La première est constituée, pour les emprunts non-thématiques, d'un forum québécois sur le divertissement (principalement télévisuel), le plus grand au Québec que nous ayons trouvé, nommé *Domaine Bleu*⁸⁶. La seconde source de constitution du corpus est, pour les emprunts thématiques, un forum de cuisine, le plus grand que nous ayons trouvé au Québec, *Recettes du Québec*⁸⁷. Les deux sites ont été exploités via le moteur de recherche Google à l'aide de la formule de

⁸⁵ La date suivra le modèle *année-mois-jour*. De plus, précisons que tous les exemples en corpus cités dans ce travail ont été choisis plus ou moins au hasard à travers toutes les occurrences possibles, sauf lorsqu'il était nécessaire de démontrer la variation en genre d'une unité ou un phénomène connexe. Nous avons coupé un minimum les commentaires cités ici par souci d'économie de longueur en conservant quelques lignes avant et après l'unité en évidence en gras.

⁸⁶ <http://www.domainebleu.ca/>.

⁸⁷ <http://www.recettes.qc.ca/forum/>.

recherche suivante « *intext:MOT site:http://www.domainebleu.ca/* », où *MOT* correspondant à l'unité recherchée et où ce qui suit *site:* correspond à l'adresse URL (ou au domaine) dans lequel la recherche est exécutée. Puisque le moteur de recherche *Google* prend en compte les caractères diacritiques, nous avons effectué chaque requête avec et sans les diacritiques nécessaires à chaque emprunt, nous permettant ainsi de prendre en compte le plus de cas possible. Pour chaque page possédant au minimum une occurrence à genre explicite, nous avons copié dans un document bloc-note (.txt) les paragraphes ou les phrases pertinentes. Nous avons fait un document .txt par unité et nous avons au total 4294 occurrences à genre explicite identifiées. Ce corpus est le plus petit des quatre corpus de l'étude, et cela s'explique partiellement par le fait que les sites Internet à l'origine du corpus sont de petites tailles (32 millions de vocables pour *Recettes du Québec*). De plus, les forums ont la particularité de permettre la reprise d'un commentaire déjà exprimé et de multiplier les vocables par commentaire. Le nombre d'échanges authentiques non repris d'autrui est souvent faible. Puisque nous avons exclu le plus de doublons possibles, nous nous retrouvons avec moins d'occurrences à genre explicite qu'un corpus sans reprise de commentaires. Malheureusement, chaque commentaire a été copié dans le document texte sans prendre en compte le nom de l'utilisateur ayant produit l'occurrence et la date de production. Chaque commentaire cité sera donc identifié par l'appellation *anonyme* et *sans date* pour indiquer le manque d'information. Suivra entre parenthèse « Corpus québécois informel », afin d'indiquer au minimum l'origine du commentaire. Voici des occurrences de *van*, *sandwich* et *ciabatta* :

- « Neil Patrick Harris tourne présentement The Smurfs 2 à Montréal! Complètement dans le mood, c'est dans une **van** bleue qu'il a été spotté! » - Anonyme, sans date (corpus québécois informel).
- « Les 2 000 personnes qui voulaient ma mort ont facilement submergé la sécurité et nous ont couru après. J'étais couverte de sang, je n'arrêtais pas de saigner, et malgré cela, les spectateurs en furie ont essayé de casser les vitres du camion. Les 3 hommes qui étaient avec moi dans le **van** étaient morts de peur et leurs mains tremblaient. Mais ce n'est pas tout, ils ont réussi à casser les vitres, alors la sécurité a mis des tables et des chaises contre les vitres brisées pour tenter de maintenir la foule dehors. Ce fut aussi terrible que l'enfer !" » - Anonyme, sans date (corpus québécois informel).
- « Depuis des années je me fais ce **sandwich** aux crevettes de Matane maintenant appelées Nordiques. » - Anonyme, sans date (corpus québécois informel).
- « Aussi j'aime bien me faire une **sandwich** rotie avec dedans des flocons de goberge mélangée avec mayo et (échalotes facultatif) » - Anonyme, sans date (corpus québécois informel).
- « Je n'ai jamais utilisé le kamut mais je serais porté à croire qu'il n'est pas aussi panifiable que d'autres farines moins rustiques. Pour réussir, comme pour la **ciabatta**, il faut que la pâte reste très collante; donc pétrir le moins possible. A manipuler du bout des doigts pour ne pas s'en mettre partout. Si vs pétrissez comme d'habitude vs ferez sortir tous les gazs nécessaires à la formation des alvéoles. » - Anonyme, sans date (corpus québécois informel).
- « j'aimerais bien avoir ta recette car j'aime beaucoup le ciabatta » - Anonyme, sans date (corpus québécois informel).

4.3.2.3 Corpus informel européen

Le corpus informel européen est constitué de commentaires électroniques diffusés sur Internet (sur les forums du site *Doctissimo*, notamment), entre le début des années 2000 et le début des années 2010. Ce corpus de commentaires est exploité dans le cadre d'un projet de recherche subventionné par le FRQSC intitulé « Appropriation du vocabulaire spécialisé par les francophones : canaux de diffusion et modes d'intégration » et dirigé par Bruno Courbon à l'Université Laval. À partir du logiciel du projet de recherche, nous avons extrait des fichiers contenant un maximum d'occurrences dans un fichier Excel (.xlsx). Nous avons consulté chaque document et nous avons conservé 100 occurrences aléatoires à genre explicite (par unité) en prenant soin de retirer les erreurs, les doublons et les résultats non pertinents⁸⁸. Nous avons exclu tout résultat comportant une unité composée de l'emprunt à l'étude et un terme qui le définissait, tel que *surprise party*. Nous avons finalement copié chaque occurrence et son contexte dans un fichier .txt pour chaque unité.

Puisque nous avons généralement cherché des formes complètes (*jam*, *risotto*) ou presque complètes (*antipast*-, *keffie*-), nous avons pu identifier des occurrences qui avaient une légère variation orthographique et les analyser correctement avec le concordancier Antconc. Nous avons fait exception pour quelques emprunts italiens qui doublent allègrement les consonnes, comme *cappuccino* ou *cannelloni*. C'est donc dire que nous n'avons pas cherché toutes les variantes graphiques d'un signe, tel *sandouiche* pour *sandwich* ou *prochiouto* pour *prosciutto*. Au final, nous avons identifié le genre des 7528 occurrences des unités que nous avons sélectionnées. Voici quelques occurrences de *party*, *risotto* et *cappuccino* :

- « Le juge David Eady, de la Haute cour de Londres, a déclaré que le tabloïd devait payer 60.000 livres sterling (environ 76.000 euros) de dommages et intérêts à Max Mosley. Le président de la FIA a reconnu avoir payé 2.500 livres pour cette "**party**" sadomasochiste, mais il a assuré que c'était privé et qu'aucune fantaisie nazie ne s'était produite. Mosley a dit que le scénario avait pour cadre une prison allemande, mais sans aspect militaire. » - jovipilu (*Doctissimo*), 2008-07-24
- « 17 Février dernier c'était l'anniversaire de paris qui vient d'avoir 27 ans! et pour remercier ses invités d'être venu pour la **party** d'anniversaire, cette dernière leur à offert un petit Strip- tease! (voir la vidéo du strip)merci Paris!! » - robertoventino (*Doctissimo*), 2008-02-22
- « J'adorais la paëlla quand je mangeais un peu de poisson. Oui j'adore aussi tes repas Delfania, surtout les poivrons miam. Hier soir j'ai mangé encore du **risotto** (super bon!), un yaourt et une compote, et il m'en reste encore pour ce midi xD. » - yannn302 (*Doctissimo*), 2013-06-25
- « Boudelaine : l'hôpital, c'était rien de grave??? Pour mes repas, ce matin, c'était comme hier matin et ce midi :- sardines à la tomates- champignons des bois et morilles- purée- yaourt- pomme Et ce soir je crois que je vais faire du **risotto**, c'est la marque Gallio je crois avec courgettes et tomates et en viande ben je ne sais pas! lol je vous dirai ça plus tard! » - didou2526 (*Doctissimo*), 2009-03-15
- « et bin pendant que zhomme était parti en courses j'ai redormi dans le canapé , je ne l'ai même pas entendu revenir , malgré une bonne nuit je n'avais pas récupéré , là ça va mieux par contre les douleurs de fibro sont là alors ça doit être lié ça + le dos on a mangé : poulet rôti , croquettes de p de t au four salade , une pêche et là je vais boire un **cappuccino** le soleil commence a sortir il serait temps pour le moment on a que

⁸⁸ Par exemple, pour l'emprunt *jam* on trouvera des formes telles que *jamais*, *jambe*, *unijambiste*, *ischiojambier*, *Jamaïque*, etc., ainsi que des formes composées non analysables comme *Pearl Jam*.

19° j'ai mis le lave vaisselle et la machine en route !!!quentin et sa copine doivent venir cet a-m elle doit couper les cheveux à clément moi j'attends avant de partir à moins qu'elle juge qu'on peut couper un peu aujourd'hui et recouper dans 3 semaines on verra , je ne fais pas de couleur avant le départ avec la cure pas possible » - edmonde59 (*Doctissimo*), 2011-06-26

- « coucou j'suis là c'est bientôt le week end yaya j'espère que tu es mieux que ce matin lolol moi je ne suis pas jalouse mais c'est pas un truc que j'apprécierai vraiment je vais me faire un **capuccino** !!! » - edmonde59 (*Doctissimo*), 2007-02-16

4.3.3 Volume réel et relatif des corpus définitifs

Corpus	Occurrences examinées ⁸⁹	Mots par article/ commentaire	Total de mots estimé	Proportion d'utilisation des occurrences sur corpus total ⁹⁰
Formel québécois	7514	436	3 276 104	1,07%
Formel européen	7874	539	4 244 086	0,59%
Informel québécois	4294	49.77 ⁹¹	213 712	0,043%
Informel européen	7528	51.07 ⁹²	384 454	0,037%

Tableau 2 : Volume réel et relatif des corpus

Les quatre corpus comportent 8 118 356 vocables au total, regroupant des articles de journaux complets (corpus formels) et des commentaires de forums entiers ou partiels (corpus informels) comportant au moins une fois chacun une unité examinée. Par rapport au volume total des vocables possibles à analyser dans les sources à partir desquelles nous avons constitué nos corpus (nous évaluons ce volume potentiel à 2,562 milliards de vocable), nos quatre corpus correspondent à 0,3% de ce volume⁹³.

⁸⁹ En prenant le nombre d'occurrences examinées, nous supposons que chaque occurrence corresponde à un seul article de journal ou commentaire de forum. Nous multiplions le nombre d'occurrences par la moyenne de mots par article ou commentaire et nous estimons ainsi la grosseur de notre corpus.

⁹⁰ Les résultats de cette colonne sont obtenus par le rapport du total de mots estimés (colonne précédente) sur le volume estimé de vocables dans les corpus potentiels entiers (tels que vus dans le tableau 1). Dans le même ordre de corpus, rappelons le nombre de vocables estimés : 306 millions, 715 millions, 462 millions (*Domaine Bleu*) et 32 millions (*Recettes Québec*), 453 millions (thématique) et 594 millions (non thématique).

⁹¹ Total obtenu en faisant la moyenne des moyennes des forums le composant, soit 56,46 mots par commentaires pour le forum *Domaine Bleu* et 43,08 pour *Recettes Québec*.

⁹² Ce nombre est à relativiser. Les commentaires de forums peuvent avoir des longueurs très variées, allant d'un mot à des centaines. Nous avons fait la moyenne du nombre de mots par commentaire par forum thématique et non-thématique, puis nous avons fait la moyenne de ces deux chiffres. La moyenne finale de 51.07 mots par commentaire est loin de la moyenne thématique (67.61 mots par commentaire) et de la moyenne non-thématique (34.53).

⁹³ C'est-à-dire que nos quatre corpus (8 118 356 vocables) représentent 0,3% du volume potentiel des sources des corpus (2,562 milliards).

Si l'on regarde chacun corpus individuellement, la proportion d'utilisation des unités examinées sur le total de chaque corpus constitués varie largement, entre 1,07% pour le corpus formel québécois et 0,037% pour le corpus informel européen. Le premier est près de 29 fois plus « présent » que le second. Les occurrences sont plus utilisées (plus « présentes ») en corpus formel qu'en corpus informel. De même, la différence entre les deux corpus formels est presque du double et nous ne pouvons expliquer ces deux constats par la fréquence. Ces différences ne s'expliquent pas par le fait que les corpus soient plus volumineux; si le corpus informel européen est le plus massif et le moins « présent », le corpus informel québécois est le deuxième plus petit et pourtant très peu « présent ». Celui-ci est 44% moins volumineux que le premier et environ aussi « présent ».

Les corpus européens totaux ont un nombre d'occurrences beaucoup plus élevé (715 millions et 1,047 milliard) que les corpus québécois totaux correspondants (respectivement 306 millions et 494 millions), et pourtant, les corpus québécois utilisent plus fréquemment les unités (surtout considérant leur petite taille) que leurs homologues européens. À cela nous n'avons également aucune piste suffisante pour expliquer ce constat.

4.4 Identification du genre

4.4.1 Genre des occurrences des emprunts en corpus

Une fois les corpus constitués et les préanalyses terminées, nous avons utilisé un concordancier (Antconc) afin d'identifier la proportion d'emploi de chaque genre pour chaque emprunt dans chaque corpus. Avant chaque analyse, nous nous sommes assuré de sélectionner un encodage permettant l'affichage des caractères diacritiques (*UTF-8* par défaut, changé pour *Western Latin (iso-8859-1)*). Plusieurs fichiers du corpus ont souvent été analysés en même temps, en petits groupes de trois à cinq unités, sauf s'il y avait une grande similitude conceptuelle entre les réalités (*burqa* et *niqab*). Cette méthode permet d'éviter un bon nombre de doublons et explique pourquoi plusieurs de nos résultats dépassent 100 occurrences par unité.

Lorsque les fichiers ont été chargés dans le corpus d'Antconc, nous avons lancé la concordance par un pivot qui isolait le contexte gauche et droit de la phrase autour du pivot. De manière générale, nous avons trié les occurrences afin que le premier mot à gauche du pivot soit identifiable rapidement (généralement un déterminant). Le deuxième point de tri était généralement le deuxième mot à gauche du pivot (plus souvent un déterminant, lorsque pertinent) et le troisième point de tri était le premier mot à droite du pivot (un adjectif). Après un tri rapide, nous avons compté le nombre d'occurrences portant la marque du masculin ou du féminin et avons entré dans notre grille de calcul le nombre approprié.

Lorsque nous avons à compter le nombre d'occurrences d'une unité présentant plus d'une graphie, nous avons utilisé un symbole non-alphanumérique (comme « ? » et « * ») qui nous permettait de chercher un ou plusieurs caractères alphanumériques. Le symbole « ? » a été utilisé pour identifier les caractères avec diacritiques, car il représente n'importe quel caractère alphanumérique⁹⁴. Le caractère « * » été utilisé afin de tronquer un mot et d'identifier un ou plusieurs caractères alphanumériques, jusqu'à ce qu'il arrive à un pause. Dans le cas des mots à large variation graphique (*cappuccino*, *cannelloni*), nous avons utilisé une expression rationnelle⁹⁵ qui nous permettait d'identifier toutes les variantes d'une unité dans la même requête.

Pour terminer, nous n'avons pas considéré les occurrences où le genre donné par le déterminant et par l'adjectif était contradictoire. Une phrase telle « la dernier clip » ne laisse pas savoir quel est le genre de *clip*, alors ces occurrences sont restées écartées.

4.4.2 Genre majoritaire pour analyse par critère

Une fois que le total des occurrences pour chaque genre de chaque unité a été calculé pour chaque corpus, nous avons décidé d'établir un genre majoritaire (GM), un total statistique indiquant à quel genre les locuteurs associent le plus souvent l'emprunt. Ce genre a été établi afin de pouvoir comparer les genres en corpus avec les genres supposés des critères à l'étude. Le genre majoritaire devait résumer l'ensemble des quatre corpus et il a été calculé selon l'algorithme suivant :

- Pour calculer et obtenir un GM qui n'est pas « variable » (V), l'unité devait avoir au moins trois genres à plus de 80% de masculin (M), ou moins de 20% (le féminin [F]), qu'on peut aussi appeler *genres décisifs*. Pour chaque corpus, nous avons regardé le pourcentage de masculin de chaque corpus.
 - A. Advenant le cas où il y a au moins trois des quatre corpus avec un genre décisif, afin d'établir le GM, on laisse de côté les cas où le genre était variable. On suit donc ces directives pour décider du GM :
 1. S'il y a autant de M que de F, le GM est considéré V.
 2. S'il y a plus de M que de F, le GM est considéré M.
 3. S'il y a plus de F que de M, le GM est considéré F.
 - B. Advenant le cas où il y a moins de trois des quatre corpus avec un genre décisif, on prend en compte le pourcentage de masculin de tous les corpus. Pour chaque corpus :
 1. Si le pourcentage de masculin oscille entre 50% et 100%, on le considère comme M.
 2. Si le pourcentage de masculin oscille entre 0% et 50%, on le considère comme F.
 3. Si le pourcentage de masculin est de 50%, on l'ignore du calcul du GM.
 4. On répète les consignes 1 à 3 qui décide du genre majoritaire.

Suivant l'algorithme ci-haut, le genre majoritaire est d'abord attribué aux unités ayant trois ou quatre corpus à genre décisif (c'est-à-dire ayant 80% d'occurrences pour le même genre), puis aux unités ayant moins de trois

⁹⁴ En effet, chercher *e* ne nous aurait pas permis de trouver *é*, *è* ou *ê*.

⁹⁵ Une expression rationnelle (ou régulière, de l'anglais *regular expression*) permet de chercher une suite de caractères spécifiques. Par exemple, l'expression « *cap(p)?uc(c)?in(n)?o* » permet de trouver l'unité *cappuccino* avec un ou deux *p*, *c* et *n*.

corpus à genre décisif. C'est ce qui explique que *ricotta* sera féminin (F+F+(V)+F), tout comme *ciabatta* (V[F]+V[F]+V[M]+V[F]), alors que *van/vanne* sera de genre variable (F+M+F+M), tout comme *minestrone* (V[F]+M+F+V[M]). Rappelons que ce que nous appelons *genre majoritaire* est un artéfact statistique ou « mathématique » que nous utilisons afin de pouvoir comparer le genre d'une unité avec le genre qu'un critère d'attribution exige afin d'être considéré comme valide. *Genre majoritaire* ne renvoie donc pas à un genre physique ou sémantique, ni à des caractéristiques sémantiques ou à un quelconque critère d'analyse.

4.5 Remplissage de la grille d'analyse

Une fois que les genres en corpus ont été identifiés, ainsi qu'un genre majoritaire a été établi pour chacune des unités, nous avons établi une grille d'analyse, aussi appelée *matrice*, qui regroupe chacun des critères liés à la forme physique et au sens, ainsi que le critère de conservation du genre. Pour chaque critère qui sera exposé dans cette section, nous avons comparé le genre majoritaire (établi plus haut) avec le genre que chaque critère dit justifier. Nous avons ensuite calculé un pourcentage de justification, pourcentage illustrant à quel point le critère peut justifier le genre d'un emprunt de notre échantillon.

Nous envisageons la possibilité que certains critères puissent se combiner et dominer sur d'autres; cependant, il n'est pas dans les objectifs de notre recherche de tester et de quantifier les interactions entre les critères d'attribution du genre. Voici maintenant comment nous avons rempli la matrice selon chaque critère.

4.5.1 Critères de forme physique

4.5.1.1 Réanalyse morphologique, aspects phonétique et graphique

Nous avons décidé de séparer le critère de la réanalyse morphologique en deux critères de forme physique, selon la forme graphique ou la forme phonétique. Pour chacun de ces critères, nous avons identifié de notre expérience, ainsi qu'à l'aide du dictionnaire *Robert Brio* (Rey-Debove 2004), les finales graphiques et phonétiques qui sont des morphèmes attestés en français sur d'autres mots que des unités à l'étude.

Les finales phonétiques retenues correspondent souvent à des morphèmes différents. Par exemple, le phonème final /a/ peut être écrit *-a* (*diva, marina, opéra*) ou *-at* (*secrétariat, syndicat, doctorat*), le phonème /o/ peut être orthographié *-o* (*cargo, bravo, commando*), *-eau* (*éléphanteau, barreau, jambonneau*) et *-ot* (*chiot, caillot, cageot*). Nous avons complété chaque case de la matrice et nous avons identifié le genre de chaque finale. Les exemples ci-dessus illustrent que la finale phonétique /a/ pourrait être réanalysée comme féminine (*-a*) ou masculine (*-at*), alors que la finale /o/ ne pourrait être réanalysée que comme masculine (*-o, -eau, -ot*).

Nous avons fait le choix de ne retenir que les formes au singulier de chaque morphème afin d'éviter de dédoubler le nombre d'occurrences d'emploi de chaque genre. De ce fait, nous avons également exclu la

finale *-aux*, malgré que sa forme au singulier soit phonétiquement différente (*-a*). Nous n'avons considéré que des morphèmes qui ont un sémantisme transparent. Par exemple, la finale *-igue* des mots *fatigue*, *gigue*, *figue* a été exclue, puisqu'elle n'est pas à priori significative (comparativement à la finale *-oune* dans la note 45), mais nous avons inclus la finale *-èse* de *genèse* ou *électrophorèse*, car on peut reconnaître dans *genèse* les formes *générer*, *génération*, et dans la forme *électrophorèse*, le préfixe *électro-* qui amène une part significative et plus transparente à un terme spécialisé. De même, la finale *-anne* de *panne*, *canne* ou *manne* n'est pas à priori significative et n'a donc pas été retenue pour *vanne*.

Dans le cas des finales graphiques, nous avons suivi la même démarche, mais avec une forme graphique uniquement. Nous avons été fidèle au *Robert Brio* en ce qui a trait aux diacritiques. L'unité *bonsaï* n'a donc pas trouvé de correspondant, car son tréma distingue la finale de l'unité des autres finales.

4.5.1.2. Syllabicit   enti  re et finale

Nous avons s  par   la syllabicit   en deux crit  res. La syllabicit   enti  re   value le nombre de syllabes dont est form  e une unit  . Selon Surridge (1982), en corpus qu  b  cois, les noms monosyllabiques portent environ autant la marque du masculin que celle du f  minin, alors qu'en corpus fran  ais, les monosyllabiques portent    environ 95% la marque du masculin. Nous avons donc uniquement consid  r   le genre dans nos deux corpus europ  ens. Pour v  rifier dans quelle mesure ce crit  re peut justifier le genre des emprunts, l'unit   devait   tre monosyllabique et devait porter la marque du masculin. Si l'unit     tait polysyllabique, son genre n'  tait pas consid  r   dans la validation de ce crit  re.

Dans le cas de la syllabicit   finale graphique, nous nous sommes bas   sur une hypoth  se de Harley (1979) et Carroll (2005) selon laquelle les noms se terminant en voyelle (phon  tique) sont davantage masculins. Nous avons invers   cette tendance en nous basant sur les exemples donn  s par les auteurs (*fran  ais-fran  aise*), exemples sugg  rant que les noms se terminant par une voyelle (graphique) sont davantage f  minins. Pour   valuer la capacit   de ce crit  re    justifier le genre d'emprunts de ce crit  re, nous avons consid  r   que les unit  s se terminant par une voyelle devaient porter la MAF et que celles se terminant par une consonne devaient porter la MAM. Advenant la possibilit   de plusieurs graphies (*van-vanne*, *bug-bogue*), le mot a   t   ignor  , car il pouvait autant se terminer par une voyelle que par une consonne.

4.5.1.3 Finale graphique

Pour ce crit  re, nous avons bas   nos chiffres sur les r  sultats obtenus par Lyster (2006). Le genre associ      chaque finale graphique a   t   compar   au genre majoritaire de chaque emprunt. Les finales identifi  es « ambig  es » par l'auteur ont   t   consid  r  es comme variables. Dans le cas d'une finale absente des donn  es de Lyster, nous avons ignor   des calculs les unit  s concern  es (*-ew* d'*interview*, *-eed* de *speed*, *-ebs* de *clebs*).

Advenant le cas d'un mot majoritairement ambigu et d'une finale ambiguë, le critère a considéré cela comme une correspondance, donc un cas suivant la règle, un cas en faveur du critère.

4.5.1.4 Similitude graphique

Pour ce critère, il a été difficile de trouver des unités présentant une similitude graphique notable, car plusieurs des emprunts à l'étude possèdent au moins un graphème ou une combinaison de graphèmes plutôt rare en français⁹⁶. Lorsqu'un mot indigène au français possède une consonne finale prononcée, il est habituel que celle-ci soit suivie d'un <e> muet ou caduc (*pape, bouffe, brousse, île, père*), malgré que les exceptions soient nombreuses (*tour, sens, il*). Or, la majorité des unités que nous avons sélectionnées (57/100) se terminent par une consonne graphique (*niqab, djihad, clebs, break, scoop, gig*) ce qui complexifie la tâche de trouver des correspondants graphiques en français.

Afin de trouver ces correspondants nous nous sommes servi de notre intuition de locuteur québécois. Ceux-ci pouvaient avoir un graphème de plus ou de moins que l'unité à l'étude, conservant la majorité des graphèmes originaux, et ceux-ci devaient rester dans le même ordre, sauf advenant le cas d'un groupe consonantique qui pourrait être inversé (*party* : <rt> devient <tr>, <y> devient <i>, ajout du <e> final, résultat : *patrie*)⁹⁷. Nous avons accepté les correspondants à plus d'un graphème de l'original, dans la mesure où l'unité se retrouvait intacte dans un autre mot plus long (*mug* dans *muguet*, *obi* dans *Obi Wan Kenobe*). Les emprunts polysyllabiques (surtout ceux provenant d'une autre langue que l'anglais) ont été les plus difficiles à combler. Au final, 26 unités ont été adjointes d'au moins un correspondant.

4.5.1.5 AP par phonèmes final et pénultième

Pour ce critère, nous avons considéré les chiffres de Tucker *et al.* (1970, 1977 :36-37) qui ont étudié le rapport entre le phonème final d'un mot et le genre de ce mot. Le taux considéré par les auteurs pour déterminer le genre correspondant à un phonème final était de 55%. Les phonèmes ayant un taux supérieur à 55% ont été considérés comme généralement masculins, les phonèmes ayant en deçà de 45% ont été considérés comme généralement féminins. Selon ces auteurs (1977:37), la majorité des mots se terminant par une voyelle portent la MAM, sauf s'ils terminent par /i/ et /ĩ/ qui sont des « voyelles féminines » et /e/ dont le taux ne permet aucune association claire à un genre (50,1% masculin). Pour ce qui est des consonnes, environ la moitié d'elles portent la marque du masculin et l'autre moitié porte celle du féminin, sauf /p/ et /t/ qui sont des phonèmes ni masculin ni féminin. Cependant, si on questionne la validité d'un taux de 55%, on est en droit de se demander la raison de ce taux et pourquoi il n'est pas plus restrictif (considérant que le taux de Lyster (2006) devait être supérieur à 90%). Nous avons essayé de modifier le taux à 60% et cela ne change que la

⁹⁶ La lettre <k> est généralement réservée aux emprunts, comme la lettre <w>. Les combinaisons <ck> ou <dj> sont elles aussi plutôt rares en dehors des emprunts, de même que les combinaisons <oo>, <ee> et <ea>.

⁹⁷ La plupart des lettres sont identiques (4/6) dans une structure syllabique similaire (CVCCV).

valeur du //, passant de masculin à « ambigu » (ni masculin, ni féminin). En montant le taux à 70%, sept consonnes féminines et deux masculines (/b/ et /k/) deviennent ambiguës. Cette mesure pénalise davantage les finales féminines que masculines, et amènent donc des résultats fortement en faveur du masculin (il reste trois phonèmes féminins pour 14 masculins). Puisque nous voulions interroger de manière la plus objective possible le critère tel qu'établi par Tucker *et al.* (1970, 1977), nous avons décidé de conserver le taux de 55%, tel que les auteurs l'ont décidé. Ont été considérés masculins les phonèmes suivants : /ʒ/, /m/, /f/, /r/, /g/, /k/, /b/, /l/, /œ̃/, /ã/, /ɛ̃/, /ø/, /o/, /ɛ/, /u/, /a/, /y/. Ont été considérés féminins les phonèmes suivants : /z/, /n/, /v/, /j/, /ʃ/, /d/, /s/, /ʒ/, /i/, /ɔ̃/. N'ont été considérés ni masculins ni féminins les phonèmes suivants : /p/, /t/, /e/.

Avec ces phonèmes assignés à un genre, nous avons procédé de manière méthodique en regardant quels étaient les phonèmes finaux et pénultièmes (selon les prononciations standards québécoises et européennes). Advenant que le phonème pénultième soit /ɔ/ (*drum*), nous avons considéré le phonème /o/, puisque celui-ci n'est pas absent des calculs de Tucker *et al.* (1970, 1977). De la même manière, /ɑ/ a été converti en /a/. Les phonèmes /œ/ et /ə/ étaient également absents. Ils ont donc été ignorés. Les phonèmes affriqués⁹⁸ (/ts/ de *shiatsu*, /dz/ de *pizza*) ont été considérés uniquement avec le dernier phonème. Les diphtongues (/ai/ de *bonsai*) ont été considérées selon chaque phonème individuel (/a/ et /i/).

4.5.1.6 AP par rime riche

Pour ce critère, nous avons consulté le *dictionnaire des rimes* de Larousse (Warnant 2007) et nous n'avons conservé que les mots dont les trois derniers phonèmes (ou davantage) étaient identiques aux unités examinées. Donc, dans le cas de *shit*, les signes *frite*, *site*, *mite*, *bite* ou *bitte* ne sont donc pas considérables, car ils ne partagent que deux phonèmes, la finale [it].

Nous n'avons considéré que les substantifs au singulier, donc les lemmes. Pour les consonnes affriquées, nous n'avons considéré que le deuxième son (donc pour *carpaccio*, [aʃo] et non [aʃo] ou [ato]). Nous avons ensuite extrait de la suite de rimes le genre le plus fréquent et nous l'avons comparé aux corpus. Nous avons considéré les prononciations québécoise et européenne (*soja* et *soya*), ainsi que la prononciation que nous avons jugée de « conservatrice » dans le cas des emprunts italiens (donc *tagliatelle* [ɛl] et [ɛle]). Lorsque notre intuition trouvait une rime absente du dictionnaire de rimes, nous l'avons ajouté à la liste. Cela n'est cependant pas arrivé plus d'une ou deux fois.

4.5.1.7 Paronymie

Pour ce critère, nous avons puisé les noms comparables à partir de notre intuition de locuteur francophone natif, ainsi que du *Grand Robert en ligne* (2014). À partir de ce dernier, nous n'avons conservé que les

⁹⁸ Ou assibilés, nous n'entrons pas dans la définition de ces termes, et nous les considérons ici comme synonymes.

homonymes (principalement homophones). Nous avons exclu les dérivés, tel que *intervieweur* pour l'emprunt *interview*.

Pour les mots que nous avons tirés de notre intuition, nous avons choisi de ne conserver que des mots qui ont le même nombre de phonèmes que l'unité originale⁹⁹, et n'ayant qu'un seul phonème de différent. Nous avons décidé, au départ, de ne conserver que les comparables suivant ces règles et ayant au minimum un trait sémantique commun, mais nous nous sommes vite rendu compte qu'il n'y aurait presque pas de résultats. Nous avons donc abandonné le critère sémantique et nous avons considéré le critère de paronymie uniquement sur le plan de la forme physique.

4.5.1.8 Initiale vocalique

En dernier lieu, nous avons comparé nos résultats en corpus québécois à ceux de Barbaud, Ducharme et Valois (1982:115). Ceux-ci ont démontré qu'en français québécois, les mots à initiale vocalique ont davantage tendance à être féminisés que masculinisés. Malheureusement, parmi nos 100 unités, on n'en compte que neuf qui commencent par une voyelle. Nous avons considéré les noms commençant par un *h* graphique comme débutant par une voyelle phonétique, mais nous avons affecté à ces unités la valeur d'un genre variable. Afin de vérifier cette hypothèse, les unités devaient porter la marque du féminin ou, au minimum, montrer une variation substantielle de genre (variable) en corpus québécois; nous basons cette méthode sur le fait que le féminin peut représenter le résultat de la féminisation des unités¹⁰⁰ et que le genre variable peut représenter le processus non accompli de celle-ci.

4.5.2 Critères sémantiques

Voyons dans cette section comment, dans notre matrice, les critères sémantiques ont été évalués.

4.5.2.1 Intégration paradigmatique par similitude conceptuelle

Afin de compléter le critère de l'intégration paradigmatique dans la matrice, nous avons dû créer un paradigme latent, formalisé à travers une conception latente¹⁰¹ à l'aide, entre autres, de traits sémantiques. Un exemple de conception latente serait, pour le cas de *slam* : « [événement] poétique/artistique parlé/scandé ». Pour chaque conception latente, nous avons assigné des unités lexicales qui selon nous intègrent le paradigme de chaque conception. Donc, pour *slam*, on retrouverait *récitation*, *happening*, *rap* et *discours*. Par la suite, nous avons cherché dans tous les corpus les combinaisons les plus fréquentes d'un verbe et de l'unité. Nous nous sommes assuré que chaque unité lexicale du paradigme latent puisse s'utiliser avec au moins un des

⁹⁹ Ce choix semble restrictif, mais il fallait établir une barrière afin de ne pas accepter n'importe quel comparable avec une unité. Cette barrière nous semblait assez souple pour les mots monosyllabiques, voir bisyllabiques, et assez rigide pour les mots à plus de deux syllabes.

¹⁰⁰ Si l'on s'en tient à Barbaud, Ducharme et Valois (1982).

¹⁰¹ Nous employons *conception latente* ici comme la formalisation de ce qui unit chaque unité lexicale et concept dans un paradigme latent.

syntagmes verbaux les plus fréquents¹⁰². Dans le cas de *slam*, les verbes les plus fréquents étaient *écouter*, *écrire*, *faire*, *découvrir*, *improviser* et *s'initier à*. Après avoir relevé les combinaisons les plus fréquentes, nous avons déduit de chaque paradigme latent le genre le plus fréquent, genre supposément associé à la valeur sémantique hypothétique prédominante du paradigme, reflétant le dynamisme de la représentation. Par exemple, le paradigme latent intégrant l'emprunt *slam* montre une valeur sémantique plutôt statique, stable et passive associée au genre masculin. Pour la dernière étape dans l'évaluation de ce critère, nous avons comparé le genre associé au paradigme avec le genre majoritaire en corpus. Dans le cas de *slam*, le genre majoritaire en corpus correspond au genre typiquement associé au paradigme.

Dans le cas particulier de la paire *jam* et *jam-session*, nous avons considéré que ce sont deux formes d'une même unité sémantique. Ils intègrent donc le même paradigme et sont combinés et représentés de la même manière. Puisque le genre majoritaire de chaque variante est différent, on ne peut évaluer correctement la validité du critère que si l'on analyse soit les deux soit aucune. Il faut donc considérer qu'il y a un total de 98 unités analysables différentes, et non 100.

4.5.2.2 Intégration paradigmatique par hyperonymie

Pour ce critère, nous avons cherché chaque unité dans le moteur de recherche *Google*¹⁰³ avec la forme suivante : « X est un type de ». La lettre X a été remplacée par l'unité désirée. Les guillemets garantissent que les mots se retrouveront dans cet ordre et qu'ils sont tous nécessaires à la requête. La formule a été utilisée au singulier et au pluriel avec les noms de pâtes italiennes, car il est fréquent de voir le nombre varier avec ces unités, considérant que ces noms sont déjà des pluriels morphologiques en italien, utilisés souvent au singulier en français. Nous avons également substitué *type* par les termes génériques *sorte* et *genre*, car ils permettent une plus grande étendue d'hyperonymes et que certains hyperonymes pourraient entrer dans des combinaisons plus fréquentes avec un des trois termes génériques¹⁰⁴. En effet, la formule *X est une sorte de Y* est une formule classique exposant l'hypo-/hyperonymie d'un réseau sémantique (Levrat & Sabah 1990:89, Murphy 2010:114-115). Or, cette formule permet d'exposer d'autres termes qui ne sont pas à priori hyperonymique, mais qui ont une fonction hyperonymique. Ces autres termes sont des correspondants qui ne sont pas proprement des classes taxonomiques supérieures, mais des cohyponymes, les plus saillants de leur paradigme, appartenant à une même classe générale qui agissent comme un hyperonyme en prenant une valeur générique, dans le cadre de la forme « X est un type de Y ». Nous avons conservé ces correspondants

¹⁰² Au-delà des syntagmes verbaux, nous avons également accepté des syntagmes prépositionnels lorsque ceux-ci étaient très fréquents (*à bord de* dans le cas de *van/vanne*). Dans les syntagmes verbaux retenus, on trouve des infinitifs (*pondre* pour *haïku*) et des participes passés passifs (*détruit par* pour *tsunami*).

¹⁰³ Nous avons uniquement consulté les dix premiers résultats par formule, bien qu'il fût rare de voir plus de 10 résultats.

¹⁰⁴ Nous avons décidé de ne pas inclure la formule « X est une espèce de », car nous avons constaté qu'elle entraînait trop de métaphores.

dans la matrice et nous les appelons des **hyperonymes analogiques**¹⁰⁵. Par exemple, on peut dire que « pancake est un type de crêpe », sans pour autant dire que *pancake* appartient à la classe *crêpe*, mais que les deux appartiennent à une classe supérieure commune, classe qu'on pourrait formaliser avec des traits sémantiques ainsi : « [comestible] [pâte chauffée] [+solide] [+plat] ».

Nous avons considéré tous les mots qui apparaissaient à la suite de la formule « X est un type de Y », sauf si ceux-ci présentaient un trop grand niveau de métaphoricité. Par exemple, « le tofu est une sorte de fromage de soja » est acceptable, car *fromage* et *tofu* partagent une apparence similaire, relevant davantage de la métonymie que de la métaphore; à l'inverse, « le tofu est une sorte de caméléon, il se mariera au goût que vous lui donnerez » n'est pas acceptable, car le lien qui unit *tofu* et *caméléon* est moins concret que le précédent lien, ne relevant pas du référent, mais du sens contextuel seulement. Nous avons donc jugé chaque hyperonyme analogique individuellement afin de voir si la relation entre les termes relevait de la réalité ou d'une métaphore trop élaborée. À noter que les hyperonymes trouvés étaient rarement des termes hypergénériques tels que *lieu*, *chose*, *aliment*, *nourriture*.

4.5.2.3 Intégration paradigmaticque par ellipse d'un signe entier

Dans le but de voir si, avec les unités à l'étude, certaines combinaisons d'unités lexicales auraient pu être réduites, nous avons, tout au long de la constitution de notre corpus, noté les termes fréquents qui apparaissaient conjointement avec l'unité, sans préposition. Nous avons également mené une recherche dans *Google Books* en cherchant des livres en français datant d'au moins 40 ans. Aucun de termes élidés identifiés ne provient de notre intuition. Au final, 34 des 100 unités possèdent au moins un terme élidé.

4.5.3 Conservation du genre

Afin de pouvoir confirmer ou infirmer l'hypothèse que nous avons faite sur la conservation du genre des emprunts venant de LAG, nous avons consulté le site des dictionnaires Collins¹⁰⁶ pour l'italien. Pour vérifier le genre d'origine des emprunts arabes, nous savons que le genre de certains mots peut changer selon les dialectes arabes (comme en français). Plutôt que de réaliser un examen précis du genre arabe, nous avons tenu à considérer le genre que nous donnerait un informateur originaire du Maghreb, puisque la France a eu de grands contacts avec ces régions depuis près de deux siècles. Par conséquent, nous avons consulté une locutrice native de l'arabe originaire du Maghreb.

¹⁰⁵ Notre appellation.

¹⁰⁶ <http://www.collinsdictionary.com/dictionary/english-italian>

5. Résultats et analyse

Considérant la somme de résultats que nous avons, nous allons procéder méthodiquement à la présentation des résultats d'un aspect, puis à l'analyse de ces résultats, pour ensuite passer à l'aspect suivant, et ainsi de suite. Nous commencerons par la présentation des résultats de manière générale pour chaque corpus, puis nous examinerons les critères d'analyse un par un, pour finalement observer un certain nombre de cas remarquables. Des tableaux en annexes montreront les résultats de chaque critère en détail.

Rappelons que nos résultats sont basés sur des usages produits par des locuteurs francophones, généralement francophones natifs, regroupés en corpus. Les résultats sont donc fonction des énoncés dont ils sont formés. Précisons que nos résultats analysent en surface les différences entre les différentes communautés ou registre de formalité. Nous avons choisi d'abord ce sujet en surface, car il ne permettait pas de tirer suffisamment de conclusions quant aux critères que nous avons choisis. Néanmoins, le survol que nous effectuons en 5.1 et 5.2 servent de lignes directrices pour les tendances observées.

5.1 Résultats généraux

5.1.1 Nombre d'occurrences d'emploi pour chaque genre

Avant de jeter un regard sur les résultats, revoyons la composition et la taille des corpus de manière simplifiée.

Corpus	Composition	Nombre d'occurrences des unités examinées	Total estimé de vocables ¹⁰⁷
Formel québécois	<i>La Presse, Le Soleil, Le Journal de Québec, Le Nouvelliste, Le Droit, Le Devoir</i>	7514	3 276 104
Formel européen	<i>Le Monde, Libération, Le Figaro, Sud Ouest, Le Temps, Le Soir</i>	7874	4 244 086
Informel québécois	<i>Domaine Bleu et Recettes du Québec</i>	4294 ¹⁰⁸	213 712
Informel européen	<i>Doctissimo</i>	7528	384 454

Tableau 3 : Composition et taille réelle des corpus

¹⁰⁷ Afin d'arriver à ce nombre, nous avons multiplié le nombre d'occurrences des unités à l'étude par le nombre moyen de mots par commentaire de forums ou articles préalablement calculé.

¹⁰⁸ Rappelons que ce corpus est le plus petit des corpus parce que ses sources sont plus petites ou comportent plus de doublons (à ratio équivalent) que les autres corpus.

Nous avons calculé le nombre d'occurrences des unités lexicales examinées portant la marque du masculin ou du féminin pour chaque corpus. Voyons le tableau quatre, il présente le nombre d'occurrences d'emploi de chaque genre par corpus :

Type de corpus	Nombre d'occurrences des unités examinées portant une MAM	Nombre d'occurrences des unités examinées portant une MAF
Formel québécois	5808 (77,3% de masculin)	1706 (22,7% de féminin)
Formel européen	5859 (74,4% de masculin)	2015 (25,6% de féminin)
Informel québécois	3449 (80,3% de masculin)	845 (19,7% de féminin)
Informel européen	5710 (75,9% de masculin)	1818 (24,1% de féminin)
<i>Total</i>	20826 (76,5% de masculin)	6384 (23,5% de féminin)

Tableau 4 : Nombre d'occurrences d'emploi de chaque genre sur toutes les unités à l'étude par corpus

Au premier regard, on remarque que le nombre d'occurrences du masculin dépasse toujours celui des occurrences du féminin. Entre 74% et 80% des occurrences des unités sélectionnées portent la marque du masculin, tendance qui suit ce que la littérature a relevé pour les emprunts à l'anglais (voir note 36). Or, puisque notre échantillon est également composé d'emprunts au japonais, à l'arabe et à l'italien, nous devons regarder quelle tendance on peut dégager lorsqu'on observe par langue et par corpus. Le tableau suivant illustre le nombre d'occurrences (et le pourcentage de masculin) pour chaque langue et chaque corpus :

Corpus	Anglais	Japonais	Arabe	Italien	Total
Québécois formel	1805 M + 343 F = 2148 (84%) ¹⁰⁹	1636 M + 27 F = 1663 (98,4%)	1186 M + 694 F = = 1880 (63,1%)	1181 M + 642 F = = 1823 (64,8%)	5808 M + 1706 F = = 7514 (77,3%)
Européen formel	1587 M + 616 F = 2203 (72%)	1769 M + 30 F = 1799 (98,3%)	1456 M + 872 F = = 2328 (62,5%)	1047 M + 497 F = = 1544 (67,8%)	5859 M + 2015 F = = 7874 (74,4%)
Québécois informel	1043 M + 200 F = 1243 (83,9%)	834 M + 3 F = 837 (99,6%)	727 M + 287 F = 1014 (71,7%)	845 M + 355 F = 1200 (70,4%)	3449 M + 845 F = 4294 (80,3%)
Européen informel	1414 M + 235 F = 1649 (85,7%)	1956 M + 23 F = 1979 (98,8%)	1256 M + 875 F = = 2131 (58,9%)	1084 M + 685 F = = 1769 (61,3%)	5710 M + 1818 F = = 7528 (75,9%)
<i>Total</i>	5849M + 1394 F = = 7243 (80,8%)	6195 M + 83 F = 6278 (98,7%)	4625M + 2728 F = = 7353 (62,9%)	4157M + 2179 F = = 6336 (65,6%)	20826 M + 6384 F = = 27210 (76,5%)

Tableau 5 : Nombre d'occurrences d'emploi de chaque genre pour chaque langue par corpus

Le tableau cinq présente des différences importantes dans la distribution du genre pour les emprunts de chaque langue. Si la majorité des emprunts (peu importe la langue ou le corpus) porte la marque du masculin, on observe des différences marquées entre les échantillons provenant des LAG et ceux des LSG. On remarque que les emprunts à l'arabe et à l'italien, sont moins masculins que les emprunts à l'anglais et au japonais. Cette observation va dans la même direction que notre troisième hypothèse, voulant que les emprunts provenant des LSG portent plus la marque du masculin (80,8% et 98,7%) que les emprunts des LAG (62,9% et 65,6%).

Ensuite, on constate une tendance à utiliser plus souvent le genre masculin dans les corpus informels (québécois 80,3%, européen 75,9%), toutes langues donneuses confondues, que dans les corpus formels (même ordre : 77,3% et 74,4%). Bien que la différence soit faible et sans doute peu significative, cela pourrait s'expliquer par le fait que le registre plus formel est de manière générale plus codifié et strict dans son usage que ne l'est le registre informel. Dans le registre informel, les locuteurs sont davantage libres d'utiliser le genre qu'ils croient mieux représenter la réalité désignée par l'emprunt qu'ils utilisent.

Le japonais se distingue des autres langues par la nette prédominance du masculin sur le féminin. En effet, son pourcentage de masculin est très élevé et très stable parmi les corpus (entre 98,3% et 99,6%). À l'inverse, les emprunts à l'arabe montrent un plus faible taux de masculin et une plus grande variabilité de la répartition

¹⁰⁹ Voici comment lire le tableau. D'abord le nombre d'occurrences des unités à l'étude portant la marque du masculin (1805 M), puis le nombre portant la marque du féminin (343 F), suivi du total des deux nombres d'occurrences (= 2148). Finalement, entre parenthèse, le nombre d'occurrences portant la marque du masculin sur le total.

du genre selon les corpus (58,9% à 71,7%). Pour expliquer cette dominance du masculin et cette stabilité du masculin en japonais, nous ne voyons qu'un critère susceptible de le faire uniformément à travers les corpus : l'attribution de l'archigenre. On peut affirmer raisonnablement que la plupart des réalités propres à la culture japonaise (le saké, le sashimi, le nunchaku) sont plus distantes et plus étrangères pour un Européen ou un Québécois que ne le sont les réalités italiennes, anglaises (ou américaines) ou arabes, que ce soit pour des raisons historiques, culturelles ou géopolitiques. Puisque ces réalités sont plus étrangères, elles ont nécessairement une conceptualisation plus vague que les conceptualisations des réalités familières et communes. En conséquence, la conceptualisation vague amène généralement l'attribution de l'archigenre, à défaut d'une conceptualisation spécifique. Thornton (2009:31) disait : « la plupart des emprunts dénote des entités particulières et très spécifiques, qui ont plutôt un statut de subordonné ». Comme nous l'avons vu auparavant (section 2.3.2), le statut taxonomique des emprunts devrait refléter le genre des hyperonymes (de niveau conceptuel de base), hyperonymes qui se partagent les deux genres. Une distinction est à faire ici : si les emprunts ont une conceptualisation vague, rien ne les empêche de dénommer des réalités spécifiques et particulières. Parce qu'on compte en moyenne 76,5% de masculin dans tous les corpus, on peut confirmer que le genre est d'abord attribué par la conceptualisation de la réalité représentée avant d'être attribué selon le niveau taxonomique de cette réalité. Puisque les emprunts au japonais sont conceptualisés de manière plus vague que les emprunts à l'italien, à l'anglais (les langues plus « proches » des réalités des francophones européens) ou à l'arabe (langue à grand contact avec la France), il est compréhensible de voir les emprunts au japonais porter la marque du masculin (et de l'archigenre) en beaucoup plus grande proportion.

5.1.2 Degré de présence des unités par langue par corpus

Regardons maintenant comment les unités sont distribuées dans nos corpus, en faisant abstraction du genre. Ce procédé nous aidera à visualiser comment certaines communautés linguistiques (Québec, Europe) ont certaines affinités avec certaines langues donneuses et leur culture. Le tableau six présente entre parenthèses le pourcentage que constitue chaque groupe de 25 unités d'une même langue et d'un même corpus sur le total du corpus.

Corpus	Anglais	Japonais	Arabe	Italien	Total
Québécois formel	2148 (28,6%)	1663 (22,1%)	1880 (25%)	1823 (24,3%)	7514 (100%)
Européen formel	2203 (28%)	1799 (22,8%)	2328 (29,6%)	1544 (19,6%)	7874 (100%)
Québécois informel	1243 (28,9%)	837 (19,5%)	1014 (23,6%)	1200 (27,9%)	4294 (100%)
Européen informel	1649 (21,9%)	1979 (26,3%)	2131 (28,3%)	1769 (23,5%)	7528 (100%)

Tableau 6 : Proportion d'utilisation des unités par langue par corpus sur le total de chaque corpus

Nous sommes d'avis que la proportion d'utilisation des unités traduit un degré de présence, une proximité culturelle entre le français et la langue donneuse. Ainsi, une plus grande fréquence indiquerait une plus grande présence. Il va de soi que cela ne peut pas se vérifier à travers seulement un tableau, mais nous croyons que cette possibilité nuance d'une manière intéressante les conclusions que nous pouvons tirer de nos résultats.

Le tableau six présente quelques tendances intéressantes. La langue donneuse la plus présente dans les corpus québécois est l'anglais, alors que dans les corpus européens, il s'agit de l'arabe. La langue donneuse la moins présente dans les corpus québécois est le japonais, alors que dans les corpus européens, il s'agit de l'italien (corpus formel) et de l'anglais (corpus informel). Ces constats laissent entendre que le Québec a eu de plus grands contacts (ou des contacts plus saillants) avec la culture anglaise (anglo-américaine ou britannique) qu'avec la culture japonaise – ce qui est vrai, sans l'ombre d'un doute –, alors que l'Europe francophone a eu de plus de contacts avec la culture arabe qu'avec la culture italienne ou la culture anglaise – ce dont on est en droit de douter. L'Europe francophone (particulièrement la France) a eu des contacts continus avec l'Angleterre et avec l'Italie (particulièrement la France et la Suisse) depuis la Renaissance et davantage. Affirmer que l'Europe francophone, particulièrement la France, a eu plus de contacts avec la culture arabe du Maghreb que la culture anglaise ou italienne, sans parler d'une période en général, est matière à débat, mais affirmer que les contacts avec la culture arabe du Maghreb ont été plus fréquents et plus saillants dans les deux derniers siècles semble plausible, ce qui pourrait expliquer la raison de la grande présence des emprunts à cette langue dans le contexte européen.

5.2 Résultats spécifiques

5.2.1 Genre majoritaire

Nous avons expliqué dans la section 4.4.2 qu'il était nécessaire de désigner le genre le plus représentatif d'un emprunt à travers tous les corpus afin de le comparer aux genres que nos critères suggèrent. Nous avons dénommé ce genre « virtuel » le genre *majoritaire*. Il est désormais pertinent à cette étape de notre recherche d'établir quelles unités ont hérité de quel genre majoritaire afin que nous sachions avec quels genres majoritaires nous avons évalué les critères et afin que nous puissions avoir une bonne idée de la distribution des genres majoritaires. Voici donc la liste complète des emprunts selon leur genre majoritaire, classés par langue, puis en ordre alphabétique.

Sont majoritairement masculins :

Anglais : *break, bug (bogue), clip, cake, drum, groove, jam, mug, pancake, party, pipeline, sandwich, scoop, shit, slam, speed, trip*;

Japonais : *bento, bonsaï, dojo, haïku, karaoké, kimono, maki, manga, matcha, miso, nunchaku, obi, saké, sashimi, seitan, soja (soya), shiatsu, sudoku, surimi, sushi, tatami, tofu, tsunami, wasabi*;

Arabe : *barda, bled, brick, chèche, clebs, couscous, djihad (jihad), hidjab (hijab), keffieh, loukoum, méchoui, niqab, sarouel, taboulé, tajine*;

Italien : *antipasto (antipasti), cannelloni, cappuccino, carpaccio, espresso (espresso), panettone, panini, prosciutto, risotto, sabayon, salami, sépia, spaghetti, , tiramisu*.

Sont majoritairement féminins :

Anglais : *dope, interview, jam-session, stout*;

Arabe : *baraka, burqa, charia, chicha, chorba, fatwa, harissa, intifada, merguez*;

Italien : *bruschetta, calzone, ciabatta, farfalle, omertà, pancetta, pizza, ricotta, tagliatelle, vendetta (vendetta)*.

Sont généralement variables :

Anglais : *gig, hype, rave, van (vanne)*;

Japonais : *tempura*;

Arabe : *kefta*;

Italien : *minestrone*.

Au total, lorsqu'on observe et qu'on quantifie le nombre d'unités pour chaque genre majoritaire apposé, 70 unités portent majoritairement le genre masculin, 23 portent majoritairement le genre féminin et 7 sont généralement variables. Sur les sept unités variables, plus de la moitié (4/7) viennent de l'anglais, et les trois autres viennent toutes de langues différentes. Des 23 unités portant la marque du féminin, la grande majorité (19/23) viennent d'une LAG, ce qui nous donne déjà un indice de conservation possible du genre des LAG. À titre de comparaison, 10 unités de l'italien et 11 de l'arabe portent la marque du féminin.

Regardons la répartition du genre selon les langues donneuses. Les emprunts au japonais sont presque uniquement masculins (24/25), sauf *tempura*, qui est de genre majoritaire variable. Par exemple, deux occurrences du même corpus :

- « Ouvert depuis quelques mois, ce restaurant dédié aux sushis, sashimis et autres makis joue la carte du cru auprès d'une clientèle raffolant des spécialités du chef qui confectionne, la main heureuse, des bouchées japonaises originales, toutes présentées aussi joliment qu'un origami ! De l'oeil du dragon, un maki frit au **tempura**, saumon et carotte, au sashimi Massago au caviar d'éperlan, certains noms évoquent ces justiciers des films d'arts martiaux bravant la mort nourris d'un seul bol de riz. » Stéphanie Bois-Houde (*Le Soleil*), 2004-12-04
- « C'est plutôt un restaurant omnivore, comme il en existe des centaines à Montréal. Comme ces restaurants japonais, marocains, libanais ou indiens où l'on peut manger de la **tempura** aux légumes, du couscous aux pois chiches ou des falafels, mais à petit prix. Tous les quartiers chinois de la planète permettent aussi de découvrir du tofu, du faux poulet ou des crevettes à base de soya. » Marjolaine Jolicœur (*Le Devoir*), 2012-12-6

C'est donc dire qu'aucun emprunt au japonais faisant partie de notre échantillon n'est majoritairement féminin, ce qui soutient notre hypothèse de l'attribution de l'archigendre. Les emprunts à l'anglais, à l'arabe et à l'italien suivent relativement le même modèle : la majorité est d'abord masculine, puis minoritairement féminine avec quelques cas de genre variable. Les emprunts à l'anglais se démarquent de ceux à l'arabe et à l'italien en ce qu'ils ont le plus de masculin (en ordre : 17-15-14) et autant d'emprunts majoritairement féminins (4-9-10) que variables (4-1-1). Les emprunts à l'arabe et à l'italien n'ont qu'un emprunt chacun à genre majoritairement variable et ont un nombre d'emprunts majoritairement féminins beaucoup plus grand que dans le cas des emprunts à l'anglais. Ce constat va dans la même direction que nos deuxième et troisième hypothèses, voulant que les emprunts des LAG conservent le plus souvent leur genre d'origine¹¹⁰, alors que les emprunts des LSG portent plus souvent la marque du masculin, à travers l'archigendre.

5.2.2 Retour sur la recherche préliminaire

Dans la section 4.2, nous avons fait une recherche préliminaire avec chacune de nos unités afin d'observer ce que la littérature lexicographique disait au sujet de notre échantillon. Parmi les 100 unités, 13 variaient en genre (*ciabatta*, *calzone*, *harissa*, *interview*, *minestrone*, *party*, *pipeline*, *rave*, *sandwich*, *stout*, *tagliatelle*, *tempura*, *van/vanne*). Or, une fois que le genre majoritaire a été calculé, seules les unités *minestrone*, *rave*, *tempura* et *van/vanne* ont été identifiées comme alternante en genre. Ajoutons *gig*, *hype* et *kefta* qui n'ont pas été relevés comme variables dans ces ouvrages. Si nous illustrons plusieurs occurrences de ces unités dans les prochaines sections, prenons le temps d'illustrer une occurrence pour chaque unité identifiée comme majoritairement variable portant chacune des marques. Commençons par *minestrone*, *tempura*, et *kefta* :

¹¹⁰ Autre constat intéressant, les emprunts se terminant en *-a* sont majoritairement féminins, comme on pourrait s'y attendre, sauf *barda* qui porte la MAM.

- « C'est le secret du **minestrone** où tout est mélangé mais rien n'est écrasé. Les coulisses du défilé Sportmax, c'est un peu pareil. Photographes, journalistes, stylistes, coiffeurs, maquilleuses, équipes de télé se croisent dans un joyeux brouhaha.! » - J.H. (*Le Soir*), 2010-09-24
- « Cette douceur compensait un peu pour le reste d'un repas somme toute bien ordinaire. La **minestrone** est bien faite, pas grasse du tout, et avec plein de légumes et de légumineuses. Mais un cuisinier a oublié de mettre les pâtes canneroni au bon moment et elles sont encore trop "al dente" quand elles sont servies. » - Pierre Jury (*Le Droit*), 2009-03-20
- « Ouvert depuis quelques mois, ce restaurant dédié aux sushis, sashimis et autres makis joue la carte du cru auprès d'une clientèle raffolant des spécialités du chef qui confectionne, la main heureuse, des bouchées japonaises originales, toutes présentées aussi joliment qu'un origami ! De l'oeil du dragon, un maki frit au **tempura**, saumon et carotte, au sashimi Massago au caviar d'éperlan, certains noms évoquent ces justiciers des films d'arts martiaux bravant la mort nourris d'un seul bol de riz. » - Stéphanie Bois-Houde (*Le Soleil*), 2004-12-04
- « «J'ai entendu parler il y a plus de vingt ans des usages possibles des huiles de friture usées, explique Takeo Someya. J'ai appris qu'il était possible de faire rouler des voitures grâce à un diesel tiré de l'huile végétale. L'idée m'est venue de fabriquer de l'essence en recyclant les huiles usées de tempura.» La **tempura**, friture enrobée dans une pâte à base de farine légère, est une spécialité gastronomique très prisée. » - Michel Temman (*Libération*), 2006-05-29
- « enfin.....alors ce soir tagine au **kefta** vite fait et super bon!!!! bon ap'!!!!!!! » - imanouar (*Doctissimo*), 2006-09-28
- « Mais ensuite je rajoute oignons hachés (juste en condiments donc meme en PP) ail, coriandre, persil plat, cumin, poivre, sel, poivre doux, paprika et gingembre. Et c'est vraiment délicieux meme avec une poele anti-adhésive! Ca s'appelle la **Kefta** » - manelle78 (*Doctissimo*), 2008-08-14

Continuons avec *rave* et *gig* :

- « Deux D.J. internationaux, Deadmus 5 et Skrillex, ont participé à ce **rave**. «Réunir ces deux D.J. à Québec un 30 décembre, c'est exceptionnel», lance M. Cinq-Mars. » - Anonyme (*Le Journal de Québec*), 2011-12-31
- « Daniel Caux : du blues à la techno en passant par les avant-gardes, le rock déjanté et la musique arabe. Le DJ Laurent Garnier lui a dédié son concert à la **rave** monstre du désert de Monegros (Aragon) le 12 juillet, jour de sa mort. » Francis Marmande (*Le Monde*), 2008-07-17
- « "C'est la plus grosse **gig** de notre vie!" s'est exclamé le chanteur Sébastien Plante joint hier à L.A. "C'est le rêve de tous les groupes de se retrouver dans une situation comme celle-là et je suis tellement fier de n'avoir jamais lâché prise!" » - Valérie Lesage (*Le Soleil*), 2010-04-27
- « Pas mal comme audience. « Quand on est venu s'installer ici, dit le chanteur Matt Balmer, 22 ans, on ne pensait pas un jour se produire devant autant de personnes. » Le « **gig** » le plus froid de l'histoire du rock aura au moins attiré l'attention sur la fonte alarmante de la banquise. » - Jean-Claude Vantroyen (*Le Soir*), 2007-07-09

Terminons avec *van-vanne* et *hype* :

- « Jack prévient Danny et Martin qui sont en train d'escorter Adisa. Alors qu'ils sont arrêtés à un feu, les portes du **van** qui les précèdent s'ouvrent. Laisant sortir Dornvald et un autre homme, qui leur tirent dessus » - isou510 (*Doctissimo*), 2006-09-05
- « "Il y avait une voie qui était bloquée, mais la personne qui descendait n'a pas tenu compte du présignaleur et elle a quand même voulu dépasser la **vanne**...ça a passé proche", a résumé son patron, Sylvain Simard, de l'entreprise Fernand Gilbert à Chicoutimi. » - Jean-Luc Lavallée (*Le Journal de Québec*), 2008-10-22

- « Dan Smith est né un 14 juillet. D'où le nom de Bastille pour ce projet mêlant pop et électro. Ce n'est que le début du **hype** pour lui, raison pour laquelle le « petit » club a été choisi mais c'est complet évidemment ! » - Anonyme (*Le Soir*), 2013-04-10
- « On pense que la **hype** n'aime que le très neuf ou le très vieux. On pense que la hype est une sale snobinarde élitiste. Et pourtant. La **hype** sait ménager ses effets. Après la sardine en boîte dans les apéros chics, voilà Jacques Chirac intronisé icône de mode. La **hype**, mi-trotskiste, mi-dadaïste, a ses raisons que la raison ne connaît pas. Dans un happening ubuesque, voilà ce bon vieux Jacquot, ce Tony Soprano de Corrèze, devenu l'objet d'un culte intimiste mais très pointu, par la grâce d'un tumblr et d'une page Facebook intitulée FuckYeahJacquesChirac. Depuis, la **hype** a essaimé et l'idée a fait son chemin. Le chanteur Psy, interrogé sur la personnalité qu'il aimerait voir danser le Gangnam Style, aurait naturellement désigné « Djak Chilak ». » - Valérie Fromont (*Le Temps*), 2013-04-20

5.3 Résultats par critères

Dans cette section, nous regarderons le résultat de chaque critère en relation avec les unités de notre étude. Les résultats complets se trouvent en annexes (section neuf). Nous avons utilisé pour cette section un code de couleur symbolique. Dans le cas des corpus, le bleu représente un pourcentage de masculin au-delà de 80%, le gris un pourcentage entre 20% et 80% et le rose un pourcentage en deçà de 20%. Voici un exemple de ce qui est développé dans la section 9.1.1 :

	A	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Unités	Corpus Québec Formel				Corpus Europe Formel				Corpus Québec informel				Corpus Europe informel				
2		Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M	Genre majoritaire
3	gig	9	25	34	26,47	14	1	15	93,33	2	2	4	50	0	0	0	-	v
4	jam	105	2	107	98,13	9	58	67	13,43	25	0	25	100	0	0	0	-	m
5	jam-session	25	11	36	69,44	2	72	74	2,703	2	0	2	100	0	1	1	0	f
6	party	183	0	183	100	1	50	51	1,961	116	0	116	100	75	23	98	76,53	m
7	rave	84	1	85	98,82	0	55	55	0	14	0	14	100	0	11	11	0	v
8	slam	132	0	132	100	141	0	141	100	17	0	17	100	87	0	87	100	m
9	hype	19	17	36	52,78	17	94	111	15,32	20	6	26	76,92	0	1	1	0	v
10	interview	0	100	100	0	0	158	158	0	8	69	77	10,39	8	114	122	6,557	f

Dans le cas des critères, le vert représente une correspondance entre le genre suggéré par le critère (un *oui*), le rouge représente une non-correspondance (un *non*), et le noir représente une correspondance ambiguë (un *peut-être*, ou un *oui* et un *non*)¹¹¹. Voici un exemple de ce qui est développé en section 9.2.1.1 :

¹¹¹ Dans le cas des correspondances ambiguës, nous voulions mettre la couleur jaune ou orange afin d'imiter le code des feux de signalisation (vert-jaune/orange-rouge), mais la couleur jaune sur blanc est difficile à lire, et la couleur orange ressemble trop au rouge. Nous avons donc utilisé le noir, qui paraît bien sur un fond blanc.

	A	U	V	W	X	
1	Unités	réanalyse morphologique phonétique				
2		faux suffixe	exemples	genre	identique?	
16	drum	<i>um</i>	<i>référendum, quorum, forum</i>	m	o	
17	groove	-	-	-	-	
18	pipeline	<i>ine</i>	<i>olivine, marine, canine, bottine, vitrine</i>	f	n	
19	dope	-	-	-	-	
20	shit	<i>ite</i>	<i>bronchite, sinusite, graphite, météorite</i>	f	n	
21	speed	<i>ide</i>	<i>acide, glucide, lipide, oxyde</i>	m	o	
22	van/vanne	<i>ane</i>	<i>octane, méthane, butane</i>	m	n	
23	pancake	-	-	-	-	
24	sandwich	<i>iche</i>	<i>pouliche, barbiche, potiche, caniche</i>	m/f	v	

L'image ci-dessus ne considère que des morphèmes à priori *transparents* et *isolables*, tels que vu en section 4.5.1.1. Puisque, dans le cas de *sandwich* par exemple, le mot peut être prononcé /itʃ/, on pourrait avoir tendance à le réanalyser avec des mots rimant, tel que *bitch*. Or, puisque la réanalyse doit se baser sur un morphème significatif existant et qu'aucune morphème en *-itche* n'existe en français, nous avons ignoré ce morphème. Ainsi il en est pour tous les autres résultats de ce critère. Il en est de même pour *dope* (qui rime de deux sons avec *flop* ou *top*) ou *groove* (avec *move* au Québec), mais leur finale phonétique ne correspond pas à des morphèmes phonétiques.

Afin de calculer dans quelle mesure les critères peuvent justifier le genre d'un emprunt (ce que nous appelons le *taux de justification* de chaque critère), nous avons calculé (en pourcentage) le rapport entre le nombre de correspondances (entre le genre majoritaire en corpus d'une unité et le genre prescrit par un critère pour cette même unité) et le nombre total de cas analysables par le critère (*i.e.* le nombre de cas où le critère est applicable à l'unité¹¹²). Nous avons pris le parti d'établir le seuil de validité des critères à 60%. Les critères ayant un taux de justification (ou de succès, d'efficacité) supérieur à 60% sont considérés comme potentiellement opératoires (*PO*), fonctionnels et méritent une analyse plus en profondeur. Les critères sous le seuil de 60% sont considérés comme faiblement opératoire (*FO*), voire non opératoires, dysfonctionnels et ne fonctionnent qu'en un nombre limité de cas.

5.3.1 Réanalyse morphologique (phonétique et graphique)

Nous avons décidé de regrouper les deux critères de réanalyse morphologique sous une même section du chapitre considérant que les conclusions que nous en tirons sont similaires. Pour ces critères, afin d'être confirmés, il fallait que le genre majoritaire en corpus soit identique au genre du ou des morphèmes (sous leur forme phonétique ou graphique). Rappelons-nous qu'ici un morphème réanalysé correspond à une finale de mot significative en français qui se retrouve dans la forme physique de l'emprunt. Lorsqu'il y avait correspondance entre le genre du morphème réanalysé et le genre majoritaire en corpus, nous avons noté *O* (pour *oui*) dans la matrice; dans le cas contraire, nous avons noté *N* (*non*). Dans les cas où des morphèmes des deux genres correspondaient à une unité, nous avons noté *V* (*variable*). Les résultats des deux critères

¹¹² Donc le critère de syllabité entière phonétique n'aura logiquement que des cas analysables monosyllabiques. Tous les polysyllabiques sont considérés comme non analysables, quand ils ne peuvent être traités par les critères.

peuvent être trouvés en annexes (section 9.2) sous forme de grille de résultats, ce que nous appelons *matrice de résultats*.

5.3.1.1 Réanalyse morphologique phonétique

Sur les 100 unités, 68 sont analysables, c'est-à-dire que leur finale correspond à au moins un morphème sous une forme phonétique¹¹³. Sur les 68, le genre majoritaire de 15 unités est identique au genre des morphèmes réanalysés, comme dans le cas de la finale phonétique de l'emprunt majoritairement féminin *merguez*, [ɛz], identique à la finale féminine –èse (de *genèse*, *diurèse*, *électrophorèse*). Par exemple :

- « « Il y a encore quelques stades qui sentent le sandwich à la **merguez** (rire) ! Moi, c'est surtout l'odeur de la pelouse qui me marque, surtout quand elle vient d'être coupée, une odeur qui rappelle l'enfance. » » - Benoît Trémoulinas (Propos recueilli par Frédéric Laharie, *Sud Ouest*), 2013-01-20.

Le genre de cinq (5) unités n'est pas identique au genre associé aux morphèmes réanalysés, comme c'est le cas de l'emprunt majoritairement masculin *pipeline*, à finale phonétique [in] identique à la finale féminine –ine (*olivine*, *canine*, *bottine*, *vitrine*). Notons que *pipeline* porte parfois la MAF, bien que l'unité soit majoritairement associée avec la MAM :

- « Ou qu'elle le fasse dans la **pipeline** . » - Anonyme, sans date (corpus québécois informel)
- « Quelque 500.000 feuilles de papier. Au minimum. Telle est l'économie que le fisc pourrait réaliser cette année grâce à un projet de loi qui est entré dans le **pipeline** parlementaire la semaine dernière. » - Joan Condijs (*Le Soir*), 2013-04-26
- « À la mi-octobre, des étudiants de l'Université Laval ont lancé une pétition sur Facebook afin d'exercer des pressions auprès de l'administration pour que celle-ci refuse le passage du **pipeline** sur les terres de la ferme. » - Annie Mathieu (*Le Soleil*), 2013-11-27
- « Les indigènes Sarayaku de l'Amazonie se sont mobilisés pour sauver leur forêt afin de protéger leur milieu de vie et ont gagné le droit de gestion de leur territoire. Les Cris au Canada luttent aussi actuellement pour le respect de leurs droits et pour protéger les lacs et les rivières. La Colombie-Britannique et les États-Unis ont refusé le passage du **pipeline** alors on veut le faire passer au Québec. Y aura-t-il une mobilisation contre ce projet de pipeline qui n'apportera rien de bon pour la population du Québec comme le souhaite Québec solidaire? Une chose est certaine ce **pipeline** rapportera de gros profits aux entreprises et pour les dgts [*sic.*] "on verra"... » - Pierre Lavergne (*Le Nouvelliste*), 2013-10-07

Les finales de 48 unités correspondent à des morphèmes des deux genres et sont donc considérées à genre variable, comme c'est le cas de l'emprunt majoritairement masculin *manga*, à finale phonétique [a] identique aux morphèmes marquant le masculin –a (*quota*, *opéra*) et –at (*secrétariat*, *doctorat*) et au morphème marquant le féminin –a (*Maria*, *diva*, *marina*). Si on considère le rapport des correspondances sur la somme

¹¹³ Un exemple de cas non analysables serait l'emprunt *nunchaku* dont les finales en [aku], [ku] ou [u] ne correspondent pas à un morphème du français.

des cas analysables comme un taux de justification du critère, on calcule alors 15 sur 68. Le taux de justification du critère est donc de 22%¹¹⁴.

5.3.1.2 Réanalyse morphologique graphique

Sur les 100 unités, 45 sont analysables, c'est-à-dire que leur finale correspond à au moins un morphème sous une forme graphique. Sur celles-ci, le genre majoritaire de 18 unités est identique au genre du morphème réanalysé, comme c'est le cas de l'emprunt majoritairement masculin *karaoké*, dont la finale *-é* correspond au morphème marquant le masculin *-é* (*passé, comté*). Voici une occurrence de *karaoké* portant la MAM et pouvant être expliqué par ce critère :

- « Samedi, salle Pierre-de-Montesquiou à Condom, la Troupendélir montera sur scène. Deux spectacles sont programmés, à 15 heures, puis à 21 heures. Entre les deux, un **karaoké** est prévu avec Franz. La troupe propose deux heures de spectacle avec chant, danse et sketches. Entrée libre. ARCH. " S O " » - Anonyme (*Sud Ouest*), 2013-10-07

Le genre de deux (2) unités n'est pas identique au genre de leur finale, *pipeline* et *tajine* étant tous deux des emprunts majoritairement masculins, mais dont la finale en *-ine* correspond au morphème marquant le féminin *-ine* (*olivine, canine, bottine, vitrine*). *Tajine* portent fréquemment la marque des deux genres, comme les occurrences suivantes le démontrent.

- « D'ailleurs, mon invité ne peut résister au **tajine** du menu 30e anniversaire (remplacé depuis par le menu Dingue de steak et de moules offrant deux tables d'hôte pour 30 \$) composé de plats qui ont connu leur gloire au fil des ans. Il débute par les rillettes de canard issues de ce menu d'une délicieuse nostalgie. » - Stéphanie Bois-Houde (*Le Soleil*), 2009-11-06
- « Outre le **tajine** tunisien, qui est une sorte de tourte aux oeufs, aux légumes et à la viande, contrairement au **tajine** marocain qui est un ragoût, l'auteur évoque la gargoulette, une spécialité originale qui est l'agneau cuit longuement dans une cruche en terre que l'on brise ensuite pour le déguster. » - Anonyme (*Sud Ouest*), 2010-01-21
- « J'ai reçu une **tajine** pour Noel et je recherche des recettes simples que je pourrais y faire. » - Anonyme, sans date (corpus québécois informel).
- « alala autant de kilos perdu aaaa quel courage moi je sais que j'ai du mal a perdre ses kilos et tu perd vite toi non ? moi non je perd soit 500g en une semaine a un kilos et des fois l'autre semaine je perd rien c'est pénible sinon la **tajine** c'est un plat avec légume viande on peu avoir tajine mouton tajine poulet toute sorte de viande dans un plat en pierre qui et mijoté avec une sauce jaune et le poulet c'est un poulet roti avec haricot vert et flageoler en général que je prend et couscous en générale je prend pas de semoule car le soir jamais de féculent même pour mon repas plaisir car jai trop peur de grossir je fait toujours ça .. » - mysterieusefr59 (*Doctissimo*), 2013-03-29

¹¹⁴ Dans la mesure où l'on exclut les cas à genre variable, le taux de justification est de 15/20 (75%), mais comme nous le mentionnerons en discussion (section 5.3.1.3), nous considérons toujours les cas à genre variable dans nos calculs.

Les finales de 25 unités correspondent à des morphèmes de deux genres et sont donc considérées à genre variable. Considérant le nombre de correspondances comme la mesure d'un taux de justification du critère, il y a 18/45 correspondances, ce qui donne un taux de 40%¹¹⁵.

5.3.1.3 Discussion

Un problème de la réanalyse morphologique est que la forme de nombreux morphèmes, particulièrement des morphèmes monophonémiques ([a], [i], [e]) ou parfois monographémiques (<a>), est commune à celles d'unités lexicales des deux genres se terminant par ces formes. Par exemple, la finale phonétique [e] peut correspondre graphiquement à plusieurs formes : -é, -er et -ée. Les deux premières finales sont typiquement associées au masculin (*passé, comté; le parler, le toucher, le manger*) alors que la dernière l'est avec le féminin (*chevauchée, arrivée, entrée, marée, soirée*). De même, le phonème [i] peut correspondre aux finales graphiques -ie et -is. La première finale est typiquement associée au féminin (*partie, acrobatie, folie*) alors que la seconde l'est avec le masculin (*clapotis, cliquetis, gribouillis*). Cette ambiguïté affecte un trop grand nombre d'unités pour permettre une conclusion ferme et décidée. Dans le cas de la réanalyse phonétique, il s'agit de 48 unités sur un total de 68 unités analysables (dont la finale correspond au moins à un faux morphème) qui pourrait confirmer ou infirmer l'efficacité de ce critère, ce qui correspond à 70,5% des cas analysables. Pour ce qui est du cas de la réanalyse graphique, il s'agit de 25 unités sur un total de 45 analysables, ce qui correspond à 55,5% des cas analysables. Dans les deux cas, la majorité des unités analysables se voit qualifiée d'ambiguë, parce qu'aucun genre n'est décisif, ce qui mine la crédibilité d'un critère comme la réanalyse morphologique. Si l'on devait exclure les cas à genre variable des calculs de justification, le taux serait très élevé, 75% dans le cas de la réanalyse morphologique de la finale phonétique et 90% dans le cas de la réanalyse de la finale graphique. Il est nécessaire de considérer les cas à genre variable dans le calcul de justification, car cela permet de prendre en considération tous les cas analysables et de dresser un portrait représentatif de la situation réelle. Les deux critères de réanalyse morphologique sont exceptionnels en ce qu'ils sont les seuls à avoir un si grand nombre de cas à genre variable. Nous avons considéré tous les cas analysables pour chaque critère dans le calcul des taux de justification.

Les finales phonétiques à genre variable étant similaires à un morphème sont [i], [iʃ], [e], [a], [ik] et [ɛl]. Les finales graphiques à genre variable étant similaires à un morphème sont <a> et <one>. Les finales phonétiques à genre non variable étant similaires à un morphème sont [jɔ̃], [ɔ̃m], [in]¹¹⁶, [it], [id], [an], [o], [ad], [ɛz], [ɔ̃]. Les finales graphiques à genre non variable étant similaires à un morphème sont <ion>, <um>, <ine>, <o>, <é>, <el>, <elle> et <on>.

¹¹⁵ Dans la mesure où l'on exclut les cas à genre variable, le taux de justification est de 18/20 (90%).

¹¹⁶ Le cas de [in] est particulier. L'unité *pipeline* ne varie pas en genre, mais *tajine* oui. Nous avons classé [in] comme à genre non variable, car cette finale est de manière générale associée au féminin (voir *supra* 4.3.1.1).

Les cas variables comportent 50% de finales monophonémiques (qui comptent pour la grande majorité des cas) et 50% de finales polyphonémiques. À l'inverse, les cas non variables ne comptent que 20% de finales monophonémiques (11 occurrences sur 20 non variables) pour 80% de finales polyphonémiques. Les cas non variables comptent 25% de finales monogaphémiques qui représentent 13 occurrences sur 20.

5.3.2 Syllabicit  finale et enti re

5.3.2.1 Syllabicit  finale graphique

Afin de tester ce crit re, le genre majoritaire des unit s devait  tre identique au genre suppos  par le crit re, c'est- -dire que les mots se terminant par une consonne graphique devaient  tre masculins, et ceux par une voyelle devaient  tre f minins. Sur les 100 unit s, 98 sont analysables. Les graphies de *bug-bogue* et *van-vanne* se terminent   la fois par une consonne et une voyelle, ce qui les rend inanalysables. Sur les 98 cas analysables, le genre majoritaire de 41 unit s est identique au genre suppos  du crit re, comme c'est le cas de l'emprunt majoritairement f minin *baraka* se terminant par la voyelle graphique <a>.   noter que *baraka* porte g n ralement la MAF, mais au Qu bec, plusieurs occurrences font r f rence au nom d'un bar ou d'un bistro, comme le montrent les exemples suivants :

- « Steve Hill, le gars des Porn Flakes avec une grosse motte de cheveux pis une grosse gang de musiciens colons qui ont vraiment envahi la terrasse du **Baraka** hier soir... :gluk: :gluk: :gluk: :gluk: Dans le genre show-off gossants, ils  taient durs   battre! :sarcastic: » - Anonyme, sans date (corpus qu b cois informel)
- « Les 3 filles de la web t l  Week-end machin...sur la terrasse du **Baraka**. » - Anonyme, sans date (corpus qu b cois informel)
- « « C'est trop tard pour aller   L'Escogriffe, le show est termin . Le **Baraka** ? » Why not, coconut ? Pas de chance, le **Baraka**  tait mort, on se serait cru chez Alfred Dallaire. Apr s quelques bises et des « Aye, Stef ! Stefie ! Comment tu vas ? », de retour sur le trottoir, direction Saint-Denis. » - Jean-Yves Girard (*Le Devoir*), 2003-09-26
- « La plus r cente transaction s'est conclue mercredi. Elle met en vedette la bande commerciale o  sont situ s entre autres le restaurant le **Baraka** et le magasin de peinture B tonel. » - Martin Lafreni re (*Le Nouvelliste*), 2012-09-08

  l'inverse d'unit s comme *baraka*, le genre de 57 unit s n'est pas identique, comme c'est le cas de l'emprunt majoritairement masculin *ch che* se terminant par la voyelle graphique <e>. Le taux de justification du crit re est de 41,8% (41/98).

Nous avons d cid    la section 2.3.1 que nous devons nous baser sur un crit re graphique, plut t que phon tique, parce que nous  tudions des unit s sous une forme graphique,   travers un corpus  crit et non oral. Nous avons cependant calcul  le taux de justification du crit re sous sa forme phon tique afin de v rifier si notre choix  tait valide. Le taux de justification est de 45% dans ce cas. La diff rence est mineure, mais doit prendre en compte les variantes phon tiques des emprunts, notamment pour les emprunts   l'italien, comme *farfalle* [faʁfale] ou [faʁfal], *sabayon* [sabajɔn] ou [sabajɔ].

Nous avons également décidé de regarder rapidement si la « règle » stipulant que les mots se terminant par un <e> graphique soient féminins pouvait être valide dans le cas des emprunts, puisqu'elle ne l'est pas totalement avec les unités du lexique indigène. Sur nos cent unités, dix-sept (17) d'entre elles se terminent par <e> (excluant la finale <é> de *karaoké*, *saké* et *taboulé*, les trois portant majoritairement la MAM). Huit (8) unités portent majoritairement la MAM, quatre (4) celle du féminin et cinq (5) sont considérées comme à genre variable. De nos 17 unités, neuf (9) proviennent de l'anglais, aucune du japonais, trois (3) de l'arabe et cinq (5). Considérant que, pour être valide, critère doit avoir 11 unités portant la MAF pour atteindre un taux de justification minimal de 64%, nous avons calculé le taux et avons obtenu un total de 23,5% de féminin pour notre petit échantillon. Nous n'avons pas retenu ce critère dans notre méthodologie, mais, une comparaison avec un échantillon différent pourrait être intéressante à effectuer, notamment avec plus d'emprunts à l'anglais.

5.3.2.2 Syllabicit   enti  re

Afin de tester ce crit  re, le genre majoritaire des unit  s en corpus europ  ens devait   tre identique au genre suppos   par le crit  re, c'est-  -dire que, les unit  s monosyllabiques devraient porter la marque du masculin. Parmi les 100 unit  s, seules 22 sont monosyllabiques, donc analysables. Le genre de 15 d'entre elles est masculin, comme c'est le cas de *slam*. Le genre de 7 unit  s est f  minin (comme *rave*, *stout* ou *dope*) ou variable (comme *drum* et *brick*). Le taux de justification du crit  re est de 68,1% (15/22). Voici quelques occurrences qui attestent le genre de chaque unit   mentionn  e ci-haut (*slam*, *rave*, *stout*, *dope*, *drum*, *brick*) :

- « Et c'est bien dommage car, toujours d'apr  s Tarik, la qualit   de vie du quartier s'est am  lior  e depuis quelques ann  es. Notamment gr  ce    des associations, comme Qu'on pose ou encore 2 H Prod, qui initient les jeunes au rap ou au **slam**. » - Daniel Charpentier (*Sud Ouest*), 2012-11-19
- « Rez est un merveilleux shoot sur rails, con  u par Tetsuya Mizuguchi, qui raconte en avoir eu l'id  e apr  s avoir particip      une **rave** m  morable et conjugu   cette exp  rience avec les principes de synesth  sie   labor  s par Kandinsky (d'o   son nom de code originel: K-project). Esth  tiquement, sensoriellement et po  tiquement, Rez HD constitue l'une des plus belles exp  riences disponibles pour les joueurs connect  s. » - Olivier S  guret (*Lib  ration*), 2008-02-09
- « Le ph  nom  ne d  fie le bon sens. Plus l  g  res que l'eau et l'alcool, les bulles de gaz doivent remonter    la surface. En fait, d  s que la **stout** est servie, il se passe toujours un moment avant que le breuvage ne trouve son   quilibre et que les nuages de bulles disparaissent. Durant ce laps de temps, une colonne de bulles de gaz ascendantes se forme au centre du verre. » - Anton Vos (*Le Temps*), 2004-04-06
- « En juillet, le Tour fait salle comble. C'est l'occasion de faire le tour, en cinq   tapes, des id  es re  ues sur le v  lo. Elles sont innombrables et touchent forc  ment la **dope** qui   clabousse tout le peloton. Cette ann  e, en l'absence de Contador puni et d'Andy Schleck bless  , deux favoris : le Britannique Bradley Wiggins (Sky) et l'Australien tenant du titre Cadel Evans (BMC). » - Jean-Louis Le Touzet (*Lib  ration*), 2012-06-30
- « " I believe I can survive on my own ", y chante-t-elle : " Je pense que c'est la chanson sur laquelle je vais rebondir, dit-elle. Depuis, je retourne puiser dans une base plus black, plus soul. J'ai envie de revenir    la percussion, au **drum**, repartir du rythme, de la basse, puis du vocal. Partir de l   plut  t que de faire gling gling avec ma guitare et de petites m  lodies cutes. Je veux partir du groove. " Doba est bel et bien de retour avec toute sa fougue et son envie de la sc  ne. » - Yves Bernard (*Le Devoir*), 2013-05-23

- « Jouer What's eating you de Airbourne à la **drum** je connais quasi le morceau par coeur niveau rythme, en une semaine c'est fait sans fausses notes, mais je n'ai pas de batterie » Profil supprimé (*Doctissimo*), 2010-11-10
- « Sami se fait d'ailleurs un devoir de recréer l'ambiance toute tunisienne de sa jeunesse, notamment en cuisinant les recettes de sa mère. La salade méchouia, la **brick** (fine enveloppe de pâte souvent fourrée d'un oeuf et d'une viande ou d'un poisson), les baklavas et jusqu'aux émissions sur la chaîne arabe ART (arab radio television), auxquelles il s'abonne pour le mois du ramadan. » - Lisa-Marie Gervais (*Le Devoir*), 2011-08-09
- « Avec, çà et là, effeuillé de pomme ou de prune d'ente. Les chefs avertis ne manquent pas de trouver à ce raffiné dessert un cousinage avec la pastilla marocaine ou le **brick** tunisien. Pour la conduire à l'autel : un verre de pacherec frais. » Anonyme (*Sud Ouest*), 2004-08-07

Au total, sur les 78 unités polysyllabiques, on compte 29 unités bisyllabiques et 49 unités à trois syllabes ou plus. Sur les 22 cas analysables à partir de ce critère, on compte 19 emprunts à l'anglais et 3 à l'arabe. Les emprunts à l'italien et au japonais que nous avons sélectionnés sont tous au minimum bisyllabiques.

5.3.2.3 Discussion

Le taux de justification du critère de syllabité entière est intéressant, mais il ne concerne qu'une petite portion des emprunts, soit les emprunts monosyllabiques, dont la majorité vient de l'anglais (19). Les trois (3) autres emprunts monosyllabiques proviennent de l'arabe, une autre langue qui permet plus facilement une structure monosyllabique relativement fréquente. De plus, dans le cas du critère de syllabité finale graphique, l'anglais et l'arabe permettent des finales consonantiques de manière régulière, possibilité que ne permet pas le japonais, outre le <n>. Il en est de même pour l'italien qui, selon notre échantillon, ne permet presque pas de finale consonantique, sauf le <n> d'une unité, *sabayon*. Cette distribution syllabique de notre échantillon influence grandement les résultats de ces critères. De tous les cas analysables, 19 emprunts à l'anglais sont monosyllabiques et 16 se terminent par une consonne graphique, ce qui relève d'une distribution équilibrée. À l'inverse, les emprunts au japonais et à l'italien sont uniquement polysyllabiques et se terminent tous, à l'exception de *seitan* et de *sabayon*, par une voyelle. Ce très grand écart de construction syllabique des mots entre les différentes langues donneuses ne rend les critères syllabiques potentiellement opératoires qu'à condition d'observer un échantillon composé d'emprunts de différentes langues, de toutes familles, de tout système graphique ou phonologique. Pour les emprunts au japonais et à l'italien, la grande majorité d'emprunts à finale vocalique supposerait toujours un genre féminin, ce qui est loin d'être le cas.

Dans le cas du critère de la syllabité entière, nous notons que nos résultats ne concordent pas avec ceux de Surridge (1982:37) pour qui les unités (empruntées à l'anglais, rappelons-le) portaient la marque du masculin dans 94% des cas. Son échantillon, composé de 275 unités monosyllabiques tirées du « *Trésor de la langue française*, [du] *Dictionnaire alphabétique et analogique* de Robert et [du] *Dictionnaire Lexis* de Larousse » (p. 29), est plus de douze fois plus volumineux que le nôtre et comportent des unités qui remonteraient jusqu'au 17^e siècle (*ibid.*). Nous avons fait le choix de nous limiter aux emprunts d'après 1800. Cependant, comme le

démontre Surridge (1984), le pourcentage d'emprunts à l'anglais portant la marque du féminin est passé de 25% au 17^e siècle à 12% au 20^e siècle. Puisque nos emprunts devraient afficher un plus grand nombre de MAM (car ils sont plus récents), et montrent pourtant un plus faible taux de masculin (68,1%), nous sommes en mesure d'affirmer que, pour notre échantillonnage restreint de 22 cas, le critère de syllabité entière est potentiellement opératoire dans le cas des corpus européens, mais pas autant que le prétend Surridge.

Ajoutons en dernier que, dans le cas des monosyllabes analysés dans les corpus québécois, Surridge (1982:39) décrivait le taux de masculin comme étant entre 60% et 50% pour ses deux corpus québécois. Dans notre cas, comme on peut le voir en annexes (section 9.1.1 et 9.1.3), 14 des 19 emprunts à l'anglais portent la MAM (*gig* et *hype* portent fréquemment les deux genres; *dope*, *van* et *stout* portent la MAF), et les trois emprunts à l'arabe portant la MAM, ce qui nous donne un taux de masculin de 77,2% (17/22), un taux supérieur à ceux de Surridge. Nos résultats contredisent ceux de Surridge pour qui les emprunts monosyllabiques à l'anglais portent nettement plus souvent la MAM en France qu'au Québec. Nous notons un plus grand pourcentage d'emprunts monosyllabiques à l'anglais au Québec qu'en Europe.

Pour conclure, bien que le taux de justification du critère de syllabité entière soit élevé, la structure syllabique de la langue d'origine des emprunts a beaucoup à jouer sur le nombre de cas monosyllabiques.

5.3.3 Finale graphique et similitude graphique

5.3.3.1 Finale graphique

Afin de tester ce critère, le genre majoritaire des unités devait être identique au genre associé à la finale graphique de l'unité dans Lyster (2006). Sur les 100 unités, les finales de 80 ont trouvé correspondance, et sont donc analysables. Sur les 80 cas, le genre majoritaire de 49 unités est identique au genre des finales graphiques relevées par Lyster, comme c'est le cas de l'emprunt majoritairement masculin *scoop* se terminant par la finale *-op*, noté par Lyster comme appartenant à des mots masculins. Illustrons le cas de *scoop* avec quelques exemples :

- « Un autre **scoop** qui fesse aussi fort que celui de Lindsay » - Anonyme, sans date (corpus québécois informel)
- « oui le vote gratuit du ballottage est peut être (surement) manipulé par la prod pour moi spa un gros **scoop** là... et de faire des suppositions pis de crier scoop cest presque pire qu'une rumeur car yen a qui peuvent prendre ça pour du cash! » - Anonyme, sans date (corpus québécois informel)
- « Londres Correspondant - Savoir ce qui se passera demain en Bourse vaut de l'or. Le détenteur d'un **scoop** financier peut facilement le monnayer. En témoigne l'inculpation, mardi 2 octobre, de quatre personnes dans le cadre de la plus complexe opération jamais menée dans la City contre des délits d'initié. » - Marc Roche (*Le Monde*), 2012-10-04
- « Nicolas avait un **scoop** pour les Rosemontois: "Nous allons bientôt ouvrir, tout près de l'ancienne (et légendaire) taverne Donald Duck, une taverne de sports qui s'appellera le Grand Lionel, en hommage au Grand Antonio et à Lionel Duval en même temps." » - Ronald King (*La Presse*), 2013-08-26

À l'inverse d'unités comme *scoop*, le genre de 31 unités est différent du genre des finales relevées par Lyster, comme c'est le cas de l'emprunt majoritairement masculin *pipeline* se terminant par la finale *-ine*, notée par Lyster comme terminant généralement des mots féminins. Le taux de justification du critère est de 61% (49/80).

5.3.3.2 Similitude graphique

Afin de tester ce critère, le genre majoritaire des unités devait être identique au genre des unités sélectionnées comme ayant un bon degré de similitude graphique. Sur les 100 unités, 26 ont trouvé correspondance, et sont donc analysables. Sur les 26 cas, le genre majoritaire de 10 était identique à celui de correspondants graphiques, comme c'est le cas de l'emprunt majoritairement masculin *dojo* ressemblant au mot masculin *mojo*¹¹⁷. Le genre de 26 unités était différent, comme c'est le cas de l'emprunt majoritairement masculin *manga* ressemblant au mot féminin *mangue*. Le taux de justification du critère est de 38,4% (10/26). Illustrons les cas de *dojo* et *manga* :

- « Vendredi dernier, les élèves de CP de l'école Marie-Curie et des grandes sections des écoles maternelles Jacques-Prévert et Jean-Moulin se sont tous retrouvés au **dojo** pour une rencontre sportive " Jeux d'opposition ". Les dix équipes constituées d'élèves des trois écoles, ont pu s'affronter en jouant successivement à quatre jeux différents. » - Anonyme (*Sud Ouest*), 2013-12-18
- « L'athlète le plus en forme dans les mois précédant ces Jeux, Antoine Valois-Fortier, un ancien du **Dojo** de Beauport, n'a pas été gâté. Son premier adversaire, mardi, est l'ancien champion olympique à 73 kilos, qui fait le saut chez les 81 kilos cette année. » - Gabriel Béland (*Le Soleil*), 2012-07-28
- « Je préfère la troisième version. Il me semble intéressant de suivre le personnage dans l'élaboration de son plan et de ses actions préméditées. ça me fais un peu penser au **manga** DN Angel en fait » - Suiginto-Rozen (*Doctissimo*), 2011-10-26
- « Connais-tu le **manga** actuel, dis-moi? Arfff! Avec un message pareil, bien sûr que non! Tu t'en es même arrêtée à DBZ! Et si tu me connaissais, tu saurais que je déteste ce **manga** justement! On ne t'en veut pas la parisienne de mon coeur! Angelique la vraie a écrit :faut pas avoir honte à dire que tu lis des mangas, moi je trouve bien l'inspiration dans les catalogues pages sous vêtements Imaginer le résultat me donne déjà des cauchemars... » - dominator63 (*Doctissimo*), 2009-11-29

5.3.3.3 Discussion

L'aspect graphique de l'attribution du genre est un volet peu étudié. Les emprunts ont la particularité de conserver le plus souvent leur graphie originale (lorsqu'elle est du même alphabet, sinon elle suit un code, une transcription bien établie). Ils ont également la particularité d'être généralement plus fidèles à la correspondance d'un graphème à un phonème (ex : *niqab*, *trip*, *karaoké* et *calzone* comportent autant de graphèmes que de phonèmes). Évidemment, ce n'est pas le cas de tous les emprunts, comme les emprunts italiens qui conservent les consonnes doubles (*cappuccino*, *cannelloni*, *bruschetta*). Ces deux particularités rendent particulièrement difficile la correspondance à des formes graphiques similaires, spécialement dans le

¹¹⁷ Également un emprunt, hors de notre échantillon.

cas des unités non monosyllabiques (78/100 des unités à l'étude). Sur les 26 unités qui ont trouvé au moins un correspondant graphique, onze (11) sont monosyllabiques, dix (10) sont bisyllabiques et cinq (5) sont trisyllabiques ou davantage. Ce constat nous mène à conclure que le critère de similitude graphique est davantage opératoire avec les emprunts mono- et bisyllabiques qu'avec les emprunts à trois syllabes ou plus. Cela est validé par le fait que la moitié des unités à l'étude (49/100) comportent trois syllabes ou plus et que seules cinq ont au moins un correspondant graphique¹¹⁸, alors que la moitié des unités monosyllabiques (11/22) et le tiers des bisyllabiques (10/29) trouvent correspondance.

En ce qui a trait au critère de la finale graphique, le nombre de cas analysables (80/100) et le taux de justification de 61% font de ce critère un cas très intéressant. Son taux est suffisamment élevé pour que nous puissions, selon nos standards, classer ce critère comme étant potentiellement opératoire (*PO*) et celui-ci s'applique à un grand nombre d'unités provenant de toutes les langues donneuses.

5.3.4 AP par phonème final et pénultième

Afin que nous puissions considérer les critères phonétiques comme *PO*, il fallait que le genre majoritaire des unités soit identique aux genres associés aux phonèmes du français par Tucker *et al.* (1977).

5.3.4.1 AP par phonème final

Sur les 100 unités, 86 se terminent par un phonème relevé par Tucker *et al.* comme étant associé soit au genre masculin, soit au genre féminin. Rappelons que les phonèmes /e/, /p/ et /t/ n'ont pas été associés à un genre selon cette étude de 1977; 14 unités (sur les 100 au total) se terminent donc par l'un de ces phonèmes, considérés « ambigus » et ne sont pas considérées comme analysables. Sur les 86 unités analysables, nous observons 35 correspondances de genre (comme l'emprunt masculin *kimono* correspond à la finale [o], généralement associé au masculin) et 49 non-correspondances (comme l'emprunt féminin *fatwa* ne correspond pas à la finale [a], généralement associée au masculin). Illustrons les cas de *kimono* et de *fatwa* :

- « Et puis, arrivent les modèles en soie, tout aussi réjouissants, comme en attestent ce **kimono** ceinturé de cuir, ou encore cette robe avec ce col amovible (de cuir encore) porté avec un pantalon (de soie). Parlons-en justement, du cuir, ces ceinturons-corsets, ces bustiers, tous sont les oeuvres de Laurence St Pierre, jeune designer qui maîtrise le cuir embossé, technique qui permet de créer du relief dans les peaux en les plongeant dans l'eau. Le mélange des genres est parfait. » - Elsa Vecchi (*La Presse*), 2011-09-09
- « S'il n'y a pas de compétitions, les différents clubs du Born et du bassin d'Arcachon organisent régulièrement des rencontres pour favoriser l'émulation et permettre aux élèves de progresser avec de nouveaux partenaires. En complément du **kimono**, certains pratiquants portent le hakama, symbole de la tenue traditionnelle des samourais : chaque pli de cette jupe en tissu noir est le rappel d'une valeur à respecter. » - Axel Frank (*Sud Ouest*), 2013-07-25
- « En 1994, ses parents, engagés, militants et aussi doués qu'elle pour la controverse, font l'objet d'une condamnation à mort décrétée par les intégristes algériens. Cette **fatwa** marquée au fer rouge sera

¹¹⁸ Ce qui correspond à 10,2% (5/49).

paradoxalement une bénédiction pour Djemila. Pas tout de suite, évidemment. La famille est obligée de quitter l'Algérie en catastrophe et de s'établir en banlieue de Paris. Mais en 1997, trois mois seulement après son arrivée au Canada, cette **fatwa** lui permettra d'obtenir un statut de réfugiée politique et de tout recommencer. "Au début, dit-elle, je voulais tout savoir sur le Québec. Je posais mille et une questions à mes amis. J'étais comme un Dracula qui veut tout absorber. Je me baladais à Montréal, à la fois assommée par le fait que personne dans la rue ne me connaissait alors qu'à Oran, c'était le contraire, mais en même temps, cet anonymat était très exaltant." » - Djemila Benhabib (*La Presse*), 2012-11-17

- « Le romancier britannique d'origine indienne Salman Rushdie a annoncé mardi qu'il comptait écrire le récit de ses 10 ans passés dans la clandestinité après la **fatwa** de l'ayatollah Khomeini appelant les musulmans à l'assassiner à cause des Versets sataniques. "C'est mon histoire et à un certain moment, il faut qu'elle soit racontée. Ce moment approche, je crois", a-t-il déclaré lors de la présentation d'une exposition consacrée à son oeuvre, qui ouvrira ses portes au public vendredi à l'Université Emory d'Atlanta, en Géorgie. Publiés à la fin de 1988, Les versets sataniques ont séduit la critique mais déclenché de vives protestations, des musulmans en colère brûlant des exemplaires en public et manifestant un peu partout dans le monde. Le 14 février 1989, l'ayatollah Khomeini a lancé la **fatwa** appelant à assassiner le romancier. Salman Rushdie, qui vivait à Londres et était déjà un écrivain reconnu, a alors été contraint d'entrer dans la clandestinité pendant une dizaine d'années, sous la protection du Royaume-Uni. Les Versets sont quant à eux devenus un succès de librairie mondial. En 1998, le gouvernement iranien a déclaré qu'il ne soutenait pas la **fatwa**, mais qu'il ne pouvait pas pour autant la lever. L'ayatollah Khomeini est mort en juin 1989 et Rushdie a depuis longtemps cessé de s'inquiéter constamment pour sa sécurité, considérant que la **fatwa** est devenue davantage "une figure rhétorique qu'une vraie menace". » - Anonyme (*La Presse*), 2010-02-25

Deux cas, ceux de *seitan* (*fāl* ou *lnl*) et d'*antipasto-antipasti*, correspondent aux deux genres. Nous avons considéré plusieurs autres emprunts selon deux prononciations, mais le genre de ceux-ci était soit identique (*tofu* prononcé [y] ou [u], masculin et masculin), soit l'une des deux finales était [e], [p] et [t] (*minestrone* [n] et [e], féminin et « ambigu »). Le taux de justification du critère est de 41,2% (35/86). Illustrons quelques cas intéressants de *seitan*¹¹⁹ et d'*antipasto-antipasti* :

- « Pour moi ce matin : thé au lait écrémé en poudre+miel, pain complet, confiture 100% fruits, carré frais 0% Clémentine en collation A midi j'ai fait du **seitan** aux échalotes et champignons, des haricots verts, une petite salade de roquette, 1 wasa fibres (IG 35) avec du petit billy, 1 carré de chocolat 70% avec mon café Et pour ce soir je sais pas non plus » - bikiny (*Doctissimo*), 2007-10-13
- C'est à base de gluten (faut pas y être allergique) Ca fait partie aussi des aliments prônés dans le régime Okinawa. J'avais posté la recette sur cuisine végétarienne là voilà Pour le **seitan** 250 gr de gluten 50 gr de farine complète 15 gr de levure alimentaire Épices et/ou fines herbes diverses comme par exemple 1/2 cs de gingembre en poudre, 1/2 cs d'ail semoule, 1/2 cs de thym. un peu de poivre 500 ml d'eau tiède 30 ml d'huile 30 ml de tamari Il faudra aussi un bon bouillon bien parfumé 2 litres d'eau 60 ml de tamari 60 ml d'huile 1 cs de miso kombu De l'ail semoule 1 cube de bouillon de légumes Thym et laurier 1/2 cc de sel poivre Dans un saladier mélanger la farine, la levure et les herbes&épices Ajouter l'eau, le tamari et l'huile et bien mélanger le tout. Verser d'un coup la farine de gluten et pétrir 5 mn avec les mains. Laisser reposer le temps de préparer le bouillon. Lorsque le bouillon est prêt, diviser la **seitan** en 4 ou 5 morceaux et faire cuire 1h à feu tout doux dans le bouillon et en retournant les morceaux de temps à autre. Laisser refroidir dans le bouillon puis conserver au frais. Se conserve 3 jours au frais. J'ai obtenu plus d'un kilo avec cette recette. J'en ai donc congelé, coupé en tranches. Et décongelé au four à 120° 30 minutes selon les besoins et c'est parfait ! » - bikiny (*Doctissimo*), 2007-03-05

¹¹⁹ Notons que les deux premiers commentaires sont produits par le même utilisateur et portent les deux genres.

- « Excellente idée que de faire un **antipasto** froid pour l'entrée. Je choisis souvent la formule cocktail quand je reçois, plutôt que de faire une entrée assis à table. Je trouve ça plus convivial, moins stressant et (surtout) pas de Si tu vas du côté italien, je te conseille des cannelonnis (j'avoue que cette photo n'est pas à son meilleur mais quand même...)et regarde dans ma catégorie ENTRÉES...Tu y trouveras une salade de tomates et boconccini, un **Antipasto** ou bien encore des bouchées de prosciutto au deux melons. » - Anonyme, sans date (corpus québécois informel)
- « De bonnes adresses pour ceux qui ont besoin de lire quelques lignes avant de fermer l'oeil... et qui aiment flâner dans les rayonnages avant de manger un **antipasti** sur le pouce, feuilleter des livres de cuisine pour se donner l'eau à la bouche ou finalement louer un DVD juste à côté ! » -

5.3.4.2 AP par phonème pénultième

Sur les 100 unités, 83 se terminent par un autre phonème que [e], [p] et [t], c'est-à-dire un phonème auquel l'étude de Tucker *et al.* (1977) ont identifié un genre. Sur les 83 cas analysables, le genre des finales de 41 unités est identique au genre en corpus, comme c'est le cas de l'emprunt majoritairement masculin *saké* au phonème pénultième [k] généralement associé au masculin. Le genre majoritaire est différent dans le cas de 38 unités, comme pour l'emprunt majoritairement féminin *baraka* au phonème pénultième [k]. Quatre (4) unités, acceptant deux prononciations (*hype*, *pipeline*, *soja/soya*, *minestrone*) sont associées à deux phonèmes à genres différents. Le taux de justification du critère est de 49,4% (41/83). Illustrons de quelques exemples les cas de *hype*, de *soja/soya* et *minestrone* :

- « Et puis dans tout cela, toutes les parties ont comme but de créer le plus gros "**hype**" autour de l'émission, alors oui, le fait que Kim soit la fille d'une personne connue de tout le Québec fait d'elle un excellent point de départ pour une série d'articles... qu'elle fut en danger ou non. » - Anonyme, sans date (corpus québécois informel)
- « Quant à mon produit... Mes idées sont autrement différentes que ce que le marché demande. Les acheteurs des magasins sont forcés de choisir la sécurité, plutôt des lignes d'accessoires griffées par des créateurs ou des maisons de mode. Avec timing de livraison respecté et vente assurée. Les boutiques préfèrent des accessoires moins coûteux, des gimmicks ou tout ce qui ressemble à la **hype** du moment. » - Anonyme (*Le Soir*), 2015-05-05.
- « L'assiette se doit d'être si possible parfumée et inédite, voire d'étonner par un assaisonnement venu de loin. C'est celui-ci qu'une équipe de Côte que coûte est allée traquer à Tokyo, au sein même des ateliers qui fournissent le meilleur wasabi. Il a tendance à prendre petit à petit le pas sur la traditionnelle sauce au **soja**. En tout cas, il reste l'apanage des cuisiniers japonais et on ne le voit guère envahir les préparations occidentales. Notre gastronomie retient de fait peu de poisson cru dans ses préparations. » - Bernard Meeus (*Le Soir*), 2007-03-28
- « Autant d'un côté (soya) que de l'autre (lait vache). C'est juste que nous avons tous été élevé au lait de vache dans ma famille et nous sommes tous en bonne santé. Le lait de soya n'est pas assez vieux pour en voir les effets!!! Et ce qui me dérange c'est le fait de mettre le **soya** sur la première marche et dire que c'est un "wondrefood" et quand on cherche on réalise que c'est peut-être pas ça!!! J'aurais bien aimé qu'on me parle des différente facette du **soya** au lieu de me dire : "Il n'y a pas de problème, donne lui du Isomil et tout ira bien"! » - Anonyme, sans date (corpus québécois informel).
- « Oui Edmonde ce matin il faisait trop froid avec ce vent ! Y'avait de sacrées rafales mine de rien & bah du coup après dur de me réchauffer. Là ça va mieux, j'ai mangé un bon **minestrone** bien chaud ça a fait son

effet !! Demain ça va peser sur la balance avec tout ce que je me suis enfilée !! » Cataluna (*Doctissimo*), 2010-10-20

- « Sinon, comme Wajomour le mentionne, dans une soupe c'est super. Particulièrement dans une bonne **minestrone**. Dans les minestrone authentique, on écrase une partie des légumineuses et on l'ajoute au bouillon pour en faire une soupe consistante et d'une belle texture. La **minestrone**, c'est tellement bon et réconfortant. C'est d'ailleurs ma soupe favorite. » - Anonyme, sans date (corpus québécois informel)

5.3.4.3 Discussion

Les critères phonétiques, basés sur les résultats de Tucker *et al.* (1977), sont très présents dans les études sur l'attribution du genre aux emprunts. Or, force est de constater qu'ils ne sont pas particulièrement opératoires dans le cas des unités que nous avons sélectionnées. Un taux de 40% démontre que le critère est faiblement opératoire (FO), et un taux de près de 50% pourrait ne relever que du hasard. Cette différence entre nos résultats et la littérature pourrait être expliquée par le fait que la majorité des études sur le sujet portent sur des emprunts à l'anglais et que notre échantillon est composé d'emprunts à plusieurs langues. Un bref regard à nos résultats nous permet d'avancer que les emprunts au japonais (et dans une moindre mesure à l'anglais et à l'arabe) répondent mieux au critère de la finale phonétique que les emprunts à l'italien. Les emprunts au japonais sont presque exclusivement masculins. De même, tel que vu dans le chapitre méthodologique, les résultats de Tucker *et al.* (1977) favorisent le masculin (en l'occurrence 17 phonèmes, 10 pour le féminin, 3 « ambigus ») et les voyelles sont presque uniquement masculines (sauf /i/ qui est associé au féminin et /e/ qui est « ambigu »). Puisque les emprunts au japonais se terminent presque tous par une voyelle et sont de genre masculin, nous pouvons déduire une certaine interaction entre ce critère et les emprunts à cette langue.

Un autre point à prendre en considération est la finale en /a/. Les résultats de Tucker et de ses collaborateurs montrent que le phonème /a/ est associé au masculin dans le lexique indigène du français. Cependant, nous savons que l'arabe et l'italien possèdent chacun un suffixe au genre féminin avec la finale /a/. Nos observations préliminaires au sujet de la conservation du genre indiquent que le genre est, de manière générale, conservé. Prenant en considération qu'en italien et en arabe, la finale –a est caractéristique du féminin, on peut facilement expliquer que 15 des 18 cas en –a soit féminins¹²⁰. Si on ajustait le critère de l'AP par phonème final en tenant en compte de la langue donneuse et de la finale en -a, on pourrait augmenter le taux de succès du critère à 58,1% (50/86). Nous avons cependant tenu à considérer le critère tel quel pour nos conclusions.

En ce qui a trait à l'AP par phonème pénultième, bien que son taux de justification soit plus élevé que le critère d'AP par phonème final, il reste trop faible pour être considéré comme réellement efficace. On peut imaginer

¹²⁰ Les exceptions : *barda* et *sépie* sont majoritairement masculins, et *kefta* est majoritairement variable.

une influence combinée des deux derniers phonèmes, mais pour être encore plus strict en ce qui a trait aux phonèmes finaux, nous avons choisi d'analyser la possibilité de la rime riche.

5.3.5 AP par rime riche

Afin de tester ce critère, le genre majoritaire des unités sélectionnées devait être identique au genre d'un ou plusieurs mots ayant au moins 3 phonèmes finaux identiques. Sur les 100 unités, 70 riment avec au moins un mot. Sur ces 70, le genre majoritaire de 30 unités est identique au genre des rimes, comme c'est le cas de *mug*, un emprunt majoritairement masculin rimant avec *smog*. Le genre majoritaire de 38 unités n'est pas identique à celui des rimes, comme l'emprunt majoritairement masculin *shit* rimant avec deux mots féminins *bronchite* et *bullshit*. Deux unités (*minestrone* et *seitan*) ont autant de rimes masculines que féminines et cela est dû aux variantes phonétiques possibles. Le taux de justification du critère est de 42,9% (30/70). Illustrons les cas de *mug* et de *shit* :

- « Jusqu'au 13 octobre se tient au hangar du 47, avenue Kennedy à Biarritz, le premier marché d'objets anciens autour de la surf culture. Du **mug** hawaïen, à la planche en séquoia massif à 10 000 euros, 250 pièces sont à saisir. Tous les jours de 10 heures à midi et de 14 heures à 19 heures. Fermé le mardi. » - Anonyme (*Sud Ouest*), 2013-10-07
- « à c'est midi je vient de me boire un **mug**, plus un du matin, puis après j'en prend a 4h si j'ai le tps. Je pensais ke le thé vert au contraire était meilleurs que les autres et boosté avant de pratiquer un sport... » - andclone (*Doctissimo*), 2008-09-19
- « Me Arianna Di Dio, toujours pour la défense, a tenté de nuancer la réalité des bandes urbaines. « Ils forment un groupe, oui, mais ils ne se trouvaient pas dans cette galerie avec un esprit guerrier. Ils voulaient acheter du **shit** et ont eu peur d'un individu qu'ils ne connaissaient pas, et qu'ils ont pris pour un espion de leurs adversaires. Et tout a dérapé. Il a frappé, mais sans intention de tuer. D'ailleurs, son coup a été porté à la jambe... » » - Jean-Pierre Borloo (*Le Soir*), 2010-06-10
- « Phénomène urbain ? Christopher Nunes, citoyen de Pauillac, ne se sent pas plus à l'aise chez lui, en zone rurale. Il s'en prend aussi aux gendarmes : « Ils m'ont arrêté un jour et m'ont demandé si je fumais du **shit**. J'ai dit non et ils m'ont ri au nez. » » - Hervé Mathurin (*Sud Ouest*), 2011-07-29

Le taux de ce critère est plutôt faible. Considérant les variantes phonétiques, les rimes des emprunts à l'anglais étaient surtout féminines (M = 5, F = 11, V = 0), alors que les rimes avec les emprunts des autres langues étaient principalement masculines¹²¹. Au total, bien que la majorité des groupes de rimes soient principalement masculins (M = 35, F = 24, V = 15) et qu'entre 70% (genre majoritaire) et 76,5% (occurrences en corpus) de nos emprunts portent aussi la marque du masculin, les emprunts des langues ayant le plus d'occurrences au féminin (italien et arabe) correspondent à des groupes de rimes majoritairement masculins. C'est cette association qui explique le faible taux de justification du critère.

¹²¹ Japonais (M = 11, F = 1, V = 6), arabe (M = 8, F = 5, V = 4), italien (M = 11, F = 7, V = 5).

5.3.6 Paronymie

Afin de tester ce critère, il fallait que le genre majoritaire de nos emprunts soit identique au genre des paronymes que nous avons identifié. Sur 100 unités, seules 26 ont trouvé au moins un correspondant paronymique. Sur ces 26 unités, onze (11) ont le même genre que leur correspondant, comme c'est le cas de l'emprunt majoritairement masculin *barda* ayant comme paronyme le mot masculin *bardeau*. 15 ont un genre différent, comme l'emprunt féminin *pizza* a pour correspondant paronymique le mot masculin *pita*. Le taux de justification de ce critère est de 42,9% (11/25). Illustrons les cas de *barda* et de *pizza* :

- « Entre avril 1974 et novembre 1975, le Portugal quitte l'Afrique, qu'il avait été le premier à coloniser (Luanda, la capitale de l'Angola, a été fondée en 1575). Le grand mouvement de décolonisation, commencé en 1960 avec, cette année-là, l'indépendance du Cameroun, du Congo, de la Côte d'Ivoire et du Tchad, s'achève en hiver 1975 sur les quais de Lisbonne encombrés du **barda** de dizaines de milliers de «retornados» d'Afrique. L'armée portugaise n'a plus voulu se battre pour garder ses possessions. Elle ne les voyait d'ailleurs plus comme des «possessions». » - Joëlle Kuntz (*Le Temps*), 2010-05-22
- « Pour ce vagabond, anciennement grand reporter et aujourd'hui fin critique gastronomique et romancier, il existe toujours un lieu où se rassembler et se ressembler, faire une pause. Voyager est un sentiment qui perdure et que l'auteur cultive amoureusement, peut-être pour encourager notre désir. Tout un **barda** d'émotions, de fantaisies mélancoliques, de parfums et de saveurs tenaces. De jolies digressions sur l'effrayant mystère des femmes, quelques velléités changées en regrets et un hommage appuyé à la plus redoutable des maîtresses, l'écriture. » - Isabelle Bunisset (*Sud Ouest*), 2011-06-19
- « Je n'hésiterais pas une minute à l'oublier en cave pendant quelques années. Toutefois, déjà, maintenant, sur une bonne **pizza** maison, à moins que vous ne préfériez un spaghetti putanesca, on s'en régale. » Claude Langlois (*Le Journal de Québec*), 2011-06-05
- « Le propriétaire d'Archibald, François Nolin, et le chef exécutif de ce restaurant, Olivier Blais, ont présenté récemment leur nouveau menu automnal. Un menu truffé de plats dont le caractère se marie parfaitement à la saison comme la tartiflette savoyarde et le fameux tomme gratiné et viande de grison. Les **pizzas** sont fort populaires et originales chez Archibald. L'équipe propose désormais une excellente **pizza** au smoked meat, une pizza à la merguez, une **pizza** au porc embrasé au gingembre et signale le retour de la populaire **pizza**-suisse, pourvue entre autres de gruyère, d'emmental, de fontina et de viande de grison ! Parmi les autres nouveautés, dans les plats principaux, un osso bucco à l'orange, une côte de cerf au stilton et un demi-poulet de Cornouaille, sauce au rhum, miel et piments de la Jamaïque. Le réputé foie de veau à la zurichoise, offert l'automne dernier, réapparaît lui aussi sur la carte tout comme le magret de canard, sauce aux framboises. » - Pierre Champagne (*Le Soleil*), 2006-12-01

Il a été particulièrement complexe de trouver des correspondants paronymiques pour les unités à l'étude. Puisqu'il n'existait pas de manière automatisée pour trouver des paronymes, nous avons consulté un dictionnaire en retenant les homonymes et nous avons utilisé notre intuition selon un ensemble de règles définis. De plus, nos emprunts sont composés en large mesure (49%) de termes trisyllabiques. Plus le nombre de phonèmes (et donc de syllabes) est élevé, plus il est complexe de trouver des paronymes réalistes. Au final, le taux de justification est faible, mais ne s'applique qu'à un nombre restreint de cas.

5.3.7 IP par similitude conceptuelle

Afin de tester ce critère, il fallait que le genre majoritaire des unités de chaque paradigme intégrant l'unité à l'étude soit identique au genre majoritaire des unités correspondantes. Sur les 100 unités, nous avons exclu la paire d'unités *jam* et *jam-session*, car nous considérons qu'il s'agit de deux formes différentes de la même unité lexicale et leur genre majoritaire en corpus est différent. Nous avons donc considéré les 98 autres unités comme analysables. Sur ce nombre, le genre majoritaire de 76 unités (77,6%) est identique au genre de leur paradigme, comme c'est le cas de l'emprunt majoritairement féminin *interview*¹²² intégrant le paradigme du même genre majoritaire composé des unités *entrevue*, *rencontre*, *entretien*, *conférence*. Le genre majoritaire de 19 unités est différent du genre de leur paradigme, comme c'est le cas de l'emprunt majoritairement masculin *scoop* intégrant le paradigme féminin composé des unités suivantes : *nouvelle*, *exclusivité*, *primeur*, *information* et *révélation*¹²³. Le genre de 3 paradigmes est variable, comme c'est le cas du paradigme composé de *vodka*, *gin*, *bière* et *whisky* intégrant l'unité *saké*. Le taux de justification du critère est de 77,6% (76/98)¹²⁴. Illustrons *interview* et *saké* de quelques exemples :

- « Dominic et Martin pour un **interview** pour le journal étudiant » - Anonyme, sans date (corpus québécois informel)
- « De passage à Paris, mercredi 24 juillet, ces deux hommes, de confession sunnite, ont rencontré le président français François Hollande. Dans la foulée de cette entrevue, ils ont accordé une **interview** exclusive au Monde. » - Propos recueilli par Benjamin Barthe, Rémy Ourdan et/ou Yves-Michel Riols (*Le Monde*), 2013-08-26
- « j'avais vu une **interview** d'un professeur ds oumma.com, où il parlait du hijab ds les sociétés occidentales.très intéressant!! et bien il disait qu'en islam, le message de nore prophète envers la femme est qu'elle s'instruise et soit savante.» - Profil supprimé (*Doctissimo*), 2007-01-07
- « Dans certains cas, des oenologues patentés participent à l'écriture du scénario. Ken-Ichi Hori, le plus grand oenologue nippon, a publié avec Araki Joh et Shinobu Kaitani les six volumes du Sommelier, le manga qui a permis au pinot noir de mettre la gueule de bois au **saké** ! Un autre manga, Les gouttes de Dieu, a été récompensé du titre de « meilleur livre du monde sur le vin » par le Gourmand World Cookbook. » - Anonyme (*Le Soir*), 2012-01-09
- « Je sers le **Saké** chaud tout au long du repas. » - Anonyme, sans date (corpus québécois informel)

De manière générale, les emprunts de toutes les langues donneuses suivent le même schéma : plus de correspondances de genre que de non-correspondances ou de cas variables, et plus ou autant de non-

¹²² Rappelons que, bien que l'unité *interview* porte majoritairement la MAF, il y a de nombreuses occurrences où elle porte la MAM.

¹²³ *Scoop* peut être à la fois conçu comme la nouvelle et comme l'événement couvert par la nouvelle. Advenant le deuxième cas, on pourrait rapprocher le concept d'*événement* ou de *fait*.

¹²⁴ Nous tenons ici à nuancer nos propos : le taux de justification des critères est calculé à partir du nombre de cas où le genre majoritaire de l'unité correspond à celui du critère. Dans le cas de l'IP par similitude conceptuelle, le paradigme est composé d'unités lexicales déterminées par un locuteur, ce qui rend la vérification auprès d'un ensemble de locuteurs impossible. Par définition, le paradigme latent est subjectif, car construit par le locuteur selon son expérience. Pour atteindre un paradigme à portée collective, il faudrait demander à plusieurs locuteurs de constituer un paradigme conceptuel ou se baser sur des ouvrages lexicographiques.

correspondances que de cas variables. Les emprunts à l'anglais forment le groupe avec le moins de correspondances¹²⁵ et le plus de non-correspondances¹²⁶. Cela peut peut-être s'expliquer par le fait que certains critères sont plus efficaces avec les emprunts à l'anglais, comme celui de la syllabité entière (19/25), car la plupart des emprunts à l'anglais que nous étudions sont monosyllabiques (19/25, sur 22 emprunts monosyllabiques).

La combinatoire des unités reflète les significations de celles-ci, car le verbe utilisé avec l'unité doit partager un minimum de traits sémantiques. Par exemple, dans le cas du dessert *tiramisu* et de l'entrée *antipasto*, on retrouve respectivement les combinaisons *terminer par* et *commencer par/avec*. Illustrons les cas de combinaisons de *tiramisu* et d'*antipasto* :

- « Sur place ou à emporter, le client peut également déguster des nuggets avant de **terminer par** un **tiramisu**, une tarte au chocolat, une glace ou un yaourt. Comme ses fameux homologues franchisés, Le Délice propose en outre un menu " kids ". » - Anonyme (*Sud Ouest*), 2013-04-03
- « Cette cité, dont le citoyen le plus célèbre a été Archimède lui-même, compte de nombreux restaurants où la carte offre d'excellents fruits de mer ainsi qu'une grande variété de pâtes et de pizzas. En fait, ce ne sont pas les restaurants qui manquent, et ce, dans toutes les catégories de prix. On y sert d'excellents antipasti (assortiments de hors-d'oeuvre), des pâtes aux sardines ou aux moules, de même que quelques bonnes pièces de viande, le tout accompagné de fromages corsés et d'un bon choix de vins. Pour la totale, il faut **commencer avec** un **antipasto**, poursuivre avec un primo piatto (pâtes ou risotto), puis un secondo piatto à base de viande ou de poisson avant de terminer avec un fruit ou un dessert. Ouf! » - Alexandre Shields (*Le Devoir*), 2008-03-29

Or, le cas de l'unité *farfalle* a révélé qu'en corpus, l'usage traitait souvent¹²⁷ de la pâte alimentaire en forme de papillon comme un bricolage. Les combinaisons les plus fréquentes avec cette unité sont *préparer*, *décorer*, *peindre* et *casser*. C'est le seul cas de nos unités où nous avons laissé de côté l'aspect combinatoire. Illustrons le cas de *farfalle* :

- « À l'aide de la perceuse et sur un support qui ne craint pas d'être abîmé (comme un morceau de bois récupéré) percer un trou dans la partie en creux de la **farfalle**, à une distance d'environ de pouce du bord. Procéder doucement en pressant légèrement sur le bouton de tension de la perceuse pour ne pas casser la **farfalle**. Répéter l'opération de l'autre côté. » - Christelle Tanielian (*La Presse*), 2013-05-11
- « 1. Disposer une grande feuille de papier journal ou une toile cirée sur la table pour la protéger. 2. Peindre une couche de peinture ou de vernis sur la **farfalle** et laisser sécher. 3. Appliquer une deuxième couche et laisser sécher de nouveau. On peut ensuite décorer la **farfalle** l'aide d'un stylo argenté, doré ou de couleur. 4. Passer un fil dans les deux trous de la **farfalle** et faire un noeud au bout. » - Christelle Tanielian (*La Presse*), 2013-05-11

¹²⁵ Correspondances : anglais = 15, japonais = 21, arabe = 21, italien = 19.

¹²⁶ Non-correspondances : anglais = 7, japonais = 3, arabe = 4, italien = 5.

¹²⁷ Surtout utilisé par Christelle Tanielian (*La Presse*), 2013-05-11.

Nous sommes d'avis que ces combinaisons ne relèvent pas de nouvelles significations, mais du discours d'usages spécifiques de la réalité en tant que telle. Il est possible, dans le cas des combinaisons avec *farfalle*, que nous ayons trouvé un article avec un grand nombre d'occurrences abordant le sujet du bricolage, ou que le sujet d'une conversation de forum culinaire ait été détourné sur le thème du bricolage. Il n'en reste pas moins que la *farfalle* est une pâte et qu'elle est généralement utilisée dans un contexte alimentaire.

5.3.8 IP par hyperonymie

Afin de tester ce critère, le genre majoritaire des emprunts devait être identique au genre d'un groupe d'hyperonymes. Chaque unité est analysable, puisqu'au moins un hyperonyme y a été associé. Sur les 100 unités, le genre majoritaire de 61 emprunts est identique au genre majoritaire du groupe d'hyperonymes, comme c'est le cas de l'emprunt majoritairement masculin *niqab* associé au groupe hyperonymique masculin composé de *voile*, *drap* et *masque*. Dans 39 cas, le genre majoritaire était différent de celui du groupe, comme c'est le cas de l'emprunt majoritairement masculin *surimi* associé au groupe hyperonymique féminin composé de *chair* et *pâte*. Le taux de justification du critère est donc de 61% (61/100). Illustrons les cas de *niqab* et de *surimi* :

- « « Un juste procès exige une ouverture et une communication totales, a estimé le juge Peter Murphy. Je suis convaincu que le port du **niqab** entrave cette ouverture et la communication. » Le Parlement italien, puis le gouvernement néerlandais ont aussi envisagé des mesures d'interdiction du voile intégral, sans aller jusqu'au bout pour l'instant. » - Sylvie Kauffmann (*Le Monde*), 2013-09-24
- « Pour que les choses soient claires, précisons qu'il s'agit bien ici d'une interdiction de la burqa et du **niqab**, soit des voiles qui recouvrent entièrement le corps et le visage. La première est reconnaissable à son grillage qui dissimule les yeux. » - Olivier Perrin (*Le Temps*), 2013-09-23
- « Quelques dés d'ananas dépayseront une simple salade de pâtes au **surimi**. » - Simone Estran (*Le Nouvelliste*), 1999-02-27
- « L'usine Protimer s'est installée sur la commune de Marans il y a dix-sept ans. Au total, soixante-et-un employés travaillent à la fabrication d'aliments à base de poisson, comme par exemple le **surimi**. La société produit chaque année environ 2000 tonnes de denrées. La société Protimer est une filiale de Bongrain SA » - Anonyme (*Sud Ouest*), 2007-05-03

Les emprunts aux LSG sont ceux avec le moins de concordances (anglais : 10; japonais : 11) et le moins de groupes masculins (anglais : 9; japonais : 10) lorsque nous comparons les emprunts aux LAG (arabe : 22 concordances, 12 groupes masculins; italien : 18 concordances, 14 groupes masculins). Ce constat va contre notre intuition, car les emprunts des LAG sont ceux présentant le plus d'occurrences du féminin en corpus, le plus de groupes hyperonymiques masculins, et le plus de concordances entre les deux.

5.3.9 IP par ellipse d'un signe entier

Afin de tester ce critère, il fallait que le genre majoritaire des unités soit identique au genre des termes élidés. Sur nos 100 unités, 33 ont au moins un signe élidé associé. Sur les 33, le genre majoritaire de 14 est

identique au genre des signes élidés, comme pour l'emprunt majoritairement masculin *panini* au signe élidé *sandwich*. Le genre de huit (8) unités est différent, comme c'est le cas de l'emprunt majoritairement féminin *ciabatta* dont on retrouve les signes élidés masculins *pain* et *sandwich*. Onze (11) unités sont considérées comme variables, car elles sont associées à plusieurs signes élidés portant les deux genres. Le taux de justification du critère est de 42,4% (14/33). Illustrons les cas de *panini*, de *sandwich* et de *ciabatta*, incidemment référant tous à un type de réalité similaire, celui des pains :

- « salut tout le monde j'espère que vous pouvez m'aider, le problème que j'ai c'est que je mange presque rien du tout, je prend pas de petit dej. juste du café au boulot, le déjeuner je prend la moitié du **panini** et le soir je mange quelques biscuits. Je suis 1m71 avec 43 kg aidez moi » - pahpolla (*Doctissimo*), 2010-04-28
- « Rehaussé d'un zeste de citron et d'ail, ce plat végétarien est composé d'épinards, de parmesan et de tomate. Une pizza, un steak frites, une bonne salade fraîche et même une poutine peuvent se retrouver dans votre assiette. Le poulet barbecue et le **panini** chaud viendront régaler vos papilles gustatives durant la période estivale. » - Nadia Bergeron (*Le Journal de Québec*), 2012-03-17
- « Celui d'Au Pied de cochon est dur à battre, avec sa poutine minute. Pas de cochon dans mon salon propose un excellent **sandwich** au porc effiloché classique. Fous Truck décline le croissant en superbes sandwiches et desserts raffinés. J'ai aussi adoré la soupe de Landry et filles, surtout quand la température baisse. Réconfortant. Longue vie à ces nouvelles adresses ambulantes, même en hiver, on espère! » - Marie-Claude Lortie (*La Presse*), 2013-11-30
- « Eh bien hurra, je ne suis pas seule. Je cherchais une réponse sur Google et c'est ici que je suis tombée. Ça m'est arrivé deux fois aujourd'hui, à une intervalle d'environ 30 minutes. La première chose que j'ai mangé était un **sandwich** à la dinde fumée avec un petit peu de mayo. J'ai été vraiment surprise de cette réaction à la **sandwich**. Ensuite, j'ai eu une réaction après avoir consommé une sucette, pas longtemps après la **sandwich**. Là, j'ai peur de manger encore! J'ai constaté à part de cela que les aliments contenant de l'Aspartame me faisaient cette réaction.» - invite_MMeliMelo (*Doctissimo*), 2006-09-26¹²⁸
- « je me fais un petit **sandwich** à la salade de jambon (jambon haché + mayo + céleri haché... à faire en petite quantité à cause du céleri sinon le lendemain c'est dégueu) » - Anonyme, sans date (corpus québécois informel).
- « D'autres spécialités ont vu le jour chez Nathalie et Pierre, comme le Waldkorn, un pain allemand multicéréales, ou encore le **Ciabatta**, un pain italien pétri à l'huile d'olive. » - Anonyme (*Sud Ouest*), 2009-06-09
- « Le deuxième choix, le burger au mignon, avec oignons et champignons, était savoureux et rendu gentiment croustillant puisque inséré dans une **ciabatta** bien craquante. Ce burger dépassait les simples règles de l'art, accompagné de grosses frites paysannes associées à un quatuor de mayonnaises (ail, curry, tomate et pesto). La grosse frite, toute simple, avait du panache. Les gourmands que nous sommes en avons repris une petite portion pour finir les mayonnaises. » - Rémy Charest (*Le Devoir*), 2004-04-23

Si le taux de justification de ce critère est faible, c'est parce qu'un tiers (11/33) des unités sont associés à un groupe de signes élidés variant en genre, comme c'est le cas des signes masculin et féminin *poudre* et *thé*, élidés avec l'emprunt *matcha*. En retirant les cas variables, on obtiendrait un taux de justification de 63,6% (14/22). Ce critère est similaire aux critères de réanalyse morphologique (voir 6.3.1) en ce que ces trois critères ont un taux de justification faible lorsqu'on considère tous les cas de figure, mais un taux passable ou

¹²⁸ Ce commentaire varie le genre de l'emprunt *sandwich*.

suffisamment fort lorsqu'on ignore les cas variables. Prenant pour acquis qu'un locuteur ne fait pas le tri parmi tous les cas et applique la règle telle quelle, il faut considérer les cas variables comme des « non » (donc dans le total seulement) afin de ne pas compromettre l'intégrité de la règle. Illustrons le cas de *matcha* :

- « C'est vrai qu'ils sont fous, ces desserts. Toutes les pâtisseries sont faites au beurre, bien sûr. On aime autant les classiques que sont les millefeuilles et les éclairs que les superbes bûches de Noël présentées cette année, dont une au thé au **matcha** et yusu, aux marrons et aux fruits de l'argousier, sans gluten, et la croquante avec chocolat et croustillant de noisettes pour les grands amateurs de plaisirs fous. » - Philippe Mollé (*Le Devoir*), 2011-12-03
- « Moi aussi je préfère le sencha mais il ne se présente qu'en feuilles, pas en poudre. L'inconvénient principal du **matcha**, c'est quand même qu'il est SUPER amer ! » - verobassu (*Doctissimo*), 2011-11-10

5.3.10 Conservation du genre

Afin de tester ce critère, il fallait que le genre majoritaire des unités provenant d'une LAG soit identique au genre de l'unité originelle dans la langue donneuse. Nous avons 50 unités provenant de l'arabe et de l'italien. Sur ces 50 unités, 42 ont le même genre que dans la langue donneuse (comme *spaghetti* et *bruschetta*) et huit (8) n'ont pas le même genre (comme *sépia* ou *barda*). Le taux de justification de ce critère est de 84% (42/50). Illustrons les cas de *spaghetti*, de *bruschetta* et de *sépia* :

- « Cette coopérative tire habilement son épingle du jeu avec des vins où le cépage s'exprime avec franchise, fraîcheur et équilibre. Il y a même ici suffisamment d'étoffe, de caractère et de vinosité pour permettre au **spaghetti** et à ses grosses boulettes de rayonner dans toute leur rusticité. » - Jean Aubry (*Le Devoir*), 2010-11-05
- « "L'hôtel où nous demeurons vient d'ouvrir ses portes. On nous a servi du **spaghetti** pour déjeuner au lieu des crêpes avec du sirop d'érable! C'est un choc de culture. Mais comme athlète, tu n'as pas le choix de t'adapter et espérer que tout sera bien lors des Jeux olympiques, a affirmé l'athlète franco-ontarien âgé de 27 ans. » - Martin Comtois (*Le Droit*), 2013-02-13
- « Entre 14 h 30 et 17 h, Yohan était aux fourneaux, sous la supervision de papa. Oignons et tomates en dés pour son potage, pesto maison pour la **bruschetta** d'artichauts et pignons, escalopes de veau aplaties entre deux feuilles de papier ciré pour le saltimbocca. Heureusement, l'ambitieuse (et franchement délicieuse) Marjolaine au chocolat et aux noisettes avait été préparée la veille. » - Ève Dumas (*La Presse*), 2013-02-23
- « Hum elle a l'air bonne ta **bruschetta** » - violette.pipelette (*Doctissimo*), 2013-03-27
- « A l'instar de ce qui a été fait pour Apocalypse, ces images ont été colorisées - mais de façon « light », respectueuse, on est parfois plus proche du **sépia** que de la couleur - afin de rendre le film plus vivant, plus accessible, notamment auprès des jeunes générations. » - Agnès Gorissen (*Le Soir*), 2010-05-29
- « La chaleur réconfortante de son propre salon suffit le soir, quand Chopin joue discrètement ses nocturnes en arrière-fond et qu'un éclairage discret ajoute au feutré de l'atmosphère. Il ne manque plus alors qu'un bon verre de porto ou un verre de vin de glace pour se replonger dans la **sépia** de ses souvenirs. Je vous en propose quelques-uns cette semaine. » - Claude Langlois (*Le Journal de Québec*), 2011-01-15

Ce critère a un taux de justification très élevé comparé aux autres critères PO (61%-77,6%), mais il ne s'applique que dans les cas où la langue donneuse a deux genres qui correspondent *grosso modo* au masculin et au féminin. Nous n'avons pas examiné le genre d'emprunts venant de langues donneuses à trois

genres (russe, allemand, latin), ni de d'autres systèmes de classement (classes nominales du zoulou, animéité en basque, etc.). Nous entrevoyons la possibilité, dans les cas où les systèmes de genre sont trop différents du système du français, que le genre ne soit pas conservé et que l'archigenre prenne le dessus.

Sur les 8 cas dont le genre de l'unité ne correspond pas dans les deux langues, cinq (5) sont des emprunts à l'arabe (*keffieh*, *barda*, *brick*, *kefta*, *merguez*) et trois (3) sont des emprunts à l'italien (*sépie*, *calzone*, *minestrone*). Sur ces huit (8) emprunts, trois (3) comportent le suffixe *-a* et portent la MAM dans la langue d'origine, mais masculins en français. Le genre de tous ces emprunts, sauf *minestrone* (car de genre majoritaire variable), peut être expliqué par des raisons sémantiques (IP par similitude conceptuelle ou par hyperonymie). Par exemple, *merguez* porte la MAM en arabe, mais féminin en français à cause de *saucisse* et de son paradigme de genre féminin :

- « Coucou les filles Je viens de rentrer, c'est pour cela que je ne suis pas passée avant poster, donc depuis ce matin que je suis partis en vadrouille et ce soir invitation, mais la seule chose qu'il y avait c'était une sandwicherie j'ai donc prétexté que je n'avais pas faim, sa sentais la **merguez** et la saucisse grillée à pleins nez mais non je n'ai pas faim bref c'est pas plus mal, j'ai réussi à la buvette à avoir 2 café donc ca m'a calmé et là en rentrant je me suis fais un chocolat chaud, donc pour ce soir ca ira enfin on est déjà demain lol» - Profil supprimé (*Doctissimo*), 2010-07-20
- « Merci fabio pour tes conseils: en ce qui concerne la saucisse et la **merguez**, ja savais mais mon mari avait fait des grillades et depuis le 28 juillet, je n'avais plus mangé la même chose que lui alors c'était juste pour la "convivialité". Ce matin, j'ai déjeûné, un café au lait, 40g de pain légèrement beurré et une orange. Pour le reste de la journée, je ne sais pas encore mais je vous tiens au courant. » - laetym (*Doctissimo*), 2006-09-17

Le cas de *brick* est particulier, car il existe sous deux formes, une qui réfère à un aliment fait avec plusieurs feuilles, et une qui réfère à une seule feuille. Le problème, c'est que pour certains individus, *brick* « aliment fait avec plusieurs feuilles » porte parfois la MAM, parfois la MAF, et vice-versa pour l'autre signification (qui porte majoritairement la MAF). De plus, au Québec, *brick* réfère à un type de fromage et porte la MAM. *Brick* réfère également à un type de bateau, qui porte également la MAM. Par exemple :

- « Ambiance feutrée, lumière tamisée, la Coquette porte bien son nom. Discret et chic, ce restaurant tunisien, retraite des noceurs de Bagdad, séduit une clientèle aisée qui dépense sans compter. Ben y sert avec faconde un **brick** au thon délicieux et son couscous de légumes change agréablement de l'incontournable riz au mouton. Ce ne sont pourtant pas ses talents culinaires qui lui valent les faveurs de la bourgeoisie irakienne.» - François Didier (*Libération*), 2004-02-18
- « Merci les filles ! Mon gouter: une **brick** cuite au four donc sans matières grasses et un yaourt nature » - Ziico92 (*Doctissimo*), 2011-04-03
- « Coucou valou ! Oui effectivement nous avons fait fort dans le papotage ce weekend.... caro si ma mémoire est bonne la distinction entre la **brick** et la filo est la suivante : brick : semoule de blé + eau + sel Filo : farine de blé + eau + sel La seconde étant plus friable que la première. Moi je l'utilisait beaucoup en pâtisserie. L'avantage de la **brick** est quelle est plus croustillante. J'ai déjà fait des aumonière avec des feuilles de brick mais je double le fond de façon à ce que l'humidité ne ramolli pas la pâte sinon tout s'écroule ! (et après tu as l'air d'une nouille c'est moi qui te le dit !) » - cleo9269 (*Doctissimo*), 2012-10-29

- « C'est sans doute effectivement le contrecoup de la crise qui te fait manger moins, si tu n'es pas partie dans l'idée du "je dois rattrapper ça, je mange pas" Ninis, ma tarte tatin, je ne l'ai pas faite avec du **brick** : avec l'oeuf battu+fromage blanc, ça se serait humidifié et ça serait resté tout mou. Enfin, honnêtement, hier ma tarte elle était moche! J'ai dû laisser cuire trop peu, du coup, les courgettes sont restées collées dans le plat... men fous, c'était bon! Le lait fermenté, c'est pareil que le kéfir non? Moi j'adore, c'est frais. LA pâte filo, c'est sans doute pareil que le **brick**. Peut-être que l'origine est différente, simplement? Je devrais essayer aussi les papillottes. J'ai déjà voulu, puis chaque fois, je fais une tarte salée en brick Le brousse, j'ai goûté il y a deux jours, c'est super bon. En plus, c'est pas pâteux, j'aime bien parce que le goût est assez léger et frais. Je crois que tu pouvais franchement ouvrir e paquet : mon paquet de brick est ouvert depuis trois semaines facilement et tout est tj's impeccable'...» - cioccolata (*Doctissimo*), 2006-09-06
- « J'utilise fréquemment le mozzarella, et aussi le **brick** dans mes pizzas traditionnelles, mais il n'y a pas de règle particulière puisque chaque fromage a son goût. » - Anonyme, sans date (corpus québécois informel)
- « Pierres roses de l'ouvrage, granit de même couleur de la cale qui lui fait suite, à gauche, et qui ferme l'anse de la Corderie. Assis face au large, le regard suit la mer qui monte par petites vaguelettes. L'image d'un **brick** sous voiles de retour de quelques courses après des galions chargés d'or vient un instant occuper l'espace, avant que le rêve se brise. » - Serge Bolloch (*Le Monde*), 2006-02-11

Brick est la seule unité que nous avons considérée à genre originellement variable, et pour cause. Finalement, les deux emprunts *minestrone* et *calzone* portent le suffixe flexionnel italien *-e*, suffixe marquant le masculin ou le féminin. Illustrons le cas de *calzone* :

- « Pas puriste pour deux sous, j'ai joyeusement mélangé les genres, ce weekend, en combinant une garniture mexicaine et une pâte plus digne du **calzone** que de l'empanada. Comme boulanger amateur, j'ai toujours du levain en route et, cette fois, l'idée d'un pâté fourré me semblait faire agréablement changement. » - Rémy Charest (*Le Journal de Québec*), 2013-06-02
- « Une cuisine simple et familiale qui ravit une clientèle d'habitueés. Les pizzas sont délicieuses et n'ont rien à envier aux napolitaines, comme la **calzone** farcie à la mozzarella et à la ricotta, une spécialité de Salerne. » - Emmanuel Tresmontant (*Le Monde*), 2012-11-01

Puisque le genre du suffixe *-e* est moins marqué que celui du *-a* féminin ou du *-o* masculin, les locuteurs du français ne savent pas automatiquement quel genre donner à l'unité et ils tendent davantage vers des critères sémantiques ou vers d'autres critères de forme physique.

5.3.11 Initiale vocalique

Afin de tester ce critère, il fallait que le genre majoritaire des unités analysables soit féminin ou variable. Sur les 100 unités, neuf (9) commencent par une voyelle (ou un <h>). Sur les neuf, cinq (5) portent la marque du féminin (*intifada*, *omertà*) et quatre (4) celle du masculin (*hidjab*, *expresso*). Le taux de justification de ce critère est de 55,5% (5/9). Illustrons les cas d'*intifada*, d'*omertà*, d'*hidjab* et d'*expresso* :

- « Le bâtiment en construction est situé à quelques centaines de mètres de la Mouqata, le quartier général de l'Autorité palestinienne. Du sommet de la tour, on aperçoit la cour rectangulaire de l'édifice, dont les murs en partie détruits par les bulldozers israéliens - et aujourd'hui reconstruits - étaient devenus un des symboles de la deuxième **Intifada**. Après ces années noires, la tour de la Palestine représente le renouveau récent de Ramallah : une ville ouverte, cosmopolite, où l'amélioration économique, au moins apparente, a permis le

retour des investisseurs étrangers et de certains hommes d'affaires palestiniens de la diaspora, malgré le maintien des restrictions israéliennes. » - Delphine Matthieussent (*Libération*), 2010-09-21

- « Il y a une évolution. Nous avons adopté une résistance non violente, surtout contre la construction du mur. Ça a déstabilisé les Israéliens, qui auraient voulu nous pousser à une nouvelle **intifada**. Ce sont encore des stratégies et des provocations qui cherchent à nous piéger. » - Propos recueilli par Hervé Chassain (*Sud Ouest*), 2016-04-01
- « Le parasite vecteur de la maladie, bénigne, qui se transmet par contact prolongé peau contre peau était donc bien « dans l'air », et au terme de sa « petite enquête », Nadine Lacoste-Barby regrette cette **omertà** à son égard : « Ça aurait dû être une question de respect des autres... À l'inverse, il faut dire merci à la maman de l'élève qui nous a prévenus en début de semaine. » Son fils présentait des plaques cutanées suspectes qui le démangeaient... » - Thomas Longué (*Sud Ouest*), 2011-12-16
- « Très peu de gens vont voter aux municipales -et encore moins s'intéresser à ce qui se passe à l'hôtel de ville. Une **omertà** très lourde règne parmi les syndiqués, les fonctionnaires, les administrateurs. » - Benoît Aubin (*Le Journal de Québec*), 2012-04-24
- « Demandons-nous maintenant quel est l'aspect qui se montre en premier lorsque nous sommes devant une personne voilée : un soi (une personne) ou une réalité donnée, dépersonnalisée par le voile ? Pour plusieurs, et nous faisons ici abstraction du **hidjab**, ce sera le voile qui primera sur la personne. » - René Bolduc (*Le Devoir*), 2013-02-02
- « Oui, le voile intégral (comme celui que portait la musulmane qui a été expulsée) est choquant. Mais le **hidjab** l'est tout autant. C'est la version light du voile intégral, c'est tout. » - Richard Martineau (*Le Journal de Québec*), 2010-03-03
- « car ce que vous ne savez pas mes amis, c'est que ça fait gonfler! une p'tite dose par-ci par-là, dans mon thé, dans mon café, ... après c'est ma mère qui m'a montré comment mieux m'en servir, dans la machine à laver, ça fait comme un bon **expresso** maintenant!!! » - ursulgazon (*Doctissimo*), 2008-11-18
- « choisissez un bon **expresso** avec un peu de lait chaud. » - Anonyme, sans date (corpus informel québécois)

Le nombre d'unités commençant par une voyelle ou un <h> est très réduit. Cela est peut-être dû à la sélection de celles-ci, mais de notre expérience, le pourcentage (10%) de noms à initial vocalique nous semble représentatif. Une étude rapide de notre préliste de 171 unités nous confirme que seules 19 (11,1%) commencent par une voyelle ou un <h>.

Nous pensons qu'il faut relativiser ce critère. Au-delà du faible échantillon analysable, les études de Barbaud *et al.* (1982) suggèrent qu'au Québec, les noms masculins ont plus tendance à se féminiser que les noms féminins à se masculiniser. Cela ne veut pas dire qu'il y aura un plus grand nombre de noms portant la marque du féminin. C'est pour cette raison que nous avons considéré les noms à genre majoritaire variable, car ils peuvent être la trace d'une féminisation ou d'une masculinisation. Les noms féminins peuvent être le résultat de la féminisation (et pas de la masculinisation), alors que les noms masculins ne peuvent pas être le résultat de la féminisation.

5.3.12 Opérationnalité simultanée de plusieurs critères

Nous avons remarqué qu'il y avait, pour la plupart de nos unités, plusieurs critères pouvant être opératoires et pouvant justifier le genre de l'unité simultanément. Nous avons calculé le nombre de critères opératoires pour

chaque unité et nous avons fait la moyenne en regroupant les emprunts selon leur langue donneuse. Considérant que le critère de la conservation du genre ne s'applique qu'à la moitié de l'échantillon, et qu'il est opératoire à 80% pour les emprunts à l'arabe et à 88% pour les emprunts à l'italien (moyenne 84%), nous avons retiré 0,80 à la moyenne obtenue pour les emprunts à l'arabe et 0,88 à la moyenne obtenue pour les emprunts à l'italien.

Voici, en ordre croissant, les moyennes (ajustées, avec 0,80 et 0,88 en moins respectivement pour l'arabe et l'italien) que nous avons obtenues pour les emprunts aux langues suivantes : anglais, 3,88 critères PO par unité; arabe, 4,12 critères PO par unité; italien, 4,16 critères PO par unité; japonais, 4,64 critères PO par unité. La moyenne de ces moyennes est de 4,2. Les emprunts sont donc généralement justifiés par environ 4 critères. L'écart de critères PO applicables simultanément se situe entre 1 et 10¹²⁹. Sur nos critères, sept s'appliquent à une minorité d'unités de notre échantillon¹³⁰. Bien que les moyennes soient très rapprochées les unes des autres, les emprunts à l'anglais forment la catégorie qui peut être la moins justifiée par nos critères, alors que les emprunts au japonais forment la catégorie qui peut être la plus justifiée par nos critères. À l'inverse, les emprunts aux LAG semblent pouvoir être justifiés par un nombre de critères près de la moyenne, relativement stable.

Une seule unité peut être justifiée par 10 critères, *dojo*. Le résultat est étonnant, considérant que ce n'est pas un emprunt d'une LAG, catégorie d'emprunts bénéficiant du critère de la conservation du genre. Notons que les emprunts masculins se terminant en [o] et <o> (*kimono, miso, carpaccio, prosciutto*, etc.) ont généralement deux critères PO de plus que les autres emprunts, car les deux critères de réanalyse morphologique (phonétique et graphique) peuvent justifier le masculin. À l'inverse de *dojo*, quatre unités ne peuvent être justifiées que par un critère PO (que nous devons supposer opératoire, à défaut de considérer l'archigenre comme un de nos 15 critères d'analyse). Ces unités sont *gig, minestrone, party* et *rave*. De ces quatre unités, quatre sont des emprunts à l'anglais, ce qui rejoint notre observation voulant que les emprunts à l'anglais aient la plus faible moyenne de critères potentiellement opérationnels simultanément. Le genre majoritaire de chacun de ces emprunts est variable¹³¹, ce qui selon nous ne fait que démontrer que les emprunts variant en genre ne sont pas en mesure d'être expliqués ou justifiés globalement par nos critères simultanément, mais uniquement par plusieurs critères pour chaque genre séparément. Ces quatre emprunts semblent être justifiés par trois critères : finale graphique (*minestrone* et *party*), syllababilité entière (*gig*) et IP par hyperonymie (*rave*).

¹²⁹ L'écart de critères PO pour les emprunts à l'anglais se situe entre 1 et 9; pour les emprunts au japonais entre 2 et 10; pour les emprunts à l'arabe entre 3 et 9; pour les emprunts à l'italien entre 1 et 8.

¹³⁰ Ces critères sont l'initiale vocalique, l'IP par ellipse d'un signe entier, la paronymie, la similitude graphique, la syllababilité entière, les réanalyses morphologiques graphique et phonétique.

¹³¹ En fait, le genre majoritaire de *party* est masculin, mais nous éprouvons quelques réserves quant à la validité de ce résultat, modelé par notre méthodologie. Selon nos observations, *party* porte la MAM au Québec et la MAF en Europe.

Nous croyons que notre matrice est un excellent outil pour mettre en évidence l'interactivité des critères pouvant justifier le genre des emprunts non variables. Cependant, les emprunts variant en genre devraient être évalués en fonction des deux genres spécifiques et non pas d'un « troisième genre » majoritaire variable. Ainsi donc, observer quels critères sont opératoires pour quel genre permettrait probablement de mettre en évidence des régularités pour certains groupes d'emprunts ou certaines communautés linguistiques. Il est possible que certains groupes d'emprunts (ou communautés linguistiques) valorisent davantage certains processus. Par exemple, le critère de la syllabité entière semble très efficace avec les emprunts à l'anglais (19 cas sur 25) et très peu avec les emprunts à l'arabe (3/25). En revanche les emprunts à l'italien ou au japonais ne sont aucunement applicables au critère, car aucun n'est monosyllabique.

Il existe probablement un emprunt (sans doute de genre variable) hors de notre échantillon qui ne serait justifié par aucun critère, un emprunt qu'il serait fort intéressant d'étudier. En ce qui a trait aux unités que nous étudions, toutes sont potentiellement justifiées par au moins un critère. Nous avons exclu l'archigenre de nos 15 critères matriciels, car la conceptualisation d'un emprunt ne s'évalue pas aussi objectivement qu'une finale phonétique, voire qu'un paradigme sémantique¹³². Il reste que l'archigenre pourrait justifier le genre d'une majorité des unités que nous étudions.

5.4 Paradigmes d'unités lexicales

Dans cette section, nous aborderons quelques paradigmes intéressants, autant pour l'aspect sémantique que pour l'aspect physique.

5.4.1 Types de voiles et couvre-chefs

Le premier paradigme examiné regroupe les noms des voiles et couvre-chefs d'origine arabe. Celui-ci est composé des unités à l'étude *hidjab*, *burqa*¹³³ et *niqab*, ainsi que des unités *tchador* et *chèche*. Toutes les unités, sauf *burqa*, portent de manière générale la marque du masculin. Mis à part la *burqa*, tous les voiles permettent de voir une partie de l'identité de l'individu. Le *hidjab*, le *chèche* et le *tchador* permettent de voir le visage entier ou partiel; le *niqab* ne laisse entrevoir que les yeux. À l'inverse, la *burqa* ne laisse rien paraître de l'identité de l'individu : elle couvre d'un grillage les yeux. Cette occultation comporte une intériorité qui sépare la représentation de *burqa* de celles des autres voiles. En cela *burqa* devient conçu comme un type de robe ou de vêtement et non plus comme un voile. Le paradigme des robes et autres vêtements longs et

¹³² Pour notre commentaire sur le calcul d'opérationnalité de l'archigenre, il en sera question dans la section 6.1.5.

¹³³ Précisons que si la *burqa* est définie dans Antidote comme un « vêtement féminin traditionnel porté dans certains pays musulmans, couvrant le corps de la tête aux pieds en ne laissant qu'une petite ouverture grillagée à la hauteur des yeux », le Robert (2011) la décrit comme un « voile épais, ajouré à hauteur des yeux qui, dans certains pays musulmans, couvre intégralement le corps des femmes jusqu'aux pieds. » La réalité est parfois conçue comme un voile et parfois comme un vêtement, et cette ambiguïté pourrait influencer le genre assigné.

amples couvrant la majorité du corps est majoritairement féminin. Celui-ci est composé de *robe*, *tunique* et *toge*. Se basant sur une série de concept différent, on comprend donc mieux la motivation du genre de *burqa*.

Illustrons avec quelques cas *niqab*, *burqa* et *chèche* :

- « Croyez-vous que la dame au **niqab** qui s'est fait expulser d'un cours de français voulait s'intégrer? » - Anonyme, sans date (corpus québécois informel)
- « Au Royaume-Uni, plusieurs personnalités politiques ont exprimé l'idée d'interdire le **niqab**. Shocking, au pays du multiculturalisme » - Eric Albert (*Le Monde*), 2013-09-28
- « Selon elle, le voile est plutôt un message. Un message de soumission, un message politique qui a cours depuis la révolution iranienne de 1979. Un recul. Dans les pays musulmans, des femmes sont violées parce qu'elles refusent de le porter, rappelle-t-elle. "Ce voile et cette **burqa** qu'on défend ici sont entachés du sang de toutes les adolescentes et les femmes qui ont voulu dire non." » - Marie Claude Lortie (*La Presse*), 2013-02-15
- « « Rappelez-vous ce Français en 2012 qui avait pris le train pour aller poignarder deux policiers à Bruxelles le jour de l'annonce de l'interdiction de la **burqa** par la Belgique », illustre le policier. Et il est tout aussi compliqué, ajoute-t-il, d'anticiper la réaction « des combattants aguerris, plus âgés et plus structurés quand ils reviendront de Syrie ». » - Jacques Follorou (*Le Monde*), 2013-05-27
- « Et comme parfois l'habit fait le moine, les membres n'hésitent pas à sortir les tenues de circonstance. Le sarouel et le **chèche** constituent l'uniforme du parfait jockey. « Nos tenues sont quand même très francisées. Par exemple, nous avons des vêtements rouges, une couleur que les Arabes ne portent pas », indique Olivier. » - Hélène Guinhut (*Sud Ouest*), 2011-08-03
- « Les fins de semaine, M. Atigh, portant un magnifique **chèche** sur la tête, sert des spécialités de son pays: le maafé, riz au boeuf, carottes, pommes de terre, laurier et sauce aux arachides, ou le thieboudienne, plat national à base de riz cassé et de poisson, relevé d'une variété de gingembre mauritanien très parfumé et très acidulé. » - Jean-Philippe Tastet (*Le Devoir*), 2003-05-16

5.4.2 Types de pâtes

Le second paradigme examiné regroupe les noms des pâtes alimentaires. Ce paradigme est composé de sous-paradigmes selon la manière dont on se représente les pâtes. Nous avons séparé les paradigmes selon la forme canonique qu'ils représentent et selon leur fonction. Nous comptons cinq catégories, comportant les unités suivantes :

1. Pâtes désignant un tube : *macaroni*, *penne*, *rigatoni*, *cannelloni*.
2. Pâtes plates : *linguine/linguini*, *fettuccine/fettuccini*, *lasagne*, *tagliatelle*.
3. Pâtes longues non plates non creuses : *spaghetti*, *spaghettini*, *capelli (d'angelo)*, *capellini*, *vermicelle*, *fusilli*, *gemelli*.
4. Pâtes courtes et à formes spécifiques : *farfalle*, *orecchiette*, *gnocchi*, *ruote/rotelle*, *cavatelli*.
5. Pâtes farcies : *tortellini*, *ravioli*, *dumpling*, *won ton*, *cannelloni*, *gnocchi*.

Les sous-paradigmes numéro un, trois et cinq ne montrent aucune unité portant la marque du féminin, alors que les sous-paradigmes deux et quatre montrent un bon nombre d'unités portant les deux marques. Le sous-paradigme des pâtes plates (le numéro deux) comporte essentiellement des noms féminins (*tagliatelle*,

lasagne), ainsi que des noms à genre variable (*linguinellinguini*, *fettuccine/fettuccini*). De même le quatrième sous-paradigme comporte de nombreuses unités portant majoritairement la marque du féminin, telles que *farfalle*, *orecchiette* et *ruotel/rotelle*. Le paradigme comporte également deux autres unités portant la marque du masculin : *cavatelli* et *gnocchi*. La majorité du quatrième paradigme porte la marque du féminin. Illustrons les cas de *cannelloni* et de *tagliatelle*, ceux de *farfalle* et de spaghetti ayant été illustrés plus haut dans ce chapitre :

- « en mettant le **cannelloni** debout dans la garniture tout en la faisant pénétrer à l'intérieur » - Anonyme, sans date (corpus québécois informel)
- « 5mn de tapis dont : les premiers 400m en marche rapide à 6km/h, puis 2000m à 9 km/h, puis 200m marche à 6km/h, 200m course à 8km/h ça deux fois, donc 800m, puis alternance 100m à 6km/h et 100m à 8km/h, 4 fois, avec une inclinaison de 8, et bien ça fait bien mal!!! puis muscu, donc étirage nuque, les dorsaux et les bras : vous voyez le truc ki remue sous le bras kan on fait coucou et bien ça, ça bouge preske pu donc 4 séries de 15 à chaque appareil avec 1mn de temps de pause entre chaque séries voilà pour moi kakett dis à ta mère de plaquer ta balance et de ne la sortir ke ;;;;;;;;;;;;;;e jeudi et le lundi, ça te fera au moins deux fois par semaine Ahh oui ça fait vraiment de bien super pour toi. est ce que vraiment faut pas se peser plusieurs fois par semaine? moi, je me pèse tous les matins ...parfois les après-midi aussi ça me motive lol j'ai avalé, tout à l'heure vers 18:30, un grand verre d'eau+ forlax + deux pruneaux sinon pour ce soir ça sera :** ma soupe d'hier** juste goûter ma **cannelloni** farcie à la viande hachée gratinée au four.** peut être un fromage blanc en plus » - ammoula12 (*Doctissimo*), 2010-04-07
- « bon ben moi tout compte fait je me suis rien pris pour le gouter, pas faim. Il me reste 9 points pour ce soir. Je fais le blé au **tagliatelle** de courgettes et poivrons. Mais c'est écrit 1,5 pt par personne, ca me parait peu, y' a 2cs d'huile et le blé. Je vais compter un peu plus, déjà je vais faire 100gr de blé chacun donc ca fait déjà 2 points pour le blé. Je m'autoriserais peut être un petit dessert » - Profil supprimé (*Doctissimo*), 2006-04-06
- « Ce soir osso bucco de veau à la sauce moutarde et à l'estragon avec des **tagliatelle** fraiches au parmesan et gorgonzola » - sabrinadiamond (*Doctissimo*), 2011-11-17

Qu'est-ce qui peut motiver le fait que ces types de pâtes soient dénommés davantage par des noms portant la marque du féminin que l'inverse? La conservation du genre et l'archigenre ont un grand rôle à jouer. La plupart des emprunts à l'italien ont conservé leur genre d'origine. Il est intéressant de constater de *fettuccine* et *linguine*, d'abord féminins, ont été masculinisés graphiquement par un <i>, phonétiquement par un [j] et par le biais de la valeur générique de l'archigenre. Une caractéristique typique des noms de pâtes est qu'ils décrivent souvent le type de pâte, ainsi que le plat entier qu'elles composent. Cette dernière conceptualisation est plus générique, plus large que celle des unités composant le plat. En conséquence, plus la conceptualisation est générique, plus l'archigenre est présent et plus il impose la marque du masculin. Il est donc fort probable que les noms de pâtes en général soient masculins par l'influence de l'archigenre et que ceux qui portent la marque du féminin le soit en vertu de leur genre originel en italien.

5.4.3 Party et les événements festifs

Pour ce qui est du troisième paradigme, nous devons d'abord expliquer le cas particulier de l'unité *party* afin de mieux comprendre la composition du paradigme. *Party* est majoritairement masculin dans les corpus, mais

nous remettons en doute ce résultat, puisque de notre connaissance, l'unité est principalement utilisée au masculin au Québec et au féminin en Europe. Les résultats de tous les corpus semblent concorder avec la réalité, sauf dans le cas du corpus européen informel (*Doctissimo*). En effet, puisque nous avons exclu toutes les combinaisons *nom + party* (comme *surprise party*)¹³⁴, combinaisons qui portaient le plus souvent la marque du féminin, nous avons conservé uniquement les cas où *party* est seul ou inscrit dans un syntagme prépositionnel (*party de Noël*). De plus, avec 76,53% de masculin dans le corpus européen informel, nous avons observé quels étaient les cas conservés et nous nous sommes rendu compte que la majorité des occurrences étaient produites par des Québécois sur le forum européen. On peut identifier l'origine d'un commentaire à partir de du nom de l'utilisateur ou du vocabulaire utilisé, comme dans l'expression « méchant party » ou « faire le party ». Voici quelques exemples d'occurrences produites par des Québécois où nous avons identifié en gras des « indicateurs dialectaux » (Courbon 2015b) :

- « Le premier Noël d'Olivier a été féérique et abondant. Le 24 chez ma mère...Olivier a **jaser** et c promener partout jusqu'a 10h pas moyen de le coucher il déconnaît dans le lit et voulais rien raté de la soirée. Le 25 Brunch a midi chez mes beau parent (père de Luc) il a dormi apres 2h30 de déballage de cadeaux pour 8 personnes ensuite il a dormit pendant qu'on préparait la soirée. Il s'est recouché à 9h le soir, il **a fait le party** à la table a nous faire faire pipi dans notre culotte....THE SHOW. Il n'a jamais **chigné**, pleurer ou montré un moindre petit mécontentement j'ai été complètement J'ai un bébé A.D.O.R.A.B.L.E !!!!! et parfois Ce soir JE recois cette fois ma Belle-Mère avec son **chum** sa fille et le **chum** de sa fille plus mon beau frère.... encore de la **maudite dinde** car elle en a pas manger encore donc je passais vous faire coucou mais je repasseraï pas avant demain car j'ai de la popotte a faire. Sur ce, je vous souhaite une merveilleuse journée ensoleillé rempli d'amour » – suzy-q (*Doctissimo*), 2005-12-27
- « **ouin !!! gros party** qui change l'ambiance de ce post » – lakebekoise (*Doctissimo*), 2009-05-09
- « suzy-q a écrit : Pres de 4+h aller et 3+ pour revieudre. on sera 9 pour lui faire l'**intertaining**, au pire on le **switch** dans la soute a bagage...de toute facon **y'aime ca** fouiller et sortir les trucs des boites, sac, armoires...il sera heureux comme un pape **Ouin**, y'aurait de quoi **faire un méchant party** dans la soute à bagage! Imagine la face du monde rendu à destination, messieurs, mesdames, retrouvez vos affaires **astheure!!** » – kiki251 (*Doctissimo*), 2006-11-02

Également à remarquer est la graphie <party> pour l'unité *partie*, telle que dans l'expression « être de la partie », comme dans l'exemple suivant :

- « Cela est complexe en effet et je tiens à te dire que j'admire ta capacité à te remettre en question. Tu **seras** donc **de la "party"** ce week-end ! Bravo. Pour le reste, je me souviens que dans mes années "ado", j'aimais aussi tout comprendre, et que les silences planant au dessus de ma tête ne m'ont jamais trop "plu" et me rendaient infernal... » – Profil supprimé (*Doctissimo*), 2011-08-26

¹³⁴ Ce type de combinaison est extrêmement productif en Europe et relève de la syntaxe de l'anglais. Au Québec, la formule la plus fréquente est *party + de + nom*. Voici une liste relevant les noms précédant *party* les plus fréquents dans le corpus informel européen : *barbecue, beach, bloc, chausson, chinois, confetti, couscous, crêpe, engueulade, fiesta, garden, gastro, grippe, hamburger, house, kébab, mammoth, ménage, merguez, microbe, murder, paella, piscine, pizza, pool, pyjama, raclette, rave, salade, sexe, surprise, sushi, tea, tisane, troc, varicelle, vomis, yoyo*.

Le paradigme qu'intègre *party* est principalement masculin (*karaoké, méchoui, slam, jam*, mais *fête*) ou variable (*gig, rave*). Nous avons ajouté *jam-session* au paradigme, la forme non réduite de *jam* qui doit son genre féminin, pensons-nous, à *session*, unité portant déjà la marque du féminin en français. La variation du genre de *party*, masculin au Québec et féminin en Europe, est identique pour l'unité *rave*. Dans le cas de *party*, nous sommes d'avis qu'il porte la MAM au Québec parce qu'il renvoie à un type de fête très générique, petite ou grande, peu importe la thématique et doit donc son genre à sa conceptualisation plus générique (l'archigenre), alors qu'en Europe, la conceptualisation la plus générale correspond au terme de niveau de base *fête* (hyperonyme : *événement*). Pour preuve, le paradigme latent de *party* est presque uniquement composé de mots féminins (*fête, soirée, veillée, réception* et le composé masculin *cinq à sept*) et les locuteurs ont identifié comme hyperonyme les termes *fête* et *soirée*. Pour ce qui est de *rave*, le genre masculin au Québec provient également d'une conceptualisation générique de la fête à musique techno, alors que le genre féminin en Europe provient aussi d'un sous-paradigme composé de son hyperonyme *fête*. Illustrons de quelques occurrences les cas de *gig*, de *rave*, de *jam*, de *méchoui* et de *slam* :

- « "C'est la plus grosse **gig** de notre vie!" s'est exclamé le chanteur Sébastien Plante joint hier à L.A. "C'est le rêve de tous les groupes de se retrouver dans une situation comme celle-là et je suis tellement fier de n'avoir jamais lâché prise!" - Valérie Lesage (*Le Soleil*), 2010-04-27
- « Le premier **gig** d'Arcadio était au spectrum à 22 ans. 1500 personnes. Ça fait 2 ans qu'il fait ça à temps plein » - Anonyme, sans date (corpus québécois informel)
- « Deux D.J. internationaux, Deadmus 5 et Skrillex, ont participé à ce **rave**. «Réunir ces deux D.J. à Québec un 30 décembre, c'est exceptionnel», lance M. Cinq-Mars. » - Anonyme (*Le Journal de Québec*), 2011-12-31
- « Alors, outre la puissance physique, le génie de Political Mother reste surtout dans la musique. Une musique de démiurge, signée Hofesh Shechter, Nell Catchpole et Yaron Engler, qui mêle en une trame d'une singularité envoûtante, jouée live par huit musiciens, le rock, la **rave**, à la rythmique de la tradition juive et aux sonorités de J.S Bach et de Verdi. Et qui ose finir par une chanson quasi fleur bleue de Joni Mitchell! Une fin comme un pied de nez à tant de noirceur. » - Aline Apostolska (*La Presse*), 2012-11-05
- « Oran Etkin n'est évidemment pas seul dans ce Rallye liégeois au fil de l'eau. On peut aussi entendre les groupes Big Noise (plus dixieland), Kind of Pink (qui réinterprète les musiques de Pink Floyd), le duo Pascal Mohy (le beau pianiste) et Mélanie De Biasio (la sublime chanteuse), Cruz Control (et leur musique fusion), Erik Vermeulen (le pianiste en solo), etc. Et ça se terminera dimanche soir, au Jacques Pelzer's Jazz Club, par une belle **jam** animée par l'OEil Collectif et le Témse Trio. On sait s'amuser sur les bords de la Meuse... » - Jean-Claude Vantroyen (*Le Soir*), 2013-08-22
- « Shawinigan - (IBB) - C'est aujourd'hui que commence la 10e édition du Festival Énergie Country de la Mauricie, avec le grand **jam** des musiciens amateurs. » - Anonyme (*Le Nouvelliste*), 2010-08-04
- « Pilier de Maranello, le siège de la marque, Todt amorce son départ en 2007 sur deux nouveaux sacres, sentant peut-être le vent tourner. « À chaque course, c'était la fin du monde si on ne gagnait pas. On ne pouvait faire que moins bien. J'ai préféré choisir plutôt qu'on choisisse pour moi, car on doit toujours rendre des comptes à quelqu'un » , explique celui qui rêve de sillonner l'Afrique avec ses amis et anciens pilotes, Jacky Ickx ou Georges Groine. « Pour bivouaquer et faire un bon **méchoui** » , précise-t-il avant d'avouer : « Je manque de temps. C'est là mon vrai problème. » » - Gilles Festor (*Le Figaro*), 2013-12-10
- « Celles-ci possèdent un savoir-faire ancestral qu'elles transmettent de mère en fille. L'homme est maître à l'extérieur. C'est lui qui s'occupe du **méchoui**, des brochettes sur le gril et prépare le rituel du thé à la menthe. Les épices doivent être achetées en petite quantité, car elles s'éventent rapidement et perdent les

- subtilités de leur parfum. Il faut les mettre dans des boîtes hermétiques, à l'abri de la lumière. » - Daniel Vézina (*La Presse*), 2004-01-30
- « La plus belle des trois c'est Noémy qui danse avec la mort. La mort de son papa qui n'est pas mort. Il est seulement junkie. Noémy parle à son père dans son **slam** à pleurer, à la fin tu veux la prendre dans tes bras, là là petite fille, et la première chose que tu sais c'est pas elle qui pleure, c'est toi. » - Pierre Foglia (*La Presse*), 2011-11-12
 - « Comment que ce **slam** me fout le blues » - rouge amour (*Doctissimo*), 2012-05-25

5.4.4 Types de saucissons

Le quatrième paradigme, celui des types de saucissons, est composé des unités à l'étude *pancetta*, *prosciutto* et *salami* auxquelles nous avons ajouté les unités lexicales *saucisson*, *pepperoni*, *mortadelle*, *chorizo* et *andouille*. De ce paradigme, seules les unités *pancetta*, *mortadelle* et *andouille* portent majoritairement la marque du féminin. Cela pourrait s'expliquer partiellement par la finale morphologique des unités associée au féminin : *-a* (en italien¹³⁵), *-elle* (*pelle*, *poubelle*, *tutelle*) et *-ouille* (*fouille*, *couille*, *trouille*). De même, nous percevons une différence sémantique entre les noms portant la MAF et ceux portant la MAM. *Prosciutto*, *salami*, *saucisson*, *pepperoni* et *chorizo* semblent faire référence à des charcuteries sèches, alors que *pancetta*, *mortadelle* et *andouille* semblent référence à des charcuteries « non sèches », généralement moins aromatisée¹³⁶. Notons que la majorité des termes (*prosciutto*, *salami*, *pepperoni* et *chorizo*) suivent le genre de l'hyperonyme *saucisson*. Illustrons de quelques occurrences les cas de *pancetta*, de *prosciutto* et de *salami* :

- « Autre petit reproche d'un repas tout de même très satisfaisant, la côte de porc avait beau être lardée au **pancetta**, la chair finissait par être un peu sèche. Une cuisson quelques minutes trop longues. » - Pierre Jury (*Le Droit*), 2013-09-06
- « Pour une bolo améliorée : ajoutez de la **pancetta** hachée bien caramélisée à la viande, beaucoup d'oignons fondus, une conserve de tomates cerises et une poêlée de chanterelles. » - Anonyme (*Le Soir*), 2013-11-16
- « Depuis des lustres, les Italiens tout comme les Français ont marié par instinct le melon brodé au **prosciutto** et au porto. Devenant l'un des antipasti les plus connus, servi partout dans le monde. Ce qui est inspirant dans la démarche harmonique (moléculaire) entre les aliments, c'est de pouvoir quelques fois faire dévier la piste en y ajoutant un ingrédient qui deviendra LA nouvelle piste à suivre. » - François Chartier (*La Presse*), 2011-09-17
- « Stefano intègre du **prosciutto** cotto dans son pain, mais du jambon pourrait aussi faire l'affaire, tout comme du **prosciutto** ordinaire ou même de la pancetta (dans ce cas, portez attention au sel). Une autre variation serait de préparer ce pain avec des tomates séchées, des olives ou des épinards (préalablement blanchis, égouttés et hachés finement). » - Stephano Faita (*Le Journal de Québec*), 2012-04-07
- « Yoshette85 a écrit : Petite question : est ce qu'on a le droit au **salami** et au bacon le soir ??? j'en ai ds le frigo j'ai lu seulement à midi la charcuterie » - nicoty42 (*Doctissimo*), 2012-11-12
- « Malgré cela, l'UDC, qui se profile comme le parti des loyalistes bernois, se dit «pleinement confiante dans la population et s'engagera sans concession afin de mettre un terme à la tactique du **salami** perfide des séparatistes». » - Serge Jubin (*Le Temps*), 2013-02-01

¹³⁵ Rappelons que *mortadelle* porte la finale *-a* en italien.

¹³⁶ Notre expérience.

5.4.5 Substituts protéinés

Le cinquième et dernier paradigme que nous allons regarder est celui des substituts protéinés. Il est composé des unités *tofu*, *surimi*, *seitan* et *miso* auxquelles nous ajoutons les unités *goberge*, *Soylent*, *tempeh* et *gluten*. Ce paradigme comporte un genre très stable en ce que la quasi-totalité des signes le composant sont du même genre. Sauf *goberge* dont le genre varie au Québec, le paradigme est entièrement au masculin. Cela relève d'une conceptualisation vague et floue, celle de l'archigenre, en ce que c'est « quelque chose » qui remplace la viande, mais il n'est pas clair de quoi la substance est faite. En ce qui a trait au mot *goberge*, au Québec, le féminin est surtout utilisé dans des contextes formels culinaires ou pour référer au poisson (vivant), alors que le masculin apparaît davantage dans les contextes informels. Nous sommes vis que les locuteurs utilisent *goberge* au masculin afin d'intégrer l'unité au paradigme, et afin d'uniformiser le genre. Nous avons préalablement illustré quelques cas des unités *surimi* et *seitan*; illustrons maintenant *tofu* et *miso* :

- « Dans les cantines des écoles auscitaines, on mange des lasagnes au **tofu**, des lentilles de Lavardens et des yaourts au lait bio. C'est une volonté politique de la municipalité relayée sur le terrain par l'envie de l'équipe en cuisine de travailler des produits « sains et bons ». » - Gaëlle Richard (*Sud Ouest*), 2012-06-02
- « dukeerc a écrit : l'origine du **tofu** est bien chinoise , mis il faut savoir que le japon a adopté ou s'est inspiré de tas de chose chinoise , sais-tu par exemple que le saké est encore une invention chinoise ? , cela s'explique par le fait que la chine a cultivé le riz bien avant le japon, d'ou le fait peut être que le saké chinois est bien plus fort que le japonais (une 50 éne de degré contre une 15 éne) Vous êtes sûrs que le saké se boit tiède ? » - laeti555 (*Doctissimo*), 2006-04-21
- « L'autre, à 800 yens (7,80 euros), change tous les jours. Rien de révolutionnaire dans l'assiette par rapport au menu des foyers japonais : une soupe au **miso** (pâte de soja), un bol de riz, un plat principal et quelques assortiments. » - Philippe Mesmer (*Le Monde*), 2012-02-18
- « - Diluez le **miso** avec un peu de bouillon dashi puis ajoutez-le au reste du bouillon avec les oignons verts coupés finement. Ne faites jamais bouillir le **miso**, rajoutez-le toujours à la fin. Il est en effet très sensible à la chaleur et a donc tendance à perdre en saveur et en texture. » - Anonyme (*Le Soir*), 2011-03-12

5.5 Bilan

Le chapitre cinq a exposé les résultats de nos analyses. D'abord, nous avons montré que le genre n'est pas attribué dans les mêmes proportions selon les emprunts des langues à l'étude, selon le niveau de formalité, ainsi que selon l'origine de l'énonciation. Par la suite, nous avons montré que la majorité des unités portaient la marque du masculin et que les emprunts des LAG avaient une plus grande chance d'être féminins que les emprunts des LSG. Ensuite, nous avons évalué dans quelle mesure chaque critère peut justifier le genre des emprunts de notre échantillon. Cela nous a révélé que tous les critères que nous avons examinés ne sont pas potentiellement opératoires au même degré (ou ne possèdent pas un taux de justification similaire). Finalement, nous avons observé quelques paradigmes comportant des similitudes et nous avons expliqué les différences qu'il est possible d'observer.

Passons maintenant à la discussion, à la comparaison des critères et au retour sur nos hypothèses.

6. Discussion et retour sur les hypothèses

Dans cette section, nous discuterons de la capacité de chaque critère examiné à justifier le genre d'emprunts, puis nous reviendrons sur nos hypothèses formulées dans l'introduction. Nous discuterons par la suite des limites de l'étude, ainsi que des causes d'erreurs inhérentes à notre méthodologie. Nous finirons par ce que l'on peut retirer de notre étude.

6.1 Retour sur les critères

Les critères qui ont été examinés sont la pierre angulaire de notre travail. La majeure partie de nos résultats et de nos conclusions se fonde sur le taux de justification de chacun¹³⁷. Il va donc de soi que nous commençons notre discussion par l'analyse du taux de justification de ceux-ci. Nous classerons d'abord les critères en ordre décroissant selon leur taux d'opérationnalité, puis nous observerons différentes classifications de critères afin de regarder quel type de critère est plus potentiellement opératoire (*PO*).

6.1.1 Taux de justification des critères

Commençons par faire une liste des critères analysés. Nous les classons ici en ordre décroissant, indiquant d'abord le taux de justification, puis, entre parenthèses, le nombre de cas dont le genre est expliqué sur le nombre total de cas analysables.

1. Conservation du genre : 84% (42/50);
2. IP par similitude conceptuelle (paradigme latent) 77,6% (76/98);
3. Syllabité entière : 68,1% (15/22);
4. Finale graphique : 61% (49/80);
5. IP par hyperonymie : 61% (61/100);
6. Initiale vocalique : 55,5% (5/9);
7. AP par phonème pénultième : 49,4% (41/83);
8. AP par rime riche : 42,9% (30/70);
9. IP par ellipse d'un signe entier : 42,4% (14/33);
10. Paronymie : 42,3% (11/26);
11. Syllabité finale graphique : 41,8% (41/98);
12. AP par phonème final : 40,7% (35/86);
13. Réanalyse graphique : 40% (18/45);
14. Similitude graphique : 38,4% (10/26);
15. Réanalyse phonétique : 22% (15/68).

¹³⁷ Rappelons-le, le taux de justification indique le pourcentage de cas où le critère peut justifier le genre d'une unité. Plus le pourcentage est haut, plus le critère peut justifier le genre d'un emprunt, et donc plus il est PO. Par exemple, le critère de la conservation du genre a un total de 50 unités analysables (car la forme correspondante de l'emprunt dans la langue donneuse comporte un genre) et le genre de 42 est identique à ce que suppose la règle (« chaque emprunt d'une langue à genre (LAG) devrait conserver son genre »); on dira donc que le taux de justification de ce critère est de 34 cas sur 50 (34/50), donc de 84%.

Au total, sur nos 15 critères, nous remarquons que seuls cinq (5) sont PO lorsqu'ils sont comparés à notre échantillon d'unités. Il va sans dire que tous les résultats des critères dépendent en large mesure de la sélection des unités, mais les unités ont été sélectionnées indépendamment des critères et devraient refléter une distribution naturelle ou usuelle d'emprunts. Le critère le plus opératoire est la conservation du genre¹³⁸ avec un taux de 84%, alors que le critère avec le plus faible taux est la réanalyse phonétique avec 22%. Aucun critère n'est en mesure de justifier le genre de l'entièreté des unités, et aucun critère n'est en mesure de justifier le genre d'aucune unité.

Rappelons que nous avons établi le taux de justification d'un critère à 60%, car il fallait que le taux soit supérieur à 50% + 1 (la majorité plus un). Il fallait également que le taux soit suffisamment haut pour ne pas être le fruit du hasard, dans le cas où un critère donnerait à 50% un genre sans autre raison sémantique ou formelle.

Rappelons également que si le critère de syllabité entière (selon lequel tous les emprunts monosyllabiques portent la marque du masculin en Europe) est considéré comme PO, il ne s'applique qu'à une partie du lexique (les monosyllabiques) et que dans une communauté (Europe). De même, le critère le plus opératoire, la conservation du genre, n'est applicable qu'avec les emprunts venant d'une LAG. On peut remarquer que la capacité à justifier le genre d'un emprunt est dépendante de l'échantillon.

6.1.2 Première classification : forme physique *versus* signification(s)

Nous avons séparé les critères selon qu'ils se basent sur la forme physique ou la (ou les) signification(s) des unités. Le critère de l'ellipse se base sur les deux plans, mais nous l'avons considéré comme un procédé d'abord sémantique.

Les critères de forme physique, en ordre décroissant de taux de justification, sont les suivants :

1. Syllabité entière : 68,1% (15/22);
2. Finale graphique : 61% (49/80);
3. Initiale vocalique : 55,5% (5/9);
4. AP par phonème pénultième : 49,4% (41/83);
5. AP par rime riche : 42,9% (30/70);
6. Paronymie : 42,3% (11/26);
7. Syllabité finale graphique : 41,8% (41/98);
8. AP par phonème final : 41,2% (35/86);
9. Réanalyse graphique : 40% (18/45);

¹³⁸ Ce qui rejoint les observations de Stolz (2009), comme étant un critère plus efficace que les critères de forme physique ou de signification relevés. Il est évident que ce résultat ne s'applique qu'à 50 des unités que nous étudions, c'est-à-dire celles venant d'une langue à genre.

10. Similitude graphique : 38,4% (10/26);
11. Réanalyse phonétique : 22% (15/68).

Les critères se basant sur la signification de l'unité, dans le même ordre, sont les suivants :

1. IP par similitude conceptuelle (paradigme latent) 77,6% (76/98);
2. IP par hyperonymie : 61% (61/100);
3. IP par ellipse d'un signe entier : 42,4% (14/33).

Nous remarquons du premier coup d'œil que seuls deux critères de forme physique, sur un total de 11 (18,2%), sont PO (syllabité entière et finale graphique), alors que dans le cas des critères sémantiques, deux critères sur trois (66,6%; IP par similitude conceptuelle et par hyperonymie) sont PO. Cela pourrait nous mener à conclure que les critères sémantiques sont opératoires à 66% contre 18,2% pour les critères de forme physique. Établissons ici une distinction entre « l'opérationnalité » réelle et « l'opérationnalité » potentielle, qui sont toutes deux applicables dans les cas de calculs de taux de justification et de comparaisons des critères.

6.1.2.1 Opérationnalité réelle et opérationnalité potentielle

Dans le cas du calcul du taux de justification, l'opérationnalité réelle (ou effective) calcule le nombre de concordances sur le nombre de cas analysables, alors que l'opérationnalité potentielle calcule le nombre de concordances sur la somme des nombres de cas analysables et inanalysables (100 pour notre échantillon). Nos taux de justification sont donc calculés en fonction de l'opérationnalité réelle de chaque critère, en laissant de côté les cas variables dans le calcul. Si nous devons évaluer l'opérationnalité potentielle, un critère comme la syllabité entière passerait d'un taux de justification de 68,1% (15 concordances sur 22 cas analysables) à 15% (15 concordances sur 100 cas analysables et non analysables), ce qui nous semble très peu représentatif de la réalité. Notre étude se concentre sur l'opérationnalité réelle de chaque critère, d'où la raison pour laquelle certains critères se basent sur une poignée de cas analysables.

Dans le cas de la comparaison des critères, l'opérationnalité réelle (ou effective) compare seulement le nombre de critères au taux de justification supérieur à 60% (jugés PO) d'une catégorie au nombre de critères PO d'une autre catégorie. L'opérationnalité potentielle calcule (sous forme de pourcentage) le nombre de critères PO sur le nombre de critères total de la catégorie (ici sémantique *versus* forme physique). Dans le premier cas, on compare deux critères dans chaque catégorie et on peut déduire que les deux catégories sont *effectivement, réellement* opératoires au même degré, ayant chacune deux critères opératoires. Dans le second cas, on compare 66,6% (2 critères PO sur un total de 3 critères sémantiques) d'opérationnalité pour les critères sémantiques à 18,2% (2 critères PO sur un total de 11 critères de forme physique) d'opérationnalité pour les critères de forme physique; de ces pourcentages, on pourrait tirer la conclusion que

les critères sémantiques sont beaucoup plus opératoires que les critères de forme physique. Nous avons choisi de calculer l'opérationnalité réelle de chaque catégorie de critères, car il nous semble plus représentatif d'évaluer le résultat que la possibilité de résultats. De plus, si on peut trouver une dizaine de manières (ou davantage) d'évaluer la forme physique d'une unité et d'en justifier (potentiellement) le genre d'un emprunt – car la forme physique est concrète et tangible, donc plus facilement observable et calculable que l'aspect sémantique d'une unité–, on ne peut trouver plus de quelques paires de critères qui pourraient justifier le genre d'un emprunt.

Considérant que le concept d'opérationnalité implique entre autres de la potentialité (dans sa capacité à pouvoir justifier le genre d'un emprunt), nous sommes face à un dilemme. Devrions-nous considérer l'opérationnalité dans sa dimension effective ou potentielle? Est-il légitime de notre part de ne considérer que l'opérationnalité effective ou réelle en ignorant l'opérationnalité potentielle? Nous sommes d'avis que pour évaluer à quel point un critère peut être efficace, il faut utiliser des données (effectives) et faire ressortir le potentiel à travers son actualisation, le réel. C'est pourquoi, afin de vérifier à quel point les critères peuvent justifier le genre d'un emprunt, il nous faut regarder son actualisation, et donc ne considérer que l'opérationnalité effective ou réelle des critères et des groupes de critères.

Puisqu'il y a autant de critères sémantiques ou de forme physique qui sont PO (deux et deux, respectivement), nous pouvons raisonnablement affirmer que, avec notre échantillon de 100 unités et notre méthodologie, les critères sémantiques peuvent justifier dans autant de cas le genre d'un emprunt que les critères de forme physique, c'est-à-dire que leur efficacité réelle est identique (deux pour deux), que les deux groupes sont autant potentiellement opératoires l'un que l'autre.

6.1.3 Deuxième classification : critères phonétiques *versus* critères graphiques

Notre deuxième classification se concentre sur les critères de forme physique et elle oppose la forme phonétique à la forme graphique. Voici, en ordre décroissant du taux de justification, les critères se basant sur la forme phonétique :

1. Syllabacité entière : 68,1% (15/22)¹³⁹;
2. Initiale vocalique : 55,5% (5/9);
3. AP par phonème pénultième : 49,4% (41/83);
4. AP par rime riche : 42,9% (30/70);
5. Paronymie : 42,3 (11/26);
6. AP par phonème final : 41,2% (35/86);
7. Réanalyse phonétique : 22% (15/68).

¹³⁹ Rappelons que le critère de la syllabacité entière, le troisième critère le plus efficace sur les 15 au total, ne s'applique qu'à une petite portion du lexique et dans une seule communauté linguistique.

Voici, dans le même ordre, les critères se basant sur la forme graphique :

1. Finale graphique : 61% (49/80);
2. Syllabité finale graphique : 41,8% (41/98);
3. Réanalyse graphique : 40% (18/45);
4. Similitude graphique : 38,4% (10/26).

Des onze (11) critères de forme physique, un seul de chaque type est PO, ce qui correspond à un taux de justification moyen de 14,2% pour les critères se basant sur la forme phonétique et à 25% pour les critères se basant sur la forme graphique. Puisqu'un seul critère de chaque groupe (phonétique et graphique) peut être considéré comme opératoire, nous pouvons raisonnablement déduire que les critères se basant sur la forme phonétique peuvent justifier le genre d'un emprunt au même titre que ceux se basant sur la forme graphique.

6.1.4 Autres constats

Il est intéressant de regarder une autre sous-catégorie de critères se basant sur la forme, celle de la réanalyse morphologique. Ces critères sont tous deux parmi les trois critères les moins opératoires. Nous tenons cependant à préciser que ce sont les critères avec le plus de cas variables, avec 48 cas variables pour le critère de réanalyse phonétique et 25 pour le critère de réanalyse graphique. Si l'on devait exclure du total analysable les cas variables, on obtiendrait respectivement 15/20 (75%, avant : 22%) et 18/20 (90%, avant : 40%), un taux de justification très élevé, mais qui devrait être considéré avec certaines précautions. D'abord, mis à part le critère de l'initiale vocalique, ces critères seraient ceux avec le moins de cas analysables. Ensuite, exclure des cas analysables à résultats neutres (ou variables) engendrerait la même procédure avec les autres critères, afin de conserver la même rigueur méthodologique¹⁴⁰. Finalement, une telle modification méthodologique devrait entraîner la révision des taux de justification vers un seuil plus haut, plus strict, car le seuil de plusieurs critères aurait augmenté et aucun n'aurait diminué. Une augmentation d'un taux de justification de 50% (comme dans le cas de la réanalyse graphique) ou plus remettrait en question entièrement la pertinence d'un tel calcul.

Finalement, en ce qui a trait au nombre de cas analysables par critère (ou facteur)¹⁴¹, seuls huit¹⁴² sur les quinze peuvent être applicables à un minimum de 60 emprunts analysables (et analysées). Sur ces facteurs,

¹⁴⁰ Les autres critères avec taux de justification modifiés (et nécessairement augmentés) : AP par phonème final (de 41,2% à 43,5% [37/85]), AP par phonème pénultième (de 49,4% à 51,9% [41/79]), AP par rime riche (de 42,9% à 44,1% [30/68]), IP par similitude conceptuelle (de 77,6% à 78,4 % [76/97]) et ellipse (de 42,4% à 63,6% [14/22]).

¹⁴¹ Nous utiliserons ici le terme *facteur* qui correspond au terme *critère d'analyse* que nous avons utilisé depuis le début dans un cadre d'analyse. En dehors de ce cadre, les critères sont en fait des facteurs, étant donné qu'ils se situent sur le plan opératif (celui de l'actualisation) et non seulement sur celui de l'analyse.

deux sont sémantiques, trois se basent sur la forme graphique et trois sur la forme phonétique. Nous appuyant sur ces résultats, nous pourrions affirmer que les facteurs sémantiques sont moins applicables à une majorité d'unités (deux facteurs sur huit étant applicables) que ceux de forme physique (six critères sur huit). Cependant, nous réfutons cela en précisant que les deux facteurs étant applicables à une majorité d'unités sur trois facteurs sémantiques représentent 66% de leur groupe, alors que les six facteurs étant applicables à une majorité d'unités sur onze facteurs de forme physique représentent 54,5% de leur groupe. Dans le cas présent, il faut regarder le potentiel et déduire qu'il y a plus de facteurs sémantiques ou se basant sur la forme graphique qui sont applicables à 60 d'unités ou plus (respectivement 66,6% et 75% des facteurs) que de facteurs se basant sur la forme phonétique (42,8%) qui sont applicables à 60 unités ou plus.

Nous concluons que, à taux de justification équivalents ou similaires, plus un facteur est applicable à un nombre élevé d'unités, plus il est influent. Donc, un critère PO à un nombre de cas élevé est plus intéressant à analyser qu'un critère à opérationnalité similaire mais à un nombre de cas moindre. De plus, les facteurs sémantiques sont généralement applicables à un plus grand nombre d'unités que les facteurs de forme physique. Par exemple, les critères d'IP par similitude conceptuelle ou par hyperonymie concernent respectivement 98 et 100 cas (car toutes les unités ont une signification et un hyperonyme), alors que dans le cas des critères de phonèmes finaux et pénultièmes, bien que toutes les unités aient au moins deux phonèmes, le nombre de cas analysable se chiffre respectivement à 86 et 83 cas¹⁴³. C'est ainsi qu'un critère comme la syllababilité entière, PO avec un taux de 68,1% et ne se basant que sur 22 cas, est moins intéressant à recommander à l'étude que le critère de la syllababilité finale graphique, FO à 41,8% mais se basant sur 98 cas.

6.1.5 Discussion

Quel type de critère peut le mieux justifier le genre d'emprunts? Avant de répondre à cette question, rappelons que la plupart des critères que nous avons trouvés dans la littérature (10/15) n'ont pas été jugés PO selon notre classement. Sur les cinq critères PO, deux sont des critères sémantiques et deux sont des critères de forme physique. Le cinquième critère, et le plus opératoire selon nos résultats, ne relève ni de la signification, ni de la forme physique de l'emprunt, mais du transfert, du processus lui-même qu'est l'emprunt. Il nous semble donc juste, de ce point de vue, que les critères de forme physique et de signification(s) soient tout

¹⁴² Ont plus de 60 cas analysables : IP par similitude conceptuelle (98 cas), finale graphique (80 cas), IP par hyperonymie (100 cas), AP par phonème final (86 cas) et pénultième (83 cas), AP par rime riche (70 cas), syllababilité finale graphique (98 cas), réanalyse morphologique phonétique (68 cas).

Ont moins de 60 cas analysables : conservation du genre (50 cas), réanalyse morphologique graphique (45 cas), syllababilité entière (22 cas), initiale vocalique (9 cas), IP par ellipse d'un signe entier (33 cas), paronymie (11 cas), similitude graphique (26 cas).

¹⁴³ Cela s'explique par le fait que nos résultats se basent sur les études de Tucker *et al.* (1977) et que certains phonèmes sont considérés comme « ambigus », donc inanalysables.

autant susceptibles de justifier le genre des emprunts. Comme nous l'avons mentionné plus haut dans la section 6.1.2, les critères sémantiques et ceux de forme physique influencent le genre des emprunts dans une mesure équivalente. Ensuite, des deux critères de forme physique, on retrouve un critère se basant sur la forme phonétique et un autre sur la forme graphique. C'est également le constat que nous avons relevé en 6.1.3.

En ce qui a trait aux critères jugés FO, ils ne peuvent pas nous permettre de tirer une quelconque conclusion, car calculer un taux d'opérationnalité pour chaque catégorie de critère impliquerait que nous calculons l'opérationnalité potentielle du groupe, et non pas son opérationnalité réelle ou effective. Rappelons que nous avons fait le choix dans la section 6.1.2 de n'évaluer que l'opérationnalité réelle. Comparer l'opérationnalité potentielle de deux groupes d'une taille si différente (trois critères contre onze) ne pourrait qu'amener de la spéculation.

Si nous voulions arriver à prévoir le genre d'un emprunt (tel que Nelson 2005 a tenté de le faire avec le lexique indigène du français), il faudrait, selon l'ordre de nos critères les plus opératoires, d'abord regarder de quelle langue il vient. Si langue donneuse possède un genre binaire, idéalement le masculin et le féminin, il est très probable (84% de probabilités) que l'emprunt porte le même genre que dans la langue donneuse, particulièrement si la langue donneuse possède des morphèmes de flexion distinctifs (italien *-o* et *-a*, arabe *-a*). Par la suite, il faut composer un paradigme d'unités conceptualisées de la même manière que l'unité dont on veut prévoir le genre (probabilités : 77,6%). Ensuite, il faut regarder si l'emprunt est monosyllabique (probabilités si applicable : 68,1%). Puis, il faut regarder si l'emprunt se termine graphiquement par une finale identique à une finale commune en français (probabilités si applicable : 61%). Finalement, il faut regarder quel pourrait être l'hyperonyme de niveau de base (probabilités : 61%). Des 100 unités que nous avons examinées, le genre de seules deux (2) unités (*jam-session* et *soja/soya*) ne peut être prévu à l'aide de ces critères, ce qui correspond à une méthode efficace à 98%. Ce résultat est intéressant, car il montre que le genre de tous les emprunts de LAG de notre échantillon peut être justifié par au moins un critère, le plus souvent par plusieurs critères.

En ce qui a trait à l'archigendre, nous ne pourrions jamais véritablement mettre un pourcentage de présence parmi notre échantillon, car sa marque est celle du masculin et on ne peut pas, dans la plupart des cas, dire avec certitude si le locuteur avait l'archigendre ou le genre masculin spécifique en tête. Cependant, on peut imaginer trois situations selon lesquelles nous pourrions déduire un pourcentage de présence pour l'archigendre. Selon le tableau cinq, le pourcentage moyen d'occurrences du masculin à travers tous les corpus est de 76,5% (donc 23,5% d'occurrence du féminin). Première situation : si l'on estimait que la moitié des cas masculins correspondait à l'archigendre, on aurait environ 38% de cas suggérant la présence d'archigendre.

Deuxièmement, si l'on parlait du postulat qu'il y a autant d'occurrences spécifiques masculines que féminines (ce qui reste hypothétique), il y aurait 53% (23,5% par genre spécifique) de cas d'archigenre. Troisièmement, si on considérait que la répartition du genre des emprunts devrait refléter la répartition du genre dans le lexique commun, et si l'on considère que le pourcentage moyen de masculin du lexique du français est de 61,2% selon les chiffres de Tucker *et al.* (1970:280)¹⁴⁴, on pourrait déduire qu'il y a 15,3% d'archigenre parmi notre échantillon (76,5% de masculin dans notre échantillon, auquel on retire le 61,2%). Lyster (2006:84) fournit lui aussi un taux de masculin à travers le nombre de substantifs appartenant à chaque genre (5019 masculins et 4630 féminins) sur le total des 9649 unités non épïcènes (312 sont épïcènes sur 9961 noms) du *Robert Junior Illustré*. À partir de ces statistiques, nous avons déduit un pourcentage de masculin de 52% (5019/9649), ce qui nous permettrait de calculer un taux d'archigenre de 24,5% (76,5% - 52%). Pour chaque situation, le pourcentage d'archigenre parmi notre échantillon varie largement de 15% à 53%. C'est pourquoi, comme nous l'avons dit plus haut, il nous est impossible de savoir avec certitude le pourcentage de présence d'archigenre dans les corpus, mais avec un tel débalancement entre les genres comme nous pouvons l'apercevoir, nous sommes certains de voir son influence. Rappelons, tel que le montre la note 36, que les études du genre des emprunts (souvent en anglais) montrent toujours un plus haut taux de masculin que de féminin en français. Il est donc monnaie courante que les études sur les emprunts recensent plus d'occurrences portant la marque du masculin, ce qui, selon nous, montre une trace du passage de l'archigenre.

Phénomène intéressant, la plupart des taux de justification des critères se situe en deux groupes, le premier entre 40%-50% (7/10 des critères FO) et le second entre 60%-70% (3/5 des critères PO). Cela est intéressant en ce que la plupart des critères PO sont à la limite du seuil de 60% que nous avons établi (sauf la conservation du genre avec 84% ou l'IP par similitude conceptuelle avec 77,6%) ou en ce que la plupart des critères FO ne le sont pas d'une bonne mesure (sauf l'initiale vocalique avec 55%). Même si nous voulions modifier les critères pour inverser le résultat de leur prédiction, en remplaçant le résultat par le genre opposé, le taux de justification de sept (7) critères serait sous le seuil de 60%. Seuls trois critères bénéficieraient d'une telle modification.

Finalement, nous sommes d'avis qu'une analyse du genre des emprunts pourrait être plus efficace si certains critères pouvaient être combinés ou adaptés pour tenir compte de plus d'un critère. Par exemple, le critère d'AP par phonème final, lorsque combiné avec la conservation du genre et le genre majoritaire en corpus en

¹⁴⁴ Cette hypothèse prend pour acquis que tous ces substantifs masculins portent le genre spécifique. Rappelons que cette étude date de plus de 45 ans, alors il est possible que les résultats soient quelque peu différents.

français, pourrait prendre en considération que le /a/ final de noms majoritairement féminins en français¹⁴⁵, venant d'une LAG, est associé normalement au féminin et non au masculin. Nous avons fait l'exercice sur notre échantillon et le taux de justification du critère passe de 41,2% (35/86) à 59,3% (51/86). D'autres combinaisons de critères pourraient être possibles, ce n'est pas le but de notre travail de continuer dans cette direction.

6.2 Retour sur les hypothèses

Revenons maintenant sur les hypothèses que nous avons formulées dans l'introduction.

6.2.1 Hypothèse 1 : forme physique et signification(s)

Notre première hypothèse était que le genre d'un emprunt morphologiquement simple serait à la fois déterminé par sa forme physique et par sa ou ses significations. Nous disions en section 1.2.1 qu'il était « possible d'observer une influence combinée par les critères de forme physique et les critères sémantiques ». Nos résultats nous ont montré que des cinq (5) critères PO, deux sont d'ordre sémantique et deux d'ordre de forme physique. Cela confirme donc l'influence combinée des deux types de critères comme étant autant déterminants l'un que l'autre dans l'attribution du genre d'emprunts différents.

6.2.2 Hypothèse 2 : conservation du genre et emprunts aux LAG

Notre deuxième hypothèse (en section 1.2.2) était que les emprunts aux LAG conserveraient pour la plupart leur genre originel. Nos résultats nous ont montré que 42 des 50 unités à l'étude ont conservé leur genre, ce qui correspond à 84% des cas. Les cas qui n'ont pas conservé leur genre trouvent tous une explication par des raisons sémantiques (paradigme latent ou hyperonymie). L'hypothèse est confirmée : la plupart des emprunts ont conservé leur genre d'origine en français. Cela pourrait être partiellement expliqué par le fait que, dans le cas des finales des emprunts italiens en *-o* et en *-a* et des emprunts arabes en *-a*, l'opposition de genre associée aux phonèmes finaux est encore partiellement ancrée dans les emprunts du lexique du français, avec des oppositions telles que *Mario-Maria*, *muchacho-muchacha*, ainsi qu'avec des unités telles qu'*Antonio* ou *mojito*, *Margarita* (genre variable au Québec) ou *diva*¹⁴⁶. L'intégration d'emprunts suivant la régularité ne ferait que renforcer cette opposition. Les finales des emprunts en *-e* et en *-i* de l'italien et en consonnes finales de l'arabe seraient beaucoup moins marquées que les finales en *-a* et en *-o*.

¹⁴⁵ Afin de ne pas prendre *barda* et *sépia* dans le calcul pour des noms féminins en français, même s'ils le sont originellement en arabe et en italien. Précisons que ces deux unités portent la marque du masculin au minimum pour des raisons sémantiques.

¹⁴⁶ Les emprunts se terminant par *-a* ne suivant pas la régularité (*sépia*, *kefta*, *barda*) peuvent être expliqués, au minimum, par des critères sémantiques (IP par similitude conceptuelle).

Nous entrevoyons également la possibilité d'une interaction entre l'archigenre et la conservation du genre des emprunts ne suivant pas la régularité en *-o* ou *-a*. En revanche, seule une étude plus poussée avec un échantillon d'emprunts plus large permettrait de mettre en lumière cette interaction potentielle.

6.2.3 Hypothèse 3 : archigenre et emprunts aux LSG

Notre troisième hypothèse (en section 1.2.3) était que la marque de l'archigenre (que nous appelions au début de ce mémoire « genre à valeur générique ») serait plus présente dans les LSG, parce que les LSG ne possèdent pas de critère d'attribution du genre ne se basant ni sur les significations, ni sur la forme physique. Autrement dit, on devrait observer un plus haut taux de masculin que dans les LAG. Le taux de masculin en corpus, pour les emprunts à l'anglais, s'élève entre 72% et 85,7% (respectivement pour les corpus européens formel et informel). Pour les emprunts au japonais, il se situe entre 98,3% et 99,6%. Les taux pour les emprunts à l'arabe et à l'italien se situent respectivement entre 58,9% et 71,7% et entre 61,3% et 70,4%. Notre hypothèse est confirmée en ce qu'aucun taux de masculin des emprunts venant d'une LAG n'est supérieur à un taux d'emprunts venant d'une LSG. Il y a également une différence notable dans le pourcentage de masculin des LSG et des LAG lorsqu'on compare les résultats d'un seul corpus. La différence de pourcentage de masculin entre les emprunts à l'anglais (la LSG au plus faible taux moyen de masculin) et le taux le plus haut des emprunts aux LAG du même corpus varie largement entre 24,4% (corpus informel européen : 85,7% (anglais) et 61,3% (italien)) et 4,2% (corpus formel européen : 72% (anglais) et 67,8% (italien)). Cette différence toujours positive confirme que les emprunts des LSG portent toujours plus la marque du masculin que les emprunts des LAG.

6.3 Ce qu'on peut retirer de cette étude

6.3.1 Neuf constats

À ce point de notre étude, nous croyons qu'il faut résumer les trouvailles que nous avons faites. Notre étude avait deux grands buts : d'abord faire une revue de la littérature en testant les critères que nous avons jugés pertinents, puis élaborer nos concepts afin de répondre aux questions de recherche que nous nous sommes posées. Revenons d'abord sur les critères que nous avons examinés. Voici, en ordre décroissant, les critères classés selon leur taux de justification :

1. Conservation du genre : 84% (42/50);
2. IP par similitude conceptuelle (Paradigme latent) 77,6% (76/98);
3. Syllabicit   enti  re : 68,1% (15/22);
4. Finale graphique : 61% (49/80);
5. IP par hyperonymie : 61% (61/100);
6. Initiale vocalique : 55,5% (5/9);
7. AP par phon  me p  nulti  me : 49,4% (41/83);

8. AP par rime riche : 42,9% (30/70);
9. IP par ellipse d'un signe entier : 42,4% (14/33);
10. Paronymie : 42,3% (11/26);
11. Syllabicit  finale graphique : 41,8% (41/98);
12. AP par phon me final : 40,7% (35/86);
13. R analyse graphique : 40% (18/45);
14. Similitude graphique : 38,4% (10/26);
15. R analyse phon tique : 22% (15/68).

Gr ce   ces r sultats, nous avons pu tirer les conclusions suivantes. Premièrement, en raison d'un m me nombre de crit res PO, les crit res s mantiques peuvent justifier autant le genre des emprunts de notre  chantillon que les crit res de forme physique. Nous basons cette conclusion sur ce que nous appelons l'op rationnalit  r elle ou effective, tel que vu en 6.1.2.1¹⁴⁷. Deuxi mement, puisque les crit res se basant sur la forme phon tique et ceux se basant sur la forme graphique ont chacun un crit re PO, ceux-ci justifient le genre des emprunts aussi bien l'un que l'autre.

Troisi mement, les emprunts aux LAG conservent g n ralement leur genre, alors que les emprunts aux LSG adoptent le plus souvent la marque associ e   l'archigenre, le masculin. Quatri mement, le crit re le plus op ratoire ne rel ve ni de la forme emprunt e du mot, ni de sa signification, mais bien du genre dans la langue donneuse. Cinqui mement, l'archigenre est pr sent, car il nous est possible de voir son influence, mais il nous est impossible de quantifier ad quatement sa pr sence. Tel que nous l'avons vu en section 6.1.5, nous pouvons estimer l'usage de l'archigenre avec notre  chantillon entre 15% et 53%. Nous sommes d'avis que les emprunts des LAG et des LSG ont des m canismes similaires (formels et s mantiques), mais aussi des m canismes fonctionnant diff remment : les emprunts des LAG conservent le plus souvent leur genre, alors que les emprunts des LSG adoptent le plus souvent la marque de l'archigenre.

Sixi mement, certains crit res fonctionnent particuli rement bien avec des emprunts   certaines langues et non pas avec des emprunts   d'autres langues. Par exemple, le crit re de la syllabicit  enti re est tr s efficace avec les emprunts   l'anglais (19 cas sur 25) et un peu avec les emprunts   l'arabe (3/25). En revanche les emprunts   l'italien ou au japonais ne pourraient  tre consid r s comme des cas analysables par ce crit re, car aucun de notre  chantillon n'est monosyllabique. Septi mement, plusieurs crit res ne concernent qu'une petite portion du lexique¹⁴⁸, mais cela ne fait pas d'eux des crit res PO ou FO. Huiti mement, les crit res phon tiques bas s sur les  tudes de Tucker *et al.* (1977), c'est- -dire l'AP par phon me final ou p nulti me,

¹⁴⁷ Afin de r sumer ce qui a  t  dit dans cette section, l'op rationnalit  effective fait la somme des crit res potentiellement op ratoires pour chaque groupe, alors que l'op rationnalit  potentielle  value (  l'aide d'un pourcentage) le nombre de crit res PO sur le total de crit res dans un groupe.

¹⁴⁸ Les crit res suivants ne s'appliquent qu'  une minorit  d'emprunts de notre  chantillon (moins de 50) : syllabicit  enti re (22 cas sur 100), initiale vocalique (9/100), IP par ellipse (33/100), paronymie (24/100), r analyse morphologique graphique (45/100) et similitude graphique (26/100).

sont des critères très présents dans la littérature sur le genre des emprunts et le genre des unités lexicales, mais nos résultats ont montré que, avec notre échantillon, ces critères ne sont pas PO avec des emprunts à l'anglais, au japonais, à l'arabe et à l'italien. Nous concluons que les critères sélectionnés sont peut-être efficaces avec le lexique entier, mais dans le cas des emprunts, leur influence est limitée. Il serait possible d'améliorer la capacité à justifier le genre de ces emprunts en combinant certains critères. Par exemple, si nous prenons en compte la langue donneuse de l'emprunt, nous pourrions modifier le critère d'AP par phonème final pour que les emprunts se terminant par /a/ et venant d'une LAG soient considérés comme féminins et non masculins.

Finalement, neuvième conclusion, plus générale que les autres, est que le genre semble d'abord attribué selon la conceptualisation avant de se baser sur le niveau taxonomique. Nous sommes venu à cette conclusion grâce au taux de masculin des emprunts au japonais et au degré de présence des emprunts comparativement aux emprunts des autres langues, notamment en considérant que le japonais était la langue la plus distante (géographiquement, culturellement, sémiologiquement) des quatre langues donneuses choisies.

Si ces conclusions que nous tirons ci-haut reflètent notre méthodologie et notre échantillon, il faut relativiser nos conclusions en fonction de l'échantillon examiné. Par exemple, une étude se concentrant sur les emprunts à l'anglais devrait ignorer le critère de la conservation du genre et se concentrer sur d'autres critères. Une étude se concentrant sur les emprunts au japonais devrait exclure le critère de syllabité entière et devrait s'attendre à observer une quasi-totalité d'emprunts portant la marque du masculin. Nous croyons cependant avoir été capable de vérifier la validité des critères trouvés en littérature à partir d'un échantillon varié, chose qu'aucune autre étude n'avait jamais fait, à notre connaissance.

6.3.2 Possibles interactions des critères

Nous avons abordé dans la section 5.3.12 la possibilité que plusieurs critères PO puissent interagir et qu'un critère ou un groupe de critères puissent déterminer le genre d'un emprunt malgré un autre critère. Prenons l'exemple du mot *slam* que nous avons utilisé en exemple dans la section 4.5.2. *Slam* est une unité qui pourrait être justifiée par sept critères. Les critères suivants sont, dans le cas de *slam*, potentiellement opératoires :

1. Syllabité finale graphique;
2. Syllabité entière;
3. Finale graphique;
4. Similitude graphique;
5. AP par phonème final;
6. AP par phonème pénultième;

7. IP par similitude conceptuelle.

Les critères suivants, dans le cas de *slam*, sont considérés comme faiblement opératoires ou non applicables :

1. Réanalyse morphologique phonétique;
2. Réanalyse morphologique graphique;
3. AP par rime riche;
4. Paronymie;
5. IP par hyperonymie;
6. IP par ellipse d'un signe entier;
7. Conservation du genre;
8. Initiale vocalique.

Ces résultats nous montrent que le genre du mot *slam* peut être justifié par des critères graphiques (3/4), par ses deux derniers phonèmes (2/2), par sa syllabité (2/2), ainsi que par ses significations (1/1). Son genre ne peut pas être justifié par une réanalyse morphologique (0/2), par une rime, un paronyme, un hyperonyme ou une ellipse. Puisque c'est un emprunt à l'anglais et qu'il commence par une consonne, les critères de conservation du genre et d'initiale vocalique ne sont pas applicables. Sur les sept critères PO, on remarque que (6/7) sont des critères de forme physique, alors qu'un seul critère est sémantique. Cela dit, le critère sémantique PO est le plus opératoire des trois critères sémantiques. Nous pouvons donc en conclure que le genre de *slam* peut à la fois être justifié par sa forme physique et par ses significations.

Prenons maintenant l'exemple de *pipeline*¹⁴⁹, un autre emprunt à l'anglais de genre non variable. Le genre de *pipeline* ne peut être justifié que par deux critères, l'IP par similitude conceptuelle et l'IP par hyperonymie. Dans ce cas, le genre de *pipeline* ne peut être justifié que par des critères sémantiques. Les cas de *dope*, *stout* sont identiques, leur genre ne peut être justifié que par les mêmes deux critères. À l'inverse de ces cas, le genre de *trip* ne peut être expliqué que par deux critères de forme physique, ceux de syllabité finale graphique et de syllabité entière. Les exemples ci-haut mentionnés mettent en évidence que dans le cas de certaines unités, le genre ne peut être justifié que par un groupe de critères ou par une combinaison de critères. Cependant, dans certains cas, un critère peut justifier le genre d'emprunts alors qu'un aspect (sémantique ou de forme physique) aurait tendance à dicter le contraire. Par exemple, nos emprunts *sépia* et *barda* devraient, si on en suit la majorité, porter la marque du féminin dictée par leur origine (LAG) et leur finale (-a). Or ces deux emprunts portent la marque du masculin malgré leur forme physique, notamment pour des raisons sémantiques. Ce sont des exemples de cas où les critères sémantiques prennent le dessus des critères formels. Le contraire est également possible. Le cas de *trip* suggère que le genre de l'unité pourrait n'être expliqué que par des critères de syllabité (graphique et entière). Ces prédominances d'un groupe sur

¹⁴⁹ Le cas de *dope* et de *stout* sont identiques, avec les critères d'IP par similitude conceptuelle et hyperonymie.

un autre nous démontrent donc que, pour pouvoir justifier adéquatement le genre d'un emprunt, il nous faut considérer plusieurs critères simultanément.

6.4 Limites et causes d'erreurs

Afin de rester dans l'éthique de notre recherche, nous prendrons le temps dans cette section de discuter des limites ou causes d'erreurs inhérentes à notre étude ou notre méthodologie.

D'abord, la première limite de notre recherche correspond au corpus choisi. Les résultats correspondent directement avec les occurrences recueillies dans les corpus. Changer de corpus, c'est automatiquement changer les résultats. Nous avons choisi d'étudier des emprunts au Québec et en Europe, car nous étions intéressé de voir comment ces deux communautés attribuaient le genre aux emprunts. Une étude se concentrant sur une seule communauté aurait nécessairement des genres majoritaires différents (seraient différents : *hype*, *minestrone*, *party*, *rave*, *tempura*, *van-vanne*). Une étude n'étudiant qu'un registre formel présenterait sûrement moins de cas à genre variable.

Ensuite, pour rester dans le thème du corpus, nous avons fait des erreurs lors de l'extraction. Les unités thématiques dans le corpus informel européen (*Doctissimo*) ont été extraites lors de deux jours différents. Nous avons inclus le premier jour les forums *nutrition-régime* et *cuisine* (ce qui représente 21 unités), et le deuxième jour nous avons ajouté le sous-forum *alimentation et santé* de la catégorie *famille* (ce qui représente 29 unités). C'est un manque de rigueur méthodologique, mais nous pensons que les conséquences sont très limitées. Si nous avons sélectionné au maximum une centaine d'occurrences par unité, nous l'avons fait de manière semi-aléatoire en ne regardant que le contexte avant (un déterminant) ou après (adjectif, participe passé), et non la provenance des commentaires ou leur contenu. Notre erreur méthodologique pourrait avoir eu une influence sur nos résultats que dans la mesure où certains emprunts n'ont pas atteint 100 occurrences. Faisons ici une liste des unités selon le forum dans lequel elles ont été cherchées :

Nutrition-régime, cuisine et alimentation et santé : *antipasto-antipasti*, *cake*, *cannelloni*, *cappuccino*, *carpaccio*, *couscous*, *espresso-expresso*, *farfalle*, *loukoum*, *maki*, *matcha*, *méchoui*, *miso*, *panini*, *pizza*, *prosciutto*, *risotto*, *sabayon*, *saké*, *salami*, *sashimi*, *seïtan*, *soja-soya*, *spaghetti*, *surimi*, *sushi*, *tagliatelle*, *tiramisu*, *tofu*;

Nutrition-régime et cuisine : *bento*, *brick*, *bruschetta*, *calzone*, *chorba*, *ciabatta*, *harissa*, *kefta*, *merguez*, *minestrone*, *mug*, *pancake*, *pancetta*, *panettone*, *ricotta*, *sandwich*, *stout*, *taboulé*, *tajine*, *tempura*, *wasabi*.

Une autre cause d'erreur est la quantité de doublons que nous avons extraite, particulièrement dans le cas des forums. Cela est le résultat de la possibilité technologique d'un forum de recopier presque instantanément les commentaires déjà publiés afin de les commenter à son tour, ce qui entraîne une rafale de citations. Nous avons sélectionné manuellement chaque occurrence afin d'éviter un nombre important de doublons, mais

nous avons préféré laisser un peu de bruit parmi les résultats et permettre à davantage de résultats intéressants de faire surface durant la phase d'analyse de corpus. Dans le cas des forums *Doctissimo*, nous avons sélectionné les commentaires intégralement (ce qui inclut les réponses aux commentaires précédents, s'emboîtant fréquemment les uns dans les autres pendant plusieurs tours de parole)¹⁵⁰. Nous estimons, une fois les corpus constitués, le nombre de doublons entre 5% et 10%. Lorsque nous avons comptabilisé les genres des occurrences, nous avons écarté autant que possible les doublons, mais il se peut qu'un petit nombre se soit infiltré dans nos résultats.

En ce qui a trait au critère de l'intégration paradigmatique par similitude conceptuelle que nous avons développé, nous croyons que son inconvénient principal est de ne pas permettre la création d'un paradigme polysémique, un paradigme dont plus d'une visée conceptuelle permettrait d'unir toutes les unités lexicales du paradigme.

On pourrait nous reprocher le fait que les paradigmes latents ont été construits *ad hoc*, sur mesure pour répondre à notre besoin. Nous répondons à cela que le paradigme ne doit être ni trop général (car il ne permettrait pas la saisie de la conceptualisation qui sous-tend le paradigme), ni trop spécifique (car il n'y aurait pas un nombre suffisant d'unités comparables). De plus, puisque la construction des paradigmes dépend de l'expérience des locuteurs, leur formation peut varier grandement selon le niveau de connaissance (du domaine d'appartenance de la réalité associée au paradigme) de l'utilisateur. Chaque paradigme est subjectif à l'individu qui le porte à l'esprit. Nous avons tenté d'objectiver nos paradigmes latents en ajoutant des unités lexicales que nous ne connaissions pas de prime abord, même si elles allaient contre le genre majoritaire du paradigme. Cela dit, nous avons également ajouté des unités que nous jugions comme personnelles, c'est-à-dire des unités qui viennent de notre perception des réalités.

Une autre limite de notre étude, réservée aux critères de forme physique, a trait à la variation graphique non standard (reflétant les différentes prononciations ou abréviations) des unités. Nous nous sommes concentré sur les prononciations et graphies standard, mais il existe fréquemment plusieurs variantes aux unités que nous avons sélectionnées et aux autres emprunts non retenus pour cette étude. Si nous voulions expliquer le genre d'une certaine unité dans une localité donnée, il nous faudrait prendre en considération la prononciation locale et la graphie locale, et ce même si notre étude porte sur des unités écrites. Cela dit, cette limite est mineure dans notre cas et ne devrait servir que d'avertissement pour quiconque voudrait continuer dans nos traces.

¹⁵⁰ Pour ce qui est du forum *Domaine Bleu*, nous avons sélectionné uniquement les phrases ou paragraphes contenant les occurrences étudiées. Nous avons ainsi grandement limité le nombre de doublons en s'assurant le mieux possible de ne pas sélectionner d'occurrences dans une citation.

Une dernière limite que porte notre analyse, et peut-être la plus importante, est que nous avons décidé de comparer des critères sémantiques à des critères de forme physique en utilisant des chiffres. Nous ne désirons pas « célébrer le culte du chiffre » (comme le dit Courbon 2015b, note 29). Or, afin de comparer le potentiel de justification de chaque critère, nous devons les comparer sur une base commune, celle des chiffres. Rappelons « [qu']en tant qu'indicateurs, les chiffres sont des outils d'interprétation, non des catégories d'analyse » (*ibid.* section 2.3.2)¹⁵¹. Ceci ne devrait pas poser problèmes avec les critères de forme physique, car leur forme est concrète, mesurable, tangible, mais dans le cas de l'évaluation de l'aspect sémantique d'une unité, les significations et affinités sémantiques sont intuitives, intangibles et abstraites. Des chiffres peuvent donc mal mesurer directement des significations. À cet effet, nous avons développé le paradigme latent pour « mesurer » le genre d'un paradigme et le comparer à l'emprunt qui l'intègre.

Si cette analyse est quelque peu « artificielle », car elle compare des facteurs d'ordres différents, elle a néanmoins le mérite de mettre à plat la suprématie d'un plan sur l'autre, *i.e.* le physique ne prédomine pas le sémantique et vice-versa. Conserver une visée large et regarder de manière objective (autant que possible) un maximum de critères était essentiel à la bonne progression de notre méthode et à l'interprétation de nos résultats. Notre étude n'est pas moins valide que celle de nos prédécesseurs (Guilford, Tucker *et al.*, Surridge, Poplack *et al.*); elle avait seulement pour objectif de valider ou infirmer ce que ceux qui nous ont inspiré ont affirmé. Elle porte une part d'interprétation en ce que les résultats sont la conséquence de la méthodologie que nous avons établie, ainsi que des unités sélectionnées et des corpus constitués. Cette interprétation devrait être relativisée en contexte d'autres unités, d'autres communautés, etc.

La portée de nos résultats est un apport aux études sur l'attribution du genre et aux études de la sémantique de corpus. Nous avons observé la fréquence de certaines formes, la proportion d'utilisation de ces formes et nous les avons comparées à ce que la littérature avait préalablement relevé. Nous avons tenté de tirer des conclusions sur les unités étudiées, sur les langues donneuses, sur les communautés qui ont produit les usages et sur les critères retenus. Nous sommes conscient que notre recherche n'a regardé que quelques aspects de quelques unités, telles quelques facettes d'un diamant. Nos résultats, à titre indicatif¹⁵² et basés sur des usages produits par des locuteurs francophones, ne visent pas à réformer les études sur le genre, mais à fournir quelques pistes supplémentaires sur le problème de l'attribution du genre. Notre étude apporte sa nuance, sa touche de couleur sur la peinture qu'est le genre. Nous avons contribué au domaine en étudiant

¹⁵¹ Nous avons décidé d'utiliser des chiffres, car ces outils d'interprétations sont « rassurants », partagés et plus concrets qu'une analyse d'abord qualitative cherchant à comparer physique et sémantique.

¹⁵² Nous sommes conscient que nos résultats sont interprétables selon le point de vue établi. Le physique et le sémantique n'étant pas accessible de la même manière, il nous a fallu les comparer sur une base commune afin de tirer des conclusions quant à leurs influences et possibles interactions. La part d'interprétation des données comporte son lot de risques d'erreurs, mais nous devons comparer les deux aspects sur un pied d'égalité afin d'obtenir un minimum d'informations crédibles.

d'autres unités, dans d'autres contextes et en essayant de fournir d'autres explications. Nous espérons avoir été capable d'atteindre nos objectifs de recherche.

7. Conclusion

Nous avons parcouru un long chemin depuis le début de cette étude. Résumons ici les principales conclusions que nous avons tirées de nos données. D'abord, de manière générale, le genre des emprunts lexicaux aux langues donneuses à genre est conservé, particulièrement lorsque le phonème (ou graphème) final est typiquement associé à un genre particulier dans la langue donneuse et receveuse (souvent –a en arabe et en italien, et –o en italien). En ce qui a trait aux emprunts lexicaux aux langues donneuses sans genre, nous avons suggéré, particulièrement dans le cas des emprunts au japonais, que leur genre était dû à la représentation mentale vague et imprécise qu'ont les locuteurs en leur assignant presque systématiquement l'archigenre sous la forme du masculin. De plus, dans le cas de toutes les langues donneuses, la marque du masculin était toujours plus fréquente que celle du féminin. Cela est sans doute dû à l'attribution de l'archigenre, comme mentionné ci-haut. Si nous sommes certain que l'archigenre est présent dans nos analyses, il nous reste difficile de mettre un chiffre sur sa présence. Selon nos calculs dans la section 6.1.3, la présence de l'archigenre pourrait être quantifiée entre 15% et 53% des cas. Ensuite, à travers les différents critères que nous avons analysés, seuls le tiers (5/15) peuvent être considérés comme potentiellement opératoires. Le critère le plus opératoire est la conservation du genre. Puis, nous avons remarqué que certains critères pouvaient particulièrement bien justifier le genre des emprunts à certaines langues alors qu'ils étaient presque totalement non opératoires dans le cas d'emprunts à d'autres langues. Des cinq critères opératoires, les critères de forme physique (graphiques et phonétiques) et les critères sémantiques ont une opérationnalité équivalente dans l'attribution du genre aux emprunts lexicaux que nous avons examinés en ce qu'il y a le même nombre de critères de chaque catégorie qui sont potentiellement opératoires (excluant la conservation du genre, qui est un critère ni sémantique, ni strictement de forme physique). Finalement, nous avons montré que le critère d'analogie phonétique (AP) par phonème final se basant sur les études de Tucker *et al.* (1977) (sans doute les plus fréquemment citées dans la littérature) ne s'appliquait qu'à une partie de notre échantillon et qu'il n'était pas potentiellement opératoire.

L'objectif principal de notre étude était de vérifier l'opérationnalité des critères que nous avons sélectionnés afin de démontrer une fois pour toutes quels étaient les critères qui pouvaient justifier le genre d'emprunts lexicaux provenant de quatre types de langue (les emprunts anglais étant souvent les seules unités examinées dans la littérature scientifique, en français). Voulant élargir nos horizons, il est possible que nos conclusions ne concordent pas avec celles que d'autres études (avec d'autres objets d'étude et d'autres objectifs) pourraient proposer. Nos résultats sont évidemment dépendants des corpus utilisés, des unités examinées et de notre méthodologie. Cependant, nous sommes confiant qu'une seconde étude avec la même méthodologie et des unités différentes pourrait arriver sensiblement aux mêmes conclusions.

En ce qui a trait aux limites de cette étude, nous sommes conscient que certaines erreurs d'extraction des données peuvent avoir influencé légèrement nos conclusions, mais nous ne croyons pas que les conséquences ont été drastiques. La quantité de doublons et l'extraction manuelle d'occurrences peuvent avoir influencé légèrement les données, sans que cela ne change nos conclusions. En définitive, considérant une méthode plus « objective » de constitution de paradigmes latents et de correspondants hyperonymiques, les critères sémantiques pourraient être différents. Cependant, si la matrice peut être remplie plus immédiatement pour les critères de forme physique, les critères sémantiques ont pour objet quelque chose d'intangible (des significations) et le remplissage de ces critères dans la matrice ne peut être accompli de manière totalement automatique.

Nous avons noté, tout au long de notre étude, de nombreuses pistes qu'il pourrait être intéressant d'étudier. D'abord, il serait possible d'étudier un grand nombre de critères (que nous n'avons pas trouvés dans la littérature de l'attribution du genre). Par exemple, il est possible que la présence d'un diacritique puisse décider du genre qu'un locuteur donne à un emprunt. Ensuite, malgré que n'ayons pas d'exemple pour ce critère probablement, on pourrait pousser la thèse de la motivation par la forme à son extrême et supposer que le nombre de graphèmes pourrait, dans des contextes écrits, être un critère potentiel ou influencer le choix du genre attribué. Les variations graphiques d'une même unité pourraient également influencer le genre. Finalement, il est possible que la présence de lettres peu fréquentes en français (<w>, <k>, <y>, voire <q> sans <u>) donnent un aspect étranger au mot et que des locuteurs attribuent davantage un genre qu'un autre. Il se pourrait même que la seule connaissance qu'un mot est un emprunt puisse entraîner un genre particulier (nous faisons l'hypothèse qu'il s'agirait de l'archigenre).

Deux autres critères qui pourraient être intéressants à étudier, bien qu'ils ne soient pas de forme physique ou entièrement sémantiques, mais purement référentiels, seraient les critères d'holonymie et de méronymie. Ceux-ci exigeraient que le genre d'un mot vienne de ses parties (les plus prototypiques) ou de son tout. Par exemple, le genre de *pizza* pourrait venir de *pâte* ou de *sauce*, mais pas de *pepperoni* ou de *fromage*. Ce critère serait particulièrement intéressant à étudier avec les emprunts thématiques en tenant compte des ingrédients les plus prototypiques, ainsi qu'avec les emprunts renvoyant à des réalités tangibles. Ainsi, le genre de *sushi* pourrait être justifié par le genre de *poisson* et de *riz*, mais pas d'*algue*.

En section 5.3.2.1, nous avons regardé si la « règle » voulant que les mots se terminant par un <e> graphique soient féminins pouvait être valide dans le cas des emprunts. Nous avons trouvé que, dans le cas de notre échantillon, ce pseudo-critère ne s'appliquait pas (ou seulement à 23,5%). En revanche, nous avons remarqué que la majorité des unités se terminant par un <e> provenaient de l'anglais. Une recherche plus en profondeur

avec un plus grand échantillon d'emprunts à l'anglais pourrait fournir une réponse davantage définitive en regard à la question que nous avons soulevée.

Un dernier critère qu'il pourrait être intéressant d'étudier est le transit linguistique de l'emprunt. Un emprunt arabe que le français a pris à l'espagnol pourrait avoir changé ou conservé son genre originel. Avec un échantillon spécifiquement constitué d'emprunts « indirects » (ayant transité par une langue à genre avant d'être empruntés), ce critère pourrait évaluer dans quelle mesure le genre du français vient réellement du français et non pas de la langue intermédiaire ou d'origine.

Au-delà d'autres critères d'attribution du genre, il pourrait être intéressant d'étudier un échantillon d'unités se basant sur d'autres thématiques ou d'autres langues. Par exemple, les emprunts à des langues à trois genres (allemand, néerlandais, russe, latin, roumain, etc.) pourraient être intéressants à analyser, particulièrement dans le cas de l'attribution du genre neutre. Les emprunts à des langues avec d'autres traits grammaticaux nominaux, comme l'animéité, pourraient également être adaptés en français d'une manière différente des emprunts aux langues à genre ou aux langues sans genre.

Afin de vérifier nos conclusions concernant l'attribution de l'archigenre et la conservation du genre, nous aimerions appliquer notre méthodologie à d'autres emprunts de langues à genre (espagnol, gallois, suédois) ou sans genre (coréen, mandarin, cantonais, vietnamien, turc, finnois, basque).

Il pourrait également être intéressant d'analyser des unités qui varient de genre à l'intérieur d'une même communauté et de regarder si la fréquence d'une unité augmente la stabilité de son genre. On pourrait supposer qu'un emprunt fréquent aurait un genre bien établi, mais nos données nous montrent que des emprunts comme *hype* (36 occurrences, corpus québécois formel), *tajine* (106, corpus québécois informel) et *vendetta* (99, corpus européen informel) ont des genres très instables, respectivement de 52,78% de masculin, de 55,66% et de 47,47%. D'expérience, nous savons que les mots *bus* « autobus » et *fan* « ventilateur » portent souvent des genres différents au Québec. D'autres mots, comme *tylenol* et *advil*, portent différents genres pour des gens originaires de la même ville et y ayant toujours vécu. Nous croyons que dans tous les cas, le genre n'est pas motivé par la forme, mais par la conceptualisation et l'hyperonyme.

Dans le cas d'emprunts, il est possible d'envisager que la connaissance qu'a un locuteur de la langue donneuse puisse avoir une influence sur l'attribution du genre par un locuteur. Cheriguen (2002:84) donne l'exemple du mot *moucharabieh*, féminin en arabe, qui est majoritairement utilisé au masculin en français, mais utilisé au féminin par l'architecte d'origine égyptienne Hassen Fathy. Il serait donc intéressant de sélectionner des unités et des locuteurs francophones connaissant à différents niveaux la langue donneuse. Il faudrait ensuite regarder dans quelle mesure le genre peut varier.

Finalement, dans le cas d'une analyse à plusieurs critères, il serait intéressant de regarder si ceux-ci peuvent avoir une influence combinée. Nous avons déjà mis en évidence le fait que la conservation du genre et la finale phonétique en /a/ pourraient être combinées pour mieux prédire le genre. Dans une analyse prédictive, de telles combinaisons pourraient être intéressantes.

8. Bibliographie

8.1 Ouvrages linguistiques et métalinguistiques

- ANDRIAMAMONJY, Pascale (2000). « Le rôle du genre grammatical au cours de la reconnaissance de noms », *L'année psychologique*, 100 :3, p.419-442.
- ARNAULD, Antoine & Claude LANCELOT (2010, 1^{ère} éd. 1997 [1660]). *Grammaire générale et raisonnée*, présentation de Jean-Marc MANDOSIO, Éd. Allia, Paris.
- ARNDT, Walter W. (1970). "Nonrandom Assingment of Loanwords: German Noun Gender", *Word*, 26:2, p.244-253.
- ARRIVÉ, Michel (1997). « Coup d'œil sur les conceptions du genre grammatical », dans *Comptes rendus des séances de l'Académie des Inscriptions et Belles-Lettres*, 141^e année, 1, p.81-96.
- AUDET, Charles-Henri & Jacques OUELLET (1984). « Considérations sur la valeur sémantique du genre », *Langues et linguistique*, 10, p.29-59.
- BADARNEH, Muhammad M. A. (2007). "Gender Assignment of Loanwords in Arabic with Special Reference to Jordanian Arabic", *Grazer linguistisches Studien*, 67-68, p.57-77.
- BARBAUD, Philippe, Charles DUCHARME & Daniel VALOIS (1982). « D'un usage particulier [sic.] du genre en canadien-français : la féminisation des noms à initiale vocalique », *Revue canadienne de linguistique*, 27 :2, p.103-133.
- BENVENISTE, Émile (1966). *Problèmes de linguistique générale*, t.2, Gallimard, Paris, 286p.
- BOULANGER, Jean-Claude & Anna MALKOWSKA (2008). « Itinéraires croisés des emprunts en alimentation : « Les années Petit Robert » », dans Jean PRUVOST (éd.), *Dictionnaires et mots voyageurs : les 40 ans du Petit Robert, de Paul Robert à Alain Rey*, Éditions des Silves, Éragny-sur-Oise, p.103-136.
- CARROLL, Susanne E. (1989). "Second-Language Acquisition and the Computational Paradigm", *Language Learning*, 39:4, p.535-594.
- CARROLL, Susanne E. (2005). "Input and SLA: Adult's Sensitivity to Different Sorts of Cues to French Gender", *Language Learning*, 55:1, p.79-138.
- CHERIGUEN, Foudil (2002). *Les mots des uns, les mots des autres : le français au contact de l'arabe et du berbère*, Casbah Éditions, Alger, 302p.
- CORBETT, Greville (1991). *Gender*, Cambridge Textbooks in Linguistics, Cambridge University Press, Cambridge, 363p.
- CORREA-ZOLI, Yole (1973). "Assignment of Gender in American Italian", *Glossa*, 7:2, p.123-128.

COURBON, Bruno (2009). "From "Luck" to "Wealth": the Stylistic (Re)Distribution of *fortuné* in Modern French", dans Kate BEECHING, Nigel ARMSTRONG & Françoise GADET (dir.), *Sociolinguistic Variation in Contemporary French*, Philadelphia et Amsterdam, John Benjamins, p.231-252.

COURBON, Bruno (2012a). « Intégration syntagmatique du sens lexical et établissement de rapports synonymiques », dans Françoise BERLAN & Gérard BERTHOMIEU, *La synonymie*, Presses de l'Université de Paris-Sorbonne, Paris, p.329-341.

COURBON, Bruno (2012b). « Une réutilisation possible du concept d'usage en sémantique diachronique? », dans Julie GLIKMAN, Leda MANSOUR et Stéphanie WEISER, *Le vocabulaire scientifique et technique en sciences du langage, Coldoc 2007*, Actes du 4e colloque des doctorants et jeunes chercheurs en Sciences du Langage du laboratoire MoDyCo, 20 et 21 juin 2007, Paris, Université Paris 10 / MoDyCo, p.102-128.

COURBON, Bruno (2015a). « Variation sémantique et jeu de la référence : le cas de la réduction de signes », dans Gaétane DOSTIE et Pascale HADERMANN (éd.), *La dia-variation en français actuel* (pp. 179-205). Peter Lang, Frankfurt.

COURBON, Bruno. (2015b). « Pratiques sémantiques et différences interindividuelles à l'ère des corpus informatisés », *Cahiers de lexicologie*, 105, p.91-126.

COURBON, Bruno (à paraître). « Questions de genre : aspects lexico-référentiels d'une catégorie linguistique », dans Laure GARDELLE et Yannick CHEVALIER (éd.), *Arbitraire et motivation du genre grammatical*, Presses universitaires de Lyon, Lyon.

DAMOURETTE, Jacques & Édouard PICHON (1968 [1911]). *Des mots à la pensée*, Tome I, D'Arthey, Paris, 674p.

DEROY, Louis (1956). *L'emprunt linguistique*, Société d'Édition Les Belles Lettres, Paris, 470p.

DE VILLERS, Marie-Éva (2009). *Multidictionnaire de la langue française*, 5^e édition, Québec Amérique, Montréal, 1707p.

DUBOIS, Jean, Mathée GIACOMO, Louis GUESPIN, Christiane MARCELLESI, Jean-Baptiste MARCELLESI & Jean-Pierre MÉVEL. *Grand dictionnaire [de] linguistique & sciences du langage*, Larousse, Paris, 2007, 516p.

ELMIGER, Daniel (2013). « Pourquoi le masculin à valeur générique est-il si tenace, en français? », *Romanica Olomucensia*, 25:2, p.113-119.

GALMICHE, Michel (1990). « Hyponymie et généricité », *Langages*, 98, p.33-49.

GAUDIN, François & Louis GUESPIN (2000). « Initiation à la lexicologie française : de la néologie aux dictionnaires », Duculot, Bruxelles, p.355.

- GEZUNDHAJT, Henriette. *Le genre*, Université de Toronto, Département d'études françaises, [en ligne : <http://www.linguistes.com/mots/genre.html>] (consulté le 23 avril 2016).
- GREVISSE, Maurice (1964, 1980, 1986, 2008). *Le Bon usage*, 8^e, 11^e, 12^e et 14^e édition, J. Duculot, Paris-Gembloux, 1194p., 1519p., 1768p. & 1600p.
- GUILBERT, Louis (1975). *La créativité lexicale*, Larousse, Paris, 285p.
- GUILBERT, Louis, René LAGANE & Georges NIOBEY (Dir.) (1971-1978). *Grand Larousse de la Langue Française* (GLLF), Éditions Larousse, Paris, 7 volumes, 6500p.
- GUILFORD, Jonathon (1999). « L'attribution du genre aux emprunts à l'anglais », *La linguistique*, 35:1, p.65-85.
- GUIRAUD, Pierre (1991). *Les gros mots*, Presses universitaires de France, Paris.
- HARLEY, Birgit (1979). "French Gender 'Rules' in the Speech of English-dominant, French-dominant and Monolingual French-speaking Children", *Working papers in bilingualism*, 19, 129-156.
- HAUGEN, Einar (1953). *The Norwegian Language in America: a Study in Bilingual behaviour*, 2 volumes, Presses de l'Université de Pennsylvanie, Philadelphia, 695p.
- HIRTLE, Walter (2007). *Language in the Mind: an Introduction to Guillaume's Theory*, McGill-Queen's University Press, Montréal-Kingston.
- HJELMSLEV, Louis (1938). « Essai d'une théorie des morphèmes », *Actes du Quatrième congrès international de linguistes*, Copenhague, Kraus reprint, Nendeln/Liechtenstein, p.140-151
- HUMBLEY, John (1974). « Vers une typologie de l'emprunt linguistique », *Cahiers de lexicologie*, 25, p.46-70.
- HUOT, Hélène (2012[2001]). *La morphologie, forme et sens des mots du français*, Armand Colin, 2^e édition, Paris.
- IBRAHIM, Muhammad Hassan (1973). "Grammatical Gender", *Janua linguarum*, Mouton, La Haye-Paris, 115p.
- KERTÉSZ, András & Csilla RÁKOSI (2012). *Data and Evidence in Linguistics : A Plausible Argumentation Model*, Cambridge University Press, Cambridge, 312p.
- KLEIBER, Georges (1990). *La sémantique du prototype*, Presses universitaires de France, Paris, 199p.
- KILARSKI, Marcin (1997). "The Case of the Unmarked Gender in Loanwords", *Studia Anglica Posnaniensia*, 32, p.89-99.
- LABROSSE, Céline (1996). *Pour une grammaire non sexiste*, Éditions du remue-ménage, Montréal, 106p.
- LABROSSE, Céline (2001). *Pour une langue française non sexiste*, Les Intouchables, Montréal, 174p.

- LAFONT, Robert (2001). *Praxématique du latin classique*, L'Harmattan, Paris.
- LAFORST, Marty & Diane VINCENT (2004). « La qualification péjorative dans tous ses états », *Langue française*, Armand Colin / Dunod, 144 :4, p.59-81. [accessible en ligne : <http://www.cairn.info/revue-langue-francaise-2004-4-page-59.htm>]
- LAFOURCADE, Cécile (2009). *Étude des emprunts faits à l'espagnol d'Amérique latine par le français : étude lexicologique et lexicographique*, Mémoire de maîtrise, Département de langues, linguistique et traduction, Université Laval, 107p.
- LEVRAT, Bernard & Gérard SABAH (1990). « 'Sorte de', une façon de rendre compte de la relation d'hyponymie/hyperonymie dans les réseaux sémantiques », *Langages*, 98, p.87-102
- LOUBIER, Christiane (2011). *De l'usage de l'emprunt linguistique*, Office québécois de la langue française, 77p.
- LOWE, Ronald (2007). *Introduction à la psychomécanique du langage*, Tome I, Presses de l'Université Laval, Québec, 570p.
- LYSTER, Roy (2006). "Predictability in French Gender Attribution: A Corpus Analysis", *French Language Studies*, 16, p.69-92.
- MALKOWSKA, Anna (2004). *La valeur sémantique du genre grammatical en français*, Mémoire de maîtrise, Département de langues, linguistique et traduction, Université Laval, 105p.
- MARTINET, André (1956). « Le genre féminin en indoeuropéen : examen fonctionnel du problème », *Bulletin de la Société de linguistique de Paris*, 52:1, p.83-95.
- MEUNIER, Jean-Guy (2006). « Le concept : de la singularité à la synthèse », *Les cahiers du LANCI*, Université du Québec à Montréal, Montréal, 5 :1, p.1-30. [accessible en ligne : <http://www.lanci.uqam.ca/wp-content/uploads/2013/10/2006-01.pdf>]
- MICHARD, Claire (2002). *Le sexe en linguistique, sémantique ou zoologie?*, L'Harmattan, volume 1, Paris.
- MICHARD, Claire (2003). « La notion de sexe en français : attribut naturel ou marque de la classe de sexe appropriée? », *Langage et société*, 106:4, p.144. [accessible en ligne : <http://www.cairn.info/revue-langage-et-societe-2003-4-p-63.htm>]
- MILNER, Jean-Claude (1989). « Genre et dimension dans les diminutifs français », *LINX*, 21:1, p.191-201. [accessible en ligne : http://www.persee.fr/doc/linx_0246-8743_1989_num_21_1_1141]
- MORTUREUX, Marie-Françoise (2008[2004]), *La lexicologie entre langue et discours*, Armand Colin, Paris, 212p.
- MURPHY, M. Lynne (2010). *Lexical Meaning*, Cambridge University Press, Cambridge, 256p.

NDIAYE, Moussa Demba (1994). « De l'articulation phonologique du genre en français », *Présence francophone*, 45, p.61-74.

NELSON, Don (2005). "French Gender Assignment Revisited", *Word*, 56:1, p.19-38.¹⁵³

NYCKEES, Vincent (1998). *La sémantique*, Belin, Paris, 365p.

NYCKEES, Vincent (2006). « Rien n'est sans raison : les bases d'une théorie continuiste de l'évolution sémantique », dans Danielle CANDEL et François GAUDIN, *Aspects diachroniques du vocabulaire*, Publications de l'Université de Rouen et du Havre, 269p.

NYCKEES, Vincent (2007). « La cognition humaine saisie par le langage : de la sémantique cognitive au médiationnisme », *Corela*, HS-6. [accessible en ligne : <http://corela.revues.org/1538>]

NYMANSSON, Karin (1995). « Le genre grammatical des anglicismes contemporains en français », *Cahiers de lexicologie*, 66:1, p.95-113.

OUELLET, Jacques (2005). « Les fondements de la sémantique conceptuelle et l'analyse des parties du discours », *Langues et linguistique*, 31, p.145-187.

OUELLET, Jacques (2014). *Sémantique conceptuelle du français*, Presses de l'Université Laval, Québec, 490p.

PISANI, Vittore (1946). *Sull'imprestito linguistico*, dans *Rendiconti dell'istituto Lombardo di Scienze e Lettere, Lett.* 73, I, Milan, 1939-1940, p.133 ss. reproduit dans le volume *Linguistica generale e indeuropea*, p.35-64.

POPLACK Shana, Alicia POUSADA & David SANKOFF (1982). "Competing Influences on Gender Assignment: Variable Process, Stable Outcome", *Lingua*, 57, p.1-28.

RASTIER, François (1987). *Sémantique interprétative*, Presses Universitaires de France, Paris.

REICHLER-BÉGUELIN, Marie-José (1993). « Motivation et remotivation des signes linguistiques », dans Alain CHRISTOL, Charles de LAMBERTERIE & Jean-Louis PERPILLOU (Éds), *Étymologie diachronique et étymologie synchronique en Grèce ancienne*, Actes du Colloque de Rouen des 21 et 22 novembre 1991 (Revue de philologie T. LXV fasc.1), Paris, Klincksieck, p.9-30 [accessible en ligne : <https://www2.unine.ch/files/content/sites/linguistique.francaise/files/shared/documents/Rouen.pdf>].

REY, Alain, Frédéric DUVAL & Gilles SIOUFFI (2007). *Mille ans de langue française : histoire d'une passion*, Perrin, Paris, 1476p.

¹⁵³ Cité comme étant de 2006 dans SurrIDGE et Lessard 2008.

REY, Alain, Marianne TOMI, Tristan HORDÉ. & Chantal TANET (2006 [1992]). *Dictionnaire Historique de la Langue Française* (DHLF), Dictionnaires Le Robert, Paris, vol. II, p.1228-1230.

RICE, Curt (2006). "Optimizing Gender", *Lingua*, 116, p.1394-1417.

RIEGEL, Martin, Jean-Christophe PELLAT & René RIOUL (2009 [1994]). *Grammaire méthodique du français*, Presses universitaires de France, Paris, 1107p.

ROCHÉ, Michel (1992). « Le masculin est-il plus productif que le féminin? », *Langue française*, 96, p.113-124. [accessible en ligne : http://www.persee.fr/web/revues/home/prescript/article/lfr_0023-8368_1992_num_96_1_5785].

ROSCH, Eleanor (1973). "On the Internal Structure of Semantic Categories", In Timothy E. Moore (Éd.), *Cognitive development and the acquisition of language*, Academic Press, New-York, p.111-144.

ROSCH, Eleanor (1978). "Principles of Categorization", dans Eleanor ROSCH & Barbara B. LLOYD (Éds), *Cognition and Categorization*, Erlbaum, Hillsdale, p.27-48.

ROSCH, Eleanor, Carolyn B. MERVIS, Wayne D. GRAY, David M. JOHNSON & Penny BOYES-BRAEM (1976). "Basic Objects in Natural Categories", *Cognitive Psychology*, 8, p.382-439.

SHIMRON, Joseph (2003). "Semitic Languages: Are They Really Root-Based?", dans Joseph SHIMRON (Éd.) *Language Processing and Acquisition in Languages of Semitic, Root-Based, Morphology*, John Benjamins, Amsterdam, p.1-28.

STOLZ, Christel (2009). "A Different Kind of Gender Problem: Maltese Loanword Gender from a Typological Perspective", dans Bernard Comrie et al., *Introducing Maltese linguistics: selected papers from the 1st International Conference on Maltese Linguistics*, 18-20 octobre 2007 à Bremen, John Benjamins, Amsterdam-Philadelphia p.321-353.

STUMP, Gregory T. (1996). "Template Morphology and Inflectional Morphology", *Yearbook of Morphology 1996*, p.217-241.

SURRIDGE, Marie E. (1982). « L'attribution du genre grammatical aux emprunts anglais en français canadien : le rôle des homologues et des monosyllabes », *Glossa*, 16:1, p.28-39.

SURRIDGE, Marie E. (1984). « Le genre grammatical des emprunts anglais en français : la perspective diachronique », *La revue canadienne de linguistique*, 29:1, p.58-72.

SURRIDGE, Marie E. (1985). « Le genre grammatical des composés en français », *La revue canadienne de linguistique*, 30:3, p.247-271.

SURRIDGE, Marie E. (1986). « Genre grammatical et dérivation lexicale en français », *La revue canadienne de linguistique*, 31:3, p.267-283.

SURRIDGE, Marie E. (1994) « Différenciation systématique des homophones nominaux non-homonymes en français », *Cahiers de lexicologie*, 64:1, p.155-175.

SURRIDGE, Marie & Greg LESSARD (2008). « Le genre en français: Quasi-régularités de l'attribution et fonction grammaticale », *Langues et linguistiques*, 32, p.181-218.

THORNTON, Anna M. (2009). "Constraining Gender Assignment Rules", *Language Sciences*, 31, p.14-32.

TOBEGY, Knud (1965). *Structure immanente de la langue française*, Librairie Larousse, Paris.

TUCKER, Richard, André RIGAULT & Wallace E. LAMBERT (1970). « Le genre grammatical des substantifs en français : Analyse statistique et étude psycholinguistique », *Actes du X^e congrès international des linguistiques*, Éditions de l'Académie de la République Socialiste de Roumanie, Bucarest, p.279-290.

TUCKER, Richard, Wallace E. LAMBERT & André RIGAULT (1977). *The French Speaker's Skill with Grammatical Gender: an Example of Rule-governed Behaviour*, Mouton, The Hague (La Haye), 132p.

TVERSKY, Barbara (1986). "Components and Categorization", dans CRAIG Colette (Éd.), *Noun Classes and Categorization*, John Benjamins, Amsterdam-Philadelphia, p.63-75.

ULLMANN, Stephen (1959). *Précis de sémantique*, Éditions A. Francke, Berne, 352p.

VIOLI, Patrizia (1987). « Les origines du genre grammatical », *Langages*, 85, p.15-34.

WARNANT, Léon (2007). *Dictionnaire des rimes*, Larousse, Paris, 572p.

WILLIAM, Joseph. M. (1975). *Origins of the English Language*, Collier Macmillan Publishers, New-York, 422p.

ZUBIN, David & Klaus-Michael KÖPCKE (1984). "Affect Classification in the German Gender System", *Lingua*, 63, p.41-96.

ZUBIN, David & Klaus-Michael KÖPCKE (1986). "Gender and Folk Taxonomy: the Indexical Relation between Grammatical and Lexical Categorization", dans CRAIG Colette (Éd.), *Noun Classes and Categorization*, John Benjamins, Amsterdam/Philadelphia, p.139-180.

8.2 Ressources textuelles exploitées

BÉLISLE, Louis-Alexandre (1958?). *Dictionnaire Bélisle de la langue française au Canada*, Société des Éditions Leland, 1390p.

Dictionnaire Usito (2014). Édition numérique. [Disponible sur cette page : <https://www.usito.com/abonnement.html>]

Dictionnaire Collins italien-anglais. Édition numérique consultée en mai 2015. [Disponible sur cette page : <http://www.collinsdictionary.com/dictionary/english-italian>].

Le Grand Robert en ligne (2014). Édition numérique. [Disponible sur cette page : <http://www.lerobert.com/espace-numerique/enligne/le-grand-robert-de-la-langue-francaise-en-ligne-12-mois.html>]

Le Petit Larousse illustré (2012), Larousse, dir. Jacques Florient, Paris, 1904p.

Le Petit Robert 2011, dictionnaire alphabétique et analogique de la langue française (2010). Josette Rey-Debove et Alain Rey (dir.), Dictionnaires Le Robert, Paris, 2837p.

Le Petit Robert de la langue française 2015 (2014). Édition numérique. [Disponible sur cette page : <https://www.lerobert.com/espace-numerique/enligne/le-petit-robert-de-la-langue-francaise-en-ligne-12-mois.html>]

Le Robert Brio, Analyse des mots et régularités du lexique. (2004). Josette Rey-Debove (dir.), Dictionnaires Le Robert, Paris, 1897p.

Nouveau Petit Larousse illustré, dictionnaire encyclopédique. (1957), Librairie Larousse, Paris, 1791p.

Petit Larousse illustré, nouveau dictionnaire encyclopédique. (1908), Librairie Larousse, dir. Claude Augé, Paris, 1664p.

9. Annexes

9.1 Nombre d'occurrences et pourcentage de masculin en corpus

9.1.1 Emprunts à l'anglais

A	D		E		F		G		H		I		J		K		L		M		N		O		P		Q		R		S		T	
	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M		
1	Unités																																	
2		9	25	34	26,47	14	1	15	93,33	2	2	4	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	v	
3	gig	105	2	107	98,13	9	58	67	13,43	25	0	25	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
4	jam	25	11	36	69,44	2	72	74	2,703	2	0	2	100	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	f	
5	jam-session	183	0	183	100	1	50	51	1,961	116	0	116	100	75	23	98	76,53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
6	party	84	1	85	98,82	0	55	55	0	14	0	14	100	0	11	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	v	
7	rave	132	0	132	100	141	0	141	100	17	0	17	100	87	0	87	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
8	slam	19	17	36	52,78	17	94	111	15,32	20	6	26	76,92	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	v	
9	hype	0	100	100	0	0	158	158	0	8	69	77	10,39	8	114	122	6,557	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	f	
10	interview	102	0	102	100	120	1	121	99,17	89	1	90	98,89	114	2	116	98,28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
11	break	141	0	141	100	115	0	115	100	111	0	111	100	105	0	105	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
12	bug/bogue	132	3	135	97,78	156	0	156	100	112	0	112	100	116	0	116	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
13	clip (musical)	103	0	103	100	108	0	108	100	112	0	112	100	109	0	109	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
14	scoop	126	0	126	100	112	0	112	100	104	0	104	100	100	1	101	99,01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
15	trip	87	0	87	100	1	0	1	100	53	0	53	100	1	5	6	16,67	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
16	drum	95	0	95	100	111	0	111	100	17	1	18	94,44	44	1	45	97,78	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
17	groove	213	0	213	100	134	0	134	100	14	0	14	100	6	0	6	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
18	pipeline	0	28	28	0	0	104	104	0	0	56	56	0	0	59	59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	f	
19	dope	4	1	5	80	117	0	117	100	0	0	0	0	107	0	107	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
20	shit	64	0	64	100	62	1	63	98,41	19	2	21	90,48	105	8	113	92,92	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
21	speed	0	97	97	0	50	0	50	100	2	45	47	4,255	81	6	87	93,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	v	
22	van/vanne	4	4	8	50	18	0	18	100	14	6	20	70	125	0	125	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
23	pancake	121	2	123	98,37	106	0	106	100	108	10	118	91,53	3	0	3	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
24	sandwich	5	0	5	100	94	1	95	98,95	3	0	3	100	123	1	124	99,19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
25	mug	17	52	69	24,64	10	21	31	32,26	0	2	2	0	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	f	
26	stout	34	0	34	100	89	0	89	100	81	0	81	100	105	0	105	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	m	
27	cake																																	

9.1.2 Emprunts au japonais

A	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Unités	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M
2																	
28	bonsai	141	0	141	100	0	0	0	0	3	0	3	100	138	0	138	100
29	dojo	133	0	133	100	126	0	126	100	5	0	5	100	25	0	25	100
30	haiku	43	0	43	100	112	0	112	100	0	0	0	0	15	0	15	100
31	karaoké	55	0	55	100	120	0	120	100	82	0	82	100	125	1	126	99,21
32	shiatsu	20	0	20	100	104	0	104	100	0	0	0	0	10	0	10	100
33	sudoku	33	0	33	100	105	0	105	100	43	0	43	100	111	0	111	100
34	tatami	63	0	63	100	116	0	116	100	1	0	1	100	26	0	26	100
35	tsunami	131	0	131	100	158	0	158	100	101	0	101	100	121	0	121	100
36	manga	144	1	145	99,31	148	1	149	99,33	17	0	17	100	126	4	130	96,92
37	nunchaku	19	0	19	100	44	0	44	100	0	0	0	0	7	0	7	100
38	kimono	64	0	64	100	113	0	113	100	4	0	4	100	92	0	92	100
39	obi	7	7	14	50	11	4	15	73,33	0	0	0	0	0	0	0	0
40	bento	22	4	26	84,62	19	4	23	82,61	14	0	14	100	106	2	108	98,15
41	tempura/tenpura	45	14	59	76,27	10	19	29	34,48	27	3	30	90	7	8	15	46,67
42	wasabi	42	0	42	100	2	0	2	100	12	0	12	100	116	2	118	98,31
43	miso	79	1	80	98,75	33	1	34	97,06	129	0	129	100	139	3	142	97,89
44	seitan	15	0	15	100	5	0	5	100	34	0	34	100	128	1	129	99,22
45	surimi	4	0	4	100	46	0	46	100	19	0	19	100	10	0	10	100
46	matcha	15	0	15	100	23	0	23	100	15	0	15	100	34	0	34	100
47	saké	145	0	145	100	137	1	138	99,28	84	0	84	100	132	0	132	100
48	soja/soya	89	0	89	100	121	0	121	100	93	0	93	100	118	1	119	99,16
49	tofu	136	0	136	100	89	0	89	100	124	0	124	100	115	0	115	100
50	maki	49	0	49	100	49	0	49	100	3	0	3	100	74	0	74	100
51	sashimi	56	0	56	100	19	0	19	100	3	0	3	100	60	0	60	100
52	sushi	86	0	86	100	59	0	59	100	21	0	21	100	121	1	122	99,18

9.1.3 Emprunts à l'arabe

A	D E F G H I J K L M N O P Q R S T															
	Corpus Québec Formel				Corpus Europe Formel				Corpus Québec informel				Corpus Europe informel			
	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M	Masc.	Fém.	Total	% de M
1	Unités															
2																
53	3	19	22	13,64	0	105	105	0	2	0	2	100	0	22	22	0 f
54	7	75	82	8,537	1	70	71	1,408	2	0	2	100	7	50	57	12,28 f
55	117	0	117	100	109	0	109	100	5	0	5	100	8	0	8	100 m
56	7	0	7	100	96	0	96	100	1	0	1	100	14	15	29	48,28 m
57	181	0	181	100	166	0	166	100	119	3	122	97,54	95	11	106	89,62 m
58	0	135	135	0	0	170	170	0	3	115	118	2,542	1	172	173	0,578 f
59	135	0	135	100	137	0	137	100	156	0	156	100	100	0	100	100 m
60	0	123	123	0	0	127	127	0	0	4	4	0	0	4	4	0 f
61	0	147	147	0	0	154	154	0	0	17	17	0	0	43	43	0 f
62	142	0	142	100	133	0	133	100	39	3	42	92,86	59	2	61	96,72 m
63	0	120	120	0	0	119	119	0	2	62	64	3,125	0	130	130	0 f
64	24	0	24	100	102	0	102	100	9	0	9	100	43	0	43	100 m
65	106	0	106	100	114	0	114	100	31	0	31	100	139	0	139	100 m
66	7	0	7	100	38	0	38	100	0	0	0	-	105	0	105	100 m
67	56	0	56	100	105	0	105	100	0	0	0	-	140	0	140	100 m
68	5	2	7	71,43	46	4	50	92	12	2	14	85,71	16	64	80	20 m
69	0	10	10	0	1	47	48	2,083	0	4	4	0	4	102	106	3,774 f
70	5	21	26	19,23	7	34	41	17,07	21	15	36	58,33	10	43	53	18,87 f
71	7	0	7	100	2	6	8	25	5	0	5	100	32	72	104	30,77 v
72	4	37	41	9,756	0	33	33	0	4	9	13	30,77	4	122	126	3,175 f
73	96	5	101	95,05	119	3	122	97,54	59	47	106	55,66	97	20	117	82,91 m
74	45	0	45	100	10	0	10	100	98	6	104	94,23	114	3	117	97,44 m
75	121	0	121	100	125	0	125	100	124	0	124	100	114	0	114	100 m
76	10	0	10	100	32	0	32	100	0	0	0	-	49	0	49	100 m
77	108	0	108	100	113	0	113	100	35	0	35	100	105	0	105	100 m

9.1.4 Emprunts à l'italien

	A	D		E		F		G		H		I		J		K		L		M		N		O		P		Q		R		S		T	
		Corpus Québec Formel		Corpus Europe Formel		Corpus Québec informel		Corpus Europe informel		Masc.		Fém.		Total		% de M		Masc.		Fém.		Total		% de M		Masc.		Fém.		Total		% de M		Genre majoritaire	
1	Unités																																		
2																																			
78	omertà	1	44	45	2,222	0	89	89	0	0	3	3	3	3	0	0	28	28	0	28	28	0	0	28	28	0	28	28	0	28	28	0	f		
79	vendetta	0	159	159	0	0	116	116	0	0	33	33	33	33	0	47	52	99	99	47	52	99	47	52	99	47	52	99	47	52	99	47	f		
80	sépia	3	2	5	60	41	3	44	93,18	1	0	1	100	100	23	3	26	88,46	m																
81	cannelloni	12	0	12	100	21	1	22	95,45	8	0	8	100	100	7	2	9	77,78	m																
82	farfalle	0	11	11	0	1	0	1	100	1	1	2	50	50	1	3	4	25	f																
83	tagliatelle	1	4	5	20	0	7	7	0	0	2	2	0	0	7	99	106	6,604	f																
84	spaghetti	128	0	128	100	46	0	46	100	107	1	108	99,07	95	8	103	92,23	m																	
85	pancetta	1	110	111	0,901	0	18	18	0	6	103	109	5,505	1	16	17	5,882	f																	
86	carpaccio	129	0	129	100	111	0	111	100	39	0	39	100	102	5	107	95,33	m																	
87	prosciutto	136	0	136	100	28	0	28	100	105	0	105	100	3	2	5	60	m																	
88	expresso/espress	113	0	113	100	99	0	99	100	25	3	28	89,29	98	3	101	97,03	m																	
89	cappuccino	119	0	119	100	102	0	102	100	116	1	117	99,15	107	0	107	100	m																	
90	panini	77	0	77	100	31	0	31	100	33	0	33	100	114	0	114	100	m																	
91	ciabatta	9	17	26	34,62	4	7	11	36,36	6	5	11	54,55	14	26	40	35	f																	
92	panettone	61	0	61	100	44	2	46	95,65	52	3	55	94,55	47	10	57	82,46	m																	
93	calzone	18	6	24	75	0	21	21	0	9	2	11	81,82	8	94	102	7,843	f																	
94	pizza	0	144	144	0	0	96	96	0	0	66	66	0	3	163	166	1,807	f																	
95	bruschetta	0	35	35	0	0	21	21	0	4	52	56	7,143	0	32	32	0	f																	
96	antipasto/i	30	0	30	100	7	0	7	100	14	0	14	100	2	0	2	100	m																	
97	minestrone	14	15	29	48,28	40	2	42	95,24	0	38	38	0	60	19	79	75,95	v																	
98	ricotta	7	95	102	6,863	1	114	115	0,87	28	41	69	40,58	3	119	122	2,459	f																	
99	risotto	162	0	162	100	119	0	119	100	108	1	109	99,08	72	0	72	100	m																	
100	sabayon	58	0	58	100	119	0	119	100	34	0	34	100	49	0	49	100	m																	
101	salami	29	0	29	100	115	0	115	100	44	0	44	100	113	0	113	100	m																	
102	tiramisu	73	0	73	100	118	0	118	100	105	0	105	100	108	1	109	99,08	m																	

9.2 Résultats obtenus à partir des critères de réanalyse morphologique

9.2.1 Réanalyse morphologique phonétique

9.2.1.1 Emprunts à l'anglais

	A		U		V		W		X	
	Unités	faux suffixe	exemples	réanalyse morphologique phonétique	genre	identique?				
1	Unités									
2										
3	gig	-	-							
4	jam	-	-							
5	jam-session	ion	partition, action, fusion			f	o			
6	party	ie/s	partie, acrobatie, courtoisie, folie, seigneurie / clapotis, cliquetis, éboulis, gribouillis			f/m	v			
7	rave	-	-							
8	slam	-	-							
9	hype	-	-							
10	interview	-	-							
11	break	-	-							
12	bug/bogue	-	-							
13	clip (musical)	-	-							
14	scoop	-	-							
15	trip	-	-							
16	drum	um	référendum, quorum, forum			m	o			
17	groove	-	-							
18	pipeline	ine	olvine, marine, canine, bottine, vitrine			f	n			
19	dope	-	-							
20	shit	ite	bronchite, sinusite, graphite, météorite			f	n			
21	speed	ide	acide, glucide, lipide, oxyde			m	o			
22	van/vanne	ane	octane, méthane, butane			m	n			
23	pancake	-	-							
24	sandwich	iche	pouliche, barbiche, potiche, caniche			m/f	v			
25	mug	-	-							
26	stout	-	-							
27	cake	-	-							

9.2.1.2 Emprunts au japonais

	A	U	V		W	X
			faux suffixe	exemples		
1	Unités			réanalyse morphologique phonétique		
2						
28	bonsai	-	-	-	-	-
29	dojo	o/eau/ot	bravo, cargo, commando, verso, trio, Mario, muchacho/éléphantreau, barreau, jambonneau/chiot, caillot, cageot, billot	m	o	
30	haiku	-	-	-	-	-
31	karaoké	é/ée/er	le parler, manger, toucher / arrivée, entrée, chevauchée, dictée, bouchée, destinée, marée, soirée/ passé, comté	m/f	v	
32	shiatsu	-	-	-	-	-
33	sudoku	-	-	-	-	-
34	tatami	ie/is	partie, acrobatie, courtoisie, folie, seigneurie / clapotis, cliquetis, éboulis, gribouillis	f/m	v	
35	tsunami	ie/is	partie, acrobatie, courtoisie, folie, seigneurie / clapotis, cliquetis, éboulis, gribouillis	f/m	v	
36	manga	a/at	diva, marina, opéra, quota, villa, Maria / secrétariat, assassinat, syndicat, doctorat	f/m	v	
37	nunchaku	-	-	-	-	-
38	kimono	o/eau/ot	bravo, cargo, commando, verso, trio, Mario, muchacho/éléphantreau, barreau, jambonneau/chiot, caillot, cageot, billot	m	o	
39	obi	ie/is	partie, acrobatie, courtoisie, folie, seigneurie / clapotis, cliquetis, éboulis, gribouillis	f/m	v	
40	bento	o/eau/ot	bravo, cargo, commando, verso, trio, Mario, muchacho/éléphantreau, barreau, jambonneau/chiot, caillot, cageot, billot	m	o	
41	tempura/tempura	a/at	diva, marina, opéra, quota, villa, Maria / secrétariat, assassinat, syndicat, doctorat	f/m	v	
42	wasabi	ie/is	partie, acrobatie, courtoisie, folie, seigneurie / clapotis, cliquetis, éboulis, gribouillis	f/m	v	
43	miso	o/eau/ot	bravo, cargo, commando, verso, trio, Mario, muchacho/éléphantreau, barreau, jambonneau/chiot, caillot, cageot, billot	m	o	
44	seitan	-	-	-	-	-
45	surimi	ie/is	partie, acrobatie, courtoisie, folie, seigneurie / clapotis, cliquetis, éboulis, gribouillis	f/m	v	
46	matcha	a/at	diva, marina, opéra, quota, villa, Maria / secrétariat, assassinat, syndicat, doctorat	f/m	v	
47	saké	é/ée/er	le parler, manger, toucher / arrivée, entrée, chevauchée, dictée, bouchée, destinée, marée, soirée/ passé, comté	m/f	v	
48	soja/soya	a/at	diva, marina, opéra, quota, villa, Maria / secrétariat, assassinat, syndicat, doctorat	f/m	v	
49	tofu	-	-	-	-	-
50	maki	ie/is	partie, acrobatie, courtoisie, folie, seigneurie / clapotis, cliquetis, éboulis, gribouillis	f/m	v	
51	sashimi	ie/is	partie, acrobatie, courtoisie, folie, seigneurie / clapotis, cliquetis, éboulis, gribouillis	f/m	v	
52	sushi	ie/is	partie, acrobatie, courtoisie, folie, seigneurie / clapotis, cliquetis, éboulis, gribouillis	f/m	v	

9.2.1.3 Emprunts à l'arabe

A		U		V		W	X
1	Unités	réanalyse morphologique phonétique					
2		faux suffixe	exemples		genre	identique?	
53	baraka	a/at	diva, marina, opéra, quoto, villa, Maria / secrétariat, assassinat, syndicat, doctorat		f/m	v	
54	chicha	a/at	diva, marina, opéra, quoto, villa, Maria / secrétariat, assassinat, syndicat, doctorat		f/m	v	
55	keffieh	é/ée/er	le parler, manger, toucher / arrivée, entrée, chevauchée, dictée, bouchée, marée, soirée, passé, comté		m/f	v	
56	chèche	-	-		-	-	
57	niqab	-	-		-	-	
58	burqa	a/at	diva, marina, opéra, quoto, villa, Maria / secrétariat, assassinat, syndicat, doctorat		f/m	v	
59	hidjab/hijab	-	-		-	-	
60	intifada	a/at	diva, marina, opéra, quoto, villa, Maria / secrétariat, assassinat, syndicat, doctorat		f/m	v	
61	fatwa	a/at	diva, marina, opéra, quoto, villa, Maria / secrétariat, assassinat, syndicat, doctorat		f/m	v	
62	djihad/fihad	ade	promenade, glissade, croisade, tornade, parade, noyade		f	n	
63	charia	a/at	diva, marina, opéra, quoto, villa, Maria / secrétariat, assassinat, syndicat, doctorat		f/m	v	
64	barda	a/at	diva, marina, opéra, quoto, villa, Maria / secrétariat, assassinat, syndicat, doctorat		f/m	v	
65	bled	-	-		-	-	
66	clebs	-	-		-	-	
67	sarouel	el(elle)	manuel, naturel / manivelle, parcelle		m/f	v	
68	brick	ique	périodique, antibiotique, physique/elliptique, chronique, éthique, optique, physique, robotique		m/f	v	
69	chorba	a/at	diva, marina, opéra, quoto, villa, Maria / secrétariat, assassinat, syndicat, doctorat		f/m	v	
70	harissa	a/at	diva, marina, opéra, quoto, villa, Maria / secrétariat, assassinat, syndicat, doctorat		f/m	v	
71	kefta/kofta/kofté	a/at/é/ée/er	le parler, manger / arrivée, chevauchée, soirée / passé, comté // diva, opéra, quoto, villa, Maria / secrétariat, doctorat		f/m	v	
72	merguez	èse	électrophorèse, diurèse, genèse, mais manganèse		f	o	
73	tajine	ine	olivine, marine, canine, bottine, vitrine		f	n	
74	taboulé	é/ée/er	le parler, manger, toucher / arrivée, entrée, chevauchée, dictée, bouchée, destinée, marée, soirée, passé, comté		m/f	v	
75	couscous	-	-		-	-	
76	loukoum	-	-		-	-	
77	méchoui	ie/is	partie, acrobatie, courtoisie, folle, seigneurie / clapotis, cliquetis, éboulis, gribouillis		f/m	v	

9.2.1.4 Emprunts à l'italien

	A		U	V		W	X
	Unités	réanalyse morphologique phonétique		genre	identique?		
1							
2		faux suffixe	exemples				
78	omertà	a/at	diva, marina, opéra, quota, villa, Maria / secrétariat, assassinat, syndicat, doctorat	f/m	v		
79	vendetta	a/at	diva, marina, opéra, quota, villa, Maria / secrétariat, assassinat, syndicat, doctorat	f/m	v		
80	sépia	a/at	diva, marina, opéra, quota, villa, Maria / secrétariat, assassinat, syndicat, doctorat	f/m	v		
81	cannelloni	ie/is	partie, acrobatie, courtoisie, folie, seigneurie / clapotis, cliquetis, éboulis, gribouillis	f/m	v		
82	farfalle	é/ée/er/al/ole	le parler, manger / arrivée, chevauchée, soirée/ passé, comté / cérémonial, festival, journal, signal / spirale, diagonale	m/f	v		
83	tagliatelle	e(elle)	manuel, naturel/ manivelle, parcelle	m/f	v		
84	spaghetti	ie/is	partie, acrobatie, courtoisie, folie, seigneurie / clapotis, cliquetis, éboulis, gribouillis	f/m	v		
85	pancetta	a/at	diva, marina, opéra, quota, villa, Maria / secrétariat, assassinat, syndicat, doctorat	f/m	v		
86	carpaccio	o/eau/ot	bravo, cargo, commando, verso, trio, Mario, muchacho/éléphanteau, barreau, jambonneau/chiot, caillot, cageot, billot	m	o		
87	prosciutto	o/eau/ot	bravo, cargo, commando, verso, trio, Mario, muchacho/éléphanteau, barreau, jambonneau/chiot, caillot, cageot, billot	m	o		
88	expresso/espress	o/eau/ot	bravo, cargo, commando, verso, trio, Mario, muchacho/éléphanteau, barreau, jambonneau/chiot, caillot, cageot, billot	m	o		
89	cappuccino	o/eau/ot	bravo, cargo, commando, verso, trio, Mario, muchacho/éléphanteau, barreau, jambonneau/chiot, caillot, cageot, billot	m	o		
90	panini	ie/is	partie, acrobatie, courtoisie, folie, seigneurie / clapotis, cliquetis, éboulis, gribouillis	f/m	v		
91	ciabatta	a/at	diva, marina, opéra, quota, villa, Maria / secrétariat, assassinat, syndicat, doctorat	f/m	v		
92	panettone	é/ée/er	le parler, manger, toucher / arrivée, entrée, chevauchée, dictée, bouchée, destinée, marée, soirée/ passé, comté	m/f	v		
93	calzone	é/ée/er	le parler, manger, toucher / arrivée, entrée, chevauchée, dictée, bouchée, destinée, marée, soirée/ passé, comté	m/f	v		
94	pizza	a/at	diva, marina, opéra, quota, villa, Maria / secrétariat, assassinat, syndicat, doctorat	f/m	v		
95	bruschetta	a/at	diva, marina, opéra, quota, villa, Maria / secrétariat, assassinat, syndicat, doctorat	f/m	v		
96	antipasto/fj	o/eau/ot	bravo, cargo, commando, verso, trio, Mario, muchacho/éléphanteau, barreau, jambonneau/chiot, caillot, cageot, billot	m	o		
97	minestrone	é/ée/er	le parler, manger, toucher / arrivée, entrée, chevauchée, dictée, bouchée, destinée, marée, soirée/ passé, comté	m/f	v		
98	ricotta	a/at	diva, marina, opéra, quota, villa, Maria / secrétariat, assassinat, syndicat, doctorat	f/m	v		
99	risotto	o/eau/ot	bravo, cargo, commando, verso, trio, Mario, muchacho/éléphanteau, barreau, jambonneau/chiot, caillot, cageot, billot	m	o		
100	sabayon	on	garçon, ourson, raton, électron, proton	m	o		
101	salami	ie/is	partie, acrobatie, courtoisie, folie, seigneurie / clapotis, cliquetis, éboulis, gribouillis	f/m	v		
102	tiramisu	-	-	-	-		

9.2.2 Réanalyse morphologique graphique

9.2.2.1 Emprunts à l'anglais

A		Y		Z	AA	AB
1	Unités	Réanalyse morphologique graphique				
2		faux suffixe	exemples		genre	identique?
3	gig	-			-	-
4	jam	-			-	-
5	jam-session	ion	<i>partition, action, fusion</i>		f	o
6	party	-			-	-
7	rave	-			-	-
8	slam	-			-	-
9	hype	-			-	-
10	interview	-			-	-
11	break	-			-	-
12	bug/bogue	-			-	-
13	clip (musical)	-			-	-
14	scoop	-			-	-
15	trip	-			-	-
16	drum	um	<i>référendum, quorum, forum</i>		m	o
17	groove	-			-	-
18	pipeline	ine	<i>olivine, marine, canine, bottine, vitrine</i>		f	n
19	dope	-			-	-
20	shit	-			-	-
21	speed	-			-	-
22	van/vanne	-			-	-
23	pancake	-			-	-
24	sandwich	-			-	-
25	mug	-			-	-
26	stout	-			-	-
27	cake	-			-	-

9.2.2.2 Emprunts au japonais

	A	Y	Z	AA	AB
1	Unités	Réanalyse morphologique graphique			
2		faux suffixe	exemples	genre	identique?
28	bonsai	-	-	-	-
29	dojo	o	<i>bravo, cargo, commando, verso, trio, Mario, muchacho</i>	m	o
30	haiku	-	-	-	-
31	karaoke	é	<i>passé, comté</i>	m	o
32	shiatsu	-	-	-	-
33	sudoku	-	-	-	-
34	tatami	-	-	-	-
35	tsunami	-	-	-	-
36	manga	a	<i>diva, marina, opéra, quota</i>	f/m	v
37	nunchaku	-	-	-	-
38	kimono	o	<i>bravo, cargo, commando, verso, trio, Mario, muchacho</i>	m	o
39	obi	-	-	-	-
40	bento	o	<i>bravo, cargo, commando, verso, trio, Mario, muchacho</i>	m	o
41	tempura/tempura	a	<i>diva, marina, opéra, quota</i>	f/m	v
42	wasabi	-	-	-	-
43	miso	o	<i>bravo, cargo, commando, verso, trio, Mario, muchacho</i>	m	o
44	seitan	-	-	-	-
45	surimi	-	-	-	-
46	matcha	a	<i>diva, marina, opéra, quota</i>	f/m	v
47	saké	é	<i>passé, comté</i>	m	o
48	soja/soya	a	<i>diva, marina, opéra, quota</i>	f/m	v
49	tofu	-	-	-	-
50	maki	-	-	-	-
51	sashimi	-	-	-	-
52	sushi	-	-	-	-

9.2.2.3 Emprunts à l'arabe

	A		Y		Z		AA	AB
	Unités		faux suffixe	exemples	Réanalyse morphologique graphique	genre		
53	baraka		a	<i>diva, marina, opéra, quota</i>		f/m	v	
54	chicha		a	<i>diva, marina, opéra, quota</i>		f/m	v	
55	keffieh		-	-		-	-	
56	chèche		-	-		-	-	
57	niqab		-	-		-	-	
58	burqa		a	<i>diva, marina, opéra, quota</i>		f/m	v	
59	hidjab/hijab		-	-		-	-	
60	intifada		a	<i>diva, marina, opéra, quota</i>		f/m	v	
61	fatwa		a	<i>diva, marina, opéra, quota</i>		f/m	v	
62	djihad/jihad		-	-		-	-	
63	charia		a	<i>diva, marina, opéra, quota</i>		f/m	v	
64	barda		a	<i>diva, marina, opéra, quota</i>		f/m	v	
65	bled		-	-		-	-	
66	clebs		-	-		-	-	
67	sarouel		el	<i>manuel, naturel</i>		m	o	
68	brick		-	-		-	-	
69	chorba		a	<i>diva, marina, opéra, quota</i>		f/m	v	
70	harissa		a	<i>diva, marina, opéra, quota</i>		f/m	v	
71	kefta/kofta/kofte		a	<i>diva, marina, opéra, quota</i>		f/m	v	
72	merguez		-	-		-	-	
73	tajine		ine	<i>olivine, marine, canine, bottine, vitrine</i>		f	n	
74	taboulé		é	<i>passé, comté</i>		m	o	
75	couscous		-	-		-	-	
76	loukoum		-	-		-	-	
77	méchoui		-	-		-	-	

9.2.2.4 Emprunts à l'italien

A		Y		Z		AA	AB
1	Unités	Réanalyse morphologique graphique					
2		faux suffixe	exemples	genre	identique?		
78	omertà	a	<i>diva, marina, opéra, quota</i>	f/m	v		
79	vendetta	a	<i>diva, marina, opéra, quota</i>	f/m	v		
80	sépia	a	<i>diva, marina, opéra, quota</i>	f/m	v		
81	cannelloni	-	-	-	-		
82	farfalle	-	-	-	-		
83	tagliatelle	elle	<i>manivelle, parcelle</i>	f	o		
84	spaghetti	-	-	-	-		
85	pancetta	a	<i>diva, marina, opéra, quota</i>	f/m	v		
86	carpaccio	o	<i>bravo, cargo, commando, verso, trio, Mario, muchacho</i>	m	o		
87	prosciutto	o	<i>bravo, cargo, commando, verso, trio, Mario, muchacho</i>	m	o		
88	expresso/espress	o	<i>bravo, cargo, commando, verso, trio, Mario, muchacho</i>	m	o		
89	cappuccino	o	<i>bravo, cargo, commando, verso, trio, Mario, muchacho</i>	m	o		
90	panini	-	-	-	-		
91	ciabatta	a	<i>diva, marina, opéra, quota</i>	f/m	v		
92	panettone	one	<i>silicone, carbone, acétone, neurone, hormone</i>	m/f	v		
93	calzone	one	<i>silicone, carbone, acétone, neurone, hormone</i>	m/f	v		
94	pizza	a	<i>diva, marina, opéra, quota</i>	f/m	v		
95	bruschetta	a	<i>diva, marina, opéra, quota</i>	f/m	v		
96	antipasto/i	o	<i>bravo, cargo, commando, verso, trio, Mario, muchacho</i>	m	o		
97	minestrone	one	<i>silicone, carbone, acétone, neurone, hormone</i>	m/f	v		
98	ricotta	a	<i>diva, marina, opéra, quota</i>	f/m	v		
99	risotto	o	<i>bravo, cargo, commando, verso, trio, Mario, muchacho</i>	m	o		
100	sabayon	on	<i>garçon, ourson, raton, électron, protron</i>	m	o		
101	salami	-	-	-	-		
102	tiramisu	-	-	-	-		

9.3 Résultats obtenus à partir des critères de syllababilité finale graphique, de syllababilité entière, de finale graphique et de similitude graphique

9.3.1 Emprunts à l'anglais

1	A	AC	AD	AE	AF	AG	AH	Finale graphique		AK	AL	AM	AN	
								finale	genre					
2	Unités	syllababilité finale graphique	syllababilité finale graphique	syllababilité finale graphique	Monosyllabique?	syllababilité entière	Genre corpus européens masculin?	finale	genre	identique?	forme(s)	Similitude graphique	genre	identique?
3	gig	c	m	n	O	m	o	ig	-	-	gig (aoctet)	m	n	
4	jam	c	m	o	O	f	n	am	m	oui	-	-	-	
5	jam-session	c	m	n	N	-	-	on	m	non	-	-	-	
6	party	v	f	n	N	-	-	y	m	oui	parti, partie, patrie	f	n	
7	rave	v	f	n	O	f	n	ave	f	non	rave (légume)	f	n	
8	slam	c	m	o	O	m	o	am	m	oui	Islam	m	o	
9	hype	v	f	n	O	f	n	ype	v	oui	-	-	-	
10	interview	c	m	n	N	-	-	ew	-	-	-	-	-	
11	break	c	m	o	O	m	o	ak	m	oui	-	-	-	
12	bug/bogue	-	-	-	O	m	o	ug/ogue v/-	-	non	bogue (noix)	f	n	
13	clip (musical)	c	m	o	O	m	o	ip	-	-	clip (pince)	m/f	n	
14	scoop	c	m	o	O	m	o	op	m	oui	scout	m	o	
15	trip	c	m	o	O	m	o	ip	-	-	tripe, trio	v	n	
16	drum	c	m	o	O	v	n	um	m	oui	-	-	-	
17	groove	v	f	n	O	m	o	oove	-	-	groupe	m	o	
18	pipeline	v	f	n	N	-	-	ine	f	non	-	-	-	
19	dope	c	m	n	O	f	n	ope	v	non	-	-	-	
20	shit	c	m	o	O	m	o	it	m	oui	-	-	-	
21	speed	c	m	o	O	m	o	eed	-	-	-	-	-	
22	van/vanne	-	-	-	O	m	o	an(ne)	m/f	non	vanne (robinet/blague)	f	n	
23	pancake	v	f	n	N	-	-	(a)ke	-	-	-	-	-	
24	sandwich	c	m	o	N	-	-	ich	-	-	-	-	-	
25	mug	c	m	o	O	m	o	ug	-	-	muguet	m	o	
26	stout	c	m	n	O	f	n	out	m	non	-	-	-	
27	cake	v	f	n	O	m	o	(a)ke	-	-	-	-	-	

9.3.2 Emprunts au japonais

A	AC	AD	AE		AF	AG	AH	AI		AJ	AK	AL	AM	AN
			syllabité finale	genre				finale	genre					
1	Unités													
2														
28	bonsaï	v	f	n	N	-	-	i	-	-	-	-	-	-
29	dojo	v	f	n	N	-	-	o	m	oui	-	mojo	m	o
30	haiku	v	f	n	N	-	-	u	m	oui	-	-	-	-
31	karaoké	v	f	n	N	-	-	é	m	oui	-	-	-	-
32	shiatsu	v	f	n	N	-	-	u	m	oui	-	-	-	-
33	sudoku	v	f	n	N	-	-	u	m	oui	-	-	-	-
34	tatami	v	f	n	N	-	-	i	m	oui	-	-	-	-
35	tsunami	v	f	n	N	-	-	i	m	oui	-	-	-	-
36	manga	v	f	n	N	-	-	a	v	non	-	mangue	f	n
37	nunchaku	v	f	n	N	-	-	u	m	oui	-	-	-	-
38	kimono	v	f	n	N	-	-	o	m	oui	-	-	-	-
39	obi	v	f	n	N	-	-	i	m	oui	-	Obi Wan Kenobe	m	o
40	bento	v	f	n	N	-	-	o	m	oui	-	-	-	-
41	tempura/tempura	v	f	n	N	-	-	a	v	oui	-	-	-	-
42	wasabi	v	f	n	N	-	-	i	m	oui	-	-	-	-
43	miso	v	f	n	N	-	-	o	m	oui	-	maso	m/f	n
44	seitan	c	m	o	N	-	-	an	m	oui	-	-	-	-
45	surimi	v	f	n	N	-	-	i	m	oui	-	-	-	-
46	matcha	v	f	n	N	-	-	a	v	non	-	match	m	o
47	saké	v	f	n	N	-	-	é	m	oui	-	-	-	-
48	soja/soya	v	f	n	N	-	-	a	v	non	-	-	-	-
49	tofu	v	f	n	N	-	-	u	m	oui	-	-	-	-
50	maki	v	f	n	N	-	-	i	m	oui	-	-	-	-
51	sashimi	v	f	n	N	-	-	i	m	oui	-	-	-	-
52	sushi	v	f	n	N	-	-	i	m	oui	-	-	-	-

9.3.3 Emprunts à l'arabe

A	AC	AD	AE	AF	AG	AH	Finale graphique		AI	AJ	AK	AL	AM	AN
							syllabité finale graphique	genre						
1	Unités													
2														
53	baraka	v	f	o	N	-	-	a	v	non	-	-	-	-
54	chicha	v	f	o	N	-	-	a	v	non	chiche (poix)	m	n	-
55	keffieh	c	m	o	N	-	-	fie(h)	-	-	-	-	-	-
56	chèche	v	f	n	N	-	-	èche	f	non	chiche (poix)	m	o	-
57	niqab	c	m	o	N	-	-	ab	-	-	-	-	-	-
58	burqa	v	f	o	N	-	-	a	v	non	-	-	-	-
59	hidjab/hijab	c	m	o	N	-	-	ab	-	-	-	-	-	-
60	intifada	v	f	o	N	-	-	a	v	non	-	-	-	-
61	fatwa	v	f	o	N	-	-	a	v	non	-	-	-	-
62	djihad/jihad	c	m	o	N	-	-	ad	-	-	-	-	-	-
63	charia	v	f	o	N	-	-	ia	v	non	chariot	m	n	-
64	barda	v	f	n	N	-	-	a	v	non	barde	m/f	n	-
65	bled	c	m	o	O	m	o	ed	-	-	-	-	-	-
66	clebs	c	m	o	O	m	o	ebs	-	-	-	-	-	-
67	sarouel	c	m	o	N	-	-	el	m	oui	-	-	-	-
68	brick	c	m	o	O	v	n	ick	-	-	brique	f	n	-
69	chorba	v	f	o	N	-	-	a	v	non	-	-	-	-
70	harissa	v	f	o	N	-	-	a	v	non	hérisson	m	n	-
71	kefta/kofta/kofte	v	f	n	N	-	-	a	v	oui	-	-	-	-
72	merguez	c	m	n	N	-	-	ez	-	-	-	-	-	-
73	tajine	v	f	n	N	-	-	ine	f	non	-	-	-	-
74	taboulé	v	f	n	N	-	-	é	m	oui	Kaboul	f/m	n	-
75	couscous	c	m	o	N	-	-	ous	-	-	-	-	-	-
76	loukoum	c	m	o	N	-	-	oum	-	-	-	-	-	-
77	méchoui	v	f	n	N	-	-	i	m	oui	-	-	-	-

9.3.4 Emprunts à l'italien

1	Unités	AC		AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN
		syllabité finale	syllabité finale graphique	genre	genre identique?	Monosyllabique?	syllabité entière	masculin?	finale	genre	genre identique?	Similitude graphique	genre	identique?
2														
78	omertà	v	f	f	N	-	-	-	a/á	v/-	non	-	-	-
79	vendetta	v	f	f	N	-	-	-	a	v	non	-	-	-
80	sépia	v	f	f	N	-	-	-	ia	v	non	-	-	-
81	cannelloni	v	f	f	N	-	-	-	i	m	oui	-	-	-
82	farfalle	v	f	f	N	-	-	-	alle/e	f/-	oui	-	-	-
83	tagliatelle	v	f	f	N	-	-	-	elle/e	f/-	oui	-	-	-
84	spaghetti	v	f	f	N	-	-	-	i	m	oui	-	-	-
85	pancetta	v	f	f	N	-	-	-	a	v	non	-	-	-
86	carpaccio	v	f	f	N	-	-	-	o	m	oui	-	-	-
87	prosciutto	v	f	f	N	-	-	-	o	m	oui	-	-	-
88	expresso/espress	v	f	f	N	-	-	-	o	m	oui	express(e) (autobus)	m/f	n
89	cappuccino	v	f	f	N	-	-	-	o	m	oui	-	-	-
90	panini	v	f	f	N	-	-	-	i	m	oui	panier	m	o
91	ciabatta	v	f	f	N	-	-	-	a	v	non	-	-	-
92	panettone	v	f	f	N	-	-	-	one/e	v/-	non	-	-	-
93	calzone	v	f	f	N	-	-	-	one/e	v/-	non	-	-	-
94	pizza	v	f	f	N	-	-	-	a	v	non	-	-	-
95	bruschetta	v	f	f	N	-	-	-	a	v	non	-	-	-
96	antipasto/f	v	f	f	N	-	-	-	o/i	m/m	oui	-	-	-
97	minestrone	v	f	f	N	-	-	-	one/e	v/-	oui	-	-	-
98	ricotta	v	f	f	N	-	-	-	a	v	non	-	-	-
99	risotto	v	f	f	N	-	-	-	o	m	oui	-	-	-
100	sabayon	c	m	m	N	-	-	-	on	m	oui	-	-	-
101	salami	v	f	f	N	-	-	-	i	m	oui	calame	m	o
102	tiramisu	v	f	f	N	-	-	-	u	m	oui	-	-	-

9.4 Résultats obtenus à partir des critères d'AP par phonème final, d'AP par phonème pénultième et d'AP par rime riche

9.4.1 Emprunts à l'anglais

	A	AO	AP	AQ	AR	AS	AT	AU	AV	AW
1	Unités	AP par Phonème final /phonème/ genre identique?	AP par Phonème pénultième /phonème/ genre identique?	AP par Rime riche	genre	identique?	genre	identique?		
3	gig	g	m	n	i	f	n	-	-	-
4	jam	m	m	o	a	m	o	-	-	-
5	jam-session	n	f	o	o	m	n	session, extraction...	f	o
6	party	i/e	f/-	n	t	-	-	part(i)e), fierté, liberté, puberté...	f	n
7	rave	v	f	n	e	-	-	rêve, trêve, grève	f	n
8	slam	m	m	o	a	m	o	lame, flamme, calame	f	n
9	hype	p	-	-	a/i	f/m	v	-	-	-
10	interview	u	m	n	j/i	f	o	-	-	-
11	break	k	m	o	e	-	-	grec, varech	m	o
12	bug/bogue	g	m	o	o	m	o	bogue	f	n
13	clip (musical)	p	-	-	i	f	n	tulipe	f	n
14	scoop	p	-	-	u	m	o	coupe, soucoupe	f	n
15	trip	p	-	-	i	f	n	grippe, tripe, fripe	f	n
16	drum	m	m	o	o	m	o	rhum, cédrôm, sérum, quorum, forum	m	o
17	groove	v	f	n	u	m	o	-	-	-
18	pipeline	n	f	n	i/a	f/m	v	berline, adrénaline, gazoline, insuline, mandoline, orpheline, naphthaline, vasoline	f	n
19	dope	p	-	-	o	m	n	-	-	-
20	shit	t	-	-	i	f	n	bronchite, bullshit	f	n
21	speed	d	f	n	i	f	n	lipide	m	o
22	van/vanne	n	f	n	a	m	n	caravane, savane, vanne	f	n
23	pancake	k	m	o	e	-	-	cake	m	o
24	sandwich	j/tj	f	n	i	f	n	-	-	-
25	mug	g	m	o	o	m	o	smog	m	o
26	stout	t	-	-	u/a	m/m	n	-	-	-
27	cake	k	m	o	e	-	-	-	-	-

9.4.2 Emprunts au japonais

A	AO	AP	AQ	AR	AS	AT	AU	AV	AW
1 Unités	AP par Phonème final	AP par Phonème/ genre identique?	AP par Phonème final	AP par Phonème/ genre identique?	AS Phonème pénultième	AT identique?	AP par Rime riche	AV genre	AW identique?
2	/phonème/ genre identique?	/phonème/ genre identique?	/phonème/ genre identique?	/phonème/ genre identique?	AS Phonème pénultième	AT identique?	Mots rimaux avec au moins trois phonèmes	AV genre	AW identique?
28	i	f n	a	m	o	-	-	-	-
29	o	m o	ʒ	m	o	-	<i>banjo, pageot, cageot</i>	m	o
30	u/y	m/m o	k	m	o	-	-	-	-
31	e	-	k	m	o	-	-	-	-
32	u/y	m/m o	s/ts	f	n	-	<i>jiu-jitsu</i>	m	o
33	u/y	m/m o	k	m	o	-	-	-	-
34	i	f n	m	m	o	-	<i>ami(e)</i>	v	n
35	i	f n	m	m	o	-	<i>ami(e)</i>	v	n
36	a	m o	g	m	o	-	-	-	-
37	u/y	m/m o	k	m	o	-	-	-	-
38	o	m o	n	f	n	-	<i>chrono, bigorneau</i>	m	o
39	i	f n	b	m	o	-	<i>phobie, lobby, hobby</i>	m	o
40	o	m o	t	-	-	-	<i>memento, lamento</i>	m	o
41	a	m n	r	m	n	-	<i>achoura</i>	f	n
42	i	f n	b	m	o	-	<i>rabbi</i>	m	o
43	o	m o	z/s	f/f	n	-	<i>ciseau, chorizo</i>	v	n
44	n/ä	f/m v	a/t	m/-	o	-	<i>gitan, titan, ni-temps/soutane, butane, gitane, octane, sultane,</i>	m/v	o/n
45	i	f n	m	m	o	-	<i>ch'timi, chimie</i>	v	n
46	a	m o	ʃ/tʃ	f	n	-	<i>chacha</i>	m	o
47	e	-	k	m	o	-	-	-	-
48	a	m o	j/ʒ	f/m	v	-	<i>séquoia</i>	m/-	o
49	y/u	m/m o	f	m	o	-	-	-	-
50	i	f n	k	m	o	-	<i>kaki</i>	m	o
51	i	f n	m	m	o	-	<i>ch'timi, chimie</i>	v	n
52	i	f n	ʃ	f	n	-	-	-	-

9.4.3 Emprunts à l'arabe

1	Unités	A		AO	AP	AQ	AR	AS	AT	AU	AV	AW
		AP par Phonème final	AP par Phonème pénultième	/phonème/ genre	identique?	identique?	Mots rimaux avec au moins trois phonèmes	AP par Rime riche	genre	identique?		
2												
53	baraka	a	m	n	k	m	n			<i>alpaca, fracas, tracas</i>	m	n
54	chicha	a	m	n	ʃ	f	o			-	-	-
55	keffieh	e	-	-	j/i	f	n			-	-	-
56	chèche	ʃ	f	n	ʃ	f	n			-	-	-
57	niqab	b	m	o	a	m	o			<i>cab</i>	m	o
58	burqa	a	m	n	k	m	n			<i>patriarcat/matriarcat, mazurka, parka</i>	v	n
59	hidjab/hijab	b	m	o	a	m	o			-	-	-
60	intifada	a	m	n	d	f	o			<i>armada, canada, dada, jangada,</i>	v	n
61	fatwa	a	m	n	w	m	n			<i>toit, putois</i>	m	n
62	djihad/jihad	d	f	n	a	m	o			<i>fusillade, grillade, noyade</i>	f	n
63	charia	a	m	n	i	f	o			<i>aria, cafétéria, pizzeria, prolétariat, secrétariat, salariat...</i>	v	n
64	barda	a	m	o	d	f	n			<i>concordat</i>	m	o
65	bled	d	f	n	d	f	n			<i>plaid, taled</i>	m	o
66	clebs	s	f	n	p/b	m/-	o			<i>biceps/triceps/quadriceps, prolepse</i>	m	o
67	sarouel	l	m	o	ε	m	o			<i>rouelle</i>	f	n
68	brick	k	m	o	i	f	n			<i>barbiturique, crique, fric</i>	m	o
69	chorba	a	m	n	b	m	n			-	-	-
70	harissa	a	m	n	s	f	o			<i>pissat</i>	m	n
71	kefta/kofta/kofta	a/e	m/-	n	t	-	-			<i>naphtha</i>	v	o
72	merguez	z	f	o	ε	m	n			-	-	-
73	tajine	n	f	n	i	f	n			<i>aubergine, frangine, androgynne, origine</i>	f	n
74	taboulé	e	-	-	l	m	o			<i>foulée, roulé(e), giboulée, goulée</i>	f	n
75	couscous	s	f	n	u	m	o			<i>rescousse, secousse</i>	f	n
76	loukoum	m	m	o	u	m	o			-	-	-
77	méchoui	i	f	n	u/w	m	o			-	-	-

9.4.4 Emprunts à l'italien

A	AO	AP	AQ	AR	AS	AT	AU	AV	AW
Unités	AP par Phonème final	AP par Phonème final	AP par Phonème final	AP par Phonème final	AP par Phonème final	AP par Phonème final	AP par Rime riche	AV	AW
	/phonème/ genre	/phonème/ genre	/phonème/ genre	/phonème/ genre	/phonème/ genre	identique?	Mots rimaux avec au moins trois phonèmes	genre	identique?
78	a	m	n	t	-	-	-	-	-
79	a	m	n	t	-	-	<i>feta, bêta, thêta, zêta</i>	m	n
80	a	m	o	i/j	f	n	-	-	-
81	i	f	n	n	f	n	<i>acrimonie, agonie, aphonie, cacophonie, cérémonie, calomnie, colonie, harmonie</i>	f	n
82	e	-	-	l	m	n	<i>azalée, vallée</i>	f	o
83	i/e	m/-	n	ε/l	m/m	n	<i>autel, hôtel, cartel, pastel, bagatelle, dentelle, jarretelle, tutelle / libellé, scellé</i>	v/m	n
84	i	f	n	t	-	-	-	-	-
85	a	m	n	t	-	-	<i>feta, bêta, thêta, zêta</i>	m	n
86	o	m	o	(t)j	f	n	<i>facho, macho</i>	m	o
87	o	m	o	t	-	-	<i>couteau</i>	m	o
88	o	m	o	s	f	n	<i>faisceau</i>	m	o
89	o	m	o	n	f	n	<i>final, casino, domino, ladino</i>	m	o
90	i	f	n	n	f	n	<i>(bi/mono)kini, infini, martini, ignominie, Virginie</i>	v	n
91	a	m	n	t	-	-	<i>duplicata, toccata</i>	v	n
92	e	-	-	n	f	n	<i>canzone, coordonnée, chaudronnée, randonnée</i>	f	n
93	e	-	-	n	f	o	<i>canzone, coordonnée, chaudronnée, randonnée</i>	f	o
94	a	m	n	z/(d)z	f	o	<i>piazza</i>	f	o
95	a	m	n	t	-	-	<i>feta, bêta, thêta, zêta</i>	m	n
96	o/i	m/f	v	t	-	-	<i>hostie, travesti, sacrisie, amnistie, eucharistie, modestie, plastie/-</i>	f/-	n
97	e/h	f/-	n	n/o	f/m	v	<i>canzone, coordonnée, chaudronnée, randonnée/neuronne, testotérone</i>	f/v	n/o
98	a	m	n	t	-	-	<i>iota, nota, quota, despotat</i>	m	n
99	o	m	o	t	-	-	<i>couteau</i>	m	o
100	ɔ	f	n	j	f	n	<i>médailon, (haillon, maillon)</i>	m	o
101	i	f	n	m	m	o	<i>amile(e)</i>	v	n
102	u	m	o	s	f	n	-	-	-

9.5 Résultats obtenus à partir du critère de paronymie

9.5.1 Emprunts à l'anglais

	A	AX	AY	AZ
1	Unités	Paronymie		
2		Paronyme	genre	identique?
3	gig	<i>gigue</i>	f	n
4	jam	<i>une djamme (Belisle 1958)</i>	f	n
5	jam-session	-	-	-
6	party	<i>partie, parti</i>	v	n
7	rave	<i>rêve</i>	m	n
8	slam	<i>flamme</i>	f	n
9	hype	-	-	-
10	interview	-	-	-
11	break	-	-	-
12	bug/bogue	<i>bogue, bog</i>	v	n
13	clip (musical)	<i>clip "pince", blip (ou bleep), clique, clic</i>	v	n
14	scoop	<i>scout</i>	m	o
15	trip	<i>tripe, trique, thrill</i>	f	n
16	drum	<i>drop, drone, drogue, drôle, -drome</i>	m	o
17	groove	<i>groupe</i>	m	o
18	pipeline	-	-	-
19	dope	-	-	-
20	shit	<i>site</i>	m	o
21	speed	-	-	-
22	van/vanne	<i>vanne "valve", val, manne, panne, canne, fan</i>	f	n
23	pancake	-	-	-
24	sandwich	-	-	-
25	mug	<i>moque</i>	f	n
26	stout	-	-	-
27	cake	-	-	-

9.5.2 Emprunts au japonais

	A	AX	AY	AZ
1	Unités	Paronymie		
2		Paronyme	genre	identique?
28	bonsaï	-	-	-
29	dojo	<i>mojo, donjon</i>	m	o
30	haiku	-	-	-
31	karaoké	-	-	-
32	shiatsu	-	-	-
33	sudoku	-	-	-
34	tatami	-	-	-
35	tsunami	-	-	-
36	manga	-	-	-
37	nunchaku	-	-	-
38	kimono	-	-	-
39	obi	-	-	-
40	bento	-	-	-
41	tempura/tenpura	-	-	-
42	wasabi	-	-	-
43	miso	<i>maso, mytho</i>	m	o
44	seitan	-	-	-
45	surimi	-	-	-
46	matcha	<i>match</i>	m	o
47	saké	-	-	-
48	soja/soya	<i>sauna</i>	m	o
49	tofu	-	-	-
50	maki	<i>Mali, mari, Marie, kaki</i>	m	o
51	sashimi	-	-	-
52	sushi	<i>souris</i>	f	n

9.5.3 Emprunts à l'arabe

	A	AX	AY	AZ
1	Unités	Paronymie		
2		Paronyme	genre	identique?
53	baraka	<i>maracas, baraque</i>	v	n
54	chicha	-	-	-
55	keffieh	-	-	-
56	chèche	-	-	-
57	niqab	-	-	-
58	burqa	-	-	-
59	hidjab/hijab	-	-	-
60	intifada	-	-	-
61	fatwa	-	-	-
62	djihad/jihad	-	-	-
63	charia	<i>Maria, chariot</i>	v	n
64	barda	<i>bardeau</i>	m	o
65	bled	<i>blaid, blair, bleime, Blake</i>	m	o
66	clebs	-	-	-
67	sarouel	-	-	-
68	brick	<i>brique, brise, trique</i>	f	n
69	chorba	-	-	-
70	harissa	-	-	-
71	kefta/kofta/kofte	-	-	-
72	merguez	-	-	-
73	tajine	-	-	-
74	taboulé	-	-	-
75	couscous	-	-	-
76	loukoum	-	-	-
77	méchoui	-	-	-

9.5.4 Emprunts à l'italien

	A	AX	AY	AZ
1	Unités	Paronymie		
2		Paronyme	genre	identique?
78	omertà	-	-	-
79	vendetta	-	-	-
80	sépia	-	-	-
81	cannelloni	-	-	-
82	farfalle	-	-	-
83	tagliatelle	-	-	-
84	spaghetti	-	-	-
85	pancetta	-	-	-
86	carpaccio	-	-	-
87	prosciutto	-	-	-
88	expresso/espress	-	-	-
89	cappuccino	-	-	-
90	panini	-	-	-
91	ciabatta	-	-	-
92	panettone	-	-	-
93	calzone	-	-	-
94	pizza	<i>pita</i>	m	n
95	bruschetta	-	-	-
96	antipasto/i	-	-	-
97	minestrone	-	-	-
98	ricotta	-	-	-
99	risotto	-	-	-
100	sabayon	-	-	-
101	salami	-	-	-
102	tiramisu	-	-	-

9.6 Résultats obtenus à partir du critère d'IP par similitude conceptuelle

Unité	Conception latente	Concepts composant le paradigme	Combinatoire de l'unité	Genre principal du paradigme	Identique au corpus?
<i>gig</i>	[événement] musical [-improvisé], gén. payé, dans le but de distraire/d'amuser	<i>spectacle, show, représentation, performance, concert</i>	<i>animer, faire, terminer par/commencer avec/par, proposer, participer à</i>	m	n
<i>jam</i>	[événement] musical improvisé	<i>pratique, spectacle, improvisation</i>	<i>animer, commencer/terminer, faire, inviter à, lever/décoller, organiser</i>	f	-
<i>jam-session</i>	[événement] musical improvisé	<i>pratique, spectacle, improvisation</i>	<i>animer/organiser, se taper/participer à/sortir de/assister à, rencontre dans</i>	f	-
<i>party</i>	[événement] festif dont la musique n'est pas le but de rencontre	<i>fête, soirée, 5 à 7, veillée, réception</i>	<i>écouter, écrire, faire, découvrir, improviser, s'initier à</i>	f	n
<i>rave</i>	[événement/lieu] musical festif nocturne à musique électronique/techno	<i>bar, club, boîte de nuit, nightlife</i>	<i>comprendre, profiter de, résister à, arrêter</i>	m	n
<i>slam</i>	[événement] poétique/artistique parlé/scandé	<i>récitation, happening, rap, discours</i>	<i>effectuer/faire/réaliser, répondre à, lire, écouter</i>	m	o
<i>hype</i>	[sentiment] [positif] médiatique, exagéré	<i>intensité, euphorie, avant-gardisme, chic</i>	<i>prendre/faire, avoir besoin de/demander</i>	v	o
<i>interview</i>	dialogue où une personne pose les questions et l'autre répond	<i>entrevue, rencontre, entretien, conférence</i>	<i>avoir, corriger, expliquer, provoquer, réparer</i>	f	o
<i>break</i>	interruption d'une activité	<i>interruption, pause, time-out, entracte, mi-temps, congé</i>	<i>faire/réaliser/tourner, écouter/voir/regarder</i>	f	n
<i>bug/bogue</i>	problème non prévu	<i>erreur, problème, imprévu, un fuck (Qc), dysfonctionnement</i>	<i>avoir/être, livrer/révéler</i>	m	o
<i>clip (musical)</i>	Segment vidéo et musical	<i>vidéo, film, (court/long) métrage</i>	<i>avoir, embarquer dans/être dans/sortir de, faire, offrir, vivre</i>	m	o
<i>scoop</i>	Information exclusive	<i>nouvelle, exclusivité, primeur, information, révélation</i>	<i>jouer, passer par</i>	f	n
<i>trip</i>	[expérience] forte en émotion [positif]	<i>euphorie, dépression, rêve, cauchemar, expérience, délire</i>	<i>avoir, aimer, donner</i>	v	v
<i>drum</i>	[Instrument] [percussion] [à peau] [cavité de résonance]	<i>tambour, batterie, (grosse) caisse (claire), conga, djembe, tamtam, bongo, tom</i>	<i>acheminer par/transporter dans, construire/raccorder, passer</i>	m	o
<i>groove</i>	rythme entraînant	<i>beat, rythme, swing</i>	<i>tomber dans/plonger dans, acheter/vendre/fournir/payer pour, fumer/consommer/prendre, avoir</i>	m	o
<i>pipeline</i>	réseau de tuyaux transportant des [matières liquides ou gazeuses]	<i>tuyau, aqueduc, gazéoduc, égout, plomberie, hose, boyau</i>	<i>proposer, distribuer/dealer/acheter/payer/vendre, trouver, fumer</i>	m	o
<i>dope</i>	[drogue] augmentant les performances	<i>stéroïde(s), créatine, hormone, drogue</i>	<i>prendre/consommer, vendre/donner, être accro à, avoir</i>	f	o
<i>shit</i>	[drogue] récréative douce (peu addictive) [relaxant]	<i>cannabis, chanvre, haschisch, pot</i>	<i>acheter, conduire, à bord de, monter dans, dépasser, endommager</i>	m	o
<i>speed</i>	drogue [excitante] possiblement hallucinogène	<i>amphétamine, champignon magique, mush(room), LSD, ecstasy</i>	<i>goûter/essayer/manger/s'offrir, partager, obtenir, préparer, retourner/tourner</i>	m	o
<i>van/vanne</i>	[gros/large] [véhicule] à grande capacité de chargement	<i>camion/camionnette, semi-remorque, fourgonnette/fourgon</i>	<i>acheter/commander/vendre, avaler/déguster/manger, chercher, mettre dans/faire/préparer,</i>	v	o

			<i>donner, prendre, refermer</i>		
<i>pancake</i>	[comestible] [pâte chauffée] [solide] [plat] [-garniture]	<i>crêpe, galette</i>	<i>faire, prendre, remplir, réchauffer</i>	f	n
<i>sandwich</i>	[comestible] [superposé] [pain]	<i>panini, club, hamburger, burger</i>	<i>boire/goûter, brasser/faire/préparer/concocter</i>	m	o
<i>mug</i>	[contenant] [avec poignée]	<i>tasse, chope, bock</i>	<i>manger/goûter, adorer/aimer/apprécier, faire/mixer dans/inclure dans, refroidir, servir</i>	f	n
<i>stout</i>	bière forte et foncée	<i>Porter, Guinness, Lager</i>	<i>acheter, avoir, créer, cultiver/entretenir/tailler, faire, offrir</i>	f	o
<i>cake</i>	[pâte] [sucré/salé][morceau à l'intérieur][massif][plat]	<i>biscuit, gâteau (aux fruits), focaccia, galette, pain (aux bananes), plum pudding</i>	<i>accueillir dans, entrer dans, s'entraîner à/fréquenter</i>	m	o
<i>bonsai</i>	[plante] [petit] [cultivé/domestiqué/-sauvage] [-comestible] à l'intérieur d'une maison	<i>bambou, cactus</i>	<i>faire, improviser/inspirer, lire/écrire/pondre/rédiger, partager</i>	m	o
<i>dojo</i>	[lieu] de pratique sportive [+terrestre/-aquatique]	<i>gymnase, gym, terrain, court (de tennis)</i>	<i>aller à, animer/organiser, chanter (à)/faire/participer à, se pratiquer</i>	m	o
<i>haiku</i>	[poétique][court][évocateur]	<i>vers, proverbe</i>	<i>pratiquer/faire, inspirer de</i>	m	o
<i>karaoké</i>	[événement/activité] musical(e) [chanté] [-improvisé]	<i>hymne national, chant, récital, opéra</i>	<i>jouer/faire/remplir, finir/résoudre/réussir/terminer, aimer</i>	m	o
<i>shiatsu</i>	[activité/technique] [relaxation/détente]	<i>méditation, yoga, massage, tai-chi, étirement</i>	<i>réunir autour de, fouler/monter sur/quitter/sauter sur, (se) retrouver sur, s'affronter sur, agenouillé sur</i>	m	o
<i>sudoku</i>	[jeu] de logique avec [éléments ordonnés] dans une grille	<i>binairo, mot caché, mot croisé</i>	<i>annoncer, déclencher/engendrer/entraîner/provoquer, causé/détruit/dévasté/emporté/provoqué/ravagé/touché/endommagé/meurtrie par</i>	m	o
<i>tatami</i>	revêtement de sol intérieur [artificiel]	<i>tapis, plancher, prélat (linoléum), parquet, moquette</i>	<i>acheter, connaître, lire, regarder</i>	m	o
<i>tsunami</i>	mouvement de l'eau [-régulier] [-normal]	<i>raz-de-marée, inondation, surcote, crue</i>	<i>menacer avec, armé de</i>	f	n
<i>manga</i>	[style] de dessin [narratif]	<i>chibi, dessin animé, bande dessinée, comic</i>	<i>enfiler/(re)mettre/revêtre/endosser/habillé de/drapé dans/sanglé dans/porter, acheter</i>	m	o
<i>nunchaku</i>	[arme] d'impact [-à feu] [flexible]	<i>fouet, fléau, lasso, éventail</i>	<i>ceinturé de/vêtu de/surmonté de/fermé de/ fermé par</i>	m	o
<i>kimono</i>	Habit ample noué à l'avant	<i>robe de chambre, peignoir, trench coat</i>	<i>apporter, mettre dans/apporter dans, faire/préparer/réaliser, manger</i>	m	o
<i>obi</i>	bande de tissu qui enrobe une partie du corps	<i>ceinture, bandeau, brassard</i>	<i>frit en/frit dans/tremper dans/enrober dans, manger/entamer, faire/préparer, façon/mode</i>	m	o
<i>bento</i>	boîte transportant la nourriture	<i>thermos, sac à lunch, boîte à lunch, tupperware</i>	<i>acheter, adorer, avaler, mettre, rajouter/hausser</i>	m	o
<i>tempura/tenpura</i>	[viande] dans une pâte frite, panée	<i>fish and chips, beignet, poulet popcorn, poulet/porc frit, croquette, nugget</i>	<i>acheter, ajouter/incorporer, bouillir, conserver, consommer, diluer, fabriquer</i>	m	n
<i>wasabi</i>	[ingrédient] [épicé]	<i>poivre, tabasco, sauce piquante, sambal oelek, cayenne, piment, jalapeño, chill, wasabi</i>	<i>couper/diviser, dorer, faire, manger, ajouter, remplacer par</i>	m	o
<i>miso</i>	[ingrédient] [salé] mis dans les soupes	<i>bouillon, base, fond, fumet</i>	<i>fouillé à, incorporer, mettre</i>	m	o

<i>seitan</i>	[comestible] substitut de protéine	<i>goberge, tofu, Soy lent, tempeh, gluten, miso, surimi</i>	<i>découvrir, parfumer à, déguster/goûter à, acheter, verser</i>	m	o
<i>surimi</i>	[comestible] substitut de protéine	<i>goberge, seitan, tofu, Soy lent, tempeh, gluten, miso</i>	<i>arroser à, boire/déguster/savourer, infuser à/mariner dans, remplacer, déglacer avec</i>	m	o
<i>matcha</i>	sorte de thé	<i>thé, chai, pu-erh, oolong, maté, rooibos</i>	<i>contenir/fait de, conserver, mélanger, préparer avec, consommer/manger, mariner à, cultiver</i>	m	o
<i>saké</i>	boisson alcoolisée à base de céréale [-liqueur]	<i>vodka, gin, whisky, bière</i>	<i>acheter, aimer/déguster, ajouter, cuisiner/préparer/frir, manger/essayer, mettre</i>	v	v
<i>soja/soya</i>	[comestible] [végétal] protéine remplaçant la viande	<i>noix, arachides, amande, tofu, pois, fève, gourgane</i>	<i>déguster/manger, donner, faire, rouler</i>	f	n
<i>tofu</i>	[comestible] substitut de protéine	<i>goberge, seitan, Soy lent, tempeh, gluten, miso, surimi</i>	<i>faire, acheter, manger, prendre</i>	m	o
<i>maki</i>	[comestible] [petit] roulé avec garniture au centre	<i>roulé, rouleau, sushi, wrap</i>	<i>accompagner/servir, adorer, avaler/déguster/goûter/manger, commander, découper, découvrir, faire</i>	m	o
<i>sashimi</i>	[comestible] viande crue apprêtée	<i>tartare, prosciutto, carpaccio</i>	<i>avoir, apporter, perdre, retrouver</i>	m	o
<i>sushi</i>	[comestible] [petit] roulé avec garniture au centre	<i>roulé, rouleau, maki, wrap</i>	<i>fumer, boire, distiller, vendre, essayer, interdire</i>	m	o
<i>baraka</i>	événement heureux opportun	<i>chance, luck</i>	<i>acheter, arborer/porter/revêtir, nouer, rajuster, caché par, coiffé de, couvert de/dissimulé/masqué par/sous, enroulé dans/recouvert de/vêtu de</i>	f	o
<i>chicha</i>	[outil] duquel on tire la fumée d'un tabac	<i>narguilé, pipe, bong, calumet, houka</i>	<i>mettre/porter, coiffé de, enrouler, dissimulé par, vêtu de</i>	m	n
<i>keffieh</i>	pièce de tissu recouvrant toute la tête	<i>foulard, bandana, bonnet, tuque, capuchon, capuche, turban</i>	<i>porter/arborer/mettre, caché sous, vêtu de, interdire/autoriser/accepter, rejeter/bannir, enlever/retirer, imposer</i>	m	o
<i>chèche</i>	pièce de tissu recouvrant la tête et le visage	<i>foulard, burqa, niqab, hidjab, tchador</i>	<i>bannir, imposer, interdire, mettre/porter, dissimulé sous</i>	m	o
<i>niqab</i>	pièce de tissu recouvrant le corps, la tête et/ou le visage	<i>voile, hidjab, tchador, chèche</i>	<i>accepter, autoriser/interdire, mettre/porter, retirer, revêtir, coiffé de, dissimulé sous</i>	m	o
<i>burqa</i>	Vêtement ample couvrant la majorité du corps	<i>toge, robe, tunique</i>	<i>commencer, depuis, durant/pendant, déclencher/provoquer, lancer</i>	f	o
<i>hidjab/hijab</i>	pièce de tissu recouvrant le corps, la tête et/ou le visage	<i>voile, niqab, tchador, chèche</i>	<i>déclencher, exécuter, lancer/ordonner/déclarer/prononcer/publier/émettre, menacée par, s'opposer à, échapper à</i>	m	o
<i>intifada</i>	conflit entre oppresseur et oppressé	<i>résistance, révolution, guerre civile, révolte, apartheid, ségrégation</i>	<i>Appeler à, exhorter à, faire, lancer, mener, s'opposer à, participer à, prôner, rejoindre, se référer à, évoquer</i>	f	o
<i>fatwa</i>	ordre religieux	<i>obligation, interdiction, jugement, condamnation</i>	<i>appliquer, défendre, imposer, importer, instaurer, interdire/permètre, prôner, fondé sur/basé sur, établir, soumis à</i>	f	o
<i>djihad/jihad</i>	conflit entre peuples	<i>guerre, bataille, intifada, ségrégation, apartheid, lutte, terrorisme</i>	<i>débarrasser de, déménager, déplacer, déposer, porter, poser, transporter, emmener, emporter, prendre</i>	f	n
<i>charia</i>	Loi(s) d'origine religieuse ou spirituelle	<i>commandement, doctrine, révélation</i>	<i>bouger, rester dans, envoyer à/de, habiter, partir à/de, quitter, rentrer à/du, débarquer dans, habiter dans, retourner</i>	f	o

<i>barda</i>	Ensemble d'[objets] indésirables encombrants	<i>désordre, bordel, chaos, cochonnerie, affaires</i>	<i>avoir, promener, punir</i>	m	o
<i>bled</i>	localité éloigné et de petite taille	<i>patelin, trou, village, hameau, fond</i>	<i>acheter, commencer, faire, finir, imposer, porter, terminer, voir</i>	m	o
<i>clebs</i>	[familier/péjoratif] chien	<i>cabot, clébard, pitou, toutou</i>	<i>fait avec, déguster, faire, manger</i>	m	o
<i>sarouel</i>	vêtement porté sur les jambes	<i>pantalon(s), jogging(s), legging(s), jean(s), culotte(s)</i>	<i>faire, finir, manger, préparer</i>	m	o
<i>brick</i>	[comestible] [pâte cuite] qui contient de la viande, fruits et/ou légumes	<i>rouleau, pochette, hors-d'oeuvre, egg-roll, strudel, cannelloni, ravioli, dumpling, tortellini, won ton, gnocchi</i>	<i>acheter, ajouter, conserver, consommer, essayer, mettre, mélanger, prendre, remplacer, manger, rehaussé de</i>	m	o
<i>chorba</i>	[comestible] [soupe] avec légumes et pâtes	<i>soupe, minestrone, potage, garbure, frik, harira, jardinière, goulash</i>	<i>aimer, cuisiner, cuire, faire, manger, mélanger, farci à</i>	f	o
<i>harissa</i>	[comestible] [ingrédient] [épice]	<i>poivre, tabasco, sauce piquante, sambal oelek, cayenne, piment, jalapeño, chili, wasabi</i>	<i>cuire, manger, prendre, sentir</i>	m	n
<i>kefta/kofta/kofte</i>	[comestible] viande en boulette ou autour d'un bâton	<i>boulette, brochette, fricadelle, falafel, kebab, yakitori</i>	<i>acheter, cuisiner/faire/préparer/réaliser/cuire, déguster, manger</i>	v	o
<i>merguez</i>	[comestible] viande de forme [cylindrique] épicée	<i>chorizo, Wurst, saucisse</i>	<i>aimer/apprécier/déguster, faire/préparer, manger</i>	f	o
<i>tajine</i>	[comestible] viande et légumes mélangés	<i>ragoût, pot-au-feu, potée, stew</i>	<i>accompagner, acheter, aimer/déguster, ajouter, conserver, cuire/cuisiner/faire/préparer, goûter/manger, offrir, remplacer, utiliser, servir</i>	m	o
<i>taboulé</i>	[comestible] salade de céréales et légumes [froid]	<i>salade</i>	<i>manger</i>	f	n
<i>couscous</i>	[comestible] féculents et céréales [massif]	<i>risotto, paella, riz, gruau, pouding au riz</i>	<i>avoir, déguster/goûter/manger, faire, organiser/préparer, servir, inviter à, participer à, réussir</i>	m	o
<i>loukoum</i>	[comestible] pâte sucrée et durcie, coupés en [morceau] carrés ou rectangulaires	<i>nougat, kiss, tire, caramel, fudge, sucre à la crème</i>	<i>briser, rompre, créer, entretenir, confronté à, dénoncer, régner</i>	m	o
<i>méchoui</i>	[événement] collectif culinaire / [animal] cuit sur broche	<i>party, épluchette, barbecue, pique-nique, dîner, dégustation, vins et fromages, Oktoberfest, agneau</i>	<i>déclencher, dénoncer, faire, lancer, mener, voir, embarquer dans</i>	m	o
<i>omertà</i>	obligation de garder le silence	<i>promesse, loi, règle, obligation, mort</i>	<i>virer à, faire, utiliser, baigner/(re)plonger dans, aimer</i>	f	o
<i>vendetta</i>	promesse solennelle de protéger ou venger l'honneur	<i>serment, promesse, pacte, représailles, vengeance</i>	<i>farcir, goûter, manger</i>	f	o
<i>sépia</i>	pigment de couleur d'origine naturelle (végétal, animal ou minéral)	<i>indigo, carmin, chlorophylle, ocre, cramoisi, pourpre, curcuma, safran</i>	<i>préparer, décorer, peindre, casser</i>	m	o
<i>cannelloni</i>	[comestible] pâte [fourrée ou tubulaire] [-plate] [±longue]	<i>macaroni, penne, rigatoni, tortellini, ravioli, dumpling, won ton, gnocchi</i>	<i>cuire, préparer, accompagner de, déguster</i>	m	o
<i>farfalle</i>	[comestible] pâte [-fourrée] [-tubulaire] [-plate] [-longue]	<i>orecchiette, gnocchi, ruote, coquille, cavatelli, papillon, oreille</i>	<i>aimer, ajouter, commander/prendre, cuire/faire/préparer, donner, manger</i>	f	o
<i>tagliatelle</i>	[comestible] pâte [plate] [longue]	<i>ruban, lasagne, linguine, fettucine</i>	<i>ajouter, couper, remplacer (par), rissoler, émietter, sauter</i>	f	o
<i>spaghetti</i>	[comestible] pâte [longue] [-plate]	<i>cappelli(ni), vermicelle, fusilli, gemelli</i>	<i>adorer, agrémente de, déguster/manger, faire/préparer, prendre, proposer, servir</i>	m	o
<i>pancetta</i>	[comestible] viande	<i>mortadelle, andouille</i>	<i>acheter, ajouter, chauffer, manger,</i>	f	o

	roulée [cylindrique] [-sec]		<i>mélanger, remplacer (par), prendre</i>		
<i>carpaccio</i>	[comestible] viande crue apprêtée	<i>tartare, bresaola, sashimi</i>	<i>ajouter, avaler/boire, commander/prendre, goûter/déguster/savourer/siroter, faire, préparer, finir</i>	m	o
<i>prosciutto</i>	[comestible] viande roulée [cylindrique], séchée, fumée ou assaisonnée	<i>saucisson, salami, pepperoni, chorizo</i>	<i>accompagner, adorer/déguster, boire/siroter, commander/prendre, faire, proposer, préparer, servir</i>	m	o
<i>expresso/ espresso</i>	[comestible] boisson caféinée qui donne de l'énergie	<i>café, red bull, thé, Monster, XS</i>	<i>croquer/manger, déguster, faire/préparer, prendre</i>	m	o
<i>cappuccino</i>	[comestible] boisson à base de café, lait et mousse	<i>latte, macchiato, mocha, flat white, misto, breve</i>	<i>inséré dans/fourré dans, faire, manger</i>	m	o
<i>panini</i>	[comestible] [entre deux tranches de pain][chaud/grillé]	<i>sandwich, BLT, hamburger, cheeseburger, grilled cheese, club sandwich, croque-monsieur, guédille, smoked meat</i>	<i>accompagner, acheter, cuire, donner, faire, manger, partager</i>	m	o
<i>ciabatta</i>	[comestible] pain [compact][±long][plat(+le vures)][large][-rond]	<i>fouée, pain bis, pain meunier</i>	<i>faire, manger, prendre/commander, préférer, acheter</i>	m	n
<i>panettone</i>	[comestible] [pâte] [sucré/salé][morceau à l'intérieur][massif][plat]	<i>gâteau (aux fruits), focaccia, galette, pain (aux bananes), brioche</i>	<i>accompagner, acheter, aimer/déguster, choisir, commander, composer/garnir, cuire, faire, livrer, manger, mettre sur, préférer, préparer</i>	m	o
<i>calzone</i>	[comestible] pâte cuite avec garniture fourrée au centre [-dessert]	<i>pochette, empanada, pâté, tourte</i>	<i>déguster/manger, goûter, faire/préparer</i>	f	o
<i>pizza</i>	[comestible] pâte cuite avec garniture sur le dessus	<i>tarte, fougasse, pissaladière</i>	<i>apporter, constituer, déguster, commencer par/avec</i>	f	o
<i>bruschetta</i>	[comestible] hors d'œuvre avec garniture sur pain	<i>crostino, canapé, tapas, croque-monsieur, crouton</i>	<i>aimer/adorer, faire, goûter/manger, préparer, cuire</i>	m	n
<i>antipasto/i</i>	[comestible] hors d'œuvre avec viande et légumes en morceaux	<i>tartare, salade, bouchée</i>	<i>ajouter, fabriquer, faire, farci à/fourré à, mélanger (avec), utiliser, préparer (avec)</i>	f	n
<i>minestrone</i>	[comestible] [soupe] avec légumes et pâtes	<i>potage, chorba, soupe, garbure, frik, harira, jardinière, goulash</i>	<i>adorer, accompagner de, déguster, faire, farci à, manger, préparer, servir</i>	f	n
<i>ricotta</i>	[comestible] fromage blanc mou et crémeux	<i>cottage, feta, brousse, bocconcini, caillebotte, caillé, mascarpone</i>	<i>confectionner, faire, farcir à, monter, préparer, présenter, servir, verser, émulsionner, nappé de</i>	m	n
<i>risotto</i>	[comestible] riz cuit, assaisonné	<i>paella, pilaf, jambalaya</i>	<i>acheter, ajouter, faire, manger, prendre</i>	m	o
<i>sabayon</i>	[comestible] dessert liquide	<i>milkshake, caramel liquide, compote, crème</i>	<i>confectionner, congeler, cuisiner, essayer, faire/réaliser, goûter, manger, terminer par</i>	v	v
<i>salami</i>	[comestible] viande roulée [cylindrique] [+sec]	<i>saucisson, pepperoni, chorizo, prosciutto</i>		m	o
<i>tiramisu</i>	[comestible] [dessert/gâteau] [à étage]	<i>gâteau (au chocolat), génoise, opéra, forêt-noire, Red Velvet, shortcake, mille-feuille</i>		m	o

9.7 Résultats obtenus à partir des critères d'IP par hyperonymie et d'IP par ellipse d'un signe entier

9.7.1 Emprunts à l'anglais

A	BF	BG	BH	BI	BJ	BK
	IP par Hyperonymie (X est un type de Y)	genre	identique?	IP par ellipse d'un signe entier	genre	identique?
1	Unités	genre	identique?	forme elidé	genre	identique?
2	Hyperonyme					
3	gig	m	n	-	-	-
4	jam	v	n	session	f	n
5	jam-session	v	n	-	-	-
6	party	f	n	-	-	-
7	rave	m/v	o	-	-	-
8	slam	f	n	-	-	-
9	hype	f	n	-	-	-
10	interview	f	o	-	-	-
11	break	f	n	-	-	-
12	bug/bogue	m	o	-	-	-
13	clip (musical)	m/v	o	vidéo	m/f	v
14	scoop	f	n	-	-	-
15	trip	v	n	-	-	-
16	drum	v	n	kit, basse	m/f	v
17	groove	v	n	-	-	-
18	pipeline	m	o	-	-	-
19	dope	f	o	-	-	-
20	shit	f	n	-	-	-
21	speed	f	n	dating, car	m	o
22	van/vanne	m	n	-	-	-
23	pancake	m	o	-	-	-
24	sandwich	m	o	pain	m	o
25	mug	v	n	-	-	-
26	stout	f	o	-	-	-
27	cake	m	o	-	-	-

9.7.2 Emprunts au japonais

A		BF		BG	BH	BI	BJ	BK
1	Unités	IP par Hyperonymie (X est un type de Y)		genre	identique?	IP par ellipse d'un signe entier	genre	identique?
2	Hyperonyme			genre	identique?	forme elidé	genre	identique?
28	bonsai	<i>plante, arbre, culture</i>		f	n	<i>ficus</i>	m	o
29	dojo	<i>gymnase</i>		m	o	-	-	-
30	haiku	<i>prose, poème, poésie</i>		f	n	-	-	-
31	karaoke	<i>divertissement</i>		m	o	<i>soirée, bar, machine</i>	m/f	v
32	shiatsu	<i>philosophie, massage, thérapie</i>		f	n	<i>massage, masseur, séance</i>	m/f	v
33	sudoku	<i>puzzle, casse-tête, jeu</i>		m	o	<i>jeu</i>	m	o
34	tatami	<i>revêtement, tapis</i>		m	o	-	-	-
35	tsunami	<i>vague, désastre</i>		v	n	-	-	-
36	manga	<i>bande dessinée, littérature, livre</i>		f	n	<i>personnage, dessin, style, m</i>	o	o
37	nunchaku	<i>fléau, arme</i>		v	n	-	-	-
38	kimono	<i>robe, costume, vêtement</i>		m	o	-	-	-
39	obi	<i>ceinture</i>		f	n	<i>ceinture</i>	f	n
40	bento	<i>boîte, plateau, vaisselle</i>		f	n	-	-	-
41	tempura/tenpura	<i>beignet, pâte</i>		v	o	<i>sauce</i>	f	n
42	wasabi	<i>raifort, moutarde, condiment, épice</i>		v	n	-	-	-
43	miso	<i>pâte, condiment, bouillon</i>		m	o	<i>soupe, pâte</i>	f	n
44	seitan	<i>galette, gluten</i>		v	n	-	-	-
45	surimi	<i>chair, pâte</i>		f	n	<i>bâton</i>	m	o
46	matcha	<i>thé</i>		m	o	<i>thé, poudre</i>	m/f	v
47	saké	<i>bière, vin, boisson</i>		f	n	-	-	-
48	soja/soya	<i>jus, purée, haricot, nutriment, sauce, protéine, légumineuse, fibre</i>		f	n	<i>sauce, crème, lait (de), ge</i>	m/f	v
49	tofu	<i>fromage, nourriture, produit</i>		m	o	-	-	-
50	maki	<i>sushi, rouleau</i>		m	o	<i>sushi/rouleau</i>	m	o
51	sashimi	<i>sushi, poisson</i>		m	o	-	-	-
52	sushi	<i>tartine, boulette, cuisine</i>		f	n	-	-	-

9.7.3 Emprunts à l'arabe

	A	BF	BG	BH	BI	BJ	BK
1	Unités	IP par Hyperonymie (X est un type de Y)	genre	identique?	IP par ellipse d'un signe entier	genre	identique?
2	Hyperonyme		genre	identique?	forme elidé	genre	identique?
53	baraka		f	o	-	-	-
54	chicha		f	o	-	-	-
55	keffieh		m	o	-	-	-
56	chèche		m	o	-	-	-
57	niqab		m	o	-	-	-
58	burqa		f	o	-	-	-
59	hidjab/hijab		m	o	-	-	-
60	intifada		f	o	-	-	-
61	fatwa		f	o	-	-	-
62	djihad/jihad		f	n	-	-	-
63	charia		f	o	-	-	-
64	barda		m	o	-	-	-
65	bled		f	n	-	-	-
66	clebs		m	o	-	-	-
67	sarouel		m	o	pantalon, robe	m/f	v
68	brick		f	n	soupe + beignet	m/f	v
69	chorba		f	o	-	-	-
70	harissa		f	o	sauce, poudre	f	o
71	kefta/kofta/kofte		v	o	tajine, boulette, kebab	m/f	v = o
72	merguez		f	o	saucisse	f	o
73	tajine		m	o	-	-	-
74	taboulé		m	o	-	-	-
75	couscous		m	o	-	-	-
76	loukoum		m	o	-	-	-
77	méchoui		m	o	-	-	-

9.7.4 Emprunts à l'italien

A	BF	BG	BH	BI	BJ	BK
Unités	IP par Hyperonymie (X est un type de Y)	genre	identique?	IP par ellipse d'un signe entier	genre	identique?
2	Hyperonyme			forme elidé		
78	<i>solidarité, intelligence</i>	f	o	-	-	-
79	<i>legs, coutume, tradition, loi</i>	f	o	-	-	-
80	<i>photographie, couleur, finition, calmar, seiche</i>	f	n	<i>tonalité, couleur, ton, asp.</i>	m/f	v
81	<i>plat, pâte, mets, repas</i>	m	o	-	-	-
82	<i>pâte</i>	f	o	<i>pâte</i>	f	o
83	<i>pâte, ruban</i>	m	n	-	-	-
84	<i>pâte, fil, repas, mets, plat</i>	m	o	<i>pâte, courge, soupe</i>	m/f	v
85	<i>lard, jambon, charcuterie, bacon, porc</i>	m	n	-	-	-
86	<i>bresaola, capicola</i>	f	n	-	-	-
87	<i>jambon</i>	m	o	-	-	-
88	<i>café</i>	m	o	-	-	-
89	<i>café</i>	m	o	<i>café/machine (à)</i>	m/f	v
90	<i>sandwich, pain</i>	m	o	<i>sandwich</i>	m	o
91	<i>pain</i>	m	n	<i>pain/sandwich</i>	m	n
92	<i>pain, brioche, bonbon</i>	m	o	<i>brioche</i>	f	n
93	<i>bresaola, capicola</i>	f	o	<i>pizza</i>	f	o
94	<i>tarte</i>	f	o	-	-	-
95	<i>tartine, sandwich, pizza, entrée</i>	f	o	<i>tartine</i>	f	o
96	<i>préparation, hors d'œuvre, entrée, buffet</i>	v	n	-	-	-
97	<i>soupe, entrée</i>	f	n	<i>soupe</i>	f	n
98	<i>fromage, pâte, mousse, brousse</i>	f	o	-	-	-
99	<i>riz, paella, plat</i>	m	o	-	-	-
100	<i>gratin, dessert</i>	m	o	<i>glace, sauce</i>	f	n
101	<i>saucisson, chorizo, saucisse</i>	m	o	-	-	-
102	<i>cheesecake, gâteau, verrine, flan, mousse, dessert</i>	m	o	-	-	-

9.8 Résultats obtenus à partir des critères de conservation du genre et de l'initiale vocalique

9.8.1 Emprunts à l'anglais

	A	BL	BM	BN	BO	BP
1	Unités	Conservation du genre		Initiale vocalique		
2		Genre d'origine	Identique?	Voyelle?	G corpus Qc F/V?	
3	gig	-	-	n	-	-
4	jam	-	-	n	-	-
5	jam-session	-	-	n	-	-
6	party	-	-	n	-	-
7	rave	-	-	n	-	-
8	slam	-	-	n	-	-
9	hype	-	-	n	-	-
10	interview	-	-	o	f	o
11	break	-	-	n	-	-
12	bug/bogue	-	-	n	-	-
13	clip (musical)	-	-	n	-	-
14	scoop	-	-	n	-	-
15	trip	-	-	n	-	-
16	drum	-	-	n	-	-
17	groove	-	-	n	-	-
18	pipeline	-	-	n	-	-
19	dope	-	-	n	-	-
20	shit	-	-	n	-	-
21	speed	-	-	n	-	-
22	van/vanne	-	-	n	-	-
23	pancake	-	-	n	-	-
24	sandwich	-	-	n	-	-
25	mug	-	-	n	-	-
26	stout	-	-	n	-	-
27	cake	-	-	n	-	-

9.8.2 Emprunts au japonais

	A	BL	BM	BN	BO	BP
1	Unités	Conservation du genre		Initiale vocalique		
2		Genre d'origine	Identique?	Voyelle?	G corpus Qc F/V?	
28	bonsaï	-	-	n	-	-
29	dojo	-	-	n	-	-
30	haïku	-	-	v	m/-	n
31	karaoké	-	-	n	-	-
32	shiatsu	-	-	n	-	-
33	sudoku	-	-	n	-	-
34	tatami	-	-	n	-	-
35	tsunami	-	-	n	-	-
36	manga	-	-	n	-	-
37	nunchaku	-	-	n	-	-
38	kimono	-	-	n	-	-
39	obi	-	-	o	v/-	o
40	bento	-	-	n	-	-
41	tempura/tenpura	-	-	n	-	-
42	wasabi	-	-	n	-	-
43	miso	-	-	n	-	-
44	seitan	-	-	n	-	-
45	surimi	-	-	n	-	-
46	matcha	-	-	n	-	-
47	saké	-	-	n	-	-
48	soja/soya	-	-	n	-	-
49	tofu	-	-	n	-	-
50	maki	-	-	n	-	-
51	sashimi	-	-	n	-	-
52	sushi	-	-	n	-	-

9.8.3 Emprunts à l'arabe

	A	BL	BM	BN	BO	BP
1	Unités	Conservation du genre		Initiale vocalique		
2		Genre d'origine	Identique?	Voyelle?	G corpus Qc F/V?	
53	baraka	f	o	n	-	-
54	chicha	f	o	n	-	-
55	keffieh	f	n	n	-	-
56	chèche	m	o	n	-	-
57	niqab	m	o	n	-	-
58	burqa	f	o	n	-	-
59	hidjab/hijab	m	o	v	m	n
60	intifada	f	o	o	f	o
61	fatwa	f	o	n	-	-
62	djihad/jihad	m	o	n	-	-
63	charia	f	o	n	-	-
64	barda	f	n	n	-	-
65	bled	m	o	n	-	-
66	clebs	m	o	n	-	-
67	sarouel	m	o	n	-	-
68	brick	v	n	n	-	-
69	chorba	f	o	n	-	-
70	harissa	f	o	v	f/v	o
71	kefta/kofta/kofte	f	n	n	-	-
72	merguez	m	n	n	-	-
73	tajine	m	o	n	-	-
74	taboulé	m	o	n	-	-
75	couscous	m	o	n	-	-
76	loukoum	m	o	n	-	-
77	méchoui	m	o	n	-	-

9.8.4 Emprunts à l'italien

	A	BL	BM	BN	BO	BP
1	Unités	Conservation du genre		Initiale vocalique		
2		Genre d'origine	Identique?	Voyelle?	G corpus Qc F/V?	
78	omertà	f	o	o	f	o
79	vendetta	f	o	n	-	-
80	sépia	(seppia) f	n	n	-	-
81	cannelloni	m	o	n	-	-
82	farfalle	f	o	n	-	-
83	tagliatelle	f	o	n	-	-
84	spaghetti	m	o	n	-	-
85	pancetta	f	o	n	-	-
86	carpaccio	m	o	n	-	-
87	prosciutto	m	o	n	-	-
88	expresso/espresso	m	o	o	m	n
89	cappuccino	m	o	n	-	-
90	panini	m	o	n	-	-
91	ciabatta	f	o	n	-	-
92	panettone	m	o	n	-	-
93	calzone	m	n	n	-	-
94	pizza	f	o	n	-	-
95	bruschetta	f	o	n	-	-
96	antipasto/i	m	o	o	m	n
97	minestrone	m	n	n	-	-
98	ricotta	f	o	n	-	-
99	risotto	m	o	n	-	-
100	sabayon	(zabaione/zabaglione) m	o	n	-	-
101	salami	m	o	n	-	-
102	tiramisu	m	o	n	-	-