

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LA MAITRISE DE L'ORTHOGRAPHE LEXICALE DU FRANÇAIS ET DE
L'ESPAGNOL

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAITRISE EN LINGUISTIQUE
CONCENTRATION DIDACTIQUE DES LANGUES

PAR
SONY MAYARD

SEPTEMBRE 2007

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

Avant-propos

L'orthographe du français a fait couler beaucoup d'encre, et il est évident que cela ne s'arrêtera pas de sitôt. À travers la francophonie, tout locuteur du français garde en mémoire *le long trajet* qui lui a permis d'avoir *une certaine maîtrise* de l'orthographe de sa langue. Pour donner une idée du parcours effectué, il suffit de dire que depuis le primaire jusqu'au secondaire, on peut compter onze années consacrées à l'étude de l'enseignement de l'orthographe. D'après certains spécialistes de l'orthographe du français comme Grevisse (1948), Beslais (1965) ou Catach (1989), plusieurs pays, comme l'Espagne la Turquie et la Roumanie, ont modifié l'orthographe de leur langue. Ces modifications permettent, semble-t-il, aux scripteurs de ces pays d'être compétents dans leur langue écrite. Ce sont ces facteurs de longue durée d'acquisition et de compétence qui ont motivé le choix de ce sujet de recherche. Comme un grand nombre de francophones, nous trouvons que l'enseignement de l'orthographe du français s'éternise et qu'il y aurait moyen de rendre cet apprentissage moins long et plus intéressant. Par conséquent, le premier but poursuivi par cette recherche, c'est d'ajouter notre voix à celles des personnes qui croient à une possible réforme de l'orthographe du français. D'après ce groupe de personnes, il serait important de débarrasser l'orthographe du français d'un grand nombre de particularités orthographiques héritées du passé. Le deuxième but visé, c'est de voir si les scripteurs espagnols sont plus performants dans leur langue maternelle que les francophones le sont dans la leur.

Cependant, cette recherche, qui a nécessité un travail de précision, se limite à un nombre restreint de sujets et à peu de données précises sur les variables humaines et sociales des élèves qui ont fourni le corpus. Ainsi, nous pensons à l'importance d'élargir, dans le futur ou dans des recherches

de plus grande envergure, le cadre de cette étude qui se veut une comparaison de la compétence orthographique entre deux populations de langue différente et surtout, de systèmes orthographiques différents, l'un étant beaucoup plus régulier que l'autre.

Remerciements

En tout premier lieu, mes sincères remerciements vont à Marie Nadeau, la directrice de ce mémoire. Dans les moments de laisser-aller, elle a su, par ses judicieux commentaires et ses mots d'encouragement, m'indiquer la voie à suivre, rendant par le fait même ce projet possible.

Je tiens également à remercier Bertrand Fournier du Service de Consultation en Analyse de Données (SCAD) pour les calculs statistiques, et le personnel du département de linguistique et de didactique des langues, en particulier, mesdames Lorraine D'aragon et Louisette Émirkanian, pour m'avoir donné la chance de relever ce grand défi. En effet, au début de ce projet, je nageais dans l'inconnu et j'ai trouvé sur ma route ces deux personnes qui ont su me guider dans la bonne direction.

Il ne me reste qu'à remercier ma petite famille : mon épouse, Nancy Leclerc, qui malgré des moments difficiles s'est montrée très patiente pendant la durée de cette traversée et qui m'a aidé à apprivoiser Excel. Ma petite fleur, Glaïdja Mayard, qui par ses sourires et sa bonne humeur a réussi à me redonner espoir quand le doute et le désespoir planaient au-dessus de ma tête. Enfin, ma tendre mère, Marie-Louise Stéphen, mon père, Alix, mes grands-parents, Edner, le modèle, Maria, Philomène qui, avec amour, ont tracé le chemin de ma vie. Sans oublier mes amis sincères qui se reconnaîtront dans ce texte.

TABLE DES MATIÈRES

AVANT-PROPOS	ii
LISTE DES TABLEAUX	vii
RÉSUMÉ	ix
INTRODUCTION	1
CHAPITRE I	7
LA PROBLÉMATIQUE	
1.1 Introduction	7
1.2 Les deux clans en présence : les conservateurs et les réformateurs	7
1.3 La difficile maîtrise de l'orthographe du français	9
1.4 La résistance aux changements	12
1.5 La réforme de l'orthographe française vue par des spécialistes	15
Conclusion	18
1.6 Les objectifs de recherche	19
CHAPITRE II	20
LE CADRE THÉORIQUE	
2.1 Introduction	20
2.2 Les notions de phonème et de graphème	20
2.3 L'orthographe grammaticale versus l'orthographe lexicale	26
2.4 L'apprentissage de l'orthographe lexicale	28
2.5 Les homophones	32
2.6 Les échelles de fréquence	36
2.6.1 L'échelle Dubois-Buyse (1989)	37
2.6.2 L'échelle Baudot (1992)	38
2.6.3 L'échelle Préfontaine (1995)	39
2.6.4 L'échelle Fortier (1993)	40
2.6.5 L'échelle Juilland (1964)	41

2.7	Les projets de réforme de l'orthographe	42
2.8	La typologie des erreurs lexicales	48
2.9	Les questions de recherche	53
2.10	Les hypothèses de recherche	54
	Conclusion	55
	CHAPITRE III	57
	LA MÉTHODOLOGIE	
3.1	Introduction	57
3.2	Les élèves	57
3.3	Les corpus	58
3.4	La fréquence des mots dans le corpus	60
3.5	La méthode d'analyse	68
3.6	Les calculs	73
3.7	Les limites de notre recherche	77
	Conclusion	79
	CHAPITRE IV	80
	L'ANALYSE DES RÉSULTATS	
4.1	Introduction	80
4.2	La maîtrise de l'orthographe	80
4.3	Le rôle de la fréquence dans la maîtrise de l'orthographe lexicale	88
4.4	Les types d'erreurs lexicales	107
4.5	Les projets de réforme et l'orthographe lexicale	123
	4.5.1 Les mots composés	126
	4.5.2 Les accents	127
	4.5.3 Les graphèmes (consonnes) x	129
	4.5.4 Les lettres grecques	130
	4.5.5 Les consonnes doublées	132
	Conclusion	137
	CONCLUSION GÉNÉRALE	141

APPENDICE A	147
Corpus Nadeau, Matute, Reyes-Gonzales	
APPENDICE B	150
Codification des graphies erronées pour chaque corpus	
APPENDICE C	187
Typologie des erreurs selon Catach (1980) dans chaque langue	
APPENDICE D	211
Tableaux des résultats	
APPENDICE E	216
Liste des mots avec le pourcentage d'erreurs en français et en espagnol	
BIBLIOGRAPHIE	226

LISTE DES TABLEAUX

Tableau		Page
3.1	Répartition des élèves par dictée	58
3.2	Nombre de mots disponibles par corpus	59
3.3	Nombre de mots du corpus par catégorie de fréquence dans chaque langue	62
3.4	Liste de mots du corpus français et leur fréquence selon Fortier (1993)	63
3.5	Liste de mots du corpus espagnol et leur fréquence selon Juilland (1964)	65
3.6	Extrait des graphies erronées et mots réussis par élève francophone	69
3.7	Extrait des graphies erronées et mots réussis par élève hispanophone	70
3.8	Classement des erreurs selon la typologie de Catach	71
4.1	Pourcentage de réussite dans chaque langue	80
4.2	Répartition des élèves selon le pourcentage de réussite par intervalle de réussite chez les francophones et les hispanophones	81
4.3	Répartition des erreurs selon leur catégorie, lexicale, grammaticale	82
4.4	Répartition des élèves selon le % d'erreurs lexicales chez les francophones et les hispanophones	85
4.5	Nombre de mots selon le pourcentage d'erreur (graphies non conformes) et la fréquence en français	89
4.6	Nombre de mots selon le pourcentage d'erreur (graphies non conformes) et la fréquence en espagnol	90
4.7	Nombre de mots selon le pourcentage d'erreur (graphies non conformes), la fréquence et le % des mots selon la catégorie de fréquence en français	91
4.8	Nombre de mots selon le pourcentage d'erreur (graphies non conformes), la fréquence et le % des mots selon la catégorie de fréquence en espagnol	91
4.9	Mots mal orthographiés par 50 % et plus des élèves par groupe comparé	94

4.10	Mots F et PF dont le pourcentage d'erreur est inférieur à 5 % chez les francophones	97
4.11	Mots PF réussis et ratés avec leur indice de fréquence selon Fortier (1993)	100
4.12	Mots F et PF dont le pourcentage d'erreur est inférieur à 5 % chez les hispanophones	101
4.13	Répartition des types d'erreurs lexicales et du pourcentage dans chaque langue	106
4.14	Erreurs phonogrammiques de lettres inaudibles ajoutées dans chaque langue	108
4.15	Erreurs phonogrammiques sur le phonème /s/ dans chaque langue	111
4.16	Nombre de graphies lexicales erronées, nombre de mots et % du corpus dans chaque langue	114
4.17	Erreurs phonogrammiques reliées aux accents dans chaque langue	116
4.18	Répartitions des sortes d'erreurs phonogrammiques dans chaque langue	118
4.19	Mots français dont les graphies erronées contiennent l'accent circonflexe	127
4.20	Erreurs lexicales dans les mots français contenant le graphème x à la finale	128
4.21	Erreurs lexicales dans les mots français contenant les digrammes grecs (th et ph)	129
4.22	Erreurs phonogrammiques liées à la double consonne en français	132

Résumé

Ce mémoire est une étude sur la maîtrise de l'orthographe lexicale du français et de l'espagnol. Des spécialistes de l'orthographe affirment que les scripteurs hispanophones sont plus compétents que les francophones. Ils ne s'appuient pourtant sur aucune étude empirique pour justifier leurs affirmations. Par cette étude qui compare la compétence orthographique de 42 élèves mexicains et de 42 élèves québécois de sixième année, nous avons essayé de vérifier la véracité de ces affirmations à l'aide de quatre dictées administrées à chaque groupe dans sa langue. Le deuxième but de cette étude est de voir si une réforme peut aider les francophones à être plus compétents dans l'orthographe de leur langue. Les différents sujets abordés dans le cadre théorique sont les graphèmes et les phonèmes, l'orthographe grammaticale et lexicale, l'apprentissage de l'orthographe lexicale, les échelles de fréquences, les projets de réforme, et la typologie des erreurs. Nous avons énoncé les hypothèses suivantes : Les hispanophones devraient présenter le plus haut taux de réussite (en lien avec les affirmations des spécialistes) ; les mots non maîtrisés se retrouveront parmi les mots peu fréquents (PF) (en lien avec les échelles de fréquence et les théories sur l'apprentissage de l'orthographe lexicale). Notre méthode repose sur le dénombrement de toutes les graphies erronées et une analyse graphique des types d'erreurs lexicales. Les résultats obtenus montrent une meilleure compétence des hispanophones sur le plan de l'orthographe grammaticale. Sur le plan de l'orthographe lexicale, nos résultats permettent de voir que les hispanophones produisent autant de graphies erronées que les francophones. Cependant, l'analyse des graphies montre une meilleure compétence lexicale des hispanophones. Cette étude cherchera également à savoir si certaines des réformes proposées du système orthographique français permettraient aux scripteurs francophones d'être plus compétents dans leur langue. Chez les hispanophones, il serait utile de poursuivre la réforme en s'attardant sur l'accent tonique.

Mots clés : Orthographe, réforme, lexicale, compétence.

Introduction

L'orthographe du français a parcouru un long itinéraire avant qu'elle ne devienne ce qu'elle est aujourd'hui. Quand nous suivons les grands moments de ce parcours, nous constatons que les réformes orthographiques ont été nombreuses pour que l'orthographe du français remplisse diverses fonctions comme la simplicité, la clarté, l'élégance et la fidélité à l'usage. Un travail qui a souvent été confié à « des humanistes imbus de latinité et soucieux de marquer fortement la filiation de notre vocabulaire du latin au français. » (Beslais, 1965, p.73).

Un grand nombre de réformes ont donc été appliquées alors que d'autres sont restées à l'état de projet. Aujourd'hui encore, devant la nécessité de rendre l'orthographe du français accessible à toutes les personnes qui utilisent le français comme langue première ou langue seconde, nous assistons dans toutes les sphères sociales à de grands débats sur la possibilité ou l'impossibilité de modifier l'orthographe du français, laquelle est sans contredit un lieu propice à la subjectivité comme l'exprime si bien Traimond :

Chacun dispose généralement sur ce point d'une conception très arrêtée de l'orthographe, sur sa dégradation, sur les causes de son irrespect. Chacun se sent capable de proposer des solutions. L'orthographe partage avec quelques autres thèmes, comme la religion musulmane, les « maths modernes » ou l'apprentissage de la lecture, le privilège de constituer un objet sur lequel quiconque a un avis (2001, p.14).

La pluralité des opinions révèle un fait indéniable, c'est le consensus entre les parties pour admettre que le français, à travers ces incohérences orthographiques, présente une orthographe des plus complexes. Citons comme exemple quelques incohérences extraites de *La nouvelle grammaire*

en tableaux de Marie-Éva de Villers : « *consonne* et *consonance*, *combattant* et *combatif*, *arôme* et *aromatique*, *affoler* et *folle* » (2003, p.19). L'ensemble nous donne donc un français ayant l'une des orthographe les plus opaques des langues d'origine latine. Selon Jaffré et Fayol (1997) une orthographe opaque (ou profonde) est « une orthographe dans laquelle la relation entre phonèmes et graphèmes est très irrégulière » (p.119). Par conséquent, l'opacité du français se traduit par le fait qu'il n'existe pas une relation biunivoque entre phonèmes et graphèmes à l'écrit et entre graphèmes et phonèmes à l'oral. Pour illustrer ce fait, prenons l'exemple du phonème /k/ qui peut être réalisé par différents graphèmes : *c* ; *k* ; *ch* ; *qu* ; *cc* ; etc. Un élève ou un apprenant ayant à transcrire un mot contenant le phonème /k/ (*amical*) se retrouve devant l'épineux problème de choisir entre ces différents graphèmes mis à sa disposition par le système orthographique du français si, à travers son apprentissage de l'orthographe, il ne maîtrise pas la graphie du mot en question. Dans le sens inverse, nous pouvons donner l'exemple du graphème **x**, « sa valeur de base est /ks/ : *extrême*, *lux*e, *lexique*, *sexuel* ; il prend fréquemment entre **e** et une autre voyelle, une valeur de position /gz/ : *examen*, *exécuter* [...] » (Blanche-Benveniste et Chervel, 1969, p.136). Le même phénomène se produit avec le graphème **t** qui en plus de sa valeur zéro dans certains mots (*petit***t**, *connait***t**) produit le son /s/ dans certains mots : *action*, *ineptie*, *inertie*, *éditions* en comparaison avec *portions* (nous), *éditions* (nous), *soutien*) (Ibid., p.337).

Ainsi, il devient facile de comprendre que les francophones et les apprenants du français langue seconde ne parviennent pas vraiment à maîtriser cette orthographe si opaque. Les cris d'alarme qui paraissent périodiquement dans les journaux, les enquêtes publiées sur la réussite et la situation de l'orthographe dans les écoles et la société sont là pour en témoigner. Prenons par exemple les écrits de nombreux chercheurs comme Guion

(1974) qui fait le procès de l'orthographe à travers son caractère élitiste ; Milot (1978) qui a présenté le résultat d'une enquête sur le niveau de l'orthographe dans les écoles du Québec et sur l'attitude des enseignants face à celle-ci ; Millet (1990) qui montre les opinions des professionnels de l'écriture et des élèves face à l'orthographe et la possibilité de la rectifier ; Roy et Lafontaine (1992) qui présentent les résultats d'une étude sur la réussite orthographique des élèves de différents niveaux scolaires ; et enfin, tout récemment, l'enquête de Manesse et Cogis (2007), en France, qui démontre une forte baisse du niveau orthographique des élèves entre 1987 et 2005.

En tenant compte de l'opacité de l'orthographe du français, il apparaît évident qu'après de longues années consacrées à l'enseignement/apprentissage de l'orthographe du français, bon nombre de scripteurs francophones produisent des graphies erronées ou hésitent à écrire un mot. Quand ils doivent composer un texte, ils doutent de l'orthographe lexicale d'un mot ou de l'accord d'un verbe ou d'un adjectif. Par conséquent, pour éviter la production d'une graphie erronée, ils doivent prendre le temps de chercher dans le dictionnaire ou dans la grammaire. La maîtrise de l'orthographe du français est à ce point inaccessible à tous qu'il arrive que des scripteurs aguerris reformulent une phrase par crainte de se tromper sur l'orthographe d'un mot. Il est facile de comprendre que l'activité d'écriture requiert toute l'attention du scripteur, car en orthographe, il semblerait que les acquis ne restent pas chez tous les utilisateurs du français. On retrouve à peu près les mêmes propos dans l'enquête de Millet :

L'orthographe réclame une attention de tous les instants, un travail assidu, et acharné qui n'aura jamais de fin puisque sa conquête n'est jamais achevée et que les acquis ne sont jamais définitivement acquis, alors même qu'on écrit tous les jours (1990, p.21).

Ce problème d'inaccessibilité ne concerne pas uniquement le français. L'anglais est également considéré comme une langue dont « l'apprentissage de l'orthographe est malaisé » (Arrivé, 1993, p.44). Cependant, « il semble que le français soit au premier rang des langues à orthographe complexe : il cumule en effet toutes les difficultés théoriquement imaginables. » (Ibid., p.44). Grevisse (1948) tient un discours tout aussi accablant :

Notre orthographe est donc difficile ; elle est savante, trop savante. Que ne s'est-elle libérée, comme l'a su faire celle des Italiens, des Espagnols, et des Portugais de certaines servitudes étymologiques, qui l'empêchent si fâcheusement de mieux se conformer à la prononciation (p.56).

À juste titre, nous retrouvons l'espagnol qui, malgré sa filiation latine comme celle du français, ne semble pas être sujet aux équivoques. En effet, dans la langue espagnole, le nombre de phonèmes ayant une seule graphie possible est beaucoup plus grand en comparaison au français, bien que les deux langues aient leur lot d'homophones (exemple : *saint, sain, sein, ceint*, en français ; *el/él, que/qué*¹, en espagnol).

Par conséquent, aux yeux de plusieurs personnes, l'orthographe de l'espagnol, contrairement à celle du français, se voit comme un système qui se rapproche de la transparence. Jaffré et Fayol (1997) définissent l'orthographe transparente (ou de surface) comme une « orthographe dans laquelle la relation entre phonèmes et graphèmes est très régulière » (p.119). Un principe de transparence qui est censé rendre l'orthographe de l'espagnol accessible à ceux dont c'est la langue maternelle et aux apprenants de l'espagnol langue seconde.

¹ En français : « le » et « il » ; « que » (pronom relatif ou conjonction) et « quel » (déterminant) ou « quoi »

En tenant compte de cette brève comparaison de deux systèmes orthographiques dans des langues de la même famille, nous nous demandons si, malgré tous les discours entendus et les ouvrages publiés sur la complexité de l'orthographe du français, la maîtrise de l'orthographe d'usage (lexicale) chez les jeunes francophones peut se comparer à celle de l'orthographe de l'espagnol chez les hispanophones. Cette question est à la base de ce mémoire.

Dans le premier chapitre, nous présentons la problématique en abordant le refus de toute modification à l'orthographe du français par les conservateurs de la norme orthographique. Nous montrerons que la maîtrise de l'orthographe du français est difficile et que malgré cette difficulté, bon nombre de francophones s'élèvent contre les projets de réforme alors que les réformateurs montrent la faisabilité d'une réforme en comparant le français avec des langues d'autres pays qui ont réformé leur orthographe. Enfin, nous montrerons que les discours sur la transparence et la facilité d'acquisition de tels systèmes orthographiques ne semblent pas s'appuyer sur des données empiriques, ce qui justifie d'autant plus la pertinence de notre recherche. Nous terminons le chapitre en présentant nos objectifs de recherche.

Dans le deuxième chapitre, nous regroupons les données théoriques importantes pour notre recherche. Nous parlons de la notion de phonème et de graphème, de l'orthographe lexicale versus l'orthographe grammaticale, du processus d'acquisition de l'orthographe lexicale, des problèmes posés par les homophones quant à leur enseignement et leur classement dans le système orthographique. Nous traitons également des échelles de fréquence qui montrent quels sont les mots que les élèves devraient apprendre les premiers, des points de l'orthographe du français à modifier selon quelques projets de réformes et de la typologie des erreurs. Finalement, nous

énonçons les questions de recherche auxquelles nous répondrons dans le cadre de cette étude, et les hypothèses à vérifier.

Dans le troisième chapitre, nous expliquons notre démarche méthodologique. Nous présentons le corpus et les élèves servant à notre étude comparative, la méthode d'analyse de nos résultats et les limites de notre recherche.

Dans le quatrième chapitre, nous présentons nos résultats et leur analyse, en lien avec nos questions de recherche :

- Nous comparons d'abord les performances générales d'élèves francophones et hispanophones en orthographe et plus particulièrement en orthographe lexicale ;
- Nous vérifions le rôle que joue la fréquence des mots dans l'acquisition de leur orthographe lexicale ;
- Nous analysons ensuite, à l'aide de la typologie des erreurs, les types d'erreur lexicale dans les deux langues ;
- Nous terminons en évaluant l'impact des projets de réforme sur la maîtrise de l'orthographe lexicale des francophones.

Dans la conclusion générale de notre étude, nous fournirons des pistes qui peuvent compléter cette recherche.

Chapitre I

La Problématique

1.1 Introduction

Cette problématique vise essentiellement à démontrer la pertinence de nos objectifs de recherche. Dans un premier temps, nous traiterons de l'ensemble des points qui reviennent sur le devant de la scène quand il est question de réforme de l'orthographe : les deux clans en présence, soit les conservateurs et les réformateurs. Nous abordons également la difficile acquisition de l'orthographe du français avant de présenter le mouvement de résistance aux changements. Dans un deuxième temps, il sera question des spécialistes de l'orthographe qui, dans leur appui à une réforme nécessaire, comparent l'orthographe du français avec celles d'autres pays qui ont procédé à une réforme de leur système orthographique. Nous verrons toutefois que leurs conclusions ne semblent pas s'appuyer sur des données empiriques, elles sont liées à l'analyse des systèmes orthographiques. Enfin, nous énoncerons nos objectifs de recherche.

1.2 Les deux clans en présence : les conservateurs et les réformateurs

L'orthographe du français se retrouve souvent au centre de plusieurs guerres des mots au sujet de sa complexité et de l'urgence ou non de procéder à la restructuration du système. En effet, dès qu'il s'agit de réforme orthographique, le ton monte entre les deux camps, celui des conservateurs et celui des réformateurs. L'exemple de confrontation le plus récent est la bataille médiatique générée par le travail du Conseil Supérieur de la Langue Française (CSLF) lors de ses propositions de rectification en 1990. « Les menaces de mort ? Elles ont été littéralement [...] proférées par ceux qui, naguère, ont menacé les « profanateurs » de la guillotine » (Arrivé, 1993,

p.73). Comme l'a démontré Guion (1974) dans une étude sur l'orthographe, la place qu'elle occupe dans la société et dans les écoles, les arguments pour le maintien de l'orthographe du français touchent d'abord à la fonction d'enculturation de celle-ci. D'ailleurs, le collectif *Contre la réforme de l'orthographe : va-t-on se laisser mordre la langue ?* (Bluau et al, 1990) est un bel exemple de bataille d'opinions où les interlocuteurs, conservateurs d'une norme sociale, montrent leur grand amour pour l'orthographe de la langue à travers toutes ses incohérences. Prenons par exemple les commentaires suivants :

Il y a des bizarreries en français parfois pourquoi une langue n'aurait-elle pas droit à l'étrangeté, c'est tellement plus drôle que les normalisations (Jacques Gaucheron, dans Bluau et al. 1990, p.47) ;

Pour ma part, incapable de clouer au pilori ceux qui veulent la réformer puisque de toute façon, elle changera, je continuerai à utiliser les singularités, les bizarreries, les étrangetés que j'ai eu tant de mal à apprendre. Et je continuerai, naturellement à faire de ces fautes délicieuses qui sont pour beaucoup dans les charmes de la grammaire (Jean D'ormesson, dans Bluau et al. 1990, p.91).

Dans l'ensemble, il est facile de constater que ces citations proviennent de la plume de ceux (écrivains, enseignants, journalistes) qui maîtrisent (ou croient maîtriser) cette orthographe.

Dès son instauration comme matière scolaire dans les écoles, l'orthographe a toujours occupé une place de choix dans le paysage de la France (Chervel, 1989). En effet, une fois rendue obligatoire dans les écoles françaises, elle a longtemps servi de critère de sélection. Blanche-Benveniste et Chervel (1969) nous disent que « la loi se fait sélective : la faute d'orthographe est pénalisée durement au certificat d'études et dans de nombreux examens qui ouvrent à des fonctions subalternes [...] » (p.95). Il devient facile d'imaginer, pour ne pas dire comprendre, le traumatisme vécu, dans le passé, par ceux

qui n'ont pas réussi aux évaluations finales (une dictée). Un échec qui attire le regard réprobateur des autres (parents et amis) et bien sûr celui du maître sur le finissant, car comme le dit si bien Chervel (1989) : « L'élève qui veut s'identifier à son maître ne peut se permettre d'être mauvais en orthographe. Être bon élève, c'est forcément être bon en orthographe [...] » (p.141).

Ce rôle discriminatoire de l'orthographe est bien décrié par des spécialistes de l'orthographe. Blanche-Benveniste et Chervel (1969) nous disent que « l'écriture reste un privilège de classe » (p.215). Pour Millet (1990), elle devient, à travers le temps, « langue, culture, histoire, pensée, barrière sociale [...] » (p.14). Selon Portebois (1998), « le dictionnaire est fait par des lettrés pour des lettrés, qui tous connaissent l'orthographe, qui créent également l'une ou l'autre tout en s'y soumettant » (p.78).

Cependant, nous remarquons que le mouvement de protection de l'orthographe tend à se généraliser, car outre cette élite qui lève la voix pour protester et freiner toute tentative de réforme de l'orthographe du français, de nombreux citoyens francophones se lancent à sa défense. Pour expliquer les bases de notre étude, nous allons montrer que l'objet tant défendu est difficile à maîtriser, que le refus global des conservateurs repose souvent sur des sentiments égoïstes (ou ethnocentristes) (Goosse, 1991) que d'autres pays ont procédé à une réforme de leur système orthographique, changement qui selon certains spécialistes de l'orthographe rendent les scripteurs de ces pays plus performants dans l'écriture de leur langue.

1.3 La difficile maîtrise de l'orthographe du français

Selon Blanche-Benveniste et Chervel (1969),

La science de l'orthographe est une science qui s'acquiert mot par mot et peut être sous cet angle assimilée à l'acquisition d'une langue

seconde. Depuis le cours élémentaire [...] c'est sept ans ou même huit qui sont consacrés à cette étude [...] (p.12).

Malheureusement, selon Brissaud et Bessonat (2001), « à la fin de l'école élémentaire, l'élève n'a parcouru que la moitié du chemin qui le conduit à une maîtrise satisfaisante de l'orthographe » (p.59).

Quand les gens relatent les souvenirs de leur apprentissage de l'orthographe dans les résultats de l'enquête de Millet (1990), nous retrouvons cette même idée d'une traversée interminable : « On avait un certain nombre de dictées, de mots à apprendre par cœur, de règles, d'exercices. » ; « J'ai peiné, j'ai recommencé des tas de choses, j'ai dû refaire plusieurs fois, bon, et puis à la longue, c'est rentré » (p.27).

Cette étude n'est jamais achevée parce que les élèves doivent parfaire leur apprentissage dans les classes supérieures. En effet, leur étude, en rapport à la maîtrise du français écrit (orthographe lexicale et grammaticale), Roy & Lafontaine (1992) constatent « une progression générale de 19,84 % à survenir de la deuxième secondaire à l'université. Ceci représente une progression de 3,3 % par année de scolarité. » (p.72). Les chercheurs nous mentionnent également les limites de ces résultats, « car, du secondaire à l'université, disparaissent ou se retirent du circuit de la scolarisation les étudiants les plus faibles » (Ibid., p.79). Ils ajoutent que pour savoir s'il y a vraiment eu amélioration, « il faudrait évaluer la maîtrise du français écrit chez les mêmes sujets, avec un même instrument et à quelques années de distance » (Ibid., p.80).

En plus des difficultés liées à l'orthographe d'usage, difficultés tributaires en grande partie de l'écart existant entre phonèmes et graphèmes sur le plan

lexical, s'ajoute la difficulté de l'orthographe grammaticale (plus précisément la marque silencieuse du pluriel des mots). À cet effet, Guyon (2003) nous dit que :

Le scripteur doit en effet effectuer des calculs presque sur chaque mot pour poser les marques de l'écrit, et de plus ce calcul nécessite des opérations complexes puisqu'il faut non seulement penser à poser ces marques, mais aussi choisir la seule marque juste (p.90).

Comme nous venons de le voir brièvement, l'apprentissage de l'orthographe du français, à travers les nombreuses difficultés sur le plan lexical et grammatical, demande beaucoup de temps. Et au grand dam des enseignants et des apprenants, certains mots maîtrisés se perdent au fil des années scolaires. Selon Brissaud et Bessonat (2001), rien n'indique dans le temps une amélioration sur le plan orthographique pour certains élèves à cause des nombreuses complexités de cette orthographe. Mentionnons à cet effet l'étude de Manesse et Cogis (2007), qui prouve bien que dans le temps, la compétence orthographique ne semble pas s'améliorer. Leur étude sur la maîtrise de l'orthographe, menée en France, montre en effet que les élèves produisent plus de graphies erronées que dans le passé.

Les scores sont beaucoup plus élevés en 2005 qu'ils ne l'étaient en 1987. le score moyen en 2005 est de 27 points, soit plus de 13 fautes « lourdes » ; il était de 16 points en 1987, soit 8 fautes. C'est dire que les élèves d'aujourd'hui font beaucoup plus de fautes qu'il y a vingt ans, et ceci, à chaque niveau de la scolarité. (Manesse et Cogis, 2007, p. 80).

Ainsi, comme le dit si bien un des répondants à l'enquête de Millet (1990), c'est « un travail journalier, régulier, et quotidien... sans arrêt faire quelque chose en orthographe... c'est jamais fini » (p.22). Mentionnons que cette enquête a été effectuée dans la région de Grenoble au cours des années 1987-1989 auprès de personnes qui ont un rapport quotidien avec l'orthographe. Il s'agit d'institutrices, de professeurs de lycées et de collèges,

de secrétaires et de personnes qui travaillent dans le métier du livre (éditeurs, librairies, imprimeurs). Le total des interviewés s'élève à 92 personnes (24 adultes, 49 enfants et 19 adolescents). En résumé, nous pouvons donc retenir de cette étude que « ce travail, ces efforts sont, au bout du compte, très mal récompensés puisqu'en définitive, l'orthographe se dérobe ; pour quelque raison que ce soit, on finit toujours par l'oublier. » (Ibid., p.22).

Nous venons de voir que malgré tous les efforts et le temps accordé à l'orthographe du français, celle-ci demeure difficile à maîtriser. Nous développerons cette problématique davantage dans le cadre théorique. Maintenant, il faut se demander quelles sont les voix qui résistent tant aux changements du système orthographique du français ?

1.4 La résistance aux changements

Comme nous l'avons déjà mentionné, nous assistons toujours à une guerre quand il s'agit de projets de réforme de l'orthographe. Donc avant de montrer la provenance de cette résistance, nous pensons qu'il importe de montrer que l'orthographe a, tout au long de son histoire, subi des modifications, souvent ignorées par la grande majorité des francophones.

Le premier document en langue française est, selon plusieurs ouvrages consultés, le *Serment de Strasbourg* qui date du IX^e siècle. À la lecture de ce texte, il est facile à tout lecteur de constater que l'orthographe du français a subi de profondes modifications pour présenter le visage que tout le monde connaît aujourd'hui. Des modifications qui ne sont pas les fruits du hasard. Et comme le dit si bien Gruaz (1999, p.61), « l'orthographe contemporaine est le résultat de débats et de décisions humaines prises au cours des siècles. » En effet, depuis le Moyen Age, l'orthographe a été manipulée, bricolée, ajustée

par des académiciens. Des modifications qu'il est possible de suivre à travers les différents dictionnaires de l'Académie française :

Les éditions de 1740 et 1762 remplacent des « s » par des accents circonflexes (cf. fenêtre). Des lettres grecques sont supprimées (« phanion, scholarité » deviennent « fanion, scolarité »), la graphie « oi » est remplacée par « ai » dans Anglais, Français, etc. L'édition de 1835, au contraire, revient parfois à plus de complication (amidgale est remplacé par amygdale). L'édition de 1878 simplifie de nouveau, contrairement à celle de 1935, qui marque dans bien des cas un retour à plus de complexité (rapsodie redevient rhapsodie). (Millet, 1990, p.75-76).

À travers les années qui ont suivi les dernières grandes rectifications de 1935, les changements opérés sur l'orthographe du français comme outil de transcription de l'oral ont été minimes. Les projets de réforme ont toujours démontré cette volonté de simplifier. Pourtant, malgré toutes les difficultés liées à l'apprentissage de l'orthographe, la résistance à toute réforme provient d'un groupe en particulier. Blanche-Benveniste et Chervel (1969) nous font le portrait du résistant dans ces propos :

L'homme qui domine, plus ou moins bien d'ailleurs, une technique délicate. Il l'a acquise, enfant, sous la férule du magister ou du régent de collège qui lui enseignaient le latin et donc les bases latines de l'orthographe, ou par l'exercice quotidien et fastidieux de nos modernes dictées (p.11).

Maintenant, pour cette classe « élitiste », accepter toute modification, ce serait renier le conditionnement de tant d'années d'effort ; surtout que selon Klinkenberg (1984), « l'orthographe est lâchement liée à la culture (et moins encore à l'intelligence), et la révérence qu'on a pour elle tend à montrer que son maniement aisé constitue aux yeux du public un préjugé social favorable » (p.560).

De plus, il semblerait qu'une pensée égoïste règne dans l'esprit de celui qui se bat contre les changements. À ses yeux, la conservation de la norme orthographique est obligatoire parce que « les peines, voire les souffrances de son jeune âge sont pour lui du passé, même si elles sont le présent de son fils ou de sa fille. Il est dans l'ordre des choses que les jeunes passent par où je suis passé. » (Blanche-Benveniste et Chervel, 1969, p.11). La roue doit continuer de tourner.

Qu'il utilise l'écriture régulièrement ou sporadiquement, chacun veut protéger l'orthographe de la langue. Blanche-Benveniste et Chervel (1969) expliquent également cette résistance aux changements en ces termes :

Le Français moyen est persuadé que l'orthographe, c'est la grammaire, ou même la langue ; il est donc, par patriotisme, naturellement hostile à toute modification. Et de fait, aucun projet de réforme n'a jamais, en France, dépassé le stade des vœux pieux (p.207).

La France n'est pas le seul pays où nous assistons à un front commun contre les projets de réforme. Ici, au Québec où l'on retrouve la plus grande concentration de francophones au Canada, il n'y a pas eu de montée aux barricades quand les rectifications ont été proposées par le Conseil Supérieur de la Langue Française en 1990, mais il y a eu des prises de position claire de la part du public et des gens de lettres. À titre d'exemple, nous prenons quelques exemples de réactions autour du travail du CSLF. Citons à ce propos Roger Lemelin, écrivain québécois :

Je suis farouchement contre. On enlève l'accent circonflexe sur août, après on enlèvera le « a », qui ne se prononce pas non plus, et le mot sera *out* ! Pour les gens simples, c'est comme si on voulait tuer leur mère. Je pense que des concessions sur une langue, ça ne se fait pas. (La Presse, 13/01/91).

Citons également Clément Marchand, poète, à propos de la réforme :

Le français, dans sa constitution actuelle, correspond à un des plus grands triomphes de l'esprit humain. Il ne faudrait surtout pas attenter à ce chef-d'œuvre, qui a atteint la pleine maturité de son sens, pour satisfaire des esprits un peu réducteurs. Ces réformes que l'on dit mineures, elles sont à rejeter tout à fait. » (Ibid., 13/01/91).

Les exemples ne manquent pas, mais il convient de préciser que c'est toujours le même scénario quand il s'agit de contrer tout projet de réforme. Nous assistons à une bataille d'opinions dans les médias où nous retrouvons surtout des gens de lettres et non des scientifiques spécialistes de l'orthographe, et où les conservateurs reconnaissent malgré tout l'imperfection du système orthographique. Une résistance qui trouve sa source dans la condescendance de tout français à l'égard de l'orthographe. Mais comme le dit si bien Millet (1990), « un outil qui glisse perpétuellement entre les doigts n'est certainement pas un bon outil, mais à l'orthographe, on le lui pardonne et l'on dira, soulagé, que personne n'est infailible » (p.23). La maîtrise de l'orthographe est, comme nous venons de le voir, difficile. Pourtant, la résistance aux changements est forte. Maintenant, regardons ce que disent et démontrent les spécialistes de l'orthographe, favorables à l'idée de procéder à une réforme du système orthographique du français.

1.5 La réforme de l'orthographe française vue par des spécialistes

La faisabilité d'une réforme de l'orthographe du français passe souvent par la comparaison entre les systèmes orthographiques d'autres pays. Beslais (1965) énonce l'exemple d'autres pays réformateurs en présentant son *Rapport général sur les modalités d'une simplification éventuelle de l'orthographe française* : « Plusieurs pays ont procédé à une réforme de leur orthographe. L'Espagne, qui avait depuis longtemps ouvert la voie, a résolu la plupart des problèmes qui nous préoccupent, consonnes doublées et

lettres grecques en particulier » (p.17). Les Danois, dont l'orthographe est comme le français (selon Catach 1980) un plurisystème, ont procédé à des modifications. « Le *ph* a partout été remplacé par *f*, et *th* par *t*, comme dans *ortografi* et *filosofi*, comme c'est aussi le cas en italien. » (Gruaz et al., 2001, p.301). Nous retrouvons les mêmes pays en exemple chez Catach (1989, p.324).

Ce parallèle est repris par d'autres spécialistes qui traitent de la nécessité de réformer l'orthographe du français. Masson (1991), un des membres du CSLF, nous dit que : « dès l'âge de sept ou huit ans, les Italiens connaissent tout de l'orthographe de leur langue [...] » (p. 17) ; selon Jaffré et Fayol (1997), « les jeunes Espagnols peuvent très précocement écrire de nombreux mots qu'ils n'ont jamais rencontrés auparavant, exploitant pour cela les régularités du système. » (p.77). Toujours pour montrer la nécessité d'une réforme, Masson (1991) précise que :

Pour notre langue, le concours (d'orthographe) n'est possible que parce que seule une minorité (petite) de gens maîtrise complètement l'orthographe, alors que tout le monde devrait pouvoir utiliser l'orthographe (écriture) aussi facilement que les Italiens, les Espagnols, ou les Allemands [...] (p.106).

Ces phrases disent l'essentiel. Cependant, nous constatons dans la plupart des cas qu'il s'agit d'affirmations logiques de la part de ces spécialistes de la langue, basées sur les systèmes orthographiques, sans aucune référence à des données empiriques provenant des scripteurs des pays mentionnés. Les auteurs ne fournissent pas de chiffres comparatifs permettant de voir effectivement que l'orthographe des autres langues d'origine latine, dont le système orthographique est jugé « transparent » comme l'espagnol, s'acquiert plus rapidement que celle du français.

Cet avis de spécialistes, tous francophones, ne semble pourtant pas être nécessairement partagé par les spécialistes hispanophones. Elena Pombo Llamas (1999), linguiste espagnole qui a également publié plusieurs ouvrages sur l'évolution du français au fil des siècles, affirme, comme les auteurs francophones cités ci-dessus, que « les conventions graphiques de l'espagnol sont proches de l'idéal de simplicité orthographique : une lettre représente un seul son et chaque son est toujours représenté par la même lettre » (Pombo, 1999, p.85). Pourtant, elle met un bémol sur cette vision de simplicité.

En effet, il semblerait que tous les hispanophones rencontrent la plupart des difficultés dont on parle en orthographe française parce qu'il n'y a pas « de correspondance totale entre sons et lettres dans l'orthographe espagnole, car, au cours de son histoire, le critère étymologique a été maintenu, malgré le critère dominant à l'oral. » (Ibid., p.86) La question que nous nous posons maintenant est la suivante : existe-t-il des études comparatives de la compétence orthographique des francophones et des hispanophones qui permettent à des spécialistes comme Jaffré et Fayol (1997) ou Masson (1991) d'affirmer que les hispanophones et les italophones sont plus compétents en orthographe que les francophones ?

Pour répondre à cette question, mentionnons que pendant notre recherche la seule étude comparative trouvée sur le sujet est celle de Vaca Uribe (2003). Il a effectué « une étude transversale et comparative entre enfants francophones et hispanophones d'âges différents (7-8 ans vs 9-10 ans) qui justifient leur écriture de diverses paires de mots homophones » (p.41). Les résultats ne montrent pas une supériorité des hispanophones face aux francophones quant à la capacité à différencier les homophones. Dans la conclusion, le chercheur affirme que :

Les enfants arrivent à être très systématiques dans l'application du principe de différenciation : les enfants trouvent, ou bien inventent, les moyens de différenciation, pour être cohérents avec un principe plérémique (logographique ou morphographique) (p.52).

Conclusion

Nous avons montré que l'orthographe du français est l'une des plus complexes des langues d'origine latine. Par la même occasion, nous avons conclu que la complexité et l'opacité de celle-ci nuisaient à son enseignement/apprentissage. Il faut compter de nombreuses années d'études et de pratiques constantes pour devenir un scripteur expert. C'est l'une des raisons pour laquelle nombre de spécialistes ne cessent de militer en faveur d'une réforme de l'orthographe du français. Devant la résistance globale aux changements d'un grand nombre de francophones, il arrive souvent que ces spécialistes présentent l'exemple d'autres pays réformateurs comme l'Espagne, l'Italie, la Roumanie, la Turquie, etc. pour appuyer leurs positions. Certains mettent de l'avant la nécessité de procéder à ces changements dans un but scolaire.

Cependant, nous avons constaté qu'en affirmant que les scripteurs des pays comme l'Espagne et l'Italie qui ont procédé à une réforme profonde de leur système orthographique sont plus compétents orthographiquement que les francophones, Masson (1991), Jaffré et Fayol (1997) ne présentent pas de chiffres sur les compétences de ces scripteurs ni aucune étude comparative sur le sujet. Devant le manque de données empiriques, nous croyons à l'importance de procéder à une telle étude sur la compétence orthographique des jeunes francophones et hispanophones du même âge pour voir si la dite transparence de l'orthographe de l'espagnol versus l'opacité du système graphique français constitue bien un avantage dans l'acquisition de l'orthographe en général (lexicale et grammaticale), d'autant plus que des

spécialistes hispanophones font état de difficultés d'acquisition du système graphique chez les scripteurs de cette langue.

1.6 Les objectifs de recherche

Le premier objectif de cette recherche est donc de comparer de jeunes francophones et hispanophones sur le plan de l'orthographe. Nous pourrions ainsi vérifier s'il existe un écart considérable sur le plan de la maîtrise de l'orthographe chez les enfants de 12 ans selon la transparence ou non du système.

Étant donné que toutes les réformes touchent exclusivement l'orthographe lexicale, nous nous pencherons plus particulièrement, sur les habiletés en orthographe lexicale des deux groupes comparés. Ainsi, nous verrons si le système orthographique du français gagnerait à être moins opaque et moins complexe ou si son apprentissage se situe au même niveau que les langues plus proches de la transparence (l'italien, l'espagnol, l'allemand, le roumain, etc.) C'est là le principal objectif de ce mémoire.

Le deuxième objectif est de questionner les rectifications de 1990 proposées par le Conseil Supérieur de la Langue Française par rapport à d'autres projets de réforme : Dauzat, Albert (1939-1952) ; Commission Beslais (1960-1965) ; Thimonnier, René (1967) ; et de voir leur impact potentiel sur la compétence orthographique des élèves.

Chapitre II

Le cadre théorique

2.1 Introduction

Le cadre théorique contient les notions essentielles à notre étude sur la maîtrise de l'orthographe. En somme, il s'agit de présenter les notions qui servent de base à notre étude comparative de l'espagnol et du français. Nous allons donc présenter les notions de phonème et de graphème, notions fondamentales dans l'étude d'un système orthographique. Nous cernerons ensuite la différence entre l'orthographe grammaticale et lexicale qui génère deux grandes catégories d'erreurs. Comme notre analyse des résultats se penche plus spécifiquement sur l'orthographe lexicale, nous devons expliquer les processus mis en œuvre pour maîtriser l'orthographe lexicale. Nous traiterons également des homophones pour mieux les situer dans le système orthographique du français et dans le traitement que nous en ferons dans cette étude. Nous présenterons diverses échelles de fréquence des mots puisque ce facteur est susceptible d'intervenir dans la maîtrise du système ainsi que quelques typologies d'erreurs orthographiques qui nous permettront de peaufiner notre analyse. Enfin, nous décrirons quelques projets de réforme afin de pouvoir mesurer l'impact qu'auraient ces projets sur la performance d'élèves francophones.

2.2 Les notions de phonème et de graphème

Nous présentons dans cette section une définition de certains concepts qui relèvent de la linguistique. Il s'agit des notions de phonème et de graphème et de leurs sous-catégories, lesquelles seront utiles pendant notre analyse linguistique de certains faits relevés dans le corpus.

Le phonème est, selon Catach (1986) « la plus petite unité sonore dépourvue de sens, mais qui a une fonction distinctive de la chaîne orale. Ensemble de sons reconnu par l'auditeur d'une même langue comme différent d'autres ensembles associés à d'autres phonèmes » (p.16). En un mot, les phonèmes ne sont pas décomposables en de plus petites unités. Exemple : les phonèmes /p/, /a/, /p/, /a/ dans le mot « papa ».

Le nombre de phonèmes pour le français diffère d'un ouvrage à un autre. Nous en comptons 40 dans le Petit Robert (1993), dont deux phonèmes pour les mots empruntés à l'anglais, à l'espagnol et à l'arabe (**camping**, **jota**, **khamsin**) ; 38 dans De Villers (2003) et Catach (1980), 37 dans Grevisse (1997) et Gey (1987) ; et 36 dans Chartrand (1999). D'après la moyenne du nombre relevé dans ces ouvrages (consultés pour comparaison), il existe en français 37 phonèmes pour les 26 lettres de l'alphabet, alors qu'on « dénombre au total 23 ou 24 phonèmes en espagnol [...] » pour un total de 32 graphèmes (Pottier & al, 2000, p.28). Nous verrons plus tard dans notre étude comparative des habiletés orthographiques si cette différence entre les deux langues peut aider les élèves hispanophones à être plus compétents dans la discrimination des sons permettant de bien orthographier un mot.

Le graphème est la notion la plus importante pour cette étude à cause de son importance dans la transcription et la maîtrise de l'orthographe. Se basant sur les travaux de Catach, Nadeau et Fisher (2006) expliquent que les graphèmes permettent de :

- Noter des phonèmes (ex. : le *t* dans *terre*, *atome*, *strict*) ;
- Noter des relations lexicales ou des catégories grammaticales :
 - dérivés et familles de mots (ex. : le *t* de *petit*, le *b* de *plomb*) ;
 - catégories de genre, nombre, personne, temps, mode, etc.
- Donner un visage unique à des mots, spécialement aux homophones (ex. : *mes/mais* ; *quand/quant*) ;

- Constituer des lettres étymologiques (ex. : le *p* de *baptême*) ou historiques (ex. : le *z* de *nez*), non fonctionnelles. (2006, p.160).

Le graphème est selon Catach (1986) :

La plus petite unité distinctive et/ou significative de la chaîne écrite, composée d'une lettre, d'un groupe de lettres (diagramme : on, an in, trigramme : ein, eau, ain), d'une lettre accentuée ou pourvue d'un signe auxiliaire, ayant une référence phonique et/ou sémique dans la chaîne parlée. Ex. : *p, ou, r, ch, a, ss, e, r*, dans *pourchasser* » (p.16).

Par contre, la notion de graphème en français n'est pas aussi simple qu'on pourrait le croire. Tout comme les phonèmes, il arrive souvent que le nombre de graphèmes diffère d'un auteur à un autre. Arrivé (1993, p.47) en trouve 175, Chartrand et al. (1999, p.16) en présentent 133, Catach (1986, p.14) en dénombre +/- 130. Dans une tentative d'exhaustivité, Giasson et Thériault (1983, p.235) présentent un total de plus de 500 graphèmes. Cette différence à l'intérieur de ces ouvrages peut s'expliquer par la reconnaissance de l'insuffisance du système graphique du français à rendre compte de tous les sons, et par le travail de découpage fait pour représenter le plus fidèlement possible ces différents sons que l'on retrouve en français. Ce travail de découpage inclut ou non dans la liste des graphèmes les marques silencieuses du pluriel ou des consonnes muettes qui servent au rapprochement avec la famille des mots. Ainsi, Giasson et Thériault (1983) incluent tous ces aspects : par exemple, pour le phonème /o/ nous retrouvons entre autres des graphèmes comme *aud, auds, ot, ots, etc.*. Précisons que pour le phonème /o/ nous comptons 30 graphèmes chez ces auteurs alors que Chartrand en donne seulement 5. La liste de Catach (1986) se limite à 45 *graphèmes de base* qui montrent un plus haut degré de fréquence, de pertinence et de stabilité. À ce nombre, elle ajoute d'autres graphèmes isolés que l'on peut retrouver dans le système graphique du français, ce qui explique les 130 graphèmes comme unités de base de son tableau. Prenons

le phonème /o/. Dans le tableau de Catach, nous retrouvons 3 graphèmes dits de base (*o*, *au*, *eau*) et deux autres graphèmes « ô » et « üm » pour le phonème /ɔ/ que nous pouvons rencontrer dans le mot *capharnaüm*. Mentionnons également que dans ce tableau, apparaît une section sous-graphèmes, ce qui semble expliquer la différence de découpage retrouvée chez les auteurs cités ci-dessus. Dans le cadre de cette étude, nous considérons, dans l'éventualité d'une analyse graphémique des erreurs orthographiques, le tableau de Chartrand et al. (1999) qui présente 133 graphèmes. Notre choix repose sur le fait que ce nombre se rapproche de celui de Catach (1986), notre principale référence, que cette grammaire est plus récente et qu'elle circule dans le milieu scolaire du Québec.

Catach (1986) divise les graphèmes en trois catégories. Il s'agit des phonogrammes auxquels elle ajoute la notion d'archigraphème, des morphogrammes et des logogrammes. Notre référence de base pour cette section demeure Catach (1986) comme nous l'avons déjà précisé. Rien de plus normal devant l'ampleur de son travail et le fait que plusieurs auteurs la citent dans leurs ouvrages. Nous pouvons nommer à titre d'exemple Gey (1987) qui reprend textuellement les définitions de Catach. La raison de la présence d'Arrivé (1993) dans cette section repose sur le fait que dans notre recherche nous avons toujours cherché la définition qui nous paraissait la plus claire, et pour nous, et pour les profanes de cette science qu'est la linguistique. C'est ce travail de vulgarisation que, justement, Arrivé (1993) a tenté de faire dans son ouvrage.

- Les phonogrammes sont des graphèmes chargés de transcrire les phonèmes. Les phonogrammes comprennent les archigraphèmes et leurs variantes positionnelles ; leur usage est réglé par les lois de position. Exemple : « g » dans *gare*, « gu » dans *guérir*. (p.16) Il est certain que nous

retrouvons la même définition chez différents auteurs comme Gey (1987), mais pour utiliser un autre langage que nous considérons plus clair, Arrivé nous dit que les phonogrammes « ont pour fonction de noter dans l'orthographe un phonème de la manifestation orale. Exemple : dans le mot « raison », le phonogramme « on » correspond au phonème /õ/ » (1993, p.46). L'archigraphème, selon Catach (1986), c'est un « graphème fondamental, représentant d'un ensemble de graphèmes, qui sont par rapport aux autres ensembles dans un rapport exclusif, correspondant au même phonème ou au même archiphonème. Ex. : O pour o, ô, au, eau, etc. » (p.17). Il s'agit pour Catach (1978, p.62) de 33 « formes maximales stables » que l'on peut retrouver dans toutes les positions, en toute place du mot. Avant de poursuivre, il importe de faire remarquer que cette notion d'archigraphème de Catach est critiquée par Arrivé (1993) qui nous dit que cette définition « semble implicitement exclure de l'inventaire des graphèmes les morphogrammes et les logogrammes qui sont pourtant visés par la définition du graphème[...] » (p.45). Dans le camp opposé, nous retrouvons Gey (1987) qui trouve que « cette notion est très utile en pédagogie, car elle permet de procéder par secteurs à l'intérieur desquels sont regroupés avec cohérence les différents problèmes posés par une même zone orthographique » (p.9).

- Les morphogrammes, selon Arrivé (1993), « marquent le genre et le nombre, [...] donnent des indications sur des catégories morphologiques passées sous silence, au sens le plus strict de l'expression, dans la manifestation orale » (p.50). Cette définition est assez claire. Cependant, les marques du nombre en français ne se prononcent pas à la fin d'un mot, mais dans cette définition, rien ne nous indique s'il faut les classer parmi les morphogrammes. De plus, dans le mot *marchand*, est-ce que le « d » est un morphogramme ?

Pour répondre à cette question, nous prenons quelques exemples de Catach (1986) : dans le mot **jolis** (« s » est un morphogramme du pluriel), dans le mot **blanc** (« c » est le morphogramme servant au rapprochement avec la famille de mots : **blanche**, **blancheur**, **blanchâtre**, **blanchir**, etc.). Catach (1986) distingue les morphogrammes grammaticaux des morphogrammes lexicaux. Sous la bannière morphogramme grammatical, nous retrouvons « les désinences supplémentaires qui s'ajoutent aux mots selon les rencontres des parties du discours (marques de genre et de nombre, flexions verbales) » (p.211). Donc dans notre exemple « **jolis** », le graphème « s » est un morphogramme grammatical tout comme le « ons » de *nous pensons*. Les morphogrammes lexicaux sont les « marques finales ou internes *fixes*, intégrées au lexème, pour établir un lien visuel avec le féminin ou les dérivés ; marques spécifiques des préfixes, des suffixes, des éléments entrant en composition, etc. » (Ibid., p.211). Ainsi, dans le mot **blanc** de notre exemple, le « c » représente une marque finale fixe qui permet de former des mots de la même famille, il est selon la définition de Catach (1986), un morphogramme lexical. Des préfixes ou des suffixes peuvent également correspondre à un morphogramme lexical. Par exemple, à propos d'un scripteur qui écrirait ainsi le mot **anterrement**, nous pouvons parler d'ignorance du préfixe « en » (comme dans **enfermer**, **encadrer**), une graphie essentielle à la reconnaissance du mot et de ses liens avec le verbe « enterrer ». C'est un cas de préfixe correspondant à un morphogramme lexical. Ainsi, le « d » final du mot *marchand* est un morphogramme lexical.

- Les logogrammes, sont, selon Catach (1986), des « notations de lexèmes ou « figures de mots », dans lesquels, à la limite, la « graphie » ne fait qu'un avec le mot, dont on ne peut la dissocier. La principale fonction des logogrammes est la distinction des homophones » (p.17). Arrivé (1993) nous

dit que les logogrammes « ont pour fonction de distinguer entre eux, dans l'orthographe, des unités linguistiques, le plus souvent des mots qui se confondent dans la manifestation orale. » Exemples : « Cru » de croire et « crû » de croître ; « faim » et « fin » (p.53). Comme nous l'avons déjà mentionné, la présence de cet auteur dans cette section sert à rendre le discours plus clair, car, outre la notion d'archigraphème, les auteurs (Arrivé et Catach) s'entendent sur les points que nous venons de présenter.

2.3 L'orthographe grammaticale versus l'orthographe lexicale

L'orthographe grammaticale concerne les marques d'accord : accord du verbe avec son sujet, accord en genre et en nombre des noms, adjectifs, déterminants, accord du participe passé. « Ainsi dans *Sa réponse est claire*, la forme *clair* relève de l'orthographe d'usage, tandis que la présence du *-e* féminin déterminée par l'accord de l'attribut avec le sujet ressortit à l'orthographe grammaticale. » Simard (1995, p.145). Par conséquent, quand nous parlons d'orthographe grammaticale, il s'agit de tout graphème qui, en situation d'écriture, demande à l'apprenant l'application d'une règle établissant une relation entre les mots d'une phrase, ce qui veut dire qu'il se retrouve dans l'obligation de recourir à ses connaissances déclaratives et procédurales de la grammaire pour effectuer l'accord requis (Nadeau, 1995). Ce travail d'accord demeure tout de même très complexe pour les scripteurs ou apprentis-scripteurs. Brissaud et Bessonat nous expliquent les raisons qui rendent l'orthographe grammaticale si compliquée pour les apprenants :

Aujourd'hui la grande difficulté, qui rend l'orthographe du français particulièrement opaque, réside dans le fait que sa morphologie est largement silencieuse : les marques ne s'entendent pas. Il s'ensuit que le scripteur privé du secours de l'oral pour décider de la distinction des lettres morphologiques le plus souvent muettes (-s, -t, -e, notamment) doit recourir à un calcul nécessairement coûteux. Avec l'habitude, l'exposition à des patrons syntaxiques dominants va le

conduire à traiter de manière automatisée la distribution des marques d'accord [...] (2001, p.159).

L'orthographe lexicale semble faire plutôt appel à la mémoire visuelle. Un mot présente aux yeux une graphie obtenue par convention que l'élève doit mémoriser. L'orthographe lexicale (d'usage) touche « l'ensemble des graphies imposées par des conventions linguistiques qui ne dépendent pas des règles d'accord. » (Simard, 1995, p.145). Elle contribue à la lisibilité des textes. Il s'agit de tout ce qui concerne l'image graphique ou la sonorité du mot (transcription du phonème en graphème). Nous parlons des phonogrammes (phonogramme « ou » dans le mot *hibou*) ; des morphogrammes lexicaux (le graphème « d » dans le mot *marchand*) ; des logogrammes (« cru » de *croire* et « crû » de *croître* sont différenciés par l'accent circonflexe) ; et des idéogrammes que Catach (1980) définit comme des signes extérieurs à l'alphabet, mais qui sont hautement informatifs (dans le mot *l'école*, l'apostrophe représente l'idéogramme qui évite la rencontre de deux voyelles).

La distinction de ces deux composantes de l'orthographe est importante parce qu'elle permet de préciser les points sur lesquels nous allons porter notre attention pour établir notre comparaison. En somme cette distinction prendra tout son sens à travers le traitement ou codification des erreurs. Prenons par exemple la phrase suivante : *n'appèle pas le chat pour séparé deux oiseu qui se batent*. En nous basant sur les définitions de l'orthographe grammaticale et lexicale, nous relevons donc deux erreurs de chaque type dans cette phrase : *n'appèle* et *batent* où les conventions orthographiques des deux mots ne sont pas respectées (erreurs lexicales) ; *séparé* et *oiseu* où nous constatons le non-respect de règles qui impliquent des relations entre des mots, dans ce cas, de la présence de la préposition devant le verbe

qui appelle l'infinitif (le scripteur s'étant ici trompé de morphogramme) et l'absence du « x » muet pour marquer l'accord du nom en nombre avec le déterminant pour le mot *oiseau*.

2.4 L'apprentissage de l'orthographe lexicale

Comme nous l'avons déjà dit, l'apprentissage (ou acquisition)² de l'orthographe est un processus long. Cela demande aux apprenants diverses opérations avant d'acquérir la maîtrise de la graphie correcte d'un mot. Dans les faits, comment se déroule ce processus ? Pour Guyon (2003) :

L'apprentissage de l'orthographe étant un processus long et dynamique, l'apprenant serait contraint en permanence de complexifier ses stratégies et donc de réguler les procédures qu'il utilise. Selon ce point de vue, l'apprenant va s'approprier peu à peu la « chose enseignée » par déformation, construction, déconstruction jusqu'à l'acquisition de l'objet visé (p.78).

De façon claire, l'apprentissage de l'orthographe demande aux apprenants diverses opérations avant de bien savoir orthographier un mot. Jaffré et Fayol (1997) nous fournissent une autre piste de réponse :

Très tôt, vers trois ou quatre ans, les enfants comprennent que l'écriture se différencie du dessin et acquièrent une idée, même vague, de ce qu'est la lecture. Cette première phase présente un certain nombre de caractéristiques. Elle est suivie d'une deuxième phase où dominerait la recherche d'un appariement systématique entre sons et lettres. Enfin, une troisième phase consisterait en la prise en compte croissante des contraintes orthographiques (par exemple de la morphologie ou de la forme spécifique de certains mots, les logogrammes* ou mots écrits) (p.78).

Ce processus d'apprentissage commence à la maison où l'enfant développe une conscience de l'écrit dans ses relations avec le monde qui l'entoure (Nadeau et Fisher, 2006). Selon Traimond (2001), cet apprentissage se

² Ces deux termes sont utilisés sans distinction dans le cadre de ce mémoire.

complète à l'école où les enfants apprennent les particularités spécifiques à l'orthographe essentiellement par la mémorisation du vocabulaire, la dictée, la copie et l'autodictée. Celle-ci, « utilisée plus récemment à partir de la fin des années 1960 [...] concilie en effet, la récitation, la dictée et la copie » (Traimond, 2001, p.161). Comme nous l'avons mentionné, le français constitue un système à orthographe opaque où l'on constate de nombreuses irrégularités (*concourir*, *concurrence*, *spacieux*, *spatial*, *nommer*, *nomination*). Plus l'apprenant utilise le système orthographique du français, plus il rencontre de la difficulté à transcrire certains mots (configurations sonores). Par conséquent, sans qu'il y ait une délimitation dans la progression des élèves « l'entrée dans la phase orthographique se marque par la prise en compte en lecture comme en écriture de contraintes que les appariements phonèmes-graphèmes ne suffisent pas à résoudre. » (Jaffré et Fayol, 1997, p. 94).

Chaque apprenant, étant confronté à cette insuffisance du code, il cherche à trouver une logique pour mieux contourner les pièges du système orthographique. Regardons les facteurs qui interviennent pendant cet apprentissage dont le but premier est la maîtrise de l'orthographe du français. Lalande (1988) nous présente une théorie sur l'acquisition de l'orthographe lexicale au primaire (la clientèle visée par notre étude). Le premier point à retenir, c'est le fait que l'orthographe « ne doit plus être vue comme une simple mémorisation de formes visuelles, mais comme une construction continue, une re-construction du système orthographique » (p.13). Ce travail est possible chez l'apprenant grâce à l'aide de deux sortes de connaissances : les connaissances spécifiques (CS) et les connaissances générales (CG).

Les connaissances spécifiques se définissent comme « toute information en mémoire, reliée à un mot particulier du lexique personnel et susceptible d'intervenir dans la lecture et la production graphique de ce mot » (Ibid., p.24). « Les connaissances générales constituent l'ensemble des connaissances reliées à la construction d'une graphie conforme à la langue. » (Ibid., p.32). Pour plus de précision, nous pouvons dire qu'il s'agit :

Des connaissances sur les graphèmes comme élément de base du système orthographique [...] et de règles particulières exprimant les contraintes ou les conditions d'utilisation de ces graphèmes dans la construction de graphies conformes au système orthographique (Ibid., p.192).

Pour illustrer les connaissances spécifiques et les connaissances générales, prenons un exemple de Lalande (1988) :

Savoir que dans *maison* /e/ s'écrit *ai* repose sur la connaissance d'une correspondance phono-graphique à l'intérieur de ce mot et constitue une CS, par contre, savoir que *ai* est une des graphies possibles pouvant en français représenter le phonème /e/ constitue une CG (Ibid., p.26).

En conclusion, l'apprentissage de l'orthographe, selon Lalande (1988) « consisterait à intégrer un système de connaissances » (p.191). Une réalité qui expliquerait sans doute le fait que l'apprentissage de l'orthographe lexicale du français soit si long pour les francophones. Vu le nombre de graphèmes possibles pour un phonème (CG), la quantité de CS nécessaire à la maîtrise de l'orthographe lexicale doit être énorme.

Par conséquent, en situation de production écrite le scripteur (ou apprenti-scripteur) doit constamment faire appel à sa mémoire pour retrouver la représentation graphique du mot parce que dans le cas de tout scripteur :

La production d'une graphie repose essentiellement sur deux types d'information en mémoire : une représentation visuelle complète ou partielle du mot permettant sa reconnaissance, et une liste de

graphèmes (phonogrammes) possibles pour chacun des phonèmes de la langue (Ibid., p.36).

Une théorie qui est confirmée par des recherches en neuropsychologie cognitive comme le dit Bonin (2005) :

Il existe deux procédures pour produire des mots : une procédure de récupération en mémoire de l'orthographe des mots ou de la prononciation et une procédure de conversion d'unités infralexicales (des phonèmes en graphèmes en production orthographique et des graphèmes en phonèmes en lecture à voix haute) (p.328).

Nous comprenons donc que l'élève n'écrit pas au petit bonheur de sa fantaisie. L'apprenant qui se retrouve en situation de production écrite choisit sa graphie erronée ou correcte selon une logique dictée par les connaissances du code orthographique emmagasinées.

L'orthographe du français, par sa complexité, demande un travail constant avant de parvenir à la maîtrise de la graphie d'un mot. Des recherches en psycholinguistique démontrent que « pour la production manuscrite, divers niveaux de traitement sont nécessaires pour la production d'une trace graphique correspondant à un mot » (Bonin, 2005, p. 326). Nous venons de voir que l'apprenant en situation d'écriture (dictée ou composition personnelle) génère une image du mot à transcrire. Les erreurs orthographiques semblent également provenir de cette opération ou association « image-mot ». En effet, comme nous le dit Lalande (1988), « obtenue par construction ou par combinaison de sources diverses, l'image graphique peut s'avérer conforme à l'orthographe ou s'en éloigner considérablement » (p.193).

En conclusion, d'après la théorie de Lalande (1988), l'apprentissage de l'orthographe se réalise par :

Un traitement en profondeur de l'objet d'apprentissage, par l'établissement de liens multiples et diversifiés entre les éléments des branches sémique, phonique, graphique du réseau des connaissances orthographiques et des liens entre ceux-ci et d'autres connaissances du savoir général de l'individu (Ibid., p.194).

Nous trouvons que cette section est importante pour notre étude. Elle nous fournit des informations relatives à l'apprentissage de l'orthographe lexicale. Un apprentissage qui est effectivement long et qui se révèle plus complexe que nous l'imaginions. Ces informations nous aident à mieux comprendre les processus qui entrent dans la maîtrise de l'orthographe. Elles nous aideront également à comprendre certaines graphies erronées produites par les élèves.

2.5 Les homophones

Le français compte un nombre considérable d'homophones. Selon Mareuil (1993), « Les homophones sont des mots qui se prononcent de la même manière, mais qui ont des sens différents » (p.156). Les textes du corpus de notre étude comparative contiennent plusieurs homophones (ces/ses, on/ont, etc.) À travers cette présentation des homophones, nous devons décider comment les traiter et où les classer. Relèvent-ils de l'orthographe grammaticale ou de l'orthographe lexicale ?

Catach (1986) nous dit que les logogrammes ont pour principale fonction de donner une image visuelle spécifique à certains mots homophones. Par conséquent, elle considère les logogrammes, dans leurs fonctions de différencier des mots entre eux, comme des homophones (p.267). Tout comme Catach (1986), Gey (1987) distingue trois types d'homophone.

- Les homophones grammaticaux : ce sont pour la plupart des mots-outils comme **a/à, ou/où, ce/se, son/sont**, etc. (p.93).

- Les homophones lexicaux : ce sont des termes hétérographes comme **repaire/repère, saut/sceau/sot** (p.95). On constate aisément que tout comme les homophones dit grammaticaux, les homophones lexicaux apparaissent tels quels dans le dictionnaire. La différence majeure entre les deux catégories repose sur la catégorie des mots mis en présence et sur une tradition scolaire.
- Les homophones de discours : ce sont des séquences homophoniques comme **débrouillard/des brouillards, latitude/l'attitude, la larme/l'alarme** (p.96). C'est le rapprochement phonétique qui permet le rapprochement des expressions en lien avec le découpage en mots.

Les homophones occupent une grande place dans l'orthographe du français. Il suffit de voir les nombreuses publications sur le sujet pour s'en convaincre : Bois (1973), McNicoll et Roy (1985), Braine (1993), Farid (1996). En ce qui a trait à leur enseignement dans les classes de français, l'engouement pour ces mots jumeaux commence dans la fin des années 1980 où « les grammaires utilitaires et encore plus les cahiers d'exercices, qu'on voit se multiplier à la fin des années 1980, vont accorder une très large place aux homophones » (Nadeau et Fisher 2006, p.57). Ces auteures nous disent également que ce sont les homophones grammaticaux qui étaient mentionnés dans les programmes des années 1980. Et dans les grammaires de cette époque, la distinction entre les paires d'homophones se fait essentiellement par des techniques de remplacement (substitution). Prenons dans une grammaire destinée à l'enseignement primaire l'exemple qui d'après nous est connu de tout locuteur francophone :

- Si l'on peut remplacer le mot prononcé [a] par **avait**, c'est le verbe **avoir**, il ne faut pas d'accent.
- **Dans tous les autres cas**, on écrit **à**, avec un accent grave. C'est un mot de relation : **une préposition**. (Mareuil, 1993, p.159).

Nadeau et Fisher (2006) trouvent que « lorsque l'opération de substitution se réduit à un truc unique, elle enferme l'élève dans l'application d'une procédure qui repose sur la mémoire sans compréhension réelle de la notion en jeu » (p.61). De plus, à l'aide de plusieurs exemples, elles nous montrent que l'efficacité des règles, qui, au départ étaient supposées aider l'élève à maîtriser l'orthographe, est loin d'être assurée. McNicoll et Roy (1985) nous disent que « la plupart des distinctions orthographiques relatives aux homophones sont redondantes au contexte ; elles surchargent la mémoire et ralentissent l'apprentissage : on pourrait tous s'en passer » (p.20).

En ce qui a trait aux homophones lexicaux, Gey (1987) nous dit que

Il est très difficile d'avoir une démarche pertinente pour enseigner ces homophones ou pour éliminer les erreurs dont ils font l'objet. En effet, en voulant éviter que des élèves ne confondent deux termes, on ne sait jamais dans quelle mesure on n'introduit pas artificiellement ce risque (p.95).

Le problème rencontré dans l'enseignement des homophones se situe donc dans le type de règles de substitution mentionné dans le paragraphe précédent et le mélange des catégories de mot dans le rapprochement des paires d'homophone. Prenons la paire homophonique *on/ont*. Pris isolément, les mots ont un sens différent : le « on » est un pronom indéfini et le « ont » est le verbe avoir à la troisième personne du pluriel. Par conséquent, la maîtrise de la règle d'accord du verbe avec le groupe nominal sujet peut aider l'élève à bien orthographier le mot, sans qu'il ait à rappeler en mémoire la technique de substitution. Nous retrouvons le même cas dans les homophones lexicaux. Dans la paire homophonique *voix/voie*, les deux mots n'ont que la prononciation en commun (ce qui est en fait la définition des homophones). Et comme le dit Gey (1987) n'est-ce pas le fait de rapprocher

ces mots qui nuisent à l'apprentissage orthographique de l'élève ? Nadeau et Fisher (2006) abondent dans le même sens :

Une des conséquences les plus déplorables de l'enseignement des homophones est que les élèves ont tendance à y voir une nouvelle classe de mots : il y aurait les noms, les adjectifs, etc., et les homophones. Devant *a* ou *sont*, ils ne reconnaissent plus un verbe, mais un homophone (p.59).

À travers ces exemples sur la différence de sens entre les paires d'homophone nous pouvons remarquer qu'en général la distinction entre homophones grammaticaux et homophones lexicaux se situe sur le plan de l'appartenance des paires d'homophones à une catégorie de mot. Dans le cas des homophones grammaticaux nous constatons que les mots appartiennent de façon régulière à des catégories différentes (déterminant vs verbe, préposition vs verbe). Prenons les exemples de notre corpus pour illustrer ce fait : *on/ont*, *a/à*, *son/sont*, *ses/s'est*, etc. Il est clair que la résolution des problèmes posés par ces paires d'homophones implique, pour l'élève, l'application de connaissances grammaticales. En effet, la tradition scolaire dans le traitement de ces paires d'homophones passe par des manipulations impliquant les catégories grammaticales des mots. Dans le cas des homophones lexicaux, nous rencontrons des mots qui appartiennent à des catégories identiques ou à des catégories différentes. Prenons encore les exemples de notre corpus pour illustrer ce fait : *coup/cou*, *pleine/plaine*, etc. Nous avons des noms qui sont rapprochés par leur prononciation (*coup/cou*) et un nom qui est rapproché d'un adjectif par la prononciation (*plaine/pleine*). Dans ce contexte, nous ne retrouvons pas les mêmes types de manipulation utilisée dans le cas des homophones grammaticaux parce que la réussite orthographique de ces mots semble reposer davantage sur la connaissance des formes graphiques en lien avec chaque sens. Ainsi, il est facile de voir le poids de la tradition scolaire qui s'occupe des homophones

quand les deux formes sont fréquentes (c'est le cas des « mots-outils »). Il est bien plus évident de composer des exercices pour *a/à* au lieu de *plaine/pleine*.

Ainsi, les homophones lexicaux et les homophones de discours relèvent du domaine de l'orthographe lexicale et rentreront dans la comparaison des erreurs lexicales entre les élèves francophones et hispanophones. Les homophones grammaticaux relèvent davantage de la prise en compte du contexte syntaxique et de la connaissance des catégories grammaticales. Il est donc justifié de les associer au domaine de l'orthographe grammaticale. Par cette classification, nous suivons, dans le cadre de cette étude, la tradition scolaire.

2.6 Les échelles de fréquence

Dans le cadre de ce mémoire, nous nous intéressons également aux échelles de fréquence parce que selon Brissaud et Bessonat (2001, p.32) « c'est un outil qui permet d'élaborer de manière intelligente des tests d'orthographe ». De plus, Bonin (2005, p. 329) nous dit que « les mots fréquents sont en général appris tôt dans l'existence et ceux plus rares appris plus tardivement ». Dans une étude sur la perception de l'orthographe chez de futurs enseignants, Noël (1987) arrive à la conclusion que « la méconnaissance [...] de l'orthographe de mots rarement employés s'avère la principale cause d'erreurs orthographiques » (p.24). Gey (1987) nous présente les résultats de ses exercices pédagogiques auprès de 476 élèves d'un collège parisien. Dans la section *Résultats des tests*, nous retrouvons dans un de ses tableaux 65 graphies pour le mot « excursion » (p.134). Pour vérifier l'affirmation de Bonin et comprendre ce phénomène, nous avons cherché l'échelle de fréquence du mot dans Préfontaine (1995). Dans le corpus dépouillé par Préfontaine pour établir son échelle de fréquence, le mot

« excursion » obtient, en français écrit, la fréquence d'utilisation « 8 » pour un indice de difficulté de 27 (l'indice de difficulté étant standardisé de 1 à 50, 50 étant le plus difficile). Le corpus de Préfontaine (1995) compte 4461 mots différents qui ont été lemmatisés. Certains mots (en caractère gras dans le livre) ont une fréquence d'apparition supérieure à 100 : exemple le verbe *dormir* dont la fréq. est 111. Donc nous prenons comme hypothèse que ce mot est considéré comme étant fréquent dans le discours. Par conséquent, il est clair pour nous qu'une étude sur la compétence orthographique doit considérer la fréquence des mots dans le discours parce qu'il semblerait que ce facteur de fréquence a une incidence sur la capacité des apprenants à maîtriser l'orthographe lexicale.

Il existe diverses échelles de fréquence en français : Dubois-Buyse (1989), Baudot (1992), Fortier (1993), Préfontaine (1995). En espagnol, nous n'avons pu consulter que celle de Juilland (1964) Les études présentées dans ces livres portent sur le même sujet (établir la liste des mots fréquents à l'écrit pour aider à l'enseignement de l'orthographe en classant les mots par niveau scolaire et par degré de difficulté). Cependant, parmi cette liste d'échelles francophones, il nous faudra faire un choix en fonction de la pertinence pour notre étude comparative. Présentons les avantages et les limites de ces échelles.

2.6.1 L'échelle Dubois-Buyse (1989)

L'échelle Dubois-Buyse est un matériel qui « donne le moyen de se constituer des tests à l'aide desquels, en un quart d'heure on peut estimer le niveau d'un élève, d'une classe, au point de vue de l'orthographe d'usage. » (Dubois-Buyse, 1989, p.82) Le corpus est construit avant tout à partir de (Ibid., p.10) « 4100 rédactions spontanées d'élèves des deux sexes, elle a abouti à établir la liste des 4239 mots les plus employés dans les écrits

d'usage courant. » Indéniablement, ce procédé a l'avantage de répertorier le vocabulaire des jeunes tout en vérifiant leurs habiletés à résoudre des problèmes en cours de rédaction. Cependant, à cause de son ancienneté (1940 étant la première date de publication de l'échelle Dubois-Buyse) elle ne touche pas le vocabulaire des jeunes québécois actuels. Nous mentionnons les jeunes d'ici, alors que Catach (1986, p.58) fait la même remarque pour la France : « le choix demanderait à être revu, rectifié et complété constamment, pour tenir compte de l'évolution rapide intervenue entre-temps dans le vocabulaire utilisé en France, pour les enfants d'aujourd'hui [...] » L'exemple le plus simple est le mot « ordinateur » qui n'apparaît pas dans la liste. Nous choisissons ce mot comme indice de renouveau à cause de la grande place qu'occupe cet outil dans la vie quotidienne des gens de cette génération depuis bientôt, une vingtaine d'année.

2.6.2 L'échelle Baudot (1992)

Baudot (1992), pour établir son échelle de fréquence, a utilisé un corpus de 803 échantillons de textes d'environ 1000 à 1500 mots. Les textes sont répartis en 15 genres littéraires et proviennent de la France (62 %) et du Canada (37 %). Il est clair que les textes ne sont pas récents. L'auteur nous mentionne l'année de publication de son corpus : « la majorité d'entre eux (726 sur 803) se situant entre 1960-1967 » (p.14) ; et qu'ils se situent hors du monde des enfants (notre clientèle étant des jeunes de 12 ans). Il ne pourrait en être autrement, car ce sont des textes d'adultes écrits pour des adultes. Mais l'intérêt premier de cette échelle réside dans le fait que ce sont des mots un peu plus contemporains que ceux de Dubois-Buyse, et que le chercheur présente une échelle qui tient compte de l'emploi du mot selon ses différentes entrées dans le dictionnaire (Exemple : *vaincu* = 23, 20 *adj* + 3 *n*). Mentionnons que le mot ordinateur n'apparaît toujours pas dans cette liste de mots.

2.6.3 L'échelle Préfontaine (1995)

Préfontaine (1995), dans un premier temps, a vérifié la validité de l'échelle Dubois-Buyse. Il a constaté que malgré son utilité, le vocabulaire restait éloigné de la clientèle québécoise. Par la suite, il a questionné la pertinence des mots à l'aide de l'échelle de Baudot (1992). Ce travail lui a permis d'épurer l'échelle Dubois-Buyse tout en la complétant à l'aide de Baudot (1992) en fonction de la fréquence d'apparition des mots (20 et plus). Elle comporte donc plus de mots récents (le mot ordinateur est présent dans la liste).

En somme, cette échelle de fréquence des mots reprend le modèle de Dubois-Buyse pour donner un outil plus récent aux enseignants qui évaluent la compétence de l'orthographe d'usage chez les élèves. Le mot « ordinateur » apparaît dans la liste des mots. Cependant, comme tous les ouvrages cités ici, l'auteur ajoute une liste complémentaire des mots variables selon leur genre, mais les mots listés avec leur fréquence sont uniquement au masculin (sans indiquer si la forme masculine englobe la forme féminine) ; et les verbes sont lemmatisés au mode infinitif. Ainsi, il est facile de comprendre que ce procédé (ramener tous les verbes à l'infinitif) représente un inconvénient quand nous nous penchons sur l'orthographe parce qu'à maints égards le « savoir orthographier » des verbes conjugués fait partie de la maîtrise de l'orthographe d'usage. Prenons l'exemple suivant : le verbe aller, contrairement aux autres verbes en « er » présente plusieurs radicaux dans sa conjugaison : *je vais, j'irai, j'allais, etc.* aussi vrai que la finale du verbe est du ressort de l'orthographe grammaticale, il est, pour nous, évident que toute modification à la graphie du radical relève de l'orthographe lexicale.

2.6.4 L'échelle Fortier (1993)

Fortier (1993) s'est intéressé au vocabulaire des adolescents et des adolescentes du Québec. Les sujets proviennent des écoles de la commission scolaire de Montréal. Un total de 1271 élèves fréquentant le premier secondaire et le cinquième secondaire ont participé au test. Des participants qui ont fourni un corpus de 259.951 mots au chercheur. En plus d'établir la fréquence des mots, son étude présente également, en pourcentage, un indice du vocabulaire disponible dans une population donnée. Cet indice, il le décrit comme étant « l'ensemble du vocabulaire en mémoire chez tous les sujets de cette population avant toute production de discours » (Fortier, 1993, p.46). Le vocabulaire en mémoire correspond aux connaissances que l'utilisateur a du système linguistique : vocabulaire, grammaire, syntaxe, morphologie, etc. Il peut donc les activer pour les utiliser dans ses activités linguistiques : parler, écrire. Par conséquent, cette étude cadre bien avec notre clientèle puisque l'ensemble des mots provient d'une association libre « continue » : « le test propose [...] au sujet une série de mots stimuli, [...] le sujet peut mentionner plusieurs réponses pour le même mot stimulus » (ibid., p.51). Ainsi, pendant notre étude, il faudra comparer la réussite des mots par rapport à leur fréquence dans le vocabulaire des francophones.

L'échelle de Fortier (1993) est retenue parce qu'elle est plus adaptée au contexte québécois. Elle permettra une analyse en profondeur des erreurs à travers la fréquence des mots. De plus, l'indice de disponibilité apporte une information supplémentaire quant à l'accessibilité des mots dans le vocabulaire des francophones d'ici.

2.6.5 L'échelle Juilland (1964)

L'échelle de fréquence consultée pour les dictées en langue espagnole est celle d'Alphonse Juilland. Ce dictionnaire de fréquence date de 1964. Le résultat final découle d'un corpus de 25.000 phrases pour un total de 500 000 mots. Les textes retenus pour cette étude datent de 1918-1934. « To insure temporal or chronological homogeneity, only works written between the two wars were admitted (...) » (Juilland, 1964).

Les textes proviennent de cinq genres d'écrits : « dramatic literature, fictional literature, essayistic literature, technical literature, journalistic literature ». Après des éliminations successives, seulement 5.024 mots sont retenus pour le dictionnaire de fréquence. La principale caractéristique qui rend cet ouvrage intéressant, c'est le fait que les mots apparaissent sous leurs formes diverses. Par exemple, un nom ou un adjectif est lemmatisé en genre et en nombre : « ciega, ciegas, ciego, ciegos ». Les verbes sont à l'infinitif et sous leurs formes conjuguées : « *alimentar, alimenta, alimentaba, alimentan, alimentando, etc.* »³. On peut ainsi savoir, par exemple, si la forme du participe passé est plus fréquente que le verbe conjugué au présent.

Les limites de ce dictionnaire de fréquence se situent d'abord dans l'ancienneté de l'ouvrage. Les réalités ayant changé, rien ne garantit que l'auteur aurait fait les mêmes choix de textes et d'auteurs. Tout comme l'a fait Préfontaine (1995), il aurait sans doute épuré ou élargi le corpus pour rendre compte des nouvelles réalités d'une époque plus récente. De plus, il est clair que le vocabulaire provient d'une littérature spécialisée. À l'image de Baudot (1992) ce sont des textes d'adultes écrits pour des adultes. Pour finir, mentionnons que l'ouvrage est sélectif et n'englobe pas tous les peuples

³ alimentar (nourrir) ; alimente (il) ; alimentait (il) ; alimenté.

hispanophones, car « to insure spatial homogeneity, only spanish peninsular sources were admitted [...] » Cependant, nous retenons ce dictionnaire dans le cadre de notre étude pour deux raisons. Premièrement, il s'agissait de la seule échelle de fréquence pour l'espagnol que nous avons pu consultée. Deuxièmement après une comparaison des mots correspondants dans les deux langues, nous constatons certaines similitudes entre cette échelle et celle de Fortier (1993). Par exemple, nous trouvons 69 mots qui dans les deux langues ont une fréquence inférieure à 100, et 32 mots qui ont une fréquence supérieure à 100. Il est clair qu'il ne peut y avoir correspondance parfaite entre les mots qui servent à notre étude. Cela s'explique par le fait que dans le cas de Juilland (1964) le vocabulaire provient d'adultes experts alors que dans le corpus de Fortier (1993) les mots proviennent des élèves du secondaire. Cependant, cette échelle de l'espagnol demeure pertinente pour notre étude par rapport à l'éventail de mots qu'elle présente.

2.7 Les Projets de réforme de l'orthographe

Le deuxième objectif de notre recherche est de voir les impacts qu'aurait un projet de réforme sur la compétence des élèves dans l'éventualité de son application. Nous pensons donc qu'il est important de brosser un portrait des différents projets publiés dans la deuxième moitié du 20^e siècle. Nous retenons ces projets parce qu'ils sont plus récents par rapport à notre époque. De plus, ils semblent avoir servi de base au travail du CSLF.

En 1990, le CSLF a fait des propositions qui ont été acceptées majoritairement (Arrivé, 1993, p.121) « 23 voix contre 6 » par les membres de l'Académie française. Mentionnons que les recommandations ne touchent que 2430 mots dont 56 sont présentés comme étant de haute fréquence. Voici quelques exemples de mots de haute fréquence tirés dans le *Vadémécum de l'orthographe recommandée*, 2004 : *aout*, *boîte*, *diner*,

révoluer, voute. Donc, pour Brissaud, (2001, p.44) « notre orthographe ne s'en trouve pas bouleversée. 20 % des graphies recommandées se trouvaient déjà dans les dictionnaires en 1990 ». Une fois encore, les rectifications, qui ne sont que des recommandations à cause de la coexistence permise des deux graphies, sont longtemps restées lettre morte. Dans les écoles, les nouvelles graphies ne sont pas enseignées, et les dictionnaires « suivent timidement les recommandations du Conseil Supérieur de la Langue Française » (Ibid., p.36).

Les données ont cependant changé avec le temps. « En 2003, les logiciels de correction *Antidote Prisme* (produit au Québec par Druide informatique inc.) et *ProLexis* (Europe) ont intégré toutes les nouvelles graphies dans leurs dictionnaires [...] Le *Robert* intègre maintenant environ 60 p. 100 de ces nouvelles graphies, le *Dictionnaire Hachette* les contient à 100 p. 100 depuis 2002, et le *Dictionnaire du français usuel* (de Boeck) est écrit en orthographe rectifiée. » (Contant, 2004). Nous constatons que les rectifications font graduellement leur chemin à travers plusieurs publications dont le *Bescherelle, l'art de conjuguer* (2006) selon le Groupe québécois pour la modernisation de la norme du français (GQMNF). Notons également que Microsoft est en train d'intégrer les rectifications orthographiques à ses outils. Essayons maintenant d'analyser les perspectives de quelques-uns des projets de réforme (les plus récents du siècle dernier).

René Thimonnier (1967), à travers une présentation détaillée du système graphique du français, montre que les véritables anomalies sont en très petit nombre. Il présente donc un « projet-limite ». Nous parlons de « projet-limite » parce qu'il s'agit selon cet auteur d'un « émondage rationnel de l'écriture. » Le but premier de l'auteur n'est pas de réformer l'orthographe du français, « mais de la protéger de réformes possibles. » Pour lui, « normaliser

l'écriture, c'est essentiellement rectifier tout ce qui s'écarte de la norme. Mais si légitime qu'elle paraisse, l'absorption de l'anormal par le normal n'est pas toujours souhaitable » (p.263). Nous trouvons dans l'argumentation de Thimonnier le souci de conserver les habitudes. Mais il propose quand même d'imposer dans les familles de mots, « autant qu'il se peut, la forme la plus générale » Un objectif à atteindre avec un grand total de 228 mots rectifiés. Par exemple, (p.370) il nous dit pour le mot *désiller* « conformément à l'étymon *cil*, *cilium*, tous les mots de la famille s'écrivent avec un *c* : *ciliaire*, *cillement*, *ciller* ; c'est donc la graphie « déciller » qui s'impose » À l'opposé, il nous dit que le mot *honneur* ne peut être simplifié en *honeur* parce que la première graphie est « incrustée dans des millions de mémoires individuelles et concrétisée dans des millions d'ouvrages » (p.369). Thimonnier explique sa position en précisant qu'une telle réforme était possible « il y a 150 ans », mais aujourd'hui « C'est assez dire, qu'il est trop tard pour opérer une simplification générale et cohérente du système » (p.370). Il convient de retenir que Thimonnier considère l'orthographe du français comme assez cohérente et que les sections à réformer sont minimales.

Aristide Beslais a présenté deux propositions de réforme orthographique, l'une en 1952, l'autre en 1965. Dans les deux cas, aucune des recommandations n'a été retenue. Pendant la préparation du second projet, la commission ministérielle d'études orthographiques, présidée par Beslais, analysa le système orthographique et les copies des élèves à différents niveaux pour identifier les aspects de l'orthographe lexicale qui occasionnent le plus d'erreur (Beslais, 1965, pp 123-125). Quels ont été les points retenus par la Commission Beslais pour formuler son projet de réforme ? Nous présentons ici les propositions principales (Ibid., p.103-106) : la réforme des mots composés, soudure pour les mots à premier élément verbal et les mots à premier élément invariable (porteplume, antédiluvien) pour les autres, la

commission propose le statu quo ; la simplification des lettres grecques (*th, rh, ch, ph, y*), la commission demande de généraliser la francisation des mots contenant ces diagrammes grecs (*téâtre, pharmacie, système*), mais de maintenir les *h* en début de mot et le *ch* dans quelques mots à consonnance religieuse et les mots où le digramme est chuintant (*Christ, archevêque*) ; le remplacement du « **x** » (muet final) par « **s** » dans les mots où il n'est pas prononcé (soit au singulier, soit au pluriel) (des *bijoux, un ou des jaloux, pris*) ; la simplification des consonnes doubles (*honeur, courier*) sauf dans la conjugaison de certains verbes irréguliers (*courrai, pourrais*), les mots où cette réduction à l'unité changerait la phonologie du mot (*moisson, corbeille, occident*), et les cas où la consonne double permet de différencier des homonymes (*arête, arrête*) ; la normalisation de l'accentuation (*évènement*), la suppression du « **e** » muet interne avant et après voyelle dans les verbes et les noms (*je sursoirai* comme *j'assoirai*), harmonisation des noms et des adjectifs en « **ant** » et « **ent** » (*exigent, exigence*). Prévoyant la réaction du public, la commission (Beslais, 1965, p.11) propose d'étaler la réforme sur une période de seize années à l'exemple des réformes réussies dans les autres pays. Ce délai tient également compte du temps qu'il faudra pour la réédition des livres scolaires. Un projet qui se montre plus audacieux que celui de Thimonnier et qui montre que l'orthographe du français peut se débarrasser de certaines contraintes héritées du passé.

En 1940, Albert Dauzat propose d'unifier les familles de mots : verbes en *eler, eter, sur modeler et acheter, verbes en otter sur le modèle de radoter, trembloter* ; les suffixes *ot, otte* ; les mots en *onner, onnance, onneux, etc.*, Selon Dauzat, ces familles de mots offrent des discordances fâcheuses, et devraient prendre, selon l'étymologie, la consonne simple. Les pluriels en « *eux* » et « *aux* » devraient disparaître au même titre que les sept pluriels en « *oux* » et quelques autres « *x* » muets (Catach, 1978, p.82). Ce projet tient

en deux pages et porte également sur les points suivants : les lettres grecques (suppression du *h* de *rh* et des *ch* prononcés /k/) ; les erreurs étymologiques dans quelques mots (*forsené*, *herce*, etc.) ; les lettres parasites, dont la présence ne repose sur aucune raison particulière dans l'orthographe du mot (*sie pour scie*, *sier pour scier*, *domter pour dompter*). L'auteur du projet ne mentionne pas cependant le nombre de mots à modifier selon le critère étymologique. Un projet de réforme qui, sans être exhaustif, touche à des points importants. Il s'agit de simplifier un grand nombre de mots. Nous citons en exemple cette série : *donner*, *ordonner*, *ordonnance*, *donneur*, etc. qui s'aligneraient sur les mots de la même famille comme *don*, *donation*, *donateur*. Nous savons à quel point les consonnes doubles causent des problèmes dans l'acquisition de l'orthographe, puisque leur disparition aurait éliminé 11,2% de fautes dans les copies des élèves du certificat d'études primaire (C.E.P) retenue pour analyse par la commission Beslais (1965, p.123). L'autre avantage d'un tel projet réside dans l'uniformisation du pluriel : disparition des pluriels en « **eux** » et « **aux** » et les sept pluriels en « **oux** » pour respecter la règle de l'ajout du morphogramme « **s** » aux noms pluriels (*des cheveux*, *des chevaux*, *des bijoux*, etc.). Un projet qui présente des qualités reconnues par Thimonnier (1967, p.361) qui affirme que « Dauzat a su nettement dégager les principes d'une réforme rationnelle de l'écriture » et par Burney (1955, p.121) qui l'associe aux propositions de Ch. Beaulieux et l'arrêté de 1901 pour nous dire que « ces textes sont susceptibles d'obtenir l'approbation d'une imposante majorité. La plupart des éducateurs en seraient partisans, dans les trois ordres d'enseignement. » Mais nous retrouvons des spécialistes de la langue qui critiquent le projet Dauzat pour son manque de rigueur. Thimonnier (1967, p.357) affirme qu'en « dépit de ses qualités, il risque fort d'être inefficace. Non seulement, il met l'accent sur des difficultés toutes secondaires, mais il néglige les principales, ou ne leur donne que des solutions contestables. »

Selon Catach (1978, p.83) : « on pourrait reprocher au projet Dauzat non seulement des points discutables (comme l'ajout de l'encombrant trait d'union à tous les composés) mais trop de dispersion. »

Pour terminer, il reste à ajouter que, hormis le projet d'émondage de Thimonnier, nous constatons que certains points reviennent constamment dans l'ensemble des projets. Il s'agit des consonnes doubles, des lettres grecques, des mots composés, des pluriels en « x » des noms et adjectifs, de l'utilisation des accents. Le rapport du CSLF contient bien le trait d'union, les accents (tréma, accent grave, accent circonflexe), les consonnes doubles, mais pas dans les mêmes proportions que les autres projets présentés ci-dessus. Nous ne retrouvons pas les autres points (consonne « x », lettres grecques, les noms en « ant » et « ent ») relevés par Beslais lors de la compilation des erreurs dans les copies des élèves à différents niveaux, et qui ont servi de base à la présentation de son projet de simplification orthographique.

Nous retenons de tous ces projets que l'orthographe du français est tout autant difficile à maîtriser qu'à réformer. Ceux qui ont proposé des projets de réforme s'accordent sur plusieurs points communs sans leur donner la même importance. Prenons l'exemple des lettres grecques, Dauzat considère uniquement les « rh et ch » alors qu'en français il existe plusieurs mots contenant d'autres diagrammes grecs (ph, th), héritages d'une époque lointaine, que d'autres pays comme l'Italie, l'Espagne ont fait disparaître de leur système orthographique. Thimonnier nous dit qu'il est trop tard pour procéder à une simplification générale du système. Une telle déclaration apparaît inconcevable devant les nombreuses incohérences que renferme l'orthographe du français. Citons à titre d'exemple le pluriel des noms en « ou ». Pourquoi doit-on continuer à encourager les élèves à mémoriser les

sept noms qui font exception à la règle de base du pluriel qui causent tant de problème à cause du caractère silencieux de la marque du pluriel en français ? Un autre exemple, c'est la différence entre la finale « ence » et « ense ». En consultant le *Code de l'orthographe française* de Maurice Grevisse (1948), nous trouvons seulement 9 mots qui prennent la finale « ense » : *défense, dense, dépense, dispense, immense, intense, offense, récompense, suspense*. Devant « e », le graphème « c » a la valeur du phonème /s/. Ainsi, nous doutons que l'unification de la série sur la finale « ence » puisse nuire à l'image de la langue française. À ce propos, le CSLF a proposé des rectifications qui passent graduellement à l'usage malgré les cris de révolte lancés dans les journaux par les conservateurs lors de la publication de ces rectifications. La preuve que « rien ne s'oppose, donc, théoriquement, à une révision prudente et modérée de notre orthographe » (Burney, 1967, p.126).

En conclusion, ces divers projets de réforme nous renseignent sur l'urgence de donner un grand coup de balai pour débarrasser l'orthographe de certaines singularités qui l'encombrent. Ils sont utiles dans le cadre de cette recherche pour voir l'impact que l'un ou l'autre aurait eu sur les textes de notre corpus. En d'autres mots, il s'agit de voir dans quelle mesure l'application des rectifications proposées aurait permis aux élèves francophones d'être plus compétents dans l'orthographe de leur langue. Dans le chapitre des résultats, nous choisirons le projet qui nous paraît le plus intéressant pour notre étude.

2.8 La typologie des erreurs lexicales

Selon Jaffré et Fayol (1997, p.102), il existe dans un système linguistique « des zones de fragilité, définies comme particulièrement propices à l'occurrence de variations non conformes aux conventions. » L'opacité du

système orthographique du français, reconnue par Catach (1986), Arrivé (1993), Jaffré et Fayol (1997), place les apprenants en présence de plusieurs de ces zones de fragilité. À titre d'exemple, nous pouvons prendre les consonnes doubles en français (Jaffré et Fayol, 1997), et la confusion *b/v* en espagnol (Pombo, 1999). Pour mieux les cerner et les codifier, il importe de recourir à une typologie des erreurs.

Beslais (1965) présente à la fin de son projet la typologie utilisée pour corriger les copies des élèves. Le tableau où il compile le pourcentage d'erreur pour chaque niveau scolaire (CEP, Baccalauréat, Propédeutique) (p.123) compte dix items : mots composés soudés ; accents circonflexes, lettres grecques omises ou simplifiées, la consonne *x*, consonnes étymologiques, consonnes doubles simplifiées, mot en *ent-ant-ance*, l'*e* parasite, homonymes, cas particulier *c = qu*. En dépit de sa simplicité d'application, cette typologie ne semble pas convenir à notre recherche. Il est vrai que tout comme notre étude, elle demeure centrée sur l'orthographe lexicale. Nous ne retrouvons donc dans son tableau aucune section qui tient compte des différents accords qu'un scripteur est appelé à faire en situation d'écriture. Il est vrai également qu'une telle typologie est facile d'accès pour les professeurs non-spécialistes en français. En effet, cette typologie ne mentionne pas les notions de graphèmes et de phonèmes. Ce côté non linguistique des corrections ne correspond pas tout à fait à notre approche et surtout, elle ne permet pas de classer toutes les erreurs possibles. Nous le constatons facilement dans le cas particulier *c = qu* qu'il présente dans son étude, alors que le phonème /k/ peut être représenté par plusieurs graphèmes : *c, cc, ch, k, qu, etc.* De plus, nous nous questionnons sur les choix faits par Beslais (1965) dans ses catégories d'erreurs. Pourquoi considérer comme catégorie d'erreurs les accents circonflexes et pas les autres ? Pourquoi les mots en *ent-ant-ance* et non toutes les confusions

an/en ? Nous nous demandons si ces erreurs sont effectivement les plus courantes dans les copies analysées ou si cette typologie n'a pas été construite en fonction des éléments qu'il cherchait à réformer. En effet, le tableau résumé des propositions présente plusieurs de ces points qui ont toujours fait l'objet des projets de réforme : mots composés, consonnes doublées, consonne « x », lettres grecques, etc.

Catach (1980) dans son ouvrage, *L'enseignement de l'orthographe*, propose une typologie qui permet de donner un caractère plus rigoureux à la compilation des erreurs de l'élève, fondé sur l'étude linguistique du système orthographique. Cette typologie contient six catégories que nous présentons sommairement ici :

1. Erreurs phonétiques : il s'agit des erreurs liées à la prononciation ou à la discrimination auditive des sons (exemple : *chauffache*). Par exemple, l'absence d'un accent en français (malgré que le mot conserve son identité pour le locuteur francophone) devient une erreur phonétique (ex : *ecole*).
2. Erreurs phonogrammiques : il s'agit de la méconnaissance de la transcription sonore entendue. Le son est bien entendu et discriminé, mais l'élève ne sait pas comment le transcrire (exemple : *leusémie*) dans cette section, il s'agit de toute transcription qui n'altère pas la valeur phonique du mot. Dans le modèle de Lalande (1988), ces erreurs sont dues à un manque de connaissances spécifiques (CS) du mot et non à un manque de connaissances générales (CG)
3. Erreurs morphogrammiques : il faut distinguer les morphogrammes grammaticaux qui touchent aux accords, à la finale des verbes, etc. (exemple : *genous*, *applaudies*), et les morphogrammes lexicaux qui touchent à la graphie consacrée du mot. Catach (1980) regroupe dans cette section l'ignorance de la famille lexicale (*inabité/inhabité*),

l'ignorance des préfixes et des suffixes (*anterremant/enterrement*) et l'ignorance des lettres finales justifiables (*marchant/marchand*). Il convient de préciser que dans la catégorie des morphogrammes lexicaux, nous regroupons les erreurs de préfixes ou de suffixes servant à la famille du mot (exemple : le « d » de marchand) et les erreurs de lettres muettes dans le mot (exemple : le « t » de petit qui ne s'entend qu'au féminin).

4. Erreurs homophoniques : il s'agit des homophones de discours (d'eux/deux), des homophones lexicaux (*pleine/plaine*) et des homophones grammaticaux selon la tradition scolaire (*sont/son*).
5. Erreurs idéogrammiques : il s'agit de tout ce qui touche à l'aspect présentatif du mot ou de la syntaxe : virgule, apostrophe, trait d'union, majuscule, etc. (*lui même ; lécole*)
6. Erreurs des lettres non fonctionnelles : il s'agit des consonnes latines (*sculter/ sculpter*), des consonnes grecques (*téâtre/théâtre*), etc.

En parcourant les tableaux de Catach (1980) nous remarquons que certains exemples peuvent se retrouver dans plus d'une catégorie. D'ailleurs, comme le dit Catach (1980), « cette typologie est perfectible ». Prenons quelques exemples pour illustrer nos propos. La graphie erronée « *combatif* » apparaît dans la catégorie 6 (lettres non fonctionnelles), mais nous trouvons que cette erreur aurait bien pu se retrouver dans la catégorie 2 (phonogramme) parce que la sonorité du mot est respectée. Le deuxième exemple est la graphie erronée « *inabité* ». Cette graphie est placée dans la catégorie des morphogrammes lexicaux (ignorance de la famille lexicale). Nous sommes d'avis que cette graphie erronée pourrait aussi se retrouver dans la catégorie 2 où est classé le mot « *aléger* » (erreurs phonogrammiques, *l'élève connaît les sons sans connaître leur transcription*). Par conséquent, il paraît évident

que la catégorie 2 peut facilement englober la catégorie 3 sur le plan de l'orthographe lexicale (erreurs morphogrammiques).

Cela dit, cette typologie a l'avantage de permettre une bonne compréhension des erreurs des élèves à travers la hiérarchisation objective des erreurs qu'elle propose. Chaque erreur pouvant être définie selon ses caractéristiques propres, cet outil permet à l'enseignant de guider l'élève dans son apprentissage du code orthographique. Il peut donc lui fournir les explications nécessaires avec des solutions pratiques.

Nous comprenons l'importance accordée au travail de Catach par ceux qui étudient les erreurs des élèves selon une approche linguistique. Citons à titre d'exemple Gey (1987) qui reprend textuellement Catach (1980) alors que Racine et al. (1986) adoptent une typologie où les erreurs ne sont pas détaillées. Dans sa catégorie orthographe, nous retrouvons les erreurs reliées aux fautes d'orthographe, aux accents, à la confusion homophonique, aux majuscules et minuscules, aux coupures d'un mot à la fin d'une ligne, à l'emploi des nombres et des abréviations. Ceci se rapproche de la typologie de Beslais (1965). Dans le cadre de notre recherche qui adopte une approche linguistique, nous retenons par conséquent les six items de Catach (1980) pour analyser les erreurs lexicales que nous relèverons pendant notre analyse fine sur le corpus de textes. Il convient de rappeler que les erreurs grammaticales (morphogrammes grammaticaux et homophones grammaticaux) rentrent dans notre résultat global de recherche, mais notre intérêt principal demeure l'orthographe lexicale (erreurs lexicales et homophones lexicaux et de discours : **l'alarme/la larme**).

Ainsi, pour notre analyse, nous conservons la typologie de Catach à laquelle nous apportons les précisions suivantes. Les erreurs phonétiques

contiennent toutes les transcriptions qui ne respectent pas la prononciation correcte du mot, les erreurs phonogrammiques englobent toutes les prononciations correctes, mais mauvaise graphie. Dans le cas des morphogrammes lexicaux, nous retenons les erreurs liées à la graphie spécifique (préfixe ou suffixe) d'un mot.

2.9 Les questions de recherche

En comparant les compétences en orthographe d'usage (lexicale) d'élèves francophones et hispanophones, notre objectif est de vérifier s'il existe un écart considérable sur le plan de la maîtrise de l'orthographe lexicale chez les enfants de 12 ans dans les deux systèmes malgré le fait que le français soit considéré comme une langue dont l'orthographe est opaque (relation entre phonèmes et graphèmes très irrégulière). Donc en lien avec cet objectif, nos questions de recherche sont les suivantes :

- Dans le corpus étudié, les enfants francophones et hispanophones du même âge présentent-ils le même niveau de maîtrise sur le plan de l'orthographe en général, et en particulier, l'orthographe lexicale de leur langue première respective ?
- Dans la maîtrise de l'orthographe lexicale, est-ce que la fréquence des mots tient effectivement un rôle important ?
 - Est-ce que la fréquence des mots joue un rôle identique dans chacun des systèmes orthographiques comparés ?
- Rencontrons-nous les mêmes types d'erreur lexicale dans les textes des élèves de chaque langue et dans quelle proportion ?

Notre deuxième objectif est de questionner les rectifications de 1990 par rapport à d'autres projets de réforme : Dauzat, Albert (1939-1952) ; Commission Beslais (1960-1965) ; Thimonnier, René (1967) ; et de voir leur

impact potentiel sur la compétence orthographique des élèves. Notre question pour atteindre cet objectif est :

- Quelle réforme parmi celles de Dauzat (1939-1952) ; de la Commission Beslais (1960-1965) ; de Thimonnier (1967) aurait permis de diminuer le plus les fautes d'orthographe dans les copies des élèves francophones ? Et dans quel pourcentage ?

2.10 Les hypothèses de recherche

Dans le cadre de notre étude comparative sur les habiletés orthographiques des francophones et des hispanophones, malgré l'absence de données empiriques, les affirmations logiques de Masson (1991) et de Jaffré et Fayol (1997) sur la maîtrise plus aisée des systèmes graphiques de l'italien, de l'espagnol, etc. par rapport au français nous permettent de prédire que le taux de réussite orthographique devrait être plus élevé chez les hispanophones que chez les francophones à cause de la régularité du système orthographique de l'espagnol.

De plus, en considérant les théories de l'apprentissage de l'orthographe (Lalande 1988, Bonin, 2005), les études sur la compétence orthographique (Gey 1987, Noël 1987), et la contribution des échelles de fréquence pour permettre un meilleur apprentissage de l'orthographe, nous sommes en mesure de prédire que les mots les moins bien réussis par les élèves de chaque population devraient se retrouver dans la catégorie des mots peu fréquents (ou rares) ; et les mots qui leur causent le moins de problème, dans la catégorie des mots fréquents.

Conclusion

Dans ce chapitre, nous avons abordé les notions de graphème et de phonème pour les situer dans le contexte de notre recherche. Nous pouvons retenir de cette section le grand rôle que joue le graphème dans le système graphique du français. Notre étude s'intéresse plus particulièrement à la maîtrise de l'orthographe lexicale en français et en espagnol, il a donc fallu établir une différence nette entre orthographe grammaticale et lexicale de manière à montrer les perspectives retenues pour le traitement de nos données.

Nous avons brossé par la suite un portrait du processus de l'acquisition de l'orthographe. Un processus qui, selon Lalande (1988), se réalise en intégrant un système de connaissances. À travers cette présentation, nous avons introduit le fait que les mots fréquents sont maîtrisés plus vite par les élèves. Ce qui implique que dans notre recherche, nous devons considérer les échelles de fréquence. Nous les avons donc abordées brièvement.

Le savoir orthographier implique une transcription phonème-graphème sans présence d'erreurs (le cas idéal). Les élèves produisent des graphies erronées et nous devons les analyser. Nous avons présenté la typologie qui nous permettra de considérer les graphies erronées selon une approche linguistique. Notre deuxième objectif de recherche étant les projets de réforme et des impacts qu'ils pourraient avoir sur la maîtrise orthographique des élèves, nous en avons présenté quelques-uns.

Dans ce chapitre, nous avons présenté les différentes questions de recherche en lien avec les objectifs de cette étude. Nous avons également présenté les hypothèses de recherche auxquelles nous devons répondre à travers cette étude comparative. Dans le chapitre suivant qui traite de la

méthodologie nous allons montrer les méthodes d'analyse qui nous permettront d'apporter des réponses à nos questions de recherche.

Chapitre III

La méthodologie

3.1 Introduction

Ce chapitre présente la méthodologie qui nous permettra de répondre aux questions de recherche énoncées précédemment dans le cadre théorique.

Dans un premier temps, nous présentons les caractéristiques des élèves qui ont fourni le corpus sur lequel nous avons travaillé. Nous présentons ensuite le corpus de dictées. Enfin, nous abordons/traitons de manière détaillée la méthode d'analyse privilégiée pour analyser le corpus, les calculs faits pour répondre à chacune de nos questions de recherche. Nous terminons ce chapitre en exposant les limites de notre recherche.

3.2 Les élèves

Le matériel de base sur lequel repose cette étude comparative est le corpus Nadeau, Matute et Gonzales-Reyes (2006). Il s'agit de 4 dictées en français et en espagnol (Voir Appendice A). Nous avons, pour l'ensemble des quatre dictées, un total de 334 mots dictés aux francophones et 333 mots aux hispanophones, nous y reviendrons. Les dictées ont été administrées en 2004 à un total de 84 élèves de niveau sixième année. De ce nombre, nous avons 42 élèves francophones (provenant du Québec) et 42 élèves hispanophones (provenant du Mexique). Bien entendu, chaque groupe a subi l'épreuve dans sa langue maternelle.

Dans les deux pays, les élèves ont le même âge, 12 ans, et il y a un même nombre de garçons et de filles dans les deux corpus (23 garçons pour 19 filles). Toutefois, dans notre analyse cette proportion change légèrement

dans le corpus français (23 garçons pour 17 filles) parce que deux filles de ce groupe qui présentaient des problèmes de surdité ont été écartées de la recherche, vu la spécificité de leurs erreurs lexicales, attribuables à leur état. Donc, nous comparons 40 élèves francophones à 42 élèves hispanophones. Les sujets (francophones et hispanophones) proviennent du même milieu socio-économique de moyen à élevé où les parents sont scolarisés (au moins un des parents ayant fait des études post-secondaires). Ce choix repose sur le fait que dans un tel milieu les risques d'éventuelles difficultés liées au niveau de la « littéracie » familiale se trouvent minimisés. D'autre part, étant donné l'âge de ces élèves de 6^e année, ce groupe ne comportait aucun doubleur. Ainsi, nous pouvons penser que ces élèves n'éprouvent pas de problèmes particuliers d'apprentissage de la langue écrite.

3.3 Les corpus

Les dictées qui composent le corpus sont conçues en premier lieu pour l'étude des connaissances grammaticales impliquées dans les règles d'accord. Elles comptent au total 334 mots pour les francophones et 333 mots pour les hispanophones. Ce nombre comprend toutes les occurrences dans les textes. Quand nous considérons les mots sans leur répétition dans les textes, le nombre diminue considérablement : 212 mots différents pour les francophones et 210 pour les hispanophones.

En fonction des objectifs de notre étude, nous avons épuré quelque peu les textes. Pour les deux langues, nous avons retranché de l'analyse les nombres comme 42, 2002, etc., les noms propres comme Llyod, Marine, Londres, etc., et les petits nombres (deux, trois et six) écrits en lettres dans le texte dicté, mais souvent écrits en chiffre par les élèves hispanophones. Cette opération nous laisse, pour l'analyse, 315 mots dictés en français et 312 pour la traduction espagnole. Ce nombre compte toutes les occurrences

du corpus, lesquelles seront nécessaires pour calculer le pourcentage de réussite dans l'ensemble du corpus, ou encore le pourcentage de réussites en moyenne par élève.

Quand nous considérons seulement le nombre de mots différents nous obtenons par groupe comparé le nombre de mots suivant : 207 en français et 205 en espagnol, mais lors de l'analyse des types d'erreurs lexicales, les homophones grammaticaux ont été retirés du corpus, ce qui nous donne en final le nombre de mots suivants pour cette analyse: 197 en français et 203 en espagnol.

Pendant la collecte de données, nous avons dû composer avec quelques absents chez les francophones. En effet, 3 élèves ont écrit seulement trois dictées sur quatre. Par contre, les hispanophones ont participé à toutes les dictées. Nous présentons dans le tableau 3.1, ci-dessous, la répartition des élèves par dictée (le taux de participation).

Tableau 3.1– Répartition des élèves par dictée

	dictée 1	dictée 2	dictée 3	dictée 4
Nombre d'élèves francophones	38	40	40	39
Nombre d'élèves hispanophones	42	42	42	42

Par conséquent, pour les élèves absents, nous avons codifié tous les mots de la dictée non faite « *abs* ». Le problème de graphies manquantes s'est présenté sous une autre forme chez les hispanophones. Dans les dictées, chez presque tous les élèves (40 sur 42), nous trouvons quelques mots manquants ou qui ont été remplacés par un synonyme ou un autre mot totalement différent de celui dicté. Au total, 115 graphies ont été codées

« absentes » dans le corpus espagnol. Chez les francophones, nous trouvons seulement 7 élèves ayant omis des mots pour un total de 19 graphies manquantes, mais avec les 3 élèves absents à une dictée, nous avons un total de 285 graphies manquantes. Il en résulte que ces données manquantes n'ont aucune incidence négative sur nos calculs parce que la moyenne de mots écrits par élève ne présente pour ainsi dire aucune différence : 308 pour les francophones et 309 pour les hispanophones. Et surtout que chaque proportion d'erreur est calculée en fonction du nombre de mots écrits par l'élève.

Tableau 3.2 – Nombre de mots disponibles par corpus

	corpus francophone	corpus hispanophone
Nombre de mots dictés	12600	13104
nombre de mots absents	285	115
nombre de mots dans le corpus	12315	12988

Nous devons préciser que ce problème de mots absents ne risque pas d'entraîner de biais dans le calcul du pourcentage de mots réussis par chaque élève puisque le calcul est fait à partir du nombre de mots écrits par élève.

3.4 La fréquence des mots dans le corpus

Dans le cadre de notre recherche, nous avons présenté des échelles de fréquence en lien avec des théories sur les mots que les élèves apprennent le plus rapidement dans le processus d'acquisition de l'orthographe lexicale. Notre étude comparative se propose donc de répondre à ces deux questions :

- Dans la maîtrise de l'orthographe lexicale, est-ce que la fréquence des mots tient effectivement un rôle important ?

- Est-ce que la fréquence des mots joue un rôle identique dans chacun des systèmes orthographiques ?

Nous avons examiné pour chaque mot du corpus son échelle de fréquence en consultant les fréquences présentées dans Fortier (1993) pour le corpus français et Juilland (1964) pour le corpus espagnol. Nous présenterons plus loin ces listes de mots avec leur fréquence.

Afin de regrouper les mots dans des catégories de fréquence : peu fréquents, assez fréquents et très fréquents, nous avons dû déterminer les limites de ces catégories. En somme, à partir de quel indice de fréquence peut-on considérer un mot comme fréquent ? En parcourant les ouvrages sur les échelles de fréquence Baudot (1992), Fortier (1993), Préfontaine (1995), nous ne trouvons aucune indication sur les critères qui font qu'un mot est fréquent ou rare dans le discours. Notre seul indice vient de l'ouvrage de Préfontaine (1995) où les mots ayant une fréquence commençant à 100 sont mis en caractère gras. Une notification qui après comparaison des autres mots semble indiquer les mots fréquents en français.

En nous basant sur ce critère, nous avons compté les mots pour lesquels nous disposons d'une fréquence, puis nous avons fractionné le nombre obtenu pour avoir à peu près le même nombre de mots par catégorie. C'est à l'aide de cette méthode que nous avons procédé à l'élaboration des trois catégories : mots peu fréquents (PF) (fréquence 1-20), mots assez fréquents (AF) (fréquence 21-99), mots fréquents (F) (fréquence 100 et plus). Cette méthode conserve une part d'arbitraire parce que nous constatons que les mots jugés de haute fréquence dans le *Vadémécum de l'orthographe recommandée* (2004) ne tombent pas automatiquement dans les mêmes

catégories de fréquence. Prenons les mots suivants : aout, boîte, diner, revolver, voute, qui sont considérés comme des mots de haute fréquence. Quand nous les cherchons dans l'échelle De Baudot (1992), 4 de ceux-ci tombent dans la catégorie des mots AF, selon notre découpage. Cela dit, cette méthode nous donne bien des catégories assez semblables en français : 66 (PF), 62 (AF), 72 (F). Le désavantage touche le corpus espagnol où nous retrouvons, selon notre délimitation, des catégories non identiques : 69 (PF), 35 (AF), 68 (F). Nous pouvons remarquer qu'à l'exception des mots AF, les deux autres catégories comportent sensiblement le même nombre de mots. Cette différence s'explique par le nombre de mots pour lesquels nous avons trouvé une fréquence dans l'ouvrage de Juilland (1964). En effet, alors que nous disposons d'un rapport de 200/207 mots avec fréquence identifiable en français, nous ne disposons que d'un rapport de 172/205 mots en espagnol.

Le moyen de corriger cette différence serait de répartir (distribuer) les mots sans fréquence en espagnol dans la section où ils apparaissent en français et vice versa. Ainsi, nous conservons les mêmes limites de fréquence pour les deux langues. Comme nous l'avons mentionné dans le cadre théorique, nous avons trouvé des similitudes entre les deux échelles : 32 mots dont l'échelle de fréquence est supérieure à 100 et 69 mots dont l'échelle de fréquence est inférieure à 100 communs aux deux langues. Il est clair également que nous avons trouvé des divergences : en français, nous trouvons 23 mots dont la fréquence est supérieure à 100, mais qui ne partagent pas cette classification avec leur équivalent en espagnol.

Nous ne disposons pas d'une correspondance parfaite dans les fréquences des mots, mais nous ne pensons pas que la distribution des mots puisse nuire à nos résultats sur le rôle de la fréquence des mots dans

l'apprentissage de l'orthographe lexicale. D'ailleurs, les 33 mots du corpus espagnol qui n'ont pas de fréquence se répartissent ainsi : 18 mots dont leur équivalent français est PF ; 7 dont leur équivalent français est AF ; 3 dont leur équivalent français est F. Pour les cinq mots, *buzo*, *trapecistas*, *carpa*, *carpas*, *reconfortantes*, pour lesquels nous ne disposons d'aucune fréquence, nous constatons la même similitude en français. Par conséquent, nous les plaçons dans la catégorie des mots PF comme nous l'avons fait pour les 7 mots (sans fréquence) du corpus français. Cette distribution fait augmenter un peu plus la différence entre les catégories, mais elle a l'avantage de nous fournir des mots avec une indication de fréquence. Une opération qui nous sera utile plus tard dans l'étude de l'importance de la fréquence dans le processus d'acquisition de l'orthographe lexicale. Après ces modifications, la catégorie F est la seule qui, dans les deux langues, contient sensiblement le même nombre de mots. Nous présentons dans le tableau 3.3 ci-dessous nos catégories de fréquence. La première section du tableau contient les mots pour lesquels nous disposons d'une fréquence alors que la deuxième section présente l'ensemble des mots du corpus après distribution des mots sans fréquence trouvée. Les tableaux 3.4 et 3.5 présentent en détail, les mots de chaque langue selon leur catégorie de fréquence.

Tableau 3.3- Nombre de mots du corpus par catégorie de fréquence dans chaque langue

Fréquence/langue	Corpus des mots avec fréquence identifiée		Corpus de tous les mots après distribution	
	Français	Espagnol	Français	Espagnol
Peu fréquents (1-20)	66	69	73	92
Assez fréquents (20-99)	62	35	62	42
Fréquents (100 +)	72	68	72	71
Total de mots	200	172	207	205

Tableau 3.4- Liste de mots du corpus français et leur fréquence selon Fortier (1993)

peu fréquents	fréq.	assez fréquents	fréq.	fréquents	fréq.
attentivement	1	plaine	21	botte	101
caresses	1	y	21	appelle	107
compétiteurs	1	minutes	21	et	117
entraîneuse	1	illuminent	22	plusieurs	118
jongleuses	1	imaginée	22	très	118
particulièrement	1	enfilé	23	sortent	119
recueillir	1	équipe	24	sur	123
lui-même	2	charmants	25	je	125
talentueuses	2	cela	26	heures	132
ces	3	foule	28	elle	149
faut	3	paraissait	28	frère	149
collective	4	poussières	31	dans	157
confie	4	réussi	32	vous	159
depuis	4	secrets	32	dirait	173
duré	4	courage	33	dire	173
flatte	4	américains	34	accident	180
leucémie	4	au	34	qu'il	184
leucémies	4	aujourd'hui	34	que	184
spectateurs	4	colorés	34	jeunes	196
acrobates	5	du	36	arrivent	198
ça	5	leur	36	quelques	210
fascinent	5	restent	40	été	221
parmi	5	sous	40	pour	223
tous	5	ici	42	bicyclette	227
commencent	6	battre	43	tout	230
festival	6	coureurs	43	son	236
festivals	6	eux	44	bleus	246
jusqu'	6	extraordinaires	46	jeu	251
applaudit	7	exécutent	47	il	261
combinaison	7	si	49	Ils	261
immobiles	7	athlètes	50	jours	263
kilomètres	7	dessin	50	vis	279
mieux	7	personnages	51	les	293
présente	7	premières	52	pesait	297
cirque	8	genou	53	bonheur	347
cirques	8	vers	53	beaucoup	368
devant	10	journées	56	va	372
public	10	avancent	57	des	373
semblent	10	allongés	62	humains	388
vraiment	10	drôle	63	en	397
inquiète	11	comme	64	bien	419
observe	11	immenses	67	font	474
lutte	12	canadiens	70	était	529
équilibre	14	blessée	72	sont	529
gauche	14	par	74	suis	529

peu fréquents	fréq.	assez fréquents	fréq.	fréquents	fréq.
vaincu	11	nerveuse	68	s'est	533
records	14	souhaitait	74	chatte	546
réels	14	lenteur	76	douces	562
dessine	15	bouge	77	courir	565
ensuite	15	bouger	77	pas	574
marathon	15	paroles	77	a	578
contre	16	ne	82	avait	578
meilleures	16	années	84	ont	578
numéro	16	faciles	84	c'est	622
québécoise	16	me	84	moi	624
chaque	17	quels	85	parle	650
clowns	17	coup	86	le	658
normal	17	pièce	91	animaux	700
participé	17	ce	93	école	742
participent	17	cancers	94	grandes	745
demeurent	19	montés	98	couleurs	806
hier	19	souvent	99	petit	920
création	20			petite	920
kilogrammes	20			enfants	1001
presque	20			la	1012
ses	20			on	1044
chapiteau	*			à	1103
chapiteaux	*			un	1202
fonds	*			une	1202
trapézistes	*			l'un	1202
réconfortantes	*			maison	1411
s'élancent	*			de	2434
scaphandrier	*				

Tableau 3.5- Liste de mots du corpus espagnol et leur fréquence selon Juilland (1964)

peu fréquents	fréq.	assez fréquents	fréq.	fréquents	fréq.
comienzan	1	correr	21	número	145
encantadores	1	muchos	21	ellas	160
extraordinarios	1	unos	21	horas	161
hirió	1	animales	22	poco	171
iluminan	1	ve	23	les	174
imaginada	1	reales	24	contra	182
mueva	1	pieza	25	soy	186
participan	1	busca	26	palabras	188
pesaba	1	kilómetros	26	hacia	203
rodilla	1	presenta	29	se	220
venció	1	mejor	30	estaba	225
atletas	2	pequeña	34	había	234
dibuja	2	traje	36	mucho	254
mejores	2	ánimo	37	ellos	255
primeras	2	juego	38	quien	259
buenos	3	parecen	39	casi	270
espectadores	3	puesto	40	días	302
inquieta	3	personajes	41	Hoy	314
lanzan	3	creación	42	Cada	344
parecería	3	minutos	43	estos	344
permanecen	3	frecuencia	48	mi	414
colocaron	4	lucha	50	todo	419
corredores	4	delante	51	te	432
moverse	4	mucha	52	años	435
participaron	4	niños	56	hace	435
confía	5	colores	57	Después	453
observa	5	fuera	60	mismo	467
quedan	5	atención	65	casa	477
accidente	6	escuela	69	decir	482
aplaude	6	público	69	fue	495
caricias	6	parecía	70	Ella	523
repente	6	hacen	77	todos	546
americanos	7	habla	84	aquí	602
dibujo	7	varios	98	hasta	613
fondos	7	Ayer	99	tiene	640
inmóviles	7	canadienses	*	Entre	703
izquierda	7	cánceres	*	eso	715
fáciles	8	coloridos	*	ser	750
grandes	8	equipo	*	Hay	764
lentitud	8	planicie	*	Son	764
avanzan	9	qué	*	me	768
deseaba	9	recostados	*	muy	787
dulces	9			sobre	854
duró	9			él	1023
inmensas	9			si	1327

peu fréquents	fréq.	peu fréquents	fréq.	fréquents	fréq.
logró	9	trapecistas	*	Yo	1348
jornadas	10			Sus	1739
dicha	11			lo	2231
colectiva	12			como	2265
equilibrio	12			más	2363
frecuentemente	13			para	2684
gata	13			una	3708
nerviosa	13			su	3802
reunir	13			con	4667
Enseguida	14			por	4700
vencer	14			Es	4731
algunos	15			un	5033
curiosa	15			del	5340
jóvenes	15			las	5360
vivo	15			los	6343
salen	16			no	6900
Bajo	17			que	9274
humanos	17			a	10851
multitud	17			en	12922
normal	17			El	13339
particularmente	17			y	14712
secretos	17			la	18237
azules	20			de	29804
Realmente	20			bicicleta	*
acaricia	*			bota	*
acróbatas	*			hermanito	*
buzo	*				
carpa	*				
carpas	*				
circo	*				
circos	*				
competidores	*				
entrenadora	*				
fascinan	*				
festival	*				
festivales	*				
kilogramos	*				
leucemia	*				
leucemias	*				
malabaristas	*				
maratón	*				
payasos	*				
quebequense	*				
reconfortantes	*				
récords	*				
talentosas	*				

3.5 La méthode d'analyse

Dans le corpus étudié, les enfants francophones et hispanophones du même âge présentent-ils le même niveau de maîtrise sur le plan de l'orthographe en général, et en particulier, l'orthographe lexicale de leur langue première respective ?

La réponse à cette question demande un relevé systématique de toutes les graphies erronées produites par les élèves de chaque groupe. La correction des copies, à l'aide des dictées originales, francophones et hispanophones respecte les règles que l'on applique dans les écoles du Québec. Pour notre corpus, toutes les graphies non conformes, autre que les erreurs dans les noms propres, les majuscules, les petits nombres écrits massivement en chiffres par les hispanophones, ont été comptabilisées dans un premier temps. Une opération qui est faite pour nous aider à évaluer le nombre de mots parfaitement orthographiés, soit le nombre de réussites tant sur le plan de l'orthographe lexicale que grammaticale. Nous présentons un extrait des données pour les élèves francophones et les élèves hispanophones dans les tableaux 3.6 et 3.7. L'inventaire de toutes les graphies erronées permettra une analyse plus fine des sortes d'erreurs lexicales et l'évaluation de l'impact de diverses réformes orthographiques sur la compétence des francophones à l'écrit.

Dans un deuxième temps, les graphies non conformes ont été classées selon les deux grandes catégories suivantes : lexicale (incluant les homophones lexicaux et de discours) et grammaticale (incluant les homophones grammaticaux). Pour les deux langues, les mêmes critères ont été appliqués : tout mot qui présente une graphie non conforme aux conventions (graphie dont l'image ne correspond d'aucune façon à son entrée dans le dictionnaire) correspond à une erreur lexicale peu importe le nombre d'erreurs que nous

retrouvons dans le mot : en français, sca**f**endrier≠scaphandrier est compté pour une erreur, et en espagnol, **can**seres≠cánceres est compté pour une erreur ; les erreurs grammaticales englobent les notions d'accord du verbe avec son sujet, accord en genre et en nombre des noms, adjectifs, déterminants, accord du participe passé. Il s'agit également des paires d'homophone comme *ces/ses*, *sont/son*, *à/a*, *etc.* que nous retrouvons dans le corpus Nadeau, Matute et Gonzales-Reyes (2006) et que nous avons classées dans la catégorie grammaticale en accord avec une certaine tradition scolaire, ce qui veut dire que dans notre étude comparative toutes les erreurs sont comptabilisées pour les résultats globaux sur la performance orthographique des élèves de chaque groupe.

Enfin, il arrive parfois qu'un mot contienne à la fois une ou des erreurs lexicales en plus d'une erreur grammaticale. Par exemple nous pouvons citer « *les spectact**e**ur* » où nous relevons dans la graphie erronée « *spectacteur* » une erreur lexicale (adjonction du graphème c) et une erreur grammaticale (omission du morphème du pluriel). Dans *leu**s**emia* (au lieu de *leucemias*) nous retrouvons la même situation qu'en français : erreur lexicale et erreur grammaticale. Les mots contenant à la fois des erreurs d'ordre lexical et grammatical ont été comptés dans chaque catégorie.

Tableau 3.7- Extrait des graphies erronées et mots réussis par élève hispanophone

Mots dictés	n° 2 (f)	n° 3 (f)	n° 6 (f)	n° 8 (f)	n° 9 (f)	n° 10 (f)
Correr						
kilómetros				kilometros		
y						
un						
poco						
más						
qué	que		que	que		
te						
parecería	pareceria	pareceria		pareceria		
muchos						
humanos						
lo						
hacen						
en	abs					
minutos						
y						
algunos						
minutos						
Qué	que		que	que		
buenos						
Atletas						
En						
competidores						
participaron						
en						
un	una					
maratón						
Légende n° = numéro attribué à l'élève f = élève de sexe féminin g = élève de sexe masculin case vide = mots bien écrits par l'élève abs = mots non écrits par l'élève						

Dans un troisième temps, nous avons procédé au classement des erreurs lexicales des élèves francophones et hispanophones (Voir Appendice C) à l'aide de la typologie de Catach (1980) déjà présentée dans le cadre théorique. Les six catégories de la typologie de Catach (1980) ont servi à notre classification (Voir les exemples du tableau 3.8). Dans la section erreur homophonique nous ne présentons que les homophones lexicaux et de discours. Ils ont été relevés pour le caractère linguistique des erreurs lexicales et parce que ce sont les seuls qui sont comptabilisés dans le cadre

de notre comparaison des habiletés de chaque groupe sur le plan de l'orthographe lexicale.

En tenant compte des nombreuses confusions « phono-graphiques » dans la prononciation espagnole, nous avons dû apporter plus de précision dans la section des erreurs phonogrammiques pour mieux rendre compte des graphies erronées que nous retrouvons dans le corpus espagnol. Ainsi, dans la section phonogrammique, nous avons une section pour les confusions b/v (*vota#bota*), une section pour les confusions s/c/z/x (*dulses#dulces, isquierda#izquierda*).

Tableau 3.8- Exemples d'erreurs selon la typologie de Catach

Type d'erreurs/langue	corpus francophone	corpus hispanophone
Erreur phonétique	spectacteur (spectateurs)	acletas (atletas)
Erreur phonogrammique	poucières (poussières)	bisicleta (bicicleta)
Erreur morphogrammique (lexicale)	anné (année)	façiles (fáciles)
Erreur homophonique	cou (coup)	
Erreur idéogrammique	jusqu'a (jusqu'à)	ninos (niños)
Erreur lettres non fonctionnelles	maraton (marathon)	

Tenant compte du fait dans une graphie erronée nous pouvons trouver plus qu'une erreur, nous avons dans un premier temps listé en français certaines graphies erronées dans plus d'une section de la typologie de Catach (1980). En espagnol, à cause de l'importance dans la confusion de graphèmes réalisant le même son pour les mots de notre corpus (*b#v, c#s, s#z, s#x*, etc.), le fait de classer un mot dans plus d'une section de la typologie de Catach (1980) est plus courant dans la section phonogrammique. Pour illustrer cette classification multiple, prenons le mot « marathon ». Dans les copies des élèves francophones nous relevons la graphie erronée « **marraton** », il y a adjonction du graphème « r » ce qui est une erreur

phonogrammique ; il y a également omission de la consonne grecque « h » ce qui correspond à une erreur de lettres non fonctionnelles. En espagnol, la graphie erronée « *esquierda* » (*izquierda*) présente la confusion « s≠z » et une erreur phonétique. Par conséquent, cette graphie erronée pourrait apparaître dans les sections erreurs phonogrammiques et phonétiques comme nous l'avons suggéré lors de la présentation de la typologie de Catach dans le cadre théorique. Cependant, nous avons travaillé de manière plus systématique en clarifiant, dans le cadre théorique, ce qui est une erreur phonogrammique, morphogrammique et phonétique. Un travail qui nous a permis d'avoir un portrait plus clair de l'ensemble des erreurs lexicales pour notre corpus. La graphie erronée « *esquierda* » étant éloignée du mot original plus par la sonorité que la confusion graphémique (s≠z), elle se retrouve dans la section des erreurs phonétiques et compte donc pour une seule erreur dans nos calculs.

3.6 Les calculs

Notre étude comparative de la maîtrise orthographique chez les francophones et les hispanophones doit considérer les types d'erreur et les sections de la typologie de Catach (1980) où nous pouvons retrouver le plus de graphies non conformes aux conventions. Nous présentons maintenant les calculs qui doivent être faits pour nous permettre de répondre à nos questions de recherche et de vérifier nos hypothèses de recherche.

Notre première question porte sur le niveau de maîtrise de l'orthographe générale dans chaque population, à savoir : Quelle est la population (francophone/hispanophone) qui présente en général le meilleur pourcentage de réussite sur le plan de l'orthographe ?

Nous devons donc mesurer la réussite chez les francophones et les hispanophones. Les premiers calculs à effectuer sont les suivants :

- Total de mots bien écrits par élève ;
- Pourcentage de mots bien écrits par élève ;
- Moyenne de mots bien écrits par élève chez les francophones et les hispanophones.

Ces résultats nous permettront également de vérifier notre première hypothèse de recherche : le taux de réussite (moins de graphies non conformes dans les dictées) devrait être plus élevé chez les hispanophones que chez les francophones à cause de la régularité du système orthographique de l'espagnol. À l'aide de la moyenne de mots bien écrits, nous pourrions voir s'il existe un écart considérable entre les deux groupes sur le plan de la maîtrise de l'orthographe de leur langue respective. Mentionnons que ces calculs tiennent compte de toutes les occurrences.

Pendant la collecte de données, comme nous l'avons déjà mentionné, nous avons pris le soin de relever toutes les graphies erronées présentes dans les textes des élèves. Nos calculs tiennent compte des deux grandes catégories définies dans le cadre théorique : orthographe lexicale, orthographe grammaticale. Les chiffres obtenus nous serviront à comparer le niveau de maîtrise de l'orthographe générale dans chaque population. Mais notre analyse s'intéressera davantage à la compétence en orthographe lexicale des élèves de 12 ans du Québec et du Mexique qui composent notre corpus. Après avoir comparé la réussite globale, nous avons cherché à savoir si le pourcentage de réussite pour le groupe le plus performant se transpose sur le plan de la moyenne de réussite en orthographe grammaticale, et plus particulièrement en orthographe lexicale. Nous présentons les calculs qui seront effectués pour nous aider dans nos comparaisons :

- Total d'erreurs lexicales, grammaticales par élève ;

- Pourcentage d'erreurs lexicales, grammaticales, par élève ;
- Total d'erreurs lexicales, grammaticales par groupe ;
- Moyenne d'erreurs lexicales, grammaticales, par groupe.

Pour étayer davantage la réponse sur la maîtrise (compétence) en orthographe lexicale, nous procéderons à différentes analyses des données du corpus. À l'aide du tableau de la typologie des erreurs lexicales nous pourrions monter le tableau du nombre de graphies erronées pour chaque mot en français et en espagnol. Ainsi, nous pourrions voir rapidement chez quel groupe (francophone ou hispanophone) nous retrouvons le plus grand nombre de graphies erronées différentes par mot. Nous vérifierons également la répartition des élèves selon les erreurs lexicales par intervalle (ordre croissant d'erreurs) et groupe comparé. Cette vérification nous donnera un portrait de la compétence en orthographe lexicale des élèves de chaque groupe comparé.

En ce qui a trait à notre question sur le rôle de la fréquence des mots dans la maîtrise de l'orthographe, nous pensons atteindre notre but en comparant les mots qui dans la langue respective de chaque groupe causent le plus de difficulté aux élèves et les mots qui peuvent être considérés comme mieux maîtrisés par les élèves selon un pourcentage d'erreurs lexicales décroissant, et ceux où ne retrouvons aucune erreur d'aucun élève. Pour notre première catégorie, nous porterons notre choix sur les mots non réussis par au minimum 50 % des élèves de chaque groupe. Pour cela, il faudra considérer les mots mal orthographiés au minimum par 20 élèves du corpus français. Et comme nous avons deux élèves hispanophones de plus, nous prendrons les mots non réussis par au minimum 21 élèves. Pour les mots ayant plus d'une occurrence, nous ferons la moyenne pour voir s'ils rentrent dans notre limite.

En faisant le même travail avec les mots considérés comme étant les mieux réussis (le seuil choisi est les mots pour lesquels nous trouvons un pourcentage d'erreurs inférieur à 5 %) et les mots sans aucune erreur d'aucun élève, nous aurons des éléments de réponse pour établir notre comparaison sur le rôle que joue la fréquence dans la réussite de l'orthographe lexicale dans chaque langue. Par la même occasion, nous pourrions vérifier l'hypothèse qui prédit que les mots les moins bien réussis par les élèves de chaque population devraient se retrouver dans la catégorie des mots peu fréquents (ou rares) ; et que les mots qui leur causent le moins de problème, dans la catégorie des mots fréquents (F).

Pour nos dernières questions de recherche sur la comparaison des habiletés en orthographe lexicale,

- Quels sont les types d'erreurs lexicales que nous retrouvons dans les textes des élèves de chaque langue et dans quelles proportions ?

Les réponses seront fournies selon une analyse fine de la typologie des erreurs. En examinant les deux tableaux de classification des erreurs lexicales (français vs espagnol), il faudra vérifier la présence d'erreurs lexicales de même type chez les deux groupes comparés et pour quel groupe chacun de ces types d'erreurs (phonogrammique, morphogrammique, phonétique) représente un problème.

Il est clair que pour vérifier si nous rencontrons les mêmes types d'erreur dans chaque population, nous devons également procéder à une analyse linguistique de nos graphies erronées. Nous verrons si dans le processus de la maîtrise orthographique, les élèves francophones et hispanophones suivent le même la même logique pour produire les graphies qu'ils ne

maitrisent pas. Rappelons Lalande (1989) qui nous explique que « la production d'une graphie repose essentiellement sur deux types d'information en mémoire : une représentation visuelle complète ou partielle du mot permettant sa reconnaissance, et une liste de graphèmes (phonogrammes) possibles pour chacun des phonèmes de la langue » (p.36). Ceci signifie que l'élève n'écrit pas au petit bonheur de sa fantaisie. L'apprenant qui se retrouve en situation de production écrite choisit sa graphie erronée ou correcte selon une logique dictée par les connaissances du code orthographique emmagasinées.

Notre dernière question porte sur les réformes et leurs impacts possibles sur la compétence orthographique des élèves francophones, nous devons, à travers une argumentation, choisir le projet de réforme qui nous semble le plus compatible avec notre corpus. Par la suite, nous dresserons des tableaux par type d'erreur en indiquant le pourcentage du corpus représenté par ces types d'erreur. Ensuite, nous procéderons à une projection des projets de réforme qui risquent d'avoir plus d'impact sur les graphies erronées de notre corpus. Nous pourrons ainsi voir quelles sont les sections de l'orthographe du français qui gagneraient à subir plusieurs rectifications.

3.7 Les limites de notre recherche

Les résultats présentés dans le cadre de cette recherche proviennent essentiellement de l'analyse faite du corpus Nadeau, Matute, Gonzales-Reyes (2006) mis à notre disposition. Ce qui signifie que nos constatations peuvent se heurter aux limites posées par la clientèle, les mots dictés, la place de l'enseignement de l'orthographe et les attitudes sociales face à la norme orthographique dans chaque population.

En effet, les enfants retenus pour cette étude ne sont pas représentatifs de la population entière. L'échantillonnage sélectif ne permet donc pas d'avoir une vision de la maîtrise orthographique globale, toutes les couches sociales réunies. Loin de remettre en question le portrait de réussite que nous montre le groupe analysé, nous pensons que notre recherche gagnerait en profondeur si des écoles de quartier moins bien nanti s'ajoutaient au corpus. Toutefois, le milieu socioéconomique, le sexe et l'âge des élèves ainsi que leur scolarité ayant été contrôlés, la comparaison de ces deux groupes présente un bon niveau de validité.

En ce qui concerne les mots de notre corpus, malgré une grande variété, ils ont été choisis par des adultes pour être dictés dans un protocole scolaire. Enfin, nous nous demandons si les résultats n'auraient pas été différents si l'élève avait pu choisir les mots qu'il écrit dans le cadre d'un sujet libre ou donné par l'enseignant. Une autre limite de ce corpus, c'est le fait qu'ayant été conçu d'abord pour présenter une variété de cas d'accords grammaticaux, les difficultés que les mots présentent sur le plan lexical n'ont pas été contrôlées.

Pour finir, nous n'avons pas contrôlé les méthodes d'enseignement de l'orthographe dans les classes des écoles concernées. Est-ce que dans les classes du Mexique les enseignants mexicains procèdent de la même manière que chez les francophones pour l'enseignement de l'orthographe (vocabulaire à mémoriser, récitation, dictée, etc.) ? Est-ce qu'au Mexique, les enseignants accordent autant de temps à l'orthographe comme matière scolaire ? Des réponses qui nous auraient peut-être aidé à comprendre le fait que les hispanophones omettent d'écrire des mots ou décident de les remplacer par des synonymes alors qu'ils sont en situation de dictée. Nous pensons également que cette observation sur le terrain nous donnerait la

chance d'identifier clairement cette attitude : méconnaissance du code ou pure désinvolture des élèves qui n'ont pas l'habitude d'être soumis à ce type d'exercice parce que l'orthographe n'est pas une matière importante.

Cependant, malgré les limites reconnues, cette étude comparative sur le corpus Nadeau, Matude, Reyes-Gonzales (2006) mettra des chiffres sur les affirmations de Masson (1991) et de Jaffré et Fayol (1997) qui affirment une meilleure compétence orthographique des enfants italophones et hispanophones. Elle nous permettra de voir, à travers nos calculs et nos analyses, si les hispanophones sont effectivement plus compétents que les francophones sur le plan de l'orthographe en général et lexicale en particulier, ou si les enfants du même âge présentent le même niveau de compétence sur le plan de l'orthographe lexicale dans les deux langues.

Conclusion

Dans ce chapitre, nous avons présenté le matériel sur lequel repose notre comparaison et le cheminement de notre analyse. Nous avons commencé par les élèves en montrant les critères sur lesquels ils ont été retenus. Nous avons ensuite présenté le corpus composé de textes en français et de la traduction en espagnol tout en précisant les modifications opérées sur le matériel de base dans le cadre de notre recherche.

Par la suite, nous avons explicité la méthode d'analyse de nos données brutes avant d'introduire les calculs qui ont fourni les résultats présentés au chapitre suivant, nous permettant ainsi de répondre à nos questions de recherche et de vérifier nos hypothèses de recherche.

Chapitre IV

L'analyse des résultats

4.1 Introduction

Dans ce chapitre, nous présentons les résultats de notre étude comparative sur la maîtrise de l'orthographe du français et de l'espagnol chez des élèves de 12 ans. Dans un premier temps, il s'agit de présenter en lien avec notre première question de recherche, les pourcentages de réussite des deux populations en orthographe générale et plus particulièrement en orthographe lexicale. Dans un deuxième temps nous présentons les résultats en lien avec le rôle des échelles de fréquence dans l'apprentissage de l'orthographe lexicale. Dans un troisième temps, nous analysons les types d'erreur et leur répartition dans les deux langues. Nous terminons ce chapitre avec l'impact d'une réforme en français, choisie parmi celles présentées dans le cadre théorique, sur les compétences orthographiques des élèves francophones avant de présenter la conclusion générale de notre recherche.

4.2 La maîtrise de l'orthographe

La présentation des résultats passe par le rappel de notre hypothèse de recherche qui s'énonce comme suit : le taux de réussite en orthographe (moins de graphies erronées dans les dictées) devrait être plus élevé chez les hispanophones que chez les francophones étant donné la régularité du système orthographique de l'espagnol. Pour démontrer ou réfuter cette hypothèse, nous commençons cette analyse en répondant à notre première question de recherche à travers les sous questions qu'elle englobe.

Dans le corpus étudié, les enfants francophones et hispanophones du même âge présentent-ils le même niveau de maîtrise sur le plan de l'orthographe en général, et de l'orthographe lexicale en particulier, dans leur langue première respective ?

Nous prévoyons donner la réponse à cette question à la fin de cette section, car nous pensons qu'il importe de procéder par différentes étapes : trouver le groupe présentant le meilleur pourcentage de réussite générale et voir si la réussite du groupe en question se transpose sur le plan de l'orthographe lexicale pour démontrer les arguments sur lesquels repose notre réponse finale. Cela dit, nous allons répondre à la sous-question suivante : **quelle est la population (francophone/hispanophone) qui présente le meilleur pourcentage de réussite globale ?**

Le tableau 4.1 ci-dessous présente le pourcentage de réussite dans chaque langue. Ces premiers résultats montrent un pourcentage de réussite élevé dans les deux populations qui, rappelons-le, appartiennent à un milieu socioéconomique relativement aisé et scolarisé (83.96 % chez les francophones, 90.72 % chez les hispanophones). Toutefois nous constatons que dans l'ensemble, les élèves mexicains réussissent mieux que les élèves québécois. Cette réussite va dans le sens de notre hypothèse. Mais que se passe-t-il sur le plan individuel ?

Tableau 4.1 – Pourcentage de réussite dans chaque langue

	français	espagnol
total des mots bien écrits/ total de mots écrits	10340/12315	11784/12988
% de réussite dans l'ensemble du corpus	83.96	90.72
% de réussite en moyenne par élève	83.82	90.73

Le pourcentage de réussite en moyenne par élève va dans le même sens que la réussite globale. Les élèves mexicains réussissent mieux que les élèves francophones selon les chiffres de notre tableau 4.1. Cependant, nous remarquons que l'écart entre les deux groupes est assez faible : 6.76 % pour les résultats globaux et 6.91 % pour la moyenne de réussite par élève.

Cet écart constaté se perçoit dans le tableau 4.2 où nous procédons à la répartition des élèves selon le pourcentage de réussite. Nous remarquons qu'il n'y a chez les hispanophones qu'un seul élève sous la barre des 80 % de moyenne de réussite alors que nous en comptons 14 (soit le tiers des élèves) chez les francophones, dont le plus faible pourcentage de réussite chez les francophones : 63.43 %.

Tableau 4.2- Répartition des élèves selon le pourcentage de réussite chez les francophones et les hispanophones

Pourcentage de réussite	corpus francophone		corpus hispanophone	
	Nbre d'élèves	% d'élèves	Nbre d'élèves	% d'élèves
100 > 90	13	32.5 %	22	52.4 %
90 > 80	13	32.5 %	19	45.2 %
80 > 70	9	22.5 %	1	2.4 %
70 > 60	5	12.5 %	0	0 %
Total	40	100 %	42	100 %

À travers ces deux tableaux, il est clair que les hispanophones présentent non seulement le meilleur pourcentage de réussite, mais aussi des performances plus homogènes au sein du groupe tandis que les performances des francophones s'avèrent plus étalées. Cet étalement montre un niveau très variable en orthographe parmi ces élèves. Ce qui confirme notre hypothèse qui prédisait que le plus haut taux de réussite devrait se retrouver chez les hispanophones. Maintenant, il convient de voir, pour les besoins de notre étude, la répartition des erreurs selon leur catégorie

lexicale et grammaticale. Le résultat obtenu à l'aide de cette répartition nous permettra de répondre à la question suivante : **Est-ce que cette plus grande réussite pour les hispanophones se transpose dans la maîtrise de l'orthographe lexicale ?**

Avant tout nous devons préciser qu'un même mot peut contenir une erreur lexicale et une erreur grammaticale. Par conséquent, un tel mot compte deux fois pendant l'étude du corpus. Le pourcentage total d'erreurs (17.51 %) est par voie de conséquence plus élevé que le pourcentage total de mots avec erreur (16.04 %) dans le corpus français. Dans le cas du corpus espagnol, les erreurs grammaticales étant moins nombreuses, l'écart entre les deux nombres est vraiment minime (9.39 % vs 9.37 %).

Tableau 4.3- Répartition des erreurs selon leur catégorie, lexicale, grammaticale

Type d'erreurs :	Dans l'ensemble des corpus		nbre d'erreurs en moyenne par élève	
	français	espagnol	élèves français	élèves espagnols
nbre d'erreurs lexicales/nbre de mots écrits	1024/12315	995/12988	25.60	23.69
% du corpus	8.32 %	7.66 %	8.79 %	7.51 %
nbre d'erreurs grammaticales/nbre de mots écrits	1133/12315	225/12988	28.83	5.36
% du corpus	9.34 %	1.73 %	9.04 %	1.86 %
<i>grand total :</i>	<i>2157/12315</i>	<i>1220/12988</i>		
%	17.51 %	9.39 %	17.83 %	9.37 %

Dans le tableau 4.3 qui montre la répartition des erreurs selon leur catégorie, nous constatons comme déjà vu avec les pourcentages de réussite que les francophones ont un cumulatif d'erreurs nettement plus élevé que les

hispanophones. En fait, le nombre d'erreurs pour les francophones correspond à 1,8 fois (presque le double) le nombre total des erreurs dénombrées chez les élèves hispanophones. Les résultats en orthographe grammaticale nous font voir une réalité connue dans le milieu scolaire du Québec. D'ailleurs, il convient de rappeler que le corpus (les dictées) ayant servi à notre étude était conçu pour la reconnaissance des marques silencieuses d'accord en français. L'orthographe grammaticale représente effectivement la pierre d'achoppement des élèves francophones. Un phénomène qui n'a rien de surprenant parce que l'étude de Manesse et Cogis (2007), en France, arrive à la même conclusion.

52 % des erreurs de 2005 résultent de règles grammaticales non appliquées ou de catégories non identifiées, contre 40 % en 1987. On peut ainsi faire l'hypothèse, [...] que ce sont essentiellement ces erreurs grammaticales qui « plombent » les résultats de 2005. (Manesse et Cogis, 2007, p.89).

Dans notre corpus, les hispanophones font quatre fois moins d'erreurs grammaticales que les francophones : 1.86 % de moyenne d'erreurs grammaticales par élève hispanophone comparativement à 9.04 % par élève francophone. Des résultats qui sont (probablement) dus au fait que l'espagnol est, contrairement au français où nous retrouvons beaucoup de consonnes muettes dans les marques d'accord, une langue où tout se prononce. De plus, en français, les formes masculines et féminines peuvent être différentes : *frère*, *sœur*, comme elles peuvent offrir qu'une simple différence graphique : *ami*, *amie*. À l'inverse, l'élève hispanophone a un indice oral du genre dont ne dispose pas toujours l'élève francophone en situation de dictée à cause de la prononciation plus régulière de l'espagnol (Denina, 2003).

Par exemple, *las palabras* doit être plus facile à écrire pour un hispanophone et un apprenant de l'espagnol langue seconde que *les paroles* pour un

francophone et un apprenant de français langue seconde. Cette difficulté réside dans le fait qu'en réception du discours l'indication du pluriel n'est donnée qu'en début du groupe nominal malgré la redondance de cette marque du pluriel à l'écrit (Guion, 1974) : *les enfants dansent tous les jours*. Les exemples de ce type ne manquent pas entre le français et l'espagnol, et le corpus Nadeau, Matute et Gonzales-Reyes (2006), qui, au départ, était conçu pour une étude des marques d'accords dans les deux langues, nous offre plusieurs correspondances où le mot français contient des indices morphogrammiques du pluriel qui ne se prononcent pas. Une particularité linguistique qui est tout à fait le contraire de l'espagnol : athlètes/*atlelas*, personnages/*personajes*, journées/*jornadas*, etc.

Revenons toutefois à notre intérêt premier : l'orthographe lexicale. Les résultats du tableau 4.3 montrent, qu'en moyenne, les hispanophones font sensiblement le même nombre d'erreurs lexicales que les francophones du même âge. Effectivement, les francophones présentent un pourcentage de 8.32 % d'erreurs lexicales, ou 25.60 erreurs en moyenne par élève et les hispanophones, 7.66 % ou 23.69 erreurs par élève.

La question à se poser est la suivante : Que nous disent réellement ces moyennes ? Regardons le tableau de la répartition des élèves selon le pourcentage d'erreurs lexicales chez les francophones et les hispanophones (voir tableau 4.4). Le tableau nous donne le portrait suivant :

Tableau 4.4 - Répartition des élèves selon le % d'erreurs lexicales chez les francophones et les hispanophones

% d'erreurs lexicales	corpus francophone		corpus hispanophone	
	nbre d'élèves	% d'élève	nbre d'élèves	% d'élève
0 < 5	11	27.5 %	15	35.7 %
5 < 10	15	37.5 %	13	31 %
10 < 15	8	20 %	12	28.5 %
15 < 20	4	10 %	2	4.8 %
20 < 25	2	5 %	0	0 %
Total	40	100 %	42	100 %

Le premier point que nous pouvons retenir dans ce tableau, c'est qu'aucun élève n'a fait plus de 25 % d'erreurs lexicales. La répartition des élèves dans le tableau 4.4 révèle que les élèves francophones et hispanophones ont des performances semblables, les hispanophones ayant des performances à peine plus homogènes que les francophones. Nous retrouvons un grand nombre d'élèves dans les bonnes performances. Environ les 2/3 des élèves tant francophones qu'hispanophones font moins de 10 % d'erreurs lexicales sur une moyenne de plus de 300 mots écrits. Et nous retrouvons également un moins grand nombre d'élèves dans les faibles performances (plus d'erreurs). Cette diminution est plus marquée chez les hispanophones où nous ne comptons que 4.8 % des élèves dans les deux sections qui comptent le plus d'erreurs lexicales alors que ce pourcentage est de 15 % chez les francophones.

Nous constatons dans ce tableau, qu'il y a 27.5 % des francophones qui ont produit moins de 5 % de graphies erronées. Il est vrai que les hispanophones présentent un léger avantage (35.7 %), mais ces résultats montrent que l'orthographe complexe du français est très bien acquise pour une grande partie des élèves du corpus. L'enseignement semble donc leur réussir alors qu'un certain nombre d'élèves produit de nombreuses graphies erronées.

Nous savons qu'en français, c'est dans la tradition scolaire d'insister sur l'orthographe, même si nous n'avons pas contrôlé l'enseignement reçu par ces élèves dans ce domaine tout au long de l'école primaire. En espagnol, vu ces résultats nous pouvons nous questionner sur l'enseignement reçu dans ce domaine et sur la valeur accordée au respect de la norme orthographique malgré la transparence du système orthographique prônée par Jaffré et Fayol (1997) et Masson (1991).

En conclusion, les résultats obtenus montrent dans un premier temps que les hispanophones réussissent mieux globalement (90.72 % vs 83.96 %) et individuellement (90.73 % vs 83.82 %) lorsque nous considérons les mots bien écrits. De plus à travers le tableau 4.3, les résultats nous montrent que les francophones font globalement presque deux fois plus d'erreurs que les hispanophones du même âge (2157 vs 1220). Des résultats qui viennent corroborer les affirmations de Masson (1991) et de Jaffré et Fayol (1997), et qui confirment notre première hypothèse de recherche : les hispanophones réussissent mieux que les francophones.

Cependant, pour aborder la maîtrise orthographique, nous pouvons affirmer que le seul point où nous constatons une supériorité sans équivoque des hispanophones, c'est dans la maîtrise de l'orthographe grammaticale. Comme nous l'avons déjà dit, les francophones font des erreurs grammaticales dans une proportion de 4 pour 1 face aux hispanophones dont les marques d'accord s'entendent à l'oral. Ces résultats expliquent assez bien le pourcentage de réussite plus faible chez les francophones parce que selon Manesse et Cogis (2007), « les connaissances nécessaires en français pour écrire sans faute couvrent en réalité une large partie de la grammaire. » (p.104). Par conséquent, en situation d'écriture,

L'élève doit connaître les lettres qui sont des marques autorisées et les marques exclues ; il faut encore passer d'un plan à l'autre, penser à la fois à ce qui s'entend et à ce qui ne s'entend pas, et à leur possible succession. Ce sont des subtiles distinctions que les élèves (...) qui apprennent à écrire en français doivent comprendre, retenir et appliquer. (Ibid., p.105)

Quant à la maîtrise de l'orthographe d'usage, les résultats obtenus nous obligent à rejeter notre hypothèse de recherche, car la réussite des hispanophones ne se transpose pas dans la maîtrise de l'orthographe lexicale. Aucune supériorité de ces derniers sur les francophones n'apparaît même si au total les hispanophones ont commis un peu moins d'erreurs lexicales que les francophones. En effet, la différence du pourcentage d'erreur lexicale par élève entre les deux groupes est minime : 7.66 % pour les hispanophones comparativement à 8.32 % pour les francophones.

En somme, les élèves francophones et hispanophones du même âge présentent le même niveau de maîtrise sur le plan de l'orthographe lexicale de leur langue respective. Une réponse valable pour le corpus étudié qui, rappelons-le, était conçu pour la reconnaissance des marques d'accord silencieux en français.

Pour faire suite à cette section, nous pensons qu'il importe de poursuivre notre étude à l'aide des analyses sur les types d'erreurs lexicales pour voir si cette maîtrise équivalente est réelle à travers les graphies erronées des élèves.

4.3 Le rôle de la fréquence dans la maîtrise de l'orthographe lexicale

Dans cette section, nos résultats seront examinés en lien avec la question de recherche suivante : ***Dans la maîtrise de l'orthographe lexicale, est-ce que la fréquence des mots tient effectivement un rôle important ?***

Nous avons vu dans le cadre théorique que selon Bonin (2005), les mots fréquents sont appris bien plus vite que les mots plus rares. De plus, nous avons présenté les travaux de certains chercheurs comme Baudot (1992), Fortier (1993), Juilland (1964) qui ont établi des échelles de fréquence pour guider les enseignants dans le choix des mots à enseigner aux élèves, suggérant de commencer par les mots plus fréquents.

Par conséquent, nous pouvons rappeler notre deuxième hypothèse de recherche : les mots qui causent le plus de difficulté aux élèves de chaque population se retrouveront dans la catégorie des mots peu fréquents (ou rares) ; et les mots qui leur causent le moins de problème, dans la catégorie des mots fréquents.

Nous pensons qu'il est possible d'avoir la réponse à notre question principale en analysant les deux extrêmes de notre corpus, ce qui veut dire les mots les mieux réussis (aucune erreur et pourcentage d'erreur de 5 %) et les mots qui ont causé le plus de difficulté (pourcentage d'erreur de 50 % et plus) aux élèves de chaque population. Par ce type d'analyse, la comparaison des résultats selon la langue des élèves, nous sommes d'avis que nous pourrions répondre à la sous-question : ***Est-ce que la fréquence des mots joue un rôle identique dans chacun des systèmes orthographiques ?***

Pour les besoins de notre étude, il convient de préciser que les analyses que nous ferons dans cette section prennent en compte les catégories de mots (Peu Fréquents, Assez Fréquents, Fréquents) obtenues lorsque nous avons subdivisé les mots du corpus français à l'aide de l'échelle de Fortier (1993) et les mots du corpus espagnol à l'aide de l'échelle de Juilland (1964). (Voir les tableaux 3.4 et 3.5 du chapitre sur la méthodologie)

Nous commençons notre analyse en présentant pour le français et l'espagnol respectivement la répartition des mots du corpus selon leur pourcentage de graphies non conformes (% d'erreur) produites par les élèves de chaque population et leur fréquence.

Tableau 4.5 - Nombre de mots selon le pourcentage d'erreur (graphies non conformes) et la fréquence en français

Corpus français				
Fréquence des mots				
% d'erreurs	Mots Fréquents	Mots Assez fréquents	Mots Peu fréquents	Total de mots
0%	41	11	11	63
< 1% < 5 %	6	9	6	21
5% < 10 %	7	12	16	35
10% < 15%	4	9	8	21
15% < 20%	5	7	3	15
20% < 25%	0	3	5	8
25% < 30%	2	4	1	7
30% < 35%	0	2	5	7
35% < 40%	0	0	3	3
40% < 45%	0	1	3	4
45% < 50%	0	2	3	5
50% et plus	0	1	7	8
Total de mots	65	61	71	197

Tableau 4.6 - Nombre de mots selon le pourcentage d'erreur (graphies non conformes) et la fréquence en espagnol

Corpus espagnol				
Fréquence des mots				
% d'erreurs	Mots Fréquents	Mots Assez fréquents	Mots Peu fréquents	Total de mots
0%	37	13	22	74
< 1% < 5 %	19	11	25	55
5% < 10 %	3	6	11	20
10% < 15%	2	2	9	13
15% < 20%	2	2	2	6
20% < 25%	2	0	3	5
25% < 30%	0	2	1	3
30% < 35%	2	1	2	5
35% < 40%	1	1	2	4
40% < 45%	1	1	2	4
45% < 50%	0	1	0	1
50% et plus	1	1	13	15
total	70	41	92	203

Le tableau 4.5 montre qu'en français plus le pourcentage d'erreurs augmente, moins on trouve de mots fréquents. On remarque par exemple que 41 mots fréquents sur 65 (soit 63 % des mots F) sont réussis par tous les élèves et que seulement 2 mots (soit 3 % des mots F) se trouvent parmi les mots ayant un taux d'erreur de 20 % ou plus. À l'inverse, les mots peu fréquents ayant un faible taux d'erreur sont moins nombreux ; ces mots se répartissent davantage dans toute la gamme des taux d'erreur. Ce premier coup d'œil confirme un effet de la fréquence des mots sur la réussite de l'orthographe en français. Pour l'espagnol, le tableau 4.6 montre un effet moins marqué de la fréquence des mots sur la réussite de l'orthographe. Si de nombreux mots fréquents ont comme en français un faible taux d'erreur, les mots peu fréquents bien réussis sont aussi très nombreux. Pour mieux comparer l'effet de la fréquence des mots dans les deux langues, les tableaux 4.7 et 4.8 présentent ces mêmes résultats en pourcentage, et pour des taux d'erreurs regroupés en tranche de 10 %.

Tableau 4.7 - Nombre de mots selon le pourcentage d'erreur (graphies non conformes), la fréquence et le % des mots selon la catégorie de fréquence en français

Corpus français							
% d'erreurs	Mots F	% des mots F	Mots AF	% des mots AF	Mots PF	% des mots PF	Total de mots
0% < 10%	54	83	32	52	33	46.4	119
10% < 20%	9	14	16	26.2	11	15.5	36
20% < 30%	2	3	7	11.5	6	8.4	15
30% < 40%	0	0	2	3.3	8	11.3	10
40% < 50%	0	0	3	5	6	8.4	9
50% et plus	0	0	1	2	7	10	8
total	65	100	61	100	71	100	197

Tableau 4.8 - Nombre de mots selon le pourcentage d'erreur (graphies non conformes), la fréquence et le % des mots selon la catégorie de fréquence en espagnol

Corpus espagnol							
% d'erreurs	Mots F	% des mots F	Mots AF	% des mots AF	Mots PF	% des mots PF	Total de mots
0% < 10%	59	84	30	73	58	63	147
10% < 20%	4	6	4	10	11	12	19
20% < 30%	2	3	2	5	4	4	8
30% < 40%	3	4.2	2	5	4	4	9
40% < 50%	1	1.4	2	5	2	2	5
50% et plus	1	1.4	1	2	13	14	15
total	70	100	41	100	92	100	203

Dans ces tableaux, nous constatons que, dans les deux populations, les mots PF sont en général bien réussis. En considérant uniquement le pourcentage de graphies non conformes produites par les élèves inférieur à 10 %, nous relevons, du côté des francophones, 33 mots PF (soit 46.4 % des mots PF du corpus), ce qui montre que pour un grand nombre de mots rares, leur basse fréquence n'affecte pas leur réussite orthographique. Cette réussite est encore plus évidente chez les hispanophones où le pourcentage des mots PF apparaissant dans cet intervalle monte à 63 %.

Chez les francophones, nous ne trouvons aucun mot dans la colonne des mots F quand le pourcentage de graphies erronées atteint 30 % et plus. La grande majorité des mots se retrouvent dans l'intervalle inférieur à 10 % d'erreurs (83 % des mots Fréquents du corpus). Nous pouvons donc affirmer que les mots fréquents sont maîtrisés plus vite par les élèves du corpus français. Pour les mots AF qui sont moins nombreux dans le corpus, la répartition dans les intervalles de pourcentage suit à peu près la même logique que celle des mots F (52 % dans l'intervalle inférieur à 10 % d'erreurs). Nous pouvons remarquer que le nombre de mots AF diminue également plus le pourcentage de graphies non conformes augmente. Dans le cas des mots PF, nous trouvons un grand nombre de mots (33) dans

l'intervalle inférieur à 10 %. Cependant, nous constatons une distribution plus constante de mots malgré le fait que le pourcentage d'erreur augmente (15 %, 8 %, 11 %, 8 %, 10 %). Comme le font voir ces pourcentages, le nombre de mots ne diminue pas aussi vite que dans le cas des deux autres catégories (F et AF). En somme, les élèves ne semblent pas maîtriser une partie des mots PF aussi vite que les mots F et AF.

Chez les hispanophones, nous retrouvons, tout comme chez les francophones, la plus grande concentration de mots F (84 % des mots Fréquents du corpus) dans l'intervalle inférieur à 10 % d'erreurs. La différence entre les deux groupes, c'est le fait que l'on relève toujours au moins un mot dans cette colonne même si le pourcentage d'erreur augmente. Les mots AF sont moins nombreux (l'écart est plus marqué qu'en français), mais ils se retrouvent majoritairement, avec un fort taux de réussite dans l'intervalle inférieur à 10 % d'erreurs (73 % des mots AF du corpus). De plus, nous constatons une certaine régularité dans la répartition des mots AF plus le pourcentage d'erreur augmente. Quant aux mots PF, nous avons 13 mots du corpus ratés (production de graphies non conformes aux conventions) par 50 % et plus des élèves hispanophones tandis que le nombre est de 8 mots pour les francophones. Rappelons que notre subdivision fait apparaître plus de mots Peu Fréquents pour les hispanophones dans le corpus Nadeau, Matute et Gonzales-Reyes (2006). Cependant, tout comme en français, les mots PF se retrouvent en très grande proportion (63 % des mots F du corpus) dans l'intervalle inférieur à 10 % d'erreurs. Comme nous pouvons le voir, à travers ces tableaux, les mots F, les mots AF et les mots PF sont, en général, bien réussis par les hispanophones. Une réussite qui est plus marquée dans les mots PF que du côté des francophones.

Cela dit, il apparaît que la fréquence joue un rôle dans l'acquisition de l'orthographe lexicale parce que plus le pourcentage d'erreur augmente, plus le nombre de mots diminue dans les colonnes des mots F et AF chez les francophones. Ce rôle semble être moins important en espagnol où les mots PF réussis (-10 % d'erreurs) sont aussi nombreux que les mots F réussis (58 et 59 mots respectivement). De plus, il est facile de remarquer, chez les hispanophones, une plus grande fluctuation (12 %, 4 %, 4 %, 2 %, 14 %) dans la colonne des mots PF plus les pourcentages d'erreur augmentent.

Devant le fait qu'autant de mots PF sont réussis dans les deux populations, (et de manière plus importante en espagnol : 63 % des mots PF réussis à moins de 10 % d'erreurs contre 46 % en français) nous pensons qu'il convient de vérifier également la transparence ou l'opacité des mots (Jaffré et Fayol, 1997) de chaque système pour comprendre ce phénomène. Nous commencerons notre analyse par les mots ayant causé le plus de problèmes aux élèves des deux groupes. Il s'agit des mots que l'on retrouve dans l'intervalle de 50 % et plus d'erreur. Ainsi, nous verrons si l'échec est dû à la rareté du mot ou aux différentes possibilités de transcription dont dispose l'élève face à celui-ci. Nous parlons essentiellement des phonèmes qui peuvent être transcrits par plus d'un graphème. Par la suite, nous présenterons les mots F et PF pour lesquels le taux d'erreur est inférieur à 5 % (mots réussis par tous les élèves ou ratés par un maximum de 5 % des élèves de chaque population). Nous arrêtons notre choix sur cet intervalle parce que nous pensons qu'il nous donnera un portrait assez clair de la situation.

Nous présentons donc le tableau des mots mal orthographiés sur le plan lexical par au minimum 50 % des élèves de chaque population.

Tableau 4.9 – Mots mal orthographiés par 50 % et plus des élèves francophones et hispanophones

français			espagnol		
<i>mot français</i>	fréquence	% d'erreurs	<i>mot espagnol</i>	fréquence	% d'erreurs
immenses	AF	50	confía	PF	50
marathon	PF	57,50	después	F	52,38
québécoise	PF	60	leucemia	PF	52,38
fascinent	PF	62,50	pareceria	PF	54,76
trapézistes	PF	65	cánceres	AF	57,14
recueillir	PF	65	récords	PF	57,14
scaphandrier	PF	67,50	hirió	PF	59,52
lui-même	PF	82,50	jóvenes	PF	59,52
			leucemias	PF	61,90
			acróbatas	PF	64,29
			buzo	PF	64,29
			fáciles	PF	64,29
			inmóviles	PF	69,05
			quebequense	PF	73,81
			fascinan	PF	90,48
Légende PF = mot peu fréquent AF = mot assez fréquent F = mot fréquent					

Pour les élèves du corpus espagnol, comme nous le montre le tableau 4.9, sur un total de 15 mots manqués par la moitié et plus des hispanophones 12 (nous considérons *leucemia* et son pluriel comme le même mot) de ces mots apparaissent dans la catégorie des mots PF. Ces mots sont PF (ou rares) dans le discours selon l'échelle de fréquence de Juilland (1964), mais il importe de rappeler, de surcroît, que, pour notre étude, les dictées ont été administrées à une clientèle de 12 ans, alors que les mots ayant servi à construire cette échelle proviennent de textes de scripteurs experts.

Par conséquent, nous allons voir si en plus de la rareté de certains mots qui sont très éloignés de la réalité des élèves hispanophones (*leucemias*, *cánceres*, *buzo*, *quebequense*, *récords*), les 15 mots du tableau présentent

des pièges pour les scripteurs. L'examen de la correspondance phonème/graphème en espagnol (dans sa variété mexicaine) nous montre que les mots contenant le phonème /s/ représente une zone sensible à cause des différents graphèmes pouvant transcrire ce son (c, s, z, x). Il convient de préciser que ce phénomène est assez rare en espagnol, car il n'existe que deux phonèmes (/k/ et /s/) qui offrent autant de choix dans la transcription d'un mot en espagnol. Ainsi, l'élève hispanophone qui doit transcrire les mots *leucemias*, *cánceres*, *quebequense*, *buzo*, *pareceria*, *fáciles*, *fascinan*, risque de produire différentes graphies. Il faut également attirer l'attention sur les mots *acró**b**atas*, ***b**uzo*, *inmó**v**iles*, *jó**v**enes* qui contiennent le phonème /b/ pouvant être représenté indifféremment par les graphèmes « b ou v » en espagnol. Dans le cas du mot *fascinan*, quand nous regardons les différentes graphies produites par les élèves [*fascinan* (27); *fasinan* (11)] et le pourcentage d'erreur (90 %), nous constatons que le digramme « sc » est un graphème très peu connu des élèves hispanophones. De plus, dans le tableau 4.9, 9 mots sur 15 contiennent l'accent tonique souvent omis par les élèves hispanophones dans les graphies erronées produites (comme nous le verrons dans la section 4.4), dont *después* qui est classé dans la catégorie de mots F. L'autre point qui attire l'attention, c'est la longueur des mots. Nous comptons 5 mots à 2 syllabes et 10 à 3 syllabes dont *cánceres* qui contient l'accent tonique et qui est classé dans la catégorie des mots AF. Nous verrons plus loin si ce facteur peut avoir une incidence sur la réussite ou la non-réussite des mots. Cela dit, il est clair que dans ce contexte, la rareté des mots dans les échelles de fréquence n'explique pas à elle seule les mots mal orthographiés par au minimum 50 % des élèves hispanophones.

Dans le cas des élèves francophones, nous retrouvons dans le tableau 4.9 à peu près la même distribution qu'en espagnol : massivement des mots PF (7 mots dans la catégorie PF, 1 seul mot AF). En reprenant l'analyse de la plus

ou moins grande biunivocité des graphèmes dans ces mots, nous remarquons que le mot « **immenses** » (classés AF) contient une double consonne en plus de la confusion existant entre les graphèmes en/an. Un doublement qui cause beaucoup de problème aux scripteurs francophones, ainsi que les graphèmes « s et c » qui sont des choix possibles dans la transcription du mot « **immenses** ». Devant le grand nombre de consonnes doubles en français, les scripteurs ne savent pas quand ils doivent écrire la simple consonne ou la consonne double selon Beslais (1965).

Pour les autres mots, nous allons les présenter un à un (et regrouper ceux qui présentent les mêmes difficultés orthographiques). Le mot *marathon* contient un « h » tout comme le mot *scaphandrier*. La différence entre les deux, c'est le fait que celui de *marathon* est muet alors que celui de *scaphandrier* offre deux possibilités graphémiques en français : « ph ou f ». Les élèves francophones qui ont à transcrire ce mot doivent choisir entre ces deux graphèmes tout en s'attardant sur d'autres difficultés comme les phonèmes /k/ et /ã/ qui offrent plus qu'un choix de graphème. Dans le cas du mot « fascinent », il apparaît évident que le digramme « sc » représente une zone sensible (possibilité de recourir à plusieurs graphies pour transcrire un seul phonème). Si les mots « trapézistes et recueillir » ne sont pas maîtrisés par les élèves francophones, nous nous posons des questions sur les mots comme « lui-même » et « québécoise » qui, contrairement aux élèves hispanophones, fait partie de la réalité quotidienne des élèves francophones. Mentionnons que pour le mot « lui-même » le trait d'union est la principale source d'erreur lexicale dans le corpus francophone. Sur le plan de la longueur, les mots à 3 syllabes sont plus nombreux (5 mots). Ainsi, comme nous l'avons vu chez les hispanophones, la rareté n'explique pas à elle seule la production de graphies erronées chez les francophones.

Maintenant, toujours pour essayer de comprendre la réussite d'autant de mots PF, nous passons à la deuxième partie de notre analyse en présentant les mots F et PF pour lesquels nous retrouvons un pourcentage d'erreur inférieur à 5 % dans les deux langues.

Tableau 4.10 – Mots F et PF dont le pourcentage d'erreur est inférieur à 5 % chez les francophones

Pourcentage d'erreurs	Mots fréquents		Pourcentage d'erreurs	Mots peu fréquents
0 %	avait	le	0 %	contre
	beaucoup	les		depuis
	bien	l'un		faut
	bleus	maison		gauche
	de	parle		hier
	des	pas		normal
	dire	petit		participé
	elle	petite		participent
	en	pour		tous
	enfants	que		demeurent
	était	qu'il		confie
	été	sortent		
	font	suis	1,25 %	semblent
	frère	sur		
	grandes	tout	2,5 %	devant
	il	très		festival
	ils	un		numéro
	je	une		public
	jeunes	va		talentueuses
	jours	vous		
	la			
0,5 %	et			
1,25 %	sont			
2,5 %	couleurs			
	dans			
	moi			
	quelques			

Comme nous l'avons déjà vu, les mots F sont bien réussis dans les deux langues. Dans le tableau 4.10, nous remarquons que, chez les francophones, nous retrouvons dans cette catégorie un grand nombre de mots outils comme : je, il, elle, le, la, un, une, pas, pour, etc. Certains de ces mots présentent bien des pièges. Prenons l'exemple du mot « enfants » à cause

de la confusion possible entre les graphèmes en/an (ou même le phonème /f/ pouvant être réalisé par les graphèmes « f et ph », et le mot « beaucoup » à cause de la consonne muette à la fin du mot et, bien entendu, à cause des différents graphèmes disponibles en français pour transcrire le phonème /o/ et le fait que ce mot contienne une consonne muette à la finale. Mais, dans l'ensemble, il s'agit de mots courts (37 mots à une syllabe) et de mots courants de tous les jours qui ne semblent pas présenter beaucoup de difficulté pour l'apprenti scripteur. Nous exposons à titre d'exemple quelques mots avec leur fréquence relevée dans Fortier (1993) : une (1202), la (1012), maison (1411).

Dans le cas des 17 mots PF du tableau 4.10, nous comptons quatre verbes conjugués (**demeurent**, **confie**, **participent**, **semblent**). Bien évidemment, la plupart des graphies erronées touchent principalement à la terminaison verbale. Nous pouvons supposer que la réussite sur le plan lexical semble s'expliquer par le fait que ces verbes présentent dans la forme dictée un radical stable. Pour les autres mots, regardons s'ils présentent une orthographe opaque telle que définie par Jaffré et Fayol (1997) : « une orthographe dans laquelle la relation entre phonèmes et graphèmes est très irrégulière » (p.119). Quand nous analysons les mots du tableau 4.10, nous constatons que certains mots comme « depuis, faut, hier et devant » ont une consonne muette à la finale que le scripteur pourrait omettre dans la transcription du mot. Pour un élève qui ne maîtrise pas l'orthographe lexicale, des mots comme « confie, contre et public » font entendre les phonèmes /k/ et /ʃ/ qui peuvent être transcrits par plus d'un graphème. De plus, le mot « participent » contient le phonème /s/ qui peut représenter une zone dangereuse pour l'apprenti scripteur. Cependant, nous pouvons constater une grande différence entre les mots PF ratés massivement et les mots PF bien réussis par les francophones. Ces derniers ne renferment pas de

consonnes doubles, pas de lettres grecques (th ou ph) ni de graphèmes rares comme « z et sc ». Des difficultés orthographiques qui nuisent à la compétence des scripteurs francophones, selon Beslais (1965) et Jaffré et Fayol (1997).

Maintenant, regardons ce qui se passe sur le plan de la longueur des mots. Dans le tableau 4.10, les mots PF bien réussis se répartissent comme suit : 3 mots à une syllabe, 9 mots à deux syllabes et 5 mots à trois syllabes et plus (participé, participent, festival, numéro, talentueuses) alors que nous en comptons 2 mots à deux syllabes et 5 mots à trois syllabes et plus dans les mots PF ratés par 50 % et plus des élèves francophones. Il y a une certaine correspondance dans les mots à trois syllabes, mais nous sommes d'avis que ces mots PF sont bien réussis parce qu'ils sont bien connus des élèves. Les quelques pièges que renferment ces mots PF réussis ne semblent pas vraiment nuire à leur maîtrise. C'est la raison pour laquelle nous nous questionnons sur la présence de certains mots comme « depuis, faut, festival, gauche, hier, tous, devant, numéro et public » dans cette catégorie. Est-ce que les échelles de fréquence consultées représentent un indice fiable pour les occurrences d'un mot dans la langue écrite ou est-ce notre découpage qui a été trop large ? Nous comparons donc, dans le tableau 4.11, les mots PF réussis et ratés avec leur indice de fréquence relevée dans Fortier (1993).

Tableau 4.11 – Mots PF réussis et ratés avec leur indice de fréquence selon Fortier (1993)

Mots Peu Fréquents réussis		Mots Peu Fréquents ratés	
Mot	Indice de fréquence	Mot	Indice de fréquence
depuis	4	marathon	15
contre	16	québécoise	16
faut	3	fascinent	5
gauche	14	trapézistes	*
hier	19	recueillir	1
normal	17	scaphandrier	*
participé	17	lui-même	2
participent	17		
tous	5		
demeurent	19		
confie	4		
semblent	10		
devant	10		
festival	6		
numéro	16		
public	10		
talentueuses	2		
Moyenne de fréquence	9,95	Moyenne de fréquence	0,15
Légende			
* = mot sans fréquence trouvée			

Dans le tableau 4.11, parmi les 9 mots (depuis, faut, festival, gauche, hier, tous, devant, numéro et public) qui nous paraissaient moins rares, nous en comptons 4 (depuis, faut, festival, tous) dont l'indice de fréquence est inférieur à 10. L'indice de fréquence moyenne ne nous offre pas assez d'éléments pour remettre en question les échelles de fréquence. Nous croyons donc que notre découpage pour définir les mots PF aurait dû être plus restreint en établissant la limite à 9 pour cette catégorie. Surtout quand nous constatons que le mot « québécoise » a un indice de fréquence supérieur à 10. D'autant plus que le corpus ayant servi de matériel à la construction de l'échelle de Fortier (1993) provient d'élèves québécois. Par conséquent, nous aurions eu 6 mots PF réussis et 5 mots PF ratés. Un résultat qui n'aurait rien changé à notre analyse précédente parce que les 5 mots PF ratés par 50 % et plus des élèves francophones renferment soit le digramme grec soit les graphèmes rares comme « sc et z », ce qui les rend

opaques pour le scripteur francophone contrairement à « depuis, festival, faut, tous, etc. »

Maintenant, nous poursuivons notre comparaison en analysant les mots PF réussis et les mots PF ratés par les hispanophones à l'aide du tableau 4.12.

Tableau 4.12 – Mots F et PF dont le pourcentage d'erreur est inférieur à 5 % chez les hispanophones

Pourcentage d'erreurs	Mots fréquents		Pourcentage d'erreurs	Mots peu fréquents	
0 %	a	me	0 %	aplaude	imaginada
	casa	mucho		carpa	inquieta
	con	no		circo	lentitud
	de	palabras		competidores	logró
	del	por		corredores	mejores
	el	se		curiosa	multitud
	ella	ser		duró	nerviosa
	ellas	si		fondos	primeras
	ellos	sobre		gata	reunir
	entre	son		grandes	rodilla
	es	su		iluminan	salen
	eso	te			
	horas	tiene	1,60 %	dibujo	
	hoy	todo			
	la	todos	2,38 %	azules	entrenadora
	las	un		carpas	humanos
	les	una		colectiva	jornadas
	lo	y		dibuja	normal
	los			dicha	quedan
				encantadores	secretos
0,60 %	sus		3,60 %	kilogramos	
1,20 %	muy				
1,50 %	en		4,74 %	permanecen	
2,38 %	años	hace	4,76 %	atlelas	inmensas
	bota	mí		buenos	mueva
	cada	mismo		circos	talentosas
	como	para		enseguida	vivo
	contra	poco		equilibrio	
	él	soy			
3,60 %	yo				
4,76 %	decir				
	estaba				
	quien				

Comme nous pouvons le voir dans le tableau 4.12, un grand nombre de mots-outils : *con, de, ella, su, sus, y, no, les, etc*⁴. se retrouvent parmi les mots F comme chez les francophones. Ils sont courts et ne renferment ni consonnes muettes ni accent tonique. Remarquons que nous comptons seulement 2 mots à 3 syllabes dans cette colonne du tableau (versus zéro du côté des francophones). Notre analyse des mots F nous révèle plusieurs mots où la relation entre phonème et graphème n'est pas biunivoque. Les mots *casa, hace, decir, eso, mismo, estaba* ont le phonème /s/ qui offre au scripteur un choix de 4 graphèmes. Les mots *hace, horas, hoy* contiennent le graphème h qui a une valeur zéro dans la chaîne de prononciation. Les mots *estaba, bota, palabras, sobre* ont pour leur part le phonème /b/ qui se transcrit indifféremment par les graphèmes « b ou v ». En somme, dans les mots F du tableau 4.12 nous retrouvons des mots transparents qui ne présentent aucune difficulté sur le plan de l'orthographe lexicale et des mots que nous pouvons définir comme plus opaques parce qu'il n'existe pas une relation biunivoque entre les phonèmes et les graphèmes. Pour ces mots, la fréquence explique bien la réussite.

L'espagnol est considéré par Jaffré et Fayol (1997) et Masson (1991) comme une langue se rapprochant de la transparence. Par conséquent, il est logique de penser que la plupart des mots PF réussis ne présentent pas de difficultés majeures pour les hispanophones parce qu'en général la transcription reflète la prononciation du mot. Pourtant, l'analyse des mots nous dévoile une réalité qui se rapproche des mots que nous retrouvons dans l'intervalle de 50 % d'erreur. Un grand nombre de mots contiennent les phonèmes /s/, /b/, /k/ et /χ/. Il s'agit des phonèmes qui peuvent être transcrits par plus d'un graphème. À titre d'exemple, nous citons les mots *vivo, mueva, buenos, colectiva,*

⁴ avec, de, elle, son, ses, et, non, leur, etc.

*pesaba, dibuja, dibujo*⁵ dans lesquels les graphèmes « b et v » servent à transcrire le phonème /b/. Nous avons également les mots *secretos, permanecen, nerviosa, inmensas*⁶ qui contiennent le phonème /s/ pouvant être transcrits par quatre graphèmes (c, s, z, x). Il convient de signaler la présence du mot *humanos* dont le graphème h a une valeur zéro dans la prononciation espagnole. Il représente donc une source d'erreur possible pour l'élève hispanophone. De plus, il importe de préciser que les mots réussis sont tout aussi longs que la plupart des mots qui ont causé le plus de problèmes aux élèves hispanophones (50 % d'erreurs et plus). Nous comptons 4 mots à 4 syllabes et 7 à 5 syllabes. Pour tout dire, les mots PF, tout comme les mots F, que nous retrouvons dans le tableau 4.12 ne sont pas transparents. Nous en comptons 20 (soit 42 % des mots du tableau) pour lesquels la transcription nécessite une bonne maîtrise du code orthographique de la part du scripteur.

Cependant, nous nous questionnons sur la pertinence du mot « *gata* » (chatte) qui se retrouve dans la catégorie des mots peu fréquents d'après les textes dépouillés par Juilland (1964) pour composer son dictionnaire de fréquence. Les textes de base ont été écrits par des adultes pour des adultes, mais en constatant la grande réussite de ce mot dans les deux groupes (seulement 6 erreurs sur la double consonne chez les francophones), nous pensons, par voie de conséquence, que son classement est inexact quant à la réalité de nos sujets. La réalité a sans doute changé depuis 1964 à cause de la grande place qu'occupe la littérature jeunesse dans les maisons d'Éditions. D'ailleurs, ce mot apparaît bien dans notre catégorie de mots fréquents en français.

⁵ vis (je), bouge, bons, collective, pesait, dessine, dessin.

⁶ secrets, restent, nerveuse, immenses.

À travers cette analyse, nous constatons que les élèves des deux groupes produisent des graphies erronées dans toutes les catégories de fréquence de mots (F, AF et PF). L'explication que nous pouvons fournir est la suivante : les graphies non conformes aux conventions montrent que le ou les mots ne sont pas maîtrisés par celui ou ceux qui produisent la graphie erronée. Nous remarquons cependant que les mots PF réussis par les hispanophones renferment beaucoup de zones dangereuses : des phonèmes qui peuvent être représentés par plus d'un graphème.

En conclusion, nous pouvons dire, à la lumière de nos résultats, que la fréquence semble jouer un certain rôle dans la maîtrise de l'orthographe lexicale. En effet, tel qu'énoncé dans notre hypothèse de recherche, les mots qui causent le plus de difficulté aux élèves de chaque population, soit les mots mal orthographiés par 50 % ou plus des élèves, se retrouvent massivement dans la catégorie des mots peu fréquents (ou rares) : il s'agit de 7 mots sur 8 en français et de 13 mots sur 15 en espagnol. Toutefois, l'inverse n'est pas vrai. De nombreux mots peu fréquents sont tout de même très bien réussis : 46 % des mots PF ont un pourcentage d'erreurs inférieur à 10 % chez les francophones et cette proportion est de 63 % chez les hispanophones.

En ce qui a trait à la deuxième partie de notre hypothèse : les mots qui causent le moins de problème aux élèves de chaque population comparée se retrouveront dans la catégorie des mots fréquents, elle est partiellement démontrée à l'aide des tableaux 4.7 et 4.8. Effectivement, les mots du corpus où nous retrouvons moins de 10 % d'erreurs se retrouvent en grande partie dans la catégorie des mots F. Pour le français, 54 mots sur 119 très réussis sont F ; ces mots F très réussis représentent 83 % des mots F du corpus. Pour l'espagnol, bien que 84 % des mots fréquents soient réussis, il ne s'agit

que de 59 mots sur les 147 mots fréquents du corpus en espagnol, soit autant que les PF réussis (58 mots).

Devant la grande réussite des mots PF, nous avons analysé les graphies pour voir si ces mots présentaient moins de difficulté sur le plan orthographique. Nous avons également comparé la longueur des mots pour voir si ce facteur aidait à la maîtrise de ces mots. L'analyse des graphies permet de constater que pour les 17 mots PF réussis en français les difficultés orthographiques sont moins importantes que dans le cas des mots PF ratés massivement par les élèves francophones. Ce facteur joue donc un rôle dans l'acquisition rapide de ceux-ci. Nous avons remarqué chez les hispanophones la présence de nombreux mots réussis qui présentent des zones sensibles, mais pas d'accent tonique comme dans le cas des mots ratés par 50 % et plus des élèves. Pour la longueur des mots, si ce facteur semble jouer en faveur des élèves francophones, pour les élèves hispanophones, il ne semble pas avoir un impact parce que les mots PF réussis sont tout aussi longs que les mots PF ratés massivement. À travers la section sur les types d'erreurs, nous essaierons de comprendre pourquoi les élèves hispanophones produisent autant de graphies non conformes à la norme orthographique alors que l'espagnol, contrairement au français, est considéré comme une langue dont le système orthographique s'approche de la transparence. Pour finir, il convient d'aborder les échelles de fréquence. Dans le cadre de ce corpus, il apparaît clair qu'elles présentent certaines limites quant à notre clientèle. Les mots proviennent de textes d'expert, ce qui veut dire qu'un grand nombre de mots tombés dans notre section des mots PF ne semblent pas si rares dans l'univers des élèves de 12 ans. Nous pensons donc qu'il serait mieux de présenter des échelles de fréquence par niveau scolaire.

Le tableau 4.13 nous montre que tous les types d'erreurs de Catach sont identifiés dans les deux langues. Nous pouvons voir que les erreurs lexicales se concentrent, pour chaque population, dans la catégorie des erreurs phonogrammiques. La catégorie des erreurs phonétiques arrive en deuxième position pour le plus grand nombre d'erreurs dans les deux langues. Nous constatons que les proportions diffèrent toutefois beaucoup d'une langue à l'autre : 83.22 % des erreurs chez les élèves hispanophones sont phonogrammiques comparativement à 55.70 % chez les francophones. Les erreurs phonétiques sont courantes en français (28.30 % des erreurs lexicales), mais deux fois moins fréquentes en espagnol (12.20 % des erreurs lexicales). En ce qui a trait aux autres catégories, trois des quatre contiennent chacune moins de 1 % d'erreur en espagnol. En français, on trouve en moyenne 4 % d'erreurs dans chacune des autres catégories.

Maintenant, nous allons procéder à une analyse plus fine des différentes sortes d'erreurs communes aux deux langues que nous retrouvons dans la catégorie des erreurs phonogrammiques. Nous pensons que cette analyse nous permettra de voir les zones sensibles en français et le pourquoi de ce grand nombre d'erreurs phonogrammiques chez les hispanophones.

La première sorte d'erreur que nous présentons concerne l'ajout de lettres inaudibles à un mot. Nous utilisons ce vocabulaire pour préciser la nature de l'erreur lexicale, mais selon la typologie de Catach (1980), ce sont des graphies erronées classées dans la section des erreurs phonogrammiques. Par ajout de lettres inaudibles, nous entendons des graphies qui s'éloignent des conventions établies uniquement par l'adjonction d'un ou de plusieurs graphèmes (consonnes ou voyelles). Dans le cadre de cette analyse, nous avons pris le soin de relever tous les mots qui contiennent une erreur lexicale de cette sorte. En procédant à cet inventaire, nous pouvons constater, dans

le tableau 4.14, que les erreurs d'ajout de lettres apparaissent très souvent chez les francophones et très rarement chez les hispanophones. Cette réalité s'explique, bien entendu, par le fait que l'espagnol est une langue sonore : tout se prononce. Par conséquent, l'adjonction d'un graphème altère la valeur phonique du mot espagnol : exemple : *contra#contras* ; contrairement à aujourd'hui#*aujourd'hui*s, en français⁷.

Tableau 4.14 – Erreurs phonogrammiques de lettres inaudibles ajoutées dans chaque langue

français			espagnol
acrobates (2)	courage (2)	immensses	<i>halgunos</i>
acrobattes (2)	courreurs (4)	jonguleuses	<i>extrahordinarios</i>
animeaux (3)	coureures	lenteure (2)	<i>recoord</i>
athelètes (5)	courreur (3)		<i>rrepente (2)</i>
aujourd'hui (3)	coureure	lenteures	
haujour'hui	courir (7)	parraissait	
aujourd'hui	courir (2)	parmis (14)	
biscyclette	durér	parmie	
bonheure	durré	parolles	
bougers	entraînneuse	parolle	
cannadien	extraordinnaires (2)	particulièrement	
carresse (4)	festivale	poussières	
carresses (3)	festivales	presque	
chaques (5)	foulle (3)	reccords	
clowne	hummain	recueuillir (2)	
clownes	illuminne	réeussie	
collorées	imaginée	talenttueuses	
compétiteures	imaginer (2)	tallentueuses	
couleures	imaginner		
Légende			
(n) = nbre d'élèves ayant produit cette graphie si plus d'un élève			

Dans le tableau 4.14, les graphies erronées relevées représentent 18.80 % (107/570) du total des erreurs phonogrammiques des élèves francophones

⁷ Le même phénomène se produit en français. Citons l'exemple du verbe « pesait », l'ajout du graphème « s » change la valeur phonique du mot « pessait ». Cette graphie est donc une erreur phonétique.

(ou 10.54 % du total des erreurs lexicales) comparativement à 0.6 % chez les hispanophones. Un résultat qui représente une proportion de 31 pour 1.

En ce qui a trait au phénomène d'ajout de lettres (consonnes ou le graphème « e ») très courant en français, nous pouvons pointer du doigt la présence dans le système orthographique du français d'un grand nombre de lettres muettes (*longtemps*, *devant*, *sursis*, *manquement*, etc.) De plus, devant le nombre élevé de mots à double consonnes, il arrive que l'élève hésite sur l'orthographe du mot et choisisse souvent la solution la plus difficile, la consonne doublée (Beslais, 1965). Les quelques mots que nous avons placés dans le tableau témoignent de l'ampleur du problème dans le système orthographique du français. En effet, dans le tableau 4.14 nous comptons 51 graphies erronées ou 8.95 % du total des erreurs phonogrammiques (à peu près la moitié des graphies erronées du tableau) où les consonnes sont doublées inutilement. Est-ce à dire que la réforme de l'orthographe du français doit s'attaquer à la réduction des doubles consonnes ? Nous essaierons de répondre à cette question pendant la projection des projets de réforme retenus dans le cadre de cette étude (section 4.5).

L'analyse des erreurs en espagnol nous montre que dans les graphies erronées, l'ajout du « h » apparaît dans deux mots. Il est vrai que la présence du « h », dont la valeur phonique est zéro, reste problématique en espagnol comme en français. Des personnages importants comme le prix Nobel de Littérature, Gabriel García Márquez, ont demandé son abolition (Pombo, 1999). Étant une lettre instable, nous pensons qu'il n'est pas étonnant de voir des élèves hispanophones l'ajouter ou le retrancher en écrivant. Nous sommes d'avis qu'une étude approfondie auprès de cette population nous permettrait de voir si les graphies erronées de notre corpus constituent un

cas isolé ou si la tendance a une certaine fréquence dans les textes des scripteurs hispanophones.

La deuxième sorte d'erreur que nous présentons se situe dans les zones de fragilité communes aux deux langues. Une zone de fragilité correspond aux endroits, dans le système orthographique, particulièrement propices à l'occurrence de variations non conformes aux conventions (Jaffré et Fayol, 1997). Nous pouvons citer à juste titre les exemples pour les phonèmes /ʒ/ en français et /x/ en espagnol : *jenoux*≠*genoux* ; *imajinada*≠*imaginada*. De la même manière, les élèves des deux groupes confondent souvent les différents graphèmes mis à leur disposition, dans leur langue respective, lors de la transcription du phonème /s/. Comme nous l'avons déjà mentionné, cette zone de difficulté phonogrammique cause beaucoup de problèmes aux élèves mexicains de notre corpus à cause du nombre de graphèmes (s, x, z, c) pouvant transcrire le phonème /s/. La particularité de ces graphèmes, c'est de ne pas altérer la valeur phonique du mot en espagnol.

En français, la situation est différente. Le graphème « s » s'il n'est pas doublé entre deux voyelles produit le son ou phonème /z/ ; et le graphème « c » produit le son /s/ devant les graphèmes « i » et « e », devant les graphèmes « o, u, a », il cède sa place à un autre graphème (exemple : reçu) pour représenter le phonème /s/. Les francophones doivent connaître les différentes règles qui régissent l'orthographe pour faire des choix selon le contexte, alors qu'en espagnol, il est vrai que le graphème « c » peut produire le phonème /k/ devant « a, o, u », mais de façon régulière, tous les graphèmes mis à la disposition du scripteur mexicain pour transcrire le phonème /s/ sont plus souvent interchangeables sans aucune incidence sur la valeur phonique du mot écrit.

Tableau 4.15 – Erreurs phonogrammiques sur le phonème /s/ dans chaque langue

français	espagnol	
avançent (2)	acarisia (9)	leusemias (19)
avense	avansan (7)	moverce
biscyclette	abansan (2)	parecen (6)
cansère	avancan	paresería (2)
cansers	bisicleta (10)	pareseria (6)
çela	buso (27)	paresia
commencent	canadiences (12)	partisipan (3)
commensses	canseres (6)	partisiparon (2)
dousses	cánseres (2)	permanesen (2)
dousse	caricias (9)	piesa (5)
façile	carpaz	piesa
fassine (2)	cací	planisie (3)
facine (7)	sircos	precenta
fassinent	comiensan (3)	quebequence (12)
façinent	desir (2)	quevequence
facinent (11)	deceaba (3)	quévequence
facinne	dulses (4)	que vequence (2)
facines	expectadores	cecretos
imence (3)	fasiles (2)	trapesistas (14)
immensses	facinan (27)	trapeziztas
immence	fasinan (11)	venser (2)
immences (10)	festibalez	benser (2)
l'accident	hasen	bensio
l'acsxident	hasia (2)	vensio
paraiceait	a sia	vensió
pièce	asia (2)	
poucières	hasia (2)	
cecrets	isquierda (7)	
c'élanssent	lansan (7)	
s'élansent	leusemia (24)	
si (3)		
Légende (n) = nbre d'élèves ayant produit cette graphie si plus d'un élève		

La difficulté des hispanophones à transcrire le phonème /s/, pour lequel ils disposent de 4 graphèmes, se vérifie à travers les nombreuses erreurs lexicales de cette nature que nous avons relevées dans notre tableau 4.15. Nous en comptons 279 sur les 828 erreurs de type phonogrammique, ce qui

représente 33.7 % de ce type et 28 % du total des erreurs lexicales du corpus. Les francophones, en égard aux erreurs relevées dans les copies, éprouvent également des difficultés dans la transcription du phonème /s/. Dans le tableau 4.15, nous comptons 63 erreurs de cette sorte sur les 570 erreurs phonogrammiques, soit 11 % (sans compter ce même problème dans les paires d'homophone Ce/se ; c'est/ses, etc. qui sont classées dans la section des erreurs grammaticales).

Avant de poursuivre notre analyse, il importe de préciser que, pour le français et pour l'espagnol, la transcription erronée du phonème /s/ n'est pas nécessairement la seule erreur lexicale que contiennent les graphies erronées réunies dans le tableau 4.15. La graphie « **imence** » (immenses) illustre bien ce point, car en plus de l'absence de la marque d'accord silencieuse (le mot est au pluriel dans le texte dicté) la double consonne « mm », essentielle à l'image orthographique du mot, et le graphème « s », important pour marquer le nombre, n'ont pas été reproduits. En espagnol, le mot ***bensio*** (*venció, vaincu en français*) est non conforme à la norme à cause de la confusion entre les graphèmes « s et c » et les graphèmes « v et b » qui ne se distinguent pas à l'oral.

Ce tableau nous montre qu'effectivement la transcription du phonème /s/ est une zone de difficulté phonogrammique plus problématique pour les hispanophones. D'après les graphies erronées du tableau 4.15, en passant sous silence le fait qu'un mot peut contenir plus d'une erreur lexicale, cette sorte d'erreur représente 33.7 % (soit 279/828) du nombre total des erreurs phonogrammiques chez ce groupe. Dans le cas des francophones, cette proportion n'est que de 11 % (soit 63/570). Contrairement à l'adjonction de lettres inaudibles qui posait nettement moins de difficulté aux hispanophones (0.60 % des erreurs de cette sorte), les erreurs phonogrammiques reliées à la

transcription du phonème /s/ représente pour eux une difficulté majeure, 3 fois plus que chez les élèves francophones, du moins dans le corpus Nadeau, Matute, Gonzales-Reyes (2006).

Nous constatons également l'apparition d'erreurs de même sorte dans des mots identiques. Il y a les mots *cansers/canseres* où dans les deux cas, l'élève a remplacé le graphème « c » par le graphème « s ». Nous retrouvons la même tendance dans des mots comme *leusémie/leusemia*, *dousses/dulses* et bien sûr le verbe « fasciner » pour lequel aucune des graphies erronées, française ou mexicaine, ne contient le digramme consonantique « **sc** » à l'intérieur du mot. Mentionnons pour la comparaison que, dans la dictée, le mot « fascinent » est réussi par 7.5 % des francophones (seulement sur le plan lexical) et 9.5 % chez les hispanophones. Une méconnaissance du code orthographique qui nous donne les graphies suivantes pour le radical du verbe : *façinent*, *fassinent*, *facinent*, *facinne*, en français ; *fasinan* et *facinan*, en espagnol.

Le dernier constat que nous pouvons faire à l'aide de ce tableau, c'est la production de plusieurs graphies différentes pour un même mot chez les francophones. Reprenons le mot « fascinent » (*fascinan* en espagnol) pour illustrer le plus grand nombre de choix disponibles dans le système graphique du français. Nous comptons 4 graphies différentes pour les francophones, alors qu'en espagnol nous n'en comptons que 2. Cela dit, poursuivons cette analyse à l'aide du tableau 4.16 où nous présentons le nombre de graphies lexicales différentes erronées avec le nombre de mots concernés pour chaque population. Il importe de préciser que le pourcentage est calculé d'après le nombre de mots retenus pour notre analyse, soit 197 en français et 203 en espagnol.

Tableau 4.16 – Nombre de graphies lexicales erronées, nombre de mots et % du corpus dans chaque langue

nombre de graphies erronées	français	% du corpus	espagnol	% du corpus
1 graphie	32 mots	16.24	72 mots	35.50
2 graphies	30 mots	15.23	32 mots	15.80
3 graphies	29 mots	14.72	21 mots	10.34
4 graphies	11 mots	5.60	10 mots	4.93
5 graphies	9 mots	4.60	3 mots	1.50
6 graphies	2 mots	1.02	3 mots	1.50
7 graphies	5 mots	2.54	*	
8 graphies	5 mots	2.54	*	
9 graphies	3 mots	1.52	*	
11 graphies	3 mots	1.52	*	
12 graphies	1 mot	0.51	*	
13 graphies	1 mot	0.51	*	
14 graphies	1 mot	0.51	*	
15 graphies	*	*	1 mot	0.50
20 graphies	1 mot	0.51	*	

Ce tableau nous montre que les hispanophones produisent également beaucoup de graphies erronées. Cependant, au nombre de graphies différentes pour un mot, nous pouvons affirmer que les hispanophones présentent une meilleure compétence sur le plan de l'orthographe d'usage. Dans les sections du tableau où nous comptons 7,8,9,11,12,13,14 et 20 graphies différentes (Voir Appendice C), nous ne retrouvons aucun mot pour les hispanophones. Le mot pour lequel les hispanophones ont produit 15 graphies différentes est « *quebequense* » qui ne représente pas une réalité tangible pour les élèves mexicains. Cet avantage peut s'expliquer par la régularité des graphies erronées chez les hispanophones. Le tableau 4.15 donne un aperçu de cette réalité. Nous pouvons prendre l'exemple du mot *buzo* manqué par 64 % des élèves hispanophones. Les 27 élèves ont remplacé le graphème « z » par le graphème « s ». Il est vrai que nous sommes en présence de deux systèmes graphiques différents, mais nous pouvons rappeler l'exemple du mot « fascinent » manqué massivement par

les deux populations : 4 graphies erronées différentes pour le radical du verbe chez les francophones, 2 pour les hispanophones. De manière évidente, les hispanophones produisent beaucoup de graphies erronées. Un résultat qui n'était pas prévisible d'après les affirmations des spécialistes comme Jaffré et Fayol (1997) et Masson (1991) qui affirment une meilleure compétence orthographique des scripteurs de langue espagnole.

Pourtant, il nous apparaît clair que la différence majeure entre les deux systèmes graphiques réside dans le nombre de phonèmes disponibles pour chaque langue. Rappelons que le français compte en moyenne 37 phonèmes pour les 26 lettres de l'alphabet, et que cela donne en moyenne 133 graphèmes, alors que l'espagnol ne compte que 23 ou 24 phonèmes pour un grand total de 32 graphèmes. Cette particularité expliquerait donc le fait que les élèves hispanophones produisent moins de graphies erronées par mot, toutefois, au moindre choix possible, une certaine proportion d'élèves produisent une erreur.

Pour poursuivre avec les sortes d'erreurs phonogrammiques, abordons le problème d'accent dans les deux langues. En français, la place des accents (aigu, grave, circonflexe, tréma) est régie par des conventions ou des règles (Thimonnier 1967). L'accent fait donc partie de l'identité du mot quant à l'orthographe lexicale. L'accentuation sert également à différencier des homophones (pêche et pêche). Ainsi, en omettant ou en ajoutant un accent (aigu, grave, circonflexe, tréma) sur un mot, il arrive souvent que cela change et le visage et la prononciation du mot (altération de la valeur phonique). Cette modification graphique, en général, ne nuit pas au décodage grapho-phonétique d'un lecteur expert, mais reste une erreur lexicale. Les cas où l'absence d'un accent altère la valeur phonique d'un mot sont classés dans la section des erreurs phonétiques. Dans les textes des

élèves francophones, nous retrouvons des cas d'ajout et des cas d'omission (absence) pour un total de 118 erreurs lexicales. De ce nombre, nous comptons 59 graphies erronées de type phonogrammique et 59 de type phonétique. Il convient de préciser que souvent l'absence ou l'ajout de l'accent n'est pas la seule erreur lexicale dans le mot. Cependant, pour les besoins de notre analyse comparative entre le français et l'espagnol, nous considérons, dans le tableau 4.17, les erreurs phonogrammiques d'accent qui auraient une entrée dans le dictionnaire de chaque langue sans la présence d'un accent (spéctateurs) ou en absence d'un accent quelconque (drole).

Tableau 4.17 - Erreurs phonogrammiques reliées aux accents dans chaque langue

français	espagnol		
bléssé	<i>acaricia</i> (2)	<i>confia</i> (17)	<i>multitud</i>
bléssée	<i>acaricia</i>	<i>creacion</i> (11)	<i>nerviosa</i>
blésser	<i>acrobatas</i> (25)	<i>despues</i> (22)	<i>numero</i> (13)
celà	<i>álgunos</i>	<i>dias</i> (14)	<i>pareceria</i> (15)
châpiteau (5)	<i>americanos</i> (2)	<i>dibúja</i>	<i>parecia</i> (17)
châpiteaux (3)	<i>americanos</i> (2)	<i>èl</i> (2)	<i>participan</i>
déssin (3)	<i>animo</i> (11)	<i>encantadores</i>	<i>participán</i>
déssine (2)	<i>animo</i>	<i>estabá</i>	<i>participarón</i> (3)
drole (3)	<i>aqui</i> (14)	<i>éstos</i> (3)	<i>pièza</i>
durér	<i>atencion</i> (7)	<i>faciles</i> (24)	<i>póco</i>
éxécutent (6)	<i>azúles</i>	<i>fué</i> (6)	<i>presentá</i>
l'obsèrve (2)	<i>canceres</i> (14)	<i>habia</i> (14)	<i>público</i>
paraissait (2)	<i>caricias</i>	<i>hirio</i> (20)	<i>publico</i> (19)
parôle	<i>casi</i> (2)	<i>inmoviles</i> (14)	<i>quien</i>
parôles	<i>círcos</i>	<i>jovenes</i> (22)	<i>quién</i>
prèsque	<i>colocarón</i> (3)	<i>kilográmos</i>	<i>records</i> (14)
secrèt	<i>colocáron</i>	<i>kilometros</i> (15)	<i>vários</i> (2)
spéctateurs	<i>coloridos</i> (3)	<i>lanzán</i>	<i>vencio</i> (13)
	<i>colóridos</i>	<i>maraton</i> (19)	<i>vencio</i>
	<i>comienzan</i>	<i>mismo</i>	
	<i>comó</i>		
Légende			
(n) = nbre d'élèves ayant produit cette graphie si plus d'un élève			

Le constat est clair. En parcourant le tableau 4.17, il apparaît évident que la proportion des erreurs d'accent est plus élevée chez les hispanophones. Illustrons ce phénomène à l'aide des résultats obtenus pendant notre analyse. Dans le tableau 4.14, le nombre d'erreurs lexicales pour les élèves hispanophones s'élève à 995. En triant les erreurs d'accent tonique qui appartiennent uniquement à la classe des erreurs phonogrammiques, nous arrivons à un total de 411 mots sur lesquels cet accent a été omis (*despues#después*), placé sur le mauvais graphème (*vencío#venció*) ou carrément ajouté inutilement (*dibúja#dibuja*), ce qui correspond à 41.30 % (411/995) du total des erreurs lexicales chez les hispanophones. Un nombre qui représente 49.64 % (411/828) du total des erreurs phonogrammiques du tableau 4.14. Retenons que pour les mots qui seraient grammaticaux en respectant l'accentuation ce pourcentage est seulement de 6.31 % (36/570) chez les francophones.

Du nombre total des erreurs phonogrammiques reliées à l'accent écrit en espagnol (411), nous comptons 353 erreurs lexicales (35.5 % du total des erreurs lexicales) où les élèves ont omis l'accent tonique sur le mot transcrit. C'est bien plus que la moitié des erreurs lexicales de cette sorte. Pour tout dire, le grand nombre d'erreurs d'accent en espagnol explique pourquoi nous retrouvons la plus grande concentration d'erreurs lexicales dans la section phonogrammique du tableau 4.14, mais comment expliquer ce nombre élevé quand il semblerait que l'accent écrit occupe une place importante dans l'orthographe de l'espagnol.

En effet, l'accentuation donne un sens aux mots en espagnol, répond à des règles fixes et joue un rôle phonologique et grammatical essentiel. Sur le plan grammatical, il s'agit de différencier des mots identiques comme des homonymes ou deux formes d'un même verbe sans pour autant intensifier

phonétiquement la syllabe à l'intérieur du mot. (Caro 1988, Bescherelle 1998). Nous pensons qu'une étude sur l'utilisation de l'accent écrit dans les compositions des hispanophones apporterait un éclaircissement sur cette constatation réalisée dans le cadre de l'étude de notre corpus.

En guise de conclusion à l'analyse des erreurs phonogrammiques, nous présentons un tableau qui fait le bilan des sortes d'erreurs qui apparaissent dans cette catégorie. Elles ne sont pas toutes traitées parce que l'étude s'est concentrée davantage sur les sortes d'erreurs qui sont communes dans les deux langues.

Tableau 4.18 – Répartition des sortes d'erreurs phonogrammiques dans chaque langue

	français	%	espagnol	%
sortes d'erreurs				
ajout de lettres muettes	107	18.8	5	0.6
transcription du phonème /s/	63	11.1	279	33.7
accent ajouté ou omis	59	10.3	411	49.6
autres	341	59.8	133	16.1
total	570	100	828	100

Ce tableau montre que le français renferme plus de « zones de fragilités » que l'espagnol. Les trois sortes d'erreurs lexicales analysées ici ne touchent pas la moitié des erreurs phonogrammiques chez les francophones alors que chez les hispanophones, elles comptent pour 83.9 % des erreurs phonogrammiques. Les erreurs liées à l'accent écrit représentent la moitié des erreurs de ce type.

Le deuxième type d'erreur se situe sur le plan de la phonétique. Le nombre est assez élevé dans les textes des deux groupes mis en présence : 28.30 % pour les francophones et 12.20 % pour les hispanophones. Bien entendu, devant les discours qui nous décrivent l'espagnol comme étant plus régulière, plus douce, plus sonore que toutes les langues européennes, ces erreurs nous paraissent « contre-nature » dans le corpus hispanophone même si

dans les deux cas (français et espagnol) nous supposons qu'elles sont reliées soit à une mauvaise discrimination auditive de la part de l'élève, soit à un problème de prononciation chez l'élève ou celui qui dicte, soit à une surcharge de l'esprit (écouter, écrire, accorder) qui peut nuire à la compréhension de l'élève. À ce stade de notre recherche, nous ne disposons pas d'indice autre que l'explication de Catach (1980) dans la description de sa typologie des erreurs pour expliquer la présence de ces erreurs lexicales dans les textes des élèves. Ainsi, pour produire les graphies erronées, « l'élève peut entendre mal ou mal prononcer les sons. »

Dans le tableau 4.13, nous comptons 290 erreurs phonétiques chez les élèves francophones, ce qui représente 28.30 % du total des erreurs lexicales du corpus. Chez les hispanophones, le nombre est de 121 erreurs phonétiques pour 12.20 % du total des erreurs lexicales du corpus.

Ces erreurs montrent en général une graphie très éloignée du mot dicté (exemple : **immece**#immenses, **ésécutent**#exécutent en français et **mejós**#mejores, **tatentosas**#talentosas en espagnol). Cependant, nous avons classé certains mots français dans la section des erreurs phonétiques à cause des règles de prononciation de la langue. Prenons par exemple les graphies erronées : ath**l**etes, comb**in**eson, comp**et**iteurs, entr**ene**use, repr**es**ente, etc. l'absence de l'accent sur le graphème « e » change la prononciation du mot. Dans d'autres cas, l'omission de ce graphème « e » en fin de mot (acrobats, b**ot**s, ens**ui**t) modifie également la prononciation du mot dicté. Il y a également les cas où les élèves doublent le graphème « l » devant le graphème « i » en milieu de mot (équ**ill**ibre, kill**om**ètres). La prononciation du mot change, donc nous les classons parmi les erreurs phonétiques. Enfin, en plus des erreurs liées à la discrimination du mot à transcrire, nous relevons les mêmes sortes d'erreurs qui témoignent d'une

méconnaissance du code orthographique : l'adjonction de graphèmes au mot (deumeurent, meuilleures, pessait) ; la confusion de graphèmes dans la transcription (amériquien, leussimies, reconfortente, scaffandier).

Les autres types d'erreurs lexicales comme les morphogrammes lexicaux, les homophones de discours et les idéogrammes n'ont pas été présentés dans cette section parce qu'ils relèvent des particularités propres à chaque système. Nous prenons l'exemple du tilde dans le mot *pequeña*. Une erreur lexicale sur ce signe graphique ne se retrouve qu'en espagnol au même titre qu'une erreur lexicale sur l'apostrophe reste un fait graphique du français. Pour les autres types d'erreurs phonogrammiques, ils n'ont pas été mentionnés pour les mêmes raisons. Il suffit de considérer la confusion entre les graphèmes « b » et « v » qui n'existe qu'en espagnol. Une erreur lexicale de ce type ne retrouve aucun écho en français. En somme, il importe de préciser que notre étude a surtout considéré les erreurs lexicales de même type que nous retrouvons dans la typologie des erreurs pour chaque groupe comparé. Ainsi, nous avons pu constater que les graphies erronées sont concentrées dans les sections phonogrammique et phonétique. En ce qui a trait aux homophones lexicaux et de discours, aucun cas n'a été trouvé en espagnol alors que nous en relevons quelques-uns en français.

Cela dit, cette démonstration nous montre que nous retrouvons les mêmes types d'erreurs dans les deux langues. Nous avons montré la présence des erreurs d'ajout de lettres inaudibles pour un pourcentage de 18.77 % des erreurs phonogrammiques (ou 10.45 % du total des erreurs lexicales) chez les francophones et de 0.60 % des erreurs phonogrammiques (ou 0.5 % du total des erreurs lexicales) chez les hispanophones. Un pourcentage qui semble expliquer les multiples graphies produites par les élèves francophones. Nous avons montré le problème que représente la

transcription « graphémique » du phonème /s/ dans les deux langues. Les hispanophones de notre corpus démontrent une plus grande difficulté que les francophones. En effet, en passant sous silence le fait qu'un mot peut contenir plus d'une erreur, les graphies non conformes à la norme découlant de la transcription du phonème /s/ correspondent à 33.7 % des erreurs phonogrammiques (ou 28.04 % du total des erreurs lexicales) chez les hispanophones en comparaison de 6.15 % chez les francophones. Nous avons montré que les erreurs phonétiques sont communes aux deux populations. Des erreurs qui semblent provenir avant tout de la discrimination auditive des élèves.

Nous avons présenté à la fin les erreurs d'accent (phonogrammiques). Une présentation qui nous a permis de constater que, pour notre corpus, c'est la deuxième section de l'orthographe lexicale qui cause le plus de problèmes aux hispanophones. Malgré des règles fixes, les élèves omettent d'ajouter l'accent écrit là où il est nécessaire (353 erreurs de cette sorte) ou le placent sur le mauvais graphème dans le mot. Ces erreurs lexicales reliées à l'accentuation représentent 49.60 % (411/828) des erreurs phonogrammiques relevées dans le corpus espagnol, soit 41.3 % du total des erreurs lexicales. Chez les francophones, ce pourcentage est de 3.51 %.

Finalement, en considérant le tableau 4.19 où nous constatons que, chez les hispanophones, seulement deux sortes d'erreurs de type phonogrammique représentent 83.7 % des erreurs de ce type, nous pouvons dire qu'en espagnol, les zones de variations possibles sont peu nombreuses même si ces « zones » entraînent un grand nombre d'erreurs. Ainsi, les erreurs sur la transcription du phonème /s/ (279), sur l'ajout ou l'omission de l'accent (411) constituent à elles seules plus de 2/3 des erreurs lexicales. Mentionnons que les erreurs d'accent modifient peu le mot visuellement tandis qu'en français,

comme on le voit dans les nombreuses graphies produites pour un même mot, les « zones » étant nombreuses, les mots prennent souvent une allure méconnaissable. Par conséquent, l'opacité se voit davantage même si les hispanophones de notre corpus font, à 12 ans, autant d'erreurs lexicales. Maintenant, nous allons regarder quelle réforme éliminerait le plus de graphies erronées dans les copies des élèves francophones.

4.5 Les projets de réforme en français et l'orthographe lexicale

Maintenant, nous passons au deuxième objectif de notre étude. Nous allons procéder à la projection des projets de réforme présentés dans le cadre théorique pour répondre à notre dernière question de recherche : ***Quelle réforme parmi celles de Dauzat (1939-1952) ; de la Commission Beslais (1960-1965) ; de Thimonnier (1967) aurait permis de diminuer le plus les graphies erronées dans les textes des élèves francophones ? Et dans quel pourcentage ?***

Avant de commencer, il convient de préciser que cette projection n'est valable que pour les mots du corpus, lesquels n'ont pas été sélectionnés pour tester des effets d'une réforme. Ainsi, un autre corpus donnerait sans doute des résultats différents de ceux que nous obtiendrons dans ces pages, mais nous pensons que cette projection montrera qu'une réforme majeure est nécessaire pour que les graphies erronées disparaissent dans une plus grande proportion dans les textes des élèves francophones.

Pour mener à bien cette projection, nous devons voir parmi les projets de réforme considérés, lequel toucherait le plus large inventaire de mots dans notre corpus. La réforme la plus récente vient des propositions du CSLF. Il s'agit plutôt de rectifications orthographiques parce que les deux graphies sont permises. Rappelons pour les besoins de notre étude que les

recommandations ne touchent que 2430 mots dont seulement 55 (selon la codification du *Vadémécum de l'orthographe recommandée*, 2004) sont considérés comme appartenant à la catégorie des mots de haute fréquence, dont voici quelques exemples avec leur échelle de fréquence tirée de Baudot (1992) : *aout* (85), *boîte* (99), *dîner* (43), *révoluer* (26), *voute* (14). Des mots qui selon notre découpage tomberaient dans la section des mots AF. Le corpus étudié vient d'un texte ordinaire, il est clair qu'il n'y aura pas un grand nombre de mots touchés surtout que, selon Goosse (1991) : « Les innovations portent sur moins de huit-cents mots, dont certains fort rares. Dans un livre, un mot en moyenne pour deux pages est touché. » (p.2)

En effet, après examen des mots du corpus français (les 315 mots retenus pour notre analyse finale), nous constatons que les rectifications du CSLF n'ont pas une grande incidence sur les graphies erronées des élèves. Les mots rares sont les mots appris tardivement par les enfants, selon Bonin (2005) et selon les résultats dans la section qui traite du rôle de la fréquence dans l'acquisition de l'orthographe. Nous avons donc comparé les 2375 mots moins fréquents rectifiés avec ceux de notre corpus pour remarquer qu'il n'y a pas vraiment de correspondance entre les deux. Dans les mots de notre dictée, nous ne trouvons que le mot « entraîneuse » qui est touché par ces rectifications. La tendance, dans les copies des élèves, est au maintien de l'accent circonflexe (c'est le cas de 35 % des élèves francophones) et les erreurs lexicales dans le mot « entraîneuse » ne sont pas reliées à l'absence ou à la présence de cet accent dans le mot, mais plutôt à d'autres difficultés (ex : entraînneuse). Les rectifications datent de 1990 alors que l'exercice a été réalisé en 2004, nous nous demandons si ce point a été souligné dans l'enseignement reçu par ces élèves. En conclusion, nous pouvons avancer que les rectifications de 1990 ne sont pas assez étendues (touchent à trop

peu de mots) pour avoir un impact majeur sur la compétence orthographique des francophones de notre corpus.

Par rapport au nombre de mots (228) touchés par le projet d'émondage de Thimmonier (1967), aucun mot du corpus français n'aurait été touché. Par conséquent, nous l'éliminons de notre liste. Il ne reste que les projets Dauzat (1939) et Beslais (1965).

En ce qui a trait au projet de Dauzat, nous ne comptons que quatre mots qui seraient touchés. Le verbe « appeler » qui se conjuguerait comme le verbe *modeler*. Une modification qui ferait disparaître une graphie erronée (*appèle*) sur les six qui sont produites par les élèves francophones (soit 0.1 % du total des erreurs lexicales) dans l'ensemble du corpus. Il y a les mots « chapiteaux et animaux » parce que Dauzat demande d'unifier le pluriel des noms par l'ajout du morphogramme « s ». Dans ce contexte, il est évident que l'application d'une telle proposition toucherait davantage l'orthographe grammaticale (0.2 % du total des erreurs grammaticales). Le dernier mot est le verbe « fascinent ». Le graphème « s » étant considéré par Dauzat comme une lettre parasite dans la graphie de ce mot, l'application de son projet de réforme le ferait donc disparaître. Par conséquent, sensiblement 2 % (19/24 des graphies erronées pour ce mot) du total des erreurs lexicales serait éliminé de notre corpus. Voilà en résumé l'impact qu'aurait le projet Dauzat (1939) sur notre corpus : un total de 2.1 % d'erreurs lexicales en moins et 0.2 % d'erreur grammaticale.

Examinons maintenant le projet de Beslais (1965) qui « parmi les projets de réforme récents, et malgré les déboires qu'il a essuyés, c'est sans doute celui [...] qui s'est trouvé, non pas le plus près de réussir, mais en meilleure posture vis-à-vis de l'opinion » (Catach, 1989, p.293). Cela dit, nous allons

voir combien d'erreurs lexicales du corpus français seraient touchées par la réforme Beslais (1965).

Pour cette projection, nous prendrons la réforme de Beslais (1965) proposition par proposition pour que notre analyse soit complète. Il convient de préciser que dans les tableaux que nous présenterons, il se pourrait que le mot contienne plus d'une erreur. Nous les comptons uniquement parce que nous supposons que le fait d'apporter certaines simplifications au système orthographique du français rendra le mot plus accessible à l'élève. Il n'aura pas à effectuer différentes combinaisons, à essayer de deviner d'après sa logique, qui tend vers une cohérence orthographique, pour rappeler à la mémoire la bonne graphie du mot qu'il doit écrire.

4.5.1 Les mots composés

Beslais (1965) propose la soudure des mots composés avec pluriel à la fin (gardechasse, antédiluvien, porteplumes, etc.), le maintien du trait d'union pour certaines catégories de mots (demi-mesure, sans-façon, chez-soi, etc.) et le statu quo pour d'autres. Dans notre corpus, il n'y a que le mot « lui-même », classé PF, qui rentre dans cette section. Ce mot est raté par 82.5 % des élèves francophones. Nous comptons 33 erreurs lexicales, dont 32 concernent l'absence du trait d'union. Nos résultats concordent avec ceux de Manesse et Cogis (2007) parce que dans leur étude : « au XIXe siècle, plus de 31% des élèves n'avaient pas noté le trait d'union dans elle-même, alors qu'ils avaient écrit correctement ses deux constituants. Ils étaient 47% au XXe siècle » (p. 171).

Tout comme les rectifications du CSLF, aucune des recommandations de Beslais (1965) ne s'intéresse au mot « lui-même » de notre corpus. L'explication de ce fait se trouve sûrement dans la règle de composition du

mot : une section grammaticale qui traite de l'accord et une section lexicale qui traite de l'image du mot (le trait d'union). En somme, à moins d'une souplesse sur la présence du trait d'union entre le pronom et l'adjectif indéfini, la réforme de Beslais (1965) n'aurait pas réussi à éliminer 3.23 % du total des erreurs lexicales de cette sorte chez les élèves francophones.

4.5.2 Les accents

Beslais (1965) recommande de placer l'accent aigu sur les mots d'origine latine ou étrangère. Notre corpus n'en compte aucun, donc cette section de la réforme n'a aucune incidence sur les erreurs lexicales des élèves francophones.

Aucune recommandation n'est faite sur l'emploi de l'accent grave. Cependant, il est ajouté dans les mots où les finales *elle* et *ette* ont été simplifiées pour **èle** et **ète**. Nous en reparlerons davantage quand nous traiterons des consonnes doublées. Dans notre corpus, le mot *cela*, serait le seul mot touché par une telle rectification. Il s'écrirait « celà » pour s'aligner sur les mots courants « à, là, ça, où, etc. » Nous tenons à préciser que ce mot est raté par seulement 5 % (2 graphies erronées) des élèves francophones, dont 2.5 % ayant produit la graphie celà (0.1 % du total des erreurs lexicales).

Beslais (1965) nous dit que l'accent circonflexe cause beaucoup de problèmes aux scripteurs francophones. Ce type d'erreur compte pour 25 % des erreurs du corpus qu'il a analysé avant de présenter son projet de réforme (p.124). Les élèves l'ajoutent ou l'omettent sur les lettres d'un mot. Dans le projet Beslais (1965) l'accent circonflexe est maintenu sur les voyelles longues (bête, faite, âne, etc.), mais supprimé sur les mots où sa

présence marque la chute d'une lettre ancienne comme dans « assidûment » (anciennement assiduellement).

Tableau 4.19 – Mots français dont les graphies erronées contiennent l'accent circonflexe

battre	chapiteau(x)	drôle	inquiète	kilomètre	parole	pièce	paraissait
bâtre	châpitaux (2)	drole (3)	inquiète (2)	kilomètre	<i>parôle</i>	pièce	<i>paraïssait</i>
batre (2)	<i>châpiteau</i> (5)		inquiette (4)	kilometres	<i>parôles</i>	pièce	
	chapitaux (3)		inquiète (2)	killomètres (4)	parolles		
	chapito		inquite		parolle		
	chapitot						
	châpitaux (3)						
	chapitaux (11)						
	<i>châpiteaux</i> (3)						
	chapitos						
	chapitiaux						
(n) : nbre d'élèves ayant produit la graphie si plus d'un élève							

Dans le tableau 4.19, il n'y a qu'un mot dont l'orthographe lexicale requiert l'accent circonflexe (drôle). L'accent marque certainement la chute du graphème « s » que contenait jadis le mot (*dro~~s~~le*), ce qui veut dire que la suppression de l'accent éliminerait 0.3 % du total des erreurs lexicales. En ce qui a trait aux autres mots, ils seraient, en cas de réforme de l'accent circonflexe, toujours des graphies erronées parce qu'ils présentent plus d'une erreur lexicale. Parmi les sortes d'erreurs du tableau 4.19, nous remarquons 11 graphies erronées qui seraient conformes à la norme sans l'ajout de l'accent circonflexe (soit 1.1 % du total des erreurs lexicales). Dans les autres cas, (inquiète, kilomètre, pièce) il s'agit d'une confusion entre deux accents. Comme nous l'avons vu avec les rectifications du CSLF, le verbe « paraître » et le mot « entraîneur » seraient touchés par l'application de cette réforme, ce qui signifie la disparition de 0.1 % du total des erreurs lexicales.

Par conséquent, pour notre corpus, les rectifications de Beslais (1965) portant sur l'accent circonflexe auraient eu peu d'effet sur les compétences

orthographiques de nos élèves, car celui-ci est maintenu sur le mot « drôle ». Pour que cette rectification ait une grande incidence sur notre corpus, il aurait fallu proposer le maintien de l'accent circonflexe uniquement dans les cas d'homonymie, car en le maintenant sur un grand nombre de mots, les élèves continueront à l'ajouter là où il est inutile et à le retrancher là où il est nécessaire. Le tableau 4.19 illustre bien ce fait à l'aide des mots « ch^âpîteau, ch^âpîteaux, par^ôle, par^ôles, para^îssait » où la principale erreur dans le mot est l'ajout de l'accent circonflexe.

4.5.3 Le graphème (consonne) x

Beslais (1965) recommande de conserver ce graphème dans les mots où il se prononce (flexion, examen, exemple, exécuter, etc.) Comme rectification, il demande de le changer pour le graphème « s » dans les mots où il n'est pas prononcé (jalous, hibous, animaux, etc.) Une rectification qui toucherait également les déterminants numéraux et ordinaux (deus, dis, deusième, etc.) Un tel changement permettrait d'harmoniser le pluriel des noms et adjectifs. En examinant les erreurs lexicales dans les mots de notre corpus qui contiennent le graphème « x » à la finale, il apparaît évident que l'application de cette proposition n'aurait aucun effet sur la compétence orthographique des élèves francophones de notre corpus.

Tableau 4.20 - Erreurs lexicales dans les mots français contenant le graphème x à la finale

mieux	animaux	chapiteaux
mieu (11)	animeaux (3)	chapitiaux (11)
mieut		châpitiaux (3)
		châpîteaux (3)
		chapitiaux
		chapitos
(n) : nbre d'élèves ayant produit la graphie si plus d'un élève		

Dans le tableau 4.20, le fait de changer le graphème « x » en « s » ne changerait rien dans les graphies erronées du mot « mieux ». En ce qui a trait

aux deux autres mots, il s'agit de la marque silencieuse du pluriel. Les erreurs lexicales dans les mots « animaux » et « chapiteaux » représentent 2.15 % du total des erreurs lexicales de notre corpus et la marque du pluriel est respectée dans la plupart des mots pour 2 % du total des erreurs lexicales. Le problème pour les élèves francophones se trouve dans les nombreuses graphies disponibles pour transcrire le phonème /o/. Rappelons que Chartrand et al. (1999) présentent 5 graphèmes alors que Giasson et Thériault (1983) en dénombrent 30 pour transcrire le phonème /o/. Pour tout dire, cette section de la réforme touche avant tout à la règle du pluriel des noms. Par conséquent, elle aurait éliminé 0.2 % du total des erreurs grammaticales du corpus tout comme pour le projet de Dauzat (1939).

4.5.4 Les lettres grecques

Les lettres grecques reviennent constamment dans les projets de réforme. Les partisans de la simplification de l'orthographe du français font souvent la comparaison avec d'autres pays qui les ont éliminées de leur orthographe (espagnol, italien, etc.). À l'intérieur de notre corpus, nous avons trois mots qui ont une racine grecque dans leur orthographe. Il s'agit de marath**on**, athl**è**tes et scaph**h**andrier. Ces mots sont mal orthographiés massivement par les élèves.

Tableau 4.21 - Erreurs lexicales dans les mots français contenant les digrammes grecs (th et ph)

athlètes		marathon	scaphandrier
<i>acclèthes</i>	<i>athètes (3)</i>	maraton (43)	scafandrier (9)
<i>aclète</i>	athletes (2)	maratons	scafandrié (4)
<i>aclettes</i>	atlètes (5)	marraton (2)	scafendrié (5)
<i>aclèttes</i>	atlèthe (2)		scafendrier (6)
<i>acthlètes</i>	atlèthes (2)		scafondrier
atelettes	atlettes		scaphandillé
<i>athelectes</i>	athètes (5)		scaphendrier
<i>athèles</i>	hathlète		
athelètes (6)	hatlète		
<i>athelte</i>			
<i>atheltes (2)</i>			
(n) : nbre d'élèves ayant produit la graphie si plus d'un élève			

Comme nous pouvons le voir dans le tableau 4.21, les francophones ont produit 20 graphies différentes pour le mot athlètes, dont une grande proportion correspond à des erreurs phonétiques (mots en italiques dans le tableau). Apparaissant deux fois dans la dictée, la moyenne nous donne un pourcentage d'erreurs de 47.5 % pour ce mot chez les élèves francophones. Nous tenons à préciser, encore une fois, que les erreurs ne portent pas exclusivement sur l'utilisation du digramme **th**. En appliquant la réforme de Beslais (1965) qui ferait disparaître le **h** muet du mot « athlète », cela nous donnerait 15 graphies « erronées correctes » : atlètes, atlèthe, atlèthes, athlètes, **hatlète** pour un pourcentage de 1.5 % du total des erreurs lexicales du corpus francophone.

Dans le cas du mot marathon (raté par 57.5 % des francophones), l'application de cette réforme aurait permis d'éliminer 43 graphies erronées ou 4.22 % du total des erreurs lexicales. Le tableau 4.21 montre bien que le fait de remplacer, dans le mot scaphandrier, le digramme **ph** par le graphème **f** n'aurait rien changé sur le plan de la compétence orthographique des

francophones de notre corpus. Un fait qui se constate à l'aide des graphies erronées du tableau 4.21 où 62.5 % des élèves francophones ont effectué ce changement dans le mot scaphandrier, mais comme les graphies erronées produites présentent d'autres erreurs lexicales, cette rectification n'aurait éliminé aucune erreur de notre corpus.

En somme, en additionnant les deux pourcentages de graphies erronées éliminées (athlètes et marathon), nous pensons qu'une simplification des lettres grecques en français (**rh**, **ph**, **th**) aurait permis d'enlever 5.7 % du total des erreurs lexicales des élèves de notre corpus. Un pourcentage qui apparaît plus intéressant que le cas de l'accent circonflexe pour les mots de notre corpus.

4.5.5 Les consonnes doublées

Dans l'étude de Beslais (1965), les consonnes doublées venaient en deuxième position dans le pourcentage d'erreurs lexicales des élèves (p.124). Plus près de nous, une étude de Jaffré & Fayol (1997) « sur les résultats d'adultes et d'élèves de la deuxième année de primaire à la classe de troisième au collège » a démontré qu'en ce qui concerne l'orthographe lexicale, « au-delà de dix ans, les erreurs se concentraient surtout sur les consonnes doubles : les élèves comme les adultes en ajoutaient ou en enlevaient de façon erronée » (p.99). Une manière de procéder qui s'explique par le fait que dans bon nombre de cas les consonnes doubles « ne marquent pas de changement phonologique par rapport au cas où une seule consonne apparaît (par exemple, en français « ballon » ou « balon ») » (Ibid., p.105)

Après l'étude des différentes erreurs lexicales, nous remarquons que rien n'a vraiment changé dans notre corpus. Les élèves francophones ajoutent des

consonnes doubles là où ce n'est pas nécessaire ou les enlèvent là où c'est nécessaire. La connaissance du système orthographique des élèves n'étant pas suffisante pour analyser les erreurs commises, il arrive parfois que l'action de redoubler ou de simplifier une consonne apporte des changements qui altèrent la valeur phonique du mot (*paraissait*≠*paraissait* ; *enfilé*≠*enfilé*). Ces sortes d'erreurs sont donc classées dans la section des erreurs phonétiques. Parmi les 315 mots de notre corpus, nous comptons 29 mots à consonne doublée. De ce nombre, nous avons 9 mots dans lesquels, la simplification de la double consonne ne peut se faire à cause de la place que la graphie occupe dans la mémoire collective, ce qui veut dire que la première graphie est « incrustée dans des millions de mémoires individuelles et concrétisée dans des millions d'ouvrages » (Thimonnier, 1967, p.369) (*elle*, *caresses*) ; et à cause de la possibilité d'altérer la valeur phonique du mot (*dessin*, *réussi*, *accident*, *poussières*, *paraissait*, *blessée*). Il reste donc 20 mots qui seraient touchés par la réforme de Beslais (1965).

Tableau 4.22 - Erreurs phonogrammiques liées à la double consonne en français

<i>acrobates</i>	accrobates (2)	acrobattes (2)			
allongés	alongé (3)	alonger	alongés (5)		
appelle	appèle	apelle (2)			
applaudit	aplaudit (4)	aplaudient			
arrivent	arive (3)				
attentivement	atantivement	atentivement			
battre	batre (2)				
botte	bote	botes			
<i>canadien</i>	cannadien				
caresses	carresse (4)	carresses (3)			
chatte	chate (6)				
collectives	colective (7)				
<i>colorés</i>	collorées				
<i>courage</i>	courrage (2)				
<i>coureurs</i>	courreur (3)	courreurs (4)			
<i>courir</i>	courrir (7)				
<i>duré</i>	durré				
<i>entraîneuse</i>	entraînneuse				
<i>extraordinaires</i>	extraordinnaires (2)				
flatte	flate (13)	flates			
<i>foule</i>	foulle (4)				
<i>humains</i>	hummain				
illuminent	ilumine (2)	iluminent			
<i>imaginée</i>	immaginée	immaginer (2)	imaginner		
immenses	imenses	imense			
immobiles	imobile (6)	imobiles (4)			
kilogrammes	kilogrames (14)	kilogramme (7)			
lutte	lute (3)				
<i>paraissait</i>	parraissait (3)				
<i>paroles</i>	parolles	parolle			
personnages	personages (3)	perssonages			
poussières	poussièrres				
<i>presque</i>	pressque				
<i>records</i>	reccords				
<i>s'élancent</i>	s'éllance				
<i>talentueuses</i>	tallentueuses	talenttueuses			
(n) : nbre d'élèves ayant produit la graphie si plus d'un élève					

Regardons les propositions de Beslais sur les consonnes doublées. Comme règle générale, après une analyse du système graphique du français, Beslais (1965) recommande la réduction à l'unité pour un grand nombre des mots

comme **honne**ur, **barique**, **courier**, etc. Ce projet recommande également de simplifier le pronom **elle** (**èle**), les verbes en « **eler** » ou en « **eter** ». Cependant, la double consonne doit être maintenue pour le graphème « **ss** » qui produit le phonème /s/, pour le cas des préfixes « **in**, **im**, **il**, **ir**, **en**, **em**, **rem** », pour la conjugaison de quelques verbes irréguliers (**courrai**, **verrai**, **pourrais**, etc.), pour les cas où le graphème « **ll** » permet de produire le phonème /ʎ/. Nous présentons dans le tableau 4.22 les graphies erronées des élèves francophones qui proviennent de l'ajout ou de la simplification d'une consonne doublée.

Dans le tableau 4.22, les mots en caractère gras sont ceux qui possèdent une consonne double dans leur graphie nommée. Les autres mots en italiques représentent des mots auxquels certains élèves ont ajouté la double consonne. Ainsi, nous constatons qu'effectivement, les francophones en ajoutent ou en enlèvent de façon erronée. En considérant tous les mots du tableau 4.22, nous arrivons à un total de 149 graphies erronées, ce qui veut dire 14.6 % du total des erreurs lexicales du corpus étudié. Cependant, il convient de préciser que bien des mots de notre tableau resteraient des graphies erronées en appliquant la réforme de Beslais (1965) à cause du maintien d'un grand nombre de mots à double consonne dans le système graphique du français. Il s'agit des mots « **courir** et **coureurs** » parce que le verbe **courir** conserve le graphème « **rr** » dans sa conjugaison, et des mots **illuminent**, **immobiles** et **immenses** parce que la double consonne est maintenue dans leur graphie.

Par conséquent, en enlevant ces 5 mots de notre projection parce que Beslais (1965) ne trouve pas impératif la simplification de la double consonne, cette réforme permettrait sans doute de faire disparaître 77 graphies erronées du tableau 4.22, soit 7.5 % du total des erreurs lexicales

de notre corpus. Ce résultat est obtenu en comptant uniquement les mots dans lesquels nous trouvons une simplification de la consonne doublée. Pour le reste de l'analyse, nous pensons que par la simplification d'un grand nombre de consonnes doubles, les scripteurs n'auront plus à douter, à hésiter pendant la transcription d'un mot pour finalement choisir la graphie la plus compliquée : la double consonne. Ainsi, nous sommes d'avis que des graphies erronées comme « **tall**entueuses, **parolles**, **cann**adien, **entraînne**use, **recc**ords, **acc**robates, etc. » auraient disparu de notre corpus pour un total de 11.45 % du total des erreurs lexicales de notre corpus en comptant les autres mots du tableau 4.22.

Les derniers points touchés par le projet de réforme de Beslais (1965) sont les finales **ant**, **ance**, **ent**, **ence**, les lettres parasites : « e » muet à l'intérieur du mot (**assoir**≠**asseoir**), « p » à l'intérieur du mot (**sculteur**≠**sclupteur**), « s » à la finale du mot (**relai**≠**relais**). Nous n'en parlons pas parce qu'aucun mot de notre corpus n'est touché par ces propositions.

À travers cette projection, nous venons de montrer l'impact que la réforme de Beslais (1965) peut avoir sur la maîtrise de l'orthographe lexicale des francophones de notre corpus. Le fait de procéder à des changements sur l'image que les mots ont obtenue par convention aidera les élèves francophones à être plus compétents dans leur langue. Le pourcentage total d'erreurs lexicales qui disparaîtrait de notre corpus s'élève à 12.12 % (digramme grec plus consonne doublée). Ce nombre se veut exclusif, mais si nous tenons compte de notre argument sur la simplification d'un grand nombre de consonnes doublées dans le système graphique du français, ce pourcentage pourrait être de 17.2 %. Par conséquent, en tenant compte de la complexité du système graphique du français, l'harmonisation des familles de mot, la francisation des lettres non fonctionnelles comme les digrammes de

racine grecque, la simplification des consonnes doubles ne peuvent avoir qu'un effet positif dans l'apprentissage de l'orthographe. Nous croyons que l'élève, en situation d'écriture, aura moins à effectuer différentes combinaisons de graphème pour réussir à produire la bonne graphie.

Conclusion du chapitre 4

Dans ce chapitre nous avons répondu à nos questions de recherche à l'aide de nos analyses du corpus Nadeau, Matute, Gonzales-Reyes (2006). Nous avons pu voir que les hispanophones présentent un meilleur taux de réussite que les francophones sur le plan de l'orthographe en général. Des résultats qui démontrent notre première hypothèse de recherche : le taux de réussite en orthographe (moins de graphies erronées dans les dictées) est plus élevé chez les hispanophones (90.72 % vs 83.96 %).

Nos résultats nous ont montré également une meilleure réussite des hispanophones en orthographe grammaticale. Les hispanophones ne respectent pas les règles d'accord pour un pourcentage de 1.86 % seulement comparativement à 9.04 % pour les francophones. Sur le plan de la moyenne d'erreurs en orthographe lexicale, la différence entre les hispanophones et les francophones, nous apparaît minime (7.66 % vs 8.32 %). Par conséquent, notre hypothèse de recherche doit être rejetée parce que ces chiffres ne montrent pas la supériorité des hispanophones vis-à-vis des francophones dans ce domaine.

En ce qui a trait au rôle de la fréquence dans la maîtrise de l'orthographe lexicale, nos différents tableaux ont confirmé la théorie de Bonin (2005) qui affirme que les mots fréquents sont ceux que les élèves maîtrisent le mieux. Notre deuxième hypothèse qui prédisait moins d'erreurs dans les mots Fréquents et beaucoup plus d'erreurs dans les mots PF a également été

démontrée, bien que de nombreux mots PF soient tout de même bien maîtrisés. Par ailleurs, notre analyse nous a permis de voir que, pour notre corpus, la fréquence ne joue pas un rôle très important dans la maîtrise orthographique des mots, surtout en espagnol où la différence est faible entre mots F et mots PF réussis. De plus, nous avons constaté que la longueur (mots courts) et la transparence des mots semblent avantager les francophones dans la maîtrise de l'orthographe lexicale des mots PF.

La suite de l'analyse a montré des zones de difficulté communes aux deux langues, donc la présence des mêmes sortes d'erreurs dans les textes des élèves de chaque population. Le pourcentage d'erreur est différent d'une population à une autre à cause de la différence en importance de la zone de difficulté dans chaque langue. Les erreurs lexicales de lettres inaudibles ajoutées à un mot sont plus nombreuses chez les francophones (10.54 % du total des erreurs lexicales comparativement à 0.6 % chez les hispanophones). Les erreurs lexicales concernant la transcription du phonème /s/ sont plus nombreuses chez les hispanophones (28 % du total des erreurs lexicales comparativement à 6.15 % chez les francophones). Dans le cas des erreurs phonétiques, elles sont plus nombreuses chez les francophones (28.32 % du total des erreurs lexicales). Chez les hispanophones, les erreurs phonogrammiques sont plus nombreuses à cause des erreurs sur l'accent écrit. Par exemple, pour notre corpus, les erreurs d'accent comptent pour 41.30 % du total des erreurs lexicales.

À l'aide d'une comparaison sur les graphies erronées produites pour un mot, nous avons pu constater que l'espagnol est une langue dont le système orthographique semble être plus régulier que celui du français. Cependant, notre analyse ne permet pas de déclarer qu'un groupe maîtrise mieux l'orthographe lexicale de sa langue contrairement aux affirmations de Masson

(1991) et de Jaffré et Fayol (1997). Nous pouvons dire, que contre toute attente, les hispanophones produisent un grand nombre de graphies erronées qui à certains égards est comparable à la performance des francophones, mais le fait que la grande majorité des erreurs lexicales en espagnol soient des erreurs d'accent tonique ou d'accent qui distingue des homophones nous permet de conclure que l'orthographe est mieux maîtrisée par les hispanophones pour ce qui concerne les graphèmes.

Notre dernier point a touché les projets de réforme et leur impact sur la maîtrise de l'orthographe lexicale en français. Nous avons porté notre choix sur le projet Beslais (1965) qui nous semble plus audacieux et mieux adapté à notre corpus. À travers un découpage de chacune des propositions faites, nous avons examiné les erreurs lexicales qui disparaîtraient du corpus francophone. Cette projection a montré que l'application d'un tel projet de réforme aurait eu de l'impact sur les lettres grecques (5.7 % du total des erreurs lexicales qui concernent uniquement la présence de la lettre grecque) et les consonnes doubles de notre corpus (7.5 % du total des erreurs lexicales). Pour reprendre nos chiffres, nous pouvons dire que la réforme Beslais (1965) aurait permis de faire disparaître au total 12.12 % des erreurs lexicales dans les textes des francophones.

En guise de conclusion à toute cette analyse, nous pensons qu'une réforme dans les deux langues serait bénéfique pour tous les élèves. Pour l'espagnol, il est clair, à l'aide de nos résultats, que le fait d'éliminer l'accent tonique est une mesure simple qui supprimerait 41.30 % du total des erreurs lexicales dans le corpus espagnol. Dans le cas du français, nous proposons comme réforme, la suppression des marques muettes d'accord pour éliminer le nombre élevé d'erreurs grammaticales que nous avons relevées dans notre corpus. Prenons le morphogramme du pluriel « s », en comptant les mots

requérant cette marque d'accord, omise par les élèves, nous arrivons à un total de 291. Ce chiffre représente à peu près 26 % du total des erreurs grammaticales du corpus francophone. D'ailleurs, l'étude de Manesse et Cogis (2007) montre bien l'orthographe grammaticale est en baisse. En ce qui a trait à l'orthographe lexicale, nos résultats montrent que la transparence du système graphique nécessite une réforme radicale pour éliminer un grand nombre d'erreurs lexicales dans les textes des scripteurs francophones. Les consonnes doubles représentent un exemple tangible parmi tant d'autres. Par conséquent, il faut passer outre l'avis de Thimonnier (1967) qui nous dit qu'il est trop tard pour opérer une simplification générale et cohérente du système ; et proposer plus que les réformes de Beslais (1965) et du CSLF (1990).

Conclusion générale

Nous avons entrepris cette recherche sur la base de la complexité de l'orthographe du français et d'un manque de données empiriques chez certains spécialistes ; par exemple : Masson (1991), Jaffré et Fayol (1997). Ces chercheurs affirment une meilleure compétence orthographique des scripteurs italophones et hispanophones par rapport aux francophones, mais sans aucune référence à des études comparatives de leurs performances respectives. Rappelons que ce sont des affirmations qui le plus souvent servent à montrer l'urgence de réformer l'orthographe du français réputée pour être complexe et opaque comme celle de l'anglais.

Le but de notre recherche était de comparer la compétence en orthographe générale, et plus particulièrement en orthographe lexicale des élèves, de niveau sixième primaire, âgés de 12 ans. Les élèves du corpus proviennent de deux populations différentes. Pour faire cette recherche, il s'agissait de voir si effectivement, des hispanophones du Mexique présentaient sur le plan lexical une meilleure compétence orthographique que les francophones du Québec. Nos données proviennent du corpus Nadeau, Matute, Gonzales-Reyes (2006). Il s'agit de 4 dictées qui ont été administrées aux élèves dans leur langue maternelle respective. Notre corpus compte 40 élèves francophones (Québec) et 42 élèves hispanophones (Mexique). Les élèves retenus pour l'étude ont en commun, en plus de l'âge et du sexe, le même milieu socio-économique. De plus, nous avons tenté, dans un deuxième temps, de voir l'impact d'une réforme sur la compétence orthographique des francophones.

Pour les besoins de notre analyse, nous avons procédé à certaines manipulations dans le corpus : une épuration des dictées pour un total de 315 mots pour les francophones et 312 pour les hispanophones, un relevé systématique de toutes les graphies erronées produites par chaque élève, une classification des mots selon les fréquences. Nous nous sommes ainsi préparés à répondre à nos questions de recherche qui sont les suivantes :

- Dans le corpus étudié, les enfants francophones et hispanophones du même âge présentent-ils, pour leur langue première respective, le même niveau de maîtrise sur le plan de l'orthographe en général, et de l'orthographe lexicale en particulier ?
- Est-ce que la fréquence des mots tient effectivement un rôle important dans la maîtrise de l'orthographe lexicale ?
 - Est-ce que la fréquence des mots joue un rôle identique dans chacun des systèmes orthographiques comparés ?
- Rencontrons-nous les mêmes types d'erreurs lexicales dans les textes des élèves de chaque langue et dans quelle proportion ?
- Quelle réforme parmi celles de Dauzat (1939-1952) ; de la Commission Beslais (1960-1965) ; de Thimonnier (1967) aurait permis de diminuer le plus les fautes d'orthographe dans les copies des élèves francophones ? Et dans quel pourcentage ?

Les résultats nous montrent que pour l'ensemble des erreurs orthographiques, les hispanophones présentent un meilleur taux de réussite que les francophones. Pour notre corpus, les hispanophones ont un taux réussite de 90.73 % par élève comparativement à 83.82 % pour les francophones. Des résultats qui confirment notre hypothèse de recherche : les élèves hispanophones écrivent en moyenne mieux que les élèves francophones.

Toutefois, cette meilleure performance s'explique par le fait qu'ils écrivent mieux sur le plan de l'orthographe grammaticale. Nous avons un pourcentage d'erreurs grammaticales chez les francophones qui équivaut à 4 fois celui des hispanophones. Cette différence s'explique par la régularité du système orthographique de l'espagnol et par l'existence de marques orales de genre et de nombre. La marque du féminin est bien marquée à l'oral (*amigo/amiga*) et les élèves disposent aussi d'un indice sonore pour la marque du pluriel (*amigo/amigos*), phénomène inexistant en français sauf dans quelques cas comme les liaisons (les élèves). Ce fait nous laisse supposer que le « s » du pluriel dans *los amigos* est plus facile à écrire que dans *les clowns*.

Pour ce qui concerne la performance en orthographe lexicale, la moyenne d'erreurs lexicales par élève nous montre qu'il n'y a presque pas de différence entre élèves francophones et hispanophones pour notre corpus : 8.79 % chez les francophones vs 7.51 % pour les hispanophones. Il est clair que nous ne constatons pas une domination des hispanophones tel qu'attendu en nous basant sur les affirmations de Masson (1991) et de Jaffré et Fayol (1997). Par conséquent, nous avons dû rejeter notre hypothèse concernant la plus grande réussite des hispanophones sur le plan de l'orthographe lexicale.

En rapport au rôle des fréquences dans la maîtrise de l'orthographe lexicale, nos résultats ont montré qu'effectivement les mots les plus fréquents sont ceux que les élèves maîtrisent le mieux dans l'apprentissage du code orthographique. Une démonstration faite à partir des mots mal orthographiés par au minimum 50 % des élèves de chaque population et à partir des mots qui ne contiennent aucune erreur lexicale d'aucun élève et des mots bien réussis dans chaque population. Les mots qui causent le plus de difficulté

aux élèves apparaissent majoritairement dans la catégorie des mots peu fréquents (PF). En poursuivant notre étude à l'aide des mots bien réussis par les élèves de chaque population, nous avons constaté que la fréquence ne joue pas un rôle identique dans les deux langues. Les mots F et PF bien réussis sont nombreux dans les deux langues : 72.30 % et 24 % pour les francophones, 80 % et 51 % pour les hispanophones. L'analyse des graphies a démontré que les mots courts et plus ou moins transparents en français expliquent la présence d'autant de mots PF parmi les mots bien réussis. Pour le corpus hispanophone, les mots PF sont tout aussi longs et opaques que les mots PF manqués massivement par 50 % et plus des élèves mexicains.

L'étude des types d'erreurs lexicales nous montre l'existence de « zones de difficulté » dans les deux langues. Nous retrouvons le plus grand nombre d'erreurs dans les catégories phonogrammique et phonétique : 55.70 % et 28.30 % chez les francophones, 83.22 % et 12.20 % chez les hispanophones. Cette analyse nous montre également des erreurs lexicales de même sorte dans les deux langues. Cependant, chaque groupe a sa zone « problématique ». Dans le cas des francophones, les erreurs phonogrammiques reliées à l'adjonction de lettres inaudibles montrent un pourcentage beaucoup plus élevé que les hispanophones (10.45 % vs 0.50 %). Chez les hispanophones, la transcription du phonème /s/ et les erreurs phonogrammiques reliées à l'accent tonique représentent la pierre d'achoppement. En effet, chez les hispanophones, nous retrouvons une grande concentration d'erreurs lexicales dans ces deux sections : 28.04 % du total des erreurs lexicales pour la transcription du phonème /s/ et 41.3 % du total des erreurs lexicales pour l'accent tonique. Il serait utile pour établir une meilleure comparaison entre les deux groupes de poursuivre cette étude sur le terrain, en comparant les méthodes d'enseignement dans chaque population. Par exemple, une enquête auprès des enseignants, des élèves et

des répondants dans la société mexicaine permettrait d'avoir une idée juste de l'importance accordée à l'orthographe comme matière scolaire et indice de réussite sociale. Nous pensons également que nous pourrions compléter cette étude sur la complexité du français à travers un échantillonnage plus large de francophones et d'hispanophones. Nous pensons que ces études donneraient des résultats intéressants pour comparer la transparence de l'espagnol versus l'opacité du français.

Pour finir, nous avons vérifié l'impact de quelques projets de réforme sur les erreurs lexicales des francophones de notre corpus. La réforme de Beslais (1965) s'avère la seule qui puisse avoir un certain impact. Notre projection montre en effet qu'il serait possible d'éliminer 17.20 % du total des erreurs lexicales relevées dans les textes des élèves. Nous tenons cependant à préciser que les mots qui contiennent ces erreurs retenues pour cette projection comportent souvent plus d'une erreur. Nous n'excluons pas les erreurs grammaticales dans ce contexte parce qu'il est déjà prouvé que la grammaire cause beaucoup de problèmes aux francophones. Rappelons à juste titre la récente étude de Manesse et Cogis (2007) qui pointe du doigt l'orthographe grammaticale.

L'enquête de 1987, si elle avait conclu à une augmentation générale de la maîtrise de l'orthographe en un siècle, c'est-à-dire depuis la tournée de l'Inspecteur Beuvain, avait néanmoins déjà pointé une tendance à la dégradation dans le domaine grammatical. L'enquête de 2005 montre que cette tendance s'est fortement accentuée. (p.104)

Cette étude montre, à travers les résultats, qu'une réforme de l'orthographe du français est souhaitable pour permettre à tout francophone d'avoir un accès plus facile à l'orthographe de la langue. Comme le dit si bien Jaffré :

Les erreurs qui durent sont révélatrices de dysfonctionnements qui pourraient annoncer des changements, en langue comme en

orthographe. Et à l'intention de ceux qui se plaignent du nombre sans cesse croissant d'erreurs, ou qui déplorent la baisse du niveau en orthographe, paraphrasons une remarque de Nina Catach dans les délires de l'orthographe : « Changez l'orthographe et ils écriront mieux ». (1999, p.154)

APPENDICE A

CORPUS NADEAU, MATUTE, GONZALES-REYES (2006)

- | | | |
|-----|--|-----|
| A.1 | Les textes dictés aux élèves québécois | 148 |
| A.2 | Les textes dictés aux élèves mexicains | 149 |

A.1 les textes dictés aux élèves québécois.

DICTÉE du texte 1

Courir 42 kilomètres et des poussières, ça vous dirait? Plusieurs humains y arrivent en deux heures et quelques minutes. Quels athlètes! En 2002, 32 000 compétiteurs ont participé au marathon de Londres. Parmi tous ces coureurs, il y avait Lloyd Scott. Il a réussi à battre tous les records de lenteur! Son marathon a duré presque six jours! Il faut dire qu'il avait enfilé une combinaison de scaphandrier. Elle pesait 54.5 kilogrammes! Chaque botte pesait plusieurs kilogrammes! Scott souhaitait recueillir des fonds pour la lutte contre les leucémies et les cancers. Il a lui-même vaincu une leucémie il y a quelques années. Il a vraiment beaucoup de courage!

DICTÉE du texte 2

Plusieurs cirques canadiens et américains participent au Festival des festivals. Six immenses chapiteaux bleus sont montés dans la plaine. Sous l'un d'eux, on présente une création collective québécoise. Elle a été imaginée par une équipe de jeunes. Des clowns très colorés font le bonheur des spectateurs. Trois grandes jongleuses fascinent la foule. Ensuite, les premières trapézistes s'élancent. Ce sont des athlètes particulièrement talentueuses. Une s'est blessée au genou gauche. Son entraîneuse inquiète l'observe attentivement.

DICTÉE du texte 3

Je suis Mimi la chatte. Depuis l'accident, je vis ici. Rita et Lucas, les enfants de la maison, sont charmants. Rita me flatte souvent. Ses caresses et ses paroles sont douces et réconfortantes. Ce sont les meilleures! Rita me confie ses secrets. Ses journées à l'école ne semblent pas faciles. Hier, elle paraissait nerveuse. Aujourd'hui, ça va mieux. Souvent, Lucas et Rita restent allongés devant moi. C'est un jeu, ils demeurent immobiles jusqu'à ce que je bouge! Ce sont des enfants extraordinaires.

DICTÉE du texte 4

Marine dessine un cirque. Sous le chapiteau, les acrobates, les animaux et la foule semblent bien réels. Les couleurs du dessin illuminent la pièce. Tout à coup, les personnages commencent à bouger! Trois clowns en équilibre sur une drôle de bicyclette sortent du dessin et avancent vers Marine. Ils exécutent leur numéro et la petite leur parle comme si tout cela était normal! Sur le dessin, le public applaudit. Marine appelle son petit frère.

A.2 Les textes dictés aux élèves mexicains.

Texto 1

Correr 42 kilómetros y un poco más, ¿qué te parecería? Muchos humanos lo hacen en dos horas y algunos minutos. ¡Qué buenos atletas! En 2002, 32 000 competidores participaron en un maratón en Londres. Entre todos estos corredores, estaba Lloyd Scott, quien logró vencer todos los récords de lentitud. ¡Su maratón duró casi seis días! Hay que decir que se había puesto un traje de buzo. ¡Pesaba 54.5 kilogramos! Cada bota pesaba varios kilogramos. Scott deseaba reunir fondos para la lucha contra las leucemias y los cánceres. El mismo venció una leucemia hace algunos años. ¡Realmente, él tiene mucho ánimo!

Texto 2

Muchos circos canadienses y americanos participan en el *Festival de festivales*. Seis inmensas carpas azules se colocaron en la planicie. Bajo una de ellas, se presenta una creación colectiva quebequense. Ella fue imaginada por un equipo de jóvenes. Payasos muy coloridos hacen la dicha de los espectadores. Tres grandes malabaristas fascinan a la multitud. Enseguida, las primeras trapezistas se lanzan. Son atletas particularmente talentosas. Una de ellas se hirió la rodilla izquierda. Su entrenadora inquieta la observa con mucha atención.

Texto 3

Yo soy la gata Mimí. Después del accidente, vivo aquí. Lupita y Lucas, los niños de la casa, son encantadores. Lupita me acaricia frecuentemente. Sus caricias y sus palabras son dulces y reconfortantes. ¡Son las mejores! Lupita me confía sus secretos. Sus jornadas en la escuela no parecen ser fáciles. Ayer, ella parecía nerviosa. Hoy, se ve mejor. Con frecuencia, Lucas y Lupita se quedan recostados delante de mí. Es un juego, ¡ellos permanecen inmóviles hasta que yo me mueva! ¡Son unos niños extraordinarios!

Texto 4

Marina dibuja un circo. Bajo la carpa, los acróbatas, los animales y la multitud parecen muy reales. Los colores del dibujo iluminan la pieza. De repente, ¡los personajes comienzan a moverse! Tres payasos en equilibrio sobre una curiosa bicicleta salen del dibujo y avanzan hacia Marina. Ellos hacen su número y la pequeña les habla como ¡si todo eso fuera normal! Sobre el dibujo, el público aplaude. Marina busca a su hermanito.

APPENDICE B

CODIFICATION DES GRAPHIES ERRONÉES POUR CHAQUE CORPUS

B.1 Les tableaux des graphies erronées pour les francophones	151
B.2 Les tableaux des graphies erronées pour les hispanophones	169

Dictée	1 (f)	2 (g)	3 (f)	4 (g)	5 (g)	6 (f)	7 (g)	8 (g)	10 (g)	12 (f)	13 (f)	14 (f)	15 (f)
s'est	3	3	3	3	0	3	3	3	3	0	3	3	0
blessée	0	12	0	12	2	12	1	0	1	0	1	12	12
au	0	0	0	0	0	0	0	0	0	0	0	0	0
genou	0	0	0	0	0	2	0	1	2	2	1	2	2
gauche	0	0	0	0	0	0	0	0	0	0	0	0	0
Son	0	0	0	3	0	0	0	0	0	0	3	0	0
entraîneuse	0	1	0	0	0	0	0	0	0	0	0	0	1
inquiète	0	0	1	0	0	0	0	1	0	0	0	0	2
l'observe	0	0	0	0	1	0	1	0	0	0	0	0	1
attentivement	0	0	0	0	0	0	0	0	1	0	1	0	0
Je	0	0	0	0	0	0	0	0	0	0	0	0	0
suis	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
chatte	0	1	1	0	1	0	0	0	0	0	0	0	0
Depuis	0	0	0	0	0	0	0	0	0	0	0	0	0
l'accident	0	0	0	0	0	1	0	0	0	0	1	0	0
je	0	0	0	0	0	0	0	0	0	0	0	0	0
vis	0	0	0	1	0	1	0	1	0	0	0	0	1
ici	0	0	0	0	0	0	0	0	0	0	0	0	0
et	0	0	0	0	0	0	0	0	0	0	0	0	0
les	0	0	0	0	0	0	0	0	0	0	0	0	0
enfants	0	0	0	0	2	0	0	0	0	0	0	0	0
de	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
maison	0	0	0	0	0	0	0	0	0	0	0	0	0
sont	3	0	0	0	0	0	0	0	3	0	0	0	0
charmants	0	12	1	2	2	0	2	0	2	0	1	0	2
me	0	0	0	0	0	0	0	0	0	0	0	0	0
flatte	0	1	1	1	1	0	1	1	1	0	1	0	0
souvent	0	0	0	0	0	0	0	0	0	0	0	0	0
Ses	0	0	0	0	0	0	0	3	3	0	0	3	0
caresses	1	1	0	0	12	12	2	12	0	0	0	0	0
et	0	0	0	0	0	0	0	0	0	0	0	0	0
ses	0	0	0	0	0	3	3	3	3	0	0	0	0
paroles	0	12	0	2	2	2	2	2	0	0	0	0	0
sont	0	0	0	0	0	0	3	0	3	0	3	0	0
douces	0	2	0	2	2	12	2	2	2	0	1	0	2
et	0	0	0	0	0	0	0	0	0	0	0	0	0
réconfortantes	0	12	0	2	2	12	2	2	1	0	12	0	2
Ce	0	3	0	0	0	3	3	0	0	0	0	0	0
sont	0	3	0	0	0	0	3	0	0	0	0	0	0
les	0	0	0	0	0	0	0	0	0	0	0	0	0
meilleures	0	2	2	2	2	12	2	2	1	2	2	2	2
me	0	0	0	0	0	0	0	0	0	0	0	0	0
confie	0	1	1	0	0	0	1	1	1	1	0	1	0
ses	0	0	0	0	0	3	0	3	0	0	0	3	0
secrets	0	2	0	2	0	2	0	2	2	0	1	0	0
Ses	0	0	0	0	0	3	0	3	0	0	0	0	0
journées	0	2	0	2	0	2	2	0	2	0	0	0	0
à	0	0	0	0	0	0	0	0	0	0	0	0	0
l'école	0	0	0	0	0	0	0	0	0	0	0	0	0
ne	0	0	0	0	0	0	0	0	1	0	0	0	0
semblent	0	2	0	2	2	2	2	2	2	0	2	0	2

Dictée	1 (f)	2 (g)	3 (f)	4 (g)	5 (g)	6 (f)	7 (g)	8 (g)	10 (g)	12 (f)	13 (f)	14 (f)	15 (f)
pas	0	0	0	0	0	0	0	0	0	0	0	0	0
faciles	0	12	0	2	2	2	2	2	2	0	12	0	0
Hier	0	0	0	0	0	0	0	0	0	0	0	0	0
elle	0	0	0	0	0	0	0	0	0	0	0	0	0
paraissait	0	12	0	1	1	1	0	0	2	0	0	0	0
nerveuse	0	1	0	0	0	0	0	0	0	0	0	0	0
Aujourd'hui	0	0	0	1	0	0	0	1	0	0	1	0	0
ça	0	1	0	1	0	0	1	1	0	0	0	1	0
va	0	0	0	0	0	0	0	0	0	0	0	0	0
mieux	0	1	0	1	1	0	1	1	0	0	1	0	0
Souvent	0	0	0	0	0	0	0	0	0	0	0	0	0
et	0	0	0	0	0	0	0	0	0	0	0	0	0
restent	0	2	0	2	2	0	2	0	2	0	2	2	2
allongés	0	2	12	2	2	2	12	2	2	2	1	1	12
devant	0	0	0	0	1	0	0	0	0	0	0	0	0
moi	0	0	0	0	0	0	0	0	0	0	0	0	0
C'est	0	0	0	0	0	0	3	0	0	0	0	0	0
un	0	0	0	0	0	0	0	0	0	0	0	0	0
jeu	2	0	2	0	2	2	2	2	2	0	0	0	2
ils	0	2	0	0	0	0	2	2	0	0	0	0	0
demeurent	0	12	0	0	0	0	2	2	0	0	12	0	0
immobiles	1	12	1	12	2	12	2	12	1	2	12	1	2
jusqu'à	0	1	0	0	0	0	1	0	0	0	0	0	0
ce	0	3	0	0	0	3	0	0	0	0	0	0	3
que	0	0	0	0	0	0	0	0	0	0	0	0	0
je	0	0	0	0	0	0	0	0	0	0	0	0	0
bouge	0	1	0	0	0	0	2	0	0	0	0	0	0
Ce	0	3	0	0	0	3	0	0	0	0	0	0	0
sont	0	3	0	0	0	0	3	0	0	0	3	0	3
des	0	0	0	0	0	0	0	0	0	0	0	0	0
enfants	0	2	0	0	0	0	0	0	0	0	0	0	0
extraordinaires	0	12	0	12	0	1	2	0	2	0	12	1	2
dessine	0	abs	0	0	0	0	0	0	0	0	0	0	0
un	0	abs	0	0	0	0	0	0	0	0	0	0	0
cirque	0	abs	0	0	0	0	0	0	1	0	0	0	0
Sous	0	abs	0	0	0	0	0	0	0	0	0	0	0
le	0	abs	0	0	0	0	0	0	0	0	0	0	0
chapiteau	0	abs	0	12	0	0	0	12	1	2	0	1	1
les	0	abs	0	0	0	0	0	0	0	0	0	0	0
acrobates	0	abs	0	0	0	1	0	0	1	0	0	0	1
les	0	abs	0	0	0	0	0	0	0	0	0	0	0
animaux	0	abs	0	0	0	0	0	0	0	0	0	0	0
et	0	abs	0	0	0	0	0	0	0	0	0	0	0
la	0	abs	0	0	0	0	0	0	0	0	0	0	0
foule	0	abs	0	0	0	0	0	1	0	0	0	0	0
semblent	0	abs	0	2	2	2	2	2	2	0	2	0	2
bien	0	abs	0	0	0	0	0	0	0	0	0	0	0
réels	0	abs	0	2	2	0	2	1	12	0	2	0	2
Les	0	abs	0	0	0	0	0	0	0	0	0	0	0
couleurs	0	abs	0	2	2	1	0	0	0	0	0	0	0
du	0	abs	0	0	0	0	0	0	0	0	0	0	0
dessin	0	abs	0	0	0	0	0	0	0	0	0	0	0
illuminent	0	abs	0	2	12	0	2	12	12	0	12	0	2

Dictée	16 (g)	17 (g)	18 (g)	19 (f)	20 (g)	21 (g)	22 (g)	23 (f)	24 (f)	25 (g)	26 (g)	27 (f)	28 (g)
festivals	0	1	0	0	0	0	0	0	0	0	0	0	0
immenses	12	12	1	0	1	1	1	0	1	0	0	0	1
chapiteaux	1	0	1	1	0	1	0	1	1	0	0	2	1
bleus	2	0	2	0	0	2	0	0	0	0	0	0	2
sont	0	0	3	0	0	0	0	0	0	0	0	0	0
montés	2	12	2	0	2	0	2	0	0	0	2	0	0
dans	0	0	0	0	0	0	0	1	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
plaine	1	0	0	0	0	0	1	1	2	0	0	0	0
Sous	0	0	0	0	0	0	0	0	0	0	0	0	0
l'un	0	0	0	0	0	0	0	0	0	0	0	0	0
d'eux	0	0	1	1	0	0	1	0	0	1	0	0	1
on	3	0	0	0	0	3	0	0	3	0	3	0	3
présente	1	2	1	0	12	1	0	0	0	0	2	0	2
une	0	0	0	0	0	0	0	0	0	0	0	0	0
création	0	0	0	0	1	0	0	0	0	0	0	0	0
collective	1	0	0	0	0	1	0	0	0	0	0	1	0
québécoise	1	1	2	1	1	1	1	1	0	0	1	1	1
Elle	0	0	0	0	0	0	0	0	0	0	0	0	0
a	0	0	0	0	0	0	0	0	0	0	0	0	0
été	0	0	0	0	0	0	0	0	0	0	0	0	0
imaginée	2	2	12	2	2	0	0	2	2	2	2	0	2
par	0	0	0	0	0	0	0	0	0	0	0	0	0
une	0	0	0	0	0	0	0	0	0	0	0	0	0
équipe	0	1	0	0	0	0	0	0	0	0	0	0	0
de	0	0	0	0	0	0	0	0	0	0	0	0	0
jeunes	0	0	2	2	0	2	0	0	0	0	0	0	2
Des	0	0	0	0	0	0	0	0	0	0	0	0	0
clowns	1	0	1	0	0	0	1	0	0	1	0	0	0
très	0	0	0	0	0	0	0	0	0	0	0	0	0
colorés	2	0	2	0	2	0	12	0	0	0	2	0	0
font	0	0	0	0	0	0	0	0	0	0	0	0	0
le	0	0	0	0	0	0	0	0	0	0	0	0	0
bonheur	0	2	0	0	0	0	0	0	0	0	0	0	1
des	0	0	0	0	0	0	0	0	0	0	0	0	0
spectateurs	0	0	0	0	0	0	0	0	0	0	0	0	0
grandes	2	2	2	0	0	0	0	0	0	0	0	0	0
jongleuses	2	2	2	1	0	0	0	0	0	0	0	0	0
fascinant	12	1	12	1	1	12	12	12	1	0	0	2	1
la	0	0	0	0	0	0	0	0	0	0	0	0	0
foule	0	2	0	0	0	0	0	0	0	0	0	0	0
Ensuite	1	0	1	0	0	0	0	0	0	0	0	0	0
les	0	0	0	0	0	0	0	0	0	0	0	0	0
premières	2	0	0	0	1	2	0	0	0	0	0	0	0
trapézistes	12	12	1	1	12	1	1	1	1	0	0	1	12
s'élancent	12	1	12	1	0	2	12	1	0	0	0	0	1
Ce	3	0	0	0	3	3	0	0	0	0	3	0	3
sont	0	0	3	0	0	0	0	0	0	0	0	0	0
des	0	0	0	0	0	0	0	0	0	0	0	0	0
athlètes	0	2	0	1	1	12	1	1	1	1	1	0	12
particulièrement	0	1	0	0	0	0	1	0	0	0	0	0	0
talentueuses	12	12	2	0	2	0	0	0	0	0	0	0	2
Une	0	0	0	0	0	0	0	0	0	0	0	0	0

Dictée	16 (g)	17 (g)	18 (g)	19 (f)	20 (g)	21 (g)	22 (g)	23 (f)	24 (f)	25 (g)	26 (g)	27 (f)	28 (g)
s'est	0	0	3	0	3	3	3	3	3	3	0	0	0
blessée	1	12	2	0	2	12	0	0	2	0	2	2	0
au	2	0	0	0	0	0	0	0	0	0	0	0	0
genou	0	2	0	0	0	12	2	0	0	0	0	0	0
gauche	0	0	0	0	0	0	0	0	0	0	0	0	0
Son	0	0	0	0	0	3	0	0	0	0	3	0	0
entraîneuse	1	0	0	0	0	0	0	0	0	0	0	0	0
inquiète	0	0	0	2	1	2	0	0	0	0	1	0	0
l'observe	0	1	0	0	0	0	0	0	0	0	0	0	0
attentivement	0	1	0	0	0	0	0	0	0	0	0	0	0
Je	0	0	0	0	0	0	0	0	0	0	0	0	0
suis	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
chatte	0	0	0	0	0	0	0	0	0	0	0	0	0
Depuis	0	0	0	0	0	0	0	0	0	0	0	0	0
l'accident	0	1	1	0	0	1	0	0	0	0	0	0	0
je	0	0	0	0	0	0	0	0	0	0	0	0	0
vis	0	1	1	0	0	1	0	0	0	0	1	0	0
ici	0	1	0	0	0	0	0	0	0	0	0	0	0
et	0	0	0	0	0	0	0	0	0	0	0	0	0
les	0	0	0	0	0	0	0	0	0	0	0	0	0
enfants	0	0	0	0	0	0	0	0	0	0	0	0	0
de	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
maison	0	0	0	0	0	0	0	0	0	0	0	0	0
sont	0	3	3	0	0	0	0	0	0	0	0	0	0
charmants	2	12	2	0	12	12	0	0	0	0	2	0	2
me	0	0	0	0	0	0	0	0	0	0	0	0	0
flatte	1	1	1	0	1	0	0	1	1	0	0	0	0
souvent	0	0	0	0	0	0	0	0	0	0	0	0	0
Ses	3	3	3	0	0	0	3	0	0	0	0	0	3
caresses	2	1	2	0	0	0	0	1	0	0	0	0	12
et	0	0	1	0	0	0	0	0	0	0	0	0	0
ses	3	0	0	0	0	0	3	0	0	0	0	0	0
paroles	0	2	2	0	0	1	0	0	0	0	0	0	12
sont	0	0	3	0	0	0	0	0	0	0	0	0	0
douces	2	2	2	0	0	2	1	0	0	0	0	0	2
et	0	0	0	0	0	0	0	0	0	0	0	0	0
réconfortantes	2	12	2	0	1	12	12	0	1	0	0	2	2
Ce	0	0	0	3	3	0	3	0	3	0	0	0	3
sont	0	0	3	0	0	0	0	0	0	0	0	0	0
les	0	0	0	0	0	0	0	0	0	0	0	0	0
meilleures	2	2	2	2	2	2	2	0	0	2	2	0	2
me	0	0	0	0	0	0	0	0	0	0	0	0	0
confie	1	1	2	0	1	1	1	1	1	0	1	1	1
ses	3	3	0	0	0	0	3	0	0	0	0	0	0
secrets	2	2	12	0	0	0	0	0	0	0	0	0	0
Ses	3	3	0	0	0	3	3	0	0	0	0	0	0
journées	12	2	0	0	0	1	0	0	0	0	0	0	0
à	0	0	0	0	3	0	0	0	0	0	0	0	0
l'école	0	1	0	0	0	0	0	0	1	0	0	0	0
ne	0	0	0	0	0	0	0	0	0	0	0	0	0
semblent	2	2	2	0	2	2	2	0	0	0	2	0	2

Dictée	16 (g)	17 (g)	18 (g)	19 (f)	20 (g)	21 (g)	22 (g)	23 (f)	24 (f)	25 (g)	26 (g)	27 (f)	28 (g)
pas	0	0	0	0	0	0	0	0	0	0	0	0	0
faciles	2	12	2	0	2	2	0	0	0	0	2	0	2
Hier	0	0	0	0	0	0	0	0	0	0	0	0	0
elle	0	0	0	0	0	0	0	0	0	0	0	0	0
paraissait	1	1	12	0	1	1	0	0	0	1	1	0	1
nerveuse	0	0	0	0	0	0	0	0	0	0	0	0	0
Aujourd'hui	0	0	0	0	0	0	0	0	0	0	0	0	1
ça	0	1	1	0	1	1	0	0	0	0	0	0	1
va	0	0	0	0	0	0	0	0	0	0	0	0	0
mieux	1	1	1	1	1	0	0	0	0	0	0	0	0
Souvent	0	0	0	1	0	0	0	0	0	0	0	0	0
et	0	0	0	0	0	0	0	0	0	0	0	0	0
restent	2	2	2	0	0	2	2	0	0	0	0	0	0
allongés	2	2	2	0	1	2	2	1	0	0	2	0	12
devant	0	0	0	0	0	0	0	0	0	0	0	0	0
moi	0	1	0	0	0	0	0	0	0	0	0	0	0
C'est	0	0	0	0	0	0	0	0	3	0	0	0	0
un	0	0	0	0	0	0	0	0	0	0	0	0	0
jeu	0	0	0	0	0	0	0	0	0	0	2	0	0
ils	0	2	0	2	0	2	0	0	0	0	0	0	0
demeurent	12	12	0	12	0	12	0	0	0	0	1	0	0
immobiles	2	12	2	12	12	12	2	0	1	0	2	12	12
jusqu'à	0	1	1	0	1	0	0	0	0	0	0	0	0
ce	3	0	0	0	3	0	0	0	0	0	0	0	0
que	0	0	0	0	0	0	0	0	0	0	0	0	0
je	0	0	0	0	0	0	0	0	0	0	0	0	0
bouge	0	0	0	0	0	0	0	0	0	0	0	0	0
Ce	0	0	0	3	3	0	3	0	0	0	0	0	3
sont	0	0	3	0	0	0	0	0	0	0	0	0	0
des	0	abs	0	0	0	0	0	0	0	0	0	0	0
enfants	0	abs	0	0	0	0	0	0	0	0	0	0	0
extraordinaires	12	1	2	0	2	0	2	0	0	0	0	2	2
dessine	1	0	0	0	0	0	0	0	0	0	0	0	0
un	0	0	0	0	0	0	0	0	0	0	0	0	0
cirque	0	0	0	0	0	0	0	0	0	0	0	0	0
Sous	0	0	0	0	0	0	0	0	0	0	0	0	0
le	0	0	0	0	0	0	0	0	0	0	0	0	0
chapiteau	0	1	0	0	0	12	0	0	1	1	0	0	1
les	0	0	0	0	0	0	0	0	0	0	0	0	0
acrobates	0	0	1	0	0	0	1	0	1	0	0	0	0
les	0	0	0	0	0	0	0	0	0	0	0	0	0
animaux	0	0	0	0	1	0	0	0	0	0	0	0	1
et	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
foule	2	2	0	0	0	0	0	0	0	0	0	0	0
semblent	2	2	2	0	0	2	0	0	0	0	0	0	0
bien	0	0	0	0	0	0	0	0	0	0	0	0	0
réels	2	1	2	0	2	2	2	0	2	0	0	2	0
Les	0	0	0	0	0	0	0	0	0	0	0	0	0
couleurs	0	2	0	0	0	0	0	0	0	0	0	0	0
du	0	0	0	0	0	0	0	0	0	0	0	0	0
dessin	0	0	0	0	0	0	0	0	0	0	0	0	0
illuminent	12	12	2	0	1	2	0	1	0	1	0	0	1

Dictée	29 (g)	30 (g)	31 (g)	32 (f)	33 (f)	34 (g)	35 (g)	36 (g)	37 (g)	38 (f)	39 (g)	40 (f)	41 (f)
pas	0	0	0	0	0	0	0	0	0	0	0	0	0
faciles	0	12	2	0	0	2	0	2	0	0	0	0	2
Hier	0	0	0	0	0	0	0	0	0	0	0	0	0
elle	0	0	0	0	0	0	0	0	0	0	0	0	0
paraissait	0	0	1	1	0	1	1	0	0	0	0	0	0
nerveuse	0	0	0	0	0	0	0	0	0	0	0	0	0
Aujourd'hui	0	0	0	0	0	0	0	1	0	0	0	0	0
ça	0	0	0	0	0	0	0	0	0	0	0	0	0
va	0	0	0	0	0	0	0	0	0	0	0	0	0
mieux	0	1	0	0	0	0	0	0	0	0	0	0	0
Souvent	0	0	0	0	0	0	0	0	0	0	0	0	0
et	0	0	0	0	0	0	0	0	0	0	0	0	0
restent	0	0	0	0	0	0	0	0	0	0	0	0	0
allongés	2	0	2	0	0	0	0	2	0	0	0	1	0
devant	0	0	0	0	0	0	0	0	0	0	0	0	0
moi	0	0	0	0	0	0	0	0	0	0	0	0	0
C'est	0	0	0	0	0	0	0	0	0	0	0	0	0
un	0	0	0	0	0	0	0	0	0	0	0	0	0
jeu	2	0	0	0	0	0	0	2	0	0	0	0	2
ils	0	0	0	0	0	0	0	0	0	0	0	0	0
demeurent	0	0	0	0	0	0	0	0	0	0	0	0	0
immobiles	0	2	2	12	0	0	0	12	0	0	0	0	1
jusqu'à	1	0	1	0	0	0	0	0	0	0	0	1	0
ce	0	0	0	0	0	0	0	0	3	0	0	0	0
que	0	0	0	0	0	0	0	0	0	0	0	0	0
je	0	0	0	0	0	0	0	0	0	0	0	0	0
bouge	0	0	0	0	0	0	0	0	0	0	0	0	0
Ce	0	0	0	3	0	0	0	0	0	0	0	0	0
sont	0	0	0	0	0	0	0	0	0	0	0	0	0
des	0	0	0	0	0	0	0	0	0	0	0	0	0
enfants	0	0	0	0	0	0	0	0	0	0	0	0	0
extraordinaires	1	2	2	0	0	0	2	2	0	0	1	0	1
dessine	0	0	0	0	0	0	0	1	0	0	0	0	0
un	0	0	0	0	0	0	0	0	0	0	0	0	0
cirque	0	0	0	0	1	0	0	0	0	0	0	0	0
Sous	0	0	0	0	0	0	0	0	0	0	0	0	0
le	0	0	0	0	0	0	0	0	0	0	0	0	0
chapiteau	1	0	0	0	1	0	0	0	0	0	0	0	0
les	0	0	0	0	0	0	0	0	0	0	0	0	0
acrobates	0	1	0	0	0	0	0	1	0	0	0	0	0
les	0	0	0	0	0	0	0	0	0	0	0	0	0
animaux	0	0	0	0	0	1	0	0	0	0	0	0	0
et	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
foule	1	0	0	0	1	0	0	1	0	0	0	0	0
semblent	1	0	2	0	0	0	0	0	0	0	0	0	0
bien	0	0	0	0	0	0	0	0	0	0	0	0	0
réels	2	2	2	0	0	0	0	2	0	0	0	0	2
Les	0	0	0	0	0	0	0	0	0	0	0	0	0
couleurs	0	0	0	0	0	0	0	0	0	0	0	0	0
du	0	0	0	0	0	0	0	0	0	0	0	0	0
dessin	0	0	0	0	0	0	0	1	0	0	0	0	0
illuminent	12	0	1	0	0	0	0	1	0	2	0	0	0

Dictée	2 (f)	3 (f)	6 (F)	8 (f)	9 (f)	10 (m)	12 (m)	13 (f)	15 (m)	16 (m)	17 (f)	18 (m)	19 (m)
la	0	0	0	0	0	0	0	0	0	2	0	0	0
planicie	0	0	0	0	0	0	0	0	0	2	0	0	0
Bajo	2	0	0	0	0	0	0	0	0	0	0	0	0
una	0	0	0	0	0	0	0	0	0	0	0	0	0
de	0	0	0	0	0	0	0	0	0	0	0	0	0
ellas	0	0	0	0	0	0	0	0	2	0	0	0	0
se	0	0	0	0	0	0	0	0	0	0	0	0	0
presenta	0	0	0	0	0	0	0	0	0	2	0	0	0
una	0	0	0	0	0	0	0	0	0	0	0	0	0
creación	0	0	0	0	0	0	0	0	0	0	1	0	0
colectiva	0	0	0	0	0	0	0	0	0	0	0	0	0
quebequense	1	0	0	1	0	0	0	1	0	1	2	1	1
Ella	0	0	0	0	0	0	0	0	0	0	2	0	0
fue	0	1	1	0	1	1	0	0	0	0	0	0	0
imaginada	0	0	0	0	0	0	0	1	0	0	1	0	0
por	0	0	0	0	0	0	0	0	0	0	0	0	0
un	0	0	0	0	0	0	0	0	0	0	0	0	0
equipo	0	0	0	0	0	0	0	0	0	0	0	0	0
de	0	0	0	0	0	0	0	0	0	0	0	0	0
jóvenes	1	1	1	1	0	0	0	1	1	1	12	0	0
Payasos	0	0	0	0	0	0	1	0	1	0	0	0	0
muy	0	0	0	0	0	0	0	0	0	0	0	0	0
coloridos	0	0	0	0	0	0	0	0	0	0	0	0	0
hacen	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
dicha	1	0	0	0	0	0	0	0	0	0	0	0	0
de	0	0	0	0	0	0	0	0	0	0	0	0	0
los	0	0	0	0	0	0	0	0	0	0	0	0	0
espectadores	0	0	0	0	0	0	0	0	0	0	0	0	0
grandes	0	0	0	0	0	0	0	0	2	0	0	0	0
malabaristas	1	0	0	1	0	0	0	0	0	0	0	0	0
fascinan	1	0	1	1	1	1	1	1	1	1	1	1	0
a	0	0	0	0	0	0	0	abs	0	abs	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
multitud	0	0	0	0	0	0	0	0	0	0	0	0	0
Enseguida	0	0	0	0	0	0	1	0	0	1	0	0	0
las	0	0	0	0	0	0	0	0	0	0	0	0	0
primeras	abs	abs	abs	abs	abs	abs	abs	abs	abs	abs	abs	abs	abs
trapecistas	1	0	0	0	0	0	1	1	1	1	0	0	0
se	0	0	0	0	0	0	0	0	0	0	0	0	0
lanzan	0	0	0	0	0	0	0	0	1	0	1	0	0
Son	0	0	0	0	0	0	0	0	0	0	0	0	0
atletas	0	0	0	0	0	0	0	0	0	0	1	0	0
particularmente	0	0	0	0	0	0	1	abs	0	0	0	0	0
talentosas	0	0	0	0	0	0	0	0	1	0	0	0	0
Una	0	0	0	0	0	0	0	0	0	0	0	0	0
de	0	0	0	0	0	0	0	0	0	0	0	0	0
ellas	0	0	0	0	0	0	0	0	0	0	0	0	0
se	0	0	0	0	0	0	0	0	0	0	0	0	0
hirió	1	0	1	1	0	1	1	1	1	1	1	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
rodilla	0	0	0	0	0	0	0	0	0	0	0	0	0
izquierda	0	0	0	0	0	0	0	0	0	1	1	0	0

Dictée	2 (f)	3 (f)	6 (F)	8 (f)	9 (f)	10 (m)	12 (m)	13 (f)	15 (m)	16 (m)	17 (f)	18 (m)	19 (m)
Su	0	0	0	0	0	0	0	0	abs	0	0	0	0
entrenadora	0	0	0	0	0	0	0	0	0	0	0	0	0
inquieta	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
observa	1	0	0	0	0	0	0	0	1	1	0	0	0
con	0	0	0	0	0	0	0	abs	0	0	0	0	0
mucha	0	0	0	0	0	0	0	abs	abs	0	0	0	0
atención	0	0	0	0	0	0	0	abs	0	1	0	0	0
Yo	0	0	0	0	0	0	0	0	0	0	0	0	0
soy	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
gata	0	0	0	0	0	0	0	0	0	0	0	0	0
Después	1	0	1	1	0	1	1	1	1	0	1	0	0
del	0	0	0	0	0	0	0	0	0	0	0	0	0
accidente	0	0	0	0	0	0	0	0	1	0	0	0	0
vivo	0	0	0	0	0	0	0	0	0	0	0	0	0
aquí	0	1	0	1	0	0	1	0	1	0	0	0	0
Y	0	0	0	0	0	0	0	0	0	0	0	0	0
los	0	0	0	0	0	0	0	0	0	0	0	0	0
niños	0	0	1	0	0	1	0	1	0	0	0	0	0
de	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
casa	0	0	0	0	0	0	0	0	0	0	0	0	0
son	0	0	0	0	0	0	0	0	0	0	0	0	0
encantadores	0	0	0	0	0	0	0	0	0	0	0	0	0
me	0	0	0	0	0	0	0	0	0	0	0	0	0
acaricia	1	0	0	0	0	0	0	0	1	1	0	0	0
frecuentemente	0	0	0	0	0	0	0	0	1	0	1	0	0
Sus	0	0	0	0	0	0	0	0	0	0	0	0	0
caricias	0	0	0	0	0	0	1	0	0	0	1	0	1
y	0	0	0	0	0	0	0	0	0	0	0	0	0
sus	0	0	0	0	0	0	0	0	0	0	0	0	0
palabras	0	0	0	0	0	0	0	0	0	0	0	0	0
son	0	0	0	0	0	0	0	0	0	0	0	0	0
dulces	0	0	0	0	0	0	0	0	1	0	0	0	0
y	0	0	0	0	0	0	0	0	0	0	0	0	0
reconfortantes	0	1	0	0	0	1	1	2	0	0	1	0	0
Son	0	0	0	0	0	0	0	0	0	0	0	0	0
las	0	0	0	0	0	0	0	0	2	0	0	0	0
mejores	0	0	0	0	0	0	0	0	0	0	0	0	0
me	0	0	0	0	0	0	0	0	0	0	0	0	0
confía	1	1	0	1	0	0	1	1	0	1	1	0	0
sus	0	0	0	0	0	0	0	0	0	0	0	abs	0
secretos	0	0	0	0	0	0	0	0	0	0	0	0	0
Sus	0	0	0	0	0	0	0	0	0	0	0	0	0
jornadas	0	0	0	0	0	0	0	0	1	0	0	0	0
en	1	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
escuela	0	0	0	0	0	0	0	0	0	0	0	0	0
no	0	0	0	0	0	0	0	0	0	0	0	0	0
parecen	0	0	0	0	0	0	0	0	abs	0	0	0	0
ser	0	0	0	0	0	abs	0	0	2	0	0	0	0
fáciles	1	0	0	1	0	0	1	1	1	1	1	0	0

Dictée	20 (m)	21 (m)	22 (f)	23 (f)	24 (m)	25 (m)	26 (m)	29 (f)	30 (f)	32 (m)	33 (m)	35 (m)	36 (m)
la	0	0	0	0	0	0	0	0	0	0	0	0	0
planicie	0	0	0	0	1	1	0	0	0	0	0	0	0
Bajo	0	0	0	0	0	0	1	0	0	0	0	0	0
una	0	0	0	0	0	0	0	0	0	0	0	0	0
de	0	0	0	0	0	0	0	0	0	0	0	0	0
ellas	0	0	0	0	0	0	0	0	0	0	0	0	0
se	0	0	0	0	0	0	0	0	0	0	0	0	0
presenta	0	0	0	0	0	0	0	0	0	0	0	0	0
una	0	0	0	0	0	0	0	0	0	0	0	0	0
creación	1	1	0	0	1	0	0	0	0	1	0	1	0
colectiva	0	0	0	0	0	0	0	0	0	0	1	0	0
quebequense	1	1	1	1	1	1	1	0	0	1	1	0	1
Ella	0	0	0	0	0	0	0	0	0	0	0	0	0
fue	0	0	0	1	0	0	0	0	0	0	0	0	0
imaginada	0	0	0	0	0	0	0	0	0	1	2	0	0
por	0	0	0	0	0	0	0	0	0	0	0	0	0
un	0	0	0	0	0	0	0	0	0	0	0	0	0
equipo	0	0	0	0	0	0	0	0	0	0	0	0	0
de	0	0	0	0	0	0	0	0	0	0	0	0	0
jóvenes	1	1	1	0	1	0	1	1	0	0	1	1	0
Payasos	0	0	0	1	0	0	0	0	0	1	0	0	1
muy	0	0	0	0	0	0	0	0	0	0	0	0	0
coloridos	0	0	0	0	0	1	0	0	0	0	0	0	1
hacen	0	0	0	0	0	0	0	0	0	1	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
dicha	0	0	0	0	0	0	0	0	0	0	0	0	0
de	0	0	0	0	0	0	0	0	0	0	0	0	0
los	0	0	0	0	0	0	0	0	0	0	0	0	0
espectadores	0	0	0	0	0	0	0	0	0	0	0	0	0
grandes	0	0	0	0	0	0	0	0	0	0	0	0	0
malabaristas	0	1	1	0	0	1	0	1	0	2	1	1	0
fascinan	1	1	1	1	1	1	1	1	1	1	1	1	1
a	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
multitud	0	0	0	0	0	0	0	0	0	0	0	0	0
Enseguida	0	0	0	0	0	0	0	0	0	0	0	0	0
las	0	0	0	0	0	0	0	0	0	0	0	0	0
primeras	abs	abs	abs	abs	abs	abs	abs	abs	abs	0	0	0	0
trapecistas	0	1	1	0	1	0	1	0	1	0	0	0	1
se	0	0	0	0	0	0	0	0	0	0	0	0	0
lanzan	0	1	0	0	0	0	1	0	0	1	0	1	1
Son	0	0	0	0	0	0	0	0	0	0	0	0	0
atletas	0	0	0	0	0	0	0	0	0	0	0	0	0
particularmente	0	0	0	0	1	0	abs	0	0	1	0	0	0
talentosas	0	0	12	0	0	0	1	0	0	0	0	0	0
Una	0	0	0	0	0	0	0	0	0	0	0	0	0
de	0	0	0	0	0	0	0	0	0	0	0	0	0
ellas	0	0	0	0	0	0	0	0	0	0	0	0	0
se	0	0	0	0	0	0	0	0	0	0	0	0	0
hirió	0	1	1	0	0	0	1	1	1	1	1	1	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
rodilla	0	0	0	0	0	0	0	0	0	0	0	0	0
izquierda	0	1	0	0	0	0	0	0	0	1	1	1	0

Dictée	20 (m)	21 (m)	22 (f)	23 (f)	24 (m)	25 (m)	26 (m)	29 (f)	30 (f)	32 (m)	33 (m)	35 (m)	36 (m)
Su	0	0	0	0	0	0	0	0	0	0	0	0	0
entrenadora	0	0	0	0	0	0	0	0	0	0	1	0	0
inquieta	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
observa	0	0	0	0	1	0	1	0	0	1	0	1	0
con	0	0	0	0	0	0	0	0	0	0	0	0	0
mucha	0	0	0	0	0	0	abs	0	0	0	0	0	0
atención	0	0	0	0	0	0	1	0	1	1	0	0	0
Yo	0	0	0	0	0	0	0	0	0	0	0	0	0
soy	0	0	0	0	0	0	0	0	0	1	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
gata	0	0	0	0	0	0	0	0	0	0	0	0	0
Después	0	0	0	0	1	0	1	0	0	1	1	1	1
del	0	0	0	0	0	0	0	0	0	0	0	0	2
accidente	0	0	1	0	0	0	0	0	0	0	0	0	0
vivo	0	0	0	0	0	0	0	0	0	0	0	0	0
aquí	0	1	0	0	0	0	1	0	1	1	0	1	0
Y	0	0	0	0	0	0	0	0	0	0	0	0	0
los	0	0	0	0	0	0	0	0	0	0	0	0	0
niños	0	0	0	0	0	0	0	0	0	0	0	0	1
de	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
casa	0	0	0	0	0	0	0	0	0	0	0	0	0
son	0	0	0	0	0	0	0	0	0	0	0	0	0
encantadores	0	0	0	0	0	0	0	0	0	0	0	0	0
me	0	0	0	0	0	0	0	0	0	0	0	0	0
acaricia	0	0	0	0	0	1	1	0	1	0	0	0	0
frecuentemente	0	0	0	0	0	0	abs	0	0	0	0	0	0
Sus	0	0	0	0	0	0	0	0	0	0	0	0	0
caricias	0	0	0	0	0	1	0	0	0	0	1	0	0
y	0	0	0	0	0	0	abs	0	0	0	0	0	0
sus	0	0	0	0	0	0	0	0	0	0	0	0	0
palabras	0	0	0	0	0	0	0	0	0	0	0	0	0
son	0	0	0	0	0	0	0	0	0	0	0	0	0
dulces	0	0	0	0	0	0	0	0	0	0	1	0	0
y	0	0	0	0	0	0	0	0	0	0	0	0	0
reconfortantes	0	0	0	0	0	0	abs	0	0	0	0	1	0
Son	0	0	0	0	0	0	0	0	0	0	0	0	0
las	0	0	0	0	0	0	0	0	0	0	0	0	0
mejores	0	0	0	0	0	0	0	0	0	0	0	0	0
me	0	0	0	0	0	0	0	0	0	0	0	0	0
confía	0	1	1	0	0	1	1	0	0	1	0	1	1
sus	0	0	0	0	0	0	0	0	0	0	0	0	0
secretos	0	0	0	1	0	0	0	0	0	0	0	0	0
Sus	0	0	0	0	0	0	0	0	0	0	0	0	0
jornadas	0	0	0	0	0	0	0	0	0	0	0	0	0
en	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
escuela	0	0	0	0	1	0	0	0	0	0	0	0	0
no	0	0	0	0	0	0	0	0	0	0	0	0	0
parecen	1	0	0	0	0	1	0	0	0	0	0	0	0
ser	0	0	0	0	0	0	0	0	0	0	0	0	0
fáciles	1	1	1	0	0	1	1	0	1	1	1	1	1

Dictée	20 (m)	21 (m)	22 (f)	23 (f)	24 (m)	25 (m)	26 (m)	29 (f)	30 (f)	32 (m)	33 (m)	35 (m)	36 (m)
la	0	0	0	0	0	0	0	0	0	0	0	0	0
pieza	0	0	0	0	0	0	1	0	0	1	1	0	1
De	0	0	0	0	0	0	0	0	0	0	0	0	0
repente	0	0	0	1	0	0	0	1	0	1	0	1	1
los	0	0	0	0	0	0	0	0	0	0	0	0	0
personajes	0	0	0	0	1	0	0	0	0	0	0	0	0
comienzan	0	0	0	0	0	0	0	0	0	1	0	0	0
a	0	0	0	0	0	0	0	0	0	0	0	0	0
moverse	0	1	0	0	0	0	0	0	0	0	0	0	0
payasos	0	0	0	0	0	0	0	0	0	1	0	0	1
en	0	0	0	0	0	0	0	0	0	0	0	0	0
equilibrio	0	0	1	0	0	0	0	0	0	0	0	0	0
sobre	0	0	0	0	0	0	0	0	0	0	0	0	0
una	0	0	0	0	0	0	0	0	0	0	0	0	0
curiosa	0	0	0	0	0	0	0	0	0	0	0	0	0
bicicleta	1	0	0	0	0	0	0	0	0	1	0	0	1
salen	0	0	0	0	0	0	0	0	0	0	0	0	0
del	0	0	0	0	0	0	0	0	0	0	0	0	0
dibujo	0	0	0	0	0	0	0	0	0	1	0	0	0
y	0	0	0	0	0	0	0	0	0	0	abs	0	0
avanzan	0	0	1	0	0	0	1	0	1	1	0	0	0
hacia	0	0	0	0	0	0	0	0	0	1	0	1	1
Ellos	0	0	0	0	0	0	0	0	0	0	0	0	0
hacen	0	0	0	0	0	0	0	0	0	2	0	2	1
su	0	0	0	0	0	0	0	0	0	0	0	0	0
número	0	0	0	0	0	0	1	0	0	1	1	0	1
y	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
pequeña	0	0	0	0	0	0	0	0	0	0	0	0	1
les	0	0	0	0	0	0	0	0	0	0	0	0	0
habla	0	0	0	0	0	0	0	0	0	1	0	0	1
como	0	0	1	0	0	0	0	0	0	0	0	0	0
si	0	0	0	0	0	0	0	0	0	0	0	0	0
todo	0	0	0	0	0	0	0	0	0	0	0	0	0
eso	0	0	0	0	0	0	0	0	0	abs	abs	0	0
fuera	0	0	0	0	0	0	0	0	0	0	0	0	0
normal	0	0	0	0	0	0	0	0	0	0	0	0	0
Sobre	0	0	0	0	0	0	0	0	0	0	0	0	0
el	0	0	0	0	0	0	0	0	0	0	0	0	0
dibujo	0	0	0	0	0	0	0	0	0	0	0	0	0
el	0	0	0	0	0	0	0	0	0	0	0	0	0
público	0	1	0	0	0	0	1	1	0	1	1	1	1
aplaude	0	0	0	0	0	0	0	0	0	0	0	0	0
busca	0	0	0	0	0	0	0	0	0	0	0	0	0
a	0	0	0	0	0	0	0	0	0	0	0	0	0
su	0	0	0	0	0	0	0	0	0	0	0	0	0
hermanito	0	2	0	0	1	0	0	0	2	0	0	1	0

Dictée	38 (m)	39 (m)	40 (f)	41 (f)	42 (m)	43 (f)	44 (f)	45 (f)	47 (m)	48 (m)	49 (f)	50 (f)	51 (m)
Su	0	0	0	0	0	0	0	0	0	0	0	0	0
entrenadora	0	0	0	0	0	0	0	0	0	0	0	0	0
inquieta	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
observa	1	1	0	0	0	0	0	0	0	0	0	0	0
con	0	0	0	0	0	0	0	0	0	0	0	0	0
mucha	0	0	0	0	0	0	0	0	0	0	0	0	0
atención	1	0	0	0	0	0	0	0	0	0	1	0	0
Yo	0	0	0	0	0	0	0	0	0	0	0	0	0
soy	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
gata	0	0	0	0	0	0	0	0	0	0	0	0	0
Después	1	0	1	1	0	0	0	0	0	1	1	0	0
del	0	0	0	0	0	0	0	0	0	0	0	0	0
accidente	0	0	0	0	1	0	0	0	0	0	0	0	0
vivo	0	0	0	0	0	0	0	0	0	0	0	0	1
aquí	1	1	0	1	0	0	0	0	0	0	1	1	0
Y	0	0	0	0	0	0	0	0	0	0	0	0	0
los	0	0	0	0	0	0	0	0	0	0	0	0	0
niños	0	0	0	0	0	0	0	0	0	0	0	0	0
de	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
casa	0	0	0	0	0	0	0	0	0	0	0	0	0
son	0	0	0	0	0	0	0	0	0	0	0	0	0
encantadores	0	0	0	0	0	0	0	0	0	1	0	0	0
me	0	0	0	0	0	0	0	0	0	0	0	0	0
acaricia	1	1	0	1	0	0	0	0	1	1	0	0	0
frecuentemente	0	0	0	0	0	0	0	0	0	1	0	0	0
Sus	0	0	0	0	0	0	0	0	0	0	0	0	0
caricias	1	1	0	1	0	0	0	0	1	0	0	0	0
y	0	0	0	0	0	0	0	0	0	0	0	0	0
sus	0	0	0	0	0	0	0	0	0	0	0	0	0
palabras	0	0	0	0	0	0	0	0	0	0	0	0	0
son	0	0	0	0	0	0	0	0	0	0	0	0	0
dulces	1	0	0	0	0	0	0	0	0	0	0	0	0
y	0	0	0	0	0	0	0	0	0	0	0	0	0
reconfortantes	0	0	0	0	0	0	0	0	0	0	0	0	0
Son	0	0	0	0	0	0	0	0	0	0	0	0	0
las	0	0	0	0	0	0	0	0	0	0	0	0	0
mejores	0	0	0	0	0	0	0	0	0	0	0	0	0
me	0	0	0	0	0	0	0	0	0	0	0	0	0
confía	1	0	1	1	0	0	0	0	0	0	1	1	1
sus	0	0	0	0	0	0	0	0	1	0	0	0	0
secretos	0	0	0	0	0	0	0	0	0	0	0	0	0
Sus	0	0	0	0	0	0	0	0	0	0	0	0	0
jornadas	0	0	0	0	0	0	0	0	0	0	0	0	0
en	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	2	0
escuela	0	0	0	0	0	0	0	0	0	0	0	0	0
no	0	0	0	0	0	0	0	0	0	0	0	0	0
parecen	0	0	0	0	0	0	0	0	0	0	0	0	0
ser	0	0	0	0	0	0	0	0	0	0	0	0	0
fáciles	1	0	1	1	1	0	0	0	0	1	1	0	1

Dictée	38 (m)	39 (m)	40 (f)	41 (f)	42 (m)	43 (f)	44 (f)	45 (f)	47 (m)	48 (m)	49 (f)	50 (f)	51 (m)
la	0	0	0	0	0	0	0	0	0	0	0	0	0
pieza	0	0	0	0	0	0	0	0	1	0	0	0	0
De	0	0	0	0	0	0	0	0	0	0	0	0	0
repente	0	0	1	1	1	0	1	0	0	1	1	1	0
los	0	0	0	0	0	0	0	0	0	0	0	0	0
personajes	0	0	0	0	0	0	0	0	0	0	0	0	0
comienzan	0	0	0	0	1	0	0	0	0	0	0	0	2
a	0	0	0	0	0	0	0	0	0	0	0	0	0
moverse	1	0	0	0	0	0	0	0	0	0	1	1	0
payasos	0	0	0	0	0	0	0	0	0	0	0	0	0
en	0	0	0	0	0	0	0	0	0	0	0	0	0
equilibrio	0	0	0	0	0	0	0	0	0	0	0	0	1
sobre	0	0	0	0	0	0	0	0	0	0	0	0	0
una	0	0	0	0	0	0	0	0	0	0	0	0	0
curiosa	0	0	0	0	0	0	0	0	0	0	0	0	0
bicicleta	1	0	0	0	0	0	0	0	0	0	1	0	0
salen	0	0	0	0	0	0	0	0	0	0	0	0	0
del	0	0	0	0	0	0	0	0	0	0	0	0	0
dibujo	0	0	0	0	0	0	0	0	0	0	0	0	0
y	0	0	0	0	0	0	0	0	0	0	0	0	0
avanzan	1	0	0	0	0	0	1	1	0	0	0	0	0
hacia	0	1	0	1	0	0	0	0	0	0	0	0	0
Ellos	0	0	0	0	0	0	0	0	0	0	0	0	0
hacen	0	0	0	0	0	0	0	0	0	0	0	0	0
su	0	0	0	0	0	0	0	0	0	0	0	0	0
número	1	0	0	1	0	0	0	0	0	0	1	1	0
y	0	0	0	0	0	0	0	0	0	0	0	0	0
la	0	0	0	0	0	0	0	0	0	0	0	0	0
pequeña	1	0	0	0	0	0	0	0	0	0	0	0	0
les	0	0	0	0	0	0	0	0	0	0	0	0	0
habla	0	0	0	0	0	0	0	1	0	0	0	0	0
como	0	0	0	0	0	0	0	0	0	0	0	0	0
si	0	0	0	0	0	0	0	0	0	0	0	0	0
todo	0	0	0	0	0	0	0	0	0	0	0	0	0
eso	0	abs	abs	abs	abs	abs	abs	abs	abs	abs	abs	abs	abs
fuera	0	0	0	0	0	0	0	0	0	0	0	0	0
normal	0	0	0	0	0	0	0	0	0	0	0	0	0
Sobre	0	0	0	0	0	0	0	0	0	0	0	0	0
el	0	0	0	0	0	0	0	0	0	0	0	0	0
dibujo	0	0	0	0	0	0	0	0	0	0	0	0	0
el	0	0	0	0	0	0	0	0	0	0	0	0	0
público	1	0	0	1	0	0	0	0	1	0	0	0	0
aplaude	0	0	0	0	0	0	0	0	0	0	0	0	0
busca	0	0	0	0	0	0	0	0	0	0	0	0	0
a	0	0	0	0	0	0	0	0	0	0	0	0	0
su	0	0	0	0	0	0	0	0	0	0	0	0	0
hermanito	0	0	0	0	0	0	0	0	0	1	0	0	0

APPENDICE C

TPOLOGIE DES ERREURS SELON CATACH (1980) DANS CHAQUE LANGUE

C.1 La typologie des erreurs lexicales pour le corpus français	188
C.2 La typologie des erreurs lexicales pour le corpus espagnol	210

Typologie des erreurs

Phonogramme	Morphogramme lexical	Homophone lexical et de discours	Idéogramme apostrophe et trait d'union	Phonétique	Lettres non fonctionnelles
acrobates (2) acrobattes (2)				acrobats (3) accrobats	
alongé (3) alonger alongés (5)					
amériquain (2) américains (4) amériquin amériquins				americain (2) américains americans amériquain amériquins américan (2) américans amériquents amériquien	
animeaux (3)					
annés (4) anné				anniés	
apelle (2) appel appèle				appelent appele	
aplaudient aplaudit (4) applaudis aplodient (2) aplodis aplodit applodie applodient					
arive (3)					

Phonogramme	Morphogramme lexical	Homophone lexical et de discours	Idéogramme apostrophe et trait d'union	Phonétique	Lettres non fonctionnelles
atelettes athelètes (6) atlètes (5) atlèthe (2) atlèthes (2) atlettes athètes (5) hathlète hatlète				acclèthes aclète aclettes aclèttes athlètes athletes (2) athelectes athèles athelte atheltes (2) athètes (3)	
atentivement atantivement				antentivement	
aujourd'hui aujourd'hui (3) haupourd'hui					
avançent (2) avencent (2) avence (3) avense					
bâtre batre (2)					
biscyclette	byciclette (5)			bicycle	
blaiser blaissé (2) blaissée (2) blaissér blésser bléssé bléssée				blésser bléssée	

Phonogramme	Morphogramme lexical	Homophone lexical et de discours	Idéogramme apostrophe et trait d'union	Phonétique	Lettres non fonctionnelles
bonheur (2) bonheure				bonnher	
bote botes				bots	
bouje					
bougers boujer					
		ça / sa (20)		ca	
cannadien				canadian	
cansère cansers quencères				concers cocerts cencers	
carresse (4) cares carresses (3)				carest	
çela celà					
châpita (2) châpiteau (5) chapitaux (3) chapito chapitot					
châpitaux (3) chapitaux (11) châpiteaux (3) chapitos				chapitiaux	

Phonogramme	Morphogramme lexical	Homophone lexical et de discours	Idéogramme apostrophe et trait d'union	Phonétique	Lettres non fonctionnelles
chaques (5)					
charment (4)					
charments (2)					
chate (6)					
				circle (4)	
				circle	
				clows (2)	
cloûnes					
clounes (3)					
cloune					
clouns (4)					
clounws					
clowne					
clownes					
colective (7)					
collorées					
combinaison				combineson	
combinèson					
commensses					
commençent					
compétiteures				compéteurs (2)	
				compititeur	
				competeurs (2)	

Phonogramme	Morphogramme lexical	Homophone lexical et de discours	Idéogramme apostrophe et trait d'union	Phonétique	Lettres non fonctionnelles
couleurs					
cout		coup / cou			
courage (2)				corrage	
courreur (3)				coueurs	
courreurs (4)					
coureure					
coureures					
courir (7)					
courir (3)					
				creation (2)	
				deurveurent	
dessin (3)					
dessine (2)					
d'eu (6)		d'eux / deux (3)			
deu					
deut					
	devans				
dirès				dax	
dirèt					
diret					
dousse				deuent	
dousses					

Phonogramme	Morphogramme lexical	Homophone lexical et de discours	Idéogramme apostrophe et trait d'union	Phonétique	Lettres non fonctionnelles
drole (3)					
durré					
durér					
				enfillé	
				enfiller (2)	
en suite (2)				ensuit (3)	
entrèneuse				entèneuse	
entraînneuse				entreneuse	
				iquilibres	
				équilibre (2)	
				équilibres	
é				ou	
exécution (5)				ésécutent	
exécutent				executent (5)	
				execute	
extraordinaire (2)				extrordinaire	
extraordinaire				extrordinaires	
extraordinaires (2)				extrodinaires	
				extrardinaires	
façile				facille	
facil (2)					
facine (7)				fasines	
fassine (2)					

Phonogramme	Morphogramme lexical	Homophone lexical et de discours	Idéogramme apostrophe et trait d'union	Phonétique	Lettres non fonctionnelles
fassinent façinent facinent (11) facinne facines					
festivale					
festivales				vestivals festifals	
flates flate (13)				falte	
fons (2)		fonds / font (3) fonds / fonts (2)			
foulle (4)					
jenou (3) jenoux					
heurs (2)					
hummain				humais	
				icit	
hiluminent illuminne ilumine (2) iluminent				ullimine élumine (3) éluminent (4)	
imaginer (2) imaginée imaginner					

Phonogramme	Morphogramme lexical	Homophone lexical et de discours	Idéogramme apostrophe et trait d'union	Phonétique	Lettres non fonctionnelles
imence (3) imense immenses immanses immence immences (10) immensses				immececs	
immobil (2) immobile (6) imobilent imobiles (4)				immobilles immobille (2) imobille	
inquiète (2) inquiette (4)				inquite inquiete (2)	
geongleuses jongueleuses jonguleuses					
journé journés					
			jusqu'a (7)	jousqua	
kilogramme (7) kilogrames (14)				killograme (2) killogrames (4) killogrammes (4)	
kilomètre				killomètres (4) kilometres	
l'acsident l'acsxident l'axident			laccident	lacidnt	

Phonogramme	Morphogramme lexical	Homophone lexical et de discours	Idéogramme apostrophe et trait d'union	Phonétique	Lettres non fonctionnelles
			l'école	l'école	
lenteure (2)					
lenteures					
lanteur (3)					
leucémi (3)			l'eucémis	lecimi	
leucémis				lecimies	
				leucimie (3)	
				leucimis	
				leusémie	
				leusémies	
				leusimie	
				leussimet	
				leussimies	
leucémis (5)			l'eucémis	leucimies (4)	
				leucimis (2)	
				leusemie	
				leusémie	
				leusémies	
				leusimies	
				leussimies	
				lucémé	
				lucimies	
l'observe (2)			l'observe	observe	
		lui m'aime / lui-même	lui même (31)	lui meme	
lute (3)				litte	
marraton (2)					maraton (43)
					maratons
				meilleurs	
				meilleures	

Phonogramme	Morphogramme lexical	Homophone lexical et de discours	Idéogramme apostrophe et trait d'union	Phonétique	Lettres non fonctionnelles
mieu (11) mieut					
		moi / mois			
nairveuse				me	
				numéro	
paraiceait paraissait (2) paressait (2) parrassait (3)				paraissait (4) paraissai (2) paréraisais parrassait parrésais	
parmis (14) parmie					
parôle parolles parolle parôles					
particulièrement particulièrements				particulirement pariculièrement	
personages (3) perssonages					
				paisais (2) pensait (2) pessait (7)	
pièce pièce					
		plaine / pleine (7)		plane	

Phonogramme	Morphogramme lexical	Homophone lexical et de discours	Idéogramme apostrophe et trait d'union	Phonétique	Lettres non fonctionnelles
	plusieur (17)			plusieus plusieux	
poucières poussières				pousières (7)	
				premier premieres (2)	
présante (3)				précente présent présenté (2) represente	
prèsque pressque				persque	
				bublique	
québecquoise (2) québéquoise (3)				québequoise québecoise (14) quebecoise (3) québecoises	
quelle				quleques	
réconfortente (4) réconfortentes				confortantes reconfortente reconfortantes reconfortantes (2) réconfortent (2)	
recors (2) reccords reccors					
recueillir (2) requueillir (3) requueillir requiere				recellir receir recuillir (2) receuillir (8)	

Phonogramme	Morphogramme lexical	Homophone lexical et de discours	Ideogramme apostrophe et trait d'union	Phonétique	Lettres non fonctionnelles
				réquellire recueillir receuilr recuellir (2) recueullir receullir	
réèles réaile				reel	
rééussie				réusit (2) réekusi	
scafandrié (4) scafendrié (5) scafendrier (6) scaphendrier				scafandier (9) scafondrier scaphandillé	
secrèt cecrets					
c'élanssent célence s'élansent sélenche (2) s'élence (2) sélencent s'élencent (6) s'ellance s'ellance s'ellanse			sélancent		
				sembleblent	
ci (3)					
sons (3)					
souétait				soitait (2)	

Phonogramme	Morphogramme lexical	Homophone lexical et de discours	Idéogramme apostrophe et trait d'union	Phonétique	Lettres non fonctionnelles
souaitait (3) souettai souhétait				shaitai souhtait souhaitait	
sout					
souvant					
spéctateurs				spectacteur	
talenttueuses tallentueuses talentueusent				talentctueuse talentueuse	
trapésiste trapésistes (12) trapeziste (2)				trapésistes (3) trapésites trapésites trapéxistes trapézise trapézistes (2) trapézizites trapisiste	
vincu (6) vincue vinquu				fincu vaicu vecu vinque vencut viencut	
		vers / ver (2) vers / vert vers / verre			
		vis / vie (11)			
				vrairement fraiment	
				ye	

Typologie des erreurs

Phonogramme	Morphogramme lexical	Homophone de discours	Idéogramme	Phonétique	Lettres non fonctionnelles
acancia (2)				acarecia	
acaricia					
acarisia (9)					
				asidente	
				axidente (4)	
acrovatas				acrobatos	
acrobatas (25)					
hålgunos					
ålgunos					
americånos (2)					
américanos (2)					
animo (11)				aninó	
animo					
			anos		
aqui (14)					
a qui (3)					
atencion (7)					
				atlelas	
				attletas	
				aclelas (2)	
abanzan (2)				abansa	
abansa				avancen	
abansan (2)					
avansan (7)					
avancan					

Phonogramme	Morphogramme lexical	Homophone de discours	Idéogramme	Phonétique	Lettres non fonctionnelles
aller (3)					
azúles					
vajo (5)					
bisicleta (10)					
vota					
vuenos (2)					
				busca	
buso (27)					
				cado	
canadiences (12)					
canceres (14)				cansers	
cánseres (2)				cancóres	
canseres (6)					
caricias				caricis	
carisias (9)					
carpaz					
caci					
casi (2)					
círcos					
sírcos					
colectiba					
colocarón (3)					
colocáron					

Phonogramme	Morphogramme lexical	Homophone de discours	Idéogramme	Phonétique	Lettres non fonctionnelles
coloridos (3)	colòridos				
comiensan (3)	comienzan				
comó					
confia (2)				confica	
confia (17)				confieza	
				contras	
creacion (11)					
				a	
desir (2)					
de lante					
deseava (2)				deciaba	
deceaba (3)				desaba	
despues (22)					
días (14)					
dibúja					
divujo (2)					
ladicha					

Phonogramme	Morphogramme lexicat	Homophone de discours	Idéogramme	Phonétique	Lettres non fonctionnelles
dulses (4)					
el					
el					
				el (5)	
				de	
encantadores					
en segida					
ensegida					
				entradora	
				equilvrio	
				equilibrio	
				ecuela	
expectadores				expertadores (2)	
estabá					
estava					
éstos (3)					
extrahordinarios				entraordinarios	
				extraudinanos (2)	
				extraordinario	
fasiles (2)	façiles				
faciles (24)					

Phonogramme	Morphogramme lexical	Homophone de discours	Idéogramme	Phonétique	Lettres non fonctionnelles
facinan (27)					
fasinan (11)					
festibal (4)					
festibales (2)				festivalles	
festibalez					
				confrecuencia	
frecuente mente				frecuentemete frecue mente	
fué (6)					
				fura	
habia (14)				haba	abia (2)
ahbia					
ahbla				desabla	abla (4)
				hacia	
hasen					acen (2)
a sia					acia (3)
asia (2)					
hasia (2)					
				tasta	asta (8)
hai					
ay					
ahy					
ai					

Phonogramme	Morphogramme lexical	Homophone de discours	Idéogramme	Phonétique	Lettres non fonctionnelles
				hermano ehermanito	ermanito (3)
hirio (20)					irio (5) irio umanos
imaginada				inmaginada	
imagina da				imaguinada	
				imensas imensans	
inmobiles (12)				imobiles	
inmoviles (14)				imbobiles enmobiles	
isquierda (7)				esquierda	
				cornadas	
jovenes (2)				jovenen	
jovenes (22)					
kilográmos				kilos	
quilogramos					
kilometros (15)					
lanzan				tranzan	
lansan (7)					
leusemia (19)				leuciamia leuusemia leucimia	

Phonogramme	Morphogramme lexical	Homophone de discours	Ideogramme	Phonétique	Lettres non fonctionnelles
leusemias (19)				leuciamia	
leusemia (5)				leusemiia	
malavaristas (12)				marabaristas (2)	
				maravaristas (2)	
maraton (19)				maratos	
		más / mas (8)			
				mejor	
my					
mismo					
moverce					
moberse (5)					
				muchonos	
mueba (2)					
multitud					
mui					
nerviosa					
		ninos (3)			
				nomal	
numero (13)					

Phonogramme	Morphogramme lexical	Homophone de discours	Idéogramme	Phonétique	Lettres non fonctionnelles
osserba (9)				por	
presen (6)					
preseria (2)					
pareceria (15)					
preseria (6)					
presia					
parecia (17)					
participan					
participân					
partisipân (3)					
partisiparon (2)					
participarôn (3)					
particular mente(2)				particurmente particularmete	
pallasos (10)				payos	
permanesen (2)					
personages (2)					
			pequena (4)		
pesava (3)					
piesa					
piesa (5)					
pièza					

Phonogramme	Morphogramme lexical	Homophone de discours	Iséogramme	Phonétique	Lettres non fonctionnelles
planisie (3)					
póco					
presentá					
precenta					
publico (19)					
público					
				puso	
que vequence (2)				cabequense	
quebequence (12)				que dequence	
quevequence				cebequence	
quévequence				que nequence	
quevequense (4)				que quence	
				quebecence	
				quebequincie	
				quedequense	
				quedequerce	
				quebequences	
				cedan	
quien					
quiên					
real mente (4)				realmen	
				recomfortes	
				reconfortantes	
				reconfortables	
				reconfotrantes	

Phonogramme	Morphogramme lexical	Homophone de discours	Idéogramme	Phonétique	Lettres non fonctionnelles
				roconfortantes	
records (14)				recrords	
recors (3)					
récors (4)					
recoord					
record					
répente (2)				dérepente	
				derrepente (15)	
cecretos					
soi					
				la	
				latentosas	
				lalenlosas	
trage				traja	
trapesistas (14)					
trapezistas					
vários (2)					
barios (3)					
be					
venser					
benser (2)					
bencer (2)					
benciö					
bensio					
vencio					
vencio (13)					
vensiö					
vensio					
bibo				digo	
llo (3)					

APPENDICE D

TABLEAUX DES RÉSULTATS

D.1	Tableau comparatif du nombre de mots écrits en français et en espagnol	212
D.2	Tableau comparatif du nombre de mots bien écrits en français et en espagnol	213
D.3	Tableau comparatif du nombre d'erreurs grammaticales en français et en espagnol	214
D.4	Tableau comparatif du nombre d'erreurs lexicales en français et en espagnol	215

D.1 Tableau comparatif du nombre de mots écrits par élève en français et en espagnol

francophone		hispanophone	
élève	mots écrits	élève	mots écrits
1	315	2	305
2	244	3	307
3	315	6	309
4	218	8	309
5	315	9	310
6	315	10	308
7	315	12	310
8	315	13	304
10	314	15	299
12	315	16	308
13	313	17	310
14	218	18	310
15	313	19	309
16	315	20	310
17	309	21	308
18	312	22	307
19	315	23	310
20	315	24	308
21	313	25	309
22	315	26	303
23	315	29	310
24	315	30	310
25	315	32	311
26	315	33	310
27	315	35	312
28	314	36	312
29	313	38	311
30	315	39	309
31	315	40	311
32	315	41	311
33	315	42	311
34	315	43	311
35	315	44	311
36	314	45	310
37	315	47	311
38	315	48	311
39	315	49	311
40	315	50	311
41	315	51	308
42	315	54	311
		55	311
		56	311
moyenne	308		309
total	12315		12988

D.2 Tableau comparatif du nombre de mots bien écrits par élève en français et en espagnol

francophone			hispanophone		
élève	mots bien écrits	%	élève	mots bien écrits	%
1	304	96.51	2	263	86.23
2	155	63.52	3	294	95.77
3	268	85.08	6	291	94.17
4	148	67.89	8	285	92.23
5	239	75.87	9	305	98.39
6	235	74.60	10	297	96.43
7	230	73.02	12	279	90.00
8	231	73.33	13	275	90.46
10	245	78.03	15	252	84.28
12	303	96.19	16	271	87.99
13	231	78.03	17	270	87.10
14	186	85.32	18	303	97.74
15	254	81.15	19	298	96.44
16	222	70.48	20	296	95.48
17	196	63.43	21	275	89.29
18	216	69.23	22	274	89.25
19	281	89.21	23	294	94.84
20	238	75.56	24	288	93.51
21	216	69.01	25	284	91.91
22	258	81.90	26	266	87.79
23	276	87.62	29	295	95.16
24	272	86.35	30	292	94.19
25	297	94.29	32	246	79.10
26	270	85.71	33	269	86.77
27	295	93.65	35	269	86.22
28	253	80.57	36	266	85.26
29	254	81.15	38	261	83.92
30	290	92.06	39	288	93.20
31	272	86.35	40	295	94.86
32	298	94.60	41	272	87.46
33	293	93.02	42	295	94.86
34	307	97.46	43	304	97.75
35	303	96.19	44	294	94.53
36	247	78.66	45	292	94.19
37	301	95.56	47	278	89.39
38	306	97.14	48	278	89.39
39	304	96.51	49	278	89.39
40	286	90.79	50	287	92.28
41	282	89.52	51	283	91.88
42	278	88.25	54	261	83.92
			55	271	87.14
			56	250	80.39
moyenne	258.50	83.82		280.57	90.73

D.3 Tableau comparatif du nombre d'erreurs grammaticales par élève en français et en espagnol

francophone			hispanophone		
élève	erreurs grammaticales	%	élève	erreurs grammaticales	%
1	5	1.58	2	11	3.62
2	52	20.50	3	2	0.66
3	22	6.67	6	5	1.62
4	51	23.39	8	5	1.94
5	54	17.14	9	1	0.32
6	46	14.61	10	2	0.64
7	51	15.87	12	5	1.62
8	41	12.70	13	6	1.98
10	45	14.01	15	8	3.01
12	9	2.54	16	5	1.62
13	41	12.78	17	7	2.26
14	16	6.42	18	3	0.97
15	38	12.14	19	5	1.61
16	58	17.78	20	3	0.97
17	67	20.72	21	7	2.59
18	64	20.52	22	5	1.96
19	14	3.53	23	5	1.61
20	41	12.70	24	2	0.64
21	63	19.17	25	4	1.29
22	33	10.16	26	7	2.31
23	15	4.12	29	4	1.30
24	17	5.08	30	5	1.62
25	8	2.22	32	9	3.22
26	27	8.25	33	7	0.97
27	14	4.13	35	6	2.56
28	36	10.83	36	9	2.88
29	25	7.35	38	8	3.22
30	17	5.39	39	3	0.97
31	29	8.89	40	3	0.96
32	12	3.81	41	7	2.57
33	8	1.90	42	2	0.64
34	5	1.27	43	2	0.64
35	7	2.22	44	4	1.60
36	36	11.15	45	6	1.94
37	8	2.22	47	7	2.25
38	8	2.54	48	7	2.58
39	5	1.59	49	7	2.57
40	7	1.90	50	6	1.93
41	19	6.03	51	7	2.87
42	19	5.71	54	7	2.56
			55	6	2.25
			56	5	2.57
moyenne	28.33	9.04		5.36	1.86

D.4 Tableau comparatif du nombre d'erreurs lexicales par élève en français et en espagnol

francophone			hispanophone		
élève	erreurs lexicales	%	élève	erreurs lexicales	%
1	6	1.90	2	33	10.82
2	56	23.77	3	11	3.58
3	26	8.57	6	13	4.21
4	27	12.39	8	19	5.83
5	30	9.52	9	4	1.29
6	43	13.65	10	9	2.92
7	37	12.06	12	27	8.71
8	52	16.83	13	23	7.57
10	29	9.55	15	39	12.71
12	3	1.27	16	32	10.39
13	54	17.57	17	34	10.97
14	18	9.18	18	4	1.29
15	26	8.31	19	7	2.26
16	48	15.88	20	11	3.55
17	64	21.69	21	26	8.12
18	42	14.11	22	29	9.12
19	22	7.61	23	12	3.87
20	41	13.34	24	18	5.84
21	53	17.89	25	21	6.80
22	28	9.21	26	31	10.23
23	27	9.20	29	11	3.55
24	27	8.89	30	13	4.19
25	10	3.49	32	56	17.68
26	18	6.03	33	34	10.32
27	7	2.54	35	37	11.22
28	32	10.83	36	38	12.18
29	36	12.14	38	42	12.92
30	9	2.86	39	18	5.83
31	15	5.08	40	13	4.18
32	6	1.91	41	33	10.29
33	14	5.08	42	14	4.50
34	3	1.27	43	5	1.61
35	5	1.59	44	13	3.86
36	38	12.42	45	13	4.19
37	6	2.22	47	27	8.68
38	1	0.32	48	26	8.04
39	7	2.22	49	27	8.36
40	22	7.30	50	18	5.79
41	16	5.07	51	18	5.84
42	20	6.66	54	46	14.46
			55	35	10.61
			56	55	17.04
moyenne	25.60	8.79		23.69	7.51

APPENDICE E

LISTE DES MOTS AVEC % D'ERREURS EN FRANÇAIS ET EN ESPAGNOL

E.1	Liste de mots avec le pourcentage d'erreurs pour le corpus français	217
E.2	Liste de mots avec le pourcentage d'erreurs pour le corpus français	222

mots dictés	fréq. Fortier	% d'échec	catégorie
au	34	0%	af
ce	93	0%	af
comme	64	0%	af
du	36	0%	af
leur	36	0%	af
me	84	0%	af
minutes	21	0%	af
montés	98	0%	af
par	74	0%	af
quels	85	0%	af
restent	40	0%	af
équipe	24	0%	af
a	578	0%	f
à	1103	0%	f
avait	578	0%	f
beaucoup	368	0%	f
bien	419	0%	f
bleus	246	0%	f
c'est	622	0%	f
de	2434	0%	f
des	373	0%	f
dire	173	0%	f
elle	149	0%	f
en	397	0%	f
enfants	1001	0%	f
était	529	0%	f
été	221	0%	f
font	474	0%	f
frère	149	0%	f
grandes	745	0%	f
il	261	0%	f
ils	261	0%	f
je	125	0%	f
jeunes	196	0%	f
jours	263	0%	f
la	1012	0%	f
le	658	0%	f
les	293	0%	f
l'un	1202	0%	f
maison	1411	0%	f
on	1044	0%	f
ont	578	0%	f
parle	650	0%	f
pas	574	0%	f
petit	920	0%	f
petite	920	0%	f
pour	223	0%	f
que	184	0%	f
qu'il	184	0%	f

mots dictés	fréq. Fortier	% d'échec	catégorie
s'est	533	0%	f
son	236	0%	f
sortent	119	0%	f
suis	529	0%	f
sur	123	0%	f
tout	230	0%	f
très	118	0%	f
un	1202	0%	f
une	1202	0%	f
va	372	0%	f
vous	159	0%	f
ces	3	0%	pf
contre	16	0%	pf
depuis	4	0%	pf
faut	3	0%	pf
gauche	14	0%	pf
hier	19	0%	pf
normal	17	0%	pf
participé	17	0%	pf
participent	17	0%	pf
ses	20	0%	pf
tous	5	0%	pf
demeurent	19	0%	pf
confie	4	0%	pf
et	117	0,50%	f
y	21	0,80%	af
sous	40	1,25%	af
souvent	99	1,25%	af
sont	529	1,25%	f
semblent	10	1,25%	pf
bouge	77	2,50%	af
colorés	34	2,50%	af
dessin	50	2,50%	af
ici	42	2,50%	af
ne	82	2,50%	af
nerveuse	68	2,50%	af
couleurs	806	2,50%	f
dans	157	2,50%	f
moi	624	2,50%	f
quelques	210	2,50%	f
devant	10	2,50%	pf
festival	6	2,50%	pf
numéro	16	2,50%	pf
public	10	2,50%	pf
talentueuses	2	2,50%	pf
bouger	77	5%	af
canadiens	70	5%	af
cela	26	5%	af
cirque	8	5%	pf

mots dictés	fréq. Fortier	% d'échec	catégorie
cirques	8	5%	pf
commencent	6	5%	pf
coup	86	5%	af
création	20	5%	pf
dessine	15	5%	pf
duré	4	5%	pf
foule	28	5%	af
heures	132	5%	f
humains	388	5%	f
journées	56	5%	af
kilomètres	7	5%	pf
l'école	742	5%	f
meilleures	16	5%	pf
pièce	91	5%	af
vraiment	10	5%	pf
spectateurs	4	5,00%	pf
animaux	700	7,50%	f
arrivent	198	7,50%	f
attentivement	1	7,50%	pf
battre	43	7,50%	af
botte	101	7,50%	f
combinaison	7	7,50%	pf
douces	562	7,50%	f
drôle	63	7,50%	af
enfilé	23	7,50%	af
festivals	6	7,50%	pf
jongleuses	1	7,50%	pf
premières	52	7,50%	af
presque	20	7,50%	pf
si	49	7,50%	af
réels	14	7,50%	pf
bonheur	347	10%	f
dirait	173	10%	f
entraîneuse	1	10%	pf
faciles	84	10%	af
genou	53	10%	af
imaginée	22	10%	af
l'observe	11	10%	pf
lutte	12	10%	pf
paroles	77	10%	af
particulièrement	1	10%	pf
records	14	10%	pf
réussi	32	10%	af
vers	53	10%	af
équilibre	14	10%	pf
allongés	62	12,50%	af
aujourd'hui	34	12,50%	af
chaque	17	12,50%	pf
courage	33	12,50%	af

mots dictés	fréq. Fortier	% d'échec	catégorie
ensuite	15	12,50%	pf
l'accident	180	12,50%	f
personnages	51	13%	af
pesait	297	13,75%	f
années	84	15%	af
appelle	107	15%	f
avancent	57	15%	af
cancers	94	15%	af
charmants	25	15%	af
chatte	546	15%	f
compétiteurs	1	15%	pf
lenteur	76	15%	af
secrets	32	15%	af
plusieurs	118	15,80%	f
bicyclette	227	17,50%	f
clowns	17	17,50%	pf
collective	4	17,50%	pf
jeu	251	18%	f
plaine	21	18%	af
acrobates	5	20%	pf
jusqu'à	6	20%	pf
présente	7	20%	pf
caresses	1	22,50%	pf
extraordinaires	46	22,50%	af
inquiète	11	22,50%	pf
poussières	31	22,50%	af
coureurs	43	25%	af
courir	565	25%	f
ça	5	26,25%	pf
blessee	72	27,50%	af
d'eux	44	27,50%	af
souhaitait	74	27,50%	af
vis	279	27,50%	f
chapiteau	*	30%	pf
fonds	*	30%	pf
mieux	7	30%	pf
réconfortantes	*	30%	pf
applaudit	7	32%	pf
exécutent	47	32,50%	af
illuminent	22	32,50%	af
kilogrammes	20	36,75%	pf
flatte	4	37,50%	pf
parmi	5	37,50%	pf
leucémie	4	40%	pf
vaincu	11	40%	pf
immobiles	7	42,50%	pf
paraissait	28	42,50%	af
s'élancent	*	45%	pf
américains	34	47,50%	af

mots dictés	fréq. Fortier	% d'échec	catégorie
athlètes	50	47,50%	af
chapiteaux	*	47,50%	pf
leucémies	4	47,50%	pf
immenses	67	50%	af
marathon	15	57,50%	pf
québécoise	16	60,00%	pf
fascinent	5	62,50%	pf
recueillir	1	65%	pf
trapézistes	*	65%	pf
scaphandrier	*	67,50%	pf
lui-même	2	82,50%	pf

mots dictés	fréq. Juilland	% d'échec	catégorie
a	10851	0%	f
animales	22	0%	af
aplaude	6	0%	pf
carpa	*	0%	pf
casa	477	0%	f
circo	*	0%	pf
colores	57	0%	af
competidores	*	0%	pf
con	4667	0%	f
corredores	4	0%	pf
correr	21	0%	af
curiosa	15	0%	pf
de	29804	0%	f
del	5340	0%	f
duró	9	0%	pf
el	13339	0%	f
ella	523	0%	f
ellas	160	0%	f
ellos	255	0%	f
entre	703	0%	f
equipo	*	0%	af
es	4731	0%	f
eso	715	0%	f
fondos	7	0%	pf
gata	13	0%	pf
grandes	8	0%	pf
horas	161	0%	f
hoy	314	0%	f
iluminan	1	0%	pf
imaginada	1	0%	pf
inquieta	3	0%	pf
juego	38	0%	af
la	18237	0%	f
las	5360	0%	f
lentitud	8	0%	pf
les	174	0%	f
lo	2231	0%	f
logró	9	0%	pf
los	6343	0%	f
lucha	50	0%	af
me	768	0%	f
mejor	30	0%	af
mejores	2	0%	pf
minutos	43	0%	af
mucha	52	0%	af
mucho	254	0%	f
multitud	17	0%	pf
nerviosa	13	0%	pf
no	6900	0%	f
palabras	188	0%	f
por	4700	0%	f

mots dictés	fréq. Juilland	% d'échec	catégorie
primeras	2	0%	pf
que	9274	0%	f
qué	*	0%	af
reales	24	0%	af
recostados	*	0%	af
reunir	13	0%	pf
rodilla	1	0%	pf
salen	16	0%	pf
se	220	0%	f
ser	750	0%	f
si	1327	0%	f
sobre	854	0%	f
son	764	0%	f
su	3802	0%	f
te	432	0%	f
tiene	640	0%	f
todo	419	0%	f
todos	546	0%	f
un	5033	0%	f
una	3708	0%	f
unos	21	0%	af
ve	23	0%	af
y	14712	0%	f
sus	1739	0,60%	f
muchos	21	1,20%	af
muy	787	1,20%	f
en	12922	1,50%	f
dibujo	7	1,60%	pf
años	435	2,38%	f
azules	20	2,38%	pf
bota	*	2,38%	f
busca	26	2,38%	af
cada	344	2,38%	f
carpas	*	2,38%	pf
colectiva	12	2,38%	pf
como	2265	2,38%	f
contra	182	2,38%	f
delante	51	2,38%	af
dibuja	2	2,38%	pf
dicha	11	2,38%	pf
él	1023	2,38%	f
encantadores	1	2,38%	pf
entrenadora	*	2,38%	pf
escuela	69	2,38%	af
frecuencia	48	2,38%	af
fuera	60	2,38%	af
hace	435	2,38%	f
humanos	17	2,38%	pf
jornadas	10	2,38%	pf
mí	414	2,38%	f
mismo	467	2,38%	f

mots dictés	fréq. Juilland	% d'échec	catégorie
normal	17	2,38%	pf
para	2684	2,38%	f
poco	171	2,38%	f
puesto	40	2,38%	af
quedan	5	2,38%	pf
secretos	17	2,38%	pf
soy	186	2,38%	f
hacen	77	3,17%	af
kilogramos	*	3,60%	pf
pesaba	1	3,60%	pf
yo	1348	3,60%	f
permanecen	3	4,74%	pf
atletas	2	4,76%	pf
buenos	3	4,76%	pf
circos	*	4,76%	pf
decir	482	4,76%	f
enseguida	14	4,76%	pf
equilibrio	12	4,76%	pf
estaba	225	4,76%	f
inmensas	9	4,76%	pf
mueva	1	4,76%	pf
personajes	41	4,76%	af
presenta	29	4,76%	af
quien	259	4,76%	f
talentosas	*	4,76%	pf
traje	36	4,76%	af
vivo	15	4,76%	pf
algunos	15	5,25%	pf
bajo	17	5,95%	pf
casi	270	6,14%	f
ayer	99	7,14%	af
espectadores	3	7,14%	pf
estos	344	7,14%	f
frecuentemente	13	7,14%	pf
parecen	39	7,14%	af
planicie	*	7,14%	af
niños	56	8,33%	af
americanos	7	9,52%	pf
colocaron	4	9,52%	pf
coloridos	*	9,52%	af
comienzan	1	9,52%	pf
dulces	9	9,52%	pf
festival	*	9,52%	pf
festivales	*	9,52%	pf
hay	764	9,52%	f
particularmente	17	9,52%	pf
pequeña	34	9,52%	af
accidente	6	11,90%	pf
extraordinarios	1	11,90%	pf
hermanito	*	11,90%	f
participan	1	11,90%	pf

mots dictés	fréq. Juilland	% d'échec	catégorie
participaron	4	11,90%	pf
realmente	20	11,90%	pf
varios	98	11,90%	af
vencer	14	11,90%	pf
payasos	*	13,10%	pf
fue	495	14,29%	f
habla	84	14,29%	af
moverse	4	14,29%	pf
reconfortantes	*	14,29%	pf
atención	65	16,67%	af
deseaba	9	16,67%	pf
pieza	25	16,67%	af
hacia	203	19,05%	f
izquierda	7	19,05%	pf
más	2363	19,05%	f
hasta	613	21,43%	f
lanzan	3	21,43%	pf
observa	5	21,43%	pf
bicicleta	*	23,81%	f
maratón	*	23,81%	pf
caricias	6	26,19%	pf
creación	42	26,19%	af
canadienses	*	28,57%	af
acaricia	*	30,95%	pf
ánimo	37	30,95%	af
número	145	30,95%	f
avanzan	9	33,33%	pf
días	302	33,33%	f
kilómetros	26	35,71%	af
trapecistas	*	35,71%	pf
aquí	602	38,10%	f
malabaristas	*	38,10%	pf
había	234	42,86%	f
parecía	70	42,86%	af
repente	6	42,86%	pf
venció	1	42,86%	pf
público	69	47,62%	af
confía	5	50%	pf
Después	453	52,38%	f
leucemia	*	52,38%	pf
parecería	3	54,76%	pf
cánceres	*	57,14%	af
récords	*	57,14%	pf
jóvenes	15	57,14%	pf
hirió	1	59,52%	pf
leucemias	*	61,90%	pf
acróbatas	*	64,29%	pf
buzo	*	64,29%	pf
fáciles	8	64,29%	pf
inmóviles	7	69,05%	pf
quebequense	*	73,81%	pf
fascinan	*	90,48%	pf

Bibliographie

Arrivé, Michel. 1993. *Réformer l'orthographe ?* Paris : Presses universitaires de France, 233p.

Baudot, Jean. 1992. *Fréquences d'utilisation des mots en français écrit contemporain*. Montréal : Les presses de l'Université de Montréal, 430p.

Blanche-Benveniste, Claire, et Chervel, André. 1969. *L'orthographe*. Paris : F. Maspero, 260p.

Bluau, G, et al. 1990. *Contre la réforme de l'orthographe : va-t-on se laisser mordre la langue ?* Paris : Autre temps, 111p.

Bois, Claude. 1968. *Mémento orthographique*. Lausanne : Payot, 400p.

Bonin, P. 2005. « Comment accède-t-on à un mot en production verbale écrite ? » *Psychologie Française*, vol. 50, n° 3, septembre 2005, pp. 323-338.

Braine, Alain. 1993. *Homophones : mots et jumeaux*. Montréal : Éditions Hurtubise, 85p.

Brissaud, Catherine, et Bessonat, Daniel. 2001. *L'orthographe au collège : pour une autre approche*. Grenoble : CRDP de l'Académie de Grenoble, 255p.

Burney, Pierre. 1967. *L'orthographe*. Paris : Presses Universitaires de France, coll. Que sais-je ? 126p.

Caffot, Philippe. 2001. *Défense et illustration de l'orthographe française*. Paris : G. de Bouillon, 56p.

Caro, O. 1988. *L'accentuation*. Aix-en-Provence : FF. Martorana, 145p.

Catach, Nina. 1978. *L'orthographe*. Paris : Presses Universitaires de France, coll. Que sais-je ? 127p.

Catach, Nina, & al. 1980. *L'enseignement de l'orthographe*. Paris : Éditions Fernand Nathan, 95p.

Catach, Nina. 1986. *L'orthographe française : Traité théorique et pratique*. Paris : Éditions Fernand Nathan, 334p.

- Catach, Nina. 1989. *Les délires de l'orthographe*. Paris : Plon, 349p.
- Catach, Nina, Honvault, Renée. 2001. *Histoire de l'orthographe française* : Édition posthume. Paris : Champion, 425p.
- Chartrand, Suzanne-G, et al. 1999. *Grammaire pédagogique du français d'aujourd'hui*. Montréal : Les Éditions Graficor, 397p.
- Cerquiglini, Bernard. 1995. *L'accent du souvenir*. Paris : Éditions de Minuit, 165p.
- Chervel, André, et Manesse, Danièle. 1989. *La dictée : Les Français et l'orthographe*. France : Calman-Lévy, 287p.
- Contant, Chantal. 2004. « Connaitre la nouvelle orthographe. » *Correspondance*, novembre 2004, vol. 10, n° 2.
- Da Silva, Monique. 1998. *La grammaire espagnole*. Paris : Hatier, Coll. Bescherelle, 352p.
- De Villers, Marie-Éva. 2003. *La nouvelle grammaire en tableaux*. Montréal : Québec Amérique, 313p.
- Denina, Carlo. 2003. *La clef des langues, ou, Observations sur l'origine et la formation des principales langues qu'on parle et qu'on écrit en Europe*. Genève : Slatkine Reprints.
- Ducard, Dominique, et al. 1995. *L'orthographe en trois dimensions*. Paris : Éditions Nathan, 303p.
- Farid, Georges. 1996. *Notions grammaticales de base : Les homonymes*. Montréal : Georges Farid, 227p.
- Fortier, Gilles. 1993. *Le vocabulaire des adolescents et des adolescentes du Québec*. Montréal : Les Éditions Logiques, 356p.
- France, Commission ministérielle d'études orthographiques. 1965. *Rapport général sur les modalités d'une simplification éventuelle de l'orthographe française*. Paris : M. Didier, coll. Linguistique appliquée, 137p.
- Gey, Michel. 1987. *Didactique de l'orthographe française*. Paris : Éditions Nathan, 175p.

Giasson, Jocelyne et Thériault, Jacqueline. 1983. *Apprentissage et enseignement de la lecture*. Montréal : Ville-Marie, pp 224-245.

Goosse, André. 1991. *La nouvelle orthographe : exposé et commentaires*. Paris : Duculot, 136p.

Grevisse, Maurice. 1948. *Code de l'orthographe française*. Paris : Les Éditions Baude, 248p.

Gruaz, Claude. 1999. « Une si longue histoire ». *L'orthographe ? C'est pas ma faute !* pp.58-61.

Gruaz, Claude, et al. 2001. *Variations sur l'orthographe et les systèmes d'écriture. Mélanges en hommage à Nina Catach*. Paris : Honoré Champion Éditeur, 401p.

Guion, Jean. 1974. *L'institution orthographe. À quoi sert l'orthographe ? À quoi sert son enseignement ?* Lyon : Éditions du Centurion, 203p.

Guyon, Odile, et al. 2003. *L'orthographe, une construction cognitive et sociale*. Toulouse : Presses Universitaires de Mirail, 137p.

Jacques, David. 1996. *L'apprentissage de l'écriture de l'école au collège*. Paris : Presses Universitaires de France, 215p.

Jacques, David. 2003. « Les procédures orthographiques dans les productions écrites des jeunes enfants ». *Revue des sciences de l'éducation*, vol. 29, n° 1, pp. 137-158.

Jaffré, Jean-Pierre. 1992. *Didactiques de l'orthographe*. Paris : Hachette, 175p.

Jaffré, Jean-Pierre, et Fayol, Michel. 1997. *Orthographe : Des systèmes aux usages*. France : Éditions Flammarion, Coll. Dominos, 127p.

Jaffré, Jean-Pierre. 1999. « Changez l'orthographe et ils écriront mieux ». *L'orthographe ? C'est pas ma faute !* pp.150-154.

Jaffré, Jean-Pierre. 2003. « La linguistique et la lecture-écriture : de la conscience phonologique à la variable orthographique ». *Revue des sciences de l'éducation*, vol. 29, n°1, pp. 37-49.

Juilland, Alphonse. 1964. *Frequency dictionary of spanish word*. London : Mouton & Co, 500p.

Klinkenberg, J.-M. 1984. « Images sociales de l'orthographe. Enquête sociolinguistique auprès d'un public scolaire ». *Au bonheur des mots: mélanges en l'honneur de Gérard Antoine* (p. 555-571). Nancy : Presses universitaires de Nancy.

Lalande, J.-R. 1988. *Élaboration d'un modèle théorique de l'apprentissage de l'orthographe lexicale au primaire*. Montréal : Les Publications de la Faculté des sciences de l'éducation de l'Université de Montréal, 358p.

Manesse, Danièle, Cogis, Danièle. 2007. *Orthographe à qui la faute ?* France, ESF Éditeur, 252p.

Martel, Réginald. 1991. « Réforme de l'orthographe : des réactions chez les écrivains d'ici ». *La Presse*, 13 janvier, p.C3.

Masson, Michel. 1991. *L'orthographe : guide pratique de la réforme*. Paris, Éditions du Seuil, 184p.

McNicoll, Louiselle, Roy, G-Raymond. 1985. *Les homophones : problèmes et solutions*. Sherbrooke : Éditions Naaman, 94p.

Millet, Agnes. 1990. *Orthographe mon amour*. Grenoble : Presses Universitaires de Grenoble, 240p.

Nadeau, Marie. 1995. « La réussite des accords grammaticaux au primaire, comment relever le défi ? ». *Pour un nouvel enseignement de la grammaire*. Montréal : Éditions logiques, pp 275-310.

Nadeau, Marie et Fisher, Carole. 2006. *La grammaire nouvelle : La comprendre et l'enseigner*. Montréal : Gaëtan Morin, 239p.

Noël, Colette, et Gervais, Flore. 1987. « La perception de l'orthographe chez de futurs enseignants : [résultats d'une recherche] ». *Interface*, 8, n°1, janv.-févr. 1987, pp 22-26.

Préfontaine, Robert-R. 1995. *Échelle de réussite orthographique*. Laval : Beauchemin, 166p.

- Pombo, Elena Llamas. 1999. « La letra con sangre. » *L'orthographe ? C'est pas ma faute !* pp.85-89.
- Portebois, Yannick. 1998. *Les saisons de la langue*. Paris : Honoré Champion, 498p.
- Pottier, Bernard, et al. 2000. *Grammaire explicative de l'espagnol*. Paris : Éditions Nathan, 318p.
- Réseau pour la nouvelle orthographe. 2004. *Vadémécum de l'orthographe recommandée : le millepatte sur un nénuphar*. Renouvo, 38p.
- Rey-Debove, Josette. 1991. *La réforme de l'orthographe au banc d'essai du Robert*. Paris : Le Robert, 78p.
- Roy, G.-R, et Lafontaine, L. 1992. *Étude de la maîtrise du français écrit à l'université*. Éditions du CRP, Université de Sherbrooke, 126p.
- Roy, Mario. 1991. « LN A U D BB ». *La Presse*, 12 janvier, p.D3.
- Simard, Claude. 1995 « L'orthographe d'usage chez les étudiants des ordres postsecondaires ». *Revue des sciences de l'éducation*, vol. 21, no 1, pp. 145-165.
- Tardif, Jacques. 1992. *Pour un enseignement stratégique. L'apport de la psychologie cognitive*. Montréal : Les Éditions Logiques, 474p.
- Thimonnier, René. 1967. *Le système graphique du français*. Paris : Plon, 387p.
- Traimond, Bernard. 2001. *Une cause nationale : l'orthographe française*. Paris : Presses Universitaires de France, 287p.