

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LES EFFETS DE FÉMINISATION ET DE FRÉQUENCE D'OCCURRENCE
SUR LA PRODUCTION DE LETTRES MUETTES DÉRIVABLES
CHEZ LES ENFANTS DU PRIMAIRE

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAITRISE EN DIDACTIQUE DES LANGUES

PAR

LISE Y. L'HEUREUX

DÉCEMBRE 2011

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.»

REMERCIEMENTS

Mes plus sincères remerciements s'adressent à Marie Labelle, pour m'avoir offert l'immense opportunité de prendre part à cette recherche. Merci pour la confiance que tu m'as accordée, tes conseils judicieux, tes commentaires réfléchis et ta grande disponibilité.

Merci également à Lyne Laplante et Lucie Godard d'avoir cordialement accepté d'être lectrices de ce mémoire.

Un merci tout spécial s'adresse aux enfants qui ont participé avec enthousiasme à cette recherche, ainsi qu'à leurs enseignantes et enseignants qui m'ont ouvert leur porte de classe. Merci aussi à Kathy, Lise et Sylvie qui ont facilité la collecte de données dans leurs écoles respectives.

Merci également à mes amis et mes proches pour leurs encouragements, leur patience et leurs corrections, et d'avoir enduré mes longs moments de réclusion et mes « discours impromptus » sur l'acquisition de l'écrit et la morphologie dérivationnelle du français. *J'vous l'ai-tu dit que j'vous aime?*

Finalement, je remercie ma mère, Yvette, pour ses encouragements constants dans la poursuite de mes études, ainsi que son soutien et sa générosité incommensurable. Je t'aime.

TABLE DES MATIÈRES

REMERCIEMENTS.....	ii
TABLE DES MATIÈRES.....	iii
LISTE DES FIGURES.....	iii
LISTE DES TABLEAUX.....	iv
RÉSUMÉ.....	viii
INTRODUCTION.....	1
CHAPITRE 1 PROBLÉMATIQUE ET OBJECTIFS DE RECHERCHE.....	3
1.1 La progression dans l'apprentissage de conventions morphologiques.....	3
1.2 L'acquisition des connaissances morphologiques.....	5
1.3 Résumé et objectifs de recherche.....	9
CHAPITRE 2 REPÈRES THÉORIQUES.....	11
2.1 Terminologie.....	11
2.2 Les caractéristiques du système orthographique français.....	14
2.2.1 L'opacité des systèmes orthographiques.....	14
2.2.2 Les conventions phonographiques du français.....	15
2.2.3 Les conventions morphologiques du français.....	17
2.2.4 Le marqueur muet du radical morphologique.....	19

2.3	Les théories et modèles développementaux en acquisition de l'écrit.....	20
2.3.1	Le modèle développemental en stades de Frith	20
2.3.2	La théorie à phases développementales d'Ehri	22
2.3.3	Le modèle à double fondation de Seymour	24
2.3.4	La convergence et la divergence des modèles développementaux.....	27
2.4	Les théories et modèles <i>usage-based</i> en acquisition de l'écrit	32
2.4.1	Les modèles intégratifs.....	34
2.4.2	Les modèles connexionnistes.....	36
2.5	Les bases de données lexicales	38
2.6	Les effets de fréquence en production de mots	41
2.6.1	La fréquence et l'analogie.....	45
2.7	L'âge d'acquisition, la fréquence cumulative et la trajectoire fréquentielle.....	46
2.8	Le développement des connaissances morphologiques.....	48
2.9	La production des lettres muettes dans les radicaux morphologiques.....	49
2.10	Les variables explicatives dans la production des consonnes muettes	52
2.11	La langue maternelle et l'acquisition du français écrit.....	54
2.11.1	La langue maternelle et le développement morphologique	56
2.12	Les questions spécifiques de recherche	59
CHAPITRE 3 MÉTHODOLOGIE		62
3.1	L'instrument de mesure.....	64
3.1.1	Les mesures de fréquence	65
3.1.2	L'indice de fréquence des stimuli	66
3.2	La procédure	67
3.3	La codification et la grille d'analyse des résultats	67

CHAPITRE 4	PRÉSENTATION DES RÉSULTATS.....	69
4.1	Recensement de la réussite des radicaux morphologiques	70
4.1.1	La progression de la réussite orthographique.....	72
4.1.2	La progression de la réussite morphologique.....	73
4.1.3	Comparaisons de la progression orthographique et morphologique	73
4.1.4	La correspondance morphologique entre les radicaux et les dérivés.....	74
4.2	La fréquence d'occurrence	77
4.2.1	La fréquence du radical	78
4.2.2	La fréquence de la famille morphologique	79
4.3	La condition morphologique du radical.....	81
4.3.1	La différence de fréquence entre le radical et le dérivé	82
4.3.2	La flexion féminisante	83
4.4	La langue parlée à la maison.....	84
4.5	L'interaction entre la langue et la fréquence.....	86
4.6	L'analyse des erreurs morphologiques	89
4.7	La typologie des erreurs morphologiques	89
4.8	La typologie des erreurs par groupe linguistique	90
4.9	La typologie des erreurs selon la fréquence des mots	91
4.10	La distribution des erreurs	92
CHAPITRE 5	DISCUSSION	96
5.1	La progression en production de consonnes muettes dérivables au primaire	96
5.2	Le lien morphologique entre le radical et son dérivé	101
5.3	L'effet de fréquence et la production de consonnes muettes	104
5.4	La féminisation et la production de consonnes muettes	109

5.5 L'effet de la langue maternelle	110
CONCLUSION.....	113
Les limites de l'étude et des pistes pour des recherches futures	114
Les recommandations didactiques et pédagogiques.....	117
APPENDICE A : LES PHONÈMES ET GRAPHÈMES DU FRANÇAIS	120
APPENDICE B : STIMULI PAR CONDITION AVEC INDICES DE FRÉQUENCE CUMULÉE (MANULEX)	121
APPENDICE C : PHRASES DE CONTEXTE POUR LA DICTÉE	122
APPENDICE D : CONSIGNES POUR PASSATION DE DICTÉES	123
APPENDICE E : FORMULAIRES DE CONSENTEMENT DES PARENTS - FRANÇAIS	124
APPENDICE F : FORMULAIRES DE CONSENTEMENT DES PARENTS - ANGLAIS	126
APPENDICE G : FORMULAIRES DE CONSENTEMENT DES ÉCOLES - FRANÇAIS	128
APPENDICE H : FORMULAIRES DE CONSENTEMENT DES ÉCOLES - ANGLAIS	130
BIBLIOGRAPHIE	132

LISTE DES FIGURES

Figure 2.1	21
Modèle à stades de Frith (1985).....	21
Figure 2.2	25
Modèle à double fondation de Seymour (1997).....	25
Figure 4.1 Progression des réussites orthographique et morphologique, des écarts-types des scores médians	72
Figure 4.2 Comparaison de la progression des écarts-types	74
Figure 4.3 Comparaison de la corrélation (r_s) réussite/fréquence du radical à la corrélation réussite/fréquence de la famille morphologique	80
Figure 4.4 Moyenne de la réussite morphologique par niveau selon la langue maternelle	85
Figure 4.5 Réussite morphologique des radicaux morphologiques plus ou moins fréquents	87
Figure 4.8 Typologie des erreurs des francophones, par niveau scolaire.....	90
Figure 4.9 Typologie des erreurs des anglophones, par niveau scolaire	91
Figure 4.10 Typologie des erreurs par niveau scolaire selon la fréquence des mots	92
Figure 4.11 Typologie des substitutions.....	94

LISTE DES TABLEAUX

Tableau 3.1	Population par niveau scolaire, selon le sexe et l'âge.....	64
Tableau 4.1	Probabilité (p) d'un effet du niveau scolaire sur la réussite des radicaux: comparaison deux à deux de l'effet entre les niveaux	71
Tableau 4.2	Synopsis de la correspondance entre les radicaux morphologiques et dérivés (CRD)	75
Tableau 4.3	Pourcentage de correspondance morphologique entre le radical et le dérivé	75
Tableau 4.4	Probabilité (p) de réussite du dérivé lorsque la réussite du radical est o rthographique	76
Tableau 4.5	Probabilité (p) de réussite du dérivé lorsque la réussite du radical est morphologique seulement	77
Tableau 4.6	Corrélation par niveau entre la fréquence du radical et le rang de réussite morphologique	79
Tableau 4.7	Corrélation par niveau, entre la fréquence de la famille morphologique et le rang de réussite morphologique	80
Tableau 4.8	P-valeurs de la fréquence d'occurrence sous les conditions de féminisation et de fréquence d'occurrence (Wilcoxon).....	82
Tableau 4.9	P-valeurs de la corrélation entre les scores de réussite et les conditions morphologiques utilisant un test bilatéral non paramétrique de Wilcoxon. .	82
Tableau 4.10	Comparaison (Wilcoxon) de la fréquence des familles morphologiques complètes (FAM) à celle des dérivés seulement (DS).....	84

Tableau 4.11	Probabilité (p) d'un effet de la langue maternelle sur la réussite morphologique lorsque la fréquence est contrôlée	86
Tableau 4.12	Probabilité (p) d'une interaction entre la langue maternelle et la fréquence.	87
Tableau 4.13	Typologie des erreurs par groupe linguistique	90

RÉSUMÉ

Cette recherche a pour objectif de recenser à quel niveau scolaire les enfants représentent les consonnes terminales dérivables qui ne sont pas sonorisées à l'oral et d'observer l'influence de trois facteurs explicatifs des scores de réussite : la féminisation du mot, sa fréquence d'occurrence et la langue maternelle des sujets. 431 élèves du primaire, de tous les niveaux, ont écrit une dictée composée de 24 paires de mots morphologiques de type lait/laitier, regroupés selon trois types de conditions morphologiques de féminisation et de fréquence d'occurrence. Les participants étaient composés de francophones (n = 278) et d'anglophones inscrits dans un programme d'immersion française (n=153). Nos données révèlent que dès la fin du 1er cycle du primaire, tous les enfants produisent des consonnes muettes terminales et que la moitié des enfants de 2^e année ont atteint un score de réussite de 70 % ou plus. Un plafonnement (< 5 % d'erreurs) est observé dès le début du 3^e cycle du primaire. Par ailleurs, la fréquence du radical et la fréquence de la famille morphologique ont des effets significatifs sur la production de la consonne muette du radical, et ce, à tous les niveaux. Cependant, contrairement aux observations empiriques antérieures où l'on avait observé des effets de féminisation, mais sans contrôler la fréquence entre les mots avec et sans féminin, dans la présente recherche aucun effet de féminisation n'a été observé. Cela suggère que l'effet de féminisation observé dans les autres études s'expliquerait par la fréquence des radicaux qui peuvent être féminisés et non par la présence de la flexion féminisante. Finalement, les résultats indiquent que la langue maternelle agit sur la production des consonnes muettes. À partir de la 3^e année, les francophones réussissent significativement mieux que les anglophones dans un programme d'immersion. De plus, on observe une interaction entre la langue et la fréquence des mots. Les deux groupes linguistiques ont des scores similaires lorsque les mots sont fréquents, mais lorsqu'ils sont rares, les anglophones n'atteignent pas le niveau de réussite des natifs du français.

MOTS CLÉS : morphologie, morphologie dérivationnelle, acquisition du langage, langage écrit, orthographe, français,

KEYWORDS: morphology, derivational morphology, language acquisition, written language, spelling, French

INTRODUCTION

Le système orthographique du français est parmi les plus opaques des langues européennes. Le manque de transparence entre les formes orale et écrite du français en ferait effectivement la langue la plus complexe à écrire (Fayol et Jaffré, 2008). Une simulation informatisée de la productivité entre le français oral et écrit a démontré qu'une cinquantaine de règles d'association entre les lettres et les sons permettait de transcrire plus de 85 % des phonèmes, mais seulement 50% des mots (Veronis, 1988). Un des facteurs explicatifs pour ce manque de transparence est la nature muette de sa morphosyntaxe. Les marqueurs grammaticaux de genre, de nombre et de personne qui ne sont pas prononcés à l'oral sont une grande source d'erreurs à l'écrit. Les marqueurs lexicaux dérivables le sont aussi, c'est-à-dire les consonnes terminales qui marquent l'association à la famille morphologique (mont/montagne), mais qui ne sont pas présentes à l'oral. Notre intérêt de recherche est centré sur la capacité des enfants du primaire à produire de tels marqueurs morphologiques muets.

Un des défis des didacticiens est de produire du matériel scolaire qui tient compte de l'âge auquel les enfants développent les capacités cognitives pour reconnaître¹ et produire des éléments orthographiques qui ne sont pas représentés à l'oral. À quel niveau scolaire est-ce que la majorité des élèves est en mesure d'écrire des mots en faisant appel à des processus non phonologiques? Pour activer la production de marqueurs qui ne sont pas représentés à l'oral, on peut envisager deux types de processus, un traitement orthographique, la reconnaissance globale du mot, ou un traitement analogique en associant la consonne muette du radical à son féminin (*droit/droite*) ou son dérivé morphologique (*lait/laitier*) pour justifier sa présence. Mais à quel niveau du primaire est-ce que les enfants représentent ces lettres muettes dans leurs écrits?

¹ Cet article adopte l'orthographe rénovée admise par l'Académie française.

La problématique du présent travail porte sur la production de ce genre de marqueurs. Dans cette perspective, un des objectifs fixés est de dresser un portrait de la progression, du début à la fin du primaire, de la production de marqueurs morphologiques muets. Les résultats seront examinés en contrôlant trois variables explicatives identifiées dans des recherches antérieures : la fréquence d'occurrence des mots, le type morphologique des mots et la langue maternelle des apprenants.

Cet ouvrage est composé de 5 chapitres. La problématique délimite les objectifs de recherche. Les fondements théoriques et les observations empiriques orientent les questions spécifiques de recherche. La méthodologie explique la démarche scientifique suivie. Les résultats expérimentaux quantifient les scores de réussite d'une épreuve de dictée. Finalement, la discussion opère un rapprochement entre les résultats observés, les fondements théoriques et les observations empiriques sur la production de mots à contenu morphologique. Une brève conclusion offrira des pistes pour les chercheurs, pour les didacticiens et pour les intervenants en milieu scolaire.

CHAPITRE 1

PROBLÉMATIQUE ET OBJECTIFS DE RECHERCHE

Ce chapitre délimite le problème qui aiguille notre recherche : l'acquisition de connaissances lexicales dérivationnelles liées à la production de marqueurs morphologiques lexicaux qui ne sont pas représentés à l'oral. On examinera d'abord la progression des apprentissages prescrite par le ministère de l'Éducation, du Loisir et du Sport du Québec (MELS) dans l'acquisition des connaissances morphologiques liées aux familles de mots. Suivra un regard théorique sur la recherche qui fait état de la production de consonnes muettes terminales dérivables, par exemple, le *p* dans le radical morphologique *camp* qui peut être déduit en le liant à un dérivé, comme *camper*. Les objectifs de cette recherche seront formulés à partir des limites observées dans certaines de ces études.

1.1 La progression dans l'apprentissage de conventions morphologiques

Dès le début de leur scolarisation, les enfants sont exposés à des mots écrits dits simples, comme *les*, *amie* et *petit*. Du point de vue d'un enfant de six ans, ces mots « simples » ne le sont pas nécessairement. Leur structure orthographique a une relation conflictuelle avec les conventions phonologiques qu'on leur enseigne puisqu'ils comportent des lettres muettes, sans représentation phonologique. Malgré leur absence à l'oral, les consonnes terminales dans les mots comme *les*, *amie* et *petit* sont remplies de sens et jouent un rôle morphologiquement essentiel dans l'acquisition du système orthographique français. Le *-s*, le *-e* et le *-t* de ces trois mots indiquent respectivement « qu'il y en a plusieurs », « que l'amie est une fille » et « qu'il y a un lien avec d'autres mots dans sa famille ». À quel âge est-ce que les enfants prennent conscience de l'information liée à ces marqueurs muets?

Il est crucial de respecter les capacités cognitives et langagières des enfants lorsque l'on développe les manuels scolaires qui les visent ou lorsque l'on prépare leurs leçons. À quel niveau scolaire doit-on explicitement enseigner les stratégies morphologiques pour orthographier ces mots? À quel cycle les conventions morphologiques du français doivent-elles apparaître dans le matériel didactique? Le programme de formation de l'école québécoise offre des pistes pour répondre à ces questions. Le MELS a mis en ligne une progression des apprentissages au primaire² qui sert à donner une ligne directrice aux enseignants dans la programmation de leurs leçons. Elle est aussi un outil de référence indispensable aux didacticiens dans l'élaboration des contenus de manuels scolaires.

Dans l'onglet *Français langue d'enseignement*, sous la rubrique *Compétence/Écrire des textes variés/Orthographe d'usage/Connaissances*, section C, *Le système orthographique*, le MELS suggère que c'est à la fin de la 3^e année que les élèves devraient pouvoir, par eux-mêmes, identifier dans un mot le rôle des lettres muettes qui marquent le lien avec des mots de la même famille morphologique (*gourmand/gourmandise*) (N^o 2b). On y précise « l'identification » des lettres muettes et non leur « production », mais la progression se trouve sous la rubrique des « compétences à écrire des textes variés », ce qui porte à confusion. On ne trouve aucune mention de la production des lettres muettes dérivables sous la rubrique *Connaissances*. Cependant, sous la rubrique *Pistes pour l'appropriation des connaissances*, section A, *Recourir à différentes stratégies pour mémoriser des mots correctement orthographiés*, on propose aux enseignants de favoriser l'enseignement de stratégies qui permettent d'« associer un mot à sa famille morphologique pour justifier l'orthographe d'une lettre muette ou d'un graphème » (N^o 6). L'indicateur de progression suggère que c'est en 2^e année que l'on doit commencer à enseigner ces stratégies morphologiques dérivationnelles et qu'à la fin du primaire, en 6^e année, la compétence devrait être acquise.

La présente démarche scientifique se penche sur la progression de la production de marqueurs morphologiques dérivationnels et considère des facteurs qui pourraient influencer

² <http://www.mels.gouv.qc.ca/progression/>

la production de telles lettres muettes chez les enfants du niveau primaire³. Dans cette perspective, nous ferons un rapprochement entre nos résultats et la progression visée par le MELS. Bien que cette progression ait été développée dans un contexte d'enseignement-apprentissage, on y prescrit « les connaissances que les élèves doivent acquérir au cours de chacune des années du primaire », c'est-à-dire un âge d'acquisition des habiletés enseignées. Il nous semble donc pertinent de faire un parallèle entre nos résultats, qui recensent la progression de la production de consonnes muettes dérivables chez des enfants du primaire, et les normes du MELS qui suggèrent un âge d'acquisition pour cette habileté.

De plus, nos résultats seront comparés aux postulats de modèles d'acquisition du langage écrit et aux résultats d'études empiriques sur le développement des connaissances orthographiques et morphologiques.

1.2 L'acquisition des connaissances morphologiques

Les connaissances morphologiques dans le développement de l'orthographe lexicale et grammaticale retiennent l'attention des chercheurs, particulièrement depuis la fin des années quatre-vingt (Bégin, Saint-Laurent et Giasson, 2010; Carlisle, 1988; Casalis, 2003; Leybaert et Alegria, 1995; Pacton, 2001; Pacton et coll., 2001; Sénéchal, 2000; Sénéchal, Basque et Leclaire, 2006; Snowling, 1994; Totereau, Thevenin et Fayol, 1997). Selon les modèles développementaux (Ehri, 1995; Frith, 1985; Seymour, 1997), l'acquisition des connaissances morphologiques se ferait tardivement, à la fin du parcours de l'apprentissage de l'écrit. Cependant, les résultats de plusieurs études (Pacton, Fayol, et Perruchet, 2005; Pacton et coll., 2001; Sénéchal, 2000; Sénéchal, Basque et Leclaire, 2006; Snowling, 1994; Totereau et coll., 1997; Treiman et Cassar, 1996) remettent en question ce postulat.

Ainsi, les résultats de Sénéchal (2000), qui a observé les effets morphologiques sur la production de consonnes muettes à la fin des mots morphologiques, contredisent cette hypothèse. Les mots morphologiques, tels que définis par Sénéchal, sont des mots lexicaux

³ Bien que cette démarche amène à observer la progression de connaissances liées au traitement morphologique en orthographe d'usage, elle ne prendra pas en compte les interventions spécifiquement reçues par la population expérimentale. L'effet de l'enseignement reçu et son influence sur le développement des connaissances seront traités lors de la discussion des limites de cette recherche.

ayant une lettre muette terminale qui peut être déduite à partir d'un mot de la même famille morphologique. Un mot morphologique est ainsi le radical d'une famille morphologique. Par exemple, les mots *droit* et *camp* sont des mots morphologiques puisqu'ils peuvent être associés à des mots de leur famille morphologique, par exemple *droite* et *camper* (dorénavant appelé les dérivés), pour justifier leur *t* et *p* muet respectif. Sénéchal (2000) a observé que dès la 2^e année les enfants franco-ontariens orthographient mieux les mots réguliers comme *journal* (correspondances lettres-son transparentes) que les mots morphologiques comme *gentil* (lettre muette dérivable), et les mots morphologiques mieux que les mots opaques comme *foulard* (lettre muette aléatoire). Ses résultats ont été reproduits dans deux recherches subséquentes (Pacton, 2001; Sénéchal, Basque et Leclaire, 2006) appuyant le postulat que les enfants francophones seraient sensibles aux représentations morphologiques dès le 1^{er} cycle du primaire. Ces observations contredisent le développement tardif du traitement morphologique postulé dans les modèles développementaux de Frith (1985), d'Ehri (1995) et de Seymour (1997). Sénéchal, Basque et Leclaire (2006) ont aussi observé, cependant, que l'information morphologique n'est toujours pas utilisée pleinement à la fin du 2^e cycle du primaire. Chez les enfants de la 4^e année, à peine la moitié des mots morphologiques étaient correctement orthographiés. Dans une étude similaire (Bégin, Saint-Laurent et Giasson, 2010), des élèves québécois de sixième ont correctement orthographié 85 % des mots morphologiques avec une flexion féminisante (*petit/petite*). Seulement 70 % des mots qui n'avaient pas de féminin ont été réussis et les élèves dont les capacités orthographiques étaient les plus faibles, n'en ont réussi que 52 %. Il nous semble pertinent de dresser un portrait plus clair de la progression en production de consonnes muettes dérivables, du début à la fin du primaire, en utilisant les mêmes stimuli expérimentaux à tous les niveaux.

On doit aussi considérer les facteurs qui auraient pu influencer les résultats observés. Une des variables explicatives dans toutes les études sur les consonnes muettes dérivables est la présence d'une flexion féminisante. Les mots qui avaient un féminin étaient significativement mieux réussis que ceux qui n'en avaient pas, suggérant que les flexions féminisantes (*droit/droite*) sont plus accessibles que les dérivés (*droit/droitier*) pour justifier la consonne muette. On peut cependant se demander si les stimuli avec un féminin dans ces trois études n'étaient pas plus fréquents que les stimuli qui n'avaient que des dérivés nominaux

rang/rangée), adjectivaux *vent/venteux* ou verbaux *camp/camper*. Si oui, on pourrait postuler que la fréquence expliquerait l'effet de féminisation observé dans les études de Sénéchal (2000), Pacton (2001) et Sénéchal, Basque et Leclaire (2006). En effet, Sénéchal (2000) avait suggéré que l'exposition des scripteurs au mot morphologique, ainsi que l'exposition à la lettre muette dans tous les mots de la famille morphologique, aurait pu rendre la production de la lettre muette plus accessible que celle des mots opaques qui sont solitaires. Elle n'a toutefois pas observé d'effet de fréquence d'exposition aux mots. De plus, la fréquence de la famille morphologique n'a pas été contrôlée. Il est aussi important de noter que la mesure de fréquence utilisée en était une où l'exposition à l'écrit est contrôlée à partir de mesures subjectives, c'est-à-dire que les enfants devaient dire si, oui ou non, ils connaissaient les titres de livres d'enfants populaires. Par ailleurs, Sénéchal a contrôlé la fréquence des stimuli expérimentaux d'après les indices de fréquence de Brulex; les fréquences de ses stimuli ne sont pas publiées. Brulex (Content, Mousty et Radeau, 1990) est une base de données dont le corpus est tiré de livres pour adultes et les fréquences, du Trésor de la grammaire française (Imbs, 1971), aussi tirées de livres d'adultes. Les fréquences tirées de données de textes écrits pour adultes ne correspondent pas nécessairement à la fréquence d'exposition aux mots des enfants.

Bégin, Saint-Laurent et Giasson (2010) ont reproduit les conditions expérimentales de Sénéchal, mais avec un contrôle objectif de la fréquence des stimuli, c'est-à-dire en choisissant des mots rares et fréquents dans la base de données Manulex (Lété, Sprenger-Charolles et Colé, 2004). Elles ont observé des effets de fréquence chez les élèves de 6^e année dans la production des trois sortes de mots : morphologiques avec un féminin, morphologiques sans féminin et lexicaux (opaques). Les mots fréquents étaient toujours mieux réussis que les mots rares. Elles ont aussi observé que l'écart de réussite était plus grand pour les mots lexicaux, qui sont solitaires, que pour les mots appartenant à une famille morphologique. Elles ont ainsi observé un effet d'interaction entre le type de mot et sa fréquence et en concluent que les enfants utiliseraient efficacement l'information morphologique contenue dans certains mots. Cependant, bien que la fréquence entre les mots du même type ait été contrôlée (fréquents et rares), la fréquence entre les types de mots (réguliers, morphologiques, lexicaux) n'a pas été contrôlée. Selon notre vérification dans

Manulex, l'ensemble de l'échantillon de type « morphologique avec féminin » était 5 fois plus fréquent que celui des mots morphologiques sans féminin et une fois et demie plus fréquent que les mots lexicaux. À notre avis, cela remet en question les conclusions qu'il y aurait une interaction entre le type de mot et la fréquence et qu'il aurait un effet de féminisation sur la production de la consonne muette d'un radical morphologique. Par ailleurs, aucun contrôle n'a été fait pour vérifier si la fréquence de la famille morphologique entière agissait sur la réussite du marqueur morphologique muet.

À la lumière des méthodes de contrôle de fréquence dans les recherches sur la production des marqueurs morphologiques muets, il nous semble important de vérifier si l'effet de féminisation observé persisterait avec un contrôle plus étroit de la fréquence entre les types de mots et un contrôle de la fréquence de la famille morphologique du mot cible.

Un dernier facteur problématique dans l'acquisition du langage écrit est celui de la langue première des apprenants. Le profil linguistique des élèves québécois est de plus en plus hétérogène, particulièrement dans la région métropolitaine de Montréal où 34,3% de la population a une langue maternelle autre que le français⁴. Sur l'île de Montréal la diversité linguistique est d'autant plus importante que dans le reste de la province. Plus de la moitié de la population a une langue maternelle autre que le français et le quart parle l'anglais à la maison. Puisque la Charte de la langue française québécoise (communément appelée la loi 101) dicte que la langue d'enseignement au primaire et au secondaire doit être le français⁵, un nombre considérable d'élèves sont des apprenants du français langue seconde (FL2) dans des programmes de français langue première (FL1).

La divergence linguistique et typographique de la langue maternelle d'un apprenant est un facteur dans l'acquisition d'une langue seconde (Genesee, 2004). En FL2, par exemple, la présence du genre nominal est une source de difficulté bien documentée (Genesee, 2007; Morris et Labelle, 2005). Il a aussi été démontré que les enfants en immersion française au Québec n'atteignent pas le rendement des enfants francophones dans la production de mots

⁴ Source : Secrétariat de la politique linguistique du Québec. Données 2006
<http://www.spl.gouv.qc.ca/documentation/rapportsondagesstatistiques/dynamiquedeslangues/tableaux/>

⁵ Une exception est accordée aux enfants anglophones. Ils peuvent fréquenter une école anglaise si un des parents a été éduqué en anglais au Québec.

qui requièrent des connaissances morphologiques flexionnelles de genre et de nombre (Morris et Labelle, 2005). Cependant, à notre connaissance, les effets de la langue maternelle sur la production de consonnes muettes à la fin des radicaux morphologiques n'ont jamais été étudiés. Vu le nombre important d'enfants en FL2 au Québec, il nous semble pertinent de considérer cette variable propre aux apprenants.

1.3 Résumé et objectifs de recherche

En résumé, cette recherche s'intéresse à la capacité des enfants du primaire à représenter les consonnes muettes à la fin des radicaux de familles morphologiques en situation d'écriture.

Contrairement à ce que suggèrent les modèles développementaux, il a été observé qu'en début de parcours de l'apprentissage formel de l'écrit, l'orthographe des jeunes enfants représente des connaissances morphologiques liées aux consonnes muettes à la fin des radicaux morphologiques (Pacton, 2001; Sénéchal, 2000; Sénéchal, Basque et Leclaire, 2006). De plus, il a été observé qu'au milieu de parcours, les enfants ne produisent pas systématiquement ces lettres muettes (Sénéchal, 2000; Sénéchal, Basque et Leclaire, 2006) et n'utilisent pas les stratégies morphologiques efficacement (Sénéchal, Basque et Leclaire, 2006), tandis qu'à la fin des études primaires, ils ne maîtrisaient toujours pas la production de ces marqueurs morphologiques muets (Bégin, Saint-Laurent et Giasson, 2010). Par ailleurs, selon la progression des apprentissages suggérés par le MELS, les enseignants devraient favoriser, dès la 2^e année, des activités visant la production des consonnes muettes en associant le mot à sa famille morphologique, mais ce n'est qu'en 6^e année que les enfants devraient maîtriser cette compétence. À notre connaissance, aucune recherche n'a observé la progression de la production des consonnes muettes du début à la fin des études de niveau primaire. Cette recherche se penchera sur cette progression.

Par ailleurs, il a été observé que la fréquence des mots agit sur la production des consonnes muettes en 6^e année (Bégin, Saint-Laurent et Giasson, 2010) donc cette recherche visera à reproduire ces observations à tous les niveaux du primaire. De plus, il a été observé que l'existence d'un féminin agit sur la production de la consonne muette (Bégin, Saint-Laurent et Giasson, 2010; Sénéchal, 2000). Cependant, les failles méthodologiques dans le contrôle

de la fréquence remettent en question les conclusions tirées par Sénéchal et Bégin. Sénéchal (2000) a contrôlé la fréquence des stimuli à partir de données adultes et n'a pas contrôlé la fréquence de la famille morphologique. Bégin, Saint-Laurent et Giasson (2010) n'ont pas contrôlé la fréquence entre les types de mots, ni la fréquence des familles morphologiques. Il est donc pertinent de tenter de reproduire leurs résultats avec un contrôle plus rigoureux de la fréquence.

Finalement, vu la diversité linguistique dans la grande région de Montréal et étant donné qu'aucune recherche ne s'est penchée sur la production de consonnes muettes à la fin des radicaux morphologiques en tenant compte de la langue de l'apprenant, l'observation de la réussite morphologique, en tenant compte de la langue maternelle, semble pertinente.

À la lumière de nos observations, les objectifs qui sous-tendent notre démarche scientifique sont les suivants :

1. Observer la progression, du début à la fin du primaire, de la production écrite de radicaux morphologiques et faire le lien avec la progression des apprentissages suggérée par le MELS.
2. Observer si la fréquence d'occurrence des mots morphologiques et de leurs familles morphologiques agit sur la production de la consonne muette à la fin des radicaux morphologiques.
3. Observer si la présence d'une flexion de féminisation agit sur la production de la consonne muette terminale des radicaux morphologiques lorsque la fréquence est contrôlée.
4. Observer si la langue maternelle agit sur la production des consonnes muettes à la fin des radicaux morphologiques

Dans le prochain chapitre, les concepts et repères théoriques, ainsi que les questions spécifiques de recherche qui guideront notre démarche scientifique, seront précisés.

CHAPITRE 2

REPÈRES THÉORIQUES

Avant de présenter les recherches empiriques qui sous-tendent cette expérimentation, quelques particularités orthographiques, phonographiques et morphologiques du système orthographique français seront exposées. Ensuite, des théories et modèles pour décrire le développement orthographique des enfants seront présentés et le rôle de la morphologie sera situé dans le développement orthographique des enfants. Après avoir tracé un portrait de la recherche empirique qui guide cette expérimentation, les questions spécifiques de recherche seront avancées. Dans un premier temps, la terminologie retenue dans ce document sera définie.

2.1 Terminologie

Pour décrire les systèmes d'écriture linguistique⁶, la terminologie peut varier d'un auteur à l'autre et le sens des termes peut varier; donc la terminologie retenue dans cet ouvrage sera définie.

Le terme *système orthographique* est utilisé pour décrire les conventions qui régissent l'écriture d'une langue alphabétique, tandis que *convention phonographique* et *convention morphologique* sont retenues pour décrire, respectivement, les règles pour représenter ses unités sonores de la langue et celles pour représenter ses unités significatives.

⁶ La notion de système d'écriture linguistique, définie par Fayol et Jaffré (2008) comme étant « des traces graphiques qui servent à rendre les formes d'une langue visibles, et lisibles », s'oppose ici à d'autres systèmes d'écriture, telle l'écriture musicale, informatique, comptable, etc. Elle englobe cependant les différents systèmes graphiques utilisés pour représenter les langues, telle l'écriture phonétique, alphabétique, logographique, etc.

Le *phonème* est la plus petite unité linguistique distinctive de la langue orale. Bien que leurs formes graphiques soient les lettres, les langues ont plus de phonèmes que de lettres dans leur alphabet donc l'unité graphique qui représente le phonème est le *graphème*. Celui-ci peut être une lettre simple ou accentuée (*a, à, â*)⁷, un *digramme* (*ai, oi, ch*) ou un *trigramme* (*eau, ain*). Le graphème peut décrire les unités graphiques qui sont phonologiquement distinctives, mais aussi des unités sémantiquement significatives (Catach, 1980). Ainsi, dans le mot *lait*, il y a deux phonèmes /l/ et /ɛ/ et trois graphèmes, *l, ai* et *t*.

Les expressions *correspondance phonographémique* (CPG) et *correspondance graphophonémique* (CGP) sont retenues pour décrire, respectivement, les graphèmes qui représentent les phonèmes dans la production de mots écrits et, à l'inverse, les phonèmes qui correspondent aux graphèmes dans des situations d'identification de mots. Pour alléger la lecture, les acronymes CPG et CGP seront favorisés. L'adjectif *phonographique* décrit à la fois les CPG et les CGP. Le terme *graphonème* décrit des unités stables dans la relation entre les représentations graphiques et la prononciation à l'oral (Jones, 1967; Véronis, 1986). Ainsi, le *-ille* du mot *fille* et le *ge-* du mot *geste* sont des graphonèmes. Les termes *attaque-rime* et *corps-coda* décrivent des parties constituantes de la syllabe. L'*attaque* est composée de la ou des consonnes qui précèdent la voyelle (chou, fleur) et la *rime* est composée de la voyelle et de toutes consonnes qui la suivent (chou, fleur). Le *corps* est composé de l'attaque et de la voyelle (chou, fleur) et la *coda* est composée de la ou des consonnes qui suivent la voyelle (chou n'a pas de coda, fleur).

Le *morphème* est la plus petite unité significative de la chaîne orale (Catach, 1980) et ce terme sera utilisé ici pour décrire les unités de sens qui peuvent ou non avoir un contenu phonique. Ainsi, le verbe *retournent*, a trois morphèmes, *re-*, *turn-* et *-ent*. Le terme *radical* est utilisé pour décrire la base des mots qui peuvent être fléchis ou dérivés. Par exemple, le radical du verbe *tourner* est *turn-*. Cependant, le nom *tour* est le radical de la famille morphologique du verbe *tourner* (*tour, tourner, tournage*, etc.). Dans cet ouvrage, le terme *radical morphologique* sera favorisé pour décrire le mot de base de la famille morphologique.

⁷ Les graphèmes présentés entre parenthèses dans cet ouvrage n'en sont que quelques exemples, pour une liste exhaustive, voir Catach, 1980, p.14

Nous sommes d'avis que l'expression « mot morphologique » utilisée par Sénéchal (2000) et Bégin, Saint-Laurent et Giasson (2010) est trop vague et pourrait décrire n'importe lequel des mots de la famille morphologique. Notons que, dans le cadre de notre recherche, le *radical morphologique* est un mot autonome, c'est-à-dire une entrée dans le dictionnaire, sur lequel on greffe un suffixe pour former une *flexion féminisante* ou un *dérivé morphologique*. Les expressions *mot court* et *mot long* seront aussi utilisées pour décrire, respectivement, un radical morphologique et un dérivé.

Le terme *conscience phonémique* décrit l'habileté à distinguer et à manipuler les phonèmes dans une syllabe ou un mot (Ehri et coll., 2001) tandis que la *conscience phonologique* fait référence à l'habileté de distinguer et manipuler différentes unités linguistiques, dont la syllabe, l'attaque et la rime, et le phonème. La *conscience morphologique* se définit comme « la conscience qu'a l'enfant de la structure morphémique des mots et sa capacité à réfléchir (sur) et à manipuler explicitement cette structure » (Carlisle, 1995, p. 194) (traduction tirée de Colé et coll., 2004, p. 709). Comme le notent Colé et ses collaborateurs, le terme *conscience*, qui est distingué (consciousness) du terme « awareness » en anglais, est utilisé pour décrire la manipulation implicite et explicite des unités linguistiques. La distinction dans les études citées n'est pas toujours faite explicitement donc les études peuvent faire référence à l'un des deux niveaux de conscience. Dans le présent travail, nous adoptons la définition élargie du terme *conscience* puisqu'il s'agit d'une évaluation quantitative en production de mots et non une évaluation de la connaissance procédurale de la règle en question.. Ainsi, la conscience, qu'elle soit phonémique, phonologique ou morphologique, renvoie à une habileté cognitive qui permet de réussir une tâche orthographique en manipulant les propriétés linguistiques des mots, sans nécessairement en comprendre la règle ou son fonctionnement. Le terme *connaissance* sera utilisé pour décrire la notion que la règle et son fonctionnement sont (ou devraient être) explicitement connus et appliqués.

Le terme *lemme* (Lété, 2005) s'oppose ici au terme *lexème*. Le *lemme* est la forme la plus simple d'un mot variable (le singulier pour les noms et adjectifs, l'infinif pour un verbe), tandis que le *lexème* correspondant à une entrée dans le dictionnaire (le masculin singulier des noms et adjectifs, l'infinif d'un verbe). Par exemple, les mots *chat* et *chats*, ainsi que les

mots *chatte* et *chattes* ont le même lexème, *chat*, mais ont deux lemmes distincts, *chat* et *chatte*, tandis que le verbe *tourner* est à la fois lemme et lexème.

Par ailleurs, l'adjectif *lexical* est utilisé dans deux différents contextes. Dans le premier, il s'oppose à l'adjectif *grammatical* pour distinguer entre la flexion et la dérivation. Dans le second, il s'oppose à *sublexical* pour distinguer deux sortes d'unités linguistiques, c'est-à-dire le mot entier (lexical) et les unités phonographiques ou orthographiques du mot (sublexical).

Finalement, puisqu'il est question de terminaisons lexicales dans ce document, et pour faciliter la lecture de certains passages, le tiret sera utilisé pour représenter une terminaison, c'est-à-dire l'orthographe de la rime finale d'un mot, par exemple, le *-ait* du mot *lait*.

2.2 Les caractéristiques du système orthographique français

Dans cette section, le système orthographique du français sera présenté, dont certaines particularités orthographiques, phonographiques et syllabiques qui en font un système complexe à acquérir. Dans un premier temps, cependant, les notions d'*opacité* et de *transparence* des systèmes orthographiques seront abordées.

2.2.1 L'opacité des systèmes orthographiques

Un système orthographique se distingue d'un système d'écriture par la notion de règles, de lois ou de conventions d'écriture (Catach, 1980; Druides, 2008; Fayol et Jaffré, 2008). Ces règles régissent les unités sonores distinctives d'une convention phonographique ainsi que des unités sémantiques qui véhiculent un sens linguistique. Les unités sémantiques, ou morphèmes ne sont pas nécessairement sonores. De ce fait, les liens entre l'oral et l'écrit ne sont pas toujours directs. Un système orthographique absolu ou « pur » en serait un où le lien entre l'oral et l'écrit serait minimal, où chaque phonème de la langue serait représenté par une seule lettre et, à l'inverse, une lettre représenterait un seul phonème (Fayol et Jaffré, 2008). Un tel système dénoterait un lien « économique » entre l'oral et l'écrit et la notion d'économie suggère un moindre effort cognitif pour représenter les unités linguistiques.

Comme le note Jaffré : « phonographic notation is, above all, useful as a recurrent infrastructure, and it must therefore be economical » (1997, p. 10).

Aucun système phonographique n'est pur, mais on décrit comme transparent un système orthographique qui s'approche de l'idéal phonographique. La transparence est déterminée par la proximité entre le nombre de phonèmes à l'oral et le nombre de graphèmes à l'écrit. Plus ils se rapprochent, plus le système est transparent (Fayol et Jaffré, 2008). Si ce nombre est très divergent, le système est dit opaque⁸. Le finnois serait la langue européenne la plus transparente (Jaffré, 1997), son système orthographique représente 23 de ses 24 phonèmes par une seule lettre (Joshi et Aaron, 2006). L'anglais, qui compte 40 phonèmes et plus de 550 graphèmes⁹, serait le plus opaque à décoder, mais le français serait le plus opaque à encoder (Fayol et Jaffré, 2008). Une des explications est que ce dernier est régi par deux niveaux de codification, phonographique et morphologique, ce qui en fait un système dit phonomorphémique (Pacton et Deacon, 2008). Ces deux conventions sont souvent conflictuelles puisque la codification morphologique du français est en grande partie muette. Ci-dessous, on décrit certaines caractéristiques des conventions phonographiques et morphographiques qui contribuent à l'opacité du système orthographique du français.

2.2.2 Les conventions phonographiques du français

Le français compte 36 phonèmes, dont 16 voyelles, 17 consonnes et 3 semi-consonnes (semi-voyelles), et environ 130 graphèmes¹⁰ (Catach, 1980; Fayol et Jaffré, 2008) (voir la liste des phonèmes du français à l'Appendice 1). Cependant, certains linguistes adopteraient une analyse encore plus élargie et compteraient jusqu'à 175 graphèmes (Fayol et Jaffré, 2008). Même si l'on s'en tient à l'analyse plus restrictive, l'écart entre l'oral et l'écrit, c'est-à-dire l'inconsistance des CPG et des CGP, est considérable. Selon Jaffré (1997), ce décalage expose « la complexité de la zone phonographique du français ».

⁸ En anglais, plusieurs chercheurs caractérisent la notion d'opacité par la profondeur : *deep or shallow orthographies*

⁹ Daigle, Mathieu et Montésinos-Gelet affirment que l'anglais aurait plus de 1000 phonogrammes (2008, p. 641)

¹⁰ Catach propose un « système phonique minimum » pour l'apprentissage de l'écrit qui aurait seulement 31 phonèmes, dont 11 voyelles (1980, p. 37). Elle propose aussi, pour l'apprentissage de l'écrit, de réduire le nombre d'unités graphiques essentielles 133 à 70 par un recensement plus restreint des graphies dans les textes, puis le réduire davantage, jusqu'à 45, en appliquant d'autres critères de filtrage. Elle appelle ce système de base le code minimal de transcription du français (1980, p. 36).

Le français est plus opaque de l'oral à l'écrit (CPG) que de l'écrit à l'oral (CGP). Ainsi, bien que le décodage du français soit complexe, son encodage l'est davantage (Fayol et Jaffré, 2008). Les CGP en lecture sont beaucoup plus régulières que les CPG lors de l'écriture (Peereeman et Content, 1999). Cela s'explique par le fait que la majorité des graphèmes ne représentent qu'un seul phonème, tandis que plusieurs phonèmes peuvent être représentés par plusieurs graphèmes. Il y a cependant un bon nombre de CGP multiples, par exemple, les graphèmes *o* (*robe/rose*), *en* (*vent/vient*) et *ch* (*chercheur/chœur*). Ce sont cependant les graphèmes contextuels comme le *s*, le *c* et le *g* (*saison, cicatrice, garage*) qui causent le plus de problèmes dans le traitement des CGP.

Inversement, le nombre de CPG multiples créant un phénomène d'interférence phonographémique lorsque les mots ne sont pas lexicalisés rend l'encodage de la langue française particulièrement difficile. Par exemple, il existe six graphèmes pour écrire le phonème /ɛ/ (*e, è, ê, ai, ei, ë*). Si l'on élargit la notion de graphème pour inclure les représentations graphiques en position finale qui ne s'expliquent pas par la morphologie, on pourrait en rajouter quatre autres (*béret, forêt, près, mais*). Si l'on inclut les lettres muettes en position initiale, on peut en rajouter encore deux (*herbe, haine*). Le choix parmi toutes les CPG peut être accablant lors de la production de mots, même pour des scripteurs avancés. De plus, au-delà de l'interférence des multiples CPG, la structure syllabique des mots peut jouer un rôle sur sa structure orthographique, plus précisément, sur la représentation graphique des phonèmes vocaliques (voyelles orales). Le choix de graphème lors de la production de phonèmes nasaux dépend aussi du contexte graphique dans lequel il se trouve (*jungle/humble*).

Quel enseignant du 1er cycle n'a pas demandé à ses élèves de taper dans les mains pour aider les jeunes enfants à prendre conscience de cette notion de syllabe? Cependant, est-ce qu'ils tapent tous le même nombre de fois pour un mot donné? Le mot *fil*le a-t-il une ou deux syllabes? Pour y répondre, il est nécessaire de distinguer entre les syllabes orales et les syllabes graphiques du français. L'unité fondamentale de la syllabe est la voyelle donc chaque phonème vocalique peut former une syllabe. Les syllabes peuvent être ouvertes ou fermées. La syllabe ouverte se termine par une voyelle (*me*) et la syllabe fermée, par une consonne (*mer*). Cependant, les lettres muettes finales changent la répartition des syllabes

entre l'oral et l'écrit. À l'oral, on compte une seule syllabe dans le mot *table* (/tabl/), mais à l'écrit, le *e muet* fait qu'il y a deux syllabes (/ta/ble/). À l'oral, compte une syllabe fermée et une syllabe ouverte dans le mot *esprit* (/es/pri/), mais à l'écrit, le *t muet* vient fermer la syllabe finale (/prit/). Cette différence cause de grandes difficultés en production de mots, en partie parce que ce sont les syllabes graphiques qui déterminent l'accentuation des voyelles en français (Catach, 1980) (*chè/vre*, *chien/ne*) et, en partie, parce qu'il est impossible de distinguer à l'oral si l'on doit produire une syllabe ouverte ou fermée à l'écrit (*cou/loup*, *peur/heure*).

Finalement, un des facteurs qui ajoutent à l'opacité du français est le nombre impressionnant d'homophones. Outre l'homophonie grammaticale que nous verrons plus loin, l'homophonie lexicale cause aussi beaucoup de confusion (*les/lait/laid/l'es/l'est*).

Ces exemples de facteurs d'opacité dans les conventions phonographiques ne sont pas exhaustifs, mais ils donnent un aperçu de la complexité du système orthographique du français. Cependant, ils ne représentent qu'un des deux niveaux de la codification représentés dans l'orthographe du français. Le niveau morphologique, qui regroupe les conventions flexionnelles et dérivationnelles du système orthographique est une source imposante d'opacité entre l'oral et l'écrit puisque les marqueurs grammaticaux et lexicaux muets ne peuvent pas être traités à l'aide de processus cognitifs qui permettent un décodage ou un encodage phonologique. Il est donc nécessaire de faire appel à des processus cognitifs d'un autre ordre pour représenter ces marqueurs dans le lexique orthographique.

2.2.3 Les conventions morphologiques du français

Un système orthographique se distingue par la notion de règles ou de conventions, et celles-ci s'étendent au-delà des unités sonores distinctives pour véhiculer un sens. Jaffré note :

« Toutes les orthographes sont mixtes, ce qui est la condition sine qua non de leur fonctionnement social : pour être faciles à apprendre, elles doivent être aussi proches que possible de la représentation des sons d'une langue. Mais pour être utiles, elles doivent également tenir compte de sa structure lexicale et grammaticale. » (Fayol et Jaffré, 2008, p. 85)

Ainsi, au-delà des unités linguistiques phonologiquement distinctives, on trouve des unités linguistiques morphologiquement significatives. Les conventions morphologiques gèrent les structures grammaticales et lexicales pour générer de nouvelles formes de mots par des procédés morphologiques : la composition, la flexion et la dérivation (Polguère, 2003). Le procédé de composition ne sera pas approfondi, car il ne fait pas l'objet de cette recherche. Il suffit ici de dire qu'il consiste à joindre deux mots du lexique pour en créer un troisième (*bon* + *homme* = *bonhomme*). Les procédés de flexion et de dérivation concernent la composition de mots de la même famille morphologique à partir des deux structures morphologiques de base : un radical et un affixe.

Deux particularités distinguent la flexion de la dérivation : la flexion est de nature grammaticale et la dérivation est de nature lexicale; la flexion ne crée pas de nouveau lexème tandis que la dérivation crée une nouvelle entrée dans le dictionnaire. Lors du procédé de flexion, le radical reste le même et les nouveaux mots créés partagent le même lexème. Les procédés flexionnels distinguent le genre et le nombre des noms et des adjectifs, et la personne, le nombre et le temps des verbes. Par exemple, dans le mot *chats*, le suffixe *-s* se greffe au radical *chat*, créant un nouveau mot qui précise qu'il y a plusieurs *chats*. L'ajout du morphogramme du pluriel *-s* se fait par un procédé flexionnel donc n'engendre pas de nouveau lexème. Cependant, le mot *chaton* est un lexème à part entière puisqu'il est formé par processus de dérivation en y ajoutant le suffixe diminutif *-on* pour créer un nouveau mot signifiant *un petit chat*. Le mot radical et ses dérivés sont des lexèmes qui se regroupent en famille morphologique.

La structure morphologique du français écrit est très riche. Par des procédés de flexion ou de dérivation, les affixes changent la forme des mots, leur sens grammatical et la syntaxe de la phrase. Une des caractéristiques qui fait du français une langue beaucoup plus complexe à écrire qu'à lire est la prévalence des marques muettes de son système morphologique (Fayol et Jaffré, 2008). Par exemple, la nature muette des suffixes flexionnels du genre, du nombre et de la personne exacerbe la complexité d'encodage du français, en créant un nombre important d'homophones grammaticaux (*joli/jolie/jolis/jolies*, *sautais/sautait/sautaient*) qui ne peuvent être distingués que par le contexte morphosyntaxique (en l'associant à un déterminant nominal ou au sujet du verbe).

2.2.4 Le marqueur muet du radical morphologique

Contrairement à la flexion, lors du procédé de dérivation, l'homophonie n'est pas un enjeu, car les suffixes qui créent de nouveaux mots sont sonorisés (*camp/camper/campeur*). Cependant, le radical se termine souvent par une consonne qui ne se prononce pas. Dans ce cas, le scripteur ne peut pas se baser sur un traitement phonologique du radical pour orthographier la consonne muette. Deux voies s'offrent à lui. Il peut retrouver en mémoire l'orthographe du radical, ou procéder par analogie et déterminer la consonne muette sur la base d'un dérivé morphologique où cette consonne est sonorisée. La flexion du genre (féminisation) permet aussi la déduction de la consonne muette lorsque la syllabe graphique finale est fermée (*laid/laide*). Pour orthographier la lettre muette par analogie, le scripteur doit faire appel à un niveau cognitif plus élevé que l'évocation d'un phonème par un processus de traitement alphabétique (Seymour, 1997; Snowling, 1994; Treiman et Cassar, 1996). La capacité de faire appel à un processus de traitement analogique est cruciale pour orthographier les radicaux morphologiques nouveaux dont la consonne terminale est muette.

En résumé, l'acquisition orthographique est influencée par l'opacité (ou la transparence) entre l'oral et l'écrit. Il est reconnu qu'un système orthographique opaque a une incidence sur l'acquisition de l'orthographe (Pacton et coll., 2005, 2001; Sénéchal, 2000). L'opacité du français émane de la complexité de son système orthographique, dit phonomorphémique (ou morphophonémique). La nature muette des conventions morphologiques impose un défi important aux enfants en apprentissage de la langue écrite. Pour décoder et encoder les mots qui contiennent de l'information morphologique, ils doivent activer des processus de reconnaissance et de production de mots autres que ceux qui font l'association entre les lettres et les sons.

« L'apprenti scripteur [...] français aura donc intérêt à comprendre qu'orthographier ne se résume pas à appliquer un ensemble de règles phonologiques (lettres-sons) et à apprendre un plus ou moins grand nombre d'exceptions. Le développement d'habiletés morphologiques est souvent utile, voire indispensable. » (Pacton, 2003, p. 27)

Cela nous mène à nous demander comment les processus cognitifs d'identification et de production de mots morphologiques se développent, à quel niveau scolaire ils se manifestent

et à quel niveau ils deviennent fonctionnels. Dans le prochain sous-chapitre, nous examinerons des modèles qui tentent d'expliquer l'acquisition du langage écrit.

2.3 Les théories et modèles développementaux en acquisition de l'écrit

Pour traiter un mot écrit sémantiquement, on doit d'abord faire appel à des processus cognitifs pour l'identifier. Il est généralement reconnu que l'acquisition du langage écrit dépend de deux processus cognitifs d'identification et de production de mots (Ehri, 1995; Ellis, 2002; Frith, 1985; Perfetti, Rieben et Fayol, 1997; Seymour, 1997; Share, 1995). Le premier processus permettrait de reconnaître et de produire un mot en ayant recours aux CGP et aux CGP. Le deuxième processus permettrait de reconnaître et produire les mots à partir d'une représentation complète du mot en mémoire. L'appellation et la définition de ces deux processus varient d'un chercheur à l'autre et la modélisation de leur développement a pris plusieurs formes. Les modèles développementaux à stades ou à phases postulent que le développement des processus phonologiques se fait avant celui des processus orthographiques, tandis que les modèles à deux voies (ou doubles voies) théorisent que les deux processus se développent simultanément. Trois modèles développementaux seront présentés dans la prochaine section, le modèle à stades de Frith (1985), le modèle à phases d'Ehri (1995) et le modèle à « double fondation » de Seymour (1997), en s'intéressant particulièrement aux zones de développement des connaissances morphologiques.

2.3.1 Le modèle développemental en stades de Frith

Pour tenter d'expliquer la « déviation cognitive » de la dyslexie développementale, Frith (1985) a proposé une théorie d'acquisition normale des processus cognitifs d'identification et de production de mots écrits chez les enfants. Son modèle alloue trois stades développementaux : logographique, alphabétique et orthographique (voir la figure 2.1).

Le premier stade de lecture est lorsqu'un enfant distingue les mots familiers de façon logographique. Il apparaît lorsque l'enfant a une conscience symbolique de l'écriture, mais n'a pas suffisamment de connaissances linguistiques pour que l'on considère cette activité comme étant de la lecture. Les processus logographiques consistent à identifier les lettres

initiales des mots à partir d'indices visuels saillants que l'on retrouve, par exemple dans les logos, mais sans considération pour l'ordre des lettres. Par exemple, dès l'âge de trois ans, un enfant pourrait reconnaître le mot *Disney*, mais répondrait le même mot à partir des stimuli *Dsneiy** ou *Dsinye**. Le processus initial de stockage de mots à partir d'indices visuels arbitraires est inefficace pour récupérer et reproduire les mots en situation d'écriture. Ce processus logographique, qui n'est observable qu'en lecture, est représenté par le stade 1A dans le modèle (voir Figure 2.1)¹¹. Au stade 1B, l'enfant commence à retenir des parties de mots, ce qui lui permet de reproduire des graphies (complètes ou partielles) à partir d'indices visuels, sans toutefois en comprendre le système de codification. Ceci expliquerait les graphies partielles ou erronées lorsque les jeunes enfants tentent d'écrire des mots, par exemple, les graphies *cht** ou *chta** au lieu de *chat*, ou *ptti** ou *ptit** pour *petit*.

	LECTURE	ÉCRITURE
1A	Logo 1	∅
1B	Logo 2	Logo 2
2A	Logo 3	Alpha 1
2B	Alpha 2	Alpha 2
3A	Ortho 1	Alpha 3
3B	Ortho 2	Ortho 2

Figure 2.1
Modèle à stades de Frith (1985)

Au stade 2A, l'enclenchement des processus alphabétiques se ferait par l'écriture puisque, selon Frith, le système alphabétique est « fait sur mesure » pour l'écrit (1985, p. 311). Au début du stade alphabétique, les processus phonologiques s'enclenchent pour produire des mots. Des CPG simples (lettre-son) sont utilisées pour produire des mots, mais les enfants continuent à lire en utilisant le traitement logographique. C'est ici, par exemple, que les

¹¹ Nous avons marqué le début de chaque stade en gris. La progression entre l'identification et la production des mots lecture et l'écriture est indiquée avec des flèches noires. Les flèches translucides marquent la progression de chacune des compétences de lecture et d'écriture.

serait donc à partir de l'utilisation CPG en production de mots que les processus de décodage se développeraient pour identifier les mots (2B). Frith appuie ce postulat par des observations réalisées auprès d'individus ayant des dyslexies acquises. La lecture phonologique de ces derniers « était toujours accompagnée d'une écriture phonologique, mais l'écriture phonographique n'était pas toujours accompagnée de l'habileté de lire phonologiquement » (1985, p. 311). Elle postule ainsi que l'identification et la production de mots sont des habiletés distinctes qui peuvent s'enclencher à différents moments, et que la production par CPG précéderait l'identification par CGP.

Au début du 3^e stade, le stade orthographique, les enfants utilisent encore les CPG pour produire les mots, mais, à la lecture, l'accès direct aux mots s'enclenche (3A), c'est-à-dire que les enfants font abstraction des CGP et reconnaissent instantanément des unités linguistiques plus larges. S'ils se servent de la phonologie, ce ne serait que pour sonoriser les unités plus larges comme les morphèmes et les mots en entier, et non pour sonoriser les phonèmes. L'écriture orthographique se développe avec l'exposition à ces unités morphologiques et lexicales plus larges (3B). C'est au stade orthographique que les enfants utilisent des stratégies contextuelles et analogiques pour orthographier les mots et, à la longue, des stratégies prédictives pour orthographier de nouveaux mots à partir des unités stockées en mémoire. C'est aussi à ce stade du développement orthographique que les connaissances morphologiques émergent. Ainsi, on pourrait postuler que c'est à ce stade que les enfants francophones seraient en mesure de produire des lettres muettes terminales morphologiques en faisant le lien avec des mots de la même famille morphologique.

Le modèle de Frith est un incontournable en recherche sur l'acquisition de la lecture, mais beaucoup d'auteurs l'ont critiqué, dont Ehri (1995) qui a formulé sa propre théorie du développement du langage écrit.

2.3.2 La théorie à phases développementales d'Ehri

Ehri (1995) a développé sa théorie pour expliquer la lecture par reconnaissance instantanée, ou par voie directe (sight words). Bien que son auteur l'ait originalement conçu pour expliquer les processus de lecture, elle la reprend dans un article subséquent (1997) où elle

postule que la lecture et l'écriture utilisent les mêmes sources d'information stockées en mémoire pour lire et écrire. Ehri soutient qu'il y a trois façons de traiter les mots. On traite par la mémoire les mots connus, et par invention ou par analogie, les mots inconnus. Le traitement par mémoire consiste à récupérer la forme d'un mot stocké dans le lexique orthographique. Le traitement par invention s'apparente au décodage, c'est-à-dire qu'on identifie les mots en assemblant la prononciation de sons qui correspondent aux lettres, et, inversement, on les produit en générant des lettres plausibles qui correspondent aux sons que l'on entend à l'oral. Le traitement analogique consiste à utiliser un mot connu, le comparant au mot inconnu que l'on tente d'identifier ou produire.

Ehri propose quatre phases de développement cognitif: préalphabétique, alphabétique partielle, alphabétique complète et alphabétique consolidée. Dans la phase préalphabétique d'Ehri, les enfants reconnaissent des mots à partir d'indices visuels seulement, par exemple des indices saillants comme les « deux yeux » du mot *look* (regarder) (1995). Ils arrivent difficilement à relire les mots, car les indices formés en mémoire sont aléatoires ou non systématiques. Les processus de reconnaissance sont inopérants pour transférer ces indices à la reconnaissance de mots nouveaux (1997). À l'écriture, si des lettres sont produites, plutôt que du simple griffonnage, elles ne sont pas reconnues puisqu'aucun lien avec le nom ou le son n'est établi.

Ce n'est qu'à la phase alphabétique partielle, lorsque les enfants ont appris l'alphabet, que les processus impliquant la phonologie s'enclenchent pour identifier et produire des mots. Lorsque les enfants commencent réellement à lire (move into reading), ils passent d'un traitement purement visuel à un traitement qui concerne des liens entre les lettres et les sons. Dans la phase alphabétique partielle, ils représentent les mots à partir de quelques lettres seulement puisque le repère des correspondances lettre-son est incomplet. En français, on pourrait imaginer qu'à partir de représentations des lettres *p* et *t* dans le mot *petit*, un enfant pourrait lire /pəti/ en voyant les mots *parti* ou *pitié* qui ont une forme visuelle similaire ou produire des mots à partir des phonèmes les plus saillants, par exemple, écrire **cha* pour le mot *chat* ou **ti* pour *petit*.

Les enfants atteignent la phase alphabétique complète lorsqu'ils réussissent à lire les mots connus instantanément. Lorsqu'ils seront exposés assez fréquemment à un mot, les apprenants réussiront à stocker sa structure complète en mémoire, créant un accès direct entre le lexique oral et le lexique orthographique. Ce lien direct entre la forme orale et la forme écrite du mot est possible, car les principales correspondances phonographiques sont maintenant connues (2005).

La phase alphabétique consolidée sera marquée par l'émergence d'un corpus de mots reconnus instantanément, « qui se développe rapidement lorsqu'il est alimenté par l'exposition à l'écrit » (2005). Les enfants retiennent désormais des structures plus larges, composées de séquences de lettres et de régularités orthographiques (digrammes, lettres doubles, muettes, contextuelles, etc.), ainsi que des morphèmes et des structures sublexicales (syllabes, attaque/rimes, affixes, etc.). Le stockage d'unités graphiques complexes permet de récupérer et produire l'orthographe des mots plus rapidement et efficacement, et de stocker l'orthographe de mots dont la structure orthographique ne suit pas les conventions phonographiques.

Ehri (1997) appelle aussi cette dernière phase dans l'apprentissage alphabétique la phase morphémique, car elle est marquée par la capacité des apprenants à utiliser des stratégies analogiques et prédictives pour orthographier de nouveaux mots. Ce serait donc à ce dernier stade de développement que les enfants seraient en mesure de justifier l'orthographe d'une lettre muette en l'associant à un dérivé ou à la flexion féminisante.

Le troisième modèle développemental présenté, celui de Seymour (1997), remet en question le déploiement en stade des processus cognitifs phonologique et orthographique des enfants, mais appuie le postulat que les connaissances morphologiques se développent à la fin du processus d'acquisition.

2.3.3 Le modèle à double fondation de Seymour

Contrairement à un déploiement successif des processus de traitement de l'écrit comme le suggèrent Frith et Ehri, Seymour (1997) préconise un développement parallèle de deux

processus fondateurs au développement orthographique : la reconnaissance et le stockage logographique (dorénavant appelé le processus logographique) et le processus alphabétique. L'activation de ces processus dépend de l'évolution de la conscience linguistique. Ensemble, ces deux processus s'activent pour former deux structures centrales, le cadre orthographique (Orthographic Framework) et le cadre morphographique (Morphographic Framework) (voir figure 2.2).

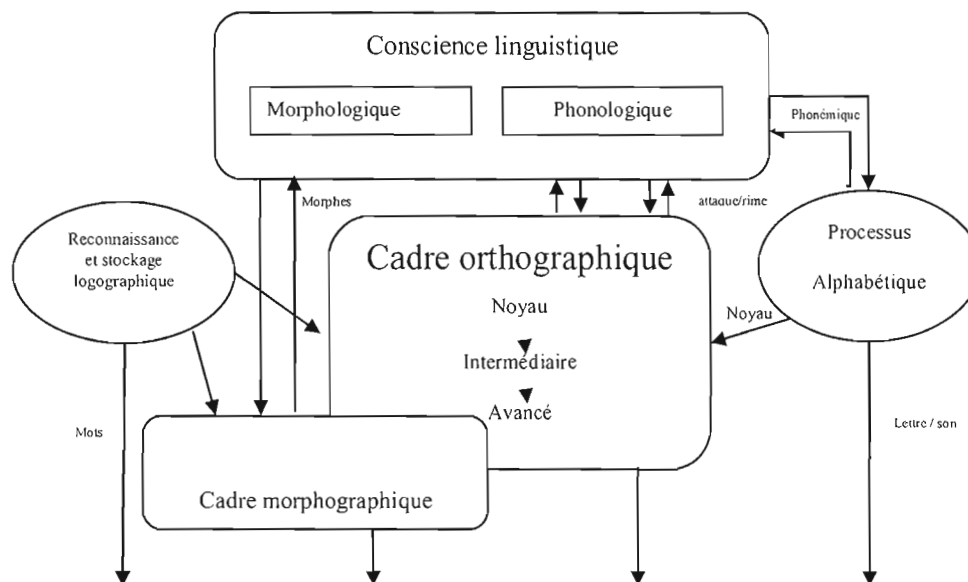


Figure 2.2
Modèle à double fondation de Seymour (1997)

Le processus logographique tel que décrit par Seymour concerne la reconnaissance et le stockage de mots en associant leur forme orale à leur forme écrite. « In this account, the logographic process deals with word-level representations and is not limited to primitive prealphabetic or 'visual' recognition. » (Duncan et Seymour, 2000, p. 146). Une schématisation lexicale (plus ou moins précise) s'active en début de parcours et est stockée en mémoire à partir des mots rencontrés. Les mots connus servent de modèles (exemplars) par lesquels l'apprenant assimile l'information orthographique nécessaire pour reconnaître et produire de nouveaux mots (1997, pp. 324–325). En contrepartie, le processus alphabétique se fonde sur la connaissance des lettres et des sons qu'ils représentent. Même s'il peut être argué que la reconnaissance de lettres est une autre forme de reconnaissance logographique

(associer un indice visuel arbitraire, une lettre, au son produit en nommant cet indice), l'essentiel du processus alphabétique est l'opération qui traduit les phonèmes en graphèmes (1997, pp. 325–326) et vice versa. Ces deux processus jouent un rôle fondateur (founding role) dans la construction du lexique orthographique. Bien qu'ils soient fondateurs, ils perdurent tout au long de l'acquisition de l'écrit.

La conscience linguistique joue le rôle de facilitateur (enabling function) en activant ces processus qui permettent de jumeler des segments orthographiques à des unités linguistiques (conscience phonologique) et de former le cadre orthographique, la structure centrale du modèle. Ainsi, l'émergence du traitement alphabétique de base et de connaissances lexicales (word specific features) marque la genèse du cadre orthographique, son noyau (core). Le cadre orthographique est ensuite alimenté par l'assimilation de correspondances de plus en plus complexes qui s'organisent selon les structures syllabiques, et par le développement de représentations orthographiques de ces structures. La structure centrale avancée est symbolisée par le traitement de mots multisyllabiques, associé au développement de la conscience morphologique, c'est-à-dire à l'association de segments orthographiques à des unités syllabiques porteuses de sens et à leurs propriétés combinatoires. C'est ainsi que le cadre morphographique (Morphographic Framework) émerge.

Le développement de la structure centrale serait cependant complexifié par la divergence entre le développement linguistique et le principe alphabétique lui-même. Le développement naturel de la conscience linguistique se ferait des unités linguistiques plus larges vers les unités plus petites, tandis que le principe alphabétique requiert un développement hiérarchique inverse.

Ainsi, un enfant prend conscience du mot avant la syllabe, de la syllabe avant l'attaque/rime et de l'attaque/rime avant le phonème. Inversement, dans un système alphabétique, la « construction » de mots part du phonème. Les phonèmes se fusionnent en unités plus larges, dont l'attaque/rime, puis la syllabe, puis la segmentation de syllabes en mots. Quoi qu'il en soit, la conscience des petites unités linguistiques et l'activation du processus alphabétique sont requises pour activer le noyau du cadre orthographique, et la formation de ce noyau est essentielle au développement de représentations orthographiques de plus en plus larges, ce

qui mène aux niveaux intermédiaire et avancé du cadre orthographique. Il est aussi présumé que la structure des mots qui ne suivent pas les conventions du système alphabétique est représentée individuellement avec un « index » interne qui signale leur divergence avec les conventions phonologiques stockées dans le cadre orthographique.

Selon Seymour, au-delà de la formation du noyau, l'évolution de l'acquisition des structures orthographiques n'est pas connue.

« The precise hierarchy of acquisition of the orthographic structure is not known and may be a matter of individual variation. However, it seems likely that the process is extended in time and may not be complete until toward the end of primary schooling. » (1997, p. 330)

Dans les étapes intermédiaire et avancée du cadre orthographique, les enfants font l'acquisition de structures sublexicales de plus en plus variables et complexes. Le cadre orthographique avancé est marqué par l'émergence d'un cadre morphographique. L'implication de la conscience syllabique et de la conscience morphémique sur le développement du cadre orthographique n'est pas établie. Cependant, Seymour postule que le cadre orthographique avancé serait accompagné du développement d'un cadre morphographique séparé qui se superpose au cadre orthographique central. Les représentations morphologiques sont dépendantes des items sublexicaux stockés en mémoire. Le cadre morphographique se formerait ainsi par l'interaction entre la conscience morphémique et les représentations logographiques d'unités syllabiques stockées en mémoire, ainsi que par la connaissance des propriétés combinatoires de ces items sublexicaux. Par extension, on pourrait imaginer que les structures syllabiques terminales complexes de la langue française, particulièrement celles qui comprennent les consonnes muettes terminales qui renferment de l'information morphologique, seraient stockées dans un cadre orthographique avancé et représentées de la façon décrite par Seymour.

2.3.4 La convergence et la divergence des modèles développementaux

Les trois auteurs des modèles développementaux présentés ci-dessus soutiennent tous l'axiome qu'il y a deux processus cognitifs impliqués dans le traitement de la langue écrite. Cependant, ils ne se rejoignent pas toujours sur la similarité des processus de lecture et

d'écriture, sur la terminologie pour décrire le développement du langage écrit ou les concepts qu'elle décrit, ni sur la façon dont les processus se déploient.

Au niveau terminologique, l'utilisation du terme « logographique » est fort divergente. Frith et Ehri convergent sur l'idée que les enfants vont produire des mots de façon purement visuelle avant de développer des connaissances du système alphabétique, mais Ehri juge que le terme « stade logographique » utilisé par Frith est trompeur puisqu'il fait référence aux systèmes d'écriture logographiques comme le chinois. Elle soutient que les représentations des enfants à ce stade primitif de la lecture ne peuvent se comparer à celles des lecteurs de systèmes logographiques. De plus, l'utilisation du mot « orthographique » est trop générale et ambiguë, vu ses maintes définitions. Elle préfère donc le terme préalphabétique (2005, p. 117). Par ailleurs, Seymour n'inclut pas une étape de prélecture non phonologique dans son modèle et le mot « logographique » prend un tout autre sens dans son modèle, que nous verrons plus loin.

Au niveau conceptuel, Frith et Ehri se rencontrent sur un déploiement successif des processus cognitifs qui permettent de traiter la langue écrite, ce qui diverge de Seymour qui postule que le déploiement des processus alphabétique et logographique se fait parallèlement et qu'ils interagissent. Par contre, Frith avance que la lecture peut se faire par décodage ou par reconnaissance visuelle, mais pour Ehri et Seymour, la lecture dépend toujours d'un traitement alphabétique quelconque. Cela implique qu'il faut nécessairement faire appel à des processus phonologiques pour lire.

De plus, Frith soutient que la lecture et l'écriture utilisent des processus cognitifs distincts qui peuvent s'activer à des moments différents. Ehri et Seymour ne font pas cette distinction. Pour Seymour, « I make the simplifying assumption that these two processes - reading and spelling- are both dependent on the development of a single central orthographic resource. » (1997, p. 319) Ehri rejoint Seymour sur ce postulat. Elle avance que la lecture et l'écriture utilisent les mêmes sources d'information stockées en mémoire, des connaissances du système alphabétique et des connaissances propres aux mots. Selon elle, la seule différence entre les deux processus est la quantité d'information qui doit être stockée en mémoire pour répondre correctement aux stimuli. L'écriture requiert un plus grand stockage que la lecture

qui ne requiert qu'une seule réponse, celle de prononcer le mot. Pour écrire, on requiert plusieurs réponses, une pour chaque lettre à orthographier, et ce, dans le bon ordre. Ainsi, la lecture et l'écriture sont semblables, bien que la seconde soit plus complexe que la première.

Malgré leur terminologie divergente, Frith et Ehri s'accordent pour dire que les jeunes enfants passent d'un traitement visuel précoce de l'écriture (stade logographique ou phase préalphabétique) à un traitement alphabétique où il y a des représentations concordantes entre les lettres et leurs sons. Pour Seymour, ce sont les processus alphabétiques qui activeraient les processus de stockage visuel d'unités lexicales pour former le cadre orthographique central. Il utilise l'expression « logographic recognition » pour décrire le stockage et la reconnaissance instantanée d'unités lexicales de taille plus ou moins large, mais qui a préalablement été traitée par des processus alphabétiques. Pour Seymour, le traitement logographique ne s'enclenche pas à une période qui précède le traitement alphabétique, pour ensuite s'estomper. La reconnaissance et le stockage logographique sont le résultat d'une convergence entre la reconnaissance visuelle et la procédure alphabétique qui se précisent, et qu'ils sont des mécanismes qui se perfectionnent tout au long du développement orthographique.

Le perfectionnement du mécanisme de reconnaissance logographique de Seymour semble rejoindre l'hypothèse plus récente d'Ehri (2005) voulant que l'identification de mots de façon instantanée ne soit pas une « stratégie » de lecture comme le sont les stratégies de décodage ou analogiques de mots nouveaux. Elle conclut que la reconnaissance instantanée entre l'écrit et l'oral (sight word reading) ne peut être stratégique puisqu'elle n'est pas issue d'un choix ou d'un raisonnement quelconque. Cette façon de traiter les mots est un réflexe automatique qui se développe à partir de connaissances phonologiques, orthographiques ou morphologiques et est le résultat d'un développement cognitif lié à la fréquence d'exposition au mot.

Un second terrain d'entente entre les trois auteurs est le postulat que l'exposition croissante au langage écrit permet aux enfants de traiter et d'emmagasiner en mémoire des unités sublexicales de plus en plus larges. L'accès direct à la forme écrite des mots rencontrés régulièrement, c'est-à-dire le traitement orthographique, marquerait l'acquisition du langage

écrit. Ce serait dans cette dernière étape, lorsque suffisamment de représentations lexicales et sublexicales sont stockées dans le lexique orthographique, que les processus de traitement morphologique émergeraient. Si l'on se base sur les prémisses des modèles développementaux, on pourrait conclure que ce serait à cette période développementale tardive que les enfants seraient en mesure de faire l'association d'un mot à sa famille pour justifier l'orthographe d'un marqueur morphologique muet.

Pour résumer, les modèles développementaux présentés ci-dessus tentent d'expliquer l'acquisition du langage écrit par des étapes développementales de processus cognitifs. Au début de leur apprentissage, les enfants sont exposés à de simples correspondances lettres-sons qui enclenchent des processus de traitement alphabétique pour décoder et encoder les mots. Avec une exposition croissante à l'écrit, ils développent des processus de traitement orthographiques qui permettent de reconnaître et reproduire de plus grandes unités sublexicales, puis, en fin de compte de traiter les mots de façon globale et instantanée, c'est-à-dire faire un lien direct entre les formes orales et écrites. La compétence est admise lorsque le traitement orthographique des mots est rapide et précis. À ce niveau de compétence, on a recours au traitement alphabétique que pour traiter les mots nouveaux ou rarement rencontrés, c'est-à-dire ceux qui respectent les conventions phonologiques stockées en mémoire. Dans cette modélisation, les mots irréguliers ou inconsistants ne peuvent être reconnus au début du processus d'acquisition puisqu'ils requièrent un traitement orthographique qui se développe vers la fin du processus.

Cela soulève la question de l'âge ou du niveau scolaire associé à l'émergence du traitement orthographique. Frith, Seymour et leurs collaborateurs respectifs ont conclu que le traitement orthographique était présent après 2 ans de scolarisation, vers l'âge de 7 ans. Dans une étude, Snowling et Frith (1981) ont noté que les bons lecteurs de cet âge avaient plus de difficultés à lire des textes qui avaient des distorsions orthographiques (ex : *b^cx** au lieu de *box*) que des textes où les distorsions étaient phonologiques (ex : *g^{aiv}** au lieu de *gave*) ou de nature purement visuelle (ex : *g^{AVE}*). Elles ont conclu que le traitement orthographique était présent, car les distorsions qui le touchaient avaient le plus d'effet sur leur identification. Dans la seconde, Seymour et Evans (1999) ont établi que la structure fondatrice du cadre orthographique, c'est-à-dire les premières représentations spécifiques aux mots, serait

fonctionnelle à la fin de la 2^e année, vers l'âge de 7 ans. Le traitement orthographique serait fonctionnel, mais il resterait moins efficace que le traitement alphabétique jusqu'au milieu du 2^e cycle, selon d'autres études menées par Seymour et ses collaborateurs (Duncan et Seymour, 2000; Duncan, Seymour et Hill, 1997; Seymour, Duncan et Bolik, 1999) qui examinaient l'effet de la conscience phonologique sur la taille des unités lexicales traitées par des enfants de différents niveaux scolaires. Lors de tâches d'identification d'unités linguistiques communes dans des paires de mots, Duncan, Seymour et Hill (1997) ont observé que les enfants du premier cycle réussissaient mieux l'appariement d'unités en début de syllabes qu'à la fin des syllabes, qu'elles soient simples (man-mat) ou complexes (cloth-clock). De plus, la rime s'est avérée particulièrement difficile (< 15% de réussite), qu'elle soit simple (mat-sat) ou complexe (paste-waste). Ces résultats ont été reproduits auprès de populations différentes. En appariant les âges de lecture, ces résultats ont été reproduits chez les enfants dyslexiques (Seymour, Duncan et Bolik, 1999; Seymour et Evans, 1999) et les enfants de milieux défavorisés (Duncan et Seymour, 2000). Seymour et Evans (1999) ont conclu que le noyau du cadre orthographique (founding structure) était fonctionnel à la fin de la 2^e année, mais que ce n'était qu'à l'âge de lecture de huit ans et demi que les enfants atteignaient des scores similaires dans le traitement des unités de petite et de grande taille ($\pm 65\%$).

Il est toutefois important de noter que dans les études menées par Seymour et Duncan, les stimuli étaient tous à très haute fréquence et qu'aucune de ces études ne s'est penchée sur les processus de traitement utilisés et sur l'âge auquel les enfants traiteraient des unités de grande taille dans les mots moyennement ou peu fréquents. Prenons l'exemple de deux mots monosyllabiques visuellement, phonologiquement et orthographiquement similaires, *sont* et *mont*. Selon les indices de fréquence de Manulex, un enfant de 1^{re} année croiserait *sont* une fois à tous les 1000 mots (SFI¹² de 73,32) et *mont* une fois à tous les 10 millions de mots (SFI de 31,86). Le mot *sont* risque donc d'être rencontré dix-mille fois plus souvent que le mot *mont*. À la lumière des difficultés observées par Seymour et ses collaborateurs dans le traitement de la rime au premier cycle, et même si l'on alloue des différences de fréquences attribuables à une base de données française plutôt que québécoise, il semble peu probable

¹² SFI = Standard Frequency Index

que ces deux mots soient traités de la même façon lorsque le traitement orthographique est en voie de développement.

Les modèles d'acquisition des processus d'identification et de production de mots présentés ci-dessus sont pertinents à la réalisation de cette étude puisqu'ils proposent une période développementale à laquelle les connaissances morphologiques émergeraient. Cela permet de faire des liens avec notre premier objectif, celui d'observer le développement des connaissances morphologiques liées à la production de marqueurs muets en position finale. Cependant, même si les modèles développementaux élucident les processus cognitifs actifs dans le traitement de mots morphologiques, l'impact de la fréquence d'occurrence des mots sur le processus d'acquisition n'est pas considéré. Par ailleurs, les observations empiriques de plusieurs chercheurs réfutent la prémisse du déploiement à stades, ou du développement tardif du traitement morphologique. Certains ont observé des représentations orthographiques (Martinet, Valdois et Fayol, 2004; Pacton et coll., 2001; Treiman et Cassar, 1996) et morphologiques (Bégin, Saint-Laurent et Giasson, 2010; Sénéchal, 2000; Sénéchal et Kearnan, 2007; Snowling, 1994; Treiman, 1993) dès le début du parcours scolaire. Certains ont aussi observé que le traitement morphologique lié à la connaissance des familles morphologiques se développe dès le 1er cycle du primaire. (Pacton et coll., 2001; Sénéchal, 2000; Sénéchal, Basque et Leclaire, 2006) Cela donne ainsi raison au besoin de théories alternatives pour modéliser le développement de l'écrit.

2.4 Les théories et modèles *usage-based* en acquisition de l'écrit

Vers la fin des années 1980, parallèlement à la remise en question des modèles à stades, « malmenés par les résultats d'études empiriques » (Share, 1995), d'autres modèles et théories d'acquisition ont fait leur apparition, mettant en évidence l'importance de la fréquence d'exposition aux mots sur le développement des représentations lexicales. Une des premières théories du genre a été la théorie des effets de Matthieu¹³ en lecture de Stanovich (1986), que ce dernier décrit comme un cadre conceptuel pour expliquer le développement du langage écrit. Il a postulé que l'habileté en lecture a un lien causal réciproque avec les

¹³ Le nom est tiré de l'évangile selon saint Matthieu, chapitre XXV:29 : « Car on donnera à celui qui a, et il sera dans l'abondance; mais, à celui qui n'a pas, on enlèvera même ce qu'il semble avoir. »

habiletés cognitives nécessaires à la lecture efficace. Cette réciprocity engendre des effets de « bootstrapping »¹⁴, c'est-à-dire des effets d'autoapprentissage. Cette théorie voulant que « les riches s'enrichissent et les pauvres s'appauvrissent » affirme que plus on lit, plus on enrichit son lexique (vocabulaire) et plus ce lexique individuel est riche, plus on a des expériences enrichissantes, ce qui nous incite à lire davantage. Dans cette synthèse de la recherche en lecture, Stanovich conclut que la lecture favorise le développement de la conscience phonologique, l'utilisation efficace de contexte, le vocabulaire et la compréhension et, qu'inversement, ces habiletés cognitives facilitent la lecture. Cet engrenage des effets de la lecture trouve sa perversion biblique chez les enfants qui, pour des raisons cognitives et environnementales, n'accumulent que peu de ces expériences enrichissantes et prennent de plus en plus de retard par rapport aux enfants qui lisent beaucoup.

Depuis Stanovich (1986), plusieurs chercheurs ont modélisé l'évolution de l'écrit à l'aide de théories basées sur la fréquence d'exposition aux mots. Quelques-uns seront présentés dans la prochaine section.

Il n'y a pas de doute que la fréquence d'exposition à l'écrit est un facteur déterminant dans son apprentissage. Les observations empiriques à cet effet abondent (voir Ellis 2002 pour un compte rendu des effets de fréquence sur les différents processus langagiers). Contrairement aux protagonistes des modèles développementaux qui postulent que la genèse de l'acquisition de l'écrit s'explique par des stades ou phases de développement cognitif des sujets, plusieurs chercheurs l'expliquent par la fréquence d'exposition aux items lexicaux (Brown et Loosemore, 1995; Ellis, 2002; Harm et Seidenberg, 2004; McClelland et coll., 1986; Monaghan et Ellis, 2010; Perfetti, 1991; Plaut et coll., 1996; Rumelhart et McClelland, 1987; Seidenberg et McClelland, 1989; Share, 1995). Deux types de modèles seront considérés, les modèles intégratifs et les modèles connexionnistes.

¹⁴ Définition de bootstrapping : « a self-sustaining process that proceeds without external help. In psycholinguistics bootstrapping refers to the question of how "gets started." [...] Bootstrap as a metaphor, meaning to better oneself by one's own unaided efforts, [...] a series of self-sustaining processes that proceed without external help [...] early knowledge would serve as bootstraps by which the child pulls himself or herself up (Wikipedia, 2011)

2.4.1 Les modèles intégratifs

Dans un modèle dit intégratif, le développement de la lecture et de l'écriture est interactif et réciproque (Lété, Peereman et Fayol, 2008). Dans son modèle RI (Restrictive-Interactive Model), Perfetti (1991) postule que l'interaction entre les items lexicaux (lettres, phonèmes et mots) lors de situations de lecture forme le lexique orthographique individuel. Selon ce modèle, il y a deux constituants de base au lexique orthographique : un lexique fonctionnel et un lexique autonome. Le lexique fonctionnel se développe à partir des unités sublexicales reconnues, puis utilisées pour identifier les mots. Le lexique autonome se compose des mots qui sont reconnus de façon autonome, c'est-à-dire sans avoir recours aux unités sublexicales du lexique fonctionnel. Ces deux éléments constitutifs s'apparentent au décodage et à la reconnaissance globale des mots que l'on retrouve dans les modèles cognitifs, mais Perfetti fait une distinction importante : « ce sont les items lexicaux, et non les enfants, qui deviennent autonomes » (1991, p.35). Chaque instance de correspondance précise (correspondance juste) et redondante (propre au mot) entre les représentations visuelles et sonores alimente le lexique orthographique. Des items lexicaux simples sont d'abord reconnus à l'aide de processus phonologiques (correspondances lettres-phonèmes) et stockés dans le lexique fonctionnel et l'on y fait appel pour se faire une représentation des mots. La taille des unités graphiques lexicalisées prend ensuite de l'expansion (digrammes, attaque/rime, syllabes, morphèmes et mots) et chaque occurrence de décodage agit sur la vitesse et la précision des représentations phonologiques. La fréquence avec laquelle les mots et les items sublexicaux sont rencontrés agit sur la vitesse avec laquelle les mots sont représentés dans le lexique autonome.

L'hypothèse d'autoapprentissage de Share (1995), qui se rapproche du modèle de Perfetti, suppose que chaque instance de recodage phonologique autonome (réussie) est une instance d'autoapprentissage orthographique, ce qui rejoint aussi l'hypothèse du « bootstrapping » de Stanovich. Ces instances agissent sur la formation du lexique orthographique et ce sont les items phonologiques qui s'autonomisent plutôt que les processus développementaux des lecteurs. Le modèle de Share diverge des théories développementales quant au rôle des processus orthographiques. Selon cet auteur, ce sont uniquement les processus phonologiques qui constituent le mécanisme pour traduire les lettres en mots. Les processus orthographiques

ne sont pas un mécanisme d'identification de mots, mais un dispositif cognitif qui permet de mettre en mémoire les représentations propres aux mots. Il postule que cette lexicalisation dépend de relativement peu d'instances de décodage phonologique autonome donc, dès un jeune âge, les mots à haute fréquence peuvent être lexicalisés, peu importe leur régularité orthographique.

Quel que soit le modèle, développemental ou intégratif, les protagonistes de modèles décrits verbalement¹⁵ présentés ci-dessus semblent généralement s'entendre sur le fait qu'il y a deux processus cognitifs qui permettent de reconnaître les mots et de les mettre en mémoire. On peut alors se demander par quel processus on traite les consonnes muettes terminales. Contrairement à l'anglais, la grande majorité des consonnes qui se trouvent en position finale ne sont pas prononcées en français, peu importe s'ils ont un rôle morphosyntaxique (tu aimais) ou non (mais). N'ayant pas d'équivalent phonologique, on peut croire que le *s* du mot, *mais* serait traité uniquement par des processus orthographiques, ou visuels. Si tel est le cas, la lettre muette terminale ne serait représentée qu'au moment où le mot entier est lexicalisé. Ainsi, les jeunes enfants qui rencontrent la graphie terminale *-ais* dans un mot nouveau la prononceraient systématiquement /es/ plutôt que /ɛ/. Cependant, il a été démontré que dès la 1^{re} année, les enfants sont sensibles aux contraintes graphotactiques, c'est-à-dire à la position que les lettres peuvent occuper dans les mots (Pacton, Fayol et Perruchet, 2002; Pacton et coll., 2001; Treiman, 1993). Par exemple, sans qu'on leur ait explicitement enseigné, les enfants francophones produisent le graphème *-eau* pour représenter le phonème /o/ en position terminale (Pacton et coll., 2002). Il est donc possible que les consonnes soient traitées différemment selon la position qu'elles occupent dans un mot. Par exemple, si la terminaison *-ais* est fréquemment rencontrée au début de l'exposition à l'écrit, il semble probable qu'elle soit lexicalisée comme un trigramme, de la même façon dont on traite le graphème complexe « eau ». Si tel est le cas, les apprentis lecteurs qui rencontrent le *-ais* à la fin d'un nouveau mot le prononceront /ɛ/ plutôt que /es/.

¹⁵ Modèles dont les caractéristiques et mécanismes sont théorisés, mais n'ont jamais été observés en soi. Expression, « verbally described », tirée de Brown et Loosemore, 1995

Peu importe les processus activés pour traiter les consonnes muettes, si l'on accepte l'hypothèse des modèles intégratifs que ce sont les items qui s'autonomisent et non les processus cognitifs des enfants, la qualité de la représentation que l'on a pour ces consonnes serait tributaire de la fréquence avec laquelle elles sont rencontrées. Ainsi, plus souvent les unités sublexicales terminales qui comprennent une lettre muette sont rencontrées, plus il y a de chances qu'elles soient reconnues instantanément. Plus les mots qui se terminent par ces unités sont rencontrés souvent, plus il y a de chance qu'ils soient reconnus par un processus de traitement orthographique.

2.4.2 Les modèles connexionnistes

Au début des années 1980, des modèles informatiques sont apparus. Selon leurs architectes, ces modèles permettaient de vérifier les hypothèses des modèles décrits verbalement en reproduisant les comportements qu'ils décrivent. Des modèles connexionnistes ont d'ailleurs réussi à reproduire l'acquisition de la lecture orale, l'acquisition de régularités orthographiques, ainsi que certains aspects de la morphologie flexionnelle, avec une précision appréciable. Avec la puissance computationnelle des ordinateurs, il était dorénavant possible de reproduire les comportements humains. « [...] children, when asked to spell the same words that the model has learned to spell, exhibit the same type of difficulty on the same word type as does the model. » (Brown et Loosemore, 1995, p.319)

La description architecturale des modèles connexionnistes est au-delà de la portée de cette recherche (voir Plaut, 1999b pour un résumé de la modélisation connexionniste), mais il suffit de savoir que ce sont des systèmes computationnels qui émulent le réseau neuronal du cerveau humain à partir d'occurrences d'évènements lexicaux. Outre leurs avantages à reproduire et, par extension, à prédire certains comportements langagiers, ils pouvaient servir à valider ou réfuter les postulats des modèles développementaux et interactifs. Par exemple, l'axiome des deux voies de traitement a été mis à l'épreuve. Certains chercheurs l'ont réfuté (Brown et Loosemore, 1995; Harm et Seidenberg, 1999; McClelland et Rumelhart, 1981; Rumelhart et McClelland, 1986), d'autres l'ont appuyé (Coltheart et coll., 2001; Coltheart et coll., 1993; Pacton, 2001; Pinker et Prince, 1988).

McClelland et ses collaborateurs ont été des innovateurs dans la modélisation connexionniste « à une voie » en reproduisant des comportements en lecture orale chez les lecteurs compétents (McClelland et Rumelhart, 1981; Plaut, 1999; Seidenberg et McClelland, 1989). Harm et Seidenberg (1999) ont simulé le rôle des représentations phonologiques dans l'acquisition de la lecture chez les normolecteurs et chez les dyslexiques. Brown et Loosemore (1995) ont reproduit le comportement d'enfants dans l'apprentissage orthographique de mots monosyllabiques phonologiquement réguliers, irréguliers et uniques. Comme les enfants, le modèle a mieux réussi les mots réguliers, c'est-à-dire ceux avec seulement des amis phonologiques (correspondances lettres-sons identiques), que les mots uniques (sans voisins phonologiques), qui étaient mieux réussis que les mots irréguliers, c'est-à-dire ceux qui avaient des « ennemis » phonologiques (correspondances lettres-sons compétitives). En acquisition de la morphologie flexionnelle, Rumelhart et McClelland (1986, 1987) ont émulé l'acquisition morphologique du passé des verbes chez les apprenants de l'anglais, langue première (L1). Leur modèle a reproduit les 3 étapes d'acquisition du passé, c'est-à-dire la courbe d'apprentissage en « U » observée chez les enfants (voir Rumelhart et McClelland, 1986, p.538-539 pour une description complète).

Le succès de ces modèles connexionnistes à une voie a donc mené leurs protagonistes à remettre en question les théories appuyant deux voies d'identification de mots. Cependant, ils ont aussi leurs limites et leurs détracteurs (Coltheart et coll., 1993; Pinker et Prince, 1988). Coltheart et ses collaborateurs (1993) ont argué que les modèles à voie unique n'arrivaient pas à simuler l'acquisition de l'oral et du passé des verbes avec la même précision que l'on observe chez les apprenants. Par exemple, selon ces derniers, le modèle de Rumelhart et McClelland (1986) ne parviendrait pas à acquérir les formes fléchies de pseudoverbes avec la même précision que les apprenants ou un modèle à deux voies, comme le leur. Le modèle DRC (Dual Route Cascaded Model) de Coltheart et ses collaborateurs (1993) semblerait avoir réussi à émuler la lecture orale de mots et de non-mots avec plus d'authenticité que les modèles connexionnistes à une voie. Plus récemment, grâce à ce modèle à deux voies, Coltheart et ses collaborateurs (2001) ont reproduit une multitude d'effets observés en lecture orale chez les adultes, dont des effets de fréquence des mots et des effets de fréquence de la

racine lexicale dans la lecture de pseudohomophones (voir p. 251 pour la liste complète des effets reproduits).

Finalement, Pacton et ses collaborateurs (2001) ont observé que leur modèle SRN (Simple Recurrent Network) ne parvenait pas à acquérir les régularités orthographiques avec la même précision qu'ils avaient observée chez les enfants francophones. Le modèle, comme les enfants, devient de plus en plus sensible à la fréquence des consonnes qui peuvent être doublées, à la position qu'elles peuvent occuper dans les mots et au fait que les voyelles ne se doublent jamais. Cependant, dans une tâche de décision lexicale de non-mots, leur modèle n'atteignait jamais la performance des enfants (56.4 % vs. 76.2 %). Ils en conclurent que la différence reflétait sans doute les connaissances phonologiques des enfants, de l'information qui n'était pas fournie au modèle, appuyant l'axiome qu'il y a deux processus de traitement de l'écrit.

Qu'ils soient à une ou à deux voies, les modèles informatiques ont reproduit des comportements observés chez les apprenants avec une précision importante. Qu'ils soient interactifs ou informatisés, les modèles *usage-based* ont consolidé le rôle de la fréquence dans l'acquisition du langage. Ils ont aussi mis en évidence l'importance de contrôler la fréquence lorsque l'on évalue les habiletés qui marquent l'acquisition du langage écrit, et évaluer le rôle de la fréquence dans la production marqueurs morphologiques muets est un des objectifs de cette recherche.

2.5 Les bases de données lexicales

Comme nous venons de le voir, l'exposition à l'écrit est un facteur fondamental en lecture et en écriture. Ainsi, la fréquence avec laquelle les mots apparaissent dans les écrits est une variable explicative dans le traitement de la langue écrite (Content et coll., 1990; Desrochers, 2006; Lété, 2005). La fréquence d'occurrence doit ipso facto être contrôlée lors d'expérimentations qui tentent d'élucider le développement et le traitement de l'écrit. Outre la nécessité de contrôler la fréquence, le désir de comprendre et d'expliquer les effets de fréquence des mots et de leurs unités sublexicales dans le traitement et l'acquisition de l'écrit a mené à la création de bases de données lexicales qui, entre autres, recensent les mots sur

une base distributionnelle, c'est-à-dire sur la base de leur fréquence d'occurrence dans les écrits.

En français, la première base de données lexicale informatisée à être mise à la disposition des chercheurs, Brulex (Content et coll., 1990), était constituée de plus de 30 000 entrées retrouvées dans le dictionnaire *Le Micro-Robert* et tirait ses mesures de fréquences des mots du *Trésor de la langue française* (Imbs, 1971). Plus récemment, on peut avoir recours à des bases de données beaucoup plus sophistiquées pour mesurer non seulement la fréquence des mots, mais aussi la fréquence de leurs constituants sublexicaux et les relations entre les unités lexicales. Par exemple, Lexique 3 (New et Pallier, 2005) comprend plus de 135 000 mots et 55 000 lemmes, tirant ses fréquences de livres, mais aussi de pages Web et de sous-titrage de films. Cette base de données comptabilise, en plus des caractéristiques catégorielles (classe, genre) et des caractéristiques structurales (nombre de phonèmes, lettres, syllabes), des caractéristiques relationnelles (nombre de voisins orthographiques et phonologiques, famille flexionnelle) et distributionnelles (fréquence cumulée des lemmes; fréquence des lettres, phonèmes, digrammes, trigrammes, syllabes, formes fléchies). Dans la base de données Omnilex (Desrochers, 2006), développée au Canada, en plus des caractéristiques catégorielles, structurales, relationnelles et distributionnelles comprises dans les autres bases, on recense des fréquences subjectives (telles qu'elles sont perçues par un échantillon de la population) et des caractéristiques sémantiques (catégorie sémantique, la typicité catégorielle, la valeur d'imagerie ou de concrétude (sic) et la valence émotionnelle) de plus de 96 000 noms, adjectifs, verbes et adverbes.

Bien que la majorité des bases de données soient développées à partir de textes destinés aux adultes, il y en a qui établissent leurs données à partir de corpus pour enfants. Novlex (Lambert et Chesnet, 2001) a été développée à partir de 38 livres scolaires et extrascolaires destinés aux enfants français de 3^e année (CE2). Il comprend un peu plus de 20 000 entrées qui représentent 9 600 racines lexicales. Les champs qu'on y retrouve comprennent, entre autres, le nombre de lettres, de phonèmes et de syllabes, le découpage et la structure syllabique, la classe grammaticale et le genre, et la présence d'un « e » muet ou d'un doublet.

Une deuxième base de données normalisée en France et qui vise l'étude de l'écrit chez les enfants est Manulex (Lété et coll., 2004). Elle fournit les fréquences d'occurrences de plus de 48 000 formes orthographiques et 23 000 lemmes catalogués à partir du recensement de 54 manuels scolaires français du primaire de la première à la 5^e année du primaire (CP à CM2). Les indices de fréquence ont été calculés pour trois catégories de lecteurs : G1 (CP ou 1^{re} année), l'âge auquel « se construit le lexique de l'enfant sur la base de la médiation phonologique »; G2 (CE1 ou 2^e année, l'âge auquel « se construit le lexique orthographique par automatisation progressive de la reconnaissance du mot écrit »; G3-5 (CE2 à CM2 ou la troisième à la 5^e année), l'âge où « se consolide et s'enrichit le stock lexical par exposition répétée à l'écrit » (Lété, 2005). Les auteurs expliquent que le vocabulaire utilisé dans les livres qui visent les enfants de première et de 2^e année diffère l'un de l'autre, et aussi de celui des textes qui visent les enfants des cycles supérieurs. Cependant, bien que le vocabulaire varie quantitativement après la 3^e année, il ne varie pas qualitativement. En autres mots, la probabilité d'exposition aux mots ne change pas d'une année à l'autre, entre le début du 2^e cycle et la fin du 3^e cycle. Cela justifie le regroupement des mesures de fréquence visant ces élèves plus âgés (Lété et coll., 2004).

La plus récente version Manulex-infra (Peereman, Lété et Sprenger-Charolles, 2007) est particulièrement intéressante, car elle offre des indices de fréquence par item (token count) et par type d'items (type count), ainsi que des indices de fréquence lexicaux et sublexicaux. Par exemple, il est possible de recenser les lettres qui sont le plus fréquemment trouvées à la fin des mots.

À défaut de pouvoir recenser la fréquence d'exposition lexicale de chaque sujet d'une recherche, les bases de données lexicales offrent aux chercheurs des données normalisées pour sélectionner les stimuli expérimentaux et mesurer les effets de fréquence dans le traitement du mot écrit. Des effets ont été observés en lecture et en orthographe, mais aussi dans toutes les sphères de la linguistique (voir Ellis 2002 pour un résumé des effets de fréquence dans le traitement du langage).

Il y a de plus en plus d'études qui se sont penchées sur les effets de fréquence en orthographe lexicale, grammaticale et morphosyntaxique. Nous présentons ci-dessous les résultats

d'études qui tentent de mettre en lumière les effets de la fréquence dans le développement orthographique des enfants du primaire.

2.6 Les effets de fréquence en production de mots

La sensibilité des enfants à la fréquence des mots et de leurs unités sublexicales a été observée dès leurs premiers mois sur les bancs d'école. Des chercheurs ont observé que les enfants du 1er cycle sont sensibles aux contraintes graphotactiques (des contraintes combinatoires et distributionnelles des lettres dans les mots) (Pacton et coll., 2001; Treiman, 1993), à la régularité orthographique et la consistance des correspondances phonographémiques (Alégria et Mousty, 1997; Lété et coll., 2008; Martinet et coll., 2004; Pacton et coll., 2001; Snowling, 1994; Sprenger-Charolles, Siegel et Bonnet, 1998) et à la fréquence d'occurrence dans l'acquisition de la morphologie flexionnelle et dérivationnelle (Bégin, Saint-Laurent et Giasson, 2010; Largy, Fayol et Cousin, 2004).

Dans des tâches de décisions lexicales et de dictées trouées, Pacton et ses collaborateurs (2001) ont constaté que les enfants francophones de la première à la 3^e année sont sensibles aux contraintes graphotactiques, c'est-à-dire la fréquence avec laquelle les lettres apparaissent, aux lettres qui peuvent ou non être doublées, ainsi qu'à la position « légale » des doublets. Un effet de fréquence de lettres simples est apparu dès la 1^{re} année, tandis que des effets de fréquence de doublets, de la position « légale » qu'ils occupent et d'incohérence (doubler des voyelles) sont apparus à partir de la 2^e année.

Les effets de fréquence ont aussi été observés au niveau lexical et sublexical. Pacton, Fayol et Perruchet (2002) ont observé, dès la 2^e année, des effets de fréquence et de lexicalité (CPG et régularités orthographiques) dans la production orthographique de non-mots. Dans l'ensemble, pour orthographier le phonème /o/, les choix orthographiques des enfants reflétaient les caractéristiques fréquentielles et distributionnelles des graphèmes, c'est-à-dire *o* et *au* étaient le plus souvent utilisés en position initiale et médiane, et *-eau*, en position finale. Dans une étude de régression à grande échelle, Lété, Peereman et Fayol (2008) ont observé des effets de fréquence lexicale (mots et voisins orthographiques) et de consistance des CPG (feedforward

consistency) dans l'orthographe de 3430 mots produits par 200 enfants de la première (CP) à la 5^e année (CM2). Dès la 1^{re} année, des effets de fréquence des mots et des effets liés aux propriétés statistiques des CPG (dorénavant appelés effets phonologiques) sont observés. Les effets phonologiques sont moins significatifs pour les mots fréquents que pour les mots rares et sont observés à deux niveaux : phonémique et syllabique. L'effet syllabique n'est cependant présent qu'à partir de la 2^e année, au moment où les enfants développent des connaissances orthographiques propres aux mots. Ceci semble rejoindre la théorie de Perfetti (1991) voulant que les instances de correspondance entre les représentations visuelles et sonores augmentent en taille avec l'usage, et que la redondance alimente le lexique orthographique autonome (effet de fréquence) pendant que les processus phonologiques restent actifs pour traiter les mots qui ne sont pas automatisés.

Des effets de fréquence ont aussi été observés dès la 1^{re} année dans l'orthographe de mots irréguliers. Dans une étude longitudinale, Sprenger-Charolles, Siegel et Bonnet (1998) ont testé la production, en janvier, puis en juin, de mots simples (correspondances lettres-sons), complexes (correspondances graphèmes-phonèmes) et phonologiquement irréguliers (ex. *piéd, femme, scie, næud*), ainsi que de pseudomots. Les effets de fréquence et de lexicalité n'étaient pas présents en janvier, mais sont apparus dès le mois de juin. Les auteurs ont conclu que « le changement dans les effets de fréquence et de lexicalité entre janvier et juin démontre que les enfants peuvent rapidement se construire un lexique orthographique » (traduction libre p. 134). Plus récemment, Martinet, Valdois et Fayol (2004) ont observé des effets de fréquence dans la production de mots phonologiquement irréguliers (composés de graphèmes non dominants, ex : *sirop*) après seulement 3 mois d'enseignement formel. Les mots irréguliers fréquents étaient mieux orthographiés que les mots irréguliers rares.

Contrairement aux études ci-dessus, Alegria et Mousty (1996) n'ont pas observé d'effets de fréquence chez les plus jeunes scripteurs dans la production de mots irréguliers. Dans une tâche de production orthographique de pseudomots ayant des graphonèmes non dominants (ex. *c* fait /s/ au début du mot *citron*) et contextuels (ex. la lettre *c* dans *ca* vs *ce*), les effets de fréquence lexicale entre les mots fréquents et rares ne sont apparus qu'à partir de la 3^e année. Cependant, dans une étude similaire subséquente (Alégria & Mousty, 1997), ils ont

observé un effet de fréquence dès la 2^e année. Dans cette dernière étude, le graphème non dominant « ain », était significativement mieux réussi dans les mots fréquents que dans les mots rares. L'effet tardif noté dans la première étude pourrait s'expliquer par la différence dans le niveau de consistance des unités lexicales dans les mots. Le nombre de représentations phonographiques à gérer pour traiter les graphonèmes contextuels ou inconsistants rend leur acquisition plus difficile. Si l'on prend l'exemple de la lettre *c*, il faut savoir quel phonème elle représente selon son contexte, /s/ pour *ce, ci, cy* et /k/ pour *ca, co, cu, cr, cl* et à la fin des syllabes fermées (*ac/tion*). De plus, en situation d'écriture, il y a plusieurs choix pour orthographier les phonèmes /k/ (*c, qu, k* et *ch*) et /s/ (*s, c, ss, ç* et *t*). Toute l'information qu'il est nécessaire de gérer pour produire correctement la lettre *c* dans des graphonèmes contextuels ou inconsistants, fait en sorte que les stimuli d'Alegria et Mousty (1996) sont, par définition, beaucoup plus complexes à maîtriser que les CPG régulières (simples et complexes) de Sprenger-Charolles et ses collaborateurs (1998) ou le graphème *ain* d'Alegria et Mousty (1997). Le *ain* est toujours prononcé /ɛ̃/ et se retrouve dans des mots facilement « imageables » et qui risque d'être rencontrés tôt dans les livres de jeunesse (ex. *main, pain, train*).

Finale­ment, des effets de fréquence ont aussi été observés dans l'étude de l'acquisition de la morphologie flexionnelle et dérivationnelle (Bégin, Saint-Laurent et Giasson, 2010; Largy et coll., 2004). Dans une étude longitudinale, Largy, Fayol et Cousin (2004) ont observé que la forme flexionnelle rencontrée la plus fréquemment influence l'accord des noms chez les enfants du 1er cycle (CP/CE1). Les sujets ont rencontré pour la première fois 40 mots expérimentaux très rares (absents de bases de données lexicales, mais facilement « imageables » - noms d'animaux, de plantes, etc.) lors d'un prétest au mois de décembre de la 1^{re} année de l'étude. Ils ont ensuite participé à deux phases d'entraînement, en mars et en juin, puis ont été soumis à des évaluations tout de suite suivant les entraînements. Ils ont subséquemment été testés en décembre et en mars de l'année suivante. Durant les phases d'entraînement, les mots ont été manipulés pour varier la condition de présentation : une moitié était présentée au singulier et l'autre au pluriel; une moitié présentée fréquemment et l'autre rarement. Les mots fréquents ont fait l'objet d'exercices de lecture deux fois par

semaine et d'écriture une fois par semaine pendant deux trimestres. Les mots rares ont fait l'objet d'un seul exercice de copie par trimestre.

Les résultats ont démontré une dominance de la forme flexionnelle la plus fréquente et ce, avant l'enseignement de la règle d'accord (1^{re} année) et suivant son enseignement (2^e année). Les mots fréquents plus souvent rencontrés au pluriel ont été mieux réussis au pluriel qu'au singulier. Les mots fréquents plus souvent rencontrés au singulier étaient moins bien réussis au pluriel que les mots plus souvent rencontrés au pluriel. Cette distinction était cependant absente pour les mots rares. Toutefois, la dominance de la forme flexionnelle n'a pas été observée à long terme, mais elle a persisté plus longtemps chez les plus jeunes enfants. L'effet de fréquence a été observé en décembre de l'année suivante pour les enfants initialement testés en 1^{re} année, mais l'effet était absent en mars. Aucun effet n'est apparu lors des épreuves de l'année suivante pour les enfants originalement testés en 2^e année. Largy, Cousin et Fayol (2004) se sont demandé si la forme lexicale globale était mieux représentée chez les plus jeunes enfants ou si cette différence serait l'effet de l'âge d'acquisition, c'est-à-dire que les mots appris plus précocement sont reconnus plus rapidement et avec plus de précision que les mots appris plus tardivement. L'âge d'acquisition a été observé chez les adultes et est étroitement lié à la fréquence d'occurrence. Nous y reviendrons plus loin.

Dans une étude récente sur le traitement morphologique chez les élèves québécois de la 6^e année, Bégin, Saint-Laurent et Giasson (2010) ont observé des effets de fréquence dans l'orthographe de mots ayant une consonne muette finale. À la dernière année du primaire, les élèves produisaient davantage la consonne muette des mots fréquents que celle des mots rares, qu'ils soient morphologiques (lettre muette pouvant être déduite par l'application d'une règle dérivationnelle) ou lexicaux (mots non dérivables). Les résultats de Bégin, Saint-Laurent et Giasson diffèrent des observations de Sénéchal (2000) qui n'avait pas observé d'effet de fréquence d'exposition aux mots dans l'orthographe de lettres muettes chez les enfants plus jeunes. L'âge, le niveau scolaire et le niveau de lecture pourraient en être la cause, mais il est important de noter que le contrôle de la fréquence dans l'étude de Sénéchal était basé sur une fréquence subjective, tandis que Bégin, Saint-Laurent et Giasson ont utilisé

les mesures de fréquence standardisées de Manulex. Nous reviendrons sur les observations de ces deux recherches lorsque nous traiterons du développement morphologique.

Les observations empiriques présentées ci-dessus appuient le postulat que le développement des représentations lexicales des jeunes enfants est tributaire de la fréquence des intrants orthographiques. Des effets de fréquence d'occurrence et de régularité ou de consistance orthographique ont été observés dès le 1er cycle du primaire. Par ailleurs, lors de la construction de leur lexique orthographique, ces jeunes enfants ont démontré une capacité à produire de nouveaux mots en faisant des liens analogiques avec des mots déjà rencontrés.

2.6.1 La fréquence et l'analogie

Dans son étude auprès des enfants de six à huit ans, Snowling (1994) a été une des premières à remettre en question l'hypothèse qu'en début de parcours, les enfants n'ont recours qu'aux processus phonologiques pour produire les mots. Elle observe que, dès l'âge de huit ans, l'orthographe de groupes consonantiques varie selon leur contexte. Les enfants épèlent des non-mots à partir d'indices dans les mots amorces et en absence d'amorces, ils utilisent les graphèmes les plus fréquents. Elle conclut que, même au début du développement orthographique, les jeunes enfants utilisent une variété de sources pour orthographier les mots, incluant des processus analogiques. Les résultats en production de non-mots dans les études de Lété et ses collaborateurs (2008), de Sprenger-Charolles et ses collaborateurs (1998) et d'Alegria et Mousty (1996, 1997) corroborent les conclusions de Snowling : les enfants ont recours à l'analogie pour orthographier des mots nouveaux. Martinet, Valdois et Fayol (2004) ont fait ces mêmes observations avec des enfants ayant à peine trois mois de scolarisation. Ces études soutiennent l'hypothèse que l'orthographe de jeunes écoliers est influencée par la fréquence des mots et par la régularité de leurs caractéristiques sublexicales, mais aussi que la fréquence d'exposition aux unités lexicales agit sur l'activation de processus analogiques lors de la production de mots nouveaux.

Une des caractéristiques fréquentes et régulières du système orthographique du français est la présence de consonnes muettes terminales. La production de ces marqueurs morphologiques muets chez les enfants pourrait ainsi se faire précocement par traitement orthographique,

emmagasinant l'orthographe globale en mémoire à la suite d'une exposition répétée au mot lui-même. Cependant, à la lumière des effets de fréquence précoces observés dans les mots qui ne suivent pas les conventions phonologiques, on peut se demander si, dès un jeune âge, les enfants peuvent produire ces consonnes muettes par analogie morphologique, c'est-à-dire en associant le mot à sa famille morphologique pour justifier la lettre muette (chaudu/chaude/chaudement). Nous reviendrons sur cette question plus loin lorsque l'acquisition des connaissances morphologiques sera abordée.

En résumé, la fréquence d'exposition aux mots, à leurs constituants sublexicaux et au contexte de ces constituants agit sur la qualité des représentations orthographiques et, par extension, sur le développement de processus de traitement analogiques pour justifier l'orthographe d'un mot. Il est donc pertinent d'évaluer l'effet de la fréquence sur la réussite orthographique des mots dont la consonne muette finale est dérivable et si le contexte sublexical agit sur la représentation qu'ont les enfants de consonnes muettes finales. On peut aussi se demander si la fréquence de la lettre muette elle-même (LM) et la structure syllabique dans laquelle elle se trouve (CV vs. CVC) agissent sur sa production.

Certains chercheurs se sont penchés sur des variables explicatives de l'effet de fréquence dans le développement des habiletés liées au langage écrit. Il est au-delà de la portée de cette recherche d'aborder ses variables explicatives en détail ou de les tester dans notre expérimentation. Nous en présentons cependant les grandes lignes.

2.7 L'âge d'acquisition, la fréquence cumulative et la trajectoire fréquentielle

L'âge d'acquisition (AoA), c'est-à-dire l'ordre dans lequel les mots sont rencontrés, a été avancé comme variable explicative pour les effets de fréquence observés dans les études empiriques. Les mots appris plus tôt seraient mieux reconnus que les mots appris plus tard. Des effets d'âge d'acquisition ont été observés dans une multitude de tâches d'identification, de production et de décision lexicale chez les adultes (Bonin et coll., 2004; Bonin et coll., 2002; Bonin et coll., 2001; Garlock et coll., 2001; Largy et coll., 2004) et dans les

simulations connexionnistes (Monaghan et Ellis, 2010), ainsi qu'en identification de mots chez les enfants au préscolaire et au primaire (Garlock et coll., 2001).

Par ailleurs Zevin et Seidenberg (2002) ont avancé deux variables explicatives pour les effets d'AoA observés : la fréquence cumulative et la trajectoire fréquentielle. La fréquence cumulative est le fait que les mots appris tôt sont généralement rencontrés plus fréquemment que les mots appris tard, tandis que la trajectoire fréquentielle est le changement de fréquence des mots entre les écrits pour enfants et les écrits pour adultes. Une trajectoire croissante favoriserait la performance tardive (puisque le mot est inconnu ou rare en début de parcours) tandis qu'une trajectoire décroissante favoriserait la performance en début de parcours (le mot est fréquent en début de parcours). En français, Bonin et ses collaborateurs (2004) ont observé que la fréquence cumulative et l'AoA avaient des effets indépendants, mais que la trajectoire fréquentielle n'agissait pas sur la reconnaissance des mots. Dans la plus récente simulation connexionniste, Monaghan et Ellis (2010) ont simulé l'identification des mots à l'aide de trois modélisations connexionnistes différentes, incluant un modèle développemental dans le but d'analyser les trois variables liées aux stimuli : la fréquence d'occurrence, l'AoA et la trajectoire fréquentielle. L'introduction des mots dans le modèle simulait l'ordre dans lequel les mots sont enseignés aux enfants. Ils ont observé que l'ordre dans lequel les mots étaient introduits avait une « profonde influence » sur leur acquisition. Ils ont aussi conclu que chacune des variables explicatives avait un effet indépendant sur la reconnaissance des mots, mais que les effets d'AoA étaient la conséquence d'une exposition croissante aux mots, de concert avec les changements de plasticité au fur et à mesure que l'apprentissage progresse, et non un effet de fréquence cumulative comme le suggéraient Zevin et Seidenberg (2010, p. 506).

Comme nous venons de voir, les théories et les variables explicatives dans le traitement de l'information lexicale abondent. Les modèles développementaux expliquent l'acquisition et par le développement de processus phonologiques, orthographiques et morphologiques des enfants, tandis que les théories *usage-based* les expliquent par la fréquence d'exposition aux mots ou à leurs caractéristiques structurales, relationnelles et distributionnelles. D'autres l'expliquent par l'ordre dans lequel les enfants sont exposés aux mots et par l'évolution de leur fréquence dans les écrits.

Dans la prochaine section, l'acquisition des connaissances morphologiques sera abordée, en ciblant particulièrement les résultats d'études empiriques qui examinent l'orthographe de marqueurs morphologiques dérivationnels, c'est-à-dire les lettres muettes à la fin des radicaux morphologiques.

2.8 Le développement des connaissances morphologiques

Depuis la fin des années quatre-vingt, les chercheurs se sont penchés sur le développement morphologique en acquisition de l'orthographe (Bégin, Saint-Laurent et Giasson, 2010; Carlisle, 1988; Casalis, 2003; Defior et coll. 2008; Leybaert et Alegria, 1995; Pacton, 2001; Pacton et coll., 2001; Sénéchal, 2000; Sénéchal, Basque et Leclaire, 2006; Snowling, 1994; Totereau et coll., 1997; Treiman et Cassar, 1996) et plusieurs ont observé que les jeunes enfants utilisaient des processus de traitement morphologique, contrairement aux postulats des modèles développementaux.

Treiman et Cassar (1996) ont observé des effets morphologiques chez des élèves anglophones dès le début du 1er cycle en comparant l'orthographe des groupes consonantiques en position finale dans les mots mono- et bimorphémiques. Les mots bimorphémiques étaient composés du radical plus un suffixe, par exemple *bars* (radical *bar* + *-s* du pluriel) et *tuned* (radical *tune* + *-ed* du passé), tandis que les mots monomorphémiques regroupaient les mêmes consonnes à l'oral (*Mars* et *brand*). Les enfants de la 1^{re} année reproduisaient mieux la consonne initiale du groupe consonantique terminal dans des mots bimorphémiques que dans des mots monomorphémiques, suggérant qu'ils pouvaient avoir des représentations séparées des deux unités morphologiques. De plus, l'analyse des erreurs a démontré que certains enfants orthographiaient les mots à partir d'indices morphologiques sans avoir maîtrisé leurs correspondances lettres-sons ce qui va à l'encontre de l'hypothèse des modèles à stades voulant que les représentations phonologiques précèdent les représentations morphologiques (Ehri, 1995; Frith, 1985).

Une différence importante entre la morphologie flexionnelle du pluriel en anglais et celle du français est son rapport avec la phonologie. En anglais, les marqueurs morphologiques du pluriel et du passé sont aussi représentés phonologiquement. On les entend, ce qui n'est pas le

cas en français. Par exemple, le pluriel nominal de l'anglais est toujours sonorisé, même si les CGP ne sont pas toujours régulières (*bars/boxes*). En français, les marqueurs morphologiques équivalents sont muets (*les boites*) ou n'apparaissent que dans la liaison (*les enfants*). Totereau, Thevenin et Fayol (1997) ont analysé la compréhension et la production du nombre nominal et verbal chez les enfants francophones du primaire. Leurs résultats démontrent que dans des situations de production de mots, la représentation morphologique du nombre des noms (-s) et des verbes (-nt) serait maîtrisée à la fin du 1er cycle (2^e année). Les auteurs notent cependant que leur méthodologie expérimentale (associer ou écrire la bonne graphie à l'aide d'images symbolisant le singulier ou le pluriel) ne reflète pas les situations d'écriture en classe, mais les résultats appuient tout de même le postulat que les jeunes enfants ont des représentations morphologiques flexionnelles dès le 1er cycle du primaire.

Dans leur étude sur les régularités graphotactiques, Pacton et ses collaborateurs (2005) ont observé les effets de régularité morphologique dérivationnelle dans l'orthographe des terminaisons en /o/ et en /et/ chez les enfants de deuxième, troisième et 5^e année. Dès le 1er cycle, les enfants utilisaient davantage les graphies *-eau* et *-ette* lorsqu'ils étaient associés à la notion de diminutif que lorsqu'ils étaient orthographiés sans contexte. Par exemple, le mot /vitaro/ était orthographié plus souvent *vitareau* lorsque le non-mot était mis en contexte par la phrase *Un petit vitar est un vitareau* que s'il apparaissait sans contexte dans une dictée de mots.

Leybaert et Alegria (1995) ont été les premiers à examiner la production des lettres muettes qui peuvent être déduites par association à un mot de la même famille morphologique.

2.9 La production des lettres muettes dans les radicaux morphologiques

Contrairement aux études présentées ci-dessus, Leybaert et Alegria (1995) n'ont pas observé d'effet morphologique avant la 4^e année (fin du 2^e cycle). Ils ont examiné l'orthographe de mots réguliers (structure orthographique régulière et consistance phonographémique), morphologiques (avec une lettre muette qui pouvait être déduite à l'aide d'un féminin) et opaques (orthographe irrégulière qui ne peut être déduite de l'information morphophonologique) d'enfants franco-belges de la deuxième et la 4^e année. À la fin du 1er

cycle (2^e année), les enfants orthographiaient généralement mieux les mots réguliers que les mots morphologiques et opaques. Cependant, l'information représentée par les lettres muettes des mots morphologiques n'a pas eu d'effet sur leur réussite orthographique, ils orthographiaient aussi bien les mots opaques que les mots morphologiques. Ce n'est qu'en 4^e année qu'une différence significative est apparue, suggérant que les représentations morphologiques dérivationnelles n'apparaissent que tardivement chez les enfants francophones. Par ailleurs, dans une étude sur l'effet de différentes méthodes d'enseignement auprès d'enfants de deuxième, quatrième et 6^e année, Leybaert et Content (1995) ont aussi conclu que l'information morphologique n'était pas exploitée par les enfants de 2^e année.

Plus récemment, des chercheurs ont documenté un effet contraire. L'orthographe des enfants du 1^{er} cycle du primaire représentait la capacité à traiter l'information morphologique contenue dans les lettres muettes de mots dérivables (Pacton, 2001; Sénéchal, 2000; Sénéchal, Basque et Leclaire, 2006). Sénéchal (2000) a repris le modèle expérimental de Leybaert et Alegria (1995), arguant que leurs mots opaques ne reflétaient pas une différence morphologique avec leurs mots morphologiques, mais bien une différence orthographique. Elle étudie donc les effets morphologiques en redéfinissant la catégorie des mots opaques, des mots dont la lettre muette terminale ne peut être déduite par analogie morphologique. Par ailleurs, deux types de mots morphologiques sont distingués : avec ou sans flexion féminisante. Une dictée de 40 mots a ainsi été administrée à 102 enfants franco-ontariens de la deuxième (n = 57) et de la quatrième (n = 55) année du primaire. Les mots sont réguliers (n = 10 sans lettre muette), morphologiques (n = 20 avec un dérivé qui permet de déduire la lettre muette, dont 10 avec un féminin et 10 sans féminin) et opaques (n = 10 dont la lettre muette ne peut être déduite). L'objectif est de mesurer si les enfants étaient sensibles à l'information morphologique qui se trouve dans les lettres muettes terminales des mots morphologiques.

Comme prévu, les résultats de Sénéchal (2000) démontrent que les enfants de la 4^e année réussissent mieux l'orthographe en général et représentent plus souvent les lettres muettes terminales que les enfants de 2^e année. Cependant, le comportement orthographique des deux groupes est similaire, c'est-à-dire qu'aux deux niveaux scolaires, les mots réguliers sont mieux orthographiés que les mots morphologiques qui, eux, sont mieux orthographiés que les

mots opaques. Ces résultats ont été reproduits dans des recherches subséquentes (Pacton et Casalis, 2006; Pacton, 2001; Sénéchal, Basque et Leclaire, 2006), rejetant l'hypothèse que les enfants francophones du 1er cycle n'exploitent pas l'information morphologique contenue dans les mots morphologiques et soutenant celle que les jeunes enfants utilisent des stratégies autres que phonologiques en début de parcours (Pacton et coll., 2005, 2001; Snowling, 1994; Totereau et coll., 1997; Treiman et Cassar, 1996).

Par ailleurs, Sénéchal (2000) a observé, lors d'une tâche d'orthographe de mots sous dictée, que les enfants en fin du 2^e cycle ont réussi à orthographier correctement 68,2 % des mots réguliers, mais seulement 52,1 % des mots morphologiques et 33,7 % des mots opaques. Dans une étude subséquente, Sénéchal, Basque et Leclaire (2006), qui examinaient le lien entre la réussite orthographique et l'utilisation explicite de stratégies morphologiques, ont aussi observé une sous-utilisation de l'information morphologique. Dans deux expérimentations séparées, les sujets de 4^e année ont relaté n'avoir utilisé une stratégie morphologique que dans 34 % et 18 % des cas, respectivement. Même si les consonnes muettes étaient représentées, l'information morphologique qu'elles symbolisent n'était toujours pas pleinement exploitée à la fin du 2^e cycle (Sénéchal, Basque et Leclaire, 2006).

Bégin, Saint-Laurent et Giasson (2010) ont étudié l'orthographe des mots morphologiques chez les élèves québécois de la 6^e année ($n = 240$). Dans des épreuves de dictées trouées, la moyenne de réussite pour la production de lettres muettes terminales dérivables était de 86,98 % (écart-type de 14,72) pour les mots ayant une flexion féminisante et 72,55 % (18,86) pour ceux qui n'en ont pas. Ce taux de réussite ne reflète pas la progression des apprentissages suggérée par le MELS quant à la production des lettres dans les radicaux morphologiques. On rappelle que c'est en 6^e année que cette habileté devrait être maîtrisée. Il est cependant intéressant de noter que les auteurs ont observé une différence significative dans le type d'erreur observé (omission et substitution) en fonction de la fréquence des mots. Il y avait plus d'erreurs de substitution dans les mots fréquents, tandis que les erreurs d'omission dominaient dans les mots rares. Ceci pourrait suggérer que les enfants sont sensibles à la présence de lettres muettes, mais n'ont pas encore maîtrisé la règle d'analogie à la famille morphologique. Par ailleurs, cela pourrait aussi refléter la connaissance des familles morphologiques des stimuli utilisés.

À la lumière des observations dans les études de Sénéchal (2000) et Bégin, Saint-Laurent et Giasson (2010), un des objectifs de cette recherche sera d'examiner la progression de la réussite orthographique et la production de la lettre muette dans les radicaux morphologiques, et ce, du début à la fin des études de niveau primaire.

Dans la prochaine section, on examine les variables explicatives dans la production des marqueurs morphologiques muets.

2.10 Les variables explicatives dans la production des consonnes muettes

Nous nous intéressons aussi aux variables explicatives dans le développement des connaissances morphologiques des enfants du primaire. Comme nous l'avons vu plus haut, des effets de régularité et de consistance orthographique ont été observés dès le 1er cycle du primaire. Sénéchal (2000) a, à posteriori, contrôlé pour des effets de consistance dans la prononciation des consonnes dans les dérivés. Par exemple, le *s* peut être sonorisé par les sons /s/ ou /z/ (*épais/épaisse, repos/reposer*). Elle n'a observé aucun lien entre la production orthographique des lettres muettes et leur prononciation dans les dérivés.

Des variables propres aux sujets ont aussi été étudiées, dont l'utilisation explicite de stratégies morphologiques (Sénéchal, Basque et Leclaire, 2006) et le niveau de lecture et d'orthographe des enfants. Pacton (2001) a reproduit les résultats de Sénéchal (2000) auprès de 42 élèves de la 3^e année (CE2) pour voir si le niveau de lecture et d'orthographe des enfants expliquerait les effets morphologiques. Les sujets ont été répartis en trois groupes : moyens, forts (+ 1 écart-type) et faibles (- 1 écart-type). Ils ont observé que même si les résultats globaux de la dictée de mots variaient en fonction du niveau de lecture et d'orthographe, l'effet morphologique (régulier > morphologique avec féminin > morphologique sans féminin > opaque) était présent, tant chez les élèves forts que chez les moyens et les plus faibles. Bégin, Saint-Laurent et Giasson (2010) ont aussi observé un effet du niveau d'habileté orthographique des scripteurs. Les élèves ayant un niveau d'habileté fort (un écart-type ou plus au-dessus de la moyenne du groupe, n=36) ont mieux réussi que les élèves faibles (un écart-type ou plus sous la moyenne, n=36) et les élèves moyens (ni forts, ni faibles, n = 170). Cependant, contrairement aux résultats de Pacton, la différence entre la

production des lettres muettes dans les mots morphologiques sans féminin et les mots lexicaux (opaques) n'était pas significative « lorsqu'on tient compte des résultats obtenus par l'ensemble des élèves » (p.287).

Par ailleurs, Sénéchal (2000) et Pacton et Casalis (2006) ont suggéré que leurs résultats pourraient s'expliquer par la fréquence d'exposition aux mots eux-mêmes, mais aussi par la fréquence de la famille morphologique. Les mots courts et leur consonne muette sont rencontrés plus souvent lorsqu'ils apparaissent aussi dans les mots longs que lorsqu'ils sont solitaires. Sénéchal suggère que cet effet de redondance pourrait avoir un effet de fréquence sur la production de la lettre muette à la fin des radicaux. Elle n'a toutefois pas contrôlé l'effet de fréquence des familles morphologiques et aucune explication du contrôle de fréquence n'est donnée. Une communication personnelle (2011) a permis d'établir que les indices de fréquence de Brulex (Content et coll., 1990) avaient été utilisés. Les indices de fréquence sont tirés de corpus destinés aux adultes.

Par ailleurs, Bégin, Saint-Laurent et Giasson (2010) ont repris la méthodologie de Sénéchal pour évaluer la production des consonnes muettes chez les enfants de la 6^e année en contrôlant la fréquence et ont effectivement observé un effet de fréquence. Les mots fréquents étaient significativement mieux réussis que les mots rares à la fin des études primaires. La fréquence de la famille morphologique des mots n'a cependant pas été contrôlée.

Un deuxième objectif sera donc d'évaluer l'effet de fréquence d'occurrence sur la réussite des marqueurs morphologiques à tous les niveaux du primaire, et ce, en contrôlant la fréquence du radical morphologique et la fréquence de la famille morphologique.

Finalement, des effets de féminisation ont été observés lors de la production de consonnes muettes à la fin des mots morphologiques (Bégin, Saint-Laurent et Giasson, 2010; Pacton et coll., 2001; Sénéchal, 2000; Sénéchal, Basque et Leclaire, 2006). Dans toutes ces études, les consonnes muettes terminales étaient mieux représentées à l'écrit lorsque le radical morphologique avait une flexion féminisante (*froid/froide*) que s'il n'y avait que des dérivés (*début/débutant*), suggérant que la flexion féminisante est plus accessible que la dérivation nominale, verbale ou adjectivale et faciliterait l'orthographe des mots morphologiques.

Cependant, des failles dans le contrôle de la fréquence lexicale ont été détectées. Le contrôle de la fréquence des mots dans Sénéchal (2000) a été fait à partir de Brulex (Content et coll., 1990) dont les indices de fréquences sont tirés du Trésor de la grammaire d'Imbs (1971), destiné aux adultes et qui date de plus de 40 ans. Par ailleurs, Bégin, Saint-Laurent et Giasson (2010) ont contrôlé la fréquence entre les mots fréquents et les mots rares, mais la fréquence entre les deux types de mots n'est pas analogue. L'ensemble des mots morphologiques qui peuvent être féminisés est 4,65 fois plus fréquent que l'ensemble des mots morphologiques sans flexion féminisante et 1,56 fois plus fréquent que les mots lexicaux (2010, p. 284). Par ailleurs, l'ensemble des mots lexicaux est presque trois fois plus fréquent que les mots morphologiques sans féminin, ce qui pourrait aussi expliquer pourquoi les scores de réussite entre les mots morphologiques sans féminin n'étaient pas statistiquement différents des mots lexicaux.

À la lumière des failles méthodologiques dans le contrôle de la fréquence, un troisième objectif sera d'évaluer l'effet de féminisation sur la production de la consonne muette en contrôlant la fréquence entre les types de mots.

Avant de présenter les questions spécifiques de recherche, une variable explicative propre aux sujets se doit d'être considérée, celle de la langue maternelle de l'apprenant. Dans la prochaine section, des études empiriques en acquisition du français langue seconde seront présentées, ainsi que l'implication de la langue maternelle dans la production de mots qui requièrent des connaissances morphologiques.

2.11 La langue maternelle et l'acquisition du français écrit

Le profil linguistique des élèves québécois est de plus en plus hétérogène, particulièrement sur l'île de Montréal (1,8 million d'habitants), où 50,2% de la population ont une langue maternelle autre que le français (17,6% anglais, 32,6% autres) et où 25% de la population parlent généralement l'anglais à la maison. Dans la région métropolitaine de Montréal, où habite près de la moitié de la population du Québec (3,6 millions sur 7,4 millions), 34,3% ont une langue maternelle autre que le français (12,5% anglais et 21,8% autres) et 17,4% parlent

généralement l'anglais à la maison¹⁶. Bien qu'il y ait une population allophone importante dans la grande région montréalaise (13,4% dans la Région et 20,6% sur l'Île), la population étudiée dans le cadre de cette recherche se limite aux enfants dont la langue maternelle est le français ou l'anglais. C'est dans cette perspective que l'acquisition du FL2 sera abordée, particulièrement les variables explicatives des résultats d'enfants qui sont éduqués dans des programmes d'immersion française au Canada.

La divergence linguistique et typographique de la langue maternelle d'un apprenant est un facteur dans l'acquisition d'une langue seconde (Genesee, 2004). Plus les langues se rapprochent du point de vue de la typographie, plus il est probable qu'un transfert se fasse entre les deux. On note ici que le type d'instruction est aussi une variable explicative importante en acquisition des langues secondes, mais que cet aspect est au-delà de la portée de cette recherche donc ne sera pas abordé.

Le français et l'anglais sont similaires sur le plan typologique (Genesee, 2004), mais partagent aussi d'autres propriétés linguistiques, dont la phonologie et la morphologie, qui seront explorées ici. Bruck et Genesee (1995) ont observé un lien entre la conscience phonologique et le niveau de lecture des enfants bilingues. Haigh et ses collaborateurs (2011) ont observé que la conscience phonémique de l'anglais jouait un rôle prédictif dans l'acquisition de la lecture en français. Effectivement, des enfants qui étaient habiles à manipuler les phonèmes de l'anglais en maternelle devenaient de meilleurs lecteurs du français en 2^e année que les enfants qui étaient moins habiles.

On peut se demander si les similitudes phonologiques et orthographiques entre certains mots de l'anglais et du français sont une source de facilitation ou de confusion en production de mots chez des enfants en immersion. Au niveau lexical, certains mots partagent la même phonologie, mais pas la même orthographe, comme le mot /bag/, *bague* en français et *bag* (*sac*) en anglais. Cette différence orthographique pourrait être une source d'erreur. Inversement, plusieurs mots partagent l'orthographe, mais pas la phonologie, comme le mot *camp*, prononcé /kamp/ en anglais. La connaissance de l'anglais faciliterait la représentation

¹⁶ Source: Secrétariat de la politique linguistique du Québec. Données 2006
<http://www.spl.gouv.qc.ca/documentation/rapportsondagesstatistiques/dynamiquedeslangues/tableaux/>

de la rime *-amp*. On peut imaginer que la connaissance orthographique du mot équivalent en anglais pourrait aider à orthographier les deux lettres qui ne sont pas prononcées en français. Il est aussi probable que l'orthographe de mots dans une langue puisse interférer avec leur production dans la seconde. Si on prend l'exemple du mot *nord*, un anglophone est-il enclin à écrire *nort** par analogie avec sa forme anglaise (*north*)?

2.11.1 La langue maternelle et le développement morphologique

Outre les liens orthographiques et phonologiques, le français et l'anglais ont des similarités morphologiques. Elles sont toutes deux des langues phonomorphémiques (ou morphophonémiques) (Deacon, Wade-Woolley et Kirby, 2007; Pacton et Deacon, 2008) dont les systèmes orthographiques représentent à la fois des unités phonémiques et morphologiques. Elles partagent ainsi certaines propriétés morphosyntaxiques, dont la suffixation. Cette similarité pourrait faciliter la production orthographique des enfants en immersion. « Readers of these two orthographies might gain an advantage if they appreciate the ways that words are made up of morphemes because these units of meaning play a role in determining the pronunciation of words in both English and French. » (Deacon et coll., 2007, p. 732).

Deacon et ses collaborateurs (2007) ont effectivement observé qu'il y avait transfert de la conscience morphologique entre le français et l'anglais. On rappelle que la conscience morphologique se définit ici comme « la prise de conscience de la structure morphémique des mots et l'habileté d'en tenir compte et de la manipuler » (Carlisle, 1995, p. 194). Ils ont observé, chez des anglophones du 1er cycle en immersion, que leur niveau de conscience des flexions verbales du passé en anglais était lié à leur niveau de lecture en français. Cependant, le lien inverse, c'est-à-dire celui entre la conscience morphologique flexionnelle en français et la lecture en anglais, ne se manifestait qu'en 3^e année. Ils en conclurent que la conscience morphologique est une connaissance linguistique qui est transférable d'un système orthographique à l'autre, mais que cette relation « interlangue » change en fonction du développement de la littéracie dans la langue seconde.

Quoique le français et l'anglais partagent la propriété morphosyntaxique de la suffixation verbale, d'autres conventions morphologiques sont fort divergentes, ce qui pourrait

désavantager les apprenants anglophones en immersion vis-à-vis de leurs compagnons francophones. La présence du genre nominal est une source de difficulté bien documentée (Genesee, 2007; Morris et Labelle, 2005), mais, à notre connaissance, le phénomène des marqueurs morphologiques muets à la fin des radicaux de familles morphologiques n'a jamais été examiné dans les études sur les enfants anglophones en immersion française.

Morris et Labelle (2005) ont observé que les enfants en immersion française n'atteignaient pas le rendement des enfants francophones dans la production de mots qui requièrent des connaissances morphologiques (mots à morphologie). Elles ont comparé les performances d'enfants québécois francophones à celle d'enfants québécois allophones (anglophones et autres langues maternelles) dans différents contextes d'immersion française à l'aide d'un C-test, c'est-à-dire un texte continu où la fin d'un certain nombre de mots est coupée. Les trois contextes d'immersion sont l'immersion partielle dans le système scolaire anglophone, soit 30% de la maternelle à la 6^e année, l'immersion totale dans le système d'éducation anglophone, soit 100% de la maternelle à la fin de la 2^e année, suivi de 50% de la 3^e à la 6^e année, et la submersion dans le système francophone. Au 1^{er} cycle, aucune différence n'est distinguée entre francophones (52%) et allophones (53%) dans la production de mots à morphologie, mais les allophones réussissent moins bien le genre des noms. C'est en 4^e, 5^e et 6^e année que l'écart se creuse. Les élèves en immersion partielle ont un taux de réussite de 26%, moins que la moitié du taux de réussite des élèves en immersion totale (57%) qui, eux, ont des résultats plus faibles que les élèves en submersion (63%) et les francophones (68% et 74%). Les auteurs notent cependant que les allophones utilisent plus leur mémoire visuelle que les francophones, qui eux se fient plus aux CGP. Ceux en immersion tendent à placer des mots fréquents « qui leur viennent à l'esprit » plutôt que des mots plus rares qui correspondraient au contexte morphosyntaxique.

« [...] cette stratégie de complétion purement lexicale, nous la retrouvons chez les élèves en immersion de 4^e à 6^e année, mais pas chez les élèves en submersion du même niveau [...]. Cela nous indique que les élèves en immersion française, contrairement à ceux en submersion, ne développent pas de stratégies de lecture efficaces en français. Le fait qu'ils tendent à copier des mots trouvés ailleurs dans le texte appuie cette observation. Ces élèves semblent privilégier des stratégies épilinguistiques de bas niveau. » (2005, p. 5)

Les observations de Morris et Labelle démontrent aussi que l'intensité d'exposition à la langue seconde aurait des effets sur sa production écrite. Cela reflète aussi les conclusions de Genesee (2007) dans son analyse sommaire de la recherche en immersion française. Il conclut que les enfants anglophones en immersion ont systématiquement de meilleurs résultats en lecture et en écriture du français que les enfants qui suivent un programme de français langue seconde (FLS), mais que leur niveau de compétence n'atteint pas toujours celui des enfants francophones. Par exemple, Genesee (2004) a observé que les anglophones en immersion réussissaient aussi bien que les francophones sur des examens de lecture standardisés. Jared (2006) a cependant observé que leurs résultats étaient significativement plus bas. Nonobstant ces observations de Jared, il semblerait que les anglophones en général, malgré un cheminement en immersion française, n'atteignent jamais le niveau d'orthographe grammaticale ou de vocabulaire des enfants francophones éduqués en L1 (Genesee, 2007).

Une explication probable pour les scores inférieurs des anglophones serait la fréquence d'exposition au français écrit. Romney et ses collaborateurs (1995) ont observé que les anglophones en immersion allouaient peu de temps à la lecture du français à l'extérieur des heures d'école. De plus, dans leurs entretiens avec des élèves de 5^e année en immersion française, ces derniers ont reconnu que leur niveau de vocabulaire insuffisant était la raison première de leur difficulté en lecture. En plus d'être issus d'un environnement familial où le français n'est pas parlé, les anglophones ont moins d'expériences de lecture du français que leurs compagnons francophones et ce, malgré un nombre équivalent d'heures d'enseignement du français. D'une perspective « usage-based », on peut déduire qu'un enfant en immersion, qui ne lit pas le français ailleurs qu'à l'école aurait un vocabulaire limité par rapport à ses compagnons francophones. Si les dérivés ne font pas partie de son vocabulaire, il ne peut pas y avoir de représentation de la lettre muette par analogie morphologique.

Pour résumer, la langue maternelle des apprenants est une variable explicative de l'acquisition du langage écrit. Il est démontré que la conscience phonologique et la conscience morphologique de l'AL1, ainsi que l'intensité de l'exposition au FL2, sont des variables prédictives en acquisition du français écrit. Les observations empiriques rapportent qu'en général, même en situation d'immersion totale, les anglophones n'atteignent pas le

niveau de succès orthographique des enfants francophones éduqués en FL1, et cela pourrait s'expliquer par un niveau inférieur d'exposition au français écrit à l'extérieur de l'école.

2.12 Les questions spécifiques de recherche

Nous rappelons que l'intérêt de cette recherche est centré sur le développement des connaissances morphologiques chez les enfants du primaire, plus spécifiquement sur la production des consonnes muettes à la fin des radicaux morphologiques. Les objectifs qui sous-tendent notre recherche sont les suivants :

- Observer la progression, du début à la fin du primaire, de la production écrite de radicaux morphologiques et la comparer à la progression des apprentissages prescrite par le MELS, ainsi qu'aux postulats des modèles développementaux;
- Observer si la fréquence d'occurrence des mots morphologiques et de leurs familles morphologiques agit sur la production de la consonne muette à la fin des radicaux morphologiques;
- Observer si la présence d'une flexion de féminisation agit sur la production de la consonne muette terminale des radicaux morphologiques lorsque la fréquence est contrôlée;
- Observer si la langue maternelle agit sur la production des consonnes muettes à la fin des radicaux morphologiques

Deux questions sont posées pour évaluer la production des consonnes muettes à la fin des radicaux mots morphologiques et la production orthographique des mots morphologiques.

1. Est-ce que le niveau scolaire agit sur la réussite orthographique et morphologique des radicaux morphologiques? Si oui, observe-t-on une progression?
 - a. À quel niveau observe-t-on que les enfants ont généralement des connaissances spécifiques aux marqueurs morphologiques muets?

- b. À quel niveau observe-t-on la maîtrise des marqueurs morphologiques muets?
 - c. Est-ce que la réussite morphologique et la réussite orthographique des radicaux morphologiques progressent au même rythme?
2. Si les enfants produisent une lettre muette, reflète-t-elle celle représentée dans l'orthographe du mot dérivé?

Dans ce contexte, deux questions sont posées pour évaluer certaines variables explicatives propres aux mots.

3. Est-ce que la réussite morphologique est liée à la fréquence des mots?
- a. La fréquence d'un mot agit-elle sur la production de sa consonne muette? Si oui, dans quelle mesure?
 - b. La fréquence de la famille morphologique agit-elle sur la production de la consonne muette? Si oui, dans quelle mesure?

Pour répondre à la prochaine question, les mots sont regroupés sous trois conditions morphologiques : le radical a un féminin et il est plus fréquent que ses dérivés (R+F), le radical n'a pas de féminin et est plus fréquent que ses dérivés (R+) et le radical n'a pas de féminin et est moins fréquent que ses dérivés (D+). La question suivante est posée:

4. Est-ce que la réussite morphologique diffère selon la condition morphologique du radical? Si oui, dans quelle mesure?
- a. Est-ce que la fréquence relative entre le radical et le dérivé morphologique agit sur la production de la consonne muette? Si oui, dans quelle mesure?
 - b. L'existence d'une flexion féminisante agit-elle sur la production de la consonne muette? Si oui, dans quelle mesure?

Une question est posée pour évaluer les variables explicatives liées aux sujets.

5. Est-ce que la réussite morphologique varie en fonction de la langue parlée à la maison? Si oui, dans quelle mesure?

Dans le prochain chapitre, nous présenterons en détail la méthodologie utilisée pour répondre à ces questions.

CHAPITRE 3

MÉTHODOLOGIE

Afin d'atteindre nos objectifs de recherche, une dictée de mots a été administrée à des enfants de la première à la sixième année du primaire. Cette démarche quantitative a pour but de permettre de tracer un portrait du développement des connaissances morphologiques liées à la production de marqueurs morphologiques muets à la fin de mots lexicaux. Notre expérience reproduit la démarche scientifique des études qui ont motivé cette recherche, c'est-à-dire Sénéchal (2000) et Bégin, Saint-Laurent et Giasson (2010). Les résultats nous permettront de mesurer les acquis orthographiques des élèves, un des objectifs de notre démarche. Les stimuli ont été choisis pour observer s'il y a un lien entre la production d'une lettre muette dans un radical et la présence de cette lettre dans un dérivé morphologique. Cet outil quantitatif permet également d'évaluer l'effet de variables explicatives en production de consonnes muettes dérivables, soit la fréquence d'occurrence, la féminisation et la langue maternelle.

Les aspects méthodologiques qui seront présentés sont la population, l'instrument de mesure, le choix de stimuli et leur mesure de fréquence, la collecte de données et la grille d'analyse des données.

La population

Un échantillon de 431 élèves de trois écoles primaires de la région de Montréal a participé à l'expérimentation. Deux des écoles sont privées, une de garçons et une de filles, et situées dans un milieu socioéconomique aisé. Elles sont des écoles anglophones qui offrent des programmes bilingues, c'est-à-dire que le niveau et le nombre d'heures d'enseignement du français sont les mêmes que dans les écoles francophones. La troisième école est une école

publique francophone située dans la banlieue sud de Montréal. Selon les indices du seuil de revenu faible (SRF) et du milieu socioéconomique (IMSE) du ministère de l'Éducation québécoise (MELS), cette école serait dans un milieu favorisé : $SRF = 3$ et $IMSE = 2$ ¹⁷.

Il n'a pas été possible, dans le cadre de cette recherche, de trouver un échantillon qui représentait la moyenne socioéconomique des enfants québécois. On considère donc que les données récoltées auprès d'enfants favorisés devraient représenter le niveau le plus élevé auquel on peut s'attendre d'enfants aux âges étudiés. Autrement dit, si des élèves issus de milieux favorisés ne réussissent pas à produire les lettres muettes ou à orthographier les mots, on s'attend à ce qu'à fortiori les élèves moins favorisés n'y arrivent pas non plus. Inversement, si des élèves de milieu favorisé peuvent produire ce genre de lettres muettes, on peut penser que cette compétence est à la portée d'élèves de milieux moins favorisés qui ont reçu le même type d'enseignement. Le type d'enseignement reçu par les sujets et le nombre d'années où les non-francophones ont été scolarisés en français n'ont pas pu être recensés dans le cadre de cette recherche. Les limites liées à ces facteurs et à la composition de l'échantillon seront discutées dans la conclusion.

L'échantillon compte 203 garçons et 228 filles, répartis en six groupes, de la première à la 6^e année. L'âge moyen est de 9 ans 9 mois, variant entre 6 ans 7 mois et 13 ans 5 mois. La collecte de données s'est déroulée à la fin de l'année scolaire. L'âge des enfants a été déterminé en calculant la différence entre la date de naissance et la journée de passation de la première des deux dictées. Nous avons regroupé les sujets en deux catégories, selon la langue généralement parlée à la maison¹⁸ : 278 sujets parlent seulement le français à la maison, 153 uniquement l'anglais. L'échantillon original avait aussi trois autres catégories selon la langue maternelle : le français et l'anglais, le français et une autre langue, une langue autre que le français ou l'anglais. Ces sujets n'ont pas été retenus puisqu'ils n'étaient pas assez nombreux pour représenter un échantillonnage significatif. Ainsi, les sujets dont la langue maternelle est l'anglais sont dans une situation d'immersion en français L2. Il est aussi important de noter

¹⁷ Au Québec, les données relatives à la défavorisation en milieu scolaire sont calculées à l'aide de deux variables socioéconomiques, le SRF et l'IMSE. Chaque école publique se voit assigner un rang de 1 à 10 pour chacune des variables, où 1 = l'indice le plus élevé ou favorisé.

¹⁸ La langue maternelle a été établie par les parents des sujets lorsqu'ils ont rempli le formulaire de consentement.

que les anglophones fréquentent tous l'une des deux écoles avec le programme bilingue, tandis que 23 % des élèves qui fréquentent les écoles bilingues ont le français comme langue maternelle. La répartition de la population par année scolaire selon le sexe, l'âge et la langue maternelle est présentée au Tableau 3.1.

Tableau 3.1
Population par niveau scolaire, selon le sexe et l'âge

NIVEAU	N	GARÇONS	FILLES	FRANÇAIS	ANGLAIS	ÂGE MOY. (<i>É-T</i>)
1 ^{re}	54	30	24	37	17	7,2 (0,28)
2 ^e	65	36	29	35	30	8,1 (0,37)
1 ^{er} CYCLE	119	66	53	72	47	7,7 (0,56)
3 ^e	78	35	43	59	19	9,1 (0,27)
4 ^e	69	33	36	43	26	10,1 (0,34)
2 ^e CYCLE	147	68	79	102	45	9,6 (0,59)
5 ^e	78	32	46	43	35	11,1 (0,29)
6 ^e	87	37	50	61	26	12,2 (0,34)
3 ^e CYCLE	165	69	96	104	61	11,7 (0,63)
TOTAL	431	203	228	278	153	9,9 (1,73)

Cette recherche s'inscrit dans le cadre du Groupe de recherche sur l'apprentissage du français du Département de linguistique et didactique des langues de l'Université du Québec à Montréal. La recherche a reçu l'approbation du Conseil de recherche en sciences humaines du Canada, qui la subventionne. Elle est approuvée par le comité d'éthique de l'université et par la direction des écoles (Appendices G et H) et seuls les résultats d'enfants qui avaient un consentement écrit de la part des parents (Appendice E et F) ont été comptabilisés.

3.1 L'instrument de mesure

La méthodologie expérimentale de Sénéchal (2000) a été adaptée pour cette recherche. L'instrument de mesure, une dictée de mots, permet de mesurer 1) si les enfants du niveau primaire représentent ou non les lettres muettes à la fin des mots courts, 2) si les lettres muettes à la fin des mots courts correspondent à la lettre sonorisée dans les dérivés, 3) si l'existence d'un féminin influence la réussite orthographique lorsque la fréquence

d'occurrence est contrôlée, 4) si la fréquence d'occurrence de la famille dérivationnelle entière influence la réussite orthographique.

Nous avons développé une dictée de 24 paires de mots, chaque paire regroupant un mot court (un radical morphologique) et un mot long (un dérivé morphologique). Les mots courts sont regroupés selon trois conditions. Dans la condition R+F, le mot court a une forme féminine et est plus fréquent que son dérivé (n = 8), dans la condition R+, le mot court est plus fréquent que son dérivé, mais qu'il n'a pas de forme féminine (n = 8) et dans la condition D+, le mot dérivé est plus fréquent que le mot court qui n'a pas de féminin (n = 8). Une quatrième catégorie, les mots dont la forme féminine est plus fréquente que la forme masculine, a été envisagée, mais éliminée par manque de mots qui répondaient à cette condition. La liste des mots par condition est à l'Appendice B.

Le choix des stimuli

Pour mesurer les effets de fréquence, l'écart entre la fréquence du mot court et celle de ses dérivés doit être aussi grand que possible. Cependant, les mots ne doivent pas être si fréquents pour qu'un niveau de plafonnement soit atteint dès le 1^{er} cycle. Par exemple, on a éliminé des mots comme *petit*, *grand* et *gros*.

3.1.1 Les mesures de fréquence

Pour mesurer la fréquence d'occurrence des mots d'épreuves administrées à des enfants québécois, il serait idéal d'utiliser des données tirées de corpus visant spécifiquement cette population et ses réalités géoculturelles uniques. Cependant, une telle base de données lexicale n'existe pas. Il existe des bases de données québécoises, mais les textes visent les adultes. On retrouve des bases de données élaborées à partir de corpus pour enfants, mais elles sont normalisées en Europe. La seule banque de données représentative de la langue française telle que parlée et écrite au Québec est la Banque de données des corpus lexicaux québécois du Secrétariat à la politique linguistique québécoise. Cependant, les corpus sur lesquels elle se fonde ne sont pas représentatifs de ceux auxquels seraient exposés des enfants de l'école primaire. Les textes sont majoritairement des écrits techniques, scientifiques, journalistiques, politiques et sociopolitiques. Il comprend aussi un recueil de textes rédigés

entre le 17^e siècle et le début du 20^e siècle. Par ailleurs, Omnilex (Desrochers, 2006) a aussi été développée au Canada (à l'Université d'Ottawa), mais les seules données normalisées au Canada sont des mesures de fréquences subjectives et de valeur d'imagerie, recueillies auprès d'adultes franco-canadiens. Selon son auteur (Desrochers, communication électronique personnelle, 2010), la base de données la plus utile pour les études visant les enfants est Manulex (Lété et coll., 2004). Il s'agit d'une base de données normalisée à partir de manuels scolaires destinés à des enfants français. Ainsi, de toutes les bases de données existantes, il semble que Manulex reflète le plus adéquatement le lexique orthographique des enfants francophones, donc c'est cette base de données qui a été retenue pour contrôler les fréquences d'occurrence des stimuli expérimentaux de cette recherche, bien qu'elle ait été normalisée en Europe.

3.1.2 L'indice de fréquence des stimuli

La mesure de fréquence retenue est l'indice SFI de Manulex. Le SFI (Standard Frequency Index)¹⁹ est dérivé directement de l'indice U, une estimation de la fréquence d'occurrence par million. Le SFI est, à notre avis, plus pratique que le U pour apprécier la fréquence d'occurrence des mots dans les textes. Un mot avec un SFI = 90 apparaît une fois à tous les 10 mots, avec un SFI = 80, tous les 100 mots, SFI = 70, tous les mille mots, ainsi de suite. Le point de référence d'une fréquence par million de mots est SFI = 40. La liste des stimuli et des SFI peut être trouvée à l'Appendice 2.2.

Les indices de Manulex pour la 3^e à la 5^e année sont les mêmes donc il y a 3 indices : G1, G2 et G3-5. Pour choisir les stimuli, le SFI_1 (1^{re} année) a été utilisé. Pour mesurer si la fréquence d'occurrence agissait sur les résultats, il nous semblait pertinent d'utiliser des indices cumulés, ce qui reflète plus justement l'exposition des enfants depuis le début de la scolarisation. Par ailleurs, Manulex ne comptabilise pas les fréquences des familles morphologiques. Celles-ci ont été calculées manuellement²⁰. Puisque les fréquences sont constantes à partir de la troisième année, il n'est donc pas nécessaire de cumuler les

¹⁹ Les formules pour calculer cet indice sont explicitées dans Lété et coll., 2004.

²⁰ Les indices U ont été additionnés et convertis en SFI selon la formule $SFI = 10 * [\log_{10}(U) + 4]$, tel que suggéré par Lété (communication électronique personnelle, 2010)

fréquences au-delà de la 3^e année. Autrement dit, les indices de fréquences cumulatifs sont les mêmes pour les niveaux 3, 4, 5 et 6.

3.2 La procédure

Les mots retenus sont divisés en deux dictées de 24 mots chacune. Chaque dictée compte 12 mots courts et 12 mots longs. Le mot court apparaît dans une dictée et son dérivé dans l'autre; dans chaque dictée, les mots sont répartis de façon aléatoire. Durant l'épreuve, le mot cible est lu et mis en contexte dans une phrase. Par exemple, le mot permis était accompagné de la phrase *Ma sœur a son permis de conduire* (la liste des mots et celle des phrases-contexte sont données à l'Appendice C). Les élèves doivent écrire seulement le mot cible sur une feuille numérotée de 1 à 24. Les épreuves durent entre 15 et 30 minutes chacune, selon le niveau scolaire.

Les deux dictées ont été administrées à une semaine d'intervalle en mai et juin. Pour écarter un effet possible de l'ordre de passation des dictées, la moitié des classes de chaque niveau scolaire a fait la dictée A en premier, l'autre moitié, la dictée B. La moitié des épreuves a été administrée par deux examinateurs, l'autre par les enseignants des classes respectives. Des consignes écrites (Appendice D) et les 2 listes de mots ont été données aux examinateurs à la dernière minute, accompagnées des feuilles de réponses.

3.3 La codification et la grille d'analyse des résultats

Pour comptabiliser les résultats, un code alphanumérique de sept unités (XXXYYZZ) est assigné à chaque sujet. Par exemple, le premier élève de la classe de 6B de l'école ABC aurait le code ABC6B01. Les données démographiques sont codées ainsi : le niveau scolaire (1 à 6), le sexe (0 = garçon, 1 = fille), la langue parlée à la maison (1 = français, 2 = anglais) et l'ordre de passation des dictées (1 = dictée A en premier, 2 = dictée B en premier). Les mots reçoivent aussi un code selon la condition morphologique du mot court : A = les mots courts qui sont plus fréquents que leurs dérivés (R+), C = les mots courts qui sont plus fréquents que leurs dérivés et qui ont une flexion féminine (R+F) et D = les mots courts qui sont moins fréquents que leurs dérivés et qui n'ont pas de féminin (D+).

La grille d'analyse comprend deux séries de codage. La série MM (radical) calcule les résultats de l'orthographe du mot court et de la représentation de sa lettre muette : 0 = items enlevés; 1 = pas de réponse; 2 = sans lettre muette (LM); 3 = avec la mauvaise LM; 4 = avec la bonne LM; 5 = avec la bonne LM + une erreur orthographique dans le mot court. La série CCL (correspondance court/long) calcule le niveau de correspondance entre la production orthographique et morphologique du mot court et l'orthographe du mot dérivé : 0 = ne s'applique pas ou pas de lettre muette; 1 = mauvaise correspondance; 2 = bonne correspondance; 3 = bonne correspondance + erreur dans le mot long.

Dans le prochain chapitre, l'analyse des résultats sera présentée.

CHAPITRE 4

PRÉSENTATION DES RÉSULTATS

Les résultats rapportés ci-dessous sont liés à la production de radicaux morphologiques et de leurs marqueurs morphologiques muets. On rappelle que les radicaux se terminent par une consonne muette dont la présence peut être déduite en associant le radical à un dérivé morphologique ou sa flexion féminisante. Les dérivés et flexions sont formés en greffant un suffixe au radical, ce qui permet de sonoriser la consonne cible et d'en déduire sa présence dans le radical. On rappelle aussi que les données analysées regroupent les résultats d'une épreuve de dictée de 24 paires de mots (ex : *renard/renardeau*) administrée à 431 enfants de la première à la 6^e année du primaire, pour un total de 10 344 stimuli morphologiques (consonnes muettes) et 20 688 stimuli orthographiques (mots).

Le recensement des scores de réussite orthographique et de réussite morphologique des radicaux sera présenté d'abord, suivi de l'analyse du niveau de correspondance entre le radical et son dérivé. Pour alléger la lecture, les diminutifs « ortho » et « morpho », ainsi que l'acronyme CRD (correspondance radical/dérivé), seront utilisés. On rappelle qu'il y a réussite morpho lorsque la bonne consonne muette est représentée à la fin du radical, peu importe si le mot lui-même est correctement orthographié (ex : *gantil** au lieu de *gentil*). Par ailleurs, la CRD est évaluée en comparant si la consonne cible produite dans le radical est la même représentée dans le dérivé. Quatre variables explicatives²¹ seront ensuite examinées pour expliquer la réussite morpho: deux variables propres aux sujets²², le niveau scolaire²³ et

²¹ Bien que d'autres variables soient envisageables (école, curriculum, enseignant, conscience morphologique, conscience phonologique, réussite orthographique générale et vocabulaire des sujets, ainsi que des facteurs catégoriels, structuraux, relationnels et sémantiques de mots) ils ne sont pas considérés dans le cadre limité de ce mémoire de maîtrise.

²² La variable sexe a aussi été considérée, mais les résultats ne seront pas commentés dans le cadre limité de cette recherche. Ils démontrent que le score moyen de la réussite morphologique des filles est significativement plus élevé que celui des garçons ($F(1, 350) = 4,91, p = 0.0274$), mais que le score moyen de la réussite orthographique ne diffère pas d'un sexe à l'autre ($F(1, 350) = 3.69, p = 0.0554$).

la langue maternelle, et deux variables propres aux mots, la fréquence d'occurrence et la condition morphologique.

4.1 Recensement de la réussite des radicaux morphologiques

Pour calculer les scores de réussite par niveau scolaire, un premier examen est effectué afin de déterminer l'ampleur des valeurs extrêmes. Ces valeurs sont définies par des résultats à plus de 3,5 écarts-types de la moyenne. On peut imaginer des élèves en grande difficulté d'apprentissage ou des jeunes qui n'ont pas fait l'épreuve de façon sérieuse obtenir de tels scores. Ces résultats hors-norme peuvent fausser l'observation d'un développement normal des représentations morphologiques à l'étude. Ainsi, les résultats de sept élèves (1 en deuxième, 1 en quatrième, 3 en cinquième et 2 en 6^e année) qui ne sont pas statistiquement fiables sont exclus des résultats par sujet.

Par ailleurs, pour des considérations d'homoscédasticité, des analyses statistiques initiales ont été effectuées en séparant les élèves du 1^{er} cycle des autres niveaux scolaires. On note que lorsqu'il y a homoscédasticité, les variances sont égales pour chaque sous-groupe à l'étude. La variance du taux moyen de réussite des sujets du 1^{er} cycle diffère de celle des autres cycles. En retirant les résultats des élèves de 1^{re} année de l'analyse, il a été possible de stabiliser la variance des autres niveaux scolaires et rendre leurs variances indépendantes de la moyenne totale et plus constantes sur l'ensemble de l'échantillon. Deux ANOVAS ont été effectuées, une incluant les élèves de 2^e année et l'autre les excluant, pour comparer les taux de variances des sujets de la deuxième à ceux des quatre autres niveaux. Les variances n'étaient pas significativement différentes ($p < 0.0001$) et les effets des variables étaient statistiquement similaires ($p < 0.05$). Les résultats des élèves de 2^e année n'ont donc pas d'effet de déstabilisation des variances sur l'ensemble de la population et sont statistiquement fiables.

D'autres considérations remettent cependant en question l'utilisation de l'échantillon des sujets de la 1^{re} année. Le nombre de participants pour le groupe anglophone de 1^{re} année est fort différent en fonction des garçons et des filles (14 garçons versus 3 filles). Aussi, le score moyen

²³ Une comparaison des résultats par âge et par niveau scolaire, avec ajustement de Tukey-Kramer, démontre une corrélation statistique entre ces deux variables ($p < 0.0001$) donc seuls les résultats par niveau scolaire ont été considérés.

des garçons anglophones est très élevé (40.2 %) par rapport à celui des filles anglophones (19.4 %) et des francophones (26.3 % pour les garçons francophones et 27.2 % pour les filles francophones), ce qui ne semble pas représentatif de l'ensemble de cette population. En conséquence, les résultats des élèves de 1^{re} année sont exclus de l'analyse des résultats.

Finalement, lorsque la variance dépend d'une moyenne et que les données sont trop étalées, comme c'est le cas dans notre expérimentation (les résultats de la dictée sont exprimés en pourcentages étalés de 0 à 100 % de réussite), une transformation arc-sinus de la racine carrée ($p' = \text{arc sin } p$) est utilisée pour normaliser les données. Les moyennes de réussite par niveau scolaire rapportées ci-dessous sont donc des données normalisées par transformation arc-sinus tandis que les écarts-types, les scores minimums, médians et maximums sont des données brutes, non transformées.

Le premier objectif de cette recherche est de documenter la progression des scores de réussite entre le premier et le 3^e cycle du primaire. La première question posée est la suivante. Est-ce que le niveau scolaire joue un rôle sur la réussite orthographique et morphologique des radicaux morphologiques? Dans cette perspective, on analyse la progression des scores de réussite du début à la fin du primaire, puis on évalue si les deux habiletés évoluent de façon similaire. Les analyses de variance utilisées pour comparer les scores moyens de réussite orthographique ($F_{19} = 16,33$, $p < 0.0001$) et de réussite morphologique ($F_{19} = 15.64$, $p < 0.0001$) montrent des différences statistiquement significatives. On compare deux à deux les scores de réussite d'un niveau scolaire à l'autre. Le tableau 4.1 présente les probabilités (p-valeurs) que le changement de niveau scolaire n'agisse pas sur les taux de réussite.

Tableau 4.1
Probabilité (p) d'un effet du niveau scolaire sur la réussite des radicaux:
comparaison deux à deux de l'effet entre les niveaux

COMPARAISON PAR NIVEAU SCOLAIRE	RÉUSSITE ORTHOGRAPHIQUE P VALEURS	RÉUSSITE MORPHOLOGIQUE P VALEURS
2 c. 3	< 0.0001	< 0.0001
3 c. 4	= 0.0040	= 0.0065
4 c. 5	= 0.0183	= 0.0058
5 c. 6	= 0.1132	= 0.1049

On observe qu'il y a une corrélation significative entre le niveau scolaire et la réussite orthographique des radicaux ($F(4, 350) = 54.46, p < 0.0001$), ainsi qu'entre le niveau scolaire et la réussite morphologique ($F(4, 350) = 55.56, p < 0.0001$) des radicaux, sauf de la cinquième à la 6^e année.

La figure 4.1 présente la progression des scores moyens, des écarts-types des moyennes et des scores médians de la réussite orthographique et morphologique, de la deuxième à la 6^e année du primaire.

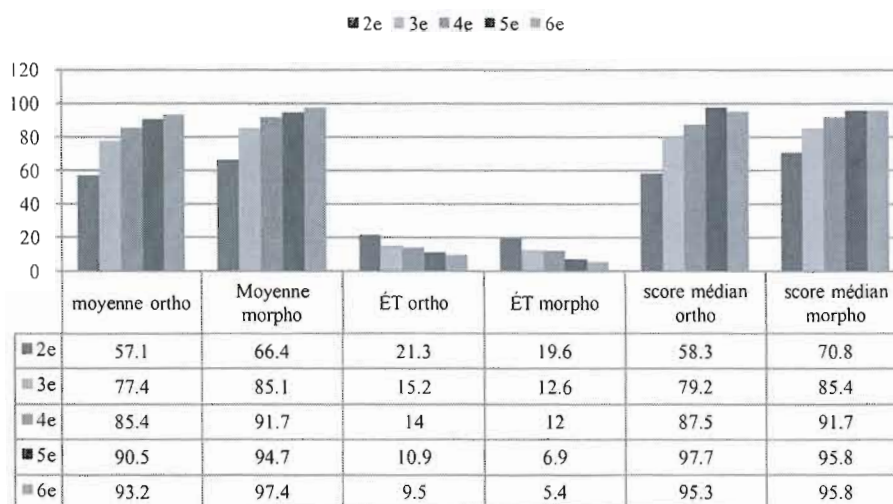


Figure 4.1
Progression des réussites orthographique et morphologique,
des écarts-types des scores médians

4.1.1 La progression de la réussite orthographique

Comme on pouvait s'y attendre, la réussite orthographique s'améliore avec le niveau de scolarisation, mais c'est entre la fin du 1^{er} cycle et le début du deuxième que l'augmentation est la plus appréciable. À la fin de la 2^e année, 57,1 % des radicaux sont correctement orthographiés, puis le score moyen de réussite fait un bond de plus de 20 points de pourcentages en 3^e année, jusqu'à 77,4 %. La progression est ensuite plus modérée. On note

niveau de plafonnement (< 5 % d'erreurs). Finalement, on observe que les écarts-types diminuent progressivement de 21,3 points en 2^e à 9,5 points en 6^e, ce qui indique que la divergence entre les résultats des élèves les plus forts et des élèves les plus faibles s'atténue avec la scolarisation.

4.1.2 La progression de la réussite morphologique

Comme pour la réussite orthographique, on observe une progression de la réussite morphologique de la deuxième à la 6^e année et l'augmentation est la plus appréciable entre la 2^e et la 3^e année. Seule l'augmentation entre la 5^e et la 6^e n'est pas significative. Dès la 2^e année, 66,4 % des consonnes muettes terminales sont correctement représentées, ce qui supplante la note de passage scolaire de l'école québécoise (>60 %), et le plafonnement est atteint dès la 5^e. On observe des scores médians similaires aux scores moyens à tous les niveaux (± 3 % de différence). Il est intéressant de noter qu'à la fin du 1er cycle, le plus faible des élèves réussit tout de même à représenter correctement une consonne muette sur cinq (score minimum de 20,8 %) et que le plus fort a déjà atteint le niveau du plafonnement. Les écarts-types passent de 19,6 points en 2^e à 5,4 % en 6^e.

4.1.3 Comparaisons de la progression orthographique et morphologique

À la figure 4.2, on observe que les écarts-types des moyennes orthographiques et morphologiques sont similaires en début de parcours, soit de 21,3 points et 19,6 points, respectivement, mais qu'ils s'éloignent en fin de parcours, soit 9,5 et 5,4 points, respectivement. Par ailleurs, les plus grandes diminutions d'écarts-types se font entre les cycles plutôt qu'à l'intérieur des cycles. Il est cependant intéressant de noter qu'à la fin du primaire, l'écart-type de la réussite orthographique est presque le double de celui de la réussite morphologique. La compétence des enfants à produire les marqueurs morphologiques muets devient plus homogène avec l'apprentissage que la compétence à orthographier correctement les mots en entier. Cela suggère que l'apprentissage des consonnes muettes est différent d'un l'apprentissage lexical.

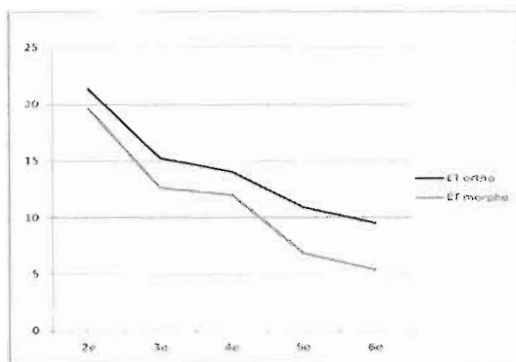


Figure 4.2
Comparaison de la progression des écarts-types

4.1.4 La correspondance morphologique entre les radicaux et les dérivés

Notre deuxième question de recherche est de déterminer si la lettre muette représentée dans le radical reflète l'orthographe du dérivé, c'est-à-dire s'il y a une correspondance morphologique entre les deux mots (CRD), par exemple *renard*/*renardeau*. Sans lettre muette, la correspondance morphologique ne se fait pas, comme *renar*/renardeau*. Cependant, si la mauvaise lettre muette est représentée à la fin du radical, mais elle est aussi produite dans l'orthographe du mot dérivé, on peut considérer que l'enfant a fait un lien morphologique entre les deux mots, et l'on considère qu'il y a CRD, par exemple **gentie/gentiesse*. Cette correspondance morphologique erronée est observée dans 4,7 % des paires de mots où la mauvaise consonne est représentée à la fin des radicaux morphologiques. Une synopsis de la CRD par niveau est présentée au tableau 4.2.

92,5 % des radicaux sont produits avec une lettre muette finale quelconque, 85,4 % avec la bonne consonne muette et 7,1 % sans lettre muette. Il y a correspondance morphologique (CRD) dans 84,8 % des paires de mots, peu importe s'ils partagent la bonne lettre cible ou si l'orthographe des mots est réussie. Lorsque le radical est orthographié correctement, 79 % des dérivés le sont aussi. Lorsqu'il y a réussite morphologique seulement, 96 % des dérivés contiennent la consonne cible, 46,1 % sont correctement orthographiés (CRD ortho) et 49,9 %

Tableau 4.2
Synopsis de la correspondance entre les radicaux morphologiques et dérivés (CRD)

RÉUSSITE RADICAL	CORRESPONDANCE RADICAL/DÉRIVÉ				TOTAL
	NE S'APPLIQUE PAS	ABSENCE DE CRD	CRD MORPHO	CRD ORTHO	
ENLEVÉ, SANS RÉPONSE, AUTRE	32				32
PAS DE LETTRE MUEFFE	645				645
MAUVAISE LETTRE MUEFFE	2	611	30		643
RÉUSSITE MORPHO SEULEMENT	2	19	260	240	521
RÉUSSITE ORTHO	9	50	1442	5706	7207
TOTAL	690	680	1732	5946	9048

La correspondance par niveau scolaire

Lorsque l'on observe la correspondance morphologique par niveau scolaire (tableau 4.3), dès la 2^e année, les deux tiers des enfants qui produisent une consonne muette produisent la même dans le dérivé. Le niveau de correspondance augmente graduellement pour atteindre plus de 90 % au 3^e cycle. Pour comparer la CRD ortho et morpho par niveau scolaire, une série d'analyses plus détaillées est effectuée.

Tableau 4.3
Pourcentage de correspondance morphologique entre le radical et le dérivé

ANNÉE	2 ^e	3 ^e	4 ^e	5 ^e	6 ^e
CORRESPONDANCE MORPHOLOGIQUE	66 %	83 %	87 %	92 %	93 %

La série d'analyses qui suit a pour but d'estimer les probabilités qu'il y ait un lien entre la réussite du radical et la réussite du dérivé. En d'autres mots, on veut savoir comment les sujets représentent les dérivés lorsqu'ils ont correctement orthographié le radical et lorsqu'ils ont simplement produit la bonne consonne muette. Pour établir ce lien par sujet, on utilise un test non paramétrique du signe qui estime la probabilité que les enfants produisent plus, moins ou autant de CRD ortho que de CRD morpho.

Les résultats de ces probabilités sont présentés au tableau 4.4 pour les radicaux qui sont correctement orthographiés et au tableau 4.5 pour les radicaux qui ont la bonne consonne muette, mais ne sont pas correctement orthographiés. Aux deux tableaux, sont indiqués le nombre de sujets (N) par niveau pour chaque classe de réussite (ortho ou morpho), la valeur statistique du signe (M) et le niveau de probabilité bilatérale (p-valeur) que le résultat de la correspondance soit positif, négatif ou neutre: > 0 est le nombre de sujets qui ont produit plus de CRD+ortho que de CRD, < 0 est le nombre de sujets qui ont produit plus de CRD que de CRD+ortho et 0 est le nombre de sujets qui en ont produit le même nombre. On note que chaque sujet de l'échantillon, peu importe son niveau scolaire, a correctement orthographié au moins un radical et, par conséquent, a aussi correctement représenté au moins un marqueur morphologique muet.

Tableau 4.4
Probabilité (p) de réussite du dérivé lorsque la réussite du radical est orthographique

NIVEAU	N	M	P-VALEUR	> 0	< 0	0
2	64	13,5	0.0007	44	17	3
3	78	2,5	<0.0001	71	6	1
4	68	30	<0.0001	63	3	2
5	75	35,5	<0.0001	73	2	0
6	85	42	<0.0001	84	0	1

On observe au tableau 4.4 que lorsque la réussite du radical est ortho, la probabilité statistique du signe est hautement significative à tous les niveaux scolaires. Cela signifie que lorsqu'un élève épèle correctement le radical, il est plus probable qu'il épèle correctement le dérivé que simplement représenter la consonne cible avec une erreur dans le dérivé. Ce n'est qu'en 2^e année que l'on observe un nombre appréciable d'élèves (>25 %) qui produisent le dérivé correspondant avec une erreur d'orthographe. Un test du signe a aussi été effectué pour vérifier la correspondance inverse, c'est-à-dire, lorsque l'élève produit correctement le dérivé, mesurer la probabilité qu'il produise le radical correctement aussi. À tous les niveaux scolaires, la probabilité est hautement significative ($p < 0.0001$) que les élèves qui épèlent le dérivé sans erreurs épèlent aussi le radical correctement.

Tableau 4.5
 Probabilité (p) de réussite du dérivé
 lorsque la réussite du radical est morphologique seulement

NIVEAU	N	M	P-VALEUR	> 0	< 0	0
2	58	-13.5	0.0001	11	38	9
3	60	3	0.4966	30	24	6
4	46	3.5	0.3240	22	15	9
5	45	4	0.2430	22	14	9
6	41	7.5	0.0081	22	7	12

Lorsque la réussite du radical est seulement morphologique, la probabilité du test du signe n'est significative qu'en 2^e et en 6^e année (tableau 4.5). En 2^e, il est trois fois plus probable qu'un élève produise le dérivé avec une erreur orthographique. En milieu de parcours, la chance de correctement orthographier les deux mots ou simplement faire un lien morpho entre les deux mots est aléatoire. En 6^e, il est trois fois plus probable que l'élève produise sans erreurs le mot dérivé, même s'il a fait une erreur dans le radical.

Pour résumer, dès la 2^e année, on observe une corrélation entre la production de la lettre muette dans le radical et sa représentation dans le dérivé, mais ce n'est qu'en 6^e année que la production de la bonne consonne muette est associée à la réussite orthographique du dérivé.

Dans les prochaines sections, nous examinerons certains facteurs qui sont susceptibles d'expliquer les résultats observés. Les analyses présentées ci-dessous ont pour but de déterminer si la fréquence d'occurrence, la condition morphologique du radical et la langue parlée à la maison expliqueraient la production de la consonne muette à la fin du radical morphologique.

4.2 La fréquence d'occurrence

Les analyses rapportées dans la présente section ont pour objectif de déterminer si certaines variables explicatives agissent sur la production du marqueur muet des radicaux. Dans cette perspective, seule la réussite morphologique des radicaux sera observée, et ce, en contrôlant trois variables propres aux mots, la fréquence du radical, la fréquence de la famille

morphologique et la condition morphologique du radical, et une variable propre aux sujets, la langue généralement parlée à la maison.

Le deuxième objectif de cette recherche est de déterminer si des facteurs propres aux mots peuvent expliquer la réussite morphologique. La première variable examinée est la fréquence d'occurrence. On rappelle que les mesures de fréquence sont tirées de la base de données Manulex et qu'un SFI de 40 = 1 occurrence sur 1 000 000 de mots, 50 = 1 sur 100 000 mots, 60 = 1 sur 10 000 mots, ainsi de suite, et que les indices utilisés sont les SFI cumulés (à moins d'indications contraires).

Trois mesures de la fréquence (token count) seront observées pour déterminer s'ils ont un effet sur la réussite morphologique: la fréquence du radical, la fréquence de la famille morphologique entière et la fréquence des dérivés et de la forme féminisée du radical. Pour faciliter la lecture, on appellera ce dernier groupe « les dérivés », même s'ils comprennent la flexion féminisante. Des tests de distribution normale de Shapiro-Wilk sont d'abord effectués pour déterminer si les fréquences des mots sont distribuées normalement. Les SFI des radicaux et les SFI des dérivés ne sont pas distribués normalement, mais les SFI des familles morphologiques complètes le sont. Pour des raisons de fiabilité statistique, lorsque la distribution des variables est normale, les résultats rapportés sont issus d'un test paramétrique de Pearson (SFI des familles morphologiques) et lorsque la distribution n'est pas normale, les résultats rapportés sont issus d'un test non paramétrique de Spearman (SFI des radicaux et des dérivés).

4.2.1 La fréquence du radical

La question que l'on se pose est la suivante. Est-ce que la fréquence du radical agit sur la production de sa consonne muette, et, si oui, dans quelle mesure? Des analyses par sujet et par item sont effectuées pour tester l'hypothèse qu'il n'y a pas de relation linéaire entre la fréquence du radical et la réussite morphologique. Dans l'analyse par sujet, on compare la réussite morphologique des 8 mots les plus fréquents aux 8 mots les moins fréquents en fonction des indices de fréquence. On observe une corrélation hautement significative ($p < 0.0001$) à tous les niveaux scolaires entre la production de la consonne muette et la fréquence

du radical. Dans l'analyse par item, on calcule d'abord le score moyen de réussite (SR) pour chacun des mots. Ensuite, en se basant sur leur rang de réussite (1 à 24), on calcule la corrélation entre la fréquence du mot et sa réussite à l'aide d'un test du coefficient de corrélation de Spearman. Les résultats présentés au tableau 4.6 donnent le niveau de corrélation entre la fréquence et la réussite morphologique (r_s), ainsi que la probabilité qu'il y ait corrélation sous l'hypothèse qu'elle ne se différencie pas de zéro (p).

Tableau 4.6
Corrélation par niveau entre la fréquence du radical et le rang de réussite morphologique

	SR_2	SR_3	SR_4	SR_5	SR_6
r_s	0.55657	0.59351	0.56763	0.21818	0.60398
p	0.0047	0.0022	0.0038	0.3057	0.0018

Coefficients de corrélation de Spearman, N = 24, Prob > |r| sous H0: Rho=0; SR = score de réussite

Les résultats indiquent qu'à tous les niveaux scolaires sauf la 5^e année, il y a corrélation entre la réussite de la lettre muette et la fréquence du radical. La lettre muette est généralement mieux représentée lorsque les mots sont fréquents et le r_s suggère une corrélation modérée ($0.50 < r_s < 0.70$) à tous les niveaux scolaires.

4.2.2 La fréquence de la famille morphologique

La famille morphologique inclut le radical lui-même (ex. *droit*), ainsi que les formes fléchies (*droits+droite+droites*) et les dérivés (*droitier+droitiers+droitière+droitières+adroit+adroits+adroite+adroites*). La question posée est : est-ce que la fréquence de la famille morphologique agit sur la production du marqueur muet dans le dérivé? Pour répondre à cette question, on teste l'hypothèse qu'il n'y a pas de relation linéaire entre la fréquence de la famille morphologique et la réussite morphologique du radical. Les résultats par sujet révèlent qu'à tous les niveaux scolaires, la corrélation entre la fréquence de la famille et la production de la consonne muette est hautement significative ($p < 0.0001$). Les résultats de l'analyse par item sont présentés au tableau 4.7. On observe une relation linéaire significative à tous les niveaux du primaire. On observe une corrélation (r_s) faible à modérée en 2^e année et une corrélation modérée à forte à partir de la 3^e année.

Tableau 4.7
Corrélation par niveau, entre la fréquence
de la famille morphologique et le rang de réussite morphologique

	SR_2	SR_3	SR_4	SR_5	SR_6
r_s	0.43179	0.49828	0.69614	0.50001	0.60899
p	0.0351	0.0132	0.0002	0.0128	0.0016

Coefficients de corrélation de Spearman, N = 24, Prob > |r| sous H0: Rho=0; SR = score de réussite

Si l'on compare les corrélations (r_s) avec la fréquence du radical aux corrélations avec la fréquence de la famille, on observe (figure 4.3) qu'en 2^e et en 3^e année, la corrélation radical/réussite a plus de poids que la corrélation famille/réussite. Inversement, en 4^e et en 5^e année, la corrélation famille/réussite est plus élevée que la corrélation radical/réussite. En 6^e, les 2 coefficients sont presque identiques.

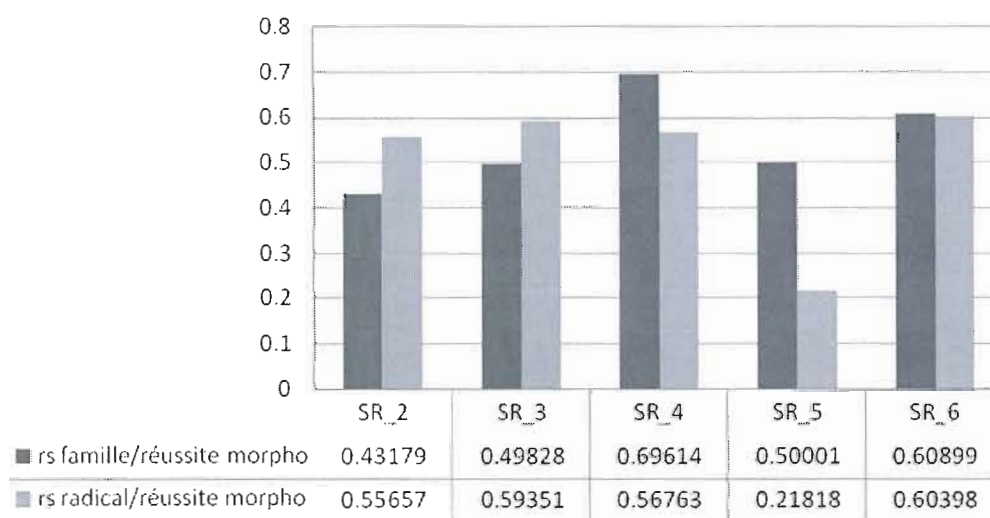


Figure 4.3
Comparaison de la corrélation (r_s) réussite/fréquence du radical
à la corrélation réussite/fréquence de la famille morphologique

On peut donc se demander si les dérivés morphologiques dans lesquels on sonorise la consonne muette ont un effet indépendant sur la production du marqueur muet dans le radical. Pour vérifier s'il existe un tel effet, on teste l'hypothèse qu'il n'y a pas de relation linéaire entre le rang de réussite morphologique du radical et la fréquence des dérivés,

linéaire entre le rang de réussite morphologique du radical et la fréquence des dérivés, excluant la fréquence du radical lui-même. On observe que les probabilités (p) de corrélation sont non significatives donc il n'existe aucune corrélation indépendante entre la fréquence des dérivés et la production de la lettre muette dans le radical.

On teste ensuite si la production du marqueur morphologique muet varie en fonction de certaines conditions morphologiques de radical.

4.3 La condition morphologique du radical

L'analyse qui suit consiste à évaluer si la production de la consonne cible dans le radical varie en fonction d'une flexion féminisante ou en fonction de la différence de fréquence entre le radical et le dérivé. On rappelle que chaque paire de mots répond à l'une des trois conditions morphologiques suivantes ($N=8$).

R+F: le radical est plus fréquent que ses dérivés et peut être féminisé (Ex : *gentil*)

R+: le radical est plus fréquent que ses dérivés, mais ne peut pas être féminisé (Ex : *nord*)

D+: les dérivés sont plus fréquents que le radical qui ne peut pas être féminisé (Ex : *mont*)

Pour évaluer si les conditions morphologiques sont des variables explicatives, deux questions sont posées. L'existence d'une flexion féminisante agit-elle sur la production des consonnes muettes? La production des consonnes muettes diffère-t-elle si les dérivés sont plus fréquents que leur radical? Si oui, dans quelle mesure? Pour y répondre, on doit d'abord valider la différence statistique entre les fréquences des radicaux morphologiques regroupés sous chacune des conditions. À l'aide d'un test bilatéral non paramétrique de Wilcoxon, on vérifie d'abord l'hypothèse que les fréquences ne sont pas statistiquement différentes sous chacune des trois conditions. On constate au tableau 4.8 que la fréquence des R+F n'est pas statistiquement différente de celle des R+ et D+ ($p > 0.05$), mais que les D+ sont statistiquement moins fréquents que les R+ et les R+F ($p < 0.05$). Cela confirme qu'il est moins probable de rencontrer les mots D+ dans les livres destinés aux enfants, mais qu'il est tout aussi probable de rencontrer les R+ que les R+F.

Tableau 4.8
P-valeurs de la fréquence d'occurrence
sous les conditions de féminisation et de fréquence d'occurrence (Wilcoxon)

CONDITION	P-VALEURS SFI-2	P-VALEURS SFI 3-5
R+F c. R+ et D+	0.3123	0.3746
D+ c. R+ et R+F	0.0008	0.0024

4.3.1 La différence de fréquence entre le radical et le dérivé

Une fois la différence de fréquence relative validée, on teste si la condition morphologique d'être moins fréquent que son dérivé agit sur la réussite de la consonne muette. Dans la mesure où les R+ sont plus fréquents que les D+ et qu'il y a corrélation établie entre la fréquence et la réussite, on considère qu'une différence entre les scores de réussite morphologique des deux groupes de mots indiquerait que la fréquence relative agit sur la production des marqueurs morphologiques muets des radicaux. Pour tester cette hypothèse, on utilise un test des rangs signés de Wilcoxon. Il mesure la probabilité que la réussite morphologique soit la même sous l'une ou l'autre des conditions morphologiques R+ et D+. Les résultats indiquent que la fréquence relative des dérivés morphologiques n'a aucun effet sur la réussite de la consonne muette. On refait le même test en comparant les D+ aux deux groupes de mots fréquents (R+ et R+F) et l'on n'observe aucune différence dans les scores de réussite morphologique (voir tableau 4.9), qu'il y ait un féminin ou non. Ainsi, le fait qu'un radical soit plus ou moins fréquent que son dérivé n'a aucun effet sur la production du marqueur morphologique muet à la fin du radical.

Tableau 4.9
P-valeurs de la corrélation entre les scores de réussite et les conditions morphologiques
utilisant un test bilatéral non paramétrique de Wilcoxon.

Niveau scolaire	2 ^e	3 ^e	4 ^e	5 ^e	6 ^e
Fréquence dérivé > fréquence radical	0.5604	0.2565	0.4991	0.6872	0.2912
Avec féminin vs sans féminin	0.6679	1.0000	0.6230	0.0631	0.5144

Coefficients de corrélation de Spearman, N = 24, Prob > |r| sous H0: Rho=0

On refait la même analyse statistique pour évaluer si la flexion féminisante agit sur la réussite.

4.3.2 La flexion féminisante

Pour vérifier si la féminisation d'un radical est liée à sa réussite morphologique, on évalue l'hypothèse qu'il n'y a pas de corrélation entre la production de la consonne muette du radical et l'existence d'une flexion féminisante. Dans la mesure où les R+F ne sont généralement pas plus fréquents que les R+, on considère qu'une différence entre les scores de réussite morphologique des R+F indiquerait que la possibilité de féminiser un radical agit sur la production de son marqueur morphologique muet. Pour tester cette hypothèse, on utilise le test de Wilcoxon qui mesure la probabilité que la réussite morphologique soit la même sous les conditions R+F et R+. Aucune différence statistique n'est observée entre les scores des mots avec ou sans féminin, et ce, à tous les niveaux scolaires. On refait le même test en comparant les R+F aux deux groupes sans féminin (R+ et D+) et l'on n'observe aucune différence dans les scores de réussite morphologique (voir tableau 4.9). Ainsi, la possibilité de féminiser les radicaux morphologiques n'a pas d'effet sur la production de leur consonne muette.

La fréquence relative des familles morphologiques complètes n'a pas été contrôlée lors du choix de stimuli donc on pourrait argüer qu'elle pourrait agir sur la réussite morphologique des radicaux. Pour tester cette hypothèse, on compare les probabilités, deux à deux, que la fréquence morphologique familiale soit la même sous chacune des conditions morphologiques. On constate au tableau 4.10 qu'il n'y a pas de différence statistique entre les fréquences des familles morphologiques (Fam), peu importe la condition morphologique. En d'autres mots, dans les livres destinés aux enfants, la probabilité de rencontrer les mots sous l'une ou l'autre des conditions morphologiques D+ serait la même. Cependant, lorsque l'on compare la fréquence des dérivés seulement (DS), il est statistiquement plus probable de rencontrer les dérivés avec un féminin que sans, plus probable de rencontrer les dérivés des radicaux R+ que ceux des D+, mais aussi probable (à la limite statistique) de rencontrer les dérivés des D+ que ceux des R+F. Ainsi, la fréquence des dérivés n'accorderait aucun

avantage statistique aux mots qui ne peuvent pas être féminisés et, à la limite, pourrait accorder un avantage à ceux qui le peuvent.

Tableau 4.10
Comparaison (Wilcoxon) de la fréquence
des familles morphologiques complètes (FAM) à celle des dérivés seulement (DS)

CONDITION	P-VALEURS	P-VALEURS	P-VALEURS	P-VALEURS
	SFI-2 FAM	SFI3-5 FAM	SFI-2 DS	SFI3-5_DS
R+F vs. R+	0.7929	0.8748	0.0173	0.0503
R+ vs. D+	0.8747	0.8748	0.0043	0.0043
R+F vs. D+	0.4309	0.4948	0.0520	0.0520

Pour résumer les effets de fréquence et de féminisation, nos résultats suggèrent que la fréquence d'occurrence a un effet significatif sur la représentation de la lettre muette à la fin des radicaux morphologiques et ce, à tous les niveaux scolaires. La corrélation entre la fréquence du radical et la réussite de sa consonne muette n'est pas significative en 5^e année, mais la corrélation entre la famille morphologique et la réussite l'est. Par ailleurs, aucune corrélation n'est établie entre la production de la consonne muette et la fréquence de la famille morphologique lorsque l'on exclut le radical. Finalement, lorsque la fréquence du radical et de la famille morphologique sont contrôlées, aucun effet de féminisation n'est observé.

4.4 La langue parlée à la maison

Le dernier objectif de cette recherche est de déterminer si des facteurs propres aux apprenants peuvent expliquer la réussite morphologique. On rappelle que le tiers de la population de l'échantillon parle uniquement l'anglais à la maison²⁴, ce qui reflète la population de la grande région de Montréal, donc il semble d'autant plus pertinent d'examiner si la langue maternelle est un facteur explicatif de la réussite morphologique. La dernière question de

²⁴ On rappelle que les échantillons des sujets bilingues et allophones (autre que le français et l'anglais) étaient trop petits pour être représentatifs donc ils n'ont pas été retenus.

recherche est la suivante. Est-ce que la réussite morphologique varie en fonction de la langue parlée à la maison, et, si oui, dans quelle mesure?

Les résultats de l'ANOVA confirment que la réussite morphologique des radicaux varie de façon significative selon la langue maternelle à la maison ($F(1, 350) = 15,46, p = 0,0001$). Les enfants anglophones n'atteignent pas le niveau de compétence des enfants francophones. La réussite moyenne des anglophones est de 86,5 %, et celle des francophones est de 91,6 %.

Pour déterminer, par niveau scolaire, si la langue maternelle explique les résultats morphologiques indépendamment de la fréquence, on fait 10 analyses de variance à mesures répétées par sujet (2 indices de fréquence x 5 niveaux) où la variable dépendante est le score de réussite morphologique et les variables explicatives sont la fréquence et la langue maternelle. La première analyse compare les scores de réussite par niveau des anglophones et des francophones en contrôlant la fréquence des radicaux morphologiques les plus fréquents ($n = 8$) et les moins fréquents ($n = 8$). La seconde analyse compare les scores de réussite des radicaux des deux groupes linguistiques en contrôlant la fréquence de la famille morphologique.

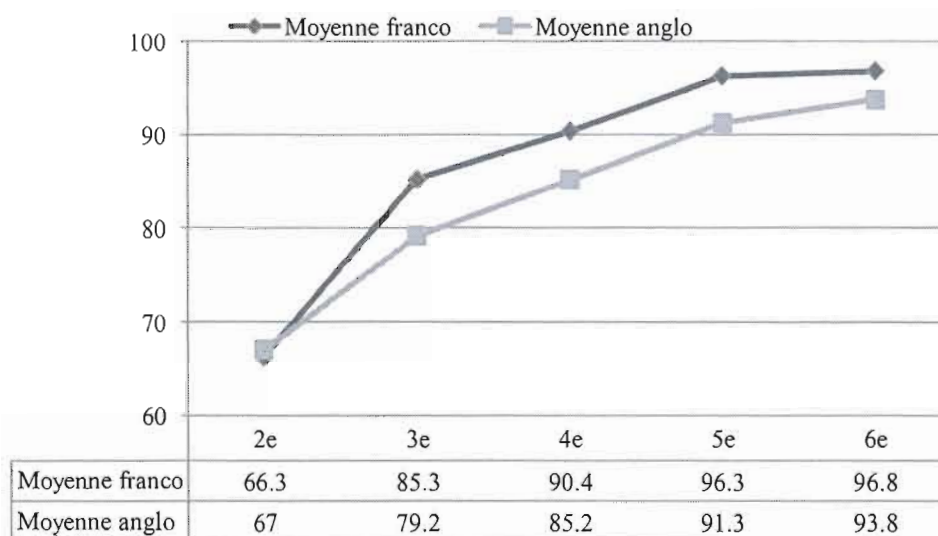


Figure 4.4
Moyenne de la réussite morphologique par niveau selon la langue maternelle

À la figure 4.4, on observe la progression de la réussite morphologique selon la langue maternelle. On observe qu'en 2^e année, la moyenne de réussite des anglophones est légèrement supérieure à celle des francophones, mais dès la 3^e année, leurs performances tombent sous celles des francophones, et ce, jusqu'à la fin du primaire.

Par ailleurs, les écarts-types des moyennes des deux groupes linguistiques diminuent avec la scolarité, de 18,4 à 4,7 points pour les francophones et de 21,2 à 6,3 pour les anglophones. L'écart entre les forts et les faibles anglophones reste toujours supérieur à l'écart entre les forts et les faibles francophones, sauf en 3^e année, mais les écarts n'étaient pas significatifs.

Le tableau 4.11 présente les probabilités qu'il n'y ait pas de différence dans les scores de réussite morphologique entre les sujets francophones et anglophones en contrôlant la fréquence du radical et de la famille morphologique.

Tableau 4.11
Probabilité (p) d'un effet de la langue maternelle
sur la réussite morphologique lorsque la fréquence est contrôlée

Mesure de fréquence contrôlée	2	3	4	5	6
Radical	0.1563	0.0943	0.1184	0.0145	0.0186
Famille morphologique	0.3481	0.0475	0.0458	0.0008	0.0305

On observe que lorsque la fréquence du radical est contrôlée, la langue a un effet significatif sur la production de sa consonne muette au dernier cycle du primaire seulement, mais lorsque la fréquence de toute la famille morphologique est contrôlée, l'effet de la langue est significatif à partir du début du 2^e cycle.

4.5 L'interaction entre la langue et la fréquence

À tous les niveaux scolaires, lorsque la valeur estimée de la variable langue est contrôlée, on observe une très haute corrélation ($p < 0.0001$) entre la réussite morphologique et l'indice de fréquence, que cela soit l'indice du radical seulement ou celui de toute la famille

morphologique. Inversement, comme nous l'observons au tableau 4.11, lorsque la fréquence est contrôlée, la corrélation entre la langue et la réussite morphologique est plus étroitement circonscrite. On peut donc se demander s'il y a interaction entre la fréquence du radical et la langue maternelle des sujets, et s'il y a interaction entre la fréquence de la famille morphologique et la langue maternelle. Le tableau 4.12 présente la probabilité des hypothèses qu'il n'y ait pas d'interaction entre la fréquence du radical et la langue ou entre la fréquence de la famille morphologique et la langue.

Tableau 4.12
Probabilité (p) d'une interaction entre la langue maternelle et la fréquence

Mesure de fréquence contrôlée	2	3	4	5	6
Radical	0.4068	0.3018	0.8826	0.0486	0.0292
Famille morphologique	<0.0001	0.0007	0.4872	0.0054	0.0162

On observe une interaction entre la fréquence du radical et la langue maternelle en 5^e et 6^e année, tandis que l'interaction entre la langue et la fréquence de la famille morphologique est présente à tous les niveaux, sauf en 4^e année.

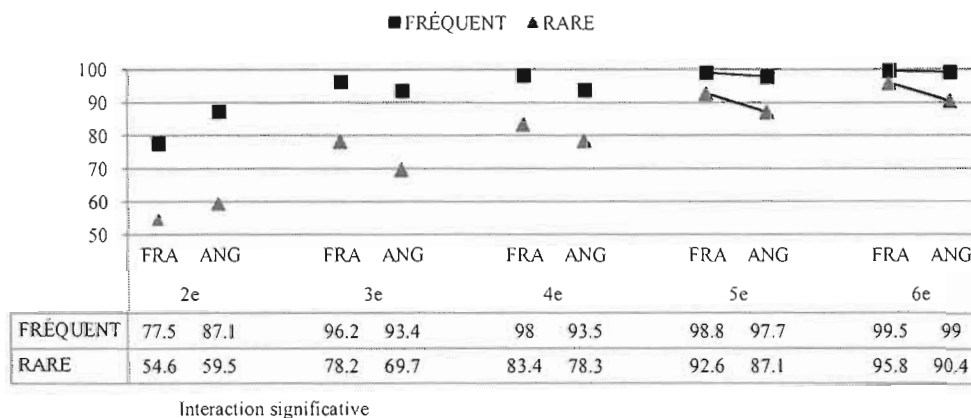


Figure 4.5
Réussite morphologique des radicaux morphologiques plus ou moins fréquents

Les figures 4.5 et 4.6 illustrent la différence entre les scores de réussite morphologique des deux groupes linguistiques. La figure 4.5 oppose les scores de réussite des radicaux

morphologiques les plus fréquents aux radicaux les moins fréquents, tandis que la figure 4.6 oppose les scores de réussites des radicaux dont les familles sont les plus ou les moins fréquentes. Les marqueurs carrés (supérieurs) représentent les scores moyens de réussite des mots fréquents et les triangles (inférieurs) représentent les scores des mots rares. Là où il y a une ligne entre les marqueurs, l'interaction langue/fréquence est significative.

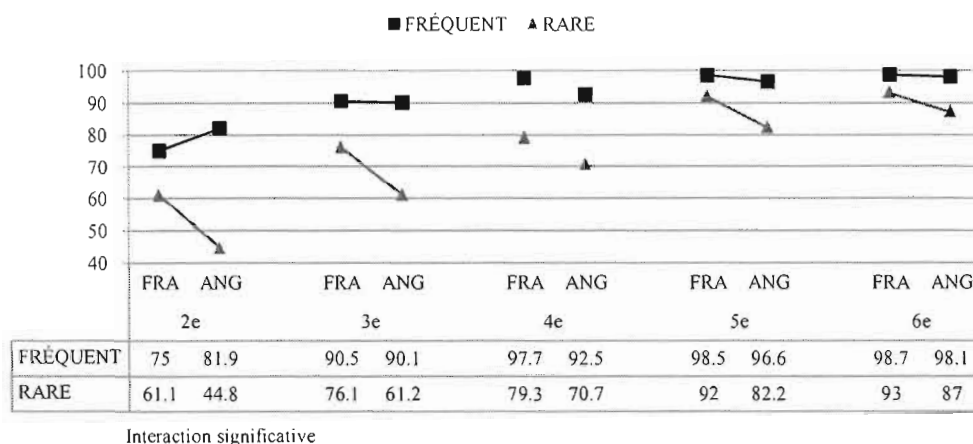


Figure 4.6
Réussite morphologique des radicaux dont les familles
sont plus ou moins fréquentes

Dans tous les cas d'interaction (ligne), on observe la même tendance: les scores des mots rares sont significativement mieux réussis par les francophones que les anglophones (ligne du bas en pente descendante) tandis que les scores des mots fréquents ne sont pas significativement différents entre les deux groupes d'élèves (ligne du haut moins pentue).

En résumé, la langue maternelle des apprenants a un effet significatif sur la représentation des consonnes muettes à la fin des radicaux morphologiques. Lorsque l'on contrôle la fréquence des mots eux-mêmes l'effet n'est observable qu'au 3^e cycle, mais lorsque l'on contrôle la fréquence de la famille morphologique entière, l'effet apparait dès le début du 2^e cycle. Par ailleurs, on observe aussi une interaction significative entre la langue et la fréquence du radical au 3^e cycle. C'est-à-dire que les francophones représentent mieux la consonne à la fin des mots rares, mais les anglophones sont tout aussi compétents avec les mots fréquents.

Cette interaction est cependant présente à tous les niveaux sauf la 4^e année lorsque l'on contrôle la fréquence cumulative de la famille morphologique entière.

Dans la prochaine section, on se penchera sur l'analyse des erreurs par niveau scolaire.

4.6 L'analyse des erreurs morphologiques

Dans cette section, on documente la typologie des erreurs en général, selon la langue maternelle et selon la fréquence des mots, et la distribution des erreurs selon la lettre de substitution, selon les mots qui causent le plus de problèmes et selon le type de syllabes terminales.

4.7 La typologie des erreurs morphologiques

Deux types d'erreurs morphologiques sont possibles dans la production des consonnes muettes terminales : l'omission ou la substitution. Sur toutes les erreurs recensées (1287), 42 % étaient des omissions (539) et 58 % des substitutions (748). À la figure 4.7, on observe la tendance dans le type d'erreur de niveau en niveau. À tous les niveaux scolaires, les substitutions sont plus fréquentes que les omissions. Entre la 2^e à la 4^e année, l'écart entre les deux types d'erreurs varie entre 5 et 15 points, mais au 3^e cycle, l'écart bondit à 42 et 50 points de pourcentage, respectivement en 5^e et 6^e.

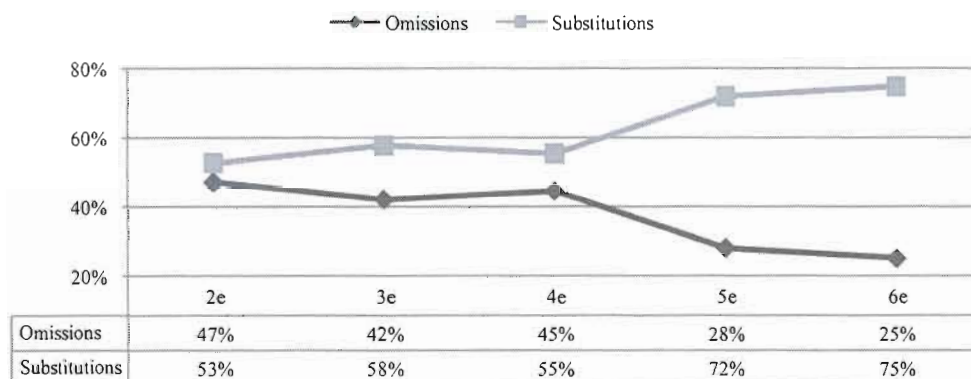


Figure 4.7
Typologie des erreurs par niveau scolaire

4.8 La typologie des erreurs par groupe linguistique

36 % des 377 sujets de la 2^e à la 6^e année sont anglophones, mais ils produisent 52 % des erreurs morphologiques recensées, plus particulièrement, 47 % des omissions et 57 % des substitutions (voir tableau 4.13).

Tableau 4.13
Typologie des erreurs par groupe linguistique

Langue	Omissions	Substitutions	Total des erreurs
Francophones	53 %	44 %	48 %
Anglophones	47%	56%	52 %

La tendance dans la progression de la typologie des erreurs est similaire pour les deux groupes linguistiques (voir figures 4.8 et 4.9). Il y a cependant quelques différences marquantes. Le pourcentage de substitutions des anglophones est plus élevé que celui des francophones, et ce, à tous les niveaux scolaires. Par ailleurs, contrairement aux anglophones qui produisent plus de substitutions que d'omissions à tous les niveaux, cette tendance n'apparaît qu'au 3^e cycle chez les francophones. On constate cependant que la différence entre le pourcentage de substitutions et d'omissions des deux groupes linguistiques fait un bond important entre les 2^e et 3^e cycles.

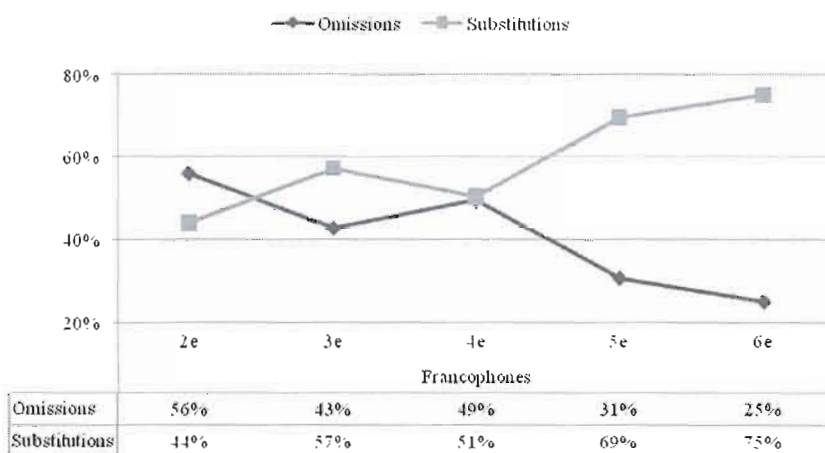


Figure 4.8
Typologie des erreurs des francophones, par niveau scolaire

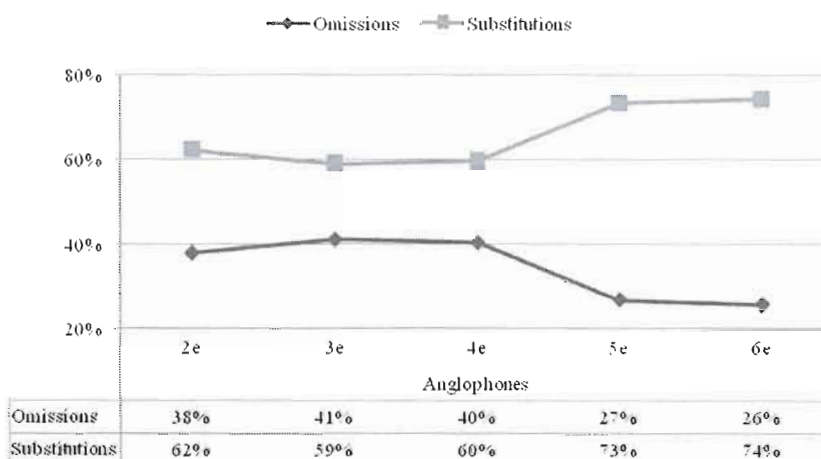


Figure 4.9
Typologie des erreurs des anglophones, par niveau scolaire

Pour vérifier si cette différence est significative, on teste l'hypothèse que la typologie des erreurs est la même pour les deux groupes linguistiques. Un test des rangs signés par sujet démontre que la probabilité qu'un élève fasse plus de substitutions que d'omissions n'est statistiquement significative que pour les anglophones du 3^e cycle ($p = 0.0032$ en 5^e année et $p = 0.0005$ en 6^e année).

4.9 La typologie des erreurs selon la fréquence des mots

Comme il a été démontré, la fréquence du radical morphologique agit sur la réussite de sa lettre muette, il y a plus d'erreurs dans les mots rares que fréquents. Plus de 75 % des erreurs (77 % des omissions et 72 % des substitutions) sont recensées dans les mots rares. Cependant, en observant la progression de la typologie des erreurs selon la fréquence (figure 4.10), on constate qu'il n'y a aucune constante entre la 2^e année et la 4^e année. Ce n'est qu'au 3^e cycle que la typologie est similaire entre les mots fréquents et rares, c'est-à-dire plus de substitutions que d'omissions dans les deux cas.

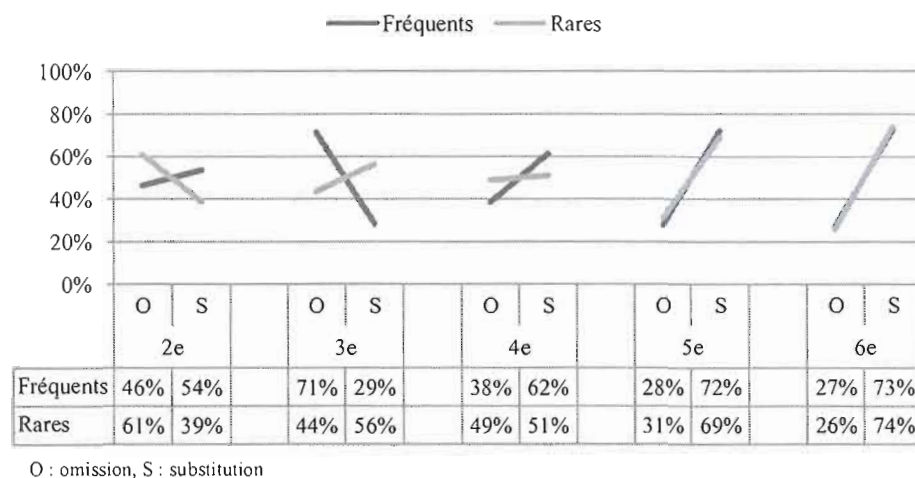


Figure 4.10

Typologie des erreurs par niveau scolaire selon la fréquence des mots

4.10 La distribution des erreurs

Finalement, quelques statistiques distributionnelles et observations qualitatives notoires sont présentées, dont les lettres les plus souvent substituées et les mots qui causent le plus d'erreurs.

Au tableau 4.14, on présente les 5 mots stimuli dans lesquels on retrouve le plus d'erreurs morphologiques au total, le plus d'omissions et le plus de substitutions. Ce « Top 5 » (20 % des mots) représente 47 % du total des erreurs morphologiques, 46 % des omissions et 58 % des substitutions. Les stimuli *permis* et *parfum* causent le plus de problèmes à tous les niveaux scolaires, représentant 25 % de toutes les erreurs. Le ratio est de 38 % et 31 %, respectivement, en 5^e et 6^e année. Par ailleurs, on observe que parmi les Top 5 omissions, trois mots se terminent par des syllabes orales fermées (CVC), *lourd*, *regard* et *bavard*. Parmi les Top 5 substitutions, deux des mots sont des homophones (*laid* et *camp*). Malgré la dominance des erreurs dans le Top 5, il est intéressant de noter que seul le mot *permis* figure parmi les mots de fréquence rare. Tous les autres stimuli de la liste des Top 5 sont de fréquence moyenne (ni parmi les 8 plus fréquentes, ni parmi les 8 moins fréquentes).

Inversement, seul le mot *lourd* ne figure pas parmi les familles morphologiques les plus rares (il arrive cependant au 9^e rang sur 24).

Tableau 4.14
Top 5 mots par types d'erreur

TOTAL DES ERREURS		OMISSIONS		SUBSTITUTIONS	
permis	15 %	permis	14 %	permis	17 %
parfum	10 %	lourd	11 %	parfum	16 %
laid	8 %	regard	8 %	laid	11 %
épais	7 %	bavard	7 %	épais	8 %
lourd	7 %	éclat	6 %	camp	6 %
TOP 5	47 %	TOP 5	46 %	TOP 5	58 %

Le mot *permis* occupe le premier rang des erreurs morphologiques à tous les niveaux scolaires, autant en omission qu'en substitution. 73 % des erreurs de substitutions dans ce mot sont avec *-t* et, dès la 3^e année, cela grimpe au-delà de 85 %. Le mot *parfum*²⁵ arrive au second rang des erreurs morphologiques à tous les niveaux scolaires, sauf en 6^e année, où il occupe le premier rang. Comme on peut s'y attendre, ce sont des erreurs de substitutions du *-um* au *-un*, dont 89 % dans la graphie **parfun* et le reste dans les graphies **parfunt*, *parfuns* et *parfune*.

Par ailleurs, à partir de la fonction « Graphemic Segmentation » de Manulex-infra (Peereman et coll., 2007), il est possible de recenser les lettres que l'on retrouve le plus fréquemment à la fin des mots. Il s'agit du *-e* (à la fin de 12 717 mots sur 48 866), du *-s* (12 172), du *t* (8254) et du *n* (1348). Ces quatre lettres occupent la position finale dans 71 % de tous les mots recensés dans cette base de données (tous niveaux scolaires confondus et incluant toutes les flexions de genre, de nombre et de personne). Ces contraintes graphotactiques sont observées dans la typologie des erreurs de substitution observées. Comme on peut le voir à la figure

²⁵ Le mot *parfum* est un cas spécial. Le *-m* n'est pas un marqueur morphologique à strictement parler, étant donné qu'il fait aussi partie du graphonème *-um = /œ/*. Il a été décidé d'inclure le mot vu la difficulté de trouver un autre mot qui répondait à toutes les conditions, et vu que sa consonne finale peut être déduite du mot dérivé.

4.11, ces quatre lettres représentent 88 % de toutes les substitutions recensées. On note, toutefois, que 97 % des substitutions au *n* sont dans la seule graphie **parfun* (parfum).

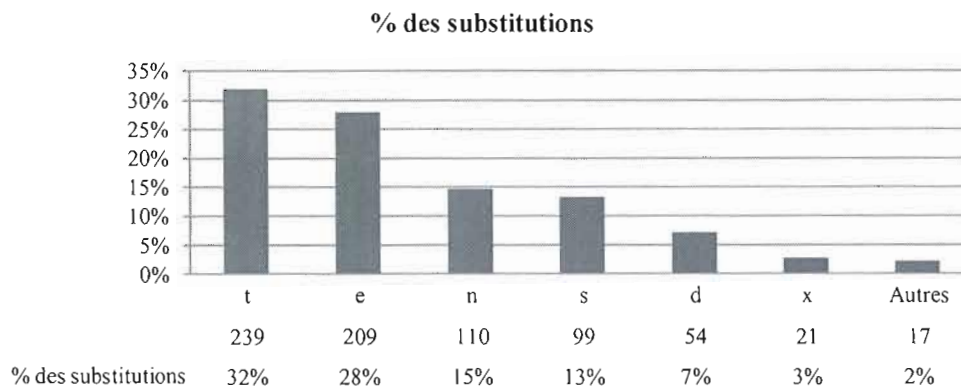


Figure 4.11
Typologie des substitutions

Par ailleurs, le Top 5 des substitutions représente 79 % de toutes les substitutions à la lettre *t*, dont 38 % se retrouve dans la graphie **permit* (permis) et 21 % dans la graphie *lait* (laid). Le Top 5 représente aussi 66 % des substitutions au *s*, dont le tiers sont dans la graphie **lais* (laid). Le Top 5 ne représente cependant que 24 % des substitutions au *e*. La majorité des substitutions au *e*, 58 %, sont produites dans les six mots qui finissent avec une syllabe orale fermée: *bavard*, *tard*, *regard*, *lourd*, *nord* et *renard* (53 % sont dans les quatre premiers).

Finalement, il semble probable que les erreurs les plus fréquentes soient attribuables aux caractéristiques phonologiques et orthographiques du français, c'est-à-dire à des régularités graphophonémiques et graphotactiques. Par exemple, le graphème *un* est dominant pour représenter le phonème /œ̃/ de */parfœ̃/* et, en plus d'être non dominant, le graphème *um* qui se prononce /œ̃/ en position terminale est unique au mot *parfum*. Par ailleurs, si la production de la graphie **permit* découle d'une analogie morphologique entre les mots *permis* et *permission*, on peut imaginer des erreurs analogiques qui découlent de la dominance de la graphie *-tion* pour représenter le graphème /sjõ/. On constate que 19 % des erreurs de substitution où l'on produit la graphie **permit* sont accompagnées d'une erreur similaire dans

représentations graphémiques du phonème /r/ terminal sont distribuées ainsi : 1620 *re*, 701 *rs*, 127 *rd* et 103 *rt*. Cela expliquerait donc le niveau élevé des substitutions au *-re* ou *-rre* dans les syllabes terminales fermées.

En résumé, l'analyse des erreurs démontre qu'il y a généralement plus de substitutions que d'omissions de la lettre muette finale à la fin des radicaux morphologiques et ce n'est qu'au 3^e cycle que cette tendance s'observe autant dans les mots rares que fréquents. Cependant, la probabilité qu'un élève produise plus de substitutions que d'omissions n'est significative que pour les enfants anglophones du 3^e cycle. Par ailleurs, la typologie des erreurs de substitution morphologique, c'est-à-dire les lettres muettes les plus souvent substituées à la fin des radicaux morphologiques, reflète les régularités graphophonémiques et les contraintes graphotactiques de la langue.

Dans la prochaine section, nous discuterons des résultats observés en faisant les liens avec la progression des apprentissages au primaire du MELS, des modèles d'acquisition du langage écrit et des observations empiriques présentées au chapitre 3.

CHAPITRE 5

DISCUSSION

Ce chapitre présente une discussion des résultats observés ci-dessus et leurs implications théoriques, méthodologiques et pratiques. La discussion s'organise en fonction des objectifs de notre recherche. Le premier est de recenser la progression des scores de réussite de radicaux morphologiques et de leur consonne terminale muette chez les enfants de l'école primaire, tandis que les autres sont d'observer si la production des consonnes muettes variait en fonction de trois variables explicatives déjà observées en acquisition du langage écrit : la fréquence d'occurrence des mots, la présence d'une flexion féminisante et la langue maternelle des enfants.

L'objectif de ce chapitre est de faire un rapprochement entre les résultats observés et ceux des recherches scientifiques qui ont aiguillé notre démarche et de les comparer à la progression des apprentissages suggérée dans le programme de formation de l'école québécoise. Les limites de nos observations seront ensuite discutées, ainsi que la portée de notre démarche scientifique dont des pistes pour les recherches futures et des implications pour les intervenants dans les écoles primaires.

5.1 La progression en production de consonnes muettes dérivables au primaire

Le premier objectif est de délimiter les niveaux auxquels les enfants représentent à l'écrit des marqueurs morphologiques muets dérivables. Ces consonnes finales à la fin de radicaux morphologiques ne sont pas représentées à l'oral, mais peuvent être justifiées par analogie

morphologique avec la lettre sonorisée dans le dérivé ou la flexion féminine. L'âge auquel les enfants sont en mesure de traiter l'information morphologique est un sujet de grands débats théoriques. Les modèles développementaux (Ehri, 1995; Frith, 1985; Seymour, 1997) avancent que les connaissances morphologiques sont traitées en fin de parcours d'acquisition de l'écrit, c'est-à-dire lorsque le traitement alphabétique est maîtrisé et le traitement orthographique des mots est fonctionnel. Plusieurs d'observations empiriques réfutent cette hypothèse (Carlisle, 1988; Pacton et coll., 2005, 2001; Sénéchal, 2000; Sénéchal, Basque et Leclaire, 2006; Totereau et coll., 1997; Treiman et Cassar, 1996).

Par ailleurs, cette hypothèse a été établie à partir du système orthographique de l'anglais dont les conventions morphologiques sont généralement représentées à l'oral, ce qui est tout le contraire du français qui a une quantité imposante d'unités sublexicales muettes qui sont plus difficiles à traiter cognitivement. Cette particularité morphologique complexifie l'apprentissage du français et pourrait laisser croire que le développement morphologique des enfants francophones serait encore plus tardif que celui des anglophones. Cependant, cela n'est pas ce que donne à entendre la progression des apprentissages proposée par le MELS qui conseille aux enfants de pratiquer, dès la 2^e année, des activités pédagogiques qui développent les connaissances analogiques morphologiques qui consistent à associer un mot à sa famille morphologique pour justifier la présence de sa lettre muette.

Nos résultats appuient la proposition implicite du MELS : les jeunes enfants sont en mesure de traiter des unités morphologiques muettes et ce, dès la fin du 1^{er} cycle. La moitié des sujets de 2^e année ont correctement produit 70 % ou plus des consonnes muettes cibles à la fin des radicaux morphologiques. Tous les enfants de l'échantillon, sans exception, ont représenté au moins une consonne muette correctement, appuyant le postulat que les enfants de 2^e année ont la capacité cognitive de traiter des unités lexicales qui ne sont pas représentées à l'oral. Ces observations s'ajoutent à un nombre croissant d'études empiriques qui réfutent la perspective développementale qui prévoit un déploiement tardif des processus cognitifs de traitement orthographique ou morphologique des mots. Les jeunes enfants, en début de formation en littéracie, auraient recours à des processus cognitifs autres que phonologiques lorsqu'ils produisent des consonnes muettes à la fin d'un radical morphologique.

Selon les recommandations du MELS, cette compétence devrait être acquise à la fin des études primaires. Les scores de réussite morphologique observés approchent le plafonnement (< 10 % d'erreurs) à la fin du 2^e cycle et plafonnent (< 5 % d'erreurs) dès la 5^e année. Un plafonnement n'a cependant pas été observé dans l'étude de Bégin, Saint-Laurent et Giasson (2010). Les scores dans les études de Sénéchal (2000) et Sénéchal, Basque et Leclaire (2006) étaient également plus faibles que ceux recensés ici. Différents facteurs pourraient expliquer ces écarts entre les différentes études sur le sujet. Nous discutons dans cette section les facteurs qui peuvent être généralisés aux apprenants (type d'épreuve, milieu socioéconomique, environnement linguistique, type d'enseignement), ainsi que la question de la fréquence des stimuli utilisés dans les différentes études.

Le type d'épreuve

Il semblerait peu probable que la variance dans les résultats soit liée aux épreuves elles-mêmes puisque la passation des épreuves et le calcul des scores de réussite étaient les mêmes. Seules les feuilles de réponses de Bégin, Saint-Laurent et Giasson (2010) étaient différentes, les élèves devaient insérer leurs réponses dans une dictée trouée, tandis que dans les autres études, les mots devaient être écrits en liste, c'est-à-dire sans contextualisation écrite. Il nous semble peu probable que la tâche de dictée soit responsable du taux de réussite plus élevé que nous avons observé.

Le milieu socioéconomique

Il est cependant envisageable qu'une partie des meilleurs résultats observés s'explique par un milieu socioéconomique favorisé. En comparant les scores des trois expérimentations, nos sujets de 2^e année, anglophones et francophones confondus atteignent les scores des sujets francophones de 4^e année dans les deux études de Sénéchal (2000) et Sénéchal, Basque et Leclaire (2006), tandis que les scores moyens de nos sujets de 4^e année dépassent ceux des élèves de 6^e de Bégin, Saint-Laurent et Giasson (2010). Bien que le niveau socioéconomique des enfants dans les expérimentations de Sénéchal ne soit pas noté, les sujets de Bégin, Saint-Laurent et Giasson étaient également issus de milieux favorisés. Le milieu socioéconomique n'expliquerait donc pas les écarts en 6^e année.

L'environnement linguistique

L'environnement linguistique différent, c'est-à-dire québécois vs franco-ontarien dans les études de Sénéchal, pourrait aussi être une variable explicative. Cette différence a été observée dans une étude qui testait l'utilisation du test de vocabulaire EVIP au Québec (Godard et Labelle, 1995). L'EVIP, une adaptation française du Peabody Picture Vocabulary Test, avait été standardisé auprès d'une population canadienne-française, mais ne reflétait pas le niveau de vocabulaire, plus élevé, observé chez des enfants québécois. Les chercheurs en concluent que « Les écarts entre les âges réels et les âges EVIP que nous avons observés laissent croire à un effet de l'environnement linguistique sur l'acquisition du vocabulaire. » (Godard et Labelle, 1995, p. 20). Indubitablement, le niveau de vocabulaire des enfants joue un rôle dans l'analogie morphologique, puisque les dérivés doivent être intégrés dans le lexique oral de l'enfant pour que le lien entre le radical et le dérivé puisse se faire. Dans la mesure où les enfants québécois semblent avoir un vocabulaire plus étendu que les enfants franco-ontariens, l'environnement linguistique pourrait expliquer une partie de la différence entre nos résultats et ceux de Sénéchal. Il n'explique toutefois pas l'écart des scores entre ceux présentés ici et ceux de Bégin.

Les approches didactiques et pédagogiques

Le type d'enseignement reçu et l'approche didactique favorisée dans les manuels scolaires des sujets pourraient être des variables explicatives des écarts observés. Bien que le contrôle de ces variables soit au-delà de la portée de cette étude, une comparaison des écarts-types des moyennes avec les études de 2000, 2006 et 2010 pourrait suggérer un enseignement différent d'une population à l'autre. Dans la présente étude, les écarts-types des scores de réussite orthographique sont similaires aux écarts-types de la réussite morphologique en début de parcours, puis ils diminuent avec la scolarisation. L'enseignement aurait ainsi un effet sur les connaissances orthographiques et morphologiques, mais l'écart-type de la moyenne orthographique régresse moins rapidement que l'écart-type de la moyenne morphologique. En 6^e année, l'écart-type de la moyenne orthographique est presque deux fois celui de la moyenne morphologique. Cela donne à penser que le type d'enseignement reçu par nos sujets aurait un effet plus marqué sur la progression de leur réussite morphologique. Cette tendance

n'est toutefois pas observée dans l'étude de Sénéchal (2000, tableau 1, p. 79), où l'écart-type de la moyenne orthographique (% correct) reste stable entre la 2^e et la 4^e année, tandis que celui de la moyenne morphologique (% final consonant correct) augmente. Contrairement à nos observations, la progression des écarts-types dans l'étude de Sénéchal (2000) semble indiquer que les connaissances morphologiques des enfants deviennent de plus en plus hétérogènes avec la scolarisation. Dans l'étude de Bégin, Saint-Laurent et Giasson (2010), où seuls les résultats d'enfants de 6^e année sont évalués, les écarts-types des moyennes orthographiques et morphologiques sont similaires, ce qui reflèterait un enseignement qui n'a pas plus d'effet sur l'acquisition des connaissances morphologiques visées que sur l'acquisition de l'orthographe d'usage. L'enseignement reçu et les manuels scolaires utilisés par les sujets de la présente étude pourraient ainsi expliquer les scores de réussite plus élevés.

La fréquence des stimuli

Finalement, nos résultats rejoignent ceux de plusieurs études où l'on a observé des effets de fréquence en production de mots. La différence de fréquence des stimuli pourrait expliquer la divergence entre nos scores de réussite morphologique et ceux des études antérieures. Pour évaluer cette hypothèse, on compare les indices de fréquence moyens des mots morphologiques des études de Sénéchal (2000), Sénéchal, Basque et Leclaire (2006) et Bégin, Saint-Laurent et Giasson (2010). À partir de Manulex-infra, nous avons comparé la fréquence moyenne (Indice G1-G5) de nos 24 mots à celle des mots aux consonnes muettes dérivables dans les trois autres études (Expérimentation 1 dans Sénéchal, Basque et Leclaire, 2006). Nous avons aussi comparé la fréquence des familles morphologiques. Ces moyennes sont présentées au tableau 5.1²⁶.

La fréquence moyenne des stimuli ne diffère pas beaucoup entre les études. Le SFI moyen est inférieur au nôtre dans les deux études de Sénéchal et supérieur dans l'étude de Bégin. Cependant, un des mots dans l'étude de Bégin, Saint-Laurent et Giasson (2010), *petit*, a un SFI de plus de 72.74, c'est-à-dire une occurrence à tous les ± 700 mots, un taux de fréquence

²⁶ L'indice « U » de la 1^{re} à la 5^e année (G1-G5) de Manulex a été recensé, puis la formule logarithmique $\text{LOG}(U+4)*10$ a été appliquée pour le traduire en SFI. On rappelle qu'un SFI de 60 est l'équivalent d'une occurrence sur 10 000 mots et un SFI de 70, une occurrence sur 1000 mots.

extrême qui biaise la fréquence moyenne des stimuli de Bégin. Aucun autre mot dans les quatre études n'approche cette fréquence.

Tableau 5.1
Comparaison des fréquences des stimuli morphologiques

	FRÉQUENCE MOYENNE DES STIMULI			
	L'HEUREUX 24 MOTS	SÉNÉCHAL 2000 20 MOTS	SÉNÉCHAL ET COLL. 2006 12 MOTS	BÉGIN ET COLL. 2010 16 MOTS (15 MOTS)
FRÉQUENCE MOYENNE DU RADICAL	60,3	59	57,9	62,8 (58,8)
FRÉQUENCE DE LA FAMILLE MORPHOLOGIQUE	63,7	61,3	62,5	65,2 (59,3)

Lorsque *petit* est exclu du calcul de la fréquence, le SFI moyen des mots de Bégin, Saint-Laurent et Giasson tombe à 58.8. On observe la même tendance dans la fréquence des familles morphologiques lorsque le mot *petit* est exclu. Même si la différence de fréquence peut sembler faible, on ne peut pas exclure qu'elle explique la variance entre les scores de réussite. Le fait qu'un mot apparaisse une fois sur 10 000 (SFI = 60) au lieu d'une fois sur 12 000 (SFI = 62) pourrait faire une grande différence dans la représentation qu'un enfant se fait de ce mot, surtout s'il n'a pas tiré avantage d'une grande exposition à l'écrit dans sa jeune carrière de lecteur.

On pourrait aussi considérer d'autres facteurs propres aux stimuli pour expliquer les différences dans les scores dont leur structure syllabique, la consistance des CGP, la dominance des CPG et le nombre de lettres, ce que nous ne ferons pas dans la limite de cette étude.

5.2 Le lien morphologique entre le radical et son dérivé

La deuxième question de recherche posée visait à examiner si la lettre muette produite à la fin du radical correspondrait à celle sonorisée dans les dérivés. Une correspondance significative a effectivement été observée. Dans 96 % des cas, lorsque la consonne cible est représentée dans le radical, elle l'est aussi dans le dérivé. Inversement, lorsque la consonne cible est représentée dans le dérivé, elle l'est dans 83 % des radicaux. Bien qu'on ne peut pas affirmer

avec certitude que les enfants font une analogie morphologique pour justifier la présence des consonnes muettes dans les radicaux (notre tâche ne portait pas spécifiquement sur le traitement analogique), les scores de réussite élevés, jumelés à une correspondance étroite entre les radicaux et les dérivés, laissent entendre que le choix de consonne muette n'est pas aléatoire.

La typologie des erreurs atteste de la sensibilité des enfants aux consonnes muettes. Deux types d'erreurs sont possibles, l'omission ou la substitution de la consonne muette. L'omission suggère que les enfants ont produit le mot en faisant abstraction de cette caractéristique lexicale particulière. Ils n'ont pas intégré le mot dans leur lexique orthographique individuel et ne connaissent pas la règle morphologique qui justifie la consonne muette. La substitution établit que le mot n'est pas lexicalisé, mais que l'enfant est sensible à la présence d'une lettre muette terminale, une régularité orthographique du français. Comme Bégin, Saint-Laurent et Giasson (2010), nous avons observé plus d'erreurs de substitution que d'omission, et ce, dès la 2^e année.

Même si les jeunes enfants n'ont pas de représentations orthographiques justes des mots, ou ne connaissent pas la règle qui justifie la consonne muette, ils démontrent une sensibilité à la régularité des syllabes terminales avec lettres muettes. Ils savent intuitivement mettre une lettre muette sans nécessairement savoir laquelle. Cette sensibilité a aussi été démontrée dans l'ajout de consonnes muettes où il ne devrait pas y en avoir (Pacton et Casalis, 2006), un phénomène que nous avons recensé dans l'orthographe du mot *parfum* (**parfunt*, **parfuns*, **parfune*). On retrouve aussi des ajouts de consonne après le *-um*, ce qui laisse croire qu'une analogie quelconque aurait été faite, même si la raison d'être de la lettre muette n'est pas connue. Par ailleurs, une erreur comme **parfune* suggère que l'élève n'a pas compris le rôle phonologique du *e muet*, qui empêche la prononciation /œ/ puisqu'il crée une syllabe écrite ouverte, *-ne* = /n/. Quelques élèves ont produit les graphies **parfume*, **parfunm* et **parfumn*, suggérant qu'ils ont des représentations morphologiques, même lorsque les conventions phonologiques et les contraintes graphotactiques ne sont pas acquises.

Entre la 2^e et la 4^e année, la sensibilité aux consonnes muettes n'est pas régularisée, c'est-à-dire que les enfants peuvent aussi bien produire des radicaux morphologiques avec ou sans

consonne muette. Au 3^e cycle, cependant, il devient statistiquement plus probable qu'un enfant soit sensible à la présence d'une consonne muette que l'inverse. Lorsqu'un mot n'est pas lexicalisé, la sensibilité des enfants du 3^e cycle qui substituent les mauvaises consonnes muettes ne serait pas propre aux mots, mais à la caractéristique sublexicale du français : une lettre muette à la fin des mots. Comme le mentionnent Bégin, Saint-Laurent et Giasson (2010),

« C'est comme si les élèves sont sensibles au fait qu'il existe une consonne finale à des mots comme « début » ou « drap » (les substantifs en français se terminent rarement par une voyelle), mais qu'ils ne recourent pas encore systématiquement à une règle morphologique telle que : « *Début* se termine par un *t* parce qu'il appartient à la même famille que *débuter*, *drap* se termine par un *p* parce qu'on dit *draper*, *draperie* ». (2010, p. 292)

On ne peut pas savoir sans recenser les enfants eux-mêmes, si ceux qui produisent la bonne consonne muette ont recours à l'analogie morphologique pour y arriver. Cependant, les instances de double substitution observées, c'est-à-dire la mauvaise lettre muette est aussi représentée dans le dérivé, confirment qu'un lien morphologique est représenté entre les mots. Par exemple, plusieurs enfants qui ont produit la graphie **gentie* ont aussi représenté le *e muet* dans le dérivé mot **gentiesse*, **gentiès*, **jantiessse*. La double substitution était fréquente dans la production de la paire *permis/permission*, orthographiée **permit/permission*. Cette sensibilité à l'information morphologique dissimulée dans les mots est présente, même lorsque le code phonologique n'est pas acquis. Par exemple, la graphie **parfume* indique que certains élèves sont sensibles aux liens morphologiques mais n'ont pas compris le rôle phonologique du *e muet*, puisque son ajout change la syllabe terminale ouverte /œ/ à la syllabe fermée /yn/.

Les trois autres objectifs de cette recherche visaient à considérer certaines variables explicatives de la réussite morphologique.

5.3 L'effet de fréquence et la production de consonnes muettes

Les résultats rapportés ci-dessus appuient le postulat central des théories *usage-based* voulant que la fréquence agisse sur la vitesse avec laquelle l'orthographe des mots est représentée en mémoire (Brown et Loosemore, 1995; McClelland et Rumelhart, 1981; McClelland et coll., 1986; Monaghan et Ellis, 2010; Perfetti, 1991; Rumelhart et McClelland, 1987; Share, 1995).

La fréquence des mots agit sur la représentation écrite de leur consonne muette, et ce, dès la 2^e année. Les mots fréquents sont mieux réussis que les mots rares, tant chez les francophones que les anglophones en immersion. Ces observations rejoignent celles de plusieurs chercheurs qui ont observé de tels effets dans l'orthographe d'usage de jeunes enfants francophones (Alégria et Mousty, 1997; Lété, Peereman, et Fayol, 2008; Martinet, Valdois, et Fayol, 2004; Pacton et coll., 2001; Sprenger-Charolles, Siegel, et Bonnet, 1998). Elles concordent aussi avec les effets de fréquence précoces observés dans la production de mots qui requièrent des connaissances morphologiques (Largy et coll., 2004), ainsi que les effets de fréquence dans la production de consonnes muettes dérivables, observés chez des enfants plus âgés (Bégin, Saint-Laurent et Giasson, 2010).

Sénéchal (2000) et Pacton et Casalis (2006) avaient aussi suggéré que la fréquence de la famille morphologique au complet pourrait être une variable explicative du taux de réussite de la consonne muette. Ils n'ont toutefois pas contrôlé cette variable, ni l'effet de redondance du radical lorsqu'il est rencontré seul et à l'intérieur des dérivés ou de la flexion féminisante. Le contrôle des deux niveaux de fréquence, celui du radical et celui de la famille morphologique, révèle que la fréquence des dérivés n'a pas d'effet indépendant sur la production de la consonne muette.

Cependant, à la fin du 2^e cycle, un effet de redondance semble apparaître. En 2^e et en 3^e année, l'ajout d'occurrences du radical, enchâssées dans les dérivés n'ajoute aucun poids au coefficient de corrélation entre la fréquence et la réussite de la consonne muette, mais en 4^e et en 5^e année, on observe l'inverse. Les occurrences de toute la famille ont plus d'effet sur la réussite de la lettre muette que les occurrences du radical seul. Toutefois, en 6^e année, l'effet s'estompe. Le coefficient de corrélation avec l'un ou l'autre des indices de fréquence est quasi identique. Cette courbe en « U » de l'effet de redondance pourrait suggérer que

l'expansion de l'exposition à l'écrit favorise le développement de l'analogie morphologique, mais que la lexicalisation des radicaux en fin de parcours supplante la nécessité de faire ce lien. Autrement dit, en 2^e-3^e année, la production de la lettre muette ne semble pas influencée par la présence de dérivés, ce qui suggère qu'elle est le résultat d'un apprentissage lexical, « usage-based ». Par contre, les enfants de 4^e-5^e année tireraient parti de l'augmentation de leur lexique orthographique, qui inclurait des mots dérivés, pour choisir la lettre muette en lien avec ces mots dérivés. Ce début de l'influence de la famille morphologique rejoint aussi les observations de Sénéchal, Basque et Leclaire (2006). Seule une minorité des participants de 4^e année, entre 18% et 34 %, disent appliquer une stratégie d'analogie morphologique pour justifier la consonne muette.

À la fin des études, lorsque le lexique orthographique prend de l'expansion, la lexicalisation est tout aussi probable que l'application de la règle. Le plafonnement que l'on observe dans les scores de réussite orthographique au 3^e cycle appuie cette hypothèse et rejoint les observations de Bégin (2008)²⁷ qui constate une relation très étroite entre les scores de réussite orthographique et les scores de réussite morphologique chez les élèves de 6^e année ($r = 0.90$).

Par ailleurs, tous les sujets de notre échantillon, sans exception, ont représenté correctement au moins une consonne muette cible et ce, même lorsqu'il y avait des erreurs phonologiques dans le mot. Si l'on endosse l'hypothèse que les enfants traitent les mots selon les conventions phonographiques et orthographiques traditionnelles (Catach, 1980), le *-ait* à la fin du mot *lait* n'est pas un graphème, mais bien deux unités linguistiques, *ai* et *t*, la première, traitée par des processus phonologiques ($ai = /ɛ/$) et la seconde, par des processus morphologiques (ajout du *t* justifié par sa présence dans le dérivé *laitier*). Par ailleurs, le *-ait* de la flexion verbale de la 3^e personne du singulier à l'imparfait requiert l'un de ces deux traitements, orthographique ou morphologique, pour être produit correctement. Dans cette perspective, les mots qui se terminent par *-ait* doivent nécessairement être lexicalisés ou traités à l'aide de processus morphologiques pour représenter le *t* muet. Dans un contexte « usage-based », la fréquence d'exposition aux unités lexicales agira sur leur acquisition, peu

²⁷ Les scores de réussite dans Bégin, Saint-Laurent et Giasson (2010) sont tirés de la thèse de doctorat de Bégin (2008)

importe le niveau d'expertise général de l'enfant dans le traitement d'autres unités écrites. Par contre, dans un contexte développemental, il semble peu probable de lexicaliser ou activer des processus morphologiques avant d'atteindre un certain niveau d'expertise, ce qui n'est pas ce que nos données indiquent.

Cependant, il a été démontré que les graphèmes peuvent être traités différemment selon leur position dans un mot (Pacton et collaborateurs, 2001, 2002). On pourrait postuler que les séquences de lettres qui se terminent par une consonne peuvent être traitées différemment selon leur position dans le mot. Si l'on reprend l'exemple du mot *lait*, la suite de lettres *ait* pourrait être traitée par processus phonologiques, c'est-à-dire comme un trigramme contextuel qui fait toujours le son /ɛ/ à la fin d'une syllabe terminale. La fréquence et la consistance des CGP et CGP du mot *lait* appuient cette thèse. Le trigramme terminal *-ait* est aussi consistant à faire le son /ɛ/ que le graphème *-eau* l'est à faire le son /o/, par exemple, mais il est beaucoup plus fréquent que *-eau*. Le *-ait* est à la fin de 1647 mots dans Manulex contre 120 pour le *-eau*, une fréquence qui approche celle du *-er* (2286 mots). Par ailleurs, la CGP du *-ait* est plus consistante que celle du *-er*, puisqu'il n'a qu'une seule CGP, /ɛ/, contrairement au *-er* qui en a deux, /e/ comme *aimer* et /ɛr/ de *mer*. De plus, *-ait* est le graphème dominant du /ɛ/ final, beaucoup plus fréquent que ses rivaux (*-aient* = 949, *-ais* = 501 et *-et* = 220), contrairement au *-er* qui est fréquent, mais non dominant pour traduire le phonème /e/ (*-er* = 2286, *-é* = 2527, *-és* = 1345, *-ées* = 862). Effectivement, les caractéristiques distributionnelles du *-ait* justifieraient un processus de traitement similaire aux terminaisons *-eau* et *-er*. Or, si les enfants traitent le *-ait* ainsi, nos résultats ne contredisent pas les hypothèses des modèles développementaux puisque les jeunes enfants pourraient utiliser des processus de traitement alphabétique pour produire les lettres muettes qui sont enchâssées dans les syllabes terminales.

De plus, même si l'on accepte que les terminaisons puissent être traitées comme des graphèmes contextuels, cela ne réfute pas nécessairement les hypothèses des modèles interactifs. D'une part, un traitement alphabétique des terminaisons complexes pourrait rejoindre l'hypothèse d'autoapprentissage de Share (1995) où seuls les processus phonologiques traduisent les lettres en mots jusqu'à ce qu'ils soient rencontrés assez

fréquemment pour être lexicalisés par un processus orthographique. D'autre part, ce postulat rejoint aussi celui du modèle R-I de Perfetti où les unités phonologiques de taille croissante sont représentées dans un lexique fonctionnel selon la fréquence avec laquelle elles sont rencontrées. Cette représentation précoce d'unités phonographiques de taille croissante a aussi été observée par Lété, Peereman et Fayol (2008), les effets de fréquence au niveau phonémique sont apparus en 1^{re} année, tandis que les effets syllabiques ne sont apparus qu'en 2^e année lorsque les enfants avaient une plus grande exposition à l'écrit et commençaient à développer des connaissances orthographiques propres aux mots. Les rimes terminales de nos stimuli sont composées de 2, 3 et 4 lettres et ce n'est pas nécessairement les plus simples qui sont les mieux réussies, même chez les jeunes enfants et chez les anglophones. La comparaison ci-dessous en témoigne.

Dans la perspective *usage-based*, la consistance et la dominance sublexicale entrent aussi en jeu pour expliquer la production orthographique selon Lété, Peereman et Fayol (2008) et Alegria et Mousty, 1996, 1997. Nos observations sur la production de consonnes muettes abondent aussi dans ce sens. Comparons les mots *lait* et *gentil*. On observe que *lait* a une CPG finale (/ɛ/ = *-ait*) fréquente et dominante (voir les fréquences ci-dessus) et une CGP consistante (*-ait* fait toujours /ɛ/), mais que la CPG est non consistante (/ɛ/ = *-ait*, *-ais*, *-aient*, *-et*, *-es*...). Le mot *gentil* a une terminaison beaucoup moins fréquente (97 mots) que *-ait* (1647), ainsi qu'une CPG non consistante et non dominante à l'écrit (/i/ = 408 *-i*, 382 *-is*, 327 *-it* et 8 *-il*) et une CGP non consistante et non dominante (/il/ = 30, /i/ = 8). Parmi les 24 radicaux, *lait* occupe le 3^e rang de fréquence et le 1^{er} rang de la réussite morphologique. *Gentil* est au 4^e rang de fréquence, mais au 11^e rang de réussite. Plus que l'écart de fréquence lexicale, les différences de consistance et de dominance sublexicales semblent expliquer l'écart entre les rangs de réussite de ces deux mots. 92 % des enfants de 2^e année ont produit le mot *lait* avec un *-ait*, tandis que 62 % ont représenté le *-il* dans *gentil*.

On observe aussi que 47% de ces mêmes élèves produisent aussi le mot *laid* sans erreur, même si le *-aid* est une terminaison rare et non dominante. Elle apparaît dans 3 mots de Manulex et dans les deux autres mots qui sont extrêmement rares, *plaid* et *raid*, le *-d* est sonorisé. Si les enfants produisaient les terminaisons complexes contextuelles en ayant

simplement recours à des processus phonologiques, les pourcentages d'enfants qui produisent les deux mots avec le *-ait* dominant devraient être similaire. S'il y a réussite de terminaisons comme /ɛ/, une des terminaisons les plus fréquentes et inconsistantes du français, il doit nécessairement y avoir une lexicalisation quelconque, particulièrement s'il y a réussite d'homophones. De plus, aucun participant n'a produit l'homophone lexical *les*, de loin le plus fréquent des trois, appuyant davantage la théorie que les jeunes enfants ont les capacités cognitives pour lexicaliser les mots et cela ne cadre pas avec la modélisation à stades de Frith (1985) ou à phases d'Ehri (1995).

Par ailleurs, un effet des consonnes muettes elles-mêmes, tant au niveau graphémique que syllabique est observé, appuyant l'hypothèse que les lettres muettes sont traitées en unités orthographiques distinctes. Les lettres muettes les plus fréquemment substituées, *-s*, *-t* et *-e*, sont dominantes en position terminale muette. De plus, la substitution du *t* et du *s* est dominante à la fin des syllabes orales ouvertes (**permit*, **lais*), tandis que la substitution du *e* est dominante à la fin de syllabes orales fermées comme dans *lourd* et *regard* (**loure*, **regare*). Cela reflète la dominance orthographique de l'un et l'autre des types de syllabes terminales. Par ailleurs, certains types d'erreurs laissent entendre qu'une sensibilité existe, non seulement à la fréquence des consonnes muettes, mais aussi à leurs contraintes graphotactiques, comme l'ont observé Pacton et ses collaborateurs (2001, 2002). Par exemple, sur 25 erreurs de substitution avec la lettre muette *-x*, 24 d'entre elles reflètent une position légitime, c'est-à-dire après les graphèmes *ai*, *au* et *eau* (**respaix*, **cheaux* et **saux*).

Les données que nous avons recueillies sur la production de consonnes muettes appuient l'hypothèse que la fréquence agit sur la lexicalisation autonome des mots, comme le suggèrent les modèles *usage-based*. Si, toutefois, la fréquence et la redondance des unités linguistiques étaient les seuls à agir sur l'acquisition de l'orthographe, les terminaisons similaires consistantes, fréquentes et redondantes seraient acquises au même rythme, peu importe si la consonne avaient une fonction lexicale (un *lit*), une fonction grammaticale (il *lit*) ou aucune fonction du tout (un *répit*). Ce n'est toutefois pas ce qui a été observé. De la 2^e à la 6^e année du primaire, les enfants produisent mieux les consonnes terminales muettes à la fin des radicaux morphologiques qu'à la fin des mots opaques (Bégin, 2010; Sénéchal, 2000; Sénéchal, Basque et Leclaire, 2006). Cela laisse entendre que la morphologie joue un rôle

dans la production de consonnes muettes. De plus, cette observation va dans le même sens que nos observations sur la diminution plus grande des écarts-types dans le cas de la réussite morphologique par rapport à la réussite orthographique. L'écart qui se creuse à la fin du primaire ne serait pas prédit si l'ajout de la lettre muette était simplement le résultat d'un apprentissage lexical.

L'exemple où l'erreur de substitution de la consonne muette est jumelé à une erreur de substitution dans le radical, comme **gentie/*gentiesse* ou **permit/*permission*, est le plus probant pour suggérer une analogie morphologique. Dans le cas du **permit/*permission*, on semble faire face à une analogie complexe, produisant le *t muet* en l'associant à la graphie dominante du suffixe /sjõ/. La dominance de la CPG du *-tion* agit sur la production du dérivé qui, lui, agit sur la production du radical.

En résumé, on peut dire que les patrons de réussite, d'omission et de substitution indiquent, en plus d'un apprentissage lexical, une prise en compte par les enfants des patrons orthographiques du français, et, dans le cas des élèves de 4^e-5^e année, de la famille morphologique.

5.4 La féminisation et la production de consonnes muettes

Les études sur la production de marqueurs muets à la fin des radicaux morphologiques rapportent des effets de féminisation (Bégin, Saint-Laurent et Giasson, 2010; Pacton, 2001; Sénéchal, 2000; Sénéchal, Basque et Leclaire, 2006), mais nous avons repéré des failles dans le contrôle de la fréquence. Sénéchal (2000) a contrôlé la fréquence des mots à partir de Brulex qui tire ses fréquences à partir d'écrits pour adultes. Pacton (2001) a repris les mêmes stimuli et, selon les indices de fréquence de Manulex, les stimuli avec féminin dans les études de Sénéchal, Basque et Leclaire (2006) et Bégin, Saint-Laurent et Giasson (2010) étaient plus fréquents que ceux sans féminin. Effectivement, les radicaux qui pouvaient être féminisés dans l'étude de Bégin, Saint-Laurent et Giasson étaient cinq fois plus fréquents que les radicaux qui n'étaient que dérivables, avec la conséquence que les effets de féminisation observés se confondaient avec les effets de fréquence. Notre 3^e objectif était ainsi de tester si les effets de féminisation dans la production de la consonne muette persistaient lorsque la

fréquence entre les deux types de radicaux morphologiques, avec et sans flexion féminisante, était appariée.

En contrôlant la fréquence entre les types de mots et entre leurs familles morphologiques respectives, nous n'avons observé aucune différence statistique dans la production des consonnes muettes à la fin des radicaux avec ou sans féminin et ce, à tous les niveaux scolaires. La présence d'une flexion féminisante ne porte pas plus de poids sur la réussite du marqueur muet à la fin d'un radical morphologique que la présence d'un dérivé lexical. On peut donc postuler que les mots les plus fréquents ont généralement des féminins et que les effets de féminisation observés dans les études précédentes s'expliqueraient par la fréquence des radicaux qui peuvent être féminisés et non par la présence de la flexion féminisante elle-même.

5.5 L'effet de la langue maternelle

La situation linguistique particulière dans la grande région de Montréal a mené plusieurs chercheurs à se pencher sur les résultats scolaires des anglophones et allophones scolarisés en français. Bien que certains aient observé des avantages pour les anglophones dans l'acquisition du français écrit (Deacon et coll., 2007), la majorité des études recensées arrivent à la conclusion opposée : les anglophones en situation d'immersion n'atteignent pas le niveau de succès orthographique des enfants francophones (Genesee, 2004, 2007; Jared, 2006; Morris et Labelle, 2005) et ce, même dans la production de mots qui sollicitent des connaissances morphologiques (mots à morphologie), c'est-à-dire, qui s'expliquent par une règle ou une régularité orthographique (Morris et Labelle, 2005). Notre dernier objectif était donc d'observer si la production des consonnes muettes variait en fonction de la langue parlée à la maison.

Lorsque la fréquence des radicaux est contrôlée, la moyenne de réussite des francophones est supérieure à celle des anglophones à tous les niveaux scolaires, mais ce n'est significatif qu'à partir du 3^e cycle. Cela se rapproche des observations de Morris et Labelle (2005), où l'écart entre les natifs et non-natifs dans la production de mots à morphologie n'était significatif qu'à partir de la fin du 2^e cycle, à la fin du cycle, en 4^e année. Toutefois, si l'on contrôle la fréquence de la famille morphologique entière, l'écart est significatif dès le début du 2^e cycle.

La redondance familiale aurait moins d'effet sur les enfants anglophones en début de parcours lorsque le vocabulaire est limité et les dérivés sont absents du lexique oral.

Nos données révèlent aussi que c'est au 3^e cycle qu'apparaît une interaction significative entre la langue maternelle et la fréquence du radical. Les francophones réussissent mieux la consonne à la fin des mots rares que les anglophones, mais les anglophones sont tout aussi compétents que les francophones lorsque les mots sont fréquents. Encore une fois, on observe que la fréquence de la famille morphologique agit dès le 1^{er} cycle, puis à tous les niveaux, sauf la 4^e année. Romney et ses collaborateurs (1995) ont observé que la fréquence d'exposition au français écrit et le niveau de vocabulaire des enfants en immersion étaient faibles comparés aux enfants francophones. Contrairement à un francophone, l'anglophone intègre souvent son milieu d'immersion avec un déficit du lexique oral de base. Si les deux enfants évoluent à la même vitesse, le vocabulaire de l'anglophone risque d'être toujours plus faible en fin de parcours. On pourrait donc s'attendre à ce que les scores de réussite soient statistiquement différents tout au long du primaire, mais ce n'est pas ce que nous observons. Nos données révèlent qu'il n'y a pas de différence statistique entre les deux groupes en début de parcours, l'écart de réussite ne se creuse au profit des francophones qu'en 5^e année. Si le vocabulaire n'est pas une variable explicative, on peut postuler que le recours aux processus morphologiques ne l'est pas non plus. L'interaction précoce observée lorsque la famille morphologique est tenue en compte ne corrobore cependant pas cette hypothèse. Les anglophones, dont le vocabulaire est considérablement plus limité en début de parcours, retirent autant de profit de la fréquence familiale que les francophones. Que la fréquence de la famille agisse tôt sur la production de la consonne muette pourrait s'expliquer par l'utilisation précoce de processus morphologiques. De plus, la fréquence des mots les rend disponibles pour enrichir le lexique oral des anglophones, ce qui avantage l'acquisition orthographique, particulièrement en fin de parcours si la règle morphologique « ajoute une consonne muette lorsqu'elle est sonorisée dans un dérivé » est maîtrisée.

La proximité linguistique de l'anglais et du français pourrait aussi faciliter la représentation précise de consonnes muettes dérivables. Deacon et ses collaborateurs (2007) ont observé un transfert de la conscience morphologique entre le français et l'anglais. Bien que dans le cas présent, ce transfert se ferait en écriture, ils concluent que la conscience morphologique est

une connaissance linguistique qui est transférable d'un système orthographique à l'autre. Dans la mesure où l'on considère que la lecture et l'écriture utilisent les mêmes processus, on peut supposer que ce transfert puisse aussi se faire pour justifier la production de consonnes muettes. Si l'on prend les mots *respect* et *camp* comme exemple, leur orthographe est identique, la connaissance de la prononciation anglaise facilite la production des consonnes finales *ct* et *mp* puisqu'elles sont sonorisées en anglais. Inversement, la prononciation du mot *permit* en anglais pourrait induire en erreur un enfant anglophone qui produirait le mot **permit* en français. Effectivement, cette erreur de substitution est plus fréquente chez les enfants anglophones que les francophones.

Nous avons comparé les scores des mots *camp*, *fruit*, *respect*, *mont* et *parfum*, les 5 stimuli dans lesquels on prononce la consonne finale en anglais²⁸. À quelques exceptions près, on observe que les scores de réussite morphologique des francophones et des anglophones ne sont pas statistiquement différents, contrairement aux scores de l'ensemble des stimuli où les francophones réussissent mieux. Là où la réussite diffère est sur l'orthographe du mot *camp*, elle est mieux réussie par les anglophones en 2^e année, mais ce sont les francophones qui réussissent mieux en 3^e. Si l'on compare le mot *permis*, qui a un rival phonologique en anglais, on observe une différence significative au 3^e cycle. Seulement 7 % des francophones ont mis un *t* à *permis*, tandis que 27 % des anglophones ont fait la même substitution. Il semble donc que les anglophones sont sensibles à la proximité orthographique de leur langue maternelle dans la production de consonnes muettes.

Qu'elle soit morphologique ou interlinguale, l'analogie est une capacité cognitive à développer et on doit encourager son enseignement dès le début de l'enseignement formel. Cependant, il faut aussi mettre les apprenants en garde contre les erreurs que l'analogie peut engendrer. Le système phonomorphémique du français écrit est l'un des plus complexes des langues européennes et ses apprenants ont intérêt à avoir une bonne mémoire orthographique visuelle. Cependant, lorsque le traitement orthographique est inaccessible, soit parce que les mots sont nouveaux, soit parce que le stockage en mémoire fait défaut, le raisonnement analogique lié aux familles morphologiques devient indispensable pour produire les lettres muettes.

²⁸ Les noms *mount* et *perfume* s'écrivent un peu différemment en anglais, mais la consonne muette y est sonorisée.

CONCLUSION

Nos intérêts de recherche étaient de recenser à quels niveaux scolaires les enfants orthographient correctement les consonnes finales qui ne sont pas représentées à l'oral, mais qui peuvent être déduites en appliquant une règle morphologique du genre « la consonne muette se justifie par sa présence (sonorisation) dans un dérivé », et d'observer si certains facteurs propres aux mots ou aux sujets agissaient sur leur représentation à l'écrit. Les données recueillies révèlent que tous les sujets testés, incluant des enfants du 1er cycle du primaire et des anglophones en immersion, représentent des lettres muettes terminales à la fin des radicaux morphologiques. La moitié des enfants de 2^e année ont atteint des scores de réussite morphologique de 70 % ou plus. Les scores ont plafonné au 3^e cycle.

Nos observations s'ajoutent à une banque croissante d'études empiriques qui contredisent la prémisse des modèles développementaux qui place le développement du traitement morphologique à la fin du processus d'acquisition de l'écrit, c'est-à-dire après la maîtrise du traitement phonologique et lorsque le traitement orthographique des mots est de niveau avancé. Or, nous avons observé que plusieurs enfants produisent correctement des marqueurs morphologiques muets à la fin de mots où il y a aussi évidence de représentations alphabétiques et orthographiques non maîtrisées (ex : *gantil**, *cheaud**).

Nos résultats corroborent cependant la modélisation *usage-based* en exposant des effets de fréquence sur la production des consonnes muettes dérivables et s'ajoutent aux nombreuses observations empiriques qui appuient l'hypothèse que les représentations orthographiques des enfants dépendent de la fréquence avec laquelle les mots sont rencontrés. De plus, les observations liées à la fréquence des familles morphologiques suggèrent aussi un effet de redondance de la représentation du radical à l'intérieur des dérivés au moment où le vocabulaire des enfants prend de l'expansion.

Ce qui n'a pas été observé, toutefois, sont les effets de féminisation rapportés dans les études de Sénéchal (2000), Pacton (2001), Sénéchal, Basque et Leclaire (2006) et Bégin Saint-Laurent et Giasson (2010). En contrôlant la fréquence entre les mots avec et sans féminin, les effets de féminisation ne sont plus apparents, et ce, dès la 2^e année. Cette observation est pertinente pour la recherche, pour les didacticiens et pour les intervenants scolaires. Elle suggère que les effets de féminisation observés dans les recherches antérieures seraient attribuables à la fréquence de la famille morphologique qui comprend un féminin plutôt qu'à la présence du féminin lui-même. De plus, les résultats suggèrent que la dérivation est tout aussi accessible par les jeunes enfants pour justifier la présence des consonnes muettes que la féminisation. Il semble donc pertinent d'intégrer ces deux habiletés analogiques au curriculum en même temps. Or, la féminisation est une notion qui apparaît généralement dans les manuels du 1^{er} cycle, mais la dérivation, qui explique la présence de lettres muettes, n'y apparaît qu'à partir du 2^e cycle ou 3^e cycle, si elle apparaît du tout.

Par ailleurs, les scores de réussite observés dans la production de consonnes muettes dérivables corroborent les observations empiriques de chercheurs en langue seconde. Les non-francophones en immersion française n'ont pas réussi au même niveau que les francophones. Certaines observations qualitatives suggèrent qu'il serait possible d'exploiter les similarités et les différences entre les deux langues pour faciliter l'apprentissage de l'orthographe lexicale, mais la fréquence d'occurrence des mots a autant d'effet sur la réussite des enfants en immersion que sur la réussite des natifs du français, soulignant l'importance du niveau d'exposition au français écrit pour les enfants issus de famille où le français oral n'est pas pratiqué.

Dans les prochaines sections, nous ferons état de certaines limites méthodologiques de la présente recherche.

Les limites de l'étude et des pistes pour des recherches futures

Nos observations de la progression des apprentissages dans la production de consonnes muettes dérivables semblent corroborer la progression prescrite par le MELS, c'est-à-dire que les enfants de 2^e année sont en mesure de faire des analogies morphologiques pour justifier des lettres muettes et que cette habileté devrait être acquise en 6^e année. Il ne nous a

malheureusement pas été possible de contrôler le type d'enseignement reçu en ce qui a trait à l'analogie morphologique. Ainsi, si nous savons que les élèves sont en mesure de faire les analogies morphologiques, nous ne savons pas dans quelle mesure cela résulte d'un apprentissage lié à l'enseignement. Comme nous l'avons noté, la progression du MELS est établie en fonction d'un contexte d'enseignement-apprentissage. Or, sans ce contrôle, il est impossible de déterminer si l'enseignement reçu agit sur la progression observée.

Ainsi, nos observations donnent un aperçu de la progression dans la production des consonnes muettes dérivables et de l'effet de certaines variables explicatives, mais d'autres expérimentations, avec un contrôle plus rigoureux des nombreuses variables susceptibles d'agir sur les items et sur les sujets seront nécessaires pour dresser un portrait généralisable de la progression des apprentissages dans la production de marqueurs muets lexicaux. Les variables explicatives par items à considérer sont la fréquence des mots, la fréquence des familles morphologiques, la fréquence par type de radical (avec ou sans lettre muette, avec ou sans féminin, dans une syllabe terminale ouverte ou fermée, etc.), la fréquence, la consistance et la dominance des consonnes muettes et des rimes terminales. Les variables explicatives à contrôler pour les sujets sont le type d'approche didactique, le type d'approche pédagogique, le niveau socioéconomique, le niveau de vocabulaire, le niveau de lecture, le niveau d'orthographe, la conscience phonologique et la conscience morphologique, la langue maternelle, le type de programme d'enseignement de la langue seconde, le nombre d'heures d'enseignement en langue seconde et le nombre d'années à parler le français. Par ailleurs, une analyse de régression aiderait à évaluer le poids des variables explicatives.

De plus, dans le but d'évaluer si la fréquence de la famille morphologique agissait sur la production des consonnes muettes, notre choix de stimuli a été fait en contrôlant l'écart de fréquence entre le radical et le dérivé, et non en fonction de l'écart entre des mots rares et fréquents, entre les familles morphologiques rares et fréquentes ou entre les mots rares avec ou sans féminin. Une manipulation des écarts entre les mots rares et fréquents pourrait aider à élucider le niveau d'exposition qui est propice à la lexicalisation. Par ailleurs, les données recueillies ne permettent pas d'évaluer l'impact de la flexion féminisante sur les mots rares, une variable distributionnelle clé. Il serait aussi intéressant d'évaluer le phénomène de la redondance familiale en profondeur. Par ailleurs, les effets de fréquence observés ici

n'incluent pas de mots opaques, dont la lettre muette n'est pas dérivable. Il serait pertinent de reproduire les observations entre mots morphologiques et mots opaques en contrôlant la fréquence des familles morphologiques et la régularité des unités sublexicales.

Une observation intéressante tirée de l'analyse des erreurs est que la connaissance de l'anglais semble agir sur la production des lettres muettes, parfois de manière positive, parfois de manière négative. Pour les chercheurs en langue seconde, il serait judicieux d'élucider l'impact des similarités et divergences orthographiques entre le français et l'anglais sur la production des consonnes muettes, dérivables ou non. On pourrait aussi étudier si la fréquence et la présence d'une flexion féminisante agiraient différemment auprès de différentes populations allophones, anglophones et francophones.

Une limite appréciable de cette expérimentation est le niveau socioéconomique des sujets. Il sera aussi indispensable de reproduire nos résultats auprès de populations de niveaux socioéconomiques divergents pour établir les convergences et divergences avec notre échantillon, particulièrement pour établir si la féminisation et la fréquence agissent différemment. D'autres études seront indispensables pour corroborer nos observations qui suggèrent que les effets de féminisation rapportés dans les études précédentes s'expliquent par la fréquence des radicaux qui peuvent être féminisés et non par la présence de la flexion féminisante elle-même.

Finalement, les indices de fréquence lexicale de Manulex que nous avons utilisés constituaient le meilleur outil dont nous pouvions disposer. Mais il n'est pas parfait. Les indices ne sont qu'une estimation de l'occurrence des stimuli expérimentaux et ne représentent pas la fréquence réelle avec laquelle les sujets ont croisé les mots de la dictée. Par ailleurs, la normalisation européenne pourrait faire en sorte que les indices de fréquence ne reflètent pas la réalité lexicale des enfants québécois. Par exemple, les divergences démographiques rendent possible que le mot *nord* et son dérivé *nordique* soient moins fréquemment rencontrés dans des livres destinés aux enfants français que dans ceux destinés aux enfants québécois qui étudient l'univers social et la géographie du Grand Nord ou qui discutent de hockey 9 mois par année (les Nordiques de Québec, que l'on voudrait faire revenir dans la LNH).

Les recommandations didactiques et pédagogiques

Nos observations se joignent à celles de Sénéchal (2000) pour indiquer une sensibilité précoce à la production de consonnes muettes dérivables, et ce, même lorsque les représentations phonologiques ne sont pas maîtrisées. De plus, si, comme nous l'avons observé, un dérivé fréquent entraîne autant d'effet sur la représentation de la consonne muette du radical qu'une flexion féminisante fréquente, il s'ensuit que l'enseignement du concept de dérivation entraînerait autant d'effet sur la représentation de la consonne muette que l'enseignement du concept de féminisation, ce que l'on fait généralement dès la première année du primaire. Il semble donc pertinent pour les didacticiens d'intégrer la dérivation au curriculum en même temps que la féminisation.

Malheureusement, le concept de dérivation pour justifier la présence d'une consonne muette n'est pas toujours explicité dans les grammaires des 2^e et 3^e cycles du primaire. Trois grammaires de 2^e et 3^e cycles ont été recensées à cet effet : les *Petit guide du 2^e cycle* et *Petit guide du 3^e cycle* de Bouchard et Lévesque (2003, 2004) la *Grammaire de base* de Chartrand et Simard (2000) et les *Grammaire du deuxième cycle* et *Grammaire du troisième cycle* de Nadeau et Trudeau (2001, 2003). Bien que le MELS prescrive le développement de cette habileté dès la fin du premier cycle, seules les grammaires de Nadeau et Trudeau expliquent que les dérivés peuvent aider à trouver les consonnes muettes des bases. Pourtant, les notions de suffixation et de famille de mots sont abordées dans les deux autres ouvrages, tous deux couramment utilisés au Québec. Chartrand et Simard y expliquent que la connaissance des mots de bases, des préfixes et des suffixes aident à orthographier les dérivés, mais aucune mention n'est faite du phénomène inverse. Bouchard et Lévesque, qui consacrent une dizaine de pages sont aux familles de mots et aux concepts d'affixation, ne font aucune référence aux liens orthographiques qui peuvent être tissés entre les mots, ni à la règle morphologique qui permet de déduire la consonne muette. Il ne faut pas s'étonner si l'habileté de produire la bonne lettre muette par analogie dérivationnelle n'est pas systématiquement enseignée au Québec.

L'ajout de l'enseignement systématique de la dérivation dans les manuels scolaires de tous les niveaux scolaires est souhaitable à notre avis. Cette manipulation dérivationnelle est intrinsèque à la notion de classes grammaticales que l'on enseigne dès le premier cycle. L'enseignement qui explicite de la dérivation stimule la conscience des liens sémantiques et

morphosyntaxiques dans les familles de mots et ces connaissances épilinguistiques contribuent au développement de l'orthographe lexicale et du vocabulaire. Une approche didactique qui exploite ces liens nous semble très pertinente. Par exemple, dans les manuels du premier cycle, la dérivation est utilisée pour enseigner le lien nom-adjectif et le rôle de ces deux classes de mots dans le groupe du nom. Au 2^e cycle, elle est utilisée pour expliquer les fonctions syntaxiques dans les phrases en faisant les liens nom-verbe et au 3^e cycle, les liens adjectif-adverbe sont rajoutés pour contraster ces deux types de modificateurs.

Si le vocabulaire et la conscience morphologique sont les connaissances qui permettent de tisser des liens orthographiques entre les familles de mots, la conscience du procédé de dérivation par suffixation est le métier sur lequel se tisseront les représentations des radicaux morphologiques et de leur consonne muette. Les jeunes enfants semblent démontrer la capacité cognitive de faire des liens morphologiques donc leurs enseignants auraient intérêt à exploiter la dérivation pour favoriser l'apprentissage orthographique. L'enseignement précoce de la dérivation serait aussi bénéfique que celui de la flexion féminisante pour stimuler la conscience morphologique. Par exemple, insérer des exercices de dérivation en début de leçons qui visent l'apprentissage de l'orthographe lexicale, l'identification des classes de mots ou l'enrichissement du vocabulaire, permettrait aux enseignants de faire usage des radicaux et des dérivés identifiés par les élèves lors de la leçon. Ils pourraient aussi servir dans des dictées subséquentes et des productions écrites, favorisant le transfert à de vraies situations d'écriture. Pour les jeunes enfants, il est aussi important de confronter les conventions phonologiques aux conventions morphologiques muettes, toujours dans le but de justifier leurs choix orthographiques.

La représentation des consonnes muettes dérivables est tributaire de la fréquence d'exposition aux mots et aux familles de mots. Un élève qui croise un ensemble de mots d'une même famille de mots a de meilleures chances de produire correctement la consonne muette de son radical que s'il croise seulement le radical. L'utilisation de mots de familles rares est donc encouragée dans les manuels scolaires, lors de leçons et dans les dictées. Les didacticiens et les intervenants scolaires peuvent donc agir sur la fréquence d'exposition aux familles de mots en sachant qu'une exposition accrue facilitera la représentation des consonnes muettes à la fin des radicaux. Cette pratique est particulièrement encouragée pour les enseignants d'élèves anglophones et allophones, désavantagés par un vocabulaire limité.

Par ailleurs, l'anglais et le français ont des similarités qui peuvent être exploitées, incluant le partage de plusieurs mots de racines latines. Les élèves anglophones en immersion française n'atteignent pas les scores de réussite morphologique des natifs de la langue française. Si un enfant a la capacité cognitive de lire deux langues et exploiter des liens entre les mots de sa langue seconde, il a probablement la capacité de tirer avantage de sa langue première pour écrire le français. Les manuels scolaires qui visent des enfants québécois et canadiens en enseignement du français L2 pourraient utiliser des mots et des familles morphologiques qui exploitent les similarités entre l'anglais et le français. La connaissance de l'anglais peut toutefois nuire à la réussite orthographique des radicaux morphologiques donc on doit aussi conscientiser les enfants concernés qu'il y a des connaissances à exploiter, mais aussi des pièges à éviter. Par exemple, le mot *camp* est identique dans les deux langues, mais *nord* s'écrit *north* en anglais, ce qui pourrait porter un enfant à écrire *nort**. De plus, la langue maternelle des nouveaux arrivants ne partage pas toujours les propriétés linguistiques du français. Il faut être sensible à ces différences. L'enseignement explicite de la suffixation dérivationnelle comme moyen de créer des familles morphologiques et de justifier les consonnes muettes sera d'autant plus important pour un allophone en immersion si le code orthographique de sa langue maternelle est éloigné du français.

En résumé, l'analogie morphologique pour justifier la présence d'une lettre muette à la fin d'un mot lexical devrait apparaître systématiquement dans les manuels scolaires. De plus, la dérivation est une connaissance morphologique qui peut être enseignée dès le premier cycle du primaire, lorsque les mécanismes de stockage en mémoire ne sont pas pleinement fonctionnels. Les didacticiens ont intérêt à mieux l'intégrer au curriculum, les enseignants du primaire ont intérêt à enseigner explicitement l'analogie dérivationnelle et les orthopédagogues doivent favoriser ce genre d'analogie avec les enfants qui ont des difficultés de mémoire visuelle.

Par ailleurs, nos observations rejoignent plusieurs autres études scientifiques qui établissent un lien entre la fréquence des mots et le succès orthographique. Plus les enfants sont exposés aux formes écrites des mots, plus il est probable qu'ils représentent correctement les lettres muettes en situation d'écriture. Ainsi, « L'usage favorise l'expertise » est un principe que les intervenants scolaires doivent impérativement faire comprendre aux élèves et à leurs parents.

APPENDICE A

LES PHONÈMES ET GRAPHÈMES DU FRANÇAIS

Les voyelles

/a/	<u>a</u> baille, à, fem <u>me</u>
/ɑ/	â <u>ne</u> , chat, là
/ə/	che <u>va</u> l, ne <u>u</u> f, œ <u>u</u> f
/e/	é <u>t</u> é, souper, nez, ma <u>i</u> , de <u>ss</u> ert
/ɛ/	e <u>ll</u> e, ch <u>è</u> v <u>r</u> e, for <u>ê</u> t, ma <u>i</u> son, pe <u>i</u> ne, cr <u>a</u> yon, cha <u>l</u> et, de <u>s</u>
/ø/	je <u>u</u> di, vœ <u>u</u> x, je <u>û</u> ne
/œ/	fl <u>eu</u> r, sœ <u>u</u> r
/i/	igloo, gymn <u>a</u> se, î <u>l</u> e, yanke <u>e</u> , je <u>a</u> n
/ɔ/	po <u>m</u> me, dinosa <u>u</u> re, maximum
/o/	ro <u>s</u> e, ja <u>u</u> ne, be <u>a</u> u, c <u>ô</u> té
/u/	lo <u>u</u> p, ao <u>û</u> t
/y/	u <u>n</u> i, s <u>û</u> r, (j' <u>a</u> i) e <u>u</u>

Les consonnes

/b/	<u>b</u> ateau, abb <u>e</u>
/d/	<u>d</u> ent, add <u>i</u> tion
/f/	fl <u>eu</u> r, ph <u>a</u> re
/g/	garage, gu <u>i</u> tare, aggr <u>a</u> ver
/ʒ/	ja <u>r</u> d <u>i</u> n, gir <u>a</u> fe
/k/	<u>c</u> adeau, cir <u>q</u> ue, acc <u>o</u> rd, cue <u>i</u> llir, ch <u>o</u> rale, ki <u>w</u> i
/l/	l <u>a</u> p <u>i</u> n, ball <u>o</u> n
/m/	m <u>a</u> man, g <u>o</u> mm <u>e</u>
/n/	n <u>i</u> d, bon <u>n</u> e
/p/	p <u>a</u> pa, grip <u>p</u> e
/r/	ra <u>d</u> io, t <u>er</u> re, rh <u>i</u> noc <u>é</u> ros
/s/	s <u>a</u> le, cer <u>i</u> se, t <u>as</u> se, gar <u>ç</u> on, sc <u>i</u> ence, d <u>i</u> x
/t/	t <u>u</u> lip <u>e</u> , pat <u>t</u> e, th <u>é</u> â <u>t</u> re
/v/	v <u>a</u> lise, w <u>a</u> gon
/z/	z <u>è</u> bre, ci <u>s</u> eau, dix <u>i</u> ème
/ʃ/	ch <u>a</u> t, sh <u>é</u> r <u>i</u> f
/p/	sign <u>e</u> , camp <u>i</u> ng

Les voyelles nasales

/ã/	en <u>f</u> ant, jam <u>b</u> e, nov <u>e</u> m <u>b</u> re, fa <u>o</u> n
/ɛ̃/	lap <u>i</u> n, ch <u>i</u> en, ma <u>i</u> n, pe <u>i</u> nture, synt <u>a</u> x <u>e</u> , sim <u>p</u> le, c <u>y</u> mb <u>a</u> le, fa <u>i</u> m
/õ/	bon <u>b</u> on, trom <u>b</u> one
/œ̃/	l <u>u</u> nd <u>i</u> , parf <u>u</u> m, à je <u>u</u> n

Les semi-consonnes

/w/	o <u>i</u> seau, Ott <u>a</u> wa
/ɥ/	fr <u>u</u> it, t <u>u</u> yau
/j/	y <u>o</u> yo, trav <u>a</u> il, bil <u>l</u> et, champ <u>i</u> on

APPENDICE B

STIMULI PAR CONDITION AVEC INDICES DE FRÉQUENCE CUMULÉE
(MANULEX)

Condition	Stimulus	Fréquence du radical		Fréquence du dérivé		Fréquence de la famille complète		Fréquence totale des dérivés	
		2 ^e	3 ^e - 5 ^e	2 ^e	3 ^e - 5 ^e	2 ^e	3 ^e - 5 ^e	2 ^e	3 ^e - 5 ^e
Condition A <i>Radical > Dérivé</i> <i>Avec féminin</i>	bavard/bavarder	53.66	56.71	42.70	50.34	54.35	57.10	46.06	46.41
	chaud/chaudement	67.55	68.65	38.39	38.39	68.33	69.59	60.50	62.49
	droit/droitier	63.34	66.00	0	42.58	66.55	69.13	63.74	66.24
	épais/épaisseur	58.28	60.63	33.42	51.25	61.70	63.83	59.07	60.99
	gentil/gentillesse	64.91	65.76	44.76	50.24	67.31	68.05	63.59	64.18
	laid/laideur	57.86	59.08	33.62	37.63	58.72	60.06	51.26	53.08
	lourd/lourdement	61.87	64.01	38.20	51.61	63.29	65.70	57.74	60.79
	renard/renardeau	68.16	69.03	42.53	44.44	68.21	69.07	48.33	48.85
Condition C <i>Radical > Dérivé</i> <i>Sans féminin</i>	camp/camper	57.91	60.61	52.38	53.57	65.42	67.18	64.57	66.10
	début/débutant	65.47	67.85	0	46.01	65.47	67.85	0	0
	fruit/fruitier	66.65	67.90	0	39.08	66.65	67.90	0	0
	lait/laitier	68.03	68.93	43.42	44.12	68.03	68.93	0	0
	nord/nordique	60.40	62.97	0.00	35.68	60.40	62.97	0	0
	parfum/parfumer	58.04	60.43	33.22	45.91	58.42	60.95	47.63	51.43
	tard/tarder	66.86	68.50	38.57	50.90	66.88	68.54	43.93	48.86
	vent/venteux	68.72	66.26	0	36.81	70.75	72.15	45.55	52.96
Condition D <i>Dérivé > Radical</i> <i>Sans féminin</i>	achat/acheter	54.62	56.86	64.08	65.50	64.54	66.06	64.08	65.50
	arrêt/arrêter	58.99	61.50	63.24	64.69	69.78	71.41	69.40	70.94
	éclat/éclater	55.70	58.39	47.72	54.97	61.72	63.96	60.47	62.55
	mont/montagne	48.31	53.67	67.07	68.37	71.12	72.63	71.10	72.57
	permis/permission	47.60	47.87	54.66	56.05	55.44	56.67	54.66	56.05
	regard/regarder	56.61	63.08	64.94	66.78	74.32	75.69	74.25	75.45
	respect/respecter	0	0	52.04	55.32	52.26	55.42	52.26	55.42
	saut/sauter	55.27	58.35	62.67	64.26	69.00	70.41	68.81	70.14

APPENDICE C

PHRASES DE CONTEXTE POUR LA DICTÉE

DICTÉE A		DICTÉE B		
MOT	CONTEXTE	MOT	CONTEXTE	
1	gentil	Mon frère est très gentil.	sauter	J'aime sauter à la corde.
2	vent	Le vent souffle fort ce matin.	lourd	Mon sac d'école est lourd.
3	respecter	Nous devons respecter l'environnement.	parfumer	Ma sœur aime se parfumer.
4	arrêter	Il ne pouvait pas arrêter de rire.	début	On met une majuscule au début des phrases.
5	laid	Ce monstre est très laid.	épais	Le carton est plus épais que le papier.
6	parfum	Ton parfum sent bon.	fruitier	Le pommier est un arbre fruitier.
7	montagne	Cette montagne est très haute.	venteux	Le haut de la montagne est toujours venteux.
8	regard	Il a un drôle de regard.	gentillesse	Elle a eu la gentillesse de s'excuser.
9	camper	Nous allons camper au Lac Massawippi	tard	Nous devons nous coucher, car il est tard.
10	saut	J'ai fait un saut lorsque la cloche a sonné.	renardeau	Le renardeau court après la queue de sa mère.
11	bavard	Mon petit cousin est bavard.	chaud	Il fait très chaud aujourd'hui.
12	éclat	Nous avons entendu un éclat de rire.	bavarder	Il ne faut pas bavarder en classe.
13	laitier	Mon grand-père était laitier.	respect	Le respect est une très belle qualité.
14	débutant	Mon frère est un skieur débutant.	camp	Cet été, je vais dans un camp de vacances.
15	tarder	Tu ne dois pas tarder à t'habiller le matin.	regarder	J'aime regarder les étoiles.
16	nordique	Le Canada est un pays nordique.	lait	Elle boit du lait au chocolat.
17	droit	Je me suis fait mal au genou droit.	nord	Notre chalet est au nord de Montréal.
18	lourdement	Il marche lourdement avec ses grosses bottes.	acheter	Mon père veut acheter une automobile.
19	renard	Le renard a une très longue queue.	arrêt	À un feu rouge, nous devons faire un arrêt.
20	achat	Nous avons fait l'achat d'une maison.	permis	Ma sœur a son permis de conduire.
21	permission	J'ai la permission d'aller chez toi.	éclater	Il a fait éclater des feux d'artifice.
22	fruit	L'orange est mon fruit préféré.	mont	J'ai fait du ski au Mont Ste-Marie
23	épaisseur	Ce livre a 5 centimètres d'épaisseur.	laideur	Cet insecte est d'une laideur effroyable.
24	chaudement	L'hiver, nous nous habillons chaudement.	droitier	Ce gant de baseball est fait pour un droitier.

APPENDICE D

CONSIGNES POUR PASSATION DE DICTÉES

Consignes aux enseignants et aux examinateurs :

Installer les écrans d'isolation ou séparer les bureaux pour que les élèves ne puissent pas voir les réponses de leurs voisins. Les élèves doivent avoir un crayon et une gomme à effacer.

En haut de sa feuille de réponse, l'élève inscrit la DATE, son PRÉNOM et NOM de famille, le NOM DE L'ÉCOLE et son GROUPE CLASSE.

L'enseignant(e) lit le mot à orthographier, ainsi que la phrase permettant aux élèves de reconnaître le mot en contexte.

Il est très important de NE PAS DONNER D'INDICES quant à l'orthographe ou la famille du mot. Au besoin, l'enseignant(e) peut répéter le mot à orthographier, ainsi que la phrase de contexte. La révision des mots à la fin de la dictée est aussi permise.

Le but de cette recherche est d'évaluer la reconnaissance des lettres muettes à la fin des mots (la conscience morphologique) donc tout indice à cet effet est interdit. L'enseignant devra noter toute situation ou anomalie où quelqu'un (un élève ou l'enseignant) aurait donné un indice lors de la passation de la dictée. Ceci permettra aux chercheurs de prendre en compte toute situation susceptible de fausser les résultats.

L'orthographe des mots NE PEUT ÊTRE DÉVOILÉE aux élèves qu'après la passation de la 2^e dictée, car les mots de la 1^{re} dictée peuvent donner des indices quant à l'orthographe des mots de la seconde.

Après la passation de la 2^e dictée, veuillez S.V.P. remettre les feuilles de réponses, ainsi que les feuilles de consentement des parents à la personne-ressource de votre école.

APPENDICE E

FORMULAIRES DE CONSENTEMENT DES PARENTS - FRANÇAIS

Cher(s) parent(s),

Notre groupe de recherche étudie l'acquisition du français en milieu scolaire chez les enfants parlant le français ou une autre langue à la maison. Nous aimerions que votre(vos) enfant(s) participe(nt) à cette étude et sollicitons votre consentement.

EXPLICATIONS DE L'ÉTUDE

- Deux dictées distinctes de 24 mots chacune (2 jours différents) sont données aux enfants de l'école primaire (de la 1^{re} à la 6^e année)
- Chaque dictée dure au maximum 15 minutes
- La participation des enfants sera sollicitée préalablement
- Le nom des enfants et toute information permettant de les identifier sont traités de manière confidentielle
- Il n'y a aucun risque ou danger pour l'enfant
- Il n'y a pas d'obligation de participer. (Les enfants qui ne participent pas ne sont pas pénalisés.)
- Les résultats sont confidentiels
- Assistante à la recherche responsable de la collecte de données : Lise L'Heureux Équipe de professeurs de l'Université du Québec à Montréal : Marie Labelle, Lucie Godard, Lori Morris, Daphnée Simard
- Recherche subventionnée par *le Conseil de recherche en sciences humaines du Canada*, approuvée par la commission scolaire (s'il y a lieu), par l'école, et par le comité d'éthique de l'université

Lise L'Heureux
 Assistante à la recherche et étudiante à la maîtrise
 Département de linguistique et de didactique des langues
 l'heureux.lise@courrier.uqam.ca
 514-768-9567

Pour informations et plaintes :

Dr Marie Labelle
 Professeure et Chercheure principale, Département de linguistique et de didactique des langues
 Université du Québec à Montréal
 CP 8888, succursale Centre-ville,
 Montréal, QC, H3C 3P8
 labelle.marie@uqam.ca
 514-987-3000 poste 7036

Comité institutionnel d'éthique de la recherche avec des êtres humains de l'Université du Québec à Montréal (secrétariat: 987-3000 poste 7753).

APPENDICE F

FORMULAIRES DE CONSENTEMENT DES PARENTS - ANGLAIS

Dear Parents,

Our research group is conducting a study into the acquisition of French language in a school setting for children who speak French or another language at home and are soliciting your child's (children's) participation.

DETAILS OF THE STUDY

- Two distinct spelling tests of 24 words each will be administered (on 2 different occasions) to students in elementary school (1st through 6th grade)
- Each test will last a maximum of 15 minutes
- The children's participation will be solicited prior to the tests
- The children's name and all information allowing to identify them will be kept strictly confidential
- There is no risk or danger to the children
- There is no obligation to participate. (Children who do not want to participate will not be penalised.)
- The results are confidential
- The Research Assistant responsible for data collection : Lise L'Heureux
- The Université du Québec à Montréal team of professors are Marie Labelle, Lucie Godard, Lori Morris, and Daphnée Simard.
- This research has the financial support of the *Social Sciences and Humanities Research Council of Canada*, and has been approved by the school board (where applicable), by the school, and by the ethics committee of the university.

Lise L'Heureux
 Research Assistant and Masters Student
 Department of Linguistics and Language Didactics
 Université du Québec à Montréal
 l'heureux.lise@courrier.uqam.ca
 514-768-9567

For information or complaints:

Dr. Marie Labelle
 Professor and Principal Researcher, Department of Linguistics and Language Didactics
 Université du Québec à Montréal
 Postal Box 8888, Centre-Ville Station
 Montreal, QC, H3C 3P8
 labelle.marie@uqam.ca
 514-987-3000 ext. 7036

Institutional Committee on Research Ethics with Human Beings of the Université du Québec à Montréal (secretariat: 987-3000 extension 7753)

DECLARATION OF INFORMED CONSENT

I hereby give my informed consent for my child/children to participate in a study aimed at assessing French language acquisition at school. I have READ and UNDERSTOOD the above information and guarantees provided by Lise L'Heureux, Research Assistant responsible for data gathering and Marie Labelle, principal researcher on this project.

I ACCEPT: YES..... NO.....

My child/children's name/s:

My child/children's age/s:

Given Name.....years.....months

Given Name.....years.....months

Given Name.....years.....months

My child/children's school:

The language generally spoken at home:

Mother's first language:

Father's first language:

Signature:

DATE:

Name of parent or guardian: (Please print)

Please have your child return this signed form to his or her classroom teacher **BEFORE May 30th, 2008**. Please keep the information page (1st page) for your records.

APPENDICE G

FORMULAIRES DE CONSENTEMENT DES ÉCOLES - FRANÇAIS

OBJET : Demande d'approbation d'une recherche auprès d'élèves du primaire

Cette recherche s'inscrit dans le cadre du Groupe de recherche sur l'apprentissage du français du Département de linguistique et didactique des langues de l'Université du Québec à Montréal.

Afin de mieux comprendre le développement des connaissances en orthographe du français, particulièrement celui du codage morphologique chez les élèves du primaire, nous avons développé un test de dictée de 48 mots. Voici les caractéristiques de la recherche :

- Le test est administré en deux dictées distinctes de 24 mots chacune (2 jours différents).
- Chaque dictée dure au maximum 15 minutes.
- Le nom des enfants et toute information permettant de les identifier sont traités de manière confidentielle.
- Il n'y a aucun risque pour l'enfant.
- La permission des parents sera obtenue.
- La participation des enfants sera sollicitée en classe. Il n'y a pas de pénalité pour ceux qui ne désirent pas participer.
- Les résultats individuels sont confidentiels

Nous désirons administrer le test aux élèves de 1^{ère} à 6^e année afin de pouvoir étudier la progression dans les connaissances entre les trois cycles du primaire.

Nous désirons tester tous les élèves, francophones et non francophones, afin de pouvoir comparer leurs connaissances en fonction de la langue maternelle. Nos recherches nous montrent que souvent les non-francophones réussissent aussi bien que les francophones, mais en utilisant des stratégies différentes. C'est un des aspects que nous désirons étudier.

La recherche a reçu l'approbation du *Conseil de recherche en sciences humaines du Canada*, qui la subventionne. Elle a également reçu un certificat d'éthique de l'UQAM.

Pour questions et informations :

Lise L'Heureux
 Assistante à la recherche et étudiante à la maîtrise,
 Département de linguistique et de didactique des langues
 lheureux.lise@courrier.uqam.ca
 514-768-9567

Dr Marie Labelle
 Professeure, Département de linguistique et de didactique des langues
 labelle.marie@uqam.ca
 514-987-3000 poste 7036

.../p.2

DÉCLARATION DE CONSENTEMENT

Je, soussigné(e), accepte que l'école (les écoles) _____ de la
Commission scolaire _____ participe(nt) à une étude sur l'acquisition
de la langue française en milieu scolaire.

OUI NON

Nom du signataire :

Poste :

Établissement scolaire:

Adresse courriel :

Numéro de téléphone :

Adresse de l'établissement :

.....

.....

Signature:..... DATE:.....

Veillez retourner ce formulaire de consentement dûment signé à l'adresse ci-dessous
AVANT LE 25 avril 2008. Veillez conserver la feuille d'informations pour vos dossiers.

APPENDICE H

FORMULAIRES DE CONSENTEMENT DES ÉCOLES - ANGLAIS

REFERENCE: Request for consent to do research with Elementary School students

This study is part of our Research Group on French Language Acquisition at the Department of Linguistics and Language Didactics of the Université du Québec à Montréal.

In order to better our knowledge of student's orthographic development we have developed a 48 word spelling test. Here are the characteristics of the research:

- Two distinct spelling tests of 24 words each will be administered (on 2 different occasions) to students in 1st through 6th grade.
- Each test will last a maximum of 15 minutes
- Parental approval will be solicited prior to testing
- The children's name and all information allowing to identify them will be kept strictly confidential
- There is no risk or danger to the children
- Children's participation will be solicited in class. Children who do not want to participate will not be penalised.
- Individual results are confidential

We wish to administer the test to students from the 1st to the 6th grades in order to study the progression of acquisition of this particular knowledge.

We wish to test all students, francophone and non-francophone, in order to compare results based on their first language. Our research has demonstrated that non-francophone students have similar results to their francophone classmates, but by using different strategies. This is one of the aspects we are studying.

This research has the financial support of *the Social Sciences and Humanities Research Council of Canada*, and has been approved by UQAM's ethics committee.

For questions or information please contact:

Lise L'Heureux
 Research Assistant and Masters Student
 Department of Linguistics and Language Didactics
 l'heureux.lise@courrier.uqam.ca
 514-768-9567

Dr. Marie Labelle
 Principal Researcher, Department of Linguistics and Language Didactics
 labelle.marie@uqam.ca
 (514) 987-3000 ext. 7036

DECLARATION OF CONSENT

I, the undersigned, accept that the following school(s) _____, part of the _____ School Board to participate in a study on the acquisition of the French language in a school environment.

YES

NO

Name of undersigned:

Title:

School or School Board:

Email address:

Telephone number:

School (Board) address:

.....

.....

Signature:..... Date:.....

Please return this signed consent form to the address below **BEFORE May 28th, 2008**. Please keep the information form (1st page) for your records.

BIBLIOGRAPHIE

- Alegria, J. et Mousty, P. (1996). The Development of Spelling Procedures in French-Speaking, Normal and Reading-Disabled Children: Effects of Frequency and Lexicality. *Journal of experimental child psychology*, 63(2), 312–338.
- Alégria, J. et Mousty, P. H. (1997). Processus lexicaux impliqués dans l'orthographe d'enfants francophones présentant des troubles de la lecture. *Rieben, L.; Fayol, M.; Perfetti, CA Des orthographes et leur acquisition*, 167–180.
- Bégin, C., Saint-Laurent, L. et Giasson, J. (2010). Le traitement morphologique dans l'écriture des mots chez les élèves de 6e année du primaire. *L'Année psychologique*, 110(02), 275–297.
- Bonin, P., Barry, C., Méot, A. et Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory and Language*, 50(4), 456–476.
- Bonin, P., Chalard, M., Méot, A. et Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, 93(1), 89–114.
- Bonin, P., Fayol, M. et Chalard, M. (2001). Age of acquisition and word frequency in written picture naming. *The Quarterly Journal of Experimental Psychology*, 54(2), 469–489.
- Bouchard, N. N. et Lévesque, C. (2003). *Le petit guide du 2e cycle. Grammaire française au primaire*. Éditions HRW.
- Bouchard, N. N. et Lévesque, C. (2004). *Le petit guide du 3e cycle. Grammaire française au primaire*. Éditions HRW.
- Brown, G. D. et Loosemore, R. P. W. (1995). Computational approaches to normal and impaired spelling. *Handbook of Spelling: Theory, Process and Intervention* (G. D. A. Brown et N. C. Ellis, Eds.). Chichester, England: John Wiley.

- Carlisle, J. F. (1988). Knowledge of Derivational Morphology and Spelling Ability in Fourth, Sixth, and Eighth Graders. *Applied Psycholinguistics*, 9(3), 247-66.
- Carlisle, J. F. (1995). Morphological awareness and early reading achievement. *Morphological aspects of language processing*. (pp. 189-209). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Casalis, S. (2003). Le codage de l'information morphologique dans l'écriture de mots chez les apprentis scripteurs. *Le Langage et l'Homme*, 38(2), 96-110.
- Catach, N. (1980). *L'orthographe Française Traite Theorique Et Pratique Avec Des Travaux D'application Et Leurs Corriges*. Nathan Université, information, formation. Paris: F. Nathan.
- Chartrand, S.-G. et Simard, C. (2000). *Grammaire de base, 2e et 3e cycle du primaire*. ERPI.
- Colé, P., Royer, C., Leuwers, C. et Casalis, S. (2004). Les connaissances morphologiques dérivationnelles et l'apprentissage de la lecture chez l'apprenti-lecteur français du CP au CE2. *L'année psychologique*, 104(4), 701-750.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R. et Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204-256.
- Coltheart, M., Curtis, B., Atkins, P. et Haller, M. (1993). Models of Reading Aloud: Dual-Route and Parallel-Distributed-Processing Approaches. *Psychological Review*, 100(4), 589-608.
- Content, A., Mousty, P. et Radeau, M. (1990). Brulex: Une base de données lexicales informatisée pour le français écrit et parlé. *L'année Psychologique*, 90, 551-556.
- Content, A., Mousty, P. H. et Radeau, M. (2011). Brulex. *Laboratoire Cognition Langage développement*. Université libre de Bruxelles, . Retrieved from <http://lclld.ulb.ac.be/outils/brulex>
- Daigle, D., Mathieu, I. et Montésinos-Gelet, I. (2008). Évaluation du traitement des correspondances phonèmes/phonogrammes chez des lecteurs/scripteurs

adultes en début d'apprentissage du français langue seconde. *The Canadian Modern Language Review / La revue canadienne des langues vivantes*, 64(4), 637-661.

- Deacon, S. H, Wade-Woolley, L. et Kirby, J. (2007). Crossover: The role of morphological awareness in French immersion children's reading. *Developmental psychology*, 43(3), 732-746.
- Defior, S., Alegría, J., Titos, R. et Martos, F. (2008). Using morphology when spelling in a shallow orthographic system: The case of Spanish. *Cognitive Development*, 23(1), 204-215.
- Desrochers, A. (2006). OMNILEX: une base de données sur le lexique du français contemporain. *Cahiers linguistiques d'Ottawa*, 34, 25-34.
- Druide, informatique inc. (2008). *Antidote*. Montréal: Druide informatique.
- Duncan, L. G. et Seymour, P. H. K. (2000). Socio-economic differences in foundation-level literacy. *British Journal of Psychology*, 91(2), 145-166.
- Duncan, L. G., Seymour, P. H. K. et Hill, S. (1997). How important are rhyme and analogy in beginning reading? *Cognition*, 63(2), 171-208.
- Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading*, 18(2), 116-125.
- Ehri, L. C. (1997). Learning to read and learning to spell are one and the same, almost. *Learning to spell: Research, theory, and practice across languages*. (pp. 237-269). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Ehri, L. C. (2005). Learning to Read Words: Theory, Findings, and Issues. *Scientific Studies of Reading*, 9(2), 167-188.
- Ehri, L., Nunes, S., Willows, D., Schuster, B. V., Yaghoub-Zadeh, Z. et Shanahan, T. (2001). Phonemic Awareness Instruction Helps Children Learn to Read: Evidence From the National Reading Panel's Meta-Analysis. *Reading Research Quarterly*, 36, 250-287.
- Ellis, N. C. (2002). Frequency Effects in Language Processing. *Studies in Second Language Acquisition*, 24(02), 143-188.

- Fayol, M. et Jaffré, J.-P. (2008). *Orthographier*. Paris: Presses universitaires de France.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In K.E Patterson, J.C. Marshall et M. Coltheart (Eds.), *Surface dyslexia* (pp. 301-330). London: Erlbaum. Glushko. *Surface Dyslexia* (Erlbaum., pp. 301-330). Londres: K.E. Patterson, J.C. Marshall et M. Coltheart.
- Garlock, V. M., Walley, A. C. et Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language*, 45(3), 468–492.
- Genesee, F. (2004). What do we Know About Bilingual Education for Majority-Language Students?
- Genesee, F. (2007). Literacy Outcomes in French Immersion. *Encyclopedia of Language and Literacy Development*. Retrieved from <http://literacyencyclopedia.ca/index.php?fa=items.show&topicId=27>
- Godard, L. et Labelle, M. (1995). Utilisation de L'ÉVIP avec une population québécoise. *Fréquences*, 7(2), 18-20.
- Harm, M. W. et Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *PSYCHOLOGICAL REVIEW-NEW YORK-*, 106, 491–528.
- Harm, M. W. et Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *PSYCHOLOGICAL REVIEW-NEW YORK-*, 111, 662–720.
- Imbs, P. (1971). Trésor de la langue Française: Dictionnaire des fréquences [Resource of the French language: Dictionary of frequencies]. Paris: Klincksieck.
- Jaffré, J.-P. (1997). From Writing to Orthography: The functions and Limits of the Notion of System. *Learning to spell: Research, theory, and practice across languages*. (Lawrence Erlbaum Associates., pp. 3-20). Londres: Charles.

- Jared, D. (2006). Becoming biliterate: A longitudinal investigation of reading growth in children in French immersion. *Conference of the Canadian Language and Literacy Research Network, Charlottetown, PEI.*
- Jones, V. W. (1967). *The Graphoneme Concept: A Systematized Approach to Initial Reading Instruction.* Northwest Regional Educational Laboratory, Portland, Oregon.
- Joshi, R. M. et Aaron, P. G. (2006). *Handbook of orthography and literacy.* Routledge.
- Lambert, E. et Chesnet, D. (2001). Novlex: une base de données lexicales pour les élèves de primaire. *L'année psychologique, 101(2), 277-288.*
- Largy, P., Fayol, M. et Cousin, M.-P. (2004). Acquérir le pluriel des noms. (C. Brissaud & C. Totereau, Eds.) *Lidil, Acquisition et enseignement de la morphographie, (30), 39-54.*
- Lété, B. (2005). MANULEX: une base de données du lexique écrit adressé aux élèves. *Didactique du lexique: contextes, démarches, supports, Bruxelles, De Boeck.*
- Lété, B., Peerean, R. et Fayol, M. (2008). Consistency and word-frequency effects on spelling among first- to fifth-grade French children: A regression-based study. *Journal of Memory and Language, 58(4), 952-977.*
- Lété, B., Sprenger-Charolles, L. et Colé, P. (2004). MANULEX: a grade-level lexical database from French elementary school readers. *Behavior Research Methods, Instruments et Computers: A Journal of the Psychonomic Society, Inc, 36(1), 156-166.*
- Leybaert, J. et Alegria, J. (1995). Spelling development in deaf and hearing children: Evidence for use of morpho-phonological regularities in French. *Reading and Writing, 7(1), 89-109.*
- Leybaert, J. et Content, A. (1995). Reading and spelling acquisition in two different teaching methods: A test of the independence hypothesis. *Reading and Writing, 7(1), 65-88.*

- Martinet, C., Valdois, S. et Fayol, M. (2004). Lexical orthographic knowledge develops from the beginning of literacy acquisition. *Cognition*, 91(2), B11-B22.
- McClelland, J. L. et Rumelhart, D. E. (1981). An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings. *Psychological Review*, 88(5), 375–407.
- McClelland, J. L., Rumelhart, D. E. et others. (1986). Parallel distributed processing.
- Monaghan, P. et Ellis, A. W. (2010). Modeling reading development: Cumulative, incremental learning in a computational model of word naming. *Journal of Memory and Language*.
- Morris, L. et Labelle, M. (2005). L'influence de l'âge et du contexte d'apprentissage sur l'utilisation des stratégies épilinguistiques en français langue seconde. Presented at the Colloque international de didactique cognitive, Toulouse, France.
- Nadeau, M. et Trudeau, S. (2001). *Grammaire du deuxième cycle. Pour apprendre, s'exercer et consulter*. Graficor.
- Nadeau, M. et Trudeau, S. (2003). *Grammaire du troisième cycle. Pour apprendre s'exercer et consulter*. Graficor.
- New, B. et Pallier, C. (2005). *Manuel de Lexique 3*. Document non publié, Laboratoire de psychologie expérimentale, Université René-Descartes, Boulogne-Billancourt, France.
- Pacton, S. (2001). Children's linguistic awareness and spelling: The case of derivational morphology. Presented at the Congress of the european society for cognitive psychology, Écosse.
- Pacton, S. (2003). Morphologie et acquisition de l'orthographe : état des recherches actuelles. *Rééducation orthophonique*, 41(213), 27-55.
- Pacton, S. et Casalis, S. (2006). L'utilisation d'informations morphologiques en production écrite rend-elle le cauchemar (d?) des lettres muettes moins cauchemardesque? *Rééducation orthophonique*, 44(225), 129–144.

- Pacton, S., Fayol, M. et Perruchet, P. (2002). The acquisition of untaught orthographic regularities in French. *Precursors of functional literacy*, 121–137.
- Pacton, S., Fayol, M. et Perruchet, P. (2005). Children's implicit learning of graphotactic and morphological regularities. *Child Development*, 76(2), 324–339.
- Pacton, S., Perruchet, P., Fayol, M. et Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General*, 130(3), 401–426.
- Pacton, Sébastien et Deacon, S. H. (2008). The timing and mechanisms of children's use of morphological information in spelling: A review of evidence from English and French. *Cognitive Development*, 23(3), 339–359.
- Peereman, R. et Content, A. (1999). LEXOP: A lexical database providing orthography-phonology statistics for French monosyllabic words. *Behavior Research Methods*, 31(2), 376–379.
- Peereman, R., Lété, B. et Sprenger-Charolles, L. (2007). Manulex-infra: Distributional characteristics of grapheme-phoneme mappings, and infralexical and lexical units in child-directed written material. *Behavior Research Methods*, 39(3), 579–589.
- Perfetti, C. A. (1991). Representations and awareness in the acquisition of reading competence. *Learning to read: Basic research and its implications* (L. Rieben et C.A Perfetti (Eds)., pp. 33–44). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Perfetti, C. A., Rieben, L. E. et Fayol, M. E. (1997). *Learning to spell: Research, theory, and practice across languages*. Lawrence Erlbaum Associates Publishers.
- Pinker, S. et Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Connections and symbols*, 73–193.

- Plaut, D. (1999). A connectionist approach to word reading and acquired dyslexia: extension to sequential processing. *Cognitive science*, 23(4), 543-568.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S. et Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103(1), 56-115.
- Polguère, A. (2003). *Lexicologie Et Sémantique Lexicale: Notions Fondamentales*. Paramètres. Montréal: Presses de l'Université de Montréal.
- Romney, J. C., Romney, D. M. et Menzies, H. M. (1995). Reading for Pleasure in French: A Study of the Reading Habits and Interests of French Immersion Children. *Canadian modern language review*, 51(3), 474-511.
- Rumelhart, D. E. et McClelland, J. L. (1986). *On learning the past tenses of English verbs, Parallel distributed processing: explorations in the microstructure of cognition, vol. 2: psychological and biological models*. MIT Press, Cambridge, MA.
- Rumelhart, D. E. et McClelland, J. L. (1987). Learning the past tenses of English verbs: Implicit rules or parallel distributed processing. *Mechanisms of language acquisition*, 195-248.
- Seidenberg, M. S. et McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4), 523.
- Sénéchal, M. (2000). Morphological effects in children's spelling of French words. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 54(2), 76-86.
- Sénéchal, M. et Kearnan, K. (2007). The role of morphology in reading and spelling. *Advances in Child Development and Behavior*, 35, 297-325.
- Sénéchal, M., Basque, M. T. et Leclaire, T. (2006). Morphological knowledge as revealed in children's spelling accuracy and reports of spelling strategies. *Journal of Experimental Child Psychology*, 95(4), 231-254.
- Seymour, P. H. K. (1997). Foundations of Orthographic Development. *Learning to spell: Research, theory, and practice across languages*. (Lawrence Erlbaum

- Associates., pp. 319-337). Londres: Charles A. Perfetti, Laurence Riebe et Michel Fayol.
- Seymour, P. H. K. et Evans, H. M. (1999). Foundation-Level Dyslexia. *Journal of Learning Disabilities*, 32(5), 394 -405.
- Seymour, P. H. K., Duncan, L. G. et Bolik, F. M. (1999). Rhymes and phonemes in the common unit task: replications and implications for beginning reading. *Journal of Research in Reading*, 22(2), 113-130.
- Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, 55(2), 151–218.
- Snowling. (1994). Towards a Model of Spelling Acquisition: The Development of Some Component Skills. *Handbook of Spelling: Theory, Process and Intervention* (John Wiley & Sons Ltd., pp. 111-128). Chichester, England: G. D. A. Brown & N. C. Ellis.
- Snowling, M. et Frith, U. (1981). The role of sound, shape and orthographic cues in early reading. *British Journal of Psychology*, 72(1), 83–87.
- Sprenger-Charolles, L., Siegel, L. S. et Bonnet, P. (1998). Reading and spelling acquisition in French: The role of phonological mediation and orthographic factors. *Journal of Experimental Child Psychology*, 68(2), 134.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360–407.
- Totereau, C., Thevenin, M.-G. et Fayol, M. (1997). The Development of the Understanding of Number Morphology in Written French. *Learning to spell: Research, theory, and practice across languages*. (Lawrence Erlbaum Associates., pp. 97-114). Londres: Charles A. Perfetti, Laurence Rieben et Michel Fayol.
- Treiman, R. (1993). *Beginning to spell: A study of first-grade children*. Oxford University Press, USA.

- Treiman, R. et Cassar, M. (1996). Effects of Morphology on Children's Spelling of Final Consonant Clusters. *Journal of Experimental Child Psychology*, 63(1), 141-170. doi:10.1006/jecp.1996.0045
- Veronis, J. (1988). From sound to spelling in French: Simulation on a computer. *Cahiers de psychologie cognitive*, 8(4), 315-334.
- Véronis, J. (1986). Étude quantitative sur le système graphique et phonographique du français. *Cahiers de psychologie cognitive*, 6(5), 501-531.