

UNIVERSITÉ DU QUÉBEC EN OUTAOUAIS

FACTEURS LINGUISTIQUES FAVORISANT L'ACQUISITION DES PHONÈMES  
MULTIGRAPHÉMIQUES CHEZ LES ÉLÈVES QUÉBÉCOIS BONS ET FAIBLES  
ORTHOGRAPHEURS DE LA TROISIÈME ANNÉE DU PRIMAIRE :  
LE CAS DU PHONÈME MULTIGRAPHÉMIQUE /Ë/

THÈSE

PRÉSENTÉE

COMME EXIGENCE PARTIELLE

DU DOCTORAT EN ÉDUCATION

PAR

MÉLANIE DUTEMPLE

DÉCEMBRE 2023

UNIVERSITÉ DU QUÉBEC À MONTRÉAL  
Service des bibliothèques

Avertissement

La diffusion de cette thèse 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.04-2020). 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

Merci à l'Équipe de recherche en littératie et inclusion (ERLI) et au Fonds de recherche société et culture du Québec (FRQSC) pour leur précieux soutien financier.

Merci aux professeurs François Vincent, Rachel Berthiaume et Anila Fejzo qui ont accepté de faire partie de mon jury. Je suis très fière et reconnaissante d'avoir eu la chance d'être évaluée par des chercheurs aussi compétents et passionnés que vous. Vos commentaires constructifs m'ont amenée à me dépasser et à présenter une thèse dont je suis particulièrement fière.

Merci du fond du cœur à mon directeur de recherche, André C. Moreau. Le rôle que tu tenais dans ma vie professionnelle depuis le début de ma maîtrise en orthopédagogie en 2015 était déjà immensément précieux, mais le rôle que tu tiens depuis 2021 est exceptionnel. C'est ton grand humanisme qui m'a permis de continuer. Ta patience, ton expertise, tes dizaines et dizaines de relecture ; il n'y a pas de mot suffisant pour t'exprimer ma gratitude !

Toute ma reconnaissance aussi à Brigitte Stanké, qui a codirigé mon doctorat pendant plusieurs années et qui m'a donné accès aux données de son groupe de recherche avec grande générosité. Sa maîtrise exceptionnelle de la linguistique m'a aidée à devenir spécialiste de mon domaine. Merci aussi à tous les professeurs qui ont pris son relai avec grande générosité à divers moments, notamment à Phaedra Royle et à Stefano Rezzonico. Enfin, merci à Myriam Fontaine, qui a codirigé mon doctorat avec attention, ouverture et confiance pendant mon énorme sprint final pendant lequel des centaines de corrections ainsi que 75 pages se sont ajoutées. Ta vision pédagogique et tes connaissances poussées en orthopédagogie m'ont permis de resserrer ma thèse et de reprendre confiance en mon projet.

Merci aux professeurs qui m'ont soutenue dans le traitement de mes données. Merci d'abord à Guillaume Loignon, statisticien, professeur et pédagogue dévoué. Merci de m'avoir accompagnée pas à pas dans ce difficile traitement statistique. Merci aussi au professeur Stefano Rezzonico, qui m'a soutenue au début du traitement de mes statistiques et qui m'a accueillie en stage de recherche doctorale. Merci d'avoir eu la brillante idée de traiter la multigraphémie à partir du phonème /ẽ/. Enfin, merci au professeur Daniel Daigle qui m'a accueillie en stage de recherche doctorale avec

grande générosité et qui a toujours continué d'être là pour me conseiller et pour m'encourager par la suite.

Merci à ma collègue de doctorat et amie, Josianne Parent. Ta présence à mes côtés m'a incitée à m'accrocher à l'idée de faire un doctorat et à me lancer pieds joints dans cette folle aventure !

Merci à tous les professionnels du centre d'apprentissage Molière et Copernic. Votre passion, votre dévouement et votre amour des jeunes me portent tous les jours. Un merci spécial à Cathie Labonté, à Johanne Lécuyer et à Sophie De Bellefeuille qui ont pris la relève de Molière et Copernic du jour au lendemain à la fin de l'été 2021 et qui ont tenu le fort pendant les derniers mois si importants accordés aux corrections de ma thèse. Merci aussi à Alexandra Leduc qui est la force tranquille de la belle aventure de Molière et Copernic depuis ses débuts.

Merci à tous les élèves que j'ai eu l'honneur d'accompagner en orthopédagogie et qui prouvent chaque jour que l'intelligence ne se mesure pas par la facilité à orthographier.

Merci à Denise Richer-Lauzon qui m'a enseigné l'amour de la langue française en sixième année du primaire et qui ne cesse, depuis, de m'encourager à me dépasser.

Merci à June Dubé, psychologue, qui m'a soutenue émotionnellement avec grande générosité pendant la tempête. Merci aussi à Vanessa Deschênes qui m'a aidée à conserver une bonne santé mentale et à prendre les bonnes décisions dans cet énorme projet qu'est le doctorat.

Merci du fond du cœur à ma famille. Naitre et grandir en étant aimée, ça donne des ailes ! Profonde gratitude à mes parents, Georges et Jocelyne Dutemple, et à mon frère Benoit Dutemple pour cet amour.

Merci à mes garçons, Mathieu et Olivier Brassard, à mes fils, Félix, Louis et Thomas Brassard, et à mon sixième, David Ladouceur, pour leur amour inconditionnel. Votre amour, votre présence et votre encouragement ont été la base de tout. Vous voilà maintenant tous dans le grand train de la vie adulte, mais sachez que je suis et que je serai toujours là, pas loin, dans le dernier wagon.

Merci à mes amis qui m'ont tellement aidée à transformer l'indicible, l'extrême douleur, en force afin que je me relève et que je continue à faire ce que j'aime et à aimer ce que je fais. Sylvie Béland, Sophie Haché, Marco Guèvremont, Julie Leblanc, René Huneault, Maxime Catellier, Sylvie Levasseur, Jacques Billette, Marie Lauzon, Isabelle Gaudet-Labine, Marilou Tanguay, Marie-Elaine Mineau, Raluca Lupu-Onet, Elaine Rochefort, Marie-Josée Pinizzotto, Monique Leclerc, Nancy Desrosiers, Randy Helmer, Danielle Dufresne, Gino Caron, Linda Poirier, Jessica Pepper et Louise Jaladeau, vous avez ma reconnaissance éternelle.

Merci du fond du cœur à toi, mon amour, Simon Fleury, qui es arrivé comme un cadeau dans ma vie alors que j'étais aux difficiles derniers milles de mon doctorat. Tu me répètes constamment combien je suis forte. Sache que ta patience, ta compréhension, ton amour et tes encouragements quotidiens ont été le moteur de ma persévérance. Merci de me prouver chaque jour que l'amour est plus fort que tout et que l'on peut aimer d'un amour tout neuf, d'un amour qui n'existe pas encore. Merci d'avoir fait de moi l'amour du reste de ta vie. Je te promets que « notre route sera grande, lumineuse et fantastique » (Ingrid St-Pierre). Enfin, merci d'avoir amené tes deux magnifiques fleurs, Marikim et Mia, dans ma vie.

Finalement, merci à André Brassard, que j'ai tant aimé et à qui je dédie cette thèse de doctorat. Le professeur Martin Lépine écrivait dernièrement cette citation si vraie à la suite du décès de son beau-frère, Karl Tremblay : « Mais ce que je comprends chaque jour, chaque minute un peu mieux, c'est qu'il nous faut tout bonnement accomplir avec intensité et ferveur ce qui nous tient à cœur, travailler fort et sans relâche à ce qu'on croit pouvoir réaliser de mieux, éviter de perdre notre temps à ne rien faire de bon pour personne, et mettre à profit généreusement, vigoureusement, farouchement, le moment qui passe. » (Laurent-Michel Vacher, *Dialogues en ruine*). Faut-il être confronté au départ précipité de quelqu'un qu'on aime profondément pour saisir réellement ces mots ? Je ne sais pas. Ce que je sais, c'est que je comprends et saisis maintenant la force de chacun d'entre eux. À la mémoire d'André, ne les oublions pas. Saisissons le jour. Saisissons la vie. Parce que la vie, c'est maintenant.

## DÉDICACE

« J'ai toujours ton cœur avec moi  
Je le garde dans mon cœur. »  
E.E. Cummings

À mon mari, mon amour, André Brassard, qui m'a soutenue pendant six ans à temps complet dans ma réorientation de carrière, dans ma maîtrise en orthopédagogie et dans mon doctorat en éducation. Jamais, le temps qui passe, depuis son décès, n'éteindra notre amour. Jamais.

## TABLE DES MATIÈRES

REMERCIEMENTS .....	ii
DÉDICACE .....	v
LISTE DES FIGURES .....	x
LISTE DES TABLEAUX .....	xi
LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES .....	xiii
LISTE DES SYMBOLES ET DES UNITÉS .....	xiv
RÉSUMÉ .....	xv
ABSTRACT .....	xvi
INTRODUCTION.....	1
CHAPITRE 1 PROBLÉMATIQUE.....	4
1.1 Importance de la littératie pour la réussite scolaire et sociale de tous les élèves .....	4
1.2 Réussite en orthographe .....	5
1.3 Facteurs jouant un rôle dans l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques .....	8
1.3.1 Facteurs individuels.....	8
1.3.2 Facteurs environnementaux .....	11
1.3.2.1 Enseignement explicite de l'orthographe lexicale .....	11
1.3.2.2 Prise en compte des phonèmes multigraphémiques dans des outils linguistiques qui régissent l'acquisition de l'orthographe lexicale.....	12
1.3.3 Facteurs linguistiques.....	16
1.4 Question de recherche .....	19
1.5 Contributions de la recherche .....	19
CHAPITRE 2 CADRE THÉORIQUE .....	22
2.1 Système orthographique du français .....	22
2.1.1 Définition de l'orthographe .....	22
2.1.2 Multigraphémie .....	23
2.1.3 Plurisystème du français .....	25
2.1.3.1 Dimension phonographémique .....	26
2.1.3.2 Dimension morphographémique.....	32
2.1.3.3 Dimension logographique.....	33
2.1.3.4 Dimension visuogrammique.....	34

2.2	Modèles théoriques de l'acquisition du langage écrit .....	35
2.2.1	Apprentissage implicite.....	36
2.2.2	Modèle de Gombert.....	38
2.2.3	Modèle à double fondation de Seymour .....	40
2.2.3.1	Traitement alphabétique .....	42
2.2.3.2	Traitement logographique.....	43
2.2.3.3	Traitement orthographique .....	44
2.2.3.4	Traitement morphographémique.....	45
2.2.4	Modèle de reconnaissance et de production des mots écrits de Daigle et al. ....	46
2.2.4.1	Procédures analogiques .....	47
2.2.4.2	Procédures d'assemblage et de décomposition.....	48
2.2.4.3	Procédures visuelles et visuo-orthographiques.....	49
2.2.4.4	Procédures morphosyntaxiques et syntaxiques.....	51
2.3	Facteurs individuels intervenant dans l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques .....	52
2.3.1	Mémoire .....	52
2.3.2	Capacités langagières et conscience phonologique.....	54
2.3.3	Empan visuoattentionnel.....	55
2.4	Facteurs linguistiques intervenant dans l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques.....	55
2.4.1	Facteurs lexicaux .....	56
2.4.1.1	Fréquence du mot.....	56
2.4.1.2	Longueur du mot.....	59
2.4.2	Facteurs sous-lexicaux intervenant dans l'acquisition de l'orthographe lexicale .....	59
2.4.2.1	Consistance phonographémique .....	60
2.4.2.2	Régularités graphotactiques.....	65
2.4.2.3	Régularités morphologiques .....	69
2.5	Facteurs environnementaux liés aux bases de données et aux échelles d'acquisition en orthographe lexicale .....	71
2.5.1	Bases de données lexicales .....	71
2.5.2	Échelles d'acquisition en orthographe lexicale .....	72
2.6	Synthèse et objectifs de recherche .....	74
2.6.1	Synthèse des recherches portant sur l'acquisition de l'orthographe lexicale .....	74
2.6.2	Objectifs et hypothèses de recherche .....	77
	CHAPITRE 3 MÉTHODOLOGIE .....	80
3.1	Posture épistémologique.....	81
3.2	Devis et variables de l'étude.....	83
3.3	Participants .....	85
3.4	Déroulement de l'étude .....	86
3.4.1	Phase d'approbation du projet de recherche.....	87
3.4.2	Phase de codage et d'entrée des données pour le développement de la base de données .....	87

3.4.3	Phase d'analyse en quatre temps.....	87
3.4.4	Phase de rédaction des chapitres sur les résultats et productions de la thèse .....	89
3.5	Instruments de collecte de données.....	89
3.6	Plan d'analyse .....	94
3.7	Enjeux déontologiques .....	99
<b>CHAPITRE 4 RÉSULTATS .....</b>		<b>101</b>
4.1	Traitement statistique .....	101
4.2	Statistiques descriptives .....	105
4.3	Analyses multiniveaux .....	108
4.4	Comparaisons multiples entre les bons et les faibles orthographieurs.....	109
4.1	.....	111
4.2	Synthèse des résultats.....	113
<b>CHAPITRE 5 DISCUSSION.....</b>		<b>118</b>
5.1	Discussion liée à l'hypothèse 1 et à l'hypothèse 3.....	122
5.1.1	Fréquence du mot.....	123
5.1.2	Consistance du mot .....	123
5.1.3	Longueur du mot (nombre peu élevé de lettres).....	123
5.1.4	Synthèse .....	124
5.2	Discussion liée à l'hypothèse 2 et à l'hypothèse 4.....	124
5.2.1	Types de graphèmes et consistance phonographémique .....	125
5.2.2	Régularité graphotactique liée au positionnement du graphème.....	127
5.2.3	Régularité graphotactique liée au contexte d'utilisation du graphème .....	128
5.2.4	Régularité graphotactique liée au voisinage du graphème .....	129
5.2.5	Régularités morphologiques dérivationnelles .....	129
5.2.6	Synthèse .....	130
5.3	Hiérarchisation des facteurs lexicaux et sous-lexicaux dans l'acquisition du phonème multigraphémique /ɛ̃/ en fonction des deux catégories de participants : les bons orthographieurs et les faibles orthographieurs .....	131
5.4	Implications des nouvelles connaissances pour la pratique .....	135
5.5	Limites.....	144
5.6	Futures études .....	146
<b>CONCLUSION.....</b>		<b>148</b>
ANNEXE A Extrait du prétest visant à analyser l'effet du nombre de difficultés par mots sur la réussite de la 1 <sup>re</sup> à la 6 <sup>e</sup> année du primaire.....		149
ANNEXE B Extrait du traitement personnel des statistiques sur Excel .....		150

ANNEXE C Table des matières de la première analyse statistique .....	153
ANNEXE D Arbres produits lors de la deuxième analyse statistique .....	155
ANNEXE E Caractéristiques distributionnelles des correspondances phonographologiques et des unités lexicales dans le lexique de l'école primaire québécoise - Extrait d'ÉQOL-Infra (Peereman et Stanké, 2022) .....	160
ANNEXE F Approbation éthique .....	162
ANNEXE G Mots qui contiennent la prononciation [ɛ̃] dans la liste orthographique à l'usage des enseignantes et des enseignants (ministère de l'Éducation et de l'Enseignement supérieur, 2014) au premier cycle du primaire .....	168
ANNEXE H Mots qui contiennent la prononciation [ɛ̃] dans la liste orthographique à l'usage des enseignantes et des enseignants (ministère de l'Éducation et de l'Enseignement supérieur, 2014) au deuxième cycle du primaire .....	171
ANNEXE I Mots qui contiennent la prononciation [ɛ̃] dans la liste orthographique à l'usage des enseignantes et des enseignants (ministère de l'Éducation et de l'Enseignement supérieur, 2014) au troisième cycle du primaire .....	178
ANNEXE J Mots qui contiennent la prononciation [ɛ̃] dans l'échelle québécoise d'acquisition de l'orthographe lexicale (ÉQOL) (Stanké et al., 2019) .....	188
RÉFÉRENCES.....	194

## LISTE DES FIGURES

Figure 2.1 L'opacité du système orthographique en schéma (reproduit et traduit de Seymour et al., 2003, p. 146).....	25
Figure 2.2 Le système initial (Demont et Gombert, 2004, p. 252) .....	39
Figure 2.3 L'apprentissage implicite de la lecture (Demont et Gombert, 2004, p. 253).....	40
Figure 2.4 Représentation du modèle de Seymour (traduit de Seymour et al., 2003) .....	42
Figure 2.5 Modèle de reconnaissance et de production des mots écrits axé sur les microtâches (Daigle et al., 2018, p. 248; Daigle et Berthiaume, 2021, p. 139).....	46
Figure 3.1 Les étapes de la démarche scientifique (Lehrer et al., 2021) .....	82
Figure 4.1 Contrastes en orthographe entre les deux profils d'élèves selon les différents graphèmes .....	111
Figure 4.2 Contrastes entre les profils de difficulté en orthographe selon la longueur du mot (nombre peu élevé de lettres) .....	113
Figure 5.1 Ordre d'acquisition des graphèmes associés au phonème / <i>ẽ</i> / .....	127
Figure 5.2 Ordre des facteurs favorisant l'acquisition des graphèmes associés au phonème / <i>ẽ</i> / pour les bons orthographes .....	132
Figure 5.3 Ordre des facteurs favorisant l'acquisition des graphèmes associés au phonème / <i>ẽ</i> / pour les faibles orthographes .....	134

## LISTE DES TABLEAUX

Tableau 1.1 Taux de réussite pour les critères liés à l'orthographe aux épreuves ministérielles de français écrit (adapté du ministère de l'Éducation, du Loisir et du Sport, 2012).....	6
Tableau 2.1 Consistances phonographémiques en français à partir des graphèmes dont le pourcentage d'utilisation est de plus de 90 % (adapté de Catach, 2016).....	27
Tableau 2.2 Inconsistances phonographémiques du français (Adapté de Catach, 2016).....	299
Tableau 2.3 Pourcentage d'utilisation des graphèmes dans les cas principaux d'inconsistances phonographémiques (adapté de Catach, 2016).....	30
Tableau 2.4 Progression pédagogique de la lecture (inspiré de Sprenger-Charolles, 2017) .....	588
Tableau 2.5 Classification des 45 graphèmes du français en fonction de leur degré de complexité (reproduit de Pérez, 2014a, p. 9) .....	64
Tableau 2.6 Synthèse des facteurs ayant un effet sur l'acquisition de l'orthographe lexicale à partir des données probantes des recherches .....	755
Tableau 3.1 Sommaire des variables de la présente recherche .....	84
Tableau 3.2 Nombre d'élèves par sexe et par profil.....	86
Tableau 3.3 Nombre d'élèves par indicateur de milieu socioéconomique.....	86
Tableau 3.4 Liste des mots du sous-corpus par facteurs sous-lexicaux .....	91
Tableau 3.5 Consistance des graphèmes selon la position qu'ils occupent dans les mots – Adapté d'ÉQOL-Infra (Peereman et Stanké, 2022) .....	93
Tableau 3.6 Nombre de mots comportant chacun des graphèmes et nombre d'observations selon le graphème et selon le profil des élèves .....	94
Tableau 3.7 Nombre de mots comportant chacun des graphèmes et nombre d'observations selon la position du graphème dans le mot et selon le profil d'élèves.....	94
Tableau 3.8 Critères infirmant ou confirmant les hypothèses, en fonction des outils .....	95
Tableau 4.1 Comparaison des modèles multiniveaux .....	103
Tableau 4.2 Modèles multivariés testés dans le cadre de la présente étude .....	104
Tableau 4.3 Statistiques descriptives de la réussite de la tâche selon le type de graphèmes et selon le profil d'élèves .....	105

Tableau 4.4 Statistiques descriptives de la réussite orthographique du graphème selon son positionnement dans le mot et selon le profil d'élèves .....	106
Tableau 4.5 Graphème choisi par les élèves en cas d'erreur des graphèmes cibles « in » et « ein » selon le profil d'élèves .....	107
Tableau 4.6 Graphème choisi par les élèves en cas d'erreur dans les graphèmes cibles « ain », « en » et « im » selon le profil d'élèves .....	107
Tableau 4.7 Tests de Wald – modèle final « minimaliste ».....	108
Tableau 4.8 Contrastes entre les profils d'élèves en orthographe, par graphème.....	110
Tableau 4.9 Contrastes entre les profils de difficulté en orthographe selon le positionnement du graphème .....	112
Tableau 4.10 Infirmité ou confirmation des hypothèses.....	114
Tableau 5.1 Synthèse des facteurs ayant un effet sur l'acquisition des phonèmes multigraphémiques à partir de la présente recherche doctorale .....	119
Tableau 5.2 Consistance des graphèmes associés au phonème multigraphémique /ɛ̃/ selon la position qu'ils occupent dans les mots – Adapté d'ÉQOL-Infra-Lemme (Peereman et Stanké, 2022) .....	139
Tableau 5.3 Mots contenant le phonème multigraphémique /ɛ̃/ à acquérir du premier au troisième cycle du primaire à partir des résultats de la présente recherche ainsi que des mots d'ÉQOL et des données d'ÉQOL-Infra .....	141

## **LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES**

ANOVA : analyse de variance

AIC : critère d'information d' Akaike

API : alphabet phonétique international

BELEC : Batterie d'évaluation du langage écrit

BIC : critère d'information bayésien

EHDAA : élèves handicapés ou en difficulté d'adaptation ou d'apprentissage

ÉOLE : Échelle d'acquisition en orthographe lexicale

ÉQOL : Échelle québécoise d'acquisition de l'orthographe lexicale

ERLI : Équipe de recherche en littératie et inclusion

FRQSC : Fonds de recherche société et culture du Québec

MEES : ministère de l'Éducation et de l'Enseignement supérieur

MELS : ministère de l'Éducation, du Loisir et du Sport

OCDE : Organisation de coopération et de développement économiques

OQLF : Office québécois de la langue française

PDA : Progression des apprentissages

SPSS : Statistical Package for the Social Sciences

UNESCO : Organisation des Nations Unies pour l'éducation

## LISTE DES SYMBOLES ET DES UNITÉS

Alphabet phonétique international (API)<sup>1</sup>

---

Voyelles	Consonnes	Semi-voyelles
/a/ <b>partie</b>	/b/ <b>bien</b>	/j/ <b>bien, soleil, yoyo</b>
/ɑ/ <b>base, mâ</b> t	/k/ <b>comme</b>	/w/ <b>oui</b>
/ə/ <b>le</b>	/d/ <b>donner</b>	/ɥ/ <b>nuit</b>
/e/ <b>mésange</b>	/f/ <b>feu</b>	
/ɛ/ <b>crème, maison</b>	/g/ <b>gomme</b>	
/i/ <b>demi</b>	/ʒ/ <b>je</b>	
/o/ <b>auto</b>	/l/ <b>loup</b>	
/ɔ/ <b>pomme</b>	/m/ <b>mouche</b>	
/y/ <b>jus</b>	/n/ <b>noix</b>	
/u/ <b>hibou</b>	/p/ <b>papa</b>	
/ø/ <b>jeu</b>	/R/ <b>rue</b>	
/œ/ <b>fleur</b>	/s/ <b>sale</b>	
/ã/ <b>maman</b>	/t/ <b>table</b>	
/ɛ̃/ <b>lapin</b>	/v/ <b>vous</b>	
/ɔ̃/ <b>bouchon</b>	/z/ <b>prison</b>	
/œ̃/ <b>lundi</b>	/ʃ/ <b>chaud</b>	
	/ɲ/ <b>gagner</b>	

---

<sup>1</sup> Les symboles retenus proviennent de l'OQLF (Office québécois de la langue française) (s. d.)

## RÉSUMÉ

### Résumé

L'opacité de la langue française a un impact négatif sur l'acquisition de l'orthographe lexicale (Seymour et al., 2003). En effet, la consistance entre les correspondances phonèmes-graphèmes est seulement de 71 % (Fayol et Jaffré, 2008). Le degré d'incertitude face à l'orthographe qui en résulte constitue l'une des principales sources de difficultés en français écrit (Fayol et Jaffré, 2008; Pacton, Foulon et al., 2005). Les facteurs qui favorisent l'acquisition du phonème multigraphémique /ɛ̃/ chez les élèves québécois de troisième année du primaire, faibles et bons orthographes sont l'objet de cette étude. Lors d'une tâche de dictée de mots, 509 participants provenant de 31 écoles du Québec devaient orthographier correctement 73 mots comportant le phonème /ɛ̃/. Nous nous sommes intéressés à la réussite orthographique des graphèmes suivants : « in », « im », « ain », « ein » et « en ». L'effet de facteurs linguistiques sur l'acquisition des différents graphèmes a été étudié selon les groupes d'élèves, faibles et bons orthographes. Les facteurs lexicaux et sous-lexicaux suivants ont été analysés : consistance, fréquence et longueur du mot; types de graphèmes; consistance phonographémique; régularités graphotactiques (position, contexte d'utilisation et voisinage du graphème) et régularités morphologiques dérivationnelles.

À la lumière de recherches théoriques et empiriques sur le sujet, la présente recherche permet de constater que la fréquence et la consistance des mots ne jouent pas un rôle significatif, alors que la consistance et le positionnement des graphèmes jouent un rôle très significatif dans l'acquisition des phonèmes multigraphémiques. De plus, les résultats suggèrent que les élèves faibles orthographes réussissent beaucoup mieux à orthographier les graphèmes associés aux phonèmes multigraphémiques lorsqu'ils sont dérivables et que les mots sont courts. Toutefois, contrairement aux élèves bons orthographes, ils peinent à intégrer les régularités graphotactiques contextuelles.

Cette thèse a des retombées importantes aux plans scientifique et social. Des propositions concrètes ressortent des résultats de recherche, tant pour les bons orthographes que pour les faibles orthographes, favorisant ainsi l'acquisition des phonèmes multigraphémiques et la réussite scolaire de tous les élèves.

Mots-clés :

Acquisition de l'orthographe lexicale, multigraphémie, régularités graphotactiques, morphologie dérivationnelle

## ABSTRACT

### Abstract

The opacity of the French language has a negative impact on the acquisition of lexical spelling (Seymour et al., 2003). Indeed, the consistency between phoneme-grapheme correspondences is only 71% (Fayol et Jaffré, 2008). The resulting degree of spelling uncertainty is one of the main sources of difficulty in written French (Fayol et Jaffré, 2008; Pacton et al., 2005). The factors that promote the acquisition of the multigraphic phoneme /ɛ̃/ in Quebec third-grade elementary school students, both weak and good spellers, were the focus of this study. In a word dictation task, 509 participants from 31 Quebec schools were asked to correctly spell 73 words containing the phoneme /ɛ̃/. We were interested in the spelling success of the following graphemes: 'in', 'im', 'ain', 'ein' and 'en'.

The effect of linguistic factors on the acquisition of the different graphemes was studied according to the groups of students, weak and good spellers. The following lexical and sub-lexical factors were analyzed: consistency, frequency and length of the word; types of graphemes; phonographic consistency; graphotactic regularities (position, context of use and neighbourhood of the grapheme) and derivational morphological regularities.

Considering theories and research findings on the subject, this research shows that word frequency and word consistency do not play a significant role, but grapheme consistency and grapheme position play a very significant role in the acquisition of multigraphic phonemes. Furthermore, the results suggest that weak spellers are much more successful in spelling the graphemes associated to multigraphic phonemes when they are derivable and when the words are short. However, unlike good spellers, they struggle to integrate contextual graphotactic regularities.

This thesis has important scientific and social implications. Concrete proposals have emerged from the research results, for both good and poor spellers, promoting the acquisition of multigraphic phonemes and the academic success of all students.

Keywords:

Acquisition of lexical spelling, multigrapheme, graphotactic regularities, derivational morphology

## INTRODUCTION

La plupart des chercheurs et des intervenants en milieu scolaire s'entendent sur l'importance de maîtriser l'orthographe lexicale, communément appelée orthographe d'usage. L'acquisition de cette compétence est parfois vue comme un fardeau par les élèves, qui l'associent à un effort cognitif démesuré.

L'orthographe lexicale du français est effectivement difficile à acquérir. Mais est-elle vraiment sertie dans un étau, cette langue belle<sup>2</sup> quand on lui demande de soigner ses lettres et ses accents ? En fait, ce n'est pas un effort cognitif qui est recherché dans cette maîtrise réquisitionnée de l'orthographe lexicale par les systèmes scolaires à travers le monde. Au contraire, plusieurs chercheurs s'entendent sur le fait que l'objectif premier de son acquisition par les élèves est d'arriver à écrire correctement les mots avec le moins d'efforts cognitifs possible (Fayol et Jaffré, 1999; Lévesque et al., 2014). En effet, les élèves qui maîtrisent l'orthographe lexicale pourront ensuite davantage se concentrer sur le processus d'écriture (Treiman et al., 2019) puisque c'est en automatisant les mécanismes de base, comme l'orthographe, qu'ils pourront mobiliser les ressources attentionnelles nécessaires pour rédiger un texte (Labrecque, 2011). De plus, l'atteinte de cet objectif entraînera des conséquences importantes dans toutes les sphères de leur vie. Car parvenir à développer les compétences nécessaires pour rédiger un texte, c'est arriver à élaborer ses idées et à les faire valoir. C'est arriver à participer activement à la société. C'est se prévaloir d'un droit fondamental qui revient à chacun (UNESCO, 2008).

Toutefois, pour arriver à cet objectif louable, il faut d'abord acquérir l'orthographe lexicale de sa langue. Il serait donc normal de s'attendre à ce qu'un grand nombre de chercheurs se soient penchés sur ce sujet. Pourtant, les recherches portant sur l'orthographe lexicale ont longtemps été laissées de côté au profit de recherches sur la lecture (Bégin et al., 2010; Fayol et Jaffré, 2016). Depuis quelques années, la situation change et de plus en plus de recherches s'intéressent à cet aspect de la langue écrite.

---

<sup>2</sup> Librement inspiré de la chanson *La Langue de chez nous* d'Yves Duteil (1985).

Bien que les chercheurs s'intéressent maintenant davantage à l'orthographe lexicale, l'origine des erreurs orthographiques est encore peu étudiée (Fayol et Jaffré, 2008). Daigle et al. (2013) ont identifié les deux sources d'erreurs en orthographe lexicale les plus fréquentes chez les élèves, soit les lettres muettes non porteuses de sens et la multigraphémie. Les lettres muettes non porteuses de sens ne s'expliquent pas par la stratégie de dérivation. Par exemple, les élèves risquent davantage de commettre une erreur en orthographiant le mot « brebis » qu'en orthographiant le mot « lit » puisqu'ils n'arriveront pas à faire référence à des mots de la même famille morphologique pour expliquer la lettre muette « s », non porteuse de sens, dans « brebis », mais qu'ils arriveront à faire dériver le mot « literie » de « lit ».

Quant à la multigraphémie, elle se définit par la présence d'un phonème (son) qui peut être transcrit par plusieurs graphèmes (une lettre ou une séquence de lettres de l'alphabet). Cette situation crée alors une inconsistance phonographémique à cause du choix entre trop de graphèmes possibles pour un même phonème et entraîne des erreurs orthographiques. Ce type de phonème est alors appelé multigraphémique. Par exemple, les élèves pourront orthographier plus facilement le mot « mouton » que le mot « agneau », puisque, hormis la lettre « t » qui pourrait être doublée, le mot « mouton » ne comporte à peu près pas de situations dans lesquelles les phonèmes, c'est-à-dire les sons, peuvent être orthographiés de plusieurs façons, donc par plusieurs graphèmes différents. En effet, si on ne tient pas compte des possibles lettres muettes finales, même les phonèmes de ce mot qui correspondent à des graphèmes de plus d'une lettre, soit le /u/ et le /ɔ̃/, ne peuvent s'orthographier qu'avec un ou deux graphèmes chacun, soit « ou » et « on » ou « om ». Mais revenons à notre « agneau » ... En l'orthographiant, les élèves risquent fort d'hésiter sur la graphie du phonème /o/ et de commettre une erreur. « Agno » ? « Agnau » ? « Agnô » ? « Agnot » ? Les élèves pourront se poser 46 questions. En effet, le phonème /o/ correspond à une situation de multigraphémie liée à une inconsistance phonographémique importante, puisqu'en comptant les situations comportant des lettres muettes finales, ce phonème /o/ pourrait s'écrire de 46 façons différentes (Dumais et al., 2014). C'est précisément à l'acquisition de tels phonèmes multigraphémiques que la présente thèse s'intéresse.

Cette thèse a pour objectif d'identifier les facteurs linguistiques qui favorisent l'acquisition des phonèmes multigraphémiques dans les productions orthographiques des élèves québécois bons et faibles orthographieurs de la troisième année du primaire. Elle compte cinq chapitres.

À partir de l'importance de la littératie pour la réussite scolaire et sociale, le premier chapitre aborde la problématique de recherche. Il expose les données actuelles de réussite en orthographe lexicale ainsi que les outils actuels qui en régissent l'acquisition. Enfin, il présente les facteurs jouant notamment un rôle dans l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques.

Le deuxième chapitre présente le cadre théorique et les objectifs de recherche. Il décrit, entre autres, le système orthographique du français et les modèles théoriques d'acquisition du langage écrit. Il définit aussi les facteurs individuels et linguistiques qui interviennent dans l'acquisition de l'orthographe lexicale et dans l'acquisition des phonèmes multigraphémiques. Enfin, il décrit les bases de données ainsi que les échelles d'acquisition en orthographe lexicale et il présente les objectifs et les hypothèses de recherche.

Le troisième chapitre s'attarde à la méthodologie qui permet d'atteindre les objectifs de recherche. Il présente la posture épistémologique ainsi que le devis et les variables de l'étude. Puis, les choix méthodologiques concernant les participants, le déroulement de l'étude, les instruments de collecte de données, le plan d'analyse ainsi que les enjeux déontologiques sont expliqués.

Le quatrième chapitre présente les résultats. À partir de statistiques descriptives, d'analyses multiniveaux et de comparaisons multiples entre les bons et les faibles orthographes, il évalue le rôle des facteurs linguistiques dans l'acquisition du phonème multigraphémique / $\tilde{\epsilon}$  /.

Finalement, le cinquième chapitre permet de discuter du rôle des facteurs linguistiques dans l'acquisition de l'orthographe lexicale chez les bons et les faibles orthographes de la troisième année du primaire. Il met en lumière l'implication des résultats obtenus pour la pratique et pour la recherche. Enfin, il présente les implications des nouvelles connaissances pour la pratique ainsi que les limites et les futures études.

# CHAPITRE 1

## PROBLÉMATIQUE

La préoccupation à l'origine de cette recherche concerne l'acquisition des phonèmes multigraphémiques, c'est-à-dire des sons pouvant être transcrits par plusieurs lettres ou séquences de lettres différentes, chez les élèves québécois bons et faibles orthographes de la troisième année du primaire. Les enjeux qui sont à l'origine de cette préoccupation sont l'importance de la littératie pour la réussite scolaire et sociale de tous les élèves, la réussite en orthographe, la prise en compte des phonèmes multigraphémiques dans divers outils linguistiques qui régissent l'acquisition de l'orthographe lexicale et les facteurs qui influencent son acquisition, notamment en contexte québécois. La problématique est donc structurée en fonction de ces enjeux majeurs en éducation. La question générale de recherche ainsi que les contributions de la recherche concluent ce chapitre.

### **1.1 Importance de la littératie pour la réussite scolaire et sociale de tous les élèves**

Plusieurs chercheurs ont travaillé à définir la littératie au cours des dernières décennies. Au Québec, Hébert et Lépine (2013) ont fait une importante recension des écrits sur le sujet. Ils ont constaté que nombre de chercheurs soulèvent l'importance de l'interdisciplinarité de ce concept évolutif lié à l'appropriation de l'écrit. En effet, précisent-ils, les compétences de ce domaine d'étude sont associées au calcul ainsi qu'à l'écriture et à la lecture, qui agissent en interaction avec l'oral. « Cette compétence multidimensionnelle et complexe de l'appropriation de la culture langagière [...] se développe tout au long du cheminement scolaire et dans différents contextes de vie », précisent Moreau et al. (2021, p. 2).

La littératie est indispensable à la réussite scolaire et sociale. Elle ouvre non seulement la porte aux autres apprentissages, mais elle est à la base des relations personnelles et sociales des élèves (ministère de l'Éducation et de l'Enseignement supérieur, 2017). L'UNESCO (2008) considère d'ailleurs que « l'alphabétisme » est un droit fondamental puisque des avantages importants en découlent, dont le fait de pouvoir prendre des décisions éclairées, de pouvoir s'enrichir et de participer à la société. De bonnes compétences en littératie, estime l'UNESCO (2016), ont des avantages humanitaires, politiques, culturels, sociaux et économiques; elles sont aussi étroitement

liées à des programmes d'enseignement de qualité, estime l'OCDE (2014). En effet, l'éducation et l'apprentissage ont un effet important sur les compétences en littératie (OCDE et Statistique Canada, 2012).

L'acquisition de l'orthographe, cette « capacité d'écrire les mots de manière normée » (Daigle et al., 2013, p. 22) est un préalable important à la réussite éducative et à la littératie (Stanké et al., 2021). Alors que l'orthographe grammaticale cible « les accords et les relations entre les constituants de la phrase » (Montésinos-Gelet et al., 2022, p. 279), l'orthographe lexicale se définit comme la « manière d'écrire les sons ou les mots d'une langue, en conformité d'une part avec le système de transcription graphique adopté à une époque donnée, d'autre part suivant certains rapports établis avec les autres sous-systèmes de la langue (morphologie, syntaxe, lexique) » (Catach, 2016, p. 16).

L'orthographe lexicale du français est difficile à acquérir, entre autres, parce que certaines correspondances entre des phonèmes et des graphèmes sont plus complexes à acquérir. C'est le cas des phonèmes multigraphémiques, c'est-à-dire des sons pouvant s'écrire de plusieurs façons. En français, si nous ne tenons pas compte des possibilités de lettres muettes finales, c'est le phonème /ɛ̃/ qui correspond au plus grand nombre de graphèmes possibles, c'est-à-dire « in », « ain », « ein », « en », « yn », « in », « im », « aim », « ym ».

Des difficultés à acquérir les phonèmes multigraphémiques et l'orthographe lexicale peuvent avoir des répercussions non négligeables sur les apprentissages des élèves (ministère de l'Éducation, du Loisir et du Sport, 2012). Ces difficultés sont présentes dans plusieurs systèmes orthographiques, dont le français. Les statistiques de réussite actuelles en français écrit en témoignent d'ailleurs.

## **1.2 Réussite en orthographe**

Les données actuelles en orthographe montrent que les difficultés sont présentes autant en anglais qu'en français et autant en Amérique du Nord qu'en France (Mauroux et al., 2018). Il est toutefois parfois difficile de bien séparer les données en orthographe grammaticale des données en orthographe lexicale. Au Québec, par exemple, les grilles d'évaluation ministérielles imposent la combinaison de ces deux domaines, faisant en sorte qu'il est ardu de distinguer ce qui appartient à l'un et à l'autre. Bien que nous arrivions à départager quelque peu les statistiques associées aux

deux types d'orthographe, nous présentons donc ici principalement des données globales sur l'orthographe.

L'orthographe représente un défi pour les élèves québécois. En 2008, un comité d'experts sur l'apprentissage de l'écriture a révélé que les élèves québécois éprouvent des difficultés importantes d'apprentissage du français écrit; les principales difficultés portent sur l'orthographe lexicale et sur la syntaxe (MELS, 2008). Plus précisément, en 2010, 11,3 % des élèves québécois de la quatrième année du primaire et 43,3 % des élèves de cinquième secondaire ont montré des acquis insuffisants en orthographe dans l'épreuve ministérielle en écriture (ministère de l'Éducation du Loisir et du Sport [MELS], 2012). Cette maîtrise insuffisante de l'orthographe pourrait se répercuter sur les apprentissages des élèves dans toutes les matières, estime le MELS (2012). Le tableau 1.1 présente les taux de réussite en orthographe des élèves de quatrième année et de sixième année du primaire ainsi que des élèves de deuxième année et de cinquième année du secondaire aux épreuves ministérielles en français écrit de 2010.

**Tableau 1.1 Taux de réussite pour les critères liés à l'orthographe aux épreuves ministérielles de français écrit (adapté du ministère de l'Éducation, du Loisir et du Sport, 2012)**

<b>Critères</b>	<b>Niveaux</b>	<b>Taux de réussite</b>
Orthographe d'usage et grammaticale	4 <sup>e</sup> primaire	88,7 %
Orthographe d'usage et grammaticale	6 <sup>e</sup> primaire	88,9 %
Orthographe d'usage et grammaticale	2 <sup>e</sup> secondaire	60,3 %
Orthographe d'usage et grammaticale	5 <sup>e</sup> secondaire	56,7 %

L'analyse des taux de réussite de ce tableau montre une régression marquée du critère « orthographe d'usage et orthographe grammaticale » entre le primaire et le secondaire. En effet,

entre la quatrième année du primaire et la cinquième année du secondaire, le taux de réussite chute de plus de 30 %. En orthographe lexicale, ces données peuvent s'expliquer par le fait que les mots faisant l'objet d'une correction dans les épreuves ministérielles en français écrit au primaire ne sont que ceux dont l'orthographe a été étudiée en classe. Plus précisément, la progression des apprentissages (PDA) exige que l'orthographe de plus de 1 500 mots soit acquise à la fin de la quatrième année du primaire et l'orthographe de plus de 3 000 mots à la fin de la sixième année du primaire (ministère de l'Éducation du Loisir et du Sport, 2011).

Le critère « orthographe d'usage et grammaticale » est un critère bien général. Boivin et Pinsonneault (2018) ont étudié plus spécifiquement le type d'erreurs que les élèves québécois commettent dans les épreuves ministérielles. Expliquant que les grilles de correction actuelles ne permettent pas de préciser suffisamment le type d'erreurs des élèves, elles ont analysé environ 1 000 productions écrites d'élèves de quatrième année et de sixième année du primaire ainsi que de deuxième année et de cinquième année du secondaire. Les résultats de leur étude montrent que ce sont les erreurs de syntaxe qui sont les plus nombreuses. Viennent ensuite les erreurs d'orthographe grammaticale, puis les erreurs d'orthographe lexicale. Ces résultats portent à croire que l'orthographe lexicale est assez bien maîtrisée par les élèves (Mauroux et al., 2018). Or, il faut tenir compte du fait que l'usage du dictionnaire est autorisé pour ces épreuves. La surcharge cognitive des élèves, qui consacrent beaucoup d'énergie à la rectitude de l'orthographe lexicale et qui n'arrivent pas, par la suite, à construire des phrases correctes et bien accordées, est aussi une hypothèse à envisager. En effet, quand une tâche cognitive demande beaucoup de ressources, il reste moins de ressources pour les autres tâches. C'est ce qu'on appelle la surcharge cognitive. La formation des lettres et l'orthographe lexicale sont deux activités de bas niveau qui peuvent nuire à des activités de haut niveau comme l'élaboration écrite (Treiman et Kessler, 2014). Ainsi, lorsqu'ils ne possèdent l'orthographe que de quelques mots, les élèves construisent leurs phrases et leurs formulations en même temps qu'ils gèrent leur orthographe, ce qui leur demande d'importantes ressources attentionnelles (Fayol et Jaffré, 1999) et occasionne de nombreuses erreurs de tous genres. Il est donc essentiel que les élèves acquièrent l'orthographe de suffisamment de mots pour développer leurs compétences en production écrite (Lévesque et al., 2014).

Les difficultés en orthographe lexicale ne sont pas propres aux élèves québécois. Des études ont, entre autres, révélé que les erreurs liées à cette compétence augmentent depuis plusieurs années

chez les élèves du primaire, et ce, tant au Québec qu'en France (Mauroux et al., 2018). Une vaste enquête de Manesse et Cogis (2007) a révélé que les compétences orthographiques ont baissé chez les apprenants depuis un siècle et que l'écart entre les élèves forts et les élèves faibles s'est creusé. En effet, entre 1987 et 2005 en France, le nombre d'élèves arrivant à faire moins de six erreurs d'orthographe dans une même dictée a chuté de 50 % à 22 %. D'ailleurs, en 2015, à partir de la plus grande base de données à propos de l'orthographe en France, le baromètre Voltaire a conclu que les Français maîtrisent de moins en moins bien l'orthographe de leur langue (Deschamps, 2015). La situation est aussi préoccupante dans la langue de Shakespeare, entre autres aux États-Unis, où les difficultés en orthographe lexicale affectent le sentiment de compétence des élèves. Un sondage mené auprès d'apprenants new-yorkais a d'ailleurs révélé que 15 à 20 % des élèves détestent écrire ou refusent même de le faire à cause de leurs difficultés en orthographe lexicale (Chandler, 2000).

Ces données mettent en évidence l'importance de s'interroger sur l'acquisition de l'orthographe lexicale et de ses phonèmes multigraphémiques. Divers facteurs influencent cette acquisition. Ils sont présentés dans la prochaine section.

### **1.3 Facteurs jouant un rôle dans l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques**

Divers facteurs influencent l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques. Ils peuvent être individuels, environnementaux ou linguistiques. Le rôle de chacun de ces facteurs est expliqué dans la présente section.

#### **1.3.1 Facteurs individuels**

Tout d'abord, des facteurs individuels, c'est-à-dire liés aux différences culturelles, à la motivation et aux capacités cognitives des apprenants, jouent un rôle important dans l'acquisition de l'orthographe lexicale (Cogis, 2005a). Les prochains paragraphes s'attardent à deux de ces facteurs individuels, soit aux facteurs conatifs et aux facteurs cognitifs.

Les facteurs conatifs ont un rôle à jouer dans l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques. Ils sont liés aux dispositions psychologiques des élèves, à la façon dont ils se perçoivent à l'école ainsi qu'à leur motivation et à leurs intérêts (Berger et Desrochers,

2011). Une étude réalisée en 2012 révèle qu'un pourcentage important d'élèves québécois ont une mauvaise perception d'eux et peu de motivation face à l'écriture (ministère de l'Éducation, du Loisir et du Sport, 2012). Plus précisément, dès la fin du troisième cycle du primaire, 28,4 % des élèves affirment ne pas être en accord avec le fait qu'ils écrivent bien et 35,2 % d'entre eux affirment être en désaccord avec le fait qu'ils aiment écrire des textes. D'ailleurs, mis à part le clavardage et les réponses aux courriels, plus de 77 % des élèves du troisième cycle du primaire affirment ne presque jamais pratiquer d'activités d'écriture à l'extérieur de l'école (ministère de l'Éducation, du Loisir et du Sport, 2012). Cette aversion des élèves pour l'écriture peut être causée par des difficultés en orthographe lexicale (Chandler, 2000).

Les facteurs cognitifs ont aussi un rôle à jouer dans l'acquisition de l'orthographe lexicale et de ses phonèmes multigraphémiques. Ils sont notamment liés aux capacités mnésiques, aux capacités attentionnelles, à l'automatisation, aux capacités langagières, dont la conscience phonologique, et aux fonctions exécutives (Berger et Desrochers, 2011; Bousquet et al., 1999; Stanké, 2009). La dénomination automatisée rapide, qui demande de nommer le plus rapidement possible des stimuli présentés de façon aléatoire, la conscience phonémique, la connaissance des lettres et des sons ainsi que l'empan de la mémoire verbale sont d'ailleurs des prédicteurs importants de l'apprentissage de la lecture et de l'écriture dans toutes les orthographe alphabétiques (Caravolas et al., 2012, 2013).

Ainsi, des difficultés ou des troubles d'apprentissage peuvent interférer avec l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques. Les troubles d'apprentissage se distinguent des difficultés d'apprentissage par leur origine neurologique ainsi que leur caractère permanent, c'est-à-dire les répercussions qu'ils auront tout au long de la vie des élèves, et ce, en dépit d'interventions adaptées (Lussier et al., 2017). Bien que les difficultés d'apprentissage soient transitoires, c'est-à-dire qu'elles sont temporaires et peuvent donc se résorber avec des interventions adaptées (Fontaine, 2019), elles sont toutes aussi importantes que les troubles d'apprentissage. En effet, dans une perspective inclusive, il importe de tenir compte du fait que les élèves à besoins particuliers peuvent éprouver des difficultés persistantes, qu'ils aient été évalués par un professionnel ou non (Moreau et al., 2005).

Les élèves handicapés ou en difficulté d'adaptation ou d'apprentissage (EHDA) sont nombreux au Québec. En 2012-2013, leur proportion était estimée à 19,1 % (commission de la culture et de

l'éducation, 2015). Ces élèves produisent davantage d'erreurs en orthographe lexicale que les élèves n'éprouvant pas de difficultés : ils font notamment plus d'erreurs phonologiques (par exemple, « trou » plutôt que « tour »), plus d'erreurs liées aux frontières lexicales (par exemple « loiseau » plutôt que l'oiseau ») et plus d'erreurs liées aux visuogrammes, comme les lettres muettes non porteuses de sens (par exemple, « jamais » plutôt que « jamais ») et aux phonèmes multigraphémiques (par exemple, « lapain » plutôt que « lapin ») (Montésinos-Gelet et al., 2022).

Le trouble des apprentissages spécifiques de la lecture et de l'expression écrite (dyslexie/dysorthographe) représente 80 % des troubles d'apprentissage chez les enfants d'âge scolaire (Neault et Guay, 2006; Shaywitz et al., 1995). C'est ce trouble d'apprentissage qui altère le plus les compétences en orthographe lexicale (Stanké et al., 2016). En effet, l'orthographe est l'une des difficultés les plus courantes chez les élèves ayant des troubles de l'apprentissage (Wanzek et al., 2006). Chez les apprenants dyslexiques/dysorthographiques, les traitements phonologique et lexical peuvent être altérés. Par ailleurs, la plupart d'entre eux présentent des difficultés dans le développement de ces deux types de traitements. Les élèves qui éprouvent des difficultés avec la procédure phonologique, peineront à lire et à écrire des mots réguliers alors que ceux qui éprouvent des difficultés avec le traitement lexical peineront à lire et à écrire des mots irréguliers (Sprenger-Charolles et Colé, 2013; Stanké et al., 2016). En plus de rencontrer des difficultés avec le traitement alphabétique, qui est lié à la conscience phonémique, ces élèves rencontrent aussi des difficultés avec le traitement orthographique, qui est lié aux règles orthographiques, leur difficulté la plus marquée étant celle de choisir le bon graphème pour un phonème donné (Fontaine, 2019). Toutefois, il semble que leur traitement morphologique, qui leur permet d'écrire correctement les mots à partir de leurs dérivations possibles, est le mieux préservé (Daigle et al., 2016). Certes, ces quelques rappels non exhaustifs à propos des élèves ayant un trouble spécifique d'apprentissage, dont la dyslexie/dysorthographe, pourraient, en soi, faire l'objet d'une étude pour mieux comprendre les difficultés à orthographier. Toutefois, la présente recherche doctorale se veut inclusive et intègre tous les élèves faibles orthographes, c'est-à-dire ceux qui éprouvent des difficultés transitoires et ceux qui éprouvent des difficultés persistantes, par exemple un trouble d'apprentissage spécifique. Ce qui importe ici est donc de s'intéresser à l'acquisition des phonèmes multigraphémiques chez les élèves faibles orthographes, qu'ils rencontrent des difficultés ponctuelles ou persistantes.

Ainsi, les facteurs individuels jouent un rôle important dans l'acquisition des phonèmes multigraphémiques. Les enseignants ont aussi un rôle à jouer dans l'enseignement d'une langue comportant autant de ces phonèmes complexes, d'autant que plus du tiers des élèves présentent une aversion pour l'écriture. Ils doivent, en effet, adapter leurs méthodes ; c'est là un facteur environnemental important.

### **1.3.2 Facteurs environnementaux**

L'enseignement que reçoivent les élèves ainsi que les outils linguistiques qui leurs sont présentés sont un facteur non négligeable dans l'acquisition des phonèmes multigraphémiques. La présente section s'intéresse à ces facteurs environnementaux.

#### **1.3.2.1 Enseignement explicite de l'orthographe lexicale**

Il est connu que l'enseignement explicite de l'orthographe lexicale aurait des effets positifs sur son acquisition et donc sur l'acquisition des phonèmes multigraphémiques. Fayol et Jaffré (1999) ont précisé, il y a plus de 20 ans, que l'acquisition de l'orthographe lexicale du français doit nécessairement passer par un enseignement explicite. D'autres chercheurs français, entre autres Pacton et Afonso Jaco (2015), et chercheurs québécois, dont Stanké et al. (2020), ont récemment souligné les bienfaits de l'enseignement explicite de l'orthographe lexicale. En effet, l'exposition à l'écrit, l'enseignement explicite, la sensibilisation aux propriétés visuelles des mots, c'est-à-dire à leurs particularités et à leur enveloppe visuelle, et la mise en place de dispositifs différenciés favorisent l'acquisition de l'orthographe lexicale (Daigle et al., 2015). De plus, l'enseignement explicite est particulièrement aidant auprès des élèves qui éprouvent des difficultés d'apprentissage (Chapleau et al., 2013; Daigle et al., 2015).

Toutefois, l'enseignement explicite de l'orthographe lexicale est encore peu réalisé dans les écoles. Des chercheurs ont constaté que les enseignants, tant au Québec qu'à l'étranger, sont peu outillés pour mettre en pratique les résultats des recherches en orthographe lexicale (Adoniou, 2013; Daigle et al., 2015; Mansour, 2012). Daigle et ses collaborateurs (2015) ont mené une étude québécoise sur l'orthographe lexicale auprès de 14 enseignants et de 143 élèves de deuxième année du primaire de 2012 à 2015. Ils ont conclu que les enseignants manquent de connaissances liées aux propriétés visuelles des mots et que leurs méthodes d'enseignement de l'orthographe sont peu variées. Ils

proposent d'adapter le programme de formation, de renouveler les contenus en formation des maîtres et de réviser les pratiques enseignantes afin de s'assurer, entre autres, que le personnel enseignant puisse transmettre les facteurs linguistiques du français écrit.

Divers outils linguistiques, comme des listes orthographiques, peuvent permettre aux élèves d'acquérir l'orthographe lexicale. Toutefois, dans la plupart des pays, ces échelles ne tiennent compte ni de l'acquisition des mots par les élèves ni des phonèmes multigraphémiques, qui sont pourtant une source importante d'erreurs. Ce facteur environnemental est présenté dans la prochaine sous-section.

### **1.3.2.2 Prise en compte des phonèmes multigraphémiques dans des outils linguistiques qui régissent l'acquisition de l'orthographe lexicale**

Les bases de données lexicales peuvent être basées sur la fréquence d'apparition des mots dans des ouvrages ou sur la réussite des mots par des populations d'élèves. Certaines donnent des informations linguistiques sur les mots (par exemple, le nombre de phonèmes qu'ils contiennent). Elles sont alors infralexicales. Ces bases de données devraient tenir une place importante dans l'enseignement de l'orthographe lexicale puisqu'elles donnent des indications précises sur les contenus à enseigner et sur les principes de base à respecter. Brissaud et Cogis (2011b) ont dressé une liste de cinq principes qui sont à la base d'un enseignement « raisonné » de l'orthographe et qui en influencent son acquisition. Dans cette liste, figure l'importance d'« établir une progression véritable [qui] doit articuler les connaissances des élèves, la complexité de l'orthographe française et le temps nécessaire à l'apprentissage » (Brissaud et Cogis, 2011b, p. 10). Certes, cette progression du simple au complexe doit tenir compte des mots et des phonèmes multigraphémiques à aborder afin de les segmenter. Brissaud et Cogis (2011b) précisent toutefois que cette même progression doit aussi pouvoir tenir compte des acquis des élèves et de la fréquence des mots pour qu'ils puissent bien orthographier les mots de chacune des disciplines scolaires. Malgré tout, il est rare que les listes orthographiques présentées aux élèves se basent sur un lexique propre aux disciplines scolaires. Elles ne se basent pas non plus sur l'acquisition des mots par les élèves. En fait, les listes orthographiques sont davantage des bases de données lexicales plutôt que des échelles d'acquisition orthographique. En effet, la plupart des listes orthographiques francophones sont bâties uniquement à partir d'échelles de fréquence. Elles ne fournissent donc pas d'informations

sur la compétence des élèves à orthographier les mots de façon conventionnelle selon leur cycle (Stanké et al., 2019).

Au Québec, la liste orthographique à l'usage des enseignantes et des enseignants (ministère de l'Éducation et de l'Enseignement supérieur, 2014), qui a pour but de fournir une liste de mots dont l'orthographe doit être acquise pour chaque cycle scolaire, est l'une des listes utilisées par les enseignants du primaire. Elle a été construite à partir d'échelles de fréquence de l'apparition des mots dans la littérature jeunesse pour chaque cycle du primaire. Bien qu'elle comporte des avantages, elle ne permet pas de savoir à quel cycle scolaire la majorité des élèves arrivent à orthographier les mots et les phonèmes multigraphémiques. Elle ne permet donc pas d'enseigner la multigraphémie selon une gradation établie. Par exemple, les élèves de deuxième année réussissent-ils à écrire les mots contenant la semi-voyelle /j/ et à faire la distinction entre « ill » et « y » afin d'écrire « vieille » ? Peuvent-ils distinguer le graphème « an » du graphème « en » et bien écrire la finale de « tellement », qui s'explique par le fait que c'est un adverbe (Stanké et Dumais, 2017a) ? Ce sont deux exemples de mots que le ministère propose présentement d'enseigner en deuxième année du primaire (ministère de l'Éducation et de l'Enseignement supérieur, 2014). Pour ce, le ministère se base uniquement sur un facteur linguistique, soit la fréquence, en faisant peu ou pas de liens avec les facteurs linguistiques de la longueur et de la consistance des mots, ni avec les facteurs liés aux phonèmes multigraphémiques, comme la position, le contexte d'utilisation, le voisinage et la dérivation des graphèmes que contiennent ces mots.

Pourtant, selon *l'Échelle d'acquisition en orthographe lexicale* en France (ÉOLE), qui a testé la réussite des élèves en orthographe lexicale (Pothier et Pothier, 2004), l'orthographe des mots « vieille » et « tellement » n'est toujours pas acquise en cinquième année. Plus précisément, la réussite des élèves de cinquième année du primaire en France est de 67 % pour l'orthographe du mot « vieille » et de 60 % pour l'orthographe du mot « tellement ». Il est certes possible de réfuter cette observation en évoquant le fait que cette dernière échelle a été développée à partir des compétences à orthographier les mots auprès d'élèves français et non québécois. Toutefois, la réussite de l'orthographe de ces mots arrive si tard en France, qu'il est difficile de croire que les élèves québécois auraient pu les maîtriser quatre ans plus tôt que leurs compatriotes français. L'élaboration d'une liste orthographique basée uniquement sur le critère de fréquence peut mener à d'autres incohérences. Par exemple, selon la liste ministérielle, l'orthographe du mot « vieille »

devrait être acquise en deuxième année alors que l'orthographe du mot « vieillesse » devrait l'être beaucoup plus tard, soit en sixième année.

Stanké et ses collaborateurs (2018) ont publié une échelle orthographique basée sur l'acquisition des mots chez les élèves. En 2018, ils ont publié l'échelle québécoise d'acquisition de l'orthographe lexicale (ÉQOL). Pour l'instant, ÉQOL permet d'obtenir des statistiques d'acquisition des mots contenant des phonèmes multigraphémiques. En effet, les mots testés contiennent des correspondances phonèmes-graphèmes consistantes, comme le phonème /u/ qui s'écrit toujours « ou » et inconsistantes, comme le phonème /ɛ̃/ qui peut être représenté à l'écrit par plusieurs graphèmes, dont le « in », le « ain » et le « ein » et il est possible de valider la réussite des mots par les élèves. Toutefois, cette échelle d'acquisition de l'orthographe lexicale ne permet pas de cibler où les erreurs ont été commises dans les mots. Par exemple, elle indique le pourcentage d'élèves de troisième année qui ont orthographié correctement le mot « princesse », mais elle ne permet pas de savoir où l'erreur est située. Comme ce mot contient plusieurs correspondances phonèmes-graphèmes inconsistantes, les élèves auraient pu faire plusieurs erreurs, notamment « praincesse », « prinsesse », « prinsaisse » ou « princèce ». Bien qu'elle fournisse des données importantes sur l'acquisition des mots entiers chez les élèves québécois du primaire, à ce jour, cette base de données ne fournit donc pas d'indices sur l'acquisition des phonèmes multigraphémiques chez les élèves québécois du primaire.

Ainsi, des recherches restent à faire pour enrichir la liste des phonèmes multigraphémiques qui sont à acquérir pour chaque cycle scolaire. L'étude de l'acquisition des phonèmes multigraphémiques permettrait, entre autres, de mieux comprendre les difficultés des élèves. En effet, même si la recherche a montré que les correspondances phonèmes-graphèmes inconsistantes sont plus difficiles à acquérir que les correspondances phonèmes-graphèmes consistantes (Morin et Montésinos-Gelet, 2005), à notre connaissance, aucune recherche ne s'est encore penchée sur les inconsistances qui posaient le plus de difficultés. Une telle recherche permettrait davantage de comprendre les difficultés rencontrées dans l'acquisition des phonèmes multigraphémiques.

L'inexistence d'une liste présentant une progression dans le cas de phonèmes pouvant être représentés par plusieurs graphèmes explique peut-être, en partie, les difficultés des enseignants à comprendre et à enseigner l'orthographe lexicale. Ces difficultés ont été soulignées par plusieurs

chercheurs, dont Adoniou (2013), Daigle et al. (2015) ainsi que Mansour (2012). Ce manque d'aisance part, entre autres, d'un doute des enseignants en lien avec leurs propres compétences en orthographe lexicale (Adoniou, 2013) et s'exprime notamment par une tendance à faire apprendre les mots à la maison sans les travailler en classe (Mansour, 2012). Cette tendance peut s'expliquer par un manque de connaissances au sujet de l'importance des propriétés visuelles des mots (par exemple, le nombre de lettres des mots et leur ordre), ce qui incite les enseignants à mettre l'accent sur la phonologie au détriment d'autres traitements importants pour l'acquisition de l'orthographe. C'est pourtant la conscience et la coordination des quatre traitements suivants qui sont à la base de l'apprentissage de la lecture et de l'écriture : logographique, phonologique, orthographique et morphologique (Écalle et Magnan, 2015). Les enseignants négligent aussi l'enveloppe visuelle des mots quand ils enseignent l'orthographe lexicale (Daigle et al., 2015). Ces propriétés visuelles sont essentielles pour acquérir les mots irréguliers, mais aussi les graphèmes plus exceptionnels liés aux phonèmes multigraphémiques. En effet, Treiman et Kessler soulignent que l'apprentissage de mots de listes orthographiques par cœur, sans faire appel aux composantes des mots, ralentit le moment où les élèves deviennent assez autonomes pour élaborer des textes par eux-mêmes puisqu'ils doivent mémoriser de nombreux mots avant d'y arriver. Ils ajoutent que l'enseignement explicite des correspondances entre les phonèmes et les graphèmes permet un transfert de règles sur bon nombre de mots, ce qui contribue à l'autonomie des élèves, qui arrivent à orthographier davantage de mots plus rapidement.

Il importe donc de bonifier les outils linguistiques régissant l'acquisition de l'orthographe lexicale. Nous sommes d'avis que le fait d'y inclure une progression basée sur l'acquisition des phonèmes multigraphémiques permettrait aux élèves de mieux acquérir l'orthographe lexicale. Toutefois, pour y arriver, il faut d'abord en apprendre davantage sur l'acquisition de ces phonèmes chez les élèves bons orthographes, mais aussi chez les élèves faibles orthographes, d'autant qu'il est connu que l'écart entre les élèves forts et les élèves faibles est de plus en plus grand (Manesse et Cogis, 2007). Une étude de l'acquisition des phonèmes multigraphémiques permettrait de comprendre quels graphèmes associés à ces phonèmes sont acquis en premier et de savoir si les difficultés liées à la multigraphémie sont les mêmes chez les élèves forts et chez les élèves faibles orthographes. Une telle étude permettrait éventuellement, tant aux enseignants qu'aux

orthopédagogues, de mieux intervenir, d'autant que les facteurs linguistiques liés à l'acquisition du français sont nombreux.

### **1.3.3 Facteurs linguistiques**

Les facteurs linguistiques exercent une influence certaine sur l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques. Dans le cas d'un système orthographique complexe comme celui de l'anglais ou du français, ces facteurs jouent un rôle encore plus grand. Écalle et Magnan (2015) soulèvent trois difficultés importantes dans le système orthographique du français, que Fayol (2003) considère d'ailleurs comme l'un des plus difficiles au monde. Tout d'abord, une même lettre peut avoir diverses valeurs. Il est possible, ici, de penser à la lettre « s » qui peut représenter les phonèmes /s/ ou /z/, par exemple, et qui peut être muette et représenter le pluriel. Le fait que la plupart des lettres sont muettes en fin de mots est une autre difficulté soulignée par Écalle et Magnan (2015). Certaines de ces lettres s'expliquent par leur dérivation comme le « t » de « fort », alors que d'autres ne sont pas porteuses de sens comme le « d » de « foulard ».

De plus, dans notre système orthographique, la majorité des phonèmes peuvent être représentés par plusieurs graphèmes. En effet, quelques consonnes ainsi que dix des seize voyelles du français nécessitent une combinaison des lettres de l'alphabet, ce qui est aussi une source d'erreurs importante pour les élèves (Dumais et al., 2014). Par exemple, comme nous l'avons vu, la voyelle nasale /ẽ/ peut correspondre à de multiples graphèmes, mais elle peut aussi être associée à des semi-voyelles et former les graphies « oin » et « ien » et même se terminer par des morphèmes dérivables comme le « gt » de « vingt » ou non dérivables comme le « g » de « shampooing ». Tant de possibilités de graphèmes pour un seul phonème rendent son acquisition plus difficile pour les élèves (Daigle et al., 2013; Fayol et Jaffré, 2008; Pacton et Afonso-Jaco, 2015; Seymour et al., 2003). Certains des graphèmes associés à un phonème sont plus susceptibles d'être difficiles à acquérir que d'autres. Par exemple, pour l'inconsistance phonographémique du /ẽ/, nous pouvons penser qu'il est plus facile d'acquérir le graphème « in », qui est très fréquent, que le graphème « ein », qui est peu fréquent, ou encore que le graphème « en », qui représente généralement un autre phonème, soit le phonème /ã/. Ce constat mène au questionnement de la présente recherche qui s'interroge sur les facteurs linguistiques qui influencent l'acquisition des phonèmes

multigraphémiques chez les élèves bons et faibles orthographes de la troisième année du primaire.

La troisième année du primaire est une année clé dans l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques. En effet, ce sont les interactions entre les facteurs individuels, linguistiques et environnementaux qui font en sorte que c'est vers l'âge de huit ans que les élèves arrivent à acquérir les phonèmes multigraphémiques, et ce, dans plusieurs langues européennes (Seymour, 2007). Au Québec, les élèves de huit ans sont habituellement en troisième année du primaire. Pour analyser convenablement l'acquisition des phonèmes multigraphémiques, il est aussi important que les élèves testés aient acquis le traitement alphabétique et qu'ils soient à développer leur traitement orthographique, ce qui est le cas en troisième année du primaire. En effet, en première année, les enseignants doivent aider tous les élèves à atteindre la phase alphabétique complète; les principales connexions graphèmes-phonèmes, en particulier celles qui impliquent des voyelles, doivent alors être apprises (Ehri, 1998). Ainsi, le traitement alphabétique, qui permet de produire des mots réguliers en faisant une correspondance entre les phonèmes et les graphèmes, s'automatise vers la fin de la deuxième année du primaire (Seymour, 2007). De plus, c'est à la fin du premier cycle du primaire que les élèves arrivent à faire un traitement qui s'appuie sur la forme orthographique des mots, car ils arrivent à enregistrer des mots en mémoire (Ehri, 1995). Enfin, Pacton et al. (2001) ont prouvé que c'est à la fin de la deuxième année du primaire que les élèves francophones arrivent à réaliser des tâches faisant appel à une sensibilité aux régularités orthographiques.

Il est donc logique de s'intéresser à l'acquisition des phonèmes multigraphémiques en troisième année du primaire puisque les élèves doivent maîtriser le traitement alphabétique pour arriver à faire appel de façon interactive au traitement orthographique. Ce traitement fait, entre autres, appel aux règles et aux régularités orthographiques, qui sont indispensables à l'acquisition de phonèmes multigraphémiques. D'ailleurs, au Québec, la progression des apprentissages (ministère de l'Éducation du Loisir et du Sport, 2011) précise que c'est à la fin de la deuxième année du primaire que les élèves connaîtront les graphèmes les plus courants pour représenter un phonème et à la fin de la troisième année du primaire qu'ils identifieront le rôle des lettres muettes qui marquent le lien morphologique ainsi que les règles de position (le « m » devant le « p » et le « b », par exemple). C'est aussi en troisième année du primaire qu'ils commenceront à apprendre les constantes

orthographiques associées à l'emploi de certains graphèmes (par exemple, le graphème « in » qui est le seul graphème possible en début de mot sauf dans le cas de « ainsi »). Ainsi, pour bien comprendre l'acquisition des phonèmes multigraphémiques chez les élèves bons et faibles orthographes, la cible de la troisième année du primaire semble la plus intéressante, d'autant que les graphèmes associés au phonème /  $\tilde{\epsilon}$  / comportent jusqu'à trois lettres. Ces trigrammes sont plus complexes à acquérir que les graphèmes simples. En effet, plus un graphème utilisé est complexe, c'est-à-dire plus il comporte de lettres, plus son acquisition est difficile (Fayol et Jaffré, 2008; Morin et Montésinos-Gelet, 2005).

En somme, il est clair qu'il faut effectuer des recherches concernant les facteurs linguistiques qui favorisent l'acquisition des phonèmes multigraphémiques chez les élèves québécois de la troisième année du primaire afin d'avoir une meilleure compréhension de l'acquisition de l'inconsistance phonographémique. D'ailleurs, à notre connaissance, peu ou pas de chercheurs se sont penchés sur ce sujet. Plusieurs chercheurs ont toutefois déjà souligné l'importance d'effectuer des recherches sur les facteurs linguistiques qui influencent l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques.

À l'international, Catach (2016) a clairement établi le pourcentage d'utilisation des graphèmes dans les mots en français, mais les données concernant leur acquisition sont inexistantes. Pérez (2014a) a mentionné l'importance de tenir éventuellement compte de l'impact des facteurs linguistiques sur l'acquisition de l'orthographe lexicale, comme la fréquence, le voisinage, la durée acoustique et la longueur orthographique. Pour sa part, Pacton (2008) a précisé que, afin de mieux comprendre l'acquisition de l'orthographe lexicale, les recherches devraient se centrer sur plusieurs types d'informations, tant graphotactiques, c'est-à-dire en se basant sur les statistiques des correspondances phonèmes-graphèmes, que morphologiques, c'est-à-dire en se basant sur les dérivations des graphèmes. De son côté, Caravolas (2019) a mentionné que plusieurs facteurs devront être analysés pour comprendre les effets déstabilisants du français sur le développement des compétences en orthographe et sur l'acquisition du code alphabétique, entre autres, chez les élèves du premier cycle du primaire.

Au Québec, Daigle et ses collègues (2015) ont suggéré des pistes de recherche à la suite de l'étude qu'ils ont réalisée sur l'orthographe lexicale auprès des élèves en difficulté dans le cadre de

*L'action concertée portant sur l'écriture.* La première piste proposée s'intéresse au rôle que tiennent les procédures visuelles ainsi que certains procédés orthographiques dans l'acquisition de l'orthographe lexicale. La deuxième piste propose de vérifier s'il est possible pour les élèves d'apprendre la forme d'un mot nouveau de façon globale. Si ce n'est pas le cas, elle propose d'identifier quelle serait « la séquence des apprentissages en fonction des marques orthographiques, de la régularité orthographique, etc. » (Daigle et al., 2015, p. 24). Stanké et ses collaborateurs (2020) viennent, pour leur part, de faire paraître leur rapport de recherche portant sur une nouvelle approche basée sur un enseignement orthographique favorisant l'apprentissage de la production écrite et de l'orthographe lexicale des élèves faibles orthographes de la troisième année du primaire. Ils y ont souligné l'importance de mieux comprendre les facteurs linguistiques qui sont liés à l'acquisition des régularités orthographiques et d'élaborer une liste hiérarchisée de ces régularités.

#### **1.4 Question de recherche**

De cet étayage des enjeux dans l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques liés aux facteurs individuels, environnementaux et linguistiques, la question de recherche émerge :

Quels sont les facteurs linguistiques qui favorisent l'acquisition des phonèmes multigraphémiques dans les productions orthographiques des élèves québécois bons et faibles orthographes de la troisième année du primaire ?

#### **1.5 Contributions de la recherche**

La présente recherche doctorale, qui se situe à la jonction des sciences de l'éducation et de la linguistique appliquée, est susceptible d'avoir des incidences tant sur le plan scientifique que sur les plans pratique et social. Sur le plan scientifique, elle pourrait permettre d'en connaître davantage sur l'acquisition des phonèmes multigraphémiques chez les élèves québécois faibles et bons orthographes. D'un point de vue pratique, ces nouvelles connaissances pourraient permettre ultérieurement de bâtir une échelle d'acquisition des phonèmes multigraphémiques. Enfin, sur le plan social, des propositions concrètes pourraient ressortir des résultats de recherche, tant pour les

bons orthographes que pour les faibles orthographes, favorisant ainsi l'acquisition de l'orthographe lexicale et la réussite scolaire de tous les élèves.

La force scientifique de cette recherche doctorale repose sur l'arrimage de tous les concepts théoriques afin de produire un apport scientifique différent. En analysant les données de Pacton, de Sprenger-Charolles, de Pothier et Pothier, de Mousty et Leybaert et de plusieurs autres chercheurs en orthographe lexicale, nous comprendrons mieux ce qui est à la base des difficultés d'acquisition des phonèmes multigraphémiques au Québec. L'étude des facteurs linguistiques liés à l'acquisition du phonème multigraphémique /*ẽ*/ pourrait, en effet, apporter des réponses intéressantes.

Socialement, ces nouvelles données québécoises en orthographe lexicale permettront éventuellement au personnel enseignant de faire des choix dans l'enseignement des phonèmes multigraphémiques. Par exemple, la liste orthographique à l'usage des enseignantes et des enseignants (ministère de l'Éducation et de l'Enseignement supérieur, 2014) compte 174 mots contenant le phonème /*ẽ*/. Ces mots ont été choisis selon la fréquence de leur apparition dans la littérature jeunesse et non selon l'acquisition de ceux-ci par les élèves. Cette prévalence de la littérature jeunesse sur la véritable acquisition des mots a, entre autres, pour effet qu'un mot aussi complexe que « princesse » doit être appris en 2<sup>e</sup> année, alors que des mots simples comme « foin », « brin » et « gamin » doivent être appris en 6<sup>e</sup> année. Le mot « invitation » doit, pour sa part, être appris en 3<sup>e</sup> année, le mot « inviter » en 4<sup>e</sup> année et le mot « invité » en 5<sup>e</sup> année.

Sans pour autant avoir la prétention de refaire toutes les listes orthographiques actuelles, la présente recherche doctorale pourrait permettre de corriger des incongruités au besoin et de mener éventuellement à une première échelle d'acquisition des phonèmes multigraphémiques basée sur une recherche doctorale au Québec. Ainsi, les enseignants seraient éventuellement mieux outillés en matière d'enseignement des phonèmes multigraphémiques. Conséquemment, l'apprentissage de l'orthographe lexicale des élèves du primaire serait facilité puisqu'ils pourraient s'approprier l'orthographe de façon mieux organisée.

En somme, la présente recherche contribue au contexte scolaire sur différents plans, notamment grâce à la jonction qu'elle fait entre les sciences de l'éducation et de la linguistique appliquée. Ces

apports sont au centre de notre démarche de recherche, dont les principales composantes sont définies dans le cadre théorique.

## **CHAPITRE 2**

### **CADRE THÉORIQUE**

Le chapitre du cadre théorique permet de définir les concepts liés à l'objectif général de recherche. Ce cadre se divise en six parties. Comme ce projet s'interroge sur les facteurs linguistiques qui favorisent l'acquisition des phonèmes multigraphémiques, il semble d'abord important de présenter le système orthographique du français, dans la première section, et les modèles théoriques d'acquisition du langage écrit dans la deuxième. Ensuite, comme ce projet est mené auprès d'élèves bons et faibles orthographes, une analyse des facteurs individuels et linguistiques qui interviennent dans l'acquisition de l'orthographe lexicale et de ses phonèmes multigraphémiques est faite. Puis, afin de clarifier les principales listes orthographiques présentement disponibles pour favoriser l'acquisition des phonèmes multigraphémiques chez les élèves québécois de la troisième année du primaire, la section cinq est consacrée aux bases de données en orthographe lexicale. Enfin, la partie six fait une synthèse des recherches relatives à l'orthographe lexicale et aux phonèmes multigraphémiques et justifie ainsi l'objectif de recherche visé.

#### **2.1 Système orthographique du français**

Afin de bien saisir l'acquisition des phonèmes multigraphémiques, il est essentiel d'en présenter le système orthographique. Pour ce, il importe d'abord de définir l'orthographe et la multigraphémie. Ensuite, la présentation de ce plurisystème pourra servir de base à la compréhension.

##### **2.1.1 Définition de l'orthographe**

L'orthographe permet de transformer des formes linguistiques en une forme codée, c'est-à-dire en mots écrits (Sprenger-Charolles et al., 1997). Cette constituante de l'écriture se compose de deux parties : l'orthographe lexicale et l'orthographe grammaticale.

L'orthographe lexicale est recensée dans les dictionnaires ou dans les bases de données informatisées. Catach (2011) la définit comme « la manière d'écrire les sons ou les mots d'une langue, en conformité, d'une part, avec le système de transcription graphique adopté à une époque

donnée, d'autre part, suivant certains rapports établis avec les autres sous-systèmes de la langue » (p. 16). Aussi appelée « orthographe d'usage », elle est liée au lexique, « lexicon » signifiant « livre de mots » (Rey, 1998). C'est la « forme droite, correcte, selon les normes définies par sa culture », dit Cogis (2005b, p. 57). Elle permet d'encoder des formes linguistiques en mots écrits selon Sprenger-Charolles (1997) et de reconnaître les mots et de les différencier, selon Fayol et Jaffré (1999).

L'orthographe grammaticale est liée aux flexions des mots, c'est-à-dire à des variations, qui sont, pour la plupart, muettes (Fayol et Jaffré, 2014). Bien qu'elle soit considérée comme une « zone à haut risque » par Manesse et Cogis (2007, p. 97), entre autres parce que ce sont les mêmes graphèmes qui codent les phonèmes en orthographe lexicale qui sont utilisés pour marquer les valeurs grammaticales, l'orthographe grammaticale n'est pas abordée dans la présente thèse.

### **2.1.2 Multigraphémie**

La linguistique cherche le moyen de représenter les sons de manière univoque (Saussure, 1985) et s'intéresse à l'analyse des unités élémentaires de la langue (Benveniste, 2006), dont les phonèmes, les graphèmes et les morphèmes. Ce sont le phonème et le graphème qui sont impliqués dans la multigraphémie étant donné le fait que différents graphèmes représentent un même phonème. Le phonème est une unité sonore composée de voyelles ou de consonnes et le graphème est une lettre ou une séquence de lettres représentant un phonème (Stanké et al., 2018). Tous deux sont les plus petites unités significatives d'une chaîne écrite, mais le phonème est à l'oral ce que le graphème est à l'écrit (Catach, 2016). Un graphème dit « simple » est représenté par une seule lettre. Par exemple, le phonème /o/ peut être représenté par une seule lettre, soit le « o ». Il peut aussi être complexe et est alors représenté par plus d'une lettre. Par exemple, le phonème /o/ peut se transcrire à l'aide du graphème « eau », qui est composé de la séquence de lettres « e » « a » « u ». Lorsqu'il est complexe, le graphème peut être formé de deux lettres ou de trois lettres; il porte alors le nom de digramme ou de trigramme, selon le cas. Enfin, un graphème peut être contextuel, c'est-à-dire que son utilisation dépend du contexte. Par exemple le phonème /g/ peut être représenté par le graphème « g » ou par le graphème « gu » en fonction de la lettre qui le suit. Tout comme le phonème et le graphème, le morphème est une unité minimale. Toutefois, le morphème est une unité minimale permettant de transmettre un sens (Pacton, 2008). Dans le cas des morphogrammes

grammaticaux, c'est le contexte syntaxique qui dicte la variation de la forme des mots (par exemple, « ils mangeNT ») alors que, dans le cas des morphogrammes lexicaux, c'est le lien avec les mots de même famille qui dicte la variation de la forme des mots (par exemple, bonD - bonDir) (Montésinos-Gelet et al., 2022).

Le phénomène de la multigraphémie est, avec les lettres muettes non porteuses de sens, la plus grande source d'erreurs en orthographe lexicale (Daigle et al., 2013). Fayol et Jaffré (2008, p. 189) identifient cette « pluralité de transcription [...] [et] les fréquences d'association entre certaines configurations sonores et certaines configurations de lettres » comme étant les facteurs principaux liés aux difficultés orthographiques. En effet, la multigraphémie implique des inconsistances phonographémiques importantes qui sont source d'erreurs. Pour éviter ces erreurs, les élèves doivent stocker visuellement les parties des mots qui contiennent des inconsistances phonographémiques dans leur lexique orthographique.

Le fait qu'un grand nombre de phonèmes puissent s'écrire à l'aide de plusieurs graphèmes différents rend le système orthographique du français opaque. En effet, comparativement au néerlandais, par exemple, langue dans laquelle chaque phonème correspond à un seul graphème, le système orthographique du français recense 130 graphèmes (133 selon Catach, 2016) pour représenter 36 phonèmes, c'est-à-dire 130 façons d'écrire 36 sons (Frost et Katz, 1992).

Le nombre de phonèmes associés à chaque graphème est peu élevé dans le système orthographique du français, ce qui rend son système assez transparent dans le sens de la lecture (Pacton, Foulin et al., 2005). Fayol et Jaffré (2008) estiment d'ailleurs que la consistance des correspondances graphèmes-phonèmes est de 96 % en lecture en français, les rares exceptions étant entre autres associées aux graphèmes contextuels et au graphème « ch » qui correspond aux deux phonèmes suivants en lecture : /ʃ/ (chat) et /k/ (chorale).

Toutefois, l'opacité de la langue française est beaucoup plus importante en orthographe. En effet, la correspondance entre les phonèmes et les graphèmes est moins consistante que la correspondance entre les graphèmes et les phonèmes en français (Fayol et Jaffré, 2008; Pacton, Foulin et al., 2005), ce qui constitue l'une des principales sources de difficultés d'acquisition de l'orthographe lexicale. En effet, seul 71 % des mots peuvent être orthographiés correctement à

partir des correspondances des phonèmes avec les graphèmes (Fayol et Jaffré, 2008). Cela fait donc en sorte que l'acquisition de l'orthographe est plus ardue que l'acquisition de la lecture en français. La figure 2.1 illustre l'opacité du système orthographique de certaines langues.

**Figure 2.1 L'opacité du système orthographique en schéma (reproduit et traduit de Seymour et al., 2003, p. 146)**

		Opacité				
		<b>Structure syllabique</b>	<b>Simple</b>	Finlandais	Grec	Portugais
	Italien					
	Espagnol					
<b>Complexe</b>			Allemand	Néerlandais	Danois	Anglais
			Norvégien			
			Islandais			

Ainsi, la précédente figure illustre bien combien le français et l'anglais sont des langues dites opaques ou profondes (« deep » en anglais). L'anglais est, par ailleurs, une langue plus opaque que le français. De plus, sa structure syllabique est plus complexe, ce qui affecte l'acquisition de sa lecture et donc de son orthographe plus tard. En effet, si les 36 phonèmes du français peuvent s'écrire de 130 façons, comme nous l'avons vu, c'est de 561 façons que peuvent s'écrire les 41 phonèmes de l'anglais (Fayol et Jaffré, 2008). Une telle opacité ralentit l'acquisition de l'orthographe lexicale et de ses phonèmes multigraphémiques.

### 2.1.3 Plurisystème du français

Malgré l'opacité de la langue française, elle n'est pas hétéroclite. Elle est, en effet, gérée par des structures rationnelles qui permettent d'acquérir l'orthographe lexicale à partir d'un ensemble de règles de correspondances entre les phonèmes et les graphèmes (Paret, 2010). C'est Catach qui a été la première chercheuse à décrire ces structures rationnelles en système orthographique du

français en 1973. Selon elle, plus le nombre de graphèmes possibles pour représenter les phonèmes est grand, plus il faut faire appel aux conventions écrites, c'est-à-dire aux normes (Catach, 2016). La constitution du système orthographique en plurisystème est présentée dans cette section.

Le système orthographique du français est défini par Catach (2016) comme un plurisystème, car les mots sont représentés à l'écrit selon trois dimensions. Tout comme plusieurs systèmes d'écriture, par exemple celui de l'anglais et de l'espagnol, il repose sur un système orthographique dit alphabétique, c'est-à-dire que sa base est phonographémique, puisque ses phonèmes se transcrivent par une ou des séquences de lettres de l'alphabet, soit les graphèmes. Une orthographe alphabétique idéale devrait contenir un ensemble de correspondances biunivoques graphème-phonème et phonème-graphème, en d'autres termes, la relation entre les lettres et les sons devrait être totalement cohérente (Caravolas, 2004), ce qui n'est pas le cas du français. Pour décrire la théorie du plurisystème développée par Catach, les dimensions phonographémique, morphographémique et logographique sont exposées dans la présente section. Une dimension visuogrammique, associée à des recherches plus récentes, est aussi présentée.

### **2.1.3.1 Dimension phonographémique**

La dimension phonographémique a pour fonction de coder les phonèmes par des graphèmes. Cette dimension est au cœur du modèle de Catach (2016), qui la qualifie de fonctionnement majeur (Catach, 2011), car ce sont les unités élémentaires du système phonologique du langage oral, les phonèmes, qui sont codées à l'écrit à partir de graphèmes.

La phonographémie, c'est-à-dire les relations entre les phonèmes et leurs graphèmes correspondants, est l'une des stratégies utilisées par les élèves pour orthographier des mots dont l'orthographe leur est inconnue en faisant appel aux correspondances phonèmes-graphèmes qu'ils connaissent. Morin et Montésinos-Gelet (2005) ont montré que, dès la maternelle, les enfants développent des connaissances sur la phonographémie. Cette stratégie est aussi très utilisée par les élèves du premier cycle du primaire, car leur lexique orthographique, c'est-à-dire le nombre de mots qu'ils savent orthographier, n'est pas encore développé. Ces novices apprenants identifient les phonèmes des mots et leur font correspondre les graphèmes les plus fréquents ou les graphèmes les plus consistants du système orthographique.

Seuls les mots constitués de ce que Mousty et ses collègues (1994) qualifient de graphèmes consistants acontextuels pourront être orthographiés de façon phonographémique en respectant les normes orthographiques. Par exemple, les graphèmes suivants sont consistants et ne dépendent pas du contexte : les graphèmes simples « m » et « i » ainsi que les graphèmes complexes « gn » et « on ». Le mot « mignon » peut donc être orthographié en faisant appel uniquement à la dimension phonographémique. Le tableau suivant présente des graphèmes consistants acontextuels, c'est-à-dire qui ne varient pas selon le contexte.

**Tableau 2.1 Consistances phonographémiques en français à partir des graphèmes dont le pourcentage d'utilisation est de plus de 90 % (adapté de Catach, 2016, p. 10 à 15)**

Phonèmes	Mots contenant les graphèmes consistants associés aux phonèmes
/b/	<b>b</b> outeille
/d/	<b>d</b> entiste
/g/	<b>g</b> randir (ou « <b>gu</b> » devant « e » et « i »)
/l/	<b>l</b> apin
/m/	<b>m</b> itaine
/n/	<b>n</b> ez
/p/	<b>p</b> einture
/ʀ/	<b>r</b> ue
/t/	<b>t</b> alon
/v/	<b>v</b> acances
/k/	<b>c</b> ou (ou « <b>qu</b> » devant « e » et « i » - 2 % des mots avec un <b>k</b> )
/a/	<b>p</b> atte
/ɑ/	ch <b>ɑ</b> t
/ə/	v <b>ə</b> nir
/i/	m <b>i</b> re
/ø/	bl <b>eu</b>
/u/	f <b>ou</b>
/y/	t <b>u</b>

Phonèmes	Mots contenant les graphèmes consistants associés aux phonèmes
/œ̃/	<b>brun</b> (rares mots en « um » et en « eun » : parfum, à jeun)
/ɔ̃/	<b>bon</b> (exception : 7 % de mots en « om » acquis à la fin du primaire)
/w a/	<b>moi</b>
/w ɛ̃/	<b>loin</b>
/ɲ/	poign <u>é</u> e
/ʃ/	<b>chandail</b> (certains mots en sch et en sh : schéma et shampoing)

Ces correspondances consistantes entre les phonèmes et les graphèmes facilitent leur acquisition. Fait à souligner, que ces graphèmes consistants soient simples ou complexes, c'est-à-dire composés d'une ou de plusieurs lettres, Catach (2016) ne se base que sur leur consistance pour considérer qu'ils pourront tous être acquis à la fin du deuxième cycle du primaire.

Les correspondances entre les phonèmes et les graphèmes ne sont toutefois pas toujours consistantes. La dimension phonographémique du système orthographique du français est qualifiée d'inconsistante (Borchardt, 2012), car les phonèmes sont multigraphémiques (Catach, 2016). C'est le cas du /ɛ̃/ qui peut être représenté par les graphèmes de plus d'une lettre suivants « in », « ain », « ein », « en », « yn » et « in », dont certains varient selon le contexte et deviennent donc contextuels devant le /p/ et le /b/ : « im », « aim », « ym ». C'est à ces différentes orthographes d'un même son (Catach, 2016), orthographes aussi appelées inconsistances phonographémiques (Borchardt, 2012), ou multigraphémie, que la présente thèse s'intéresse.

Le tableau 2.2 présente des cas d'inconsistances phonographémiques ainsi que leur archigraphème, que Catach définit comme le graphème principal, ou le plus fréquemment utilisé, parmi l'ensemble des graphèmes plausibles. Les exemples de ce tableau illustrent bien l'inconsistance de la dimension phonographémique du français, ce qui rend l'acquisition de l'orthographe lexicale plus difficile que dans d'autres systèmes plus transparents. En effet, l'utilisation des correspondances phonèmes-graphèmes ne permet d'orthographier que la moitié des mots monosyllabiques du français (Véronis, 1988).

**Tableau 2.2 Inconsistances phonographémiques du français (Adapté de Catach, 2016, p. 10 à 15)**

Phonèmes	Archigraphèmes	Exemples de mots contenant une inconsistance phonographémique
f	<b>f</b>	facile ; <b>ph</b> araon
k	<b>c</b>	carotte ; <b>ch</b> orale ; koala
s	<b>s</b>	son ; tasse ; science ; cèleri ; option
ʃ	<b>ch</b>	<b>ch</b> anter ; <b>sch</b> éma ; <b>sh</b> ampooing
v	<b>v</b>	valse ; wagon
z	<b>s</b> <sup>3</sup>	rasoir ; zoo
ʒ	<b>j</b>	<b>j</b> eu ; <b>g</b> eler ; mang <b>e</b> ait
e	<b>e</b>	les ; blé ; chez ; aimer
ɛ	<b>e</b>	terre ; père ; traite
i	<b>i</b>	vie ; bicyclette
œ	<b>eu</b>	<b>j</b> eune ; <b>œ</b> uf
o	<b>o</b>	pot ; chaud ; bateau
ɔ	<b>o</b>	sort ; rhum
ã	<b>an</b>	rang ; vent ; <b>faon</b> ; entendre
ẽ	<b>in</b>	individu ; <b>païn</b> ; <b>faim</b> ; ceinture ; <b>lynx</b> ; <b>thym</b> ; examen
œ̃	<b>un</b>	brun ; parf <b>um</b>
j	<b>i</b> <sup>4</sup>	miel ; balayer ; fille ; travail
w	<b>oi</b>	<b>roi</b> ; <b>oui</b> ; noyer ; watt

Catach (2016) a établi manuellement un pourcentage approximatif d'utilisation de chacun des graphèmes associés aux phonèmes. Ces pourcentages permettent d'en saisir le degré de consistance. Elle a aussi divisé ces correspondances phonèmes-graphèmes en trois niveaux. Le premier niveau

<sup>3</sup> Ici, Catach (2011, p. 13), identifie le Z comme étant l'archigraphème. Pourtant, elle identifie le pourcentage d'utilisation du graphème intervocalique / s / à 90 %, en faisant clairement l'archigraphème du phonème [z]

<sup>4</sup> Ici, Catach (2011, p. 13), identifie le IL ou le Y comme étant l'archigraphème. Pourtant, elle identifie le pourcentage d'utilisation du graphème / i / à 86 %, en faisant clairement l'archigraphème du phonème [j]

comporte les graphèmes de base les plus fréquents associés aux phonèmes, par exemple, le graphèmes « in » et « en » pour le phonème /ɛ̃/. La chercheuse considère que ce niveau devrait être acquis aisément en fin de deuxième cycle du primaire. Quant à lui, le deuxième niveau devrait être acquis en fin de scolarité du primaire puisqu'il est associé à des graphèmes moins fréquents, par exemple les graphèmes « im », « ain » et « ein » pour le phonème /ẽ/. Enfin, le troisième niveau contient des consonnes doubles et beaucoup de lettres exceptionnelles logographiques, par exemple le « aim », le « yn » et le « ym ». Comme ces graphèmes doivent être traités globalement, Catach considère qu'ils devraient être acquis au cours du secondaire. Ce classement basé sur la fréquence d'utilisation des graphèmes, ou sur leur consistance, ne permet pas toujours de faire une hypothèse quant à la facilité d'acquisition d'un graphème par rapport à un autre; par exemple, le phonème /ã/ s'écrit « an » dans 44 % des cas et « en » dans 47 % des cas. Dans d'autres cas, la consistance permet d'avoir une meilleure idée de la facilité d'acquisition d'un graphème. Par exemple, le phonème /ẽ/ s'écrit « in » dans 45 % des cas et « ein » dans moins de 11 % des cas, ce qui suggère que l'acquisition des mots comportant le graphème « in » sera probablement plus facile pour les élèves que l'acquisition des mots comportant le graphème « ein ». Le tableau 2.3 présente le pourcentage de l'utilisation des graphèmes et l'année scolaire de leur acquisition selon Catach.

**Tableau 2.3 Pourcentage d'utilisation des graphèmes dans les cas principaux d'inconsistances phonographémiques (adapté de Catach, 2016, p. 10 à 15)**

Phonèmes	Graphèmes de base	Année scolaire d'acquisition - Catach	Pourcentage approximatif d'utilisation
/ɛ/	e è	Fin 2 <sup>e</sup> c. primaire	68 %
	ai	Fin 2 <sup>e</sup> c. primaire	30 %
	ei	Fin du primaire	2 %
/o/ /ɔ/	o	Fin 2 <sup>e</sup> c. primaire	75 %
	au	Fin 2 <sup>e</sup> c. primaire	21 %
	eau	Fin 2 <sup>e</sup> c. primaire	3 %
	ô	Fin du primaire	Moins de 1 %
	u (m)	Fin du primaire	Moins de 1 %
/ã/	an	Fin 2 <sup>e</sup> c. primaire	44 %

<b>Phonèmes</b>	<b>Graphèmes de base</b>	<b>Année scolaire d'acquisition - Catach</b>	<b>Pourcentage approximatif d'utilisation</b>
/ɛ/	e è	Fin 2 <sup>e</sup> c. primaire	68 %
	ai	Fin 2 <sup>e</sup> c. primaire	30 %
	ei	Fin du primaire	2 %
/o/ /ɔ/	o	Fin 2 <sup>e</sup> c. primaire	75 %
	au	Fin 2 <sup>e</sup> c. primaire	21 %
	eau	Fin 2 <sup>e</sup> c. primaire	3 %
	ô	Fin du primaire	Moins de 1 %
	u (m)	Fin du primaire	Moins de 1 %
	en	Fin 2 <sup>e</sup> c. primaire	47 %
	am et em	Fin du primaire	9 %
/ɛ̃/	in	Fin 2 <sup>e</sup> c. primaire	45 %
	en	Fin 2 <sup>e</sup> c. primaire	23 %
	ain	Fin du primaire	21 %
	ein	Fin du primaire	Moins de 11 %
	im	Fin du primaire	Moins de 11 %
/j/	i ï	Fin 2 <sup>e</sup> c. primaire	86 %
	il ill	Fin 2 <sup>e</sup> c. primaire	10 %
	y	Fin 2 <sup>e</sup> c. primaire	3 %
/s/	s ss	Fin 2 <sup>e</sup> c. primaire	69 %
	c ç	Fin 2 <sup>e</sup> c. primaire	26 %
	t	Fin du primaire	3 %
/z/	s	Fin 2 <sup>e</sup> c. primaire	90 %
	z	Fin 2 <sup>e</sup> c. primaire	10 %
/ʒ/	j	Fin 2 <sup>e</sup> c. primaire	49 %
	g ge	Fin 2 <sup>e</sup> c. primaire	51 %
/f/	f	Fin 2 <sup>e</sup> c. primaire	95 %
	ph	Fin du primaire	5 %
/ks/	x	Fin 2 <sup>e</sup> c. primaire	84 %
	cc xc	Fin du primaire	16 %

Catach (2016) considère également que la plupart des graphèmes présentés seront acquis à la fin du deuxième cycle du primaire. Le degré d'inconsistance ne semble pas un facteur considéré dans l'acquisition des différents graphèmes; seuls les cas se situant autour de 5 % d'utilisation et la graphie contextuelle du « m » devant le « b » et le « p » sont présentés comme pouvant être acquis plus tard, soit à la fin du primaire.

Ainsi, bien que la dimension phonographémique soit considérée comme étant essentielle, elle est insuffisante pour acquérir l'orthographe des mots. Plus que de savoir utiliser correctement les correspondances phonèmes-graphèmes à partir du traitement alphabétique, les élèves doivent développer des connaissances lexicales, orthographiques et morphologiques (Écalte et Magnan, 2015). La dimension morphographémique du système orthographique du français est présentée dans la prochaine section.

### **2.1.3.2 Dimension morphographémique**

La dimension morphographémique est reliée à la morphologie, c'est-à-dire à la façon dont les mots sont formés (Pothier, 2011). Cette dimension a pour fonction de coder à l'écrit le sens des mots à partir de marques graphiques constituant les plus petites unités de sens : les *morphèmes*. La morphologie peut être flexionnelle; les morphogrammes sont alors grammaticaux et ils transmettent les marques du genre, du nombre, du temps, etc. (Catach, 2016). La morphologie peut aussi être dérivationnelle; les morphèmes sont alors lexicaux et fournissent des informations syntaxiques et sémantiques à partir d'affixes (préfixes ou suffixes) ou de morphogrammes lexicaux, soit de lettres ou de groupes de lettres le plus souvent en fin de mots, qui sont presque toujours muettes en français et qui servent à faire un lien visuel entre les mots d'une même famille morphologique (par exemple, le « d » du mot « blond » qui sert à faire un lien avec les dérivés du mot : « blonde », « blondeur » et « blondinet »).

En orthographe lexicale, l'utilisation de la morphologie dérivationnelle permet aux élèves de réduire l'ambiguïté graphémique. Par exemple, le morphème « ette », signifiant « petit », permet rapidement d'identifier le bon graphème correspondant au phonème /ɛ/ dans le mot « fourchette »

(Chapleau, 2013a) parmi toutes les représentations graphémiques possibles, soit « è », « e », « ai », « aî », « ay », « ei », « ë », « ê ».

Les erreurs qui sont associées à la dimension morphographémique relèvent d'un fonctionnement second selon Catach (2011). Elles peuvent être à dominante morphographémique lexicale et concerner les morphèmes dérivationnels (par exemple, « accros » pour « accroc ») ou de type flexionnel (par exemple, « ils mange » pour « ils mangent »).

Toujours selon Catach (2011), les lettres non dérivables par la morphologie ne peuvent s'acquérir que par un enseignement explicite menant à la mémorisation de l'orthographe des mots; par exemple, le « s » de « velours » ne s'explique pas par un mot dérivable de la même famille. Ces cas relèvent davantage de l'étymologie et sont donc associés à un fonctionnement que Catach qualifie de "hors système". Daigle et al. (2013) classent d'ailleurs ces erreurs comme des erreurs orthographiques visuelles et non comme des erreurs morphologiques. La mémorisation des mots contenant de tels graphèmes est donc essentielle. La mémoire est aussi directement impliquée dans la dimension logographique du système orthographique du français. Cette dimension est présentée dans la prochaine section.

### **2.1.3.3 Dimension logographique**

Selon Catach (2011), la dimension lexicale logographique permet de récupérer les formes orthographiques lexicales directement en mémoire. Les logogrammes donnent « une image visuelle spécifique à certains mots homophones, afin d'aider à une reconnaissance rapide de leur sens » (Catach, 2016, p. 262). Ils permettent, par exemple, de faire la distinction entre les homophones lexicaux « paire » ou « père » ou les homophones grammaticaux « ce » et « se ». Le mot connu est alors lu d'un coup d'œil sans porter attention aux lettres; l'image du mot ayant une valeur que Saussure (1985) qualifie d'idéographique.

Daigle et ses collègues (2013) apportent une précision à la dimension logographique de Catach en ajoutant une dimension qu'ils nomment dimension supralexicale idéogrammique. Ils mentionnent que cette dimension permet, elle aussi, de faire la distinction entre deux homophones (Catach, 2011), mais qu'elle peut être réalisée à partir d'idéogrammes, c'est-à-dire de traits d'union, d'apostrophes et de majuscules. Ainsi, en l'absence d'acquisition de cette dimension, l'apprenant

pourrait faire l'erreur d'écrire « qu'elle » plutôt que « quelles », erreur liée à l'idéogramme de l'apostrophe.

La dimension supralexicale idéogrammique n'est pas la seule dimension que Daigle et ses collègues (2013) précisent. En fait, ils s'appuient sur les dimensions du plurisystème de Catach et divisent le système orthographique du français en deux propriétés : les propriétés associées à des graphèmes transmettant un son ou un sens et les propriétés visuogrammiques. Les propriétés associées à des graphèmes transmettant un son ou un sens se rattachent respectivement aux dimensions phonographémique et morphographémique de Catach, alors que les propriétés visuogrammiques se rattachent à la dimension logographique de Catach. Daigle et al. (2013) précisent que 50 % des erreurs en orthographe lexicale proviennent de difficultés avec les propriétés visuelles et sont donc liées à la dimension visuogrammique des mots. Ils exploitent la dimension visuogrammique en la divisant en plusieurs sous-dimensions rattachées à ce qu'ils appellent les propriétés visuelles des mots. Ces propriétés sont importantes pour comprendre l'acquisition de l'orthographe lexicale et de ses phonèmes multigraphémiques et méritent de s'y attarder.

#### **2.1.3.4 Dimension visuogrammique**

Comme les correspondances phonèmes-graphèmes ne permettent d'écrire que la moitié des mots monosyllabiques dans le système orthographique du français (Véronis, 1988), l'orthographe de certains mots s'acquiert en les voyant, en y étant exposé fréquemment et en s'attardant à leur particularité ou à leur enveloppe visuelle (Daigle et al., 2013). Plus précisément, les élèves ont recours à l'aspect général des mots à partir de leurs propriétés visuelles pour la lecture et visuo-orthographiques pour la production de mots (Daigle et al., 2018).

Les propriétés visuelles et visuo-orthographiques sont nécessaires pour traiter les homophones lexicaux, les lettres muettes non porteuses de sens, les lettres géminées et les cas de multigraphémie (Daigle et al., 2018), cas qui nous intéressent particulièrement dans le cadre de cette recherche doctorale. Par exemple, les élèves ont recours à l'aspect général des mots et utilisent leurs propriétés visuelles et visuo-orthographiques pour faire la différence entre les homophones « père », « pair » et « paire », pour bien traiter la lettre muette non porteuse de sens du mot « parapluie »,

pour traiter les lettres géminées de « ballon » et pour traiter le cas de multigraphémie du « an » du mot « antilope » qui aurait aussi pu s'écrire « en » ou du « eau » du mot « bateau », qui aurait pu s'écrire « o » ou « au ». Valdois et ses collègues (2017) associent les propriétés visuelles lexicales au traitement visuoattentionnel, plus spécifiquement à l'empan visuoattentionnel, c'est-à-dire au nombre d'éléments visuels que l'œil peut traiter de façon simultanée. Elles notent d'ailleurs un déficit sur ce plan chez les élèves dyslexiques-dysorthographiques.

Pour orthographier rapidement des mots selon la norme, les élèves doivent donc bien s'en représenter mentalement l'orthographe et stocker leur représentation orthographique dans leur lexique orthographique (Daigle et al., 2013). Le lexique mental est présent lorsque la production des mots est automatisée et que les élèves y ont accès de façon fluide (Handy, 2009). Ce processus visuel d'encodage des représentations orthographiques des mots peut leur éviter des erreurs possibles, entre autres, parce qu'ils comparent la représentation mentale qu'ils se font du mot avec ce qu'ils ont écrit.

Divers modèles théoriques proposent de mieux comprendre comment les élèves arrivent à acquérir ces irrégularités ainsi que les phonèmes multigraphémiques dans un contexte aussi difficile. La prochaine section s'y intéresse.

## **2.2 Modèles théoriques de l'acquisition du langage écrit**

Pour acquérir les phonèmes multigraphémiques, les élèves doivent se référer à toutes les dimensions du système orthographique du français. Considérant l'interaction de toutes ces dimensions, il importe de comprendre les modèles d'acquisition du langage écrit. En effet, pour comprendre comment les élèves apprennent à utiliser les systèmes orthographiques, il faut d'abord comprendre ces systèmes ainsi que les processus cognitifs liés à l'apprentissage (Treiman et Kessler, 2014).

Différents chercheurs ont établi des modèles s'intéressant aux traitements cognitifs impliqués dans la lecture et dans l'écriture. Ces modèles théoriques cognitivistes et connexionnistes d'acquisition du langage écrit décrivent l'acquisition de la lecture et de l'orthographe chez les élèves. Ils permettent aussi de cibler les difficultés en reconnaissance et en production des mots écrits chez les élèves afin de mieux cibler les interventions (Daigle et al., 2018) et même d'émettre des

hypothèses quant au fonctionnement normal ou pathologique de l'acquisition de la lecture et de l'orthographe (Berger et Desrochers, 2011; Stanké, 2009). En effet, ces modèles permettent de rendre compte du cheminement scolaire des élèves et, s'il y a lieu, d'identifier les traitements déficitaires qui sont à l'origine de leurs difficultés (Demont et Gombert, 2004).

Convenons qu'il existe de nombreux modèles théoriques de l'acquisition de l'écrit, certains plus orientés vers la lecture et d'autres vers l'orthographe (Hazard, 2009); leur synthèse serait une étude en soi et n'est pas l'objet de cette thèse. Si les différentes théories sur l'acquisition du langage écrit ont longtemps privilégié un modèle séquentiel, ou développemental, présentant des stages à acquérir en lecture et en orthographe, ils ont par la suite été remis en question parce que cette vision successive d'acquisition de la lecture et de l'orthographe rejetait l'idée de coexistence de divers traitements dès le début de l'acquisition du langage écrit (Daigle et al., 2018; Écalte et Magnan, 2015; Martinet et al., 1999, 2004; Stanké, 2009). La présente thèse s'appuie sur une perspective interactive de l'acquisition de la lecture et de l'orthographe, c'est-à-dire la façon dont les traitements se mobilisent ensemble pour permettre aux élèves de lire et d'orthographier les mots. L'apport de trois de ces modèles, soit celui de Gombert (2003), celui de Seymour (1997) et celui de Daigle et ses collaborateurs (2018), est ici présenté. Ce choix se justifie, entre autres, parce que Daigle et collaborateurs ont bâti leur modèle à partir de l'orthographe du français et que, bien qu'il ait d'abord été validé en anglais, Seymour a aussi validé son modèle en français. Comme la présente recherche s'intéresse à l'acquisition des phonèmes multigraphémiques, c'est-à-dire à leur apprentissage implicite, et non à leur apprentissage explicite, ce type d'apprentissage et le modèle de Gombert (2003), qui s'intéresse à l'apprentissage implicite, sont d'abord présentés.

### **2.2.1 Apprentissage implicite**

L'apprentissage implicite est un traitement qui s'effectue sans intention (Brissaud et Cogis, 2011a; J.-E. Gombert, 2003). Les élèves s'adaptent alors à la situation sans comprendre d'où vient cette adaptation et ce qui l'a causée (Perruchet et Nicolas, 1998). Selon Écalte et Magnan (2015), c'est, entre autres, la répétition du contact avec les mots qui est à l'origine de l'apprentissage implicite de l'orthographe lexicale. Demont et Gombert (2004) considèrent aussi que c'est la fréquence qui est le moteur de l'apprentissage implicite. Nous pourrions donc nous attendre à ce que les

phonèmes multigraphémiques fréquents ainsi que les phonèmes multigraphémiques compris dans les mots fréquents soient acquis plus rapidement.

En effet, l'apprentissage implicite joue un rôle majeur dans l'acquisition des phonèmes multigraphémiques puisqu'il permet aux élèves de faire des analogies en lecture et en écriture de façon précoce (Gombert et al., 1997). De prime abord, les personnes sont sensibles aux caractéristiques de leur environnement à un point tel que leur comportement s'y adapte graduellement sans qu'ils aient à utiliser des connaissances explicites liées à cet environnement; les nombreuses régularités du système orthographique mènent au même comportement (Gombert, 2003). Ce sont les rencontres répétées des élèves avec ces régularités qui leur permettent de les acquérir (Demont et Gombert, 2004) et d'acquérir les phonèmes multigraphémiques. Par exemple, les élèves, qui lisent à répétition des mots contenant le graphème « in » en début de mot et très rarement le graphème « ain » dans cette position devraient être sensibles à cette fréquence et utiliser le graphème « in » plutôt que le « ain » sans avoir à recevoir d'enseignement explicite à ce propos. Cette sensibilité aux régularités orthographiques est présente dès la première année du primaire (Pacton et al., 2001). Bref, le système cognitif s'adapte à la fréquence des correspondances entre les phonèmes et les graphèmes (Gombert, 2003). Les élèves font appel à leurs connaissances implicites, tant pour le traitement des régularités orthographiques que pour le traitement des correspondances entre les graphèmes et les phonèmes (Gombert, 2003). Il en va de même quand les élèves établissent des correspondances entre les phonèmes et les graphèmes. Selon Écalle et Magnan (2015), deux catégories de connaissances orthographiques peuvent se développer de façon précoce grâce à l'apprentissage implicite qui se fait au contact des régularités : les connaissances liées à l'orthographe des mots, qu'ils appellent les « représentations graphiques mentales », et les connaissances liées à des séquences de lettres possibles en français, qu'ils appellent les « patterns orthographiques ».

Pour la majorité des élèves, un processus d'apprentissage implicite, ou inconscient, commence dès qu'ils font leurs premiers contacts avec l'écrit, ce qui leur permet d'acquérir des connaissances importantes concernant la structure du système alphabétique, expliquent Demont et Gombert (2004). Ils ajoutent que les traitements alphabétiques et orthographiques doivent pouvoir s'appuyer sur cet apprentissage implicite. Or, chez les élèves éprouvant des difficultés d'apprentissage, précisent-ils, cet apprentissage implicite fait défaut. Cela cause, entre autres, une incapacité à

emmagasiner inconsciemment les régularités graphotactiques, c'est-à-dire certaines structures répétitives qui sont utilisées dans les mots, ce qui nuit à l'acquisition des phonèmes multigraphémiques.

Le terme « acquisition » est associé aux expériences implicites alors que le terme apprentissage est associé aux expériences explicites, dont l'enseignement. Dans cette thèse, c'est le terme « acquisition » qui est adopté. Ce choix a été fait, car il est impossible de savoir si, dans leur parcours scolaire, les participants de la recherche ont vécu des expériences linguistiques et ont reçu un enseignement explicite menant à l'apprentissage des phonèmes multigraphémiques évalués. Cela s'explique par le fait que les élèves provenaient de plusieurs écoles et qu'ils n'ont pas été testés à la suite d'un programme d'enseignement explicite des phonèmes multigraphémiques.

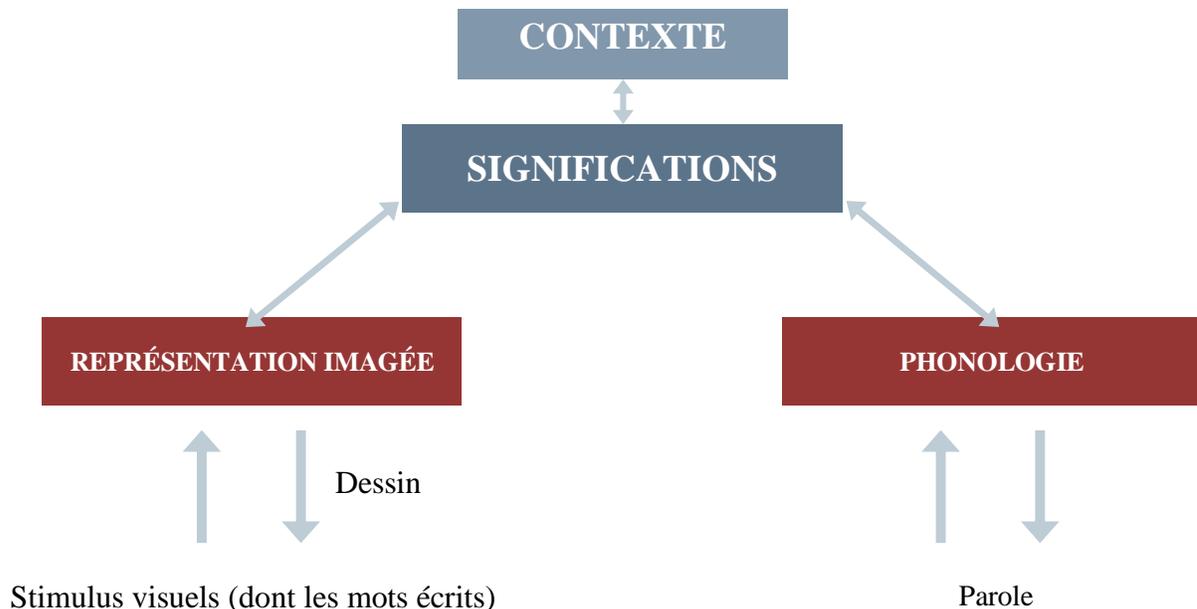
### **2.2.2 Modèle de Gombert**

Le modèle de Gombert donne une place importante aux connaissances implicites. Elles se construisent, estime le chercheur, dès le contact de l'enfant avec l'écrit, aussi précoce soit-il, et ce, même avant l'enseignement de la lecture. Gombert (2003) estime que les régularités orthographiques, alphabétiques et morphologiques jouent un rôle important dans l'acquisition de ces connaissances. Bien que ce modèle s'intéresse à l'acquisition de la lecture seulement, il permet de comprendre l'acquisition de l'écriture puisque c'est la lecture qui est à la base de toutes les autres acquisitions qui sont faites dès les premières années de la scolarité (Demont et Gombert, 2004).

Gombert estime que l'enfant fera inconsciemment très tôt des analogies entre les nouveaux mots qu'il essaie de lire et ceux qu'il connaît. Son système initial, présenté dans la figure 2.2, traite de l'apprentissage implicite de l'écrit chez l'enfant dès l'âge préscolaire. Selon lui, le traitement du langage oral que fait l'enfant servira de base à la façon dont il traitera l'écrit, soit de façon picturale, phonologique, sémantique et contextuelle (Demont et Gombert, 2004). Par exemple, le non-lecteur qui entend le mot « chien » le traitera oralement de façon phonologique et s'en fera une représentation imagée, c'est-à-dire picturale. Il en trouvera aussi le sens grâce au traitement sémantique et le traitement contextuel lui permettra de tenir compte des informations externes. En

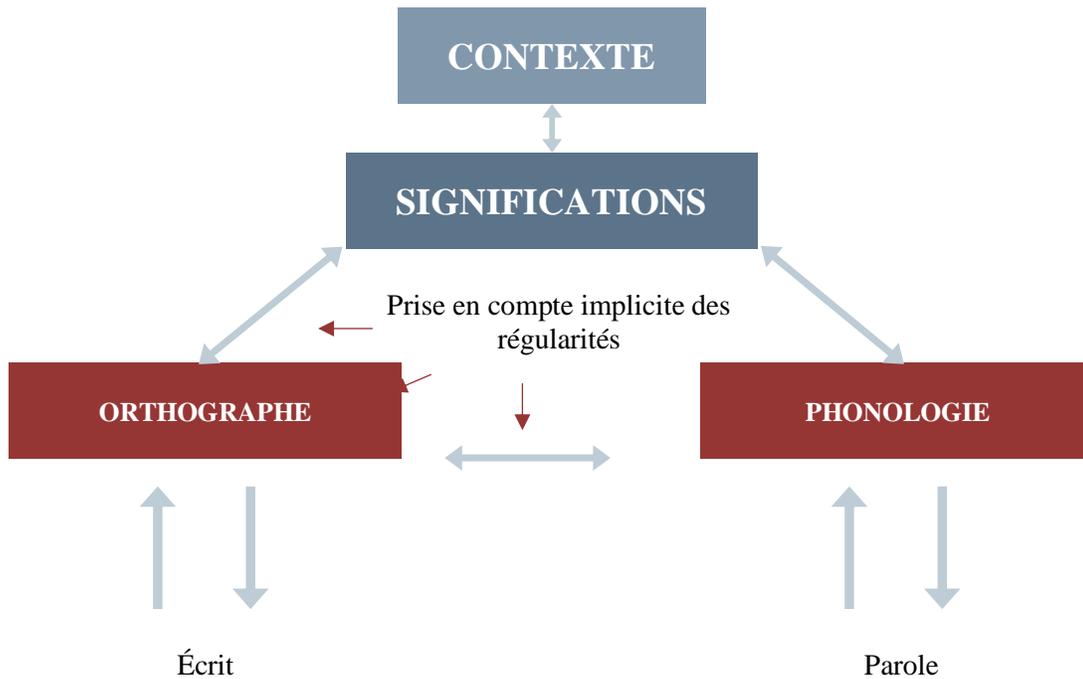
lecture, les stimuli visuels, ou picturaux, passeront des images aux lettres. La phonologie passera de la parole aux correspondances graphèmes-phonèmes.

**Figure 2.2 Le système initial (Demont et Gombert, 2004, p. 252)**



Ainsi, le modèle de Gombert donne une place importante à l'apprentissage implicite, qui continue toute la vie de l'élève. Ce dernier lira et orthographiera mieux s'il lit et écrit beaucoup (Gombert, 2003). Toutefois, Gombert ne néglige pas l'apport de l'apprentissage explicite pour autant. En effet, il considère que l'apprentissage explicite doit compléter les traitements automatiques de l'apprentissage implicite. Ces derniers traitements sont insuffisants pour lire et pour orthographier d'une façon habile. En faisant des liens entre les traitements pictural, phonologique et sémantique du système initial, l'élève arrivera à traiter les lettres de manière orthographique et se familiarisera aux régularités liées aux traitements visuo-orthographique, phonologique, lexical et morphologique (Demont et Gombert, 2004) (voir figure 2.3 ci-dessous). Ces traitements sont définis dans la prochaine section, qui présente le modèle de Seymour.

Figure 2.3 L'apprentissage implicite de la lecture (Demont et Gombert, 2004, p. 253)



Si Gombert (2003) a donné une place importante à l'apprentissage implicite dans son modèle, Seymour a, pour sa part, bâti le sien à partir de l'apprentissage explicite seulement. La lecture répétée des mots n'est effectivement pas suffisante à l'apprentissage de l'orthographe lexicale (Demont et Gombert, 2004; Pacton et Afonso-Jaco, 2015). Karmiloff-Smith (1994) considère d'ailleurs que les représentations implicites sont une première phase avant que les représentations explicites ne se développent. Ainsi, un enseignement planifié ayant pour effet un apprentissage explicite (Perruchet et Pacton, 2004) est aussi nécessaire et l'élève doit alors faire des efforts cognitifs importants.

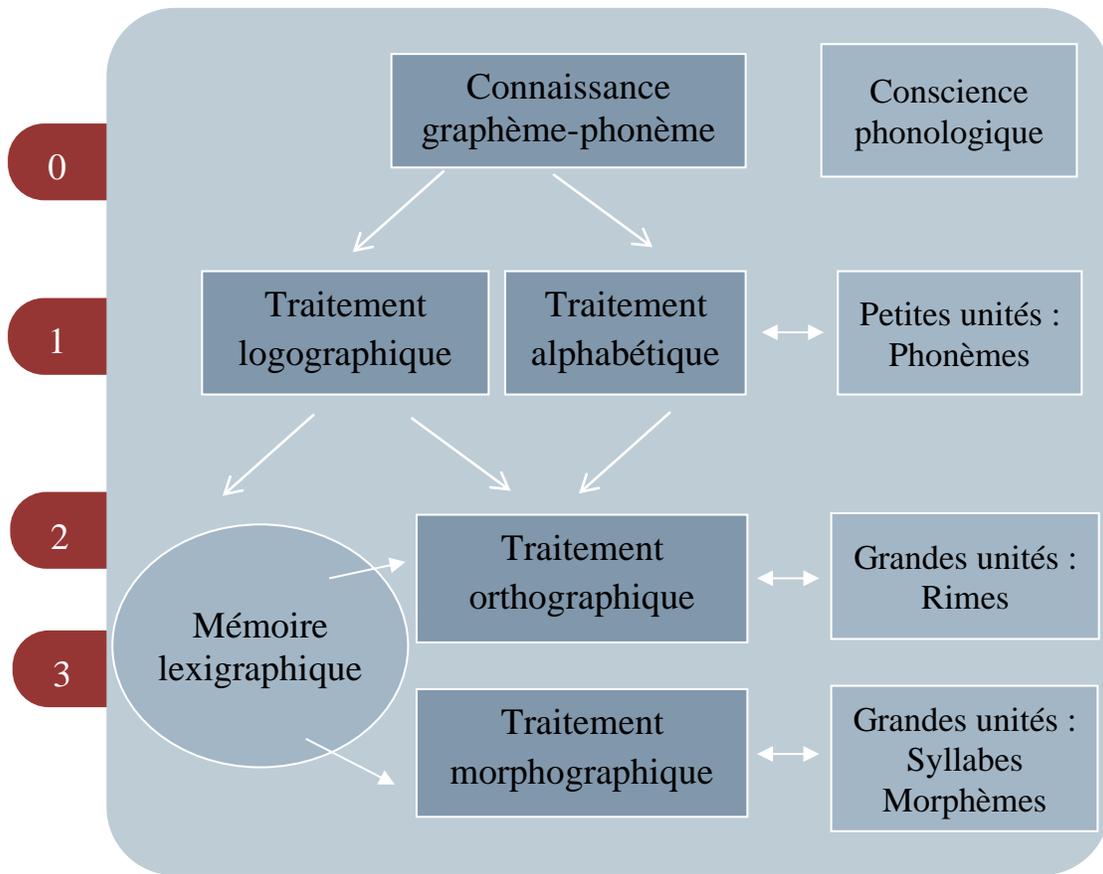
### 2.2.3 Modèle à double fondation de Seymour

Seymour (1997) a proposé un modèle d'acquisition du langage écrit dit à « double fondation ». Ce modèle se base sur quatre composantes n'agissant pas en stages, mais bien en interaction, pour la reconnaissance et pour la production des mots écrits. Le modèle théorique de Seymour s'appuie sur le fait que l'acquisition de l'orthographe lexicale doit passer par des composantes tant lexicales que phonologiques, morphologiques et alphabétiques (Seymour, 1997).

Ce modèle privilégie le codéveloppement des traitements logographique et alphabétique. En effet, Seymour et ses collaborateurs (2003) ont étudié l'acquisition de la lecture et de l'écriture dans 13 langues européennes, dont le français et l'anglais, et ils ont constaté que, dans les langues transparentes, l'acquisition de la lecture et de l'écriture se base sur le traitement alphabétique alors que, dans les langues opaques, cette acquisition doit se baser sur une double fondation, c'est-à-dire sur le traitement logographique et orthographique. Le fait de faire appel à deux traitements, en début de scolarisation, sollicite beaucoup plus les habiletés cognitives et ralentit l'acquisition de la lecture et de l'écriture, estiment Seymour et ses collaborateurs. Cette importante sollicitation des habiletés cognitives pourrait nuire à l'acquisition des phonèmes multigraphémiques, surtout chez les faibles orthographes.

Dans le chapitre *Foundations of Orthographic Development*, Seymour présente l'apport concomitant de ce qu'il appelle, en anglais, des « processes » qui interviennent dans l'acquisition du langage écrit (Seymour, 1997). Le mot « process » est traduit de diverses manières par les chercheurs francophones. Certains le traduisent par « processeurs » (Morin, 2009; Stanké, 2009), alors que d'autres le traduisent par « procédures » (Daigle et al., 2013; Écalle et Magnan, 2015). Berger et Desrochers (2011) utilisent, pour leur part, le mot « traitement » qu'ils relient au processus spécifique d'acquisition du langage écrit, ce qui est le sujet de la présente thèse. C'est donc cette traduction qui est retenue. Les quatre traitements impliqués dans le processus spécifique d'acquisition du langage écrit, selon Seymour, sont présentés dans cette section. La figure 2.4 les illustre.

**Figure 2.4 Représentation du modèle de Seymour (traduit de Seymour et al., 2003)**



### 2.2.3.1 Traitement alphabétique

Le traitement alphabétique tient une place importante dans tous les modèles développementaux. Il est associé au décodage séquentiel des mots à partir de correspondances graphèmes-phonèmes régulières (Seymour et al., 2003). Il permet d'orthographier des mots consistants phonographémiquement, c'est-à-dire des mots dont chacun des phonèmes ne peuvent correspondre qu'à un seul phonème.

La conscience phonémique, qui permet de manipuler les phonèmes des mots, est essentielle au développement du traitement alphabétique (Seymour, 1997). Cette aptitude à bien segmenter le mot de façon phonémique évite, entre autres, les omissions ou les inversions de graphèmes simples, c'est-à-dire de lettres. C'est ce qui permet d'identifier que les phonèmes du mot « pinceau » sont les suivants, par exemple : /p/, /ɛ/, /s/, /o/. En fin de première année, les élèves devraient pouvoir

lire et orthographier convenablement de manière phonologique (Sprenger-Charolles, 2012). Ainsi, on s'attend à ce qu'ils puissent écrire « pinceau » en utilisant le traitement alphabétique, peu importe les variantes orthographiques choisies. Par exemple, « pinso », « painceau » et « peinessau » correspondent à un traitement alphabétique correct.

Cela dit, chez les élèves faibles orthographieurs, un déficit phonologique est souvent présent et empêche d'accéder au traitement alphabétique. Un déficit phonologique peut s'expliquer par un déficit de l'organisation épiphonologique (Gombert, 2003). L'organisation épiphonologique est liée aux processus cognitifs inconscients qui mènent aux habiletés phonologiques (Écalte et Magnan, 2015). Une rééducation métaphonologique explicite peut alors être nécessaire pour combler le manque de sensibilité phonologique qui mène à ce déficit.

Selon Seymour (1997), l'acquisition des règles orthographiques et morphographémiques ne pourra réellement commencer que lorsque le traitement alphabétique sera acquis. Ainsi, les élèves qui n'arrivent pas bien à identifier la structure phonémique d'un mot à l'oral peineront à accéder aux autres traitements du système orthographique.

### **2.2.3.2 Traitement logographique**

Seymour (1997) considère que le traitement logographique est aussi important que le traitement alphabétique pour l'acquisition du langage écrit en début de la scolarisation. En effet, les deux traitements sont utilisés en alternance selon les expériences qui se présentent aux élèves. Le traitement logographique permet l'identification et le stockage de mots familiers (Seymour et al., 2003). Il permet aux élèves d'acquérir des unités du langage écrit par reconnaissance globale de toutes les lettres ou d'une partie des mots. Les mots sont alors acquis par ce que Demont et Gombert (2004) appellent des logogrammes, c'est-à-dire des unités globales qui se reconnaissent facilement, de la même manière qu'un dessin. Ce traitement prend toute son importance en début de scolarité, permettant aux élèves de stocker des mots spécifiques et fréquents dans leur lexique avant même d'arriver à en comprendre la construction; plus tard, lorsque cette compréhension de la construction de l'orthographe se fera, c'est le processus orthographique qui sera sollicité (Seymour, 1997).

### 2.2.3.3 Traitement orthographique

Le traitement orthographique est l'aspect central du modèle d'acquisition du langage écrit de Seymour (1997). Il a pour fonction d'encoder les caractéristiques orthographiques spécifiques des mots. Ce traitement se développe progressivement et ne s'achève pas avant la fin du primaire (Seymour, 1997). Son acquisition diffère de l'acquisition logographique principalement par le fait qu'un acquis orthographique est transférable alors qu'un acquis logographique ne l'est pas (Seymour, 1997). Par exemple, une fois qu'un élève aura appris que le graphème « g » transcrit le son /g/ devant les voyelles « a », « o » et « u », il pourra transférer cette connaissance pour lire et pour orthographier tous les mots contenant cette structure. Toutefois, si ce même élève s'appelle « Gaston » et qu'il a appris à lire et à écrire son prénom de façon logographique sans enclencher les traitements cognitifs nécessaires à l'acquisition orthographique de la règle du « g » dur, il ne pourra pas transférer cet apprentissage pour les autres mots ayant une structure similaire. Afin de transférer les apprentissages, il importe donc que les phonèmes multigraphémiques soient enseignés et que les mots des listes orthographiques ne soient pas seulement appris par cœur.

Le traitement des structures irrégulières de certains mots et des graphies contextuelles se fait par le traitement orthographique (Sprenger-Charolles et Colé, 2013). On peut penser ici au cas du « m » devant le « b » et le « p », au doublement du « s » entre deux voyelles, au « g » et au « c » durs et doux, mais aussi à la reconnaissance du mot « chorale » qui se prononce exceptionnellement /koral/ et non /foral/ et à la production du mot « paon » qui se prononce /pã/, mais qui s'écrit « paon ». C'est aussi ce qui permet aux élèves de comprendre que, exceptionnellement, le phonème /a/ correspond au graphème /e/ dans le mot « femme ». Enfin, c'est ce qui permet de traiter les inconsistances phonographémiques associées à la multigraphémie.

Si l'on prend l'exemple du mot « certain », les règles orthographiques suivantes doivent être acquises : le graphème « e » est utilisé pour représenter le phonème /ɛ/ lorsqu'il est l'avant-dernière lettre d'une syllabe fermée, c'est-à-dire fermée par une consonne, en l'occurrence « r », et le graphème « ain » est utilisé pour représenter le phonème /ɛ̃/ en fin de mot parce qu'une dérivation en « ai » est possible, c'est-à-dire « certaine » (Stanké et Dumais, 2017b). Cette dernière dérivation est d'ailleurs associée au traitement morphographémique.

#### 2.2.3.4 Traitement morphographémique

Le traitement morphographémique permet d'acquérir l'orthographe des mots contenant des morphèmes fléchis et dérivés. Considérant qu'en français, 80 % des mots contiennent des morphèmes (Gombert, 2003), il est essentiel de sensibiliser les élèves à la structure morphologique des mots. Ce traitement peut notamment faciliter l'acquisition de la lecture et de l'orthographe pour les faibles lecteurs et pour les faibles orthographes (Berthiaume et Daigle, 2014). Toutefois, ce traitement est peu considéré dans les modèles développementaux d'acquisition de la lecture et de l'écriture. D'ailleurs, la plupart des modèles développementaux considèrent que l'acquisition du traitement morphographémique se fait plus tardivement que l'acquisition des autres traitements; cela s'explique entre autres par le fait que les flexions et les dérivations sont la plupart du temps non marquées phonologiquement dans le système orthographique du français, ce qui en accentue la difficulté (Morin et Montésinos-Gelet, 2005; Mousty et Leybaert, 1999). En anglais, toutefois, ces flexions et dérivations s'acquièrent plus facilement, car elles sont marquées phonologiquement. Par exemple, « hasard » se prononce « hazarD » et « immigrant » se prononce « immigranT » en anglais. Le fait que les élèves doivent connaître plusieurs dérivations lexicales pour bien identifier les lettres muettes finales constitue donc une difficulté importante. C'est pour cette raison, et aussi parce que ce traitement sollicite des connaissances complexes comme les racines des mots et les affixes, que Seymour (1997) affirme que les élèves doivent avoir commencé à développer une certaine aisance avec le traitement orthographique avant d'accéder au traitement morphographémique.

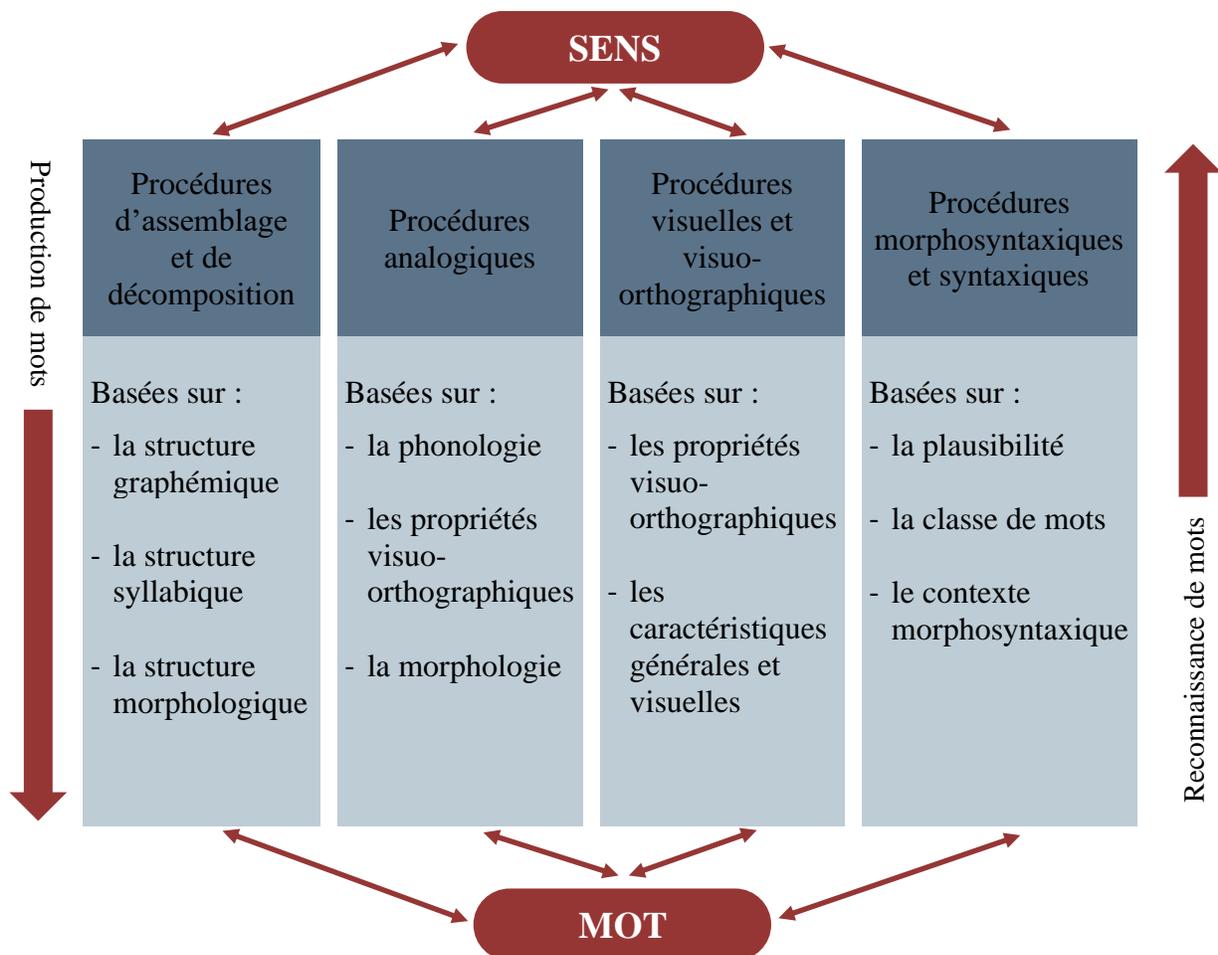
Pourtant, certains chercheurs ont constaté une sensibilité implicite à la morphologie dès le préscolaire. C'est le cas de Gombert (2003), entre autres, qui a constaté que, dès l'âge de 5 ans, les enfants savent reconnaître oralement la plausibilité de certains pseudomots contenant des préfixes réels. Par exemple, il a constaté que des mots inventés contenant le préfixe « pre » (ce qui correspond au préfixe « pré » en français) ont semblé plus réels que des mots inventés contenant le préfixe « pra », et ce, chez 67,25 % des élèves du préscolaire.

C'est, entre autres, pour reconnaître la place importante que tient la morphologie dans chacun des traitements, qui agissent en interaction, que Daigle et ses collaborateurs ont proposé un modèle renouvelé d'acquisition du langage écrit en 2018.

## 2.2.4 Modèle de reconnaissance et de production des mots écrits de Daigle et al.

Le modèle de Daigle et al. (2018) est utile pour l'enseignement et la rééducation de la lecture et de l'orthographe ainsi que pour l'évaluation des difficultés en lecture et en écriture. Il se base sur quatre procédures, qui correspondent au terme « traitements » préalablement utilisé. Comme le modèle de Daigle et ses collaborateurs est présenté en français, l'appellation « procédures » est ici conservée. Les procédures de ce modèle agissent en interaction de façon non hiérarchique. La figure 2.5 présente les microtâches associées aux procédures analogiques, aux procédures de décodage et d'encodage, aux procédures visuelles et visuo-orthographiques ainsi qu'aux procédures morphosyntaxiques et syntaxiques de ce modèle.

**Figure 2.5** Modèle de reconnaissance et de production des mots écrits axé sur les microtâches (Daigle et al., 2018, p. 248; Daigle et Berthiaume, 2021, p. 139)



### 2.2.4.1 Procédures analogiques

Pour reconnaître et pour produire des mots écrits, les élèves doivent faire appel aux procédures analogiques. Ces procédures d'acquisition du langage écrit font partie du modèle de Gombert et al. (1997) et de Daigle et al. (2018), mais elles ne font pas partie de celui de Seymour (1997). Le recours à l'analogie présuppose la présence d'un lexique orthographique. Ce lexique fait appel tant au traitement phonologique, qu'aux traitements orthographique et morphologique chez les élèves, puisqu'ils doivent avoir en mémoire un bagage de mots connus faisant référence à tous ces traitements afin de faire une analogie (Pacton, Foulin et al., 2005). Les élèves utilisent cette procédure pour lire ou pour prononcer un mot connu ou un mot inconnu qui inclut un mot connu. Par exemple, ils peuvent arriver à lire le mot inconnu « choriste » à partir du mot connu « chorale » et à orthographier correctement le phonème /ã/ dans le mot « grand » parce qu'ils connaissent le mot « grandeur ».

L'analogie permet aussi aux élèves de se référer à une forme orthographique connue pour en produire une autre. C'est le cas, par exemple, des élèves qui se réfèrent au mot « château » pour écrire le mot nouveau « râteau ». Ne pouvant s'appuyer sur le lexique orthographique qu'ils ont en mémoire puisqu'ils ne connaissent pas le mot à orthographier, les élèves utilisent la stratégie de l'analogie orthographique (Fayol et Jaffré, 2008), c'est-à-dire qu'ils se réfèrent à une forme orthographique connue pour en produire une autre.

Les analogies que les élèves peuvent faire grâce à leurs connaissances orthographiques jouent rapidement un rôle important dans l'acquisition de l'orthographe de mots nouveaux. En effet, les élèves se réfèrent à la forme connue d'un mot pour en produire un autre dès le début de leur scolarité. Ils arrivent à recourir à cette stratégie de façon très précoce, voire dès le mois de janvier de leur première année du primaire (Martinet et al., 1999) ou le mois d'avril de cette première année (Sprenger-Charolles et al., 1998). Cette stratégie à laquelle les élèves ont recours de façon précoce est toutefois limitée par le nombre insuffisant de mots qui figurent dans leur lexique orthographique (Martinet et al., 1999). Il peut aussi être source d'erreurs : par exemple, les élèves qui voient le mot « science » tous les jours affiché à l'horaire pourraient orthographier le mot nouveau « siècle » de façon erronée (« sciècle »).

Pour faire une analogie, les élèves peuvent se baser sur les propriétés phonologiques des mots. Par exemple, s'ils connaissent le mot « pomme » et s'ils comprennent que le « o » de ce mot est une voyelle ouverte /ɔ/, ils comprendront que les mots de la même famille contiennent aussi un « o » ouvert. Ils pourront alors appliquer une analogie phonologique en orthographiant le graphème « o » dans le mot « pommade » puisque la voyelle ouverte /ɔ/ s'orthographie toujours avec un « o »<sup>5</sup>. Les élèves peuvent aussi faire des analogies basées sur les propriétés visuo-orthographiques des mots. Par exemple, à partir du mot connu « pomme », ils peuvent orthographier le mot « pommier » avec lettres géminées « mm ». Enfin, les élèves peuvent faire des analogies en se basant sur les propriétés morphologiques des mots. Par exemple, ils peuvent baser leur réflexion sur l'analogie du mot connu « dentiste » afin de comprendre que le mot « dent » se termine avec le morphogramme « t ».

#### **2.2.4.2 Procédures d'assemblage et de décomposition**

Pour lire et pour orthographier les mots, les élèves font aussi appel aux procédures d'encodage et de décodage, qui sont basées sur les structures graphémique, syllabique, subsyllabique et morphémique. Tant pour le décodage, qui permet de lire les mots, que pour l'encodage, qui permet de les orthographier, il importe de segmenter les mots en unités significatives, c'est-à-dire en syllabes et en phonèmes. Les procédures de décodage et d'encodage du modèle de Daigle et al. (2018) correspondent donc, en partie, au traitement alphabétique de Seymour (1997).

Pour Daigle et al., les morphèmes jouent, ici aussi, un rôle important. En effet, ce sont autant les habiletés morphologiques que phonologiques qui permettent de traiter les lettres muettes. Lorsque les élèves lisent, l'assemblage des unités permet de prononcer le mot, à voix haute ou silencieusement. Lorsqu'ils orthographient, les élèves associent chaque phonème à une séquence orthographique et produisent cette séquence à l'écrit. Grâce à ces habiletés, les élèves comprendront que, si plusieurs chatons mangent, les lettres « nt » doivent être ajoutées au verbe, « mangent », et ce, même si elles ne se prononcent pas. Le même lien entre les habiletés morphologiques et phonologiques devra être fait pour orthographier le mot « bout », dont la lettre finale reste muette même lors du décodage et s'explique par la dérivation vers le mot « couture ».

---

<sup>5</sup> Il est à noter que cette stratégie fonctionne avec la prononciation québécoise seulement.

### 2.2.4.3 Procédures visuelles et visuo-orthographiques

Lorsque le traitement phonologique ou le traitement morphologique est impossible pour lire ou pour orthographier, les élèves ont recours à l'aspect général des mots à partir de leurs propriétés visuelles en lecture et visuo-orthographiques en production de mots (Daigle et al., 2018). Ils se basent alors sur le nombre de lettres, sur l'ordre des lettres ainsi que sur des caractéristiques visuo-orthographiques spécifiques pour traiter les mots et pour les garder en mémoire. C'est de cette façon qu'ils traiteront les homophones lexicaux, les lettres muettes non porteuses de sens, les lettres géminées ou non ainsi que les cas de multigraphémie.

Les propriétés visuelles peuvent être lexicales, c'est-à-dire qu'elles peuvent se rapporter au mot dans son ensemble. Elles peuvent aussi être supralexicales et dépasser alors les frontières lexicales des mots (Daigle et al., 2013). Pensons par exemple aux élèves qui saisissent mal la frontière du mot « oiseau » et qui écrivent « loiseau ». C'est surtout la dimension visuogrammique sublexicale, ou sous-lexicale, qui est utilisée en orthographe lexicale puisqu'elle permet de s'attarder aux particularités orthographiques qui sont situées à l'intérieur du mot (Daigle et al., 2013). En effet, elle concerne une partie du mot et non le mot dans son ensemble. Selon Daigle et al. (2013), les règles de positionnement, les lettres muettes non porteuses de sens, la légalité orthographique, l'irrégularité orthographique et la multigraphémie sont des composantes de la dimension visuogrammique sublexicale, ou sous-lexicale, du code orthographique.

Un processus visuel doit être fait par les élèves afin d'acquérir l'orthographe des mots contenant des règles de positionnement (Daigle et al., 2013), qui sont aussi appelées régularités contextuelles. En effet, il est nécessaire de porter une attention visuelle aux lettres qui varient en fonction de leur position. On y retrouve les graphèmes consistants contextuels, c'est-à-dire ceux qui ne varient pas selon le contexte, comme le « m » devant le « b » ou le « p ». On y retrouve aussi les graphèmes inconsistants contextuels, c'est-à-dire ceux qui varient selon le contexte, comme le double « s » entre deux voyelles. Par exemple, dans le mot « bambou », il faut employer le graphème « am » et non le graphème « an » parce que la voyelle nasale /ã/ est suivie d'un « b » et il faut doubler le « s » pour faire le son /s/ entre deux voyelles dans le mot « poiSSon ».

Les mots contenant des « lettres muettes non porteuses de sens » doivent aussi être stockés dans le lexique orthographique des élèves grâce aux procédés visuels (Daigle et al., 2013). Catach classe plusieurs de ces cas dans la dimension étymologique et précise que les erreurs qui y sont associées sont à dominance non fonctionnelle (Catach, 2011) puisque les lettres ne sont pas porteuses de sens. Par exemple, certains mots comme « habit » contiennent des lettres muettes qui ne sont pas porteuses de sens puisqu'aucune dérivation n'explique l'emploi de la lettre t, « habit » ne venant pas du mot « habiter ». La présence de ces lettres dans les mots est à l'origine de l'une des deux plus grandes sources d'erreurs en orthographe lexicale selon Daigle et ses collègues (2013).

La légalité orthographique se définit comme ce qui est possible en orthographe lexicale et se rapporte, entre autres, aux doubles consonnes, qui ne peuvent se doubler qu'en milieu de mot. C'est ce que Pacton nomme « connaissances générales de l'orthographe » (Pacton et al., 2001). Ces connaissances sont liées aux régularités graphotactiques (Pacton et Afonso-Jaco, 2015). C'est, en effet, en se représentant visuellement les mots à l'écrit que les élèves parviennent à juger de leur légalité orthographique. Par exemple, les mots suivants relèvent d'illégalités orthographiques : « ccatalogue » et « fleur ».

L'irrégularité orthographique se présente lorsqu'un mot contient des séquences de lettres qui ne sont pas typiques; ces séquences de lettres inattendues, par exemple des doubles voyelles, peuvent entre autres s'expliquer par le fait qu'un mot a été emprunté à une autre langue (Daigle et al., 2013). Nous pouvons, par exemple, penser au mot « zoo » qui se termine par deux voyelles. L'irrégularité orthographique est aussi présente lorsque le graphème utilisé ne correspond pas au phonème entendu (Seymour, 1997). Nous pouvons penser, bien sûr, aux mots « femme » et « monsieur », qui comportent des séquences graphémiques ne correspondant pas aux phonèmes entendus; mais aussi aux mots « second » dont le graphème « c » se prononce /g/ et non /k/; au mot « orchidée » dont le graphème « ch » se prononce /k/ et non /ʃ/; au mot « aquarium » dont le graphème « qu » se prononce /kw/ et non /k/; au mot « aiguille » dont le graphème « gu » se prononce /gw/ et non /g/; ou au mot « examen » dont le graphème « en » se prononce /ɛ̃/ et non /ã/. Les lettres muettes finales exceptionnellement non prononcées peuvent aussi être considérées comme des irrégularités orthographiques par certains chercheurs (Boutin, 2016; Jacquier-Roux et al., 2010; Lenfant et al., 2009). Par exemple, on peut penser au « g » qui n'est pas prononcé dans le mot « bourg »; au « c » qui n'est pas prononcé dans les mots « caoutchouc », « tabac » et « tronc »; au « m » qui n'est pas

prononcé dans « parfum »; au « p » qui n'est pas prononcé dans « galop »; et au « f » qui n'est pas prononcé dans « nerf ». Les irrégularités orthographiques sont donc une source d'erreurs importantes. Selon Daigle et ses collègues (2013), le fait d'encoder la partie irrégulière du mot de façon visuelle peut améliorer la compétence orthographique.

#### **2.2.4.4 Procédures morphosyntaxiques et syntaxiques**

Enfin, pour lire et pour orthographier, les élèves ont recours aux procédures morphosyntaxique et syntaxique. Ils se basent alors sur la plausibilité lexicale, sur la classe du mot et sur le contexte morphosyntaxique (Daigle et al., 2018). Ainsi, à partir du contexte entourant le mot et de l'analyse syntaxique, les élèves peuvent savoir si le mot est plausible sémantiquement. Par exemple, à partir du contexte, ils sauront que Mathieu et Olivier grimpent dans les « pommiers » et non dans les « pommes » ou les « pomiculteurs ».

Le contexte et les informations morphologiques permettent aussi aux élèves de savoir si le mot est acceptable grammaticalement. Ils se basent alors sur les propriétés grammaticales du mot. Par exemple, ils accordent le verbe avec le sujet et appliquent la conjugaison appropriée selon le verbe dans les phrases suivantes : « Félix, Louis et Thomas joueNT à cache-cache dans le verger » et « David finiT de manger sa pomme ».

Bref, les procédures analogiques, d'assemblage et de décomposition, ainsi que les procédures visuelles, visuo-orthographiques, morphosyntaxiques et syntaxiques agissent en interrelation et accordent toutes une place importante à la morphologie. Certaines de ces procédures sont faites de façon implicite, mais elles peuvent aussi être faites de façon explicite, par exemple en écrivant le mot et en se demandant s'il semble bien orthographié ou non dans le cas d'une analogie (Sprenger-Charolles et al., 1997).

Ainsi, cette section a présenté un modèle de référence pour l'enseignement et pour l'évaluation de la lecture et de l'orthographe qui tient compte du développement des connaissances orthographiques par apprentissage explicite, soit le modèle de Seymour. Elle a aussi présenté des modèles qui tiennent compte de l'apprentissage explicite et implicite, soit les modèles de Daigle et al. et de Gombert. Ces deux derniers modèles tiennent aussi compte d'un autre type de traitement dont ne tient pas compte Seymour, soit le traitement analogique. Le traitement analogique sollicite

l'attention visuelle, mais aussi la mémoire. En effet, il est difficile pour les élèves francophones d'acquérir l'orthographe lexicale et ses phonèmes multigraphémiques sans une intervention importante de la mémoire, de la vue et de l'audition. Ces facteurs individuels importants sont présentés dans la prochaine section.

### **2.3 Facteurs individuels intervenant dans l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques**

Dans un système orthographique opaque contenant autant de phonèmes multigraphémiques que le français, les élèves doivent développer diverses stratégies cognitives pour acquérir un lexique orthographique suffisant. D'ailleurs, l'opacité de l'orthographe est un facteur aggravant pour les faibles orthographes : lorsque des élèves ont un trouble d'acquisition de la lecture et de l'écriture (dyslexie/dysorthographe), le degré de sévérité de leur dyslexie est directement lié au degré d'opacité du système orthographique (Sprenger-Charolles et Colé, 2013). Daigle et Berthiaume (2021) présentent neuf processus cognitifs qui contribuent à la lecture et à l'écriture : la mémoire, la vue, l'audition, l'attention, la planification, l'inhibition, le contrôle, la flexibilité mentale et les stratégies. Cette section présente des facteurs individuels qui sont liés aux trois premiers de ces facteurs, c'est-à-dire la mémoire, les capacités langagières et la conscience phonologique ainsi que l'empan visuoattentionnel.

#### **2.3.1 Mémoire**

La structure cognitive liée à l'acquisition du langage est complexe (Chomsky, 2006). Fayol et Jaffré (2008) identifient quatre stratégies pour orthographier : les stratégies phonographique, morphologique dérivationnelle et analogique orthographique ainsi que la récupération directe en mémoire des formes lexicales.

La mémoire intervient lorsque les élèves récupèrent directement des formes orthographiques stockées. C'est la stratégie des scripteurs experts, qui ont réussi à mémoriser un nombre important de mots et à y avoir recours rapidement, et c'est la stratégie que doivent utiliser les élèves débutants qui ne peuvent s'appuyer sur les correspondances phonèmes-graphèmes pour acquérir un mot irrégulier fréquent, dont les lettres ne correspondent pas aux sons.

Plusieurs chercheurs ont identifié le rôle important de la mémoire dans l'acquisition de l'orthographe. Stanké (2009) a montré l'apport de la mémoire lexicale orthographique pour l'acquisition de l'orthographe chez les enfants d'âge préscolaire. Fayol (2011) a aussi identifié le rôle central que jouent les différents systèmes mnésiques dans l'acquisition de l'orthographe et en a présenté quatre catégories : la mémoire de travail, la mémoire épisodique, la mémoire sémantique et la mémoire procédurale. La première catégorie est liée à la mémoire à court terme alors que les trois autres sont liées à la mémoire à long terme (Lussier et al., 2017).

La mémoire de travail permet de retenir l'information à court terme et de la manipuler. Par exemple, elle permet de traiter les informations de façon cognitive complexe pour acquérir l'orthographe lexicale (Fayol, 2011). Au début de l'apprentissage, elle est essentielle pour lire et pour orthographier les mots de façon phonologique. Chez les élèves faibles orthographes, un déficit de mémoire de travail peut notamment résulter en des difficultés phonologiques (Stanké et al., 2016).

La mémoire épisodique permet aux élèves d'encoder des informations à partir de leur vie personnelle. De façon générale, elle peut être associée à une émotion positive ressentie avec un enseignant (Fayol, 2011). En orthographe, elle pourrait probablement être associée à la rétention de mots significatifs pour les élèves. On n'a qu'à penser aux jeunes scripteurs qui retiennent rapidement les structures orthographiques des mots présentant des syllabes communes à celles de leur nom ou qui retiennent rapidement l'orthographe du mot « maman ».

La mémoire sémantique permet l'acquisition des connaissances sur le monde, des connaissances scolaires (Fayol, 2011) et celle du sens des mots (Daigle et al., 2013). Elle est associée au répertoire langagier, à un niveau phonologique, lexical ou conceptuel (Lussier et al., 2017). En orthographe, cette mémoire pourrait les aider à catégoriser les mots, à distinguer les homophones lexicaux et syntaxiques et à prévoir des séquences syntaxiques, par exemple.

Pour sa part, la mémoire procédurale est associée au savoir-faire et aux traitements automatisés et est sollicitée lorsque des gestes complexes doivent être faits (Lussier et al., 2017). Elle est donc utile à l'écriture des lettres, à la production des mots connus et à la gestion des accords (Fayol, 2011).

Les élèves faibles orthographes peinent à récupérer des formes orthographiques stockées, car cette stratégie, basée sur le facteur cognitif de la mémoire, fait défaut (Scarborough, 1998). Pour acquérir l'orthographe lexicale, les élèves font implicitement des statistiques mentales, qui pourraient parfois compenser pour des problèmes de mémoire (Treiman et Kessler, 2014). Par exemple, le contact avec les mots pourrait leur permettre d'acquérir le fait que le phonème /f/ est représenté beaucoup plus souvent par le graphème « f » que par le graphème « ph ». Toutefois, bon nombre d'observations sont nécessaires pour y arriver et, comme les faibles orthographes sont très souvent de faibles lecteurs, ce traitement statistique ne peut être enclenché. Ainsi, à cause des défis liés à la mémoire et au traitement statistique, les faibles orthographes pourraient ne pas autant bénéficier de la fréquence des mots que les bons orthographes pour acquérir les phonèmes multigraphémiques.

### **2.3.2 Capacités langagières et conscience phonologique**

L'audition et les capacités langagières permettent aux élèves de développer leur conscience phonologique, qui est essentielle pour orthographier correctement les mots et les phonèmes multigraphémiques qu'ils contiennent. La conscience phonologique se définit comme la capacité à identifier les unités sonores d'un mot (Fayol et Jaffré, 2008), c'est-à-dire les phonèmes, les syllabes, l'attaque et la rime (Sprenger-Charolles et Colé, 2013) et à les manipuler intentionnellement (Stanké et Flessas, 2013).

La connaissance des lettres et les capacités de conscience phonologique ont été identifiées comme les meilleurs prédicteurs de compétence en lecture par la méta-analyse de Scarborough (1998), mais aussi comme un prédicteur important des compétences en orthographe lexicale (Stanké et Flessas, 2013). Selon Hawken (2009), la correspondance entre les phonèmes et les graphèmes consistants devrait être acquise à la fin de la première année du primaire. Ces connaissances devraient permettre aux élèves d'orthographier les mots de façon phonologique.

Toutefois, en raison de faibles capacités de conscience phonologique, certains élèves ne parviennent pas à identifier adéquatement tous les phonèmes des mots et n'arrivent pas à faire correspondre un graphème phonologiquement plausible à un graphème pour représenter un mot à l'écrit. De faibles capacités de conscience phonologique entraînent des erreurs dites d'orthographe

phonologique. Certains phonèmes peuvent être assimilés à d'autres (/t/ – /d/) ou intervertis (/sp/-/ps/). La sonorité (/f/ – /v/) ou la nasalité (/an/ – /a/) peuvent aussi ne pas être perçues. Ces erreurs risquent de s'observer plus longtemps chez les faibles orthographes, que ce soit parce ces apprenants ont une difficulté transitoire, comme ceux qui apprennent une langue seconde, ou parce qu'ils ont des difficultés persistantes ou un trouble d'apprentissage (Fayol et Jaffré, 2008).

### **2.3.3 Empan visuoattentionnel**

Bien que la dyslexie/dysorthographe ait souvent uniquement été associée à un déficit phonologique, des chercheurs associent depuis peu la dyslexie/dysorthographe à un autre facteur cognitif : l'empan visuoattentionnel (Bosse et al., 2009; Valdois, 2014; Valdois et al., 2017). L'empan visuoattentionnel est associé au nombre d'éléments visuels distincts que l'œil peut traiter simultanément, en une seule fixation, à travers plusieurs éléments (Valdois, 2014). En langage écrit, cela correspond au nombre de lettres que l'œil peut traiter à la fois.

Valdois et al. (2017) précisent que ce sont les capacités visuo-attentionnelles qui permettent de mémoriser les graphèmes et les syllabes qui composent les mots ainsi que les séquences des lettres dans les mots (Valdois et al., 2017). Ces capacités, ajoutent les chercheurs, augmentent avec les années, l'élève ne pouvant traiter qu'une lettre à la fois en début d'apprentissage. Chez certains élèves, les capacités visuo-attentionnelles sont faibles. Ces derniers ne peuvent alors pas saisir plusieurs lettres à la fois, ce qui nuit à leur capacité de saisir les règles contextuelles, qui font varier certains graphèmes selon les lettres qui se trouvent à proximité (par exemple, le « s » ou le « ss » ou le « m » devant le « p » et le « b »). Ces élèves peinent aussi à mémoriser des mots entiers.

Certes, les facteurs individuels jouent un rôle important dans l'acquisition de l'orthographe lexicale, mais plusieurs facteurs linguistiques ont aussi un rôle important à jouer. Ils sont présentés dans la prochaine section.

## **2.4 Facteurs linguistiques intervenant dans l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques**

L'acquisition de l'orthographe lexicale est influencée par plusieurs facteurs linguistiques. Ces facteurs peuvent être lexicaux ou sous-lexicaux. Cette section les présente donc en ces deux sous-catégories.

## **2.4.1 Facteurs lexicaux**

Certains facteurs lexicaux peuvent complexifier l'acquisition de l'orthographe lexicale et de ses phonèmes multigraphémiques. Ces facteurs sont associés aux mots en entier. Les deux facteurs suivants sont approfondis dans la présente section : la fréquence des mots ainsi que leur longueur. Des recherches portant sur chacun des sujets ainsi que leurs limites sont présentées.

### **2.4.1.1 Fréquence du mot**

Tout d'abord, le lexique orthographique que les élèves développent est influencé par l'effet de la fréquence (Fayol et Jaffré, 2008; Hazard, 2009; Martinet et al., 2004; Mousty et Leybaert, 1999; Pacton, 2018; Sprenger-Charolles et al., 1998; Stanké et al., 2019). En effet, alors que les mots consistants peuvent être orthographiés de manière phonographémique, les élèves doivent avoir recours à d'autres stratégies pour orthographier les mots inconsistants. Bien qu'elle soit insuffisante, la fréquence à laquelle ils ont vu les mots influence l'acquisition de ceux-ci puisqu'elle en favorise la mémorisation. L'effet de la fréquence du mot exercerait une influence sur l'acquisition de l'orthographe lexicale dès la première année du primaire (Lété et al., 2008; Martinet et al., 2004).

La recherche a montré que les mots les plus fréquents sont acquis plus rapidement; la mémorisation de l'orthographe de certains mots étant étroitement liée à la fréquence à laquelle ces mots sont vus (Fayol et Jaffré, 2008). Ainsi, plus les élèves voient un mot souvent, plus vite ils arrivent à bien le lire et à bien l'orthographier. En effet, l'acquisition d'un nouveau savoir passe par l'exposition fréquente des élèves à ce savoir (Chartrand et al., 2016).

Pour mieux comprendre l'effet de la fréquence des mots, Martinet et ses collègues (2004) ont étudié la performance orthographique d'élèves francophones de première année du primaire. Les participants devaient orthographier des mots fréquents et peu fréquents ainsi que des pseudomots qui avaient des voisins lexicaux. Les pseudomots sont des mots inventés respectant les structures et les règles orthographiques de véritables mots. Les chercheurs ont observé un effet de la fréquence sur l'apprentissage de l'orthographe des mots et un effet d'analogie dans l'orthographe des pseudomots après seulement trois mois. Les résultats ont aussi montré que les élèves acquièrent des connaissances orthographiques dès le début de leur scolarisation.

Sprenger-Charolles et ses collègues (1998) ont aussi précisé l'importance de la fréquence des mots dans l'acquisition de l'orthographe lexicale, mais, cette fois, pour les mots irréguliers. L'objectif de leur recherche était d'étudier le développement de la lecture et de l'orthographe en français chez des élèves de première année du primaire. Ils ont remarqué que, pour les mots réguliers, fréquents ou non, les élèves s'en remettaient à leurs connaissances phonologiques de base. Toutefois, pour les mots irréguliers, ils arrivaient à établir un lexique mental selon la fréquence des mots.

En 2017, une autre étude de Sprenger-Charolles s'est intéressée à l'effet qu'a la fréquence des mots, mais, dans ce cas-ci, pour acquérir la lecture. La chercheuse a présenté une progression pédagogique de la lecture, construite à partir de statistiques sur l'orthographe du français d'après Manulex-Morpho (Sprenger-Charolles, 2017). À partir, entre autres, de la fréquence, de la consistance et de la régularité des correspondances graphèmes-phonèmes dans les livres, Sprenger-Charolles a établi une échelle de fréquence et a dressé une liste hiérarchisée des phonèmes qui devraient être enseignés en lecture de la première à la cinquième année du primaire. L'auteure souligne toutefois n'avoir tenu compte que de la fréquence d'apparition des mots dans le corpus de textes analysés pour établir cette liste, ce qui constitue une limite. Cela dit, la recherche Sprenger-Charolles, bien qu'elle s'intéresse à la lecture, se rapproche de l'intérêt de la présente recherche doctorale, entre autres, parce que la chercheuse analyse certains facteurs linguistiques qui ont un effet sur l'acquisition des mots et qu'elle s'intéresse à la constitution d'un lexique mental, aussi important pour l'acquisition de la lecture que de l'orthographe lexicale.

Le tableau 2.4 résume ce que Sprenger-Charolles (2017) propose comme progression pédagogique pour acquérir la lecture. Selon elle, les correspondances graphèmes-phonèmes les plus fréquentes et les plus consistantes doivent être enseignées en premier.

**Tableau 2.4 Progression pédagogique de la lecture (inspiré de Sprenger-Charolles, 2017)**

Difficulté	Caractéristiques
1	Voyelles orales, consonnes continues et « s » en début de mot, mots monosyllabiques, structure consonne (C) - voyelle (V) : CV et quelques CVC (ex : ou, le, mou)
2	Consonnes p m d v et, ensuite, b n ch et t ; structures syllabiques complexes CCV et CCVC (ex : plu, grade) ; mots très fréquents qui contiennent des correspondances phonèmes-graphèmes irrégulières (ex : elle, les), graphèmes muets fréquents de fin de mots (dos, lit)
3	Voyelles nasales (ex : an, on, in, un), graphie du « qu »
4	Confusions phonologiques et visuelles (ex. : b, p, d, q)
5	Semi-voyelles /w/ et /ɥ/, consonne muette h et ph
6	Règles contextuelles : nasales (ex. : mb, mp) et autres (ex. : s/ss, c/g durs doux)
7	Allographes (multiples graphies) des voyelles orales (ex. : multigraphémie du phonème /o/)
8	Semi-voyelle /j/ (yod)
9	Pièges (ex. : mots irréguliers qui ne se lisent pas comme ils s'écrivent (femme), le x en milieu et en fin de mot ; le tion, les mots se terminant par s, t et r qui se prononcent ; le ch qui se prononce /k/)

Bien que Sprenger-Charolles (2017) ne tienne pas compte des statistiques de réussite des phonèmes choisis en lecture, les variables qu'elle a choisies pour bâtir cette liste hiérarchisée basée sur la correspondance graphèmes-phonèmes en lecture permettent de mieux comprendre les facteurs qui pourraient avoir un effet sur l'acquisition des phonèmes multigraphémiques. La consistance des correspondances graphèmes-phonèmes sur laquelle elle se base se traduit en consistance phonographémique en orthographe. Ce facteur, qui pourrait jouer un rôle important dans l'acquisition des phonèmes multigraphémiques des élèves, sera présenté dans la section « facteurs sous-lexicaux ».

### **2.4.1.2 Longueur du mot**

La mémoire est souvent sollicitée lorsque les élèves orthographient. Daigle et al. (2013) ont montré que la longueur du mot a un impact sur la mémoire puisque les élèves doivent, entre autres, encoder les structures graphémiques et l'ordre des lettres des mots dans leur lexique orthographique. Pour leur part, lors de la création de la base de données ÉQOL, Stanké et al. (2020) ont conclu que le nombre réduit de syllabes du mot était directement lié à sa réussite orthographique.

Bien qu'elle se soit intéressée à la longueur des graphèmes plus qu'à la longueur des mots, la recherche de Kandel et Spinelli (Kandel et Spinelli, 2010) a aussi apporté certaines réponses à ce sujet. Les chercheurs se sont intéressés à l'impact que la complexité graphémique peut avoir auprès d'adultes experts. Pour ce faire, ils se sont intéressés au temps de traitement. Les participants étaient 36 adultes droitiers de langue maternelle française (21 femmes et 15 hommes) de 20 à 26 ans. Ils ignoraient le but de l'expérience. Les adultes devaient copier en lettres majuscules trois mots contenant des graphèmes simples, des digrammes et des trigrammes et dont une même lettre occupe la même position (par exemple, clAvier, prAirie, plAinte). Les participants devaient écrire le mot qu'ils voyaient à l'écran de l'ordinateur sur le numériseur avec un stylo sur un papier ligné collé au numériseur. Les données concernant les durées de trait des lettres ont été soumises à une ANOVA (analyse par participants [F1] et par éléments [F2]), par position de la lettre (lettre précédant le graphème cible, première lettre du graphème cible) et par longueur du graphème. L'analyse a révélé que le temps de traitement est directement proportionnel à la longueur des graphèmes. Ainsi, les mots contenant des graphèmes plus complexes sont plus longs à écrire. Par exemple, les élèves devraient mettre plus de temps à écrire le mot « trAIN » que le mot « matIN ». Ce temps associé au traitement des correspondances phonèmes-graphèmes plus longs risque fort de créer une surcharge cognitive susceptible d'interférer avec l'acquisition des phonèmes multigraphémiques et de mener à l'erreur, notamment chez les élèves faibles orthographes. La longueur du mot jouerait, en effet, un rôle dans l'acquisition de l'orthographe lexicale (Lété et al., 2008).

### **2.4.2 Facteurs sous-lexicaux intervenant dans l'acquisition de l'orthographe lexicale**

Tout autant que les facteurs lexicaux, les facteurs sous-lexicaux complexifient l'acquisition de l'orthographe. Ils sont associés à des parties de mots, par exemple à des phonèmes

multigraphémiques. L'une des réactions des élèves qui se butent à la complexité d'un système orthographique inconsistant est de développer une sensibilité aux régularités graphotactiques (Borchardt et al., 2012; Pacton et al., 2005; Pacton et al., 2014; Pacton et Casalis, 2006; Sénéchal, 2000) et aux régularités morphologiques (Pacton et Afonso-Jaco, 2015). Ce sont ces facteurs sous-lexicaux qui sont abordés dans la présente section. La consistance phonographémique est aussi un facteur sous-lexical qui joue un rôle important. Elle est d'abord présentée.

#### **2.4.2.1 Consistance phonographémique**

La consistance phonographémique, qui est associée au degré de certitude face à l'orthographe, a aussi un effet sur l'acquisition de l'orthographe des mots. Pérez (2014b) associe la consistance au fait qu'un phonème ne dispose que d'une possibilité de transcription. Hazard (2009) donne l'exemple de la rime /yb/ qui s'écrit toujours « ube » et qui est donc un exemple de consistance phonèmes-graphèmes. Par exemple, on écrira « tube » et « cube ». Au contraire, les mots qui ont un faible pourcentage de correspondances entre les phonèmes et les graphèmes sont des mots dits inconsistants en orthographe. Par exemple, le mot "manteau" est inconsistant. En effet, il est possible de noter deux éléments de consistance sur quatre, puisque le phonème /m / ne peut correspondre qu'au graphème « m » et que le phonème /t / ne peut correspondre qu'au graphème « t », mais que le phonème /ã/ peut correspondre aux graphèmes « an » et « en » et le phonème /o/ peut, entre autres, correspondre aux graphèmes « o », « au » et « eau ».

La consistance peut se calculer à partir du nombre de mots dans lesquels un phonème est représenté par le même graphème par rapport au total de mots comprenant ce phonème; le rapport de consistance qui en résulte se situant entre 0 et 1 (Lété et al., 2008). Selon Stanké et ses collaborateurs (2018), plus un mot comporte de phonèmes qui peuvent correspondre à des graphèmes différents, plus il est inconsistant. Ils considèrent qu'un mot est inconsistant s'il a une valeur inférieure à 60 %.

Seymour et ses collègues (2003) ont étudié l'effet de la consistance des correspondances graphèmes-phonèmes sur les compétences des élèves en lecture. Les résultats ont mis en évidence que la reconnaissance des mots et le décodage se produisent plus lentement dans les langues opaques et inconsistantes et que ce phénomène est particulièrement élevé en anglais, langue dans

laquelle les lecteurs mettent environ deux ans et demi à acquérir des compétences en lecture alors que les lecteurs de langues transparentes y arrivent la première année.

L'effet de la transparence de la langue sur les compétences en lecture a maintes fois été prouvé. Toutefois, la consistance phonographémique en orthographe a été beaucoup moins étudiée que la consistance graphophonémique en lecture (Bonin et al., 2008). La preuve avait déjà été faite que la lecture était acquise plus rapidement dans les langues transparentes : les élèves finnois arrivant à lire 98 mots en fin de première année du primaire alors que les francophones arrivent à en lire 79 et les anglophones 34 au même moment (Seymour et al., 2003). Caravolas (2019) a mené une étude permettant de mieux comprendre l'effet de la transparence de la langue sur l'orthographe à l'université de Bangor au Royaume-Uni. Elle a analysé les effets que peut avoir l'inconsistance orthographique sur le développement des habiletés en lecture et en orthographe dans les langues européennes. Les participants (970 enfants de 5 pays) ont été suivis aux six mois de la maternelle à la deuxième année du primaire. L'étude a conclu que l'acquisition du code orthographique pour une langue non transparente, c'est-à-dire qui a une faible consistance phonographémique, comme le français ou l'anglais, serait fragilisée, surtout en début de scolarité. Les recherches de Caravolas ont donc montré que le degré de consistance d'une langue joue non seulement un rôle important dans l'acquisition de sa lecture, mais aussi dans l'acquisition de son orthographe. Il est donc fort probable que la consistance des mots et des graphèmes joue un rôle dans l'acquisition des phonèmes multigraphémiques dans les productions orthographiques des élèves québécois.

Pothier et Pothier (2004) se sont, entre autres, intéressés à l'effet de la consistance phonographémique lors de la création de l'Échelle d'acquisition en orthographe lexicale (ÉOLE), conçue en France pour hiérarchiser le lexique orthographique à apprendre. Après avoir testé des mots auprès d'élèves de la première à la cinquième année du primaire, ils ont conclu que, plus les élèves connaissent de façons d'orthographier un phonème, plus le risque d'erreurs est grand, car ils ont tendance à complexifier la graphie. Par exemple, lorsque les élèves apprennent que le phonème /o/ peut correspondre au graphème plus rare « ô », ils risquent de complexifier la graphie de nouveaux mots en les orthographiant avec ce graphème. Pothier et Pothier ont aussi fait plusieurs autres constats lors de la création d'ÉOLE. Tout d'abord, ils ont remarqué que l'aspect sémantique ne semble pas avoir d'effet sur la capacité à orthographier un mot. En effet, selon leur recherche, connaître le sens d'un mot ne signifie pas nécessairement qu'on l'orthographiera mieux. Les

chercheurs ont aussi observé que, dès qu'une semi-voyelle apparaît, il y a un risque accru d'erreurs, car l'épellation par segmentation du mot est impossible. De plus, ils ont constaté que la lecture ne permettrait pas de mémoriser l'orthographe des mots. Enfin, ils ont conclu que le phonème /ã/ ne pourrait pas être maîtrisé par les enfants du primaire. Ils expliquent cette incapacité à acquérir le phonème multigraphémique /ã/ par le fait que les graphèmes « an » et « en » sont utilisés presque au même pourcentage, soit respectivement 49 % et 51 % et que la graphie du digramme « an » est plus répandue dans l'environnement des plus jeunes, peut-être à cause du mot « maman », soulignant que les mots ayant une forte signification émotive pourraient donc être acquis plus rapidement. Bien que cette recherche arrive à des conclusions importantes en orthographe lexicale, elle ne permet toutefois pas de répondre à la question de recherche de la présente thèse, d'abord parce qu'elle ne s'attarde pas spécifiquement à l'acquisition des inconsistances phonographémiques et, ensuite, parce que les élèves impliqués sont français et non québécois.

La consistance phonographémique a servi de base à l'élaboration de la Batterie d'évaluation du langage écrit (BELEC), outil conçu en Belgique pour diagnostiquer les troubles spécifiques de l'apprentissage (Mousty et al., 1994). Pour créer la BELEC, Mousty et Leybaert (1999) expliquent avoir mené une étude longitudinale afin d'analyser la dépendance des graphèmes à leur contexte d'utilisation par des tâches d'orthographe de dictées de mots et de pseudomots à deux périodes d'une année scolaire, soit en janvier et en juin. Ils ont analysé la maîtrise de quatre catégories de graphies : graphies consistantes acontextuelles, graphies consistantes contextuelles, graphies inconsistantes contextuelles et graphies dérivables par la morphologie. L'étude s'est déroulée auprès d'élèves de la deuxième année et de la quatrième année du primaire en Belgique. Les chercheurs ont rendu compte des phases d'acquisition de l'orthographe lexicale qui sont ressorties de cette étude. En effet, ils sont arrivés à la conclusion que les élèves maîtrisent d'abord les graphies consistantes acontextuelles, c'est-à-dire qui sont stables et qui ne dépendent pas du contexte (le phonème /ɔ̃/ correspond toujours au graphème /on/, par exemple) : leur réussite est de 89,9 % dès la deuxième année. Les graphies consistantes contextuelles (le graphème /n/ se transforme toujours en graphème /m/ devant le graphème /b/, par exemple) sont, pour leur part, acquises beaucoup plus tardivement : les élèves atteignent 72,9 % de réussite seulement en quatrième année. Pour les graphies inconsistantes contextuelles (le phonème /s/ pouvant correspondre au graphème /t/ dans les mots en « tion », par exemple), l'amélioration des performances est plus grande quand ces

graphies font partie de mots fréquents. Ainsi, selon Mousty et Leybaert (1999), la compétence des élèves à orthographier une graphie contextuelle n'est pas que relative à l'application d'une règle; elle dépend aussi du nombre de fois où les élèves ont rencontré des mots qui contiennent la graphie contextuelle en question. Si le lexique orthographique des élèves n'est pas suffisamment riche, Mousty et Leybaert (1999, p. 334) affirment que les élèves « appliquer [ont] la règle dominante de façon quasi systématique ». Ainsi, ils pourraient écrire « citron » plutôt que « citron », puisque la graphie « si » est plus fréquente que la graphie « ci ». Mousty et Leybaert ont, en effet, observé que les graphies dominantes sont mieux orthographiées que les graphies minoritaires, observant un effet important de la fréquence d'exposition aux mots. Enfin, ils ont conclu que les stratégies morphologiques se développent, pour leur part, plus tardivement, les élèves de quatrième année n'utilisant toujours pas les liens morphologiques d'une façon explicite pour écrire les mots. Bien qu'elle soit très aidante pour classer les différentes graphies et qu'elle permette de dresser des parallèles importants avec notre objet d'étude, cette recherche ne permet pas, non plus, de répondre à la question de la présente thèse qui s'interroge sur l'acquisition des inconsistances phonographémiques et non de l'orthographe lexicale en général, chez les élèves québécois et non chez les élèves belges.

Pérez (2014b) a aussi étudié l'effet de la consistance sur l'acquisition de l'orthographe des mots. En fait, il s'est questionné sur les variables à considérer pour définir la complexité orthographique des mots. Il a dressé une liste de variables qui pourraient permettre de mieux définir cette complexité. Selon lui, la consistance phonographémique (le phonème /l/ étant toujours représenté par le graphème « l », par exemple), la complexité graphémique (nombre de lettres composant un graphème) et le contexte d'utilisation des graphèmes (graphies consistantes et inconsistantes; contextuelles et acontextuelles) doivent constituer les variables principales pour déterminer la complexité orthographique des mots. À partir de ces variables, Pérez (2014a) a donc établi une classification de la difficulté d'acquisition des graphèmes du français. Pour ce, il s'est inspiré des pourcentages d'utilisation des graphèmes selon Catach (2016), qui avait présenté 45 graphèmes devant être acquis prioritairement, mais qui ne les avait pas classés. Il s'est aussi inspiré des recherches de Mousty et Leybaert (1999), qui ont choisi les graphèmes à évaluer dans le cadre de la création de la BELEC à partir des variables suivantes : la complexité des graphèmes, le contexte

de leur utilisation et leur consistance. Le tableau 2.5 présente cette classification des 45 graphèmes de base de Catach (2016), hiérarchisation que Pérez a établie selon leur complexité orthographique.

**Tableau 2.5 Classification des 45 graphèmes du français en fonction de leur degré de complexité (reproduit de Pérez, 2014a, p. 9)**

Difficulté	Caractéristiques	Graphèmes
1	Graphèmes simples acontextuels consistants	a i u p b d f v l r m n
2	Graphèmes complexes acontextuels consistants	eu ou on un oi oin ch gn
3	Graphèmes simples acontextuels inconsistants	é è o ï y
4	Graphèmes complexes acontextuels inconsistants	ai au eau an en in il (l) qu
5	Graphèmes simples contextuels inconsistants	e g s c ç z j x
6	Graphèmes complexes contextuels inconsistants	gu ss ge

Ainsi, selon la classification de Pérez (2014a), les correspondances phonèmes-graphèmes consistantes et ne variant pas selon le contexte seraient acquises en premier; les graphèmes simples étant acquis une étape avant les graphèmes complexes. Puis, viendrait l'acquisition des correspondances phonèmes-graphèmes qui ne varient toujours pas selon le contexte, mais qui ne sont pas consistantes; les graphèmes simples étant ici aussi acquis une étape avant les graphèmes complexes. Enfin, les correspondances phonèmes-graphèmes qui varient selon le contexte seraient acquises plus tard selon cette classification, surtout si elles ont plus d'une lettre. Toutefois, Pérez n'a pas testé l'acquisition des graphèmes de sa classification, ce qui est une limite à sa recherche. Il mentionne d'ailleurs que, quoiqu'indispensables, les trois variables retenues pour établir la hiérarchisation qu'il a établie sont insuffisantes. Il mentionne l'importance de tenir aussi compte des variables lexicales (qui concernent le mot), de la fréquence, du voisinage, de la durée acoustique et de la longueur orthographique dans les recherches portant sur l'orthographe lexicale (Pérez, 2014a).

### 2.4.2.2 Régularités graphotactiques

Les élèves acquièrent de façon implicite des caractéristiques de leur système orthographique, comme la distribution des correspondances phonèmes-graphèmes. Ces connaissances graphotactiques leur permettent, entre autres, d'utiliser les graphèmes les plus fréquents pour orthographier un mot dont ils ne maîtrisent pas l'orthographe.

Les régularités graphotactiques ont un rôle à jouer dans l'acquisition des phonèmes multigraphémiques. Ces régularités sont basées sur les probabilités de l'orthographe lexicale. Elles sont liées à « la fréquence de combinaisons des graphèmes » de l'orthographe lexicale (Borchardt, 2012, p. 6), c'est-à-dire à leur structure répétitive (Demont et Gombert, 2004) et sont relatives à la combinaison légale des lettres (Pacton, 2008). Elles peuvent être positionnelles, contextuelles et distributionnelles.

Les régularités graphotactiques positionnelles sont associées au fait que certains graphèmes varient selon leur position. Par exemple, dès la deuxième année du primaire, les élèves acquièrent le fait que le phonème /o/ ne s'écrit jamais « eau » en début de mot, et ce, sans qu'on ait à leur enseigner (Pacton et al., 2002).

Il est possible de parler de régularités graphotactiques contextuelles lorsque l'utilisation de certains graphèmes dépend du contexte. Cette régularité a été précédemment décrite comme une règle de positionnement dans le modèle de Daigle et ses collaborateurs (2013). C'est, entre autres, le cas du graphème « n » qui se transforme en « m » devant le « b » et le « p » et du graphème « s » qui doit être doublé entre deux voyelles pour représenter le phonème /s/. Catach (2016) considère que les graphèmes contextuels ne sont acquis qu'à la fin du primaire, ce qui a été corroboré par Mousty et Leybaert (1999). Ils ont en effet montré que, plus les élèves rencontrent de mots qui contiennent un graphème contextuel, plus ils ont de chances de bien l'orthographier, ce qui nécessite donc du temps. Ils ont d'ailleurs conclu que ce n'est qu'en quatrième année du primaire que 72,9 % des élèves parvenaient à orthographier ces graphèmes contextuels.

Enfin, les régularités graphotactiques distributionnelles sont associées au voisinage d'un graphème. Ce voisinage peut en faciliter l'acquisition. Par exemple, les élèves pourraient acquérir implicitement le fait que c'est le graphème « an » et non le graphème « en », qui est utilisé après

les lettres « ch ». Le voisinage peut aussi nuire à l'acquisition d'un graphème. Pothier et Pothier (2004) ont remarqué que la présence d'une semi-voyelle, comme le /j/ de « bail » ou le /w/ de « oiseau », augmentait les chances d'erreurs dans l'orthographe du mot. En effet, la segmentation phonémique est plus difficile pour les mots comportant des semi-voyelles, car elles sont plus difficiles à identifier dans le mot, ce qui entraîne des difficultés en orthographe.

Les chercheurs s'intéressent à l'influence des régularités graphotactiques sur l'acquisition de l'orthographe lexicale depuis une vingtaine d'années. Pacton est une figure de proue de ces recherches. Pour évaluer la sensibilité des élèves à certaines régularités graphotactiques, il utilise des épreuves de jugement de non-mots (Pacton, 2008). Les paragraphes suivants présentent ses recherches importantes sur le sujet.

En 2001, Pacton et ses collaborateurs (2001) ont testé la sensibilité des élèves de l'école primaire à la fréquence du doublement des consonnes. Pour ce faire, ils ont, entre autres, testé 18 paires de non-mots de six lettres contenant une lettre double en position centrale auprès d'élèves de la première à la quatrième année du primaire. Chaque paire de non-mots contenant un choix de consonnes fréquemment doublées (mm ou ss, par exemple) et un choix de consonnes rarement doublées (dd ou vv, par exemple) figurait sur chaque page. La tâche des élèves consistait à examiner chaque paire de non-mots et à encercler le non-mot qui leur semblait le plus ressembler à un mot. L'hypothèse voulant que les élèves soient sensibles à des formats fréquents de doubles lettres s'est confirmée puisqu'ils ont, en grande majorité, choisi la structure du mot la plus probable dès la première année du primaire.

Ces mêmes chercheurs (Pacton et al., 2001) ont aussi vérifié si les élèves sont sensibles à la position légale des doubles lettres. Pour ce faire, ils ont utilisé deux types de pseudomots. En entrevue, Pacton (communication personnelle, 2019) a expliqué que des tests de Student ont été faits pour déterminer si les performances étaient attribuables au hasard pour les items pour lesquels il y avait une bonne réponse. Par exemple, des paires comme « tunnor » – « ttunor » ont permis de tester l'illégalité des doublets en début de mots. Pour certaines propriétés, comme la fréquence de doublement des consonnes, il s'agissait de la comparaison de deux types de paires, par exemple des choix entre « imose » et « idose » et entre « immose » et « iddose » et les chercheurs observaient si le choix d'une structure comme « d » était plus probable que « dd », qui est plus rare.

Ils ont observé que le choix du doublet était plus fréquent pour un choix entre m et mm que pour un choix entre d et dd. L'avantage de cette situation, a affirmé Pacton, est que les deux items se prononcent pareillement, mais l'inconvénient est que certains enfants choisissent systématiquement la consonne simple, ce qui n'est pas faux, car même une consonne comme « m » doit être encore plus fréquente en simple qu'en double. Quoiqu'il en soit, dès la première année du primaire, les élèves ont montré qu'ils étaient sensibles à la position légale des doubles lettres, ce qui a permis de confirmer l'hypothèse des chercheurs qui ont découvert qu'il y a davantage de probabilités que les élèves francophones de la première à la quatrième année du primaire identifient des non-mots comme des mots probables quand ils contiennent des structures de doubles consonnes fréquentes. Cela montre donc que les jeunes enfants acquièrent déjà, de façon implicite, des connaissances concernant les lettres qui peuvent être doublées ou non. Les chercheurs associent cette maîtrise à ce qu'ils appellent l'expérience implicite linguistique des élèves qui mémorisent les mots lus et écrits avec lesquels ils sont en contact.

Pacton et ses collaborateurs (2002) ont aussi évalué la capacité implicite de jeunes élèves à rédiger correctement le phonème multigraphémique /o/ dans des non-mots selon la position que ce phonème occupe. Sachant que la graphie « eau » est fréquente en position finale, mais inexistante en position initiale et qu'elle est plus fréquente après un « v » qu'après un « f », ils ont donné des non-mots contenant ces structures en dictée. Les chercheurs ont utilisé des pseudomots avec le phonème /o/ en début, en milieu et en fin de mots (Pacton, 2019). De plus, en position finale, le phonème /o/ était précédé de consonnes après lesquelles « eau » est soit fréquent, soit rare, voire impossible (par exemple, « reau » est fréquent et « feau » est illégal ou, en tout cas, très rare). Les chercheurs ont vérifié si le graphème « eau » variait en fonction de la consonne de gauche, ce qui a été le cas. En effet, les enfants utilisent plus le graphème « eau » pour « bilareau » que pour « bilafeau ». Ainsi, dès la deuxième année du primaire, les élèves arrivent à écrire le phonème multigraphémique /o/ en fonction de son environnement consonantique et de sa position. Cette recherche a montré la sensibilité précoce des jeunes élèves aux régularités graphotactiques, et ce, sans que ces régularités leur aient été explicitement enseignées. Les réponses des élèves, dans le cadre de cette recherche, sont étonnantes : de façon très précoce, ils arrivent à affirmer que le non-mot « opécile » est plus vraisemblable que « eaupécile » (Fayol et Jaffré, 2016).

En 2012, dans le cadre d'une thèse de recherche doctorale dirigée par Pacton, Borchardt (2012) a étudié le rôle des connaissances graphotactiques dans l'acquisition de l'orthographe lexicale chez les élèves de la troisième et de la cinquième année du primaire et chez les adultes en France. Pour ce faire, à l'aide d'analyses de corrélation et de régression, elle a notamment évalué si les connaissances graphotactiques pouvaient prédire l'acquisition de non-mots. Des épreuves de jugement de paires de non-mots contenant des situations possibles ou non de doublement de lettres ont permis de remarquer que les sujets sont plus sensibles à l'illégalité du doublement de consonnes en début de mots qu'en milieu de mots. Borchardt a aussi étudié le rôle des connaissances graphotactiques dans l'acquisition de mots nouveaux. Pour ce faire, elle a fait lire des histoires contenant des non-mots. Ces non-mots ont ensuite été testés par deux épreuves : une épreuve de dictées et une épreuve de jugement de paires de non-mots. Cela a permis d'établir des corrélations significatives entre les connaissances graphotactiques et la maîtrise de l'orthographe lexicale, montrant que les connaissances graphotactiques influencent positivement la mémorisation des mots. Enfin, Borchardt (2012) a fait apprendre des non-mots qui comprenaient des doubles consonnes aux élèves et aux adultes. Un groupe a appris ces non-mots de façon implicite, en les lisant simplement, alors que l'autre groupe a eu droit à un enseignement explicite des régularités graphotactiques associées aux doubles consonnes. L'étude a prouvé que l'enseignement explicite des régularités graphotactiques était corrélé à une meilleure acquisition de l'orthographe lexicale, du moins dans le cas de non-mots. Ainsi, les recherches de Pacton et de ses collaborateurs ont montré que les élèves sont sensibles aux régularités graphotactiques lorsqu'ils acquièrent l'orthographe lexicale. Il est fort probable que ces régularités graphotactiques jouent aussi un rôle dans l'acquisition des phonèmes multigraphémiques.

L'enseignement explicite des régularités graphotactiques peut contribuer à leur acquisition. Stanké et ses collègues (2020) ont mené une recherche analysant l'impact d'un enseignement explicite de l'utilisation des régularités orthographiques, tant graphotactiques que morphologiques, sur l'acquisition de l'orthographe lexicale. L'échantillon était constitué d'élèves québécois provenant de classes de troisième année du primaire. Chaque semaine, pendant huit semaines, les élèves des classes expérimentales ont appris dix mots associés à l'une des quatre régularités graphotactiques ou à l'une des quatre régularités morphologiques retenues. Chaque régularité était enseignée de façon explicite et révisée tous les jours à partir d'un programme d'entraînement. Chez les élèves

bons orthographes, les résultats ont montré une amélioration de 35 % après chaque séance et de 13 % après 45 jours. Les résultats des élèves faibles orthographes ont aussi été significatifs, ces derniers dépassant les résultats des élèves du groupe contrôle. La recherche a donc conclu que l'enseignement explicite des régularités orthographiques, tant graphotactiques que morphologiques, favorise le développement des compétences en orthographe lexicale de tous les élèves, tant les forts que les faibles orthographes (Stanké et al., 2020). Ainsi, plus que de prouver le rôle important de l'enseignement des régularités graphotactiques, Stanké et al. (2020) ont montré que l'enseignement explicite des régularités morphologiques facilite l'acquisition des phonèmes multigraphémiques puisque les élèves peuvent se baser sur le sens des morphèmes pour orthographier les mots. La prochaine section s'intéresse à ces régularités morphologiques.

### **2.4.2.3 Régularités morphologiques**

Bien que certains chercheurs aient conclu que le traitement morphographémique était acquis plus tardivement que les autres traitements (Morin et Montésinos-Gelet, 2005; Mousty et Leybaert, 1999; Seymour, 1997), des recherches ont permis de comprendre que son acquisition pouvait se faire assez rapidement. Sénéchal (2000) a étudié la possibilité pour les élèves de deuxième année du primaire de se représenter des informations morphologiques en orthographiant. Les élèves devaient représenter des informations morphologiques en orthographiant des mots français qui avaient des terminaisons à consonance silencieuse. Ils devaient épeler des mots réguliers, morphologiques (présence d'une lettre dérivationnelle) et profonds (absence de lettre dérivationnelle). La recherche a confirmé que les mots réguliers étaient les plus faciles à épeler. De plus, elle a conclu que les mots morphologiques contenant une consonne silencieuse dérivable (par exemple, le mot « vent ») étaient plus faciles à épeler que les mots n'ayant pas de dérivation possible et pour lesquels la consonne silencieuse devait donc être mémorisée (par exemple, le mot « jamais »). Les connaissances linguistiques que les enfants avaient de la morphologie ont donc contribué à ce qu'ils orthographient correctement des mots morphologiques, et ce, indépendamment de leur expérience de lecture, de leur vocabulaire, de leur capacité d'orthographe et de la conscience des phonèmes.

Une recherche de Pacton et ses collègues (2005) a, pour sa part, montré que les élèves étaient sensibles aux régularités morphologiques et aux régularités graphotactiques dès l'âge de 7 ans. En

effet, quand on leur a enseigné que le son /ɛt/ s'écrivait toujours « ette » quand il était employé comme diminutif et qu'il était plus fréquemment utilisé quand il était précédé du graphème « r » que du graphème « ch », les élèves ont réussi à rédiger des non-mots faisant appel à ces deux règles.

Casalis et al. (2011) ont aussi montré que les élèves étaient sensibles aux régularités morphologiques. Plus précisément, ils ont montré que, dans le cas de phonèmes multigraphémiques, les élèves arrivaient mieux à orthographier les graphèmes inconsistants contenus dans les mots dérivés que dans les autres mots. Par exemple, la connaissance du mot « lait » peut permettre d'orthographier le phonème multigraphémique /ɛ/ dans le mot « laitier » alors qu'il sera plus difficile d'orthographier ce même phonème multigraphémique dans le mot « falaise » puisqu'il ne dérive d'aucun mot. Cela dit, leur étude a porté sur les mots dérivés et non sur les mots dérivables. Ils n'ont donc pas évalué si le fait qu'un phonème multigraphémique en fin de mot soit dérivable a un impact sur la réussite orthographique.

Toutefois, bien que le traitement morphographémique puisse se faire assez rapidement lorsque les règles sont enseignées explicitement, des recherches ont prouvé que ce n'était pas le cas pour les faibles orthographieurs et que ce n'était pas toujours bénéfique pour les jeunes élèves. En effet, comme les faibles orthographieurs sont moins exposés à l'écrit, Bégin et al. (2010) ont observé qu'ils éprouvaient davantage de difficulté avec la morphologie dérivationnelle. Pour leur part, Borchardt et al. (2012) ont observé que les élèves, en début d'apprentissage, ont tendance à généraliser la règle des morphèmes dérivables, ce qui fait en sorte qu'ils ajoutent des lettres muettes inutiles en fin de mot.

Enfin, à notre connaissance, si ce n'est de Stanké et ses collègues (2021), qui ont intégré la dérivabilité dans leur nouvelle approche basée sur un enseignement explicite de l'orthographe lexicale, peu de chercheurs se sont intéressés à l'influence d'un morphogramme en fin de mot qui n'est pas muet sur l'acquisition de leur orthographe. Par exemple, nous pourrions nous demander si certains mots en « in » qui dérivent en « ine », comme « lutin/lutine », et certains mots en « ain » qui dérivent en « aine », comme « certain/certaine », sont acquis plus rapidement que ceux dont le graphème ne s'explique pas par leur dérivation.

Ainsi, tous ces facteurs lexicaux et sous-lexicaux pourraient avoir un rôle à jouer dans l'acquisition des phonèmes multigraphémiques dans les productions orthographiques des élèves québécois bons et faibles orthographes de la troisième année du primaire. Ces phonèmes multigraphémiques font partie de mots à acquérir qui sont souvent présentés dans des listes orthographiques au primaire. Il importe de présenter ces listes, qui sont des facteurs environnementaux importants qui permettent de mieux comprendre à quels mots correspondent ces phonèmes complexes à acquérir au primaire.

## **2.5 Facteurs environnementaux liés aux bases de données et aux échelles d'acquisition en orthographe lexicale**

Il existe une différence importante entre les bases de données lexicales et les échelles d'acquisition de l'orthographe. Les bases de données lexicales sont issues de différentes sources et ont, entre autres, été créées à partir des mots les plus fréquents dans les livres. Certaines échelles orthographiques tiennent compte de la compétence à orthographier les mots en évaluant l'acquisition de l'orthographe lexicale. Elles sont basées sur la capacité des élèves d'un cycle donné à orthographier des mots. Cette capacité se mesure en pourcentage. Pothier et Pothier (2004) considèrent que l'orthographe d'un mot est acquise quand elle est maîtrisée par plus de 75 % des élèves pendant deux années consécutives. La présente section présente des bases de données lexicales ainsi que des échelles d'acquisition de l'orthographe lexicale en France et au Québec.

### **2.5.1 Bases de données lexicales**

Plusieurs bases de données lexicales existent à travers la francophonie. En France, Novlex (Lambert et Chesnet, 2001), Manulex et Manulex-Morpho (Lété et al., 2004) sont des bases de données lexicales qui fournissent des informations sur les caractéristiques des mots (par exemple, la longueur du mot, le nombre de phonèmes, le nombre de graphèmes) auxquels les enfants sont exposés en littérature et dans les manuels de français. Par exemple, Manulex, base de données consultée pour la présente recherche doctorale, est une base informatique de données qui a été créée à partir de 54 manuels de français du CP au Cycle 3 en France, soit de la première à la sixième année au Québec. Cette liste présente, entre autres, la fréquence de 48 900 mots et flexions, c'est-à-dire qu'ils sont présentés avec leurs dérivés (Gala et Zock, 2013). Comme mentionné, cette base ne fournit pas d'informations sur l'acquisition des mots par les élèves.

Au Québec, la liste orthographique à l'usage des enseignantes et des enseignants a été bâtie, entre autres, à partir des statistiques d'apparition de mots dans un corpus d'environ 250 ouvrages, dont la majorité était des œuvres de fiction (ministère de l'Éducation et de l'Enseignement supérieur [MEES], 2014). En se basant sur la fréquence des mots pour cibler les 3 000 mots les plus fréquents à acquérir par les élèves à la fin du primaire, le MEES a fait un pas important puisqu'aucun outil commun n'était utilisé en orthographe lexicale avant 2014 au Québec. Toutefois, cette liste ne présente pas de gradation des mots contenant des phonèmes multigraphémiques. Cela dit, le MEES a commencé à s'adapter aux résultats des recherches récentes. En effet, dans la progression des apprentissages, le Ministère a précisé l'importance d'enseigner l'orthographe des mots en fonction des régularités orthographiques afin que les élèves connaissent les constantes orthographiques liées, entre autres, à l'emploi de certains phonèmes multigraphémiques (ministère de l'Éducation, du Loisir et du Sport, 2011) et un moteur de recherche permet de trier les mots par graphèmes ou par phonèmes. Toutefois, le MEES n'a pas évalué la compétence des élèves à orthographier chacun des mots de cette liste ainsi que chacun des phonèmes multigraphémiques contenus dans ces mots selon les cycles scolaires (Tremblay et al., 2013). Cela confirme que la liste orthographique à l'usage des enseignantes et des enseignants est une base de données et non une échelle d'acquisition de l'orthographe lexicale. En effet, les échelles d'acquisition de l'orthographe lexicale tiennent compte de la compétence des élèves à orthographier. Utilisées depuis longtemps en France, elles se développent présentement au Québec.

### **2.5.2 Échelles d'acquisition en orthographe lexicale**

En France, l'*Échelle d'acquisition en orthographe lexicale* pour l'école élémentaire (ÉOLE) a été élaborée à partir de 11 979 mots les plus fréquents dans des journaux pour adultes (Pothier et Pothier, 2004). Ces termes ont été divisés en 240 listes comportant 50 mots chacune. Chaque mot a été orthographié par 40 élèves par niveau scolaire, des traitements statistiques ayant permis d'établir cette norme pour que ce soit recevable. Cinq niveaux scolaires, soit de la fin de première à la fin de la cinquième année du primaire, ont été retenus; chaque mot a donc été orthographié 200 fois. Un prétest a permis de valider le protocole de passation. Puis, chacun des 50 mots a été évalué dans un contexte scolaire. Au total, ce sont 48 902 élèves fréquentant 464 écoles différentes qui ont participé à la recherche, produisant l'orthographe de 2 368 885 mots. L'analyse quantitative de l'orthographe des mots des dictées a permis de conclure que les élèves peuvent maîtriser le nombre

de mots suivants pour chaque année scolaire en France et qu'il est donc possible de les leur enseigner : 347 mots en première année, 1 013 mots en deuxième année, 2 188 mots en troisième année, 3 606 mots en quatrième année et 5 054 mots en cinquième année. Cela dit, bien que *l'Échelle d'acquisition en orthographe lexicale* pour l'école élémentaire ait tenu compte de la compétence des élèves à orthographier des mots, les documents à partir desquels les mots sont issus, proviennent de la presse internationale d'adultes et d'adolescents et non de livres auxquels les enfants sont exposés. De plus, ils ont été élaborés à partir des compétences des élèves français et non d'élèves québécois. La liste orthographique ÉOLE pourrait donc être peu pertinente pour permettre aux élèves québécois d'acquérir l'orthographe lexicale. En effet, Stanké et al. (2019) ont montré que les compétences orthographiques des élèves québécois sont meilleures que celles des élèves français pour plus de 60 % des mots orthographiés d'ÉOLE.

Au Québec, une première échelle de langage et d'acquisition orthographique a été créée par Préfontaine en 1979 à partir de 4 673 mots. Son objectif était de créer une échelle commune entre la fréquence orale des mots et leur utilisation écrite (Préfontaine, 1979). Mille-cinq-cent-trente mots ont été séparés en 12 listes d'environ 125 mots. Chacune des 12 listes a été dictée à 300 enfants. Puis, 3 064 mots ont aussi été séparés en 12 listes, cette fois de 255 mots chacune. Chacune de ces listes de mots a été donnée en dictée à 300 enfants de la deuxième à la sixième année du primaire. Bien que cette liste ne soit plus utilisée dans le système scolaire, la recherche qui a permis de la construire a confirmé une hypothèse importante : plus un mot est fréquent à l'oral, plus son orthographe est réussie à l'écrit (Préfontaine, 1979).

Toujours au Québec, Stanké et ses collaborateurs (2018) ont publié l'échelle québécoise d'acquisition de l'orthographe lexicale (ÉQOL) en mars 2018. Cette échelle, qui est utilisée dans le cadre de la présente recherche doctorale, a été construite à partir de 14 800 mots différents extraits de 28 manuels scolaires, de 24 livrets et de 90 extraits de littérature jeunesse. Elle tient compte de la capacité des élèves à orthographier un mot à un cycle donné. En effet, elle a été bâtie à partir de 1 200 mots les plus fréquents d'un répertoire de 14 800 mots. Ces mots ont été dictés à 4 733 élèves de la première à la sixième année, fréquentant 47 écoles primaires du Québec. Le taux de réussite lié aux caractéristiques des mots a montré que les facteurs suivants sont déterminants pour l'acquisition de l'orthographe lexicale : « le niveau élevé de consistance phonème/graphème; la fréquence élevée des mots; le nombre réduit de syllabes du mot » (Stanké et al., 2020, p. 9). Cette

échelle est innovante au Québec puisqu'elle inclut, entre autres, des mots qui ont été tirés de manuels scolaires et qui représentent les disciplines enseignées au primaire et que ces mots ont été testés chez les élèves québécois. Cela dit, cette échelle a testé l'acquisition des mots dans leur ensemble et non l'acquisition des phonèmes multigraphémiques.

Bref, nombre de facteurs individuels, linguistiques et environnementaux influencent l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques. Les facteurs linguistiques, qu'ils soient lexicaux ou sous-lexicaux, jouent un rôle important dans cette acquisition. Il convient maintenant de présenter une synthèse des recherches liées à ces facteurs.

## **2.6 Synthèse et objectifs de recherche**

La présente section permet de faire une synthèse de différentes recherches. Ces recherches figurent parmi celles qui ont donné les résultats les plus significatifs sur l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques. Les objectifs de recherche ainsi que les hypothèses de recherche de la présente recherche doctorale sont aussi exposés.

### **2.6.1 Synthèse des recherches portant sur l'acquisition de l'orthographe lexicale**

Différentes recherches ont étudié le rôle des facteurs linguistiques sur l'acquisition de l'orthographe lexicale. Cela dit, les recherches présentées dans le tableau 2.6 se questionnent peu sur l'acquisition des phonèmes multigraphémiques.

Certes, grâce à leurs recherches portant sur les régularités graphotactiques, Pacton et ses collaborateurs ont pu répondre à certaines questions sur les inconsistances phonographémiques. Ces recherches ont d'ailleurs été très aidantes pour cibler les facteurs étudiés dans la présente thèse. Toutefois, elles ont toutes été faites à partir de non-mots et non de véritables mots.

De plus, le tableau 2.6 permet de constater que peu de recherches se sont intéressées à l'acquisition de l'orthographe lexicale chez des élèves canadiens francophones. Or, il importe de baser les interventions didactiques et orthopédagogiques sur les résultats de recherches canadiennes francophones. Ce dernier point est majeur puisque, comme l'a montré l'Organisation de coopération et de développement économiques (OCDE), il existe une grande disparité dans les

compétences en lecture des élèves de pays différents, compétences intimement reliées à l'orthographe (Dumais et al., 2014).

**Tableau 2.6 Synthèse des facteurs ayant un effet sur l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques à partir des données probantes des recherches**

Facteurs ayant un effet sur l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques	Recherches
<b><u>Fréquence des mots</u></b>	
Orthographe des mots fréquents acquise plus rapidement	(Martinet et al., 2004; Mousty et Leybaert, 1999; Sprenger-Charolles et al., 1998; Stanké et al., 2019)
<b><u>Longueur des mots</u></b>	
Orthographe des mots comptant peu de syllabes, acquise plus rapidement	(Stanké et al., 2019)
<b><u>Consistance des mots</u></b>	
Le degré de consistance d'une langue joue non seulement un rôle important dans l'acquisition de sa lecture, mais aussi dans l'acquisition de son orthographe.	(Caravolas 2019; Treiman 2019)
<b><u>Consistance phonographémique</u></b>	
Application de la graphie dominante par les élèves lorsque le lexique orthographique n'est pas assez riche (ex. : utilisation du graphème /o/ plutôt que /eau/) et	(Mousty et Leybaert, 1999)
Acquisition rapide des graphèmes consistants acontextuels (ex. : /l/ et /ou/)	(Mousty et Leybaert, 1999)

---

**Facteurs ayant un effet sur l'acquisition de  
l'orthographe lexicale et des phonèmes  
multigraphémiques**

---

**Recherches**

**Régularités graphotactiques de positionnement**

Sensibilité implicite des jeunes élèves à la position  
légale des lettres dans l'acquisition de l'orthographe

(Pacton et al., 2001)

---

Sensibilité implicite des jeunes élèves à  
l'environnement consonantique et à la position  
d'une correspondance phonème-graphème (ex. :  
« eau » impossible en début de mot)

(Pacton et al., 2002)

**Régularités graphotactiques contextuelles**

Acquisition tardive des graphies consistantes  
contextuelles, c'est-à-dire des graphèmes qui  
varient selon le contexte (ex. : « c » devant « a »,  
« o », « u » et « qu » devant « e » et « i »)

(Mousty et Leybaert, 1999)

Risque accru d'erreurs selon le voisinage du  
graphème, plus précisément en présence d'une  
semi-voyelle

(Pothier et Pothier, 2004)

**Régularités morphologiques**

Acquisition tardive des stratégies morphologiques

(Mousty et Leybaert 1999)

Les faibles orthographes éprouvent davantage de  
difficulté avec la morphologie dérivationnelle

(Bégin et al., 2010)

Mots se terminant par une consonne silencieuse  
dérivable plus faciles à acquérir que si la consonne  
silencieuse n'a pas de dérivation possible

(Sénéchal, 2000)

Graphèmes inconsistants contenus dans les mots  
dérivés mieux orthographiés que dans les autres  
mots

(Casalis et al., 2011)

---

Les recherches qui ont été faites à ce jour ont donc contribué à l'avancement des connaissances en ce qui concerne l'acquisition de l'orthographe lexicale, mais elles ne répondent que partiellement à la question de la présente thèse qui cherche à savoir quels sont les facteurs linguistiques qui favorisent l'acquisition des phonèmes multigraphémiques dans les productions orthographiques des élèves québécois bons et faibles orthographes de la troisième année du primaire.

### **2.6.2 Objectifs et hypothèses de recherche**

Bref, bien que des recherches se soient intéressées aux facteurs linguistiques qui influencent l'acquisition de l'orthographe lexicale, à notre connaissance, très peu se sont intéressées aux facteurs linguistiques qui favorisent l'acquisition des phonèmes multigraphémiques dans les productions orthographiques. D'ailleurs, les phonèmes multigraphémiques sont très peu pris en compte dans les listes orthographiques et dans les bases de données linguistiques. De plus, les recherches québécoises à ce sujet sont peu nombreuses et tiennent rarement compte des élèves faibles orthographes. Pourtant, les phonèmes multigraphémiques sont une source majeure de difficulté chez les faibles orthographes (Stanké et al., 2020). Chez ces élèves, des facteurs individuels liés aux processus cognitifs de la mémoire, de la conscience phonologique et de l'empan visuoattentionnel sont souvent atteints, ce qui résulte en de faibles compétences en orthographe lexicale (Stanké et al., 2021).

L'analyse des recherches antérieures, qui précisent les facteurs linguistiques ayant un effet sur l'acquisition de l'orthographe lexicale et sur l'acquisition des phonèmes multigraphémiques, permet de préciser l'objectif principal de recherche.

#### **Objectif général**

La présente recherche doctorale a pour objectif général d'identifier les facteurs linguistiques qui favorisent l'acquisition des phonèmes multigraphémiques dans les productions orthographiques des élèves québécois bons et faibles orthographes de la troisième année du primaire.

Les facteurs linguistiques qui ont été retenus sont lexicaux, c'est-à-dire qu'ils concernent le mot en entier, et sous-lexicaux, c'est-à-dire qu'ils concernent les graphèmes à l'intérieur du mot. Les facteurs lexicaux sont liés à la consistance, à la fréquence et à la longueur du mot (nombre peu

élevé de lettres). Quant à eux, les facteurs sous-lexicaux sont liés à la consistance phonographémique, aux régularités graphotactiques positionnelles, contextuelles et distributionnelles (voisinage du graphème) ainsi qu'aux régularités morphologiques dérivationnelles.

Afin de pouvoir répondre à notre question de recherche, il a fallu choisir un phonème multigraphémique qui répondait à tous ces facteurs. C'est le phonème multigraphémique / $\tilde{\epsilon}$ / qui a été retenu pour la présente recherche. Ce phonème / $\tilde{\epsilon}$ / est apparu comme étant le seul qui permettait de traiter de tous les facteurs. Par exemple les facteurs sous-lexicaux ayant un effet sur l'acquisition des phonèmes multigraphémiques, soit les graphèmes associés au phonème / $\tilde{\epsilon}$ /, varient selon leur position (toujours « in » en début de mot sauf pour le mot « ainsi »), le contexte d'utilisation (« im » devant le « b » et le « p ») et selon leur voisinage (« in » suivant le « o » dans « oin » et « en » suivant le « i » dans « ien » et le « y » dans « yen »). Les graphèmes associés au phonème / $\tilde{\epsilon}$ / varient aussi selon la morphologie dérivationnelle : (« in » dans le cas d'une dérivation possible en « i » comme dans « patin-patiner » et « ain » lors d'une dérivation possible en « a » comme dans « main-manuel » ou en « ai » comme dans « certain-certaine »). De plus, le phonème / $\tilde{\epsilon}$ / est associé à de nombreux graphèmes, dont la consistance varie beaucoup : « in », « ain », « ein », « en », « im », « aim », « ym ». Notre objectif spécifique est donc directement lié au phonème multigraphémique / $\tilde{\epsilon}$ /.

### **Objectif spécifique**

L'objectif spécifique de la présente recherche doctorale est d'établir les facteurs linguistiques lexicaux et sous-lexicaux qui ont un effet sur l'acquisition du phonème multigraphémique / $\tilde{\epsilon}$ / dans les productions orthographiques des élèves québécois faibles et bons orthographes de la troisième année du primaire.

### **Hypothèses de recherche**

Comme les phonèmes multigraphémiques sont une source importante de difficulté chez les faibles orthographes et que ces derniers ne semblent pas acquérir l'orthographe de la même façon que les bons orthographes (Bégin et al., 2010; Sprenger-Charolles et Colé, 2013; Stanké et al., 2016,

2020), les hypothèses de recherche se divisent entre les bons et les faibles orthographes. Plus précisément, selon les résultats des recherches présentées dans la problématique et dans le cadre théorique, les hypothèses retenues se formulent comme suit :

Hypothèse 1. Les facteurs lexicaux associés à la fréquence, à la longueur (nombre peu élevé de lettres) <sup>6</sup> et à la consistance du mot ont un effet positif sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois bons orthographes de la troisième année du primaire.

Hypothèse 2. Certains facteurs sous-lexicaux (types de graphèmes, consistance phonographique, positionnement du graphème et régularités morphologiques dérivationnelles) ont un effet positif et certains autres facteurs sous-lexicaux (contexte d'utilisation et voisinage du graphème) ont un effet négatif sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois bons orthographes de la troisième année du primaire.

Hypothèse 3. Les facteurs lexicaux associés à la fréquence, à la longueur (nombre peu élevé de lettres) et à la consistance du mot ont un effet positif sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois faibles orthographes de la troisième année du primaire,

Hypothèse 4. Certains facteurs sous-lexicaux (types de graphèmes, consistance phonographique et positionnement du graphème) ont un effet positif et certains autres facteurs sous-lexicaux (voisinage du graphème, contexte d'utilisation du graphème et régularités morphologiques dérivationnelles) ont un effet négatif sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois faibles orthographes de la troisième année du primaire.

---

<sup>6</sup> Dans les hypothèses de la présente thèse, la longueur du mot signifie sa « brièveté », c'est-à-dire le peu de lettres qu'il contient.

### CHAPITRE 3

#### MÉTHODOLOGIE

La présente thèse vise à identifier les facteurs linguistiques qui favorisent l'acquisition des phonèmes multigraphémiques dans les productions orthographiques des élèves québécois, bons et faibles orthographes de la troisième année du primaire. Plus spécifiquement, elle cherche à établir les facteurs lexicaux et sous-lexicaux qui ont un effet sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois faibles et bons orthographes de la troisième année du primaire. Les facteurs lexicaux ciblés sont la fréquence, la longueur (nombre peu élevé de lettres) et la consistance du mot. Les facteurs sous-lexicaux retenus sont les types de graphèmes, la consistance phonographémique, les régularités graphotactiques (positionnement, contexte d'utilisation et voisinage du graphème) et les régularités morphologiques dérivationnelles.

Afin de présenter la méthodologie retenue, la posture épistémologique ainsi que le devis et les variables de l'étude sont d'abord exposés. Puis, les choix méthodologiques concernant les participants, le déroulement de l'étude, les instruments de collecte de données, le plan d'analyse ainsi que les enjeux déontologiques sont expliqués.

Il importe d'abord de préciser que notre recherche complète une recherche plus large menée par Stanké et al. (2020) et s'inspire donc de la méthodologie de cette recherche, dont elle utilise des données brutes non analysées. Intitulée *Nouvelle approche basée sur un enseignement orthographique favorisant l'apprentissage de la production écrite et de l'orthographe lexicale des élèves faibles orthographes de 3<sup>e</sup> année du primaire en milieu inclusif*, la recherche de Stanké et al. (2020) avait comme objectif principal de montrer qu'une approche basée sur un entraînement explicite de l'utilisation des régularités graphotactiques ou morphologiques favorisait non seulement l'apprentissage de l'orthographe lexicale, mais également celui de la production de textes écrits chez les élèves faibles orthographes. Pour ce faire, Stanké et ses collègues ont élaboré un programme d'entraînement à l'utilisation des régularités orthographiques. Afin d'élaborer ce programme, les chercheurs ont bâti une liste hiérarchisée mesurant la compétence à orthographier : l'échelle québécoise d'acquisition en orthographe lexicale (ÉQOL). L'une des

limites de cette échelle d'acquisition est le fait qu'elle traite de l'acquisition des mots dans leur ensemble, mais ne traite pas de l'acquisition des phonèmes multigraphémiques compris dans les mots. Le groupe de recherche s'était donné l'objectif de le faire en élaborant une échelle intégrant des données d'acquisition des régularités orthographiques. Toutefois, les chercheurs n'ont pas pu atteindre cet objectif avant le dépôt des résultats de la recherche. C'est à cet objectif important que nous avons décidé de répondre partiellement en établissant les facteurs linguistiques lexicaux et sous-lexicaux qui ont un effet sur l'acquisition du phonème multigraphémique /*ẽ*/ dans les productions orthographiques des élèves québécois faibles et bons orthographes de la troisième année du primaire. Pour ce, nous avons dû analyser et recoder chacune des productions orthographiques des élèves afin de savoir où se situaient exactement les erreurs. De plus, lorsque les erreurs se trouvaient dans le phonème multigraphémique /*ẽ*/, nous avons analysé les facteurs qui auraient pu influencer ces erreurs afin de valider ou d'invalider nos hypothèses.

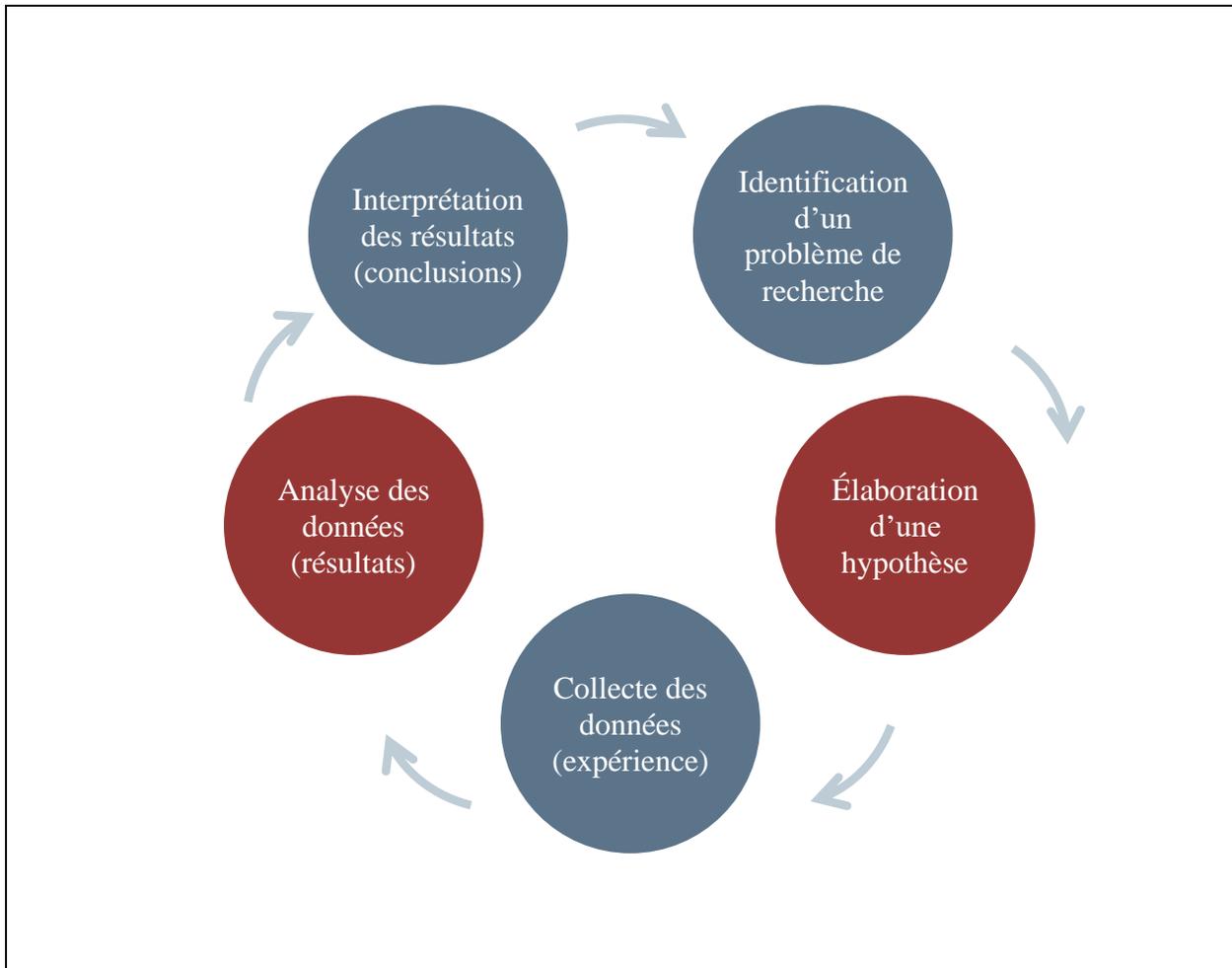
### **3.1 Posture épistémologique**

Toute recherche scientifique se fonde sur une vision du monde, c'est-à-dire sur un paradigme ou sur un modèle de référence, qui oriente la recherche et qui détermine les buts à atteindre et la façon de percevoir les phénomènes (Fortin et Gagnon, 2022). La position épistémologique du chercheur peut jouer un rôle dans le choix de sa méthodologie. Tant la formation première du chercheur, que ses croyances, son vécu professionnel et ses expériences de succès et d'échec ont un rôle à jouer (Pinard et al., 2004). De plus, certains courants peuvent influencer le chercheur. C'est le cas, par exemple, de la popularité de plus en plus grande du courant voulant que l'éducation soit basée sur la preuve (Lehrer et al., 2021). Enfin, lorsque le chercheur fait partie d'une équipe multidisciplinaire, il adopte nécessairement la posture de l'équipe. C'est le cas cette recherche doctorale pour laquelle la chercheuse a développé sa base de données à partir d'un corpus existant en s'appropriant d'abord toute la démarche de la recherche de base.

La présente recherche est de nature hypothéticodéductive, c'est-à-dire qu'elle a pour objectif de confirmer ou d'infirmer des hypothèses de recherche (Lehrer et al., 2021; Van der Maren, 2003). Une méthodologie quantitative s'est donc imposée. La recherche quantitative vise à généraliser les principes tirés des résultats à d'autres participants en validant des hypothèses à partir d'instruments de mesure (Lehrer et al., 2021). Debout (2012) précise que cette approche est bien adaptée pour les

thèmes qui ont été peu étudiés par des travaux scientifiques et qu'elle permet d'explorer des phénomènes décrits sous forme de variables qui peuvent être mesurées. Il ajoute qu'elle permet d'identifier les invariants qui influencent les phénomènes, de développer et de valider des théories explicatives, de modifier le cours d'un phénomène et de retenir la théorie la plus pertinente sur un phénomène. La figure 3.1 illustre les étapes d'une recherche scientifique quantitative.

**Figure 3.1 Les étapes de la démarche scientifique (Lehrer et al., 2021)**



Les questions, les objectifs et hypothèses de la présente recherche doctorale l'inscrivent essentiellement dans le paradigme postpositiviste. Les chercheurs qui ont une posture positiviste ne s'impliquent pas personnellement, recherchant le plus de neutralité possible dans la collecte de données (Fortin et Gagnon, 2022; Karsenti et Savoie-Zajc, 2018). Comme nous avons bâti notre base de données à partir d'un corpus existant, cette méthode nous a permis d'avoir une distance

avec les participants et de demeurer objectifs. De plus, les chercheurs qui ont une posture postpositiviste cherchent à induire des règles générales à partir de cas particuliers de la façon la plus objective possible; toutefois, ils reconnaissent que les valeurs et l'expérience du chercheur jouent un rôle dans leur recherche (Lehrer et al., 2021). Par exemple, dans le cas de la présente recherche doctorale, outre le sujet de recherche qui se prêtait à une analyse quantitative, l'expérience d'évaluation orthopédagogique de la chercheuse, qui travaille avec les élèves dyslexiques/dysorthographiques depuis nombre d'années, a pu influencer le choix de faire de la recherche quantitative. En effet, les outils d'évaluation standardisés permettant d'identifier les élèves à risque de présenter une dyslexie-dysorthographie sont basés sur des épreuves quantitatives.

### **3.2 Devis et variables de l'étude**

Le devis de recherche s'inscrit aussi dans cette posture épistémologique postpositiviste. C'est un devis corrélationnel qui permet de vérifier les hypothèses soulevées. L'objectif d'une recherche corrélationnelle peut être de déterminer les relations entre les variables ou d'utiliser les relations entre les variables afin de faire des prédictions (Gay et al., 2012). La présente recherche est de type corrélationnel prédictif. La recherche corrélationnelle prédictive consiste à collecter des données quantifiables à partir de variables choisies afin de déterminer l'influence des variables les unes sur les autres (Fortin et Gagnon, 2022). C'est l'existence d'une forte corrélation entre les variables qui permet une prédiction juste (Gay et al., 2012). Plus précisément, des corrélations entre le pourcentage de réussite des graphèmes associés au phonème /*ẽ*/ et certaines variables permettent d'évaluer l'effet des différents facteurs linguistiques dans l'acquisition des phonèmes multigraphémiques.

Ce sont les connaissances sur le sujet étayées dans le cadre théorique qui guident le choix des variables. Afin de confirmer ou d'infirmer les hypothèses qui découlent du cadre théorique, le devis corrélationnel mis en place mène à examiner le lien d'une variable indépendante à partir des données reliées aux facteurs linguistiques retenus, c'est-à-dire des variables dépendantes. Les variables utilisées sont décrites au tableau 3.1, qui précise aussi les niveaux des variables catégorielles. La composante aléatoire « élèves » a été choisie afin de tenir compte des caractéristiques propres aux élèves. Cela permet implicitement que d'autres variables propres aux élèves, et non au groupe, soient incluses, comme le genre et le milieu socioéconomique. De plus,

comme nos effets fixes portent sur le mot, le modèle ne convergerait pas si la composante aléatoire était le mot. Enfin, le modèle ne convergerait pas non plus pour l'école, car les écoles différentes (pensons au nombre d'effectifs, par exemple) surchargeaient le modèle.

**Tableau 3.1 Sommaire des variables de la présente recherche**

Variables	Description
<b>Variable dépendante</b>	
Réussite du graphème	Indique si les élèves ont correctement orthographié le graphème en / <i>ɛ̃</i> / dans le mot. 0 : le graphème n'a pas été correctement orthographié; 1 : le graphème a été correctement orthographié.
<b>Variables indépendantes explicatives</b>	
<b>Composante aléatoire</b>	
Élèves	Identifiant anonyme des participants ( $n = 509$ ).
<b>Composantes fixes</b>	
Profil	Profil orthographique des élèves, déterminé par une tâche de dépistage préalable. <ul style="list-style-type: none"> <li>• FAIBLES ORTHOGRAPHEURS : les élèves ont manifesté une difficulté en orthographe lors de la tâche préalable.</li> <li>• SANS DIFF : les élèves n'ont pas manifesté de difficulté en orthographe (référence).</li> </ul>
Graphème	Graphème correspondant au phonème / <i>ɛ̃</i> / dans le mot : EIN (facteur de référence), AIN, EN, IM, IN, AIN-dérivable, IN-dérivable. La mention « dérivable » indique un graphème dérivable ( <i>matin/matinée</i> ; <i>cert<del>ain</del>/cert<del>aine</del></i> ).
Position	Positionnement dans le mot du graphème correspondant au / <i>ɛ̃</i> / : INITIALE, INTERNE ou FINALE (facteur de référence). Par exemple, dans <i>bientôt</i> , le son / <i>ɛ̃</i> / est interne.
Longueur	Nombre de lettres dans le mot.

<b>Variabiles</b>	<b>Description</b>
Fréquence	Fréquence du mot selon le lexique ÉQOL (Stanké et al., 2019), transformée sur l'échelle SFI (Lété et al., 2004) afin que la distribution s'approche d'une courbe normale.
Consistance	Consistance phonème-graphème du mot, d'après le lexique ÉQOL. Cet indice quantifie la fréquence des correspondances phonèmes-graphèmes du mot (voir Peereman et Stanké, 2022). Nous avons utilisé les consistances calculées sur les <i>tokens</i> .

### 3.3 Participants

L'échantillon minimal acceptable d'une recherche corrélacionnelle devrait être de 30 participants (Gay et al., 2012), ce que respecte la présente recherche. Plus précisément, le nombre de mots comportant le phonème /*ẽ*/ dans le corpus de Stanké et al. (2020) était de 5049 observations (523 élèves, 73 mots), chaque observation correspondant à un mot orthographié par un élève. Une analyse de puissance statistique préalable suggérait une taille d'échantillon minimale de 808 observations et que chacun des élèves ait orthographié au moins 2 mots de la liste. Les données des 14 élèves n'ayant orthographié qu'un seul mot de la liste ont été exclues de l'étude, portant ainsi le nombre d'observations à 5035.

Après le retrait de ces 14 participants, les données portent sur 509 élèves provenant de 31 écoles du Québec (moyenne : 16,4 participants par école, écart-type : 7,0). Le tableau 3.2 présente la répartition des élèves selon le genre et le profil de difficulté en orthographe. Chaque élève a vu, en moyenne, 9,9 mots (écart-type : 3,0) sur une possibilité de 73 mots. Chaque mot a été orthographié, en moyenne, par 69,0 élèves (écart-type : 12,7). Cette particularité liée au nombre de mots testés par les participants a influencé notre choix d'analyse.

Les élèves faibles orthographieurs sont ceux qui se sont classés sous le 10<sup>e</sup> percentile, en orthographe phonologique ou lexicale, d'une dictée de quarante mots et non-mots élaborée par Stanké et ses collègues (2020), soit 8 non-mots, 8 mots irréguliers, 5 mots consistants et 19 mots inconsistants.

**Tableau 3.2 Nombre d'élèves par sexe et par profil**

<b>Genre</b>	<b>Bons orthographieurs</b>	<b>Faibles orthographieurs</b>	<b>Total</b>
Garçons	197	52	249
Filles	221	38	259
Sans réponse	1	0	1
Total	419	90	509

Le choix du groupe d'âge se base, entre autres, sur le modèle de Seymour (1997). Ce dernier précise que la double fondation de son système d'écriture, soit les traitements logographique et alphabétique, se développe vers l'âge de sept ans; viennent ensuite les traitements orthographique et morphographique (Seymour et Evans, 1999). De plus, les résultats de la recherche doctorale de Hazard (2009) confirment que 99 % des élèves offriront des réponses qui sont plausibles phonologiquement dès la troisième année. Cette recherche s'intéressant aux traitements orthographique et morphographique, le groupe d'âge retenu a donc été de 8 ans, ce qui correspond à la troisième année du primaire au Québec et au cours élémentaire de 2<sup>e</sup> année en France.

Afin de s'assurer de la diversité de l'échantillonnage, notre recherche s'est déroulée dans des milieux socioéconomiques variés. Le tableau 3.3 les présente, le critère 10 étant accordé au milieu socioéconomique le plus défavorisé.

**Tableau 3.3 Nombre d'élèves par indicateur du milieu socioéconomique**

Milieu socioéconomique	3	6	7	8	9	10	ND
N. d'élèves	85	127	93	38	62	96	8

*Note.* L'indice socioéconomique est disponible selon le secteur scolaire fréquenté par l'élève, soit l'école. Cette information était non disponible (ND) pour 2 des 31 écoles.

### 3.4 Déroulement de l'étude

Des phases importantes ont structuré le déroulement de la présente recherche. En effet, comme les mots du corpus n'avaient pas été donnés en dictée un même nombre de fois dans la recherche de

Stanké et al. (2020), un défi de taille s'est imposé. De plus, le nombre élevé de variables indépendantes rendait la tâche ardue. L'étude s'est déroulée en quatre phases importantes.

### **3.4.1 Phase d'approbation du projet de recherche**

La phase d'appropriation du projet de recherche a eu lieu de septembre 2017 à août 2019. L'analyse théorique des écrits sur le sujet et la production du projet doctoral ont mené à actualiser la méthodologie privilégiée.

### **3.4.2 Phase de codage et d'entrée des données pour le développement de la base de données**

Une préexpérimentation a eu lieu de septembre à décembre 2019. Un premier prétest a alors été réalisé. Il visait à analyser l'effet du nombre de difficultés par mots sur la réussite de la première à la sixième année du primaire (voir annexe A). Ce prétest a permis d'établir que les élèves du premier cycle n'arrivaient pas à orthographier les mots comportant plus de quatre facteurs de difficulté, liés aux facteurs lexicaux et sous-lexicaux, et a donc confirmé le choix de la troisième année du primaire comme groupe-cible.

Par la suite, de janvier à décembre 2020, la base de données a été créée. Pour ce, nous avons ciblé les facteurs linguistiques à partir du cadre théorique de la recherche et nous avons codé la réussite ou non des graphèmes associés au phonème /ɛ̃/. Un extrait du tableau Excel de 38 colonnes et de 5292 lignes de la base de données que nous avons bâtie et que nous avons utilisée pour réaliser les analyses est présenté en annexe (voir annexe B).

### **3.4.3 Phase d'analyse en quatre temps**

De janvier 2021 à juin 2022, trois plans d'analyse de statisticiens ont été réalisés. Bien que nous ne les ayons pas retenues, les deux premières ont permis d'améliorer notre base de données et ont mené au choix d'un modèle d'analyse statistique. La proposition du modèle d'analyse qui a été retenue est celle qui nous a semblé la plus susceptible de bien tester nos hypothèses afin d'engendrer de nouvelles connaissances scientifiques. Afin de répondre le plus correctement possible à la question de recherche et d'en tirer les meilleures conclusions, une phase d'analyse en quatre temps a été appliquée. Dans un premier temps, une vérification de la base de données a été

faite. Puis, nous avons fait une vérification des facteurs significatifs et non significatifs. S'en est suivie une vérification des plans d'analyse. Enfin, une production d'analyses répondant aux hypothèses soulevées a été réalisée.

Un premier plan d'analyse statistique a été fait de janvier à avril 2021. Des analyses univariées et multivariées ont alors mené à 87 pages de résultats (voir annexe C) que la chercheuse a analysées de mai 2021 à août 2021. Malgré l'évidente qualité du traitement statistique avec le logiciel Statistical Package for the Social Sciences (SPSS) (IBM Corp., 2020), trop de biais demeuraient, notamment du côté des intervalles de confiance. Cette analyse statistique a tout de même permis de constater que l'année scolaire à laquelle le mot aurait pu être enseigné selon la liste orthographique à l'usage des enseignantes et des enseignants (ministère de l'Éducation et de l'Enseignement supérieur, 2014) n'était pas un facteur significatif, ce qui nous a incités à retirer ce facteur pour les analyses subséquentes, d'autant qu'il était trop hypothétique, les ressources enseignantes ayant souvent recours à d'autres listes orthographiques. De plus, après cette analyse, nous avons décidé de séparer la dérivabilité des morphèmes en deux catégories, soit « in » et « ain », afin de préciser nos résultats.

En mars 2022, nous avons étudié un plan d'analyse par arbres de régression. Ce deuxième plan a permis d'analyser l'interaction des facteurs linguistiques retenus (voir annexe D). Nous avons analysé tous les nœuds des arbres afin de cibler la réussite exacte des graphèmes. Cette étape a permis d'améliorer la clarté de notre fichier Excel de façon considérable. Toutefois, le nombre peu élevé de faibles orthographes ne permettait pas d'obtenir des résultats précis avec ce traitement statistique. De plus, ce modèle ne permettait pas de contrôler adéquatement la convergence. Force a donc été d'admettre que nous avions des défis de taille, notamment le nombre élevé de variables retenues, le fait que les mots n'avaient pas tous été orthographiés le même nombre de fois et le nombre peu élevé de faibles orthographes en comparaison avec le nombre de bons orthographes.

En mai 2022, nous avons appliqué un troisième plan d'analyse, soit un modèle multiniveaux, permettant de contrôler la composante « élèves ». Le modèle retenu a réussi à tenir compte du fait que les élèves ne venaient pas tous du même milieu et qu'ils n'avaient pas tous vu les mêmes mots puisque ce modèle est conçu pour tenir compte des cellules dans lesquelles il y a moins d'effectifs

et pour ramener les résultats vers la moyenne. C'est toutefois à ce moment que nous avons constaté que les données liées à la consistance des graphèmes, que nous avons puisées dans une autre source, étaient erronées. Une appropriation d'ÉQOL-Infra (Peereman et Stanké, 2022) a alors été nécessaire. Cela a permis d'identifier les caractéristiques distributionnelles des correspondances phonographologiques et des unités lexicales dans le lexique de l'école primaire québécoise et d'apporter certaines corrections à la base de données (voir annexe E).

#### **3.4.4 Phase de rédaction des chapitres sur les résultats et productions de la thèse**

Enfin, les chapitres de résultats et de discussion ont été rédigés de juin 2022 à septembre 2023. En somme, le déroulement de la recherche, dont les phases se sont étalées de septembre 2017 à aout 2023, a permis de préciser la méthode d'analyse de données. Ce déroulement a impliqué un engagement important de la part de la chercheuse, qui a discuté de chaque décision avec les statisticiens afin de s'assurer de la fiabilité de la présente recherche.

### **3.5 Instruments de collecte de données**

L'instrument de collecte de données est une dictée de mots en contexte de phrases. Le corpus utilisé est l'échelle québécoise d'acquisition en orthographe lexicale (ÉQOL) (Stanké et al., 2019), qui est constituée de 1200 mots. Les mots choisis sont les plus fréquents par niveau scolaire des ouvrages suivants de la maison d'édition Erpi : 28 manuels scolaires (mathématiques, univers social et français), 24 livres de littérature jeunesse et 90 extraits de littérature jeunesse. Ces 1200 mots ont été collectés en trois temps distincts : mai et septembre 2015 ainsi que mai 2016. Comme l'année scolaire venait tout juste d'être amorcée et qu'il était important que les élèves aient eu l'occasion d'être en contact avec les mots de leurs manuels scolaires et de leurs livres de littérature pendant une année scolaire, les élèves qui étaient en quatrième année en septembre 2015 ont été considérés comme des élèves de troisième année. Pendant trois semaines consécutives, à partir d'un corpus de 62 dictées de 20 phrases chacune, 10 dictées de 20 mots cibles à orthographier, en contexte de phrases (par exemple, « Le chien gruge un os ». « Écris le mot *chien* »), ont été données à un minimum de 50 élèves par niveau scolaire. Les dictées ont été administrées par les enseignants à partir de consignes précises fournies par les chercheurs. La méthode choisie s'inspire de celle développée par Pothier et Pothier (2004). Pour développer la base de données ÉOLE en France,

ces chercheurs ont jugé que chaque mot devait être orthographié par un minimum de 40 élèves par niveau pour que ce soit recevable.

Le corpus de la présente thèse est constitué de 73 mots. Au total, 363 mots des 1200 mots du corpus ÉQOL, contenaient le phonème /ɛ̃/. Parmi ceux-ci, 75 avaient été donnés en dictée. L'un de ces mots, soit le mot « amérindien », a été traité comme deux mots puisqu'il contenait deux fois le phonème /ɛ̃/, menant le total à 76 mots. Les trois mots suivants ont été retirés : « ainsi », « européens » et « symbole », car le pourcentage de réussite des graphèmes associés au phonème /ɛ̃/ était respectivement de 7,8 %, 9,8 % et 21,2 % pour les bons orthographieurs, mais de 0 % pour les faibles orthographieurs. Une nouvelle cotation de l'orthographe des 73 mots contenant le phonème multigraphémique /ɛ̃/ a donc été réalisée dans le cadre de la présente recherche afin de tenir compte de la réussite orthographique des différents graphèmes dans les mots. Chaque mot a été orthographié, en moyenne, par 69 élèves pour un total de 5035 productions orthographiques analysées.

Le corpus final retenu est présenté dans le tableau 3.4. Les mots y sont présentés en fonction des différents facteurs sous-lexicaux étudiés. Divers facteurs linguistiques ont été pris en compte dans le cadre de cette recherche. Les facteurs lexicaux, qui s'intéressent aux mots entiers, ont été les suivants : la fréquence, la longueur et la consistance du mot. Les données concernant la consistance proviennent d'ÉQOL-Infra (Peereman et Stanké, 2022) et les données concernant la fréquence et la longueur du mot (nombre de lettres) proviennent d'ÉQOL (Stanké et al., 2019). Les facteurs sous-lexicaux, c'est-à-dire liés aux graphèmes, ont été les suivants : le type de graphème (« in », « ain », « ein », « en », « im »), la consistance phonographémique, la position du graphème (début, milieu, fin), les régularités morphologiques dérivationnelles (certAIN/certAIne; mAIN/mAnuel fIN/fINale) ainsi que les régularités graphotactiques liées au positionnement (« in » en début de mot sauf dans le cas du mot « ainsi » : « INdicatif »), aux graphèmes contextuels (« IMPossible ») et au voisinage du graphème par une semi-voyelle (lOin, rIen).

**Tableau 3.4 Liste des mots du corpus par facteurs sous-lexicaux**

<b>Graphèmes</b>	<b>Mots</b>
Graphème « in » non associé aux régularités graphotactiques ou aux régularités morphologiques dérivationnelles retenues	Afin, amérindiens, cinq, cinquante, creuse-méninges, enfin, prince, princesse, principal, principaux, printemps, province, singulier, vingt
Graphème « ain » non associé aux régularités graphotactiques ou aux régularités morphologiques dérivationnelles retenues	Ainsi (mot retiré), demain, maintenant
Graphème « ein »	Atteindre, peinture
Graphème « en » non lié à la régularité graphotactique du voisinage du graphème « en » par une semi-voyelle (le « i » ou le « y »)	Européens (mot retiré)
Graphème « in » lié à la régularité graphotactique de la position en début de mot	Incas, indicatif, indices, indique, industrie, inférieur, infinitif, informations, inscrits, insectes, intendant, intérieur, intertitres, intrus, inventions
Graphème « in » lié à la régularité graphotactique du voisinage par une semi-voyelle (le « o »)	Besoin, coin, loin, moins, points
Graphème « en » lié à la régularité graphotactique du voisinage par une semi-voyelle (le « i » et le « y »)	Amérindiens, appartient, bien, bientôt, cartésien, chien, combien, contient, devient, lien, obtient, rien, vient, moyen
Graphème « im » lié à la régularité graphotactique contextuelle du « m » devant le « b » et le « p »	Imparfait, important, impossible, simple
Graphème « ym » lié à la régularité graphotactique contextuelle du « m » devant le « b » et le « p »	Symbole (mot retiré)
Graphème « in » lié à la régularité morphologique dérivationnelle vers un mot contenant un « i »	Chemin, dessin, féminin, fin, lutin, marins, masculin, matin, médecin, sapin
Graphème « ain » lié à la régularité morphologique dérivationnelle vers un mot contenant un « a » ou « ai »	Certains, humain, main, pain, soudain, terrain, train

La cotation s'est basée uniquement sur les différents graphèmes associés au phonème /ɛ̃/. Le score de 1 a été donné lorsque le graphème associé au phonème /ɛ̃/ dans chaque mot était adéquat et le score de 0 a été donné lorsqu'il était erroné. Les autres erreurs produites dans les mots n'ont pas été comptabilisées. Dans le cas du mot « vingt », par exemple, le graphème « in » de la production « vintg » a été coté comme étant réussi puisqu'il était correctement orthographié. Il en va de même pour les semi-voyelles précédant le phonème /ɛ̃/ dans « oin » et dans « ien ». Enfin, le graphème « en » du mot « amérindien », orthographié « amérindillen », a été coté comme réussi puisqu'il a été correctement orthographié.

Comme c'est le cas dans la langue française, le pourcentage des mots retenus pour la présente recherche a varié selon le nombre de mots existant pour chacun des graphèmes associés au phonème /ɛ̃/. Le tableau 3.5 présente la consistance de chacun des graphèmes associés au phonème /ɛ̃/ selon la position de chacun et explique, par le fait même, le peu de mots de notre recherche qui sont associés à certaines consistances plus faibles. La consistance est ici calculée à partir du nombre de mots dans lesquels le phonème /ɛ̃/ est représenté dans le corpus ÉQOL (Perremen et Stanké, 2022). Nous remarquons que le graphème « ym », que nous avons rejeté à cause de son taux d'échec à 100 % pour les élèves faibles orthographeurs, est consistant à 12 %, ce qui signifie que le phonème /ɛ̃/ ne s'écrit « ym » que pour 12 % des mots. Cela s'explique par le fait que le calcul de la consistance du graphème tient compte de la fréquence des mots de l'ensemble des manuels et des livres du primaire de l'étude. Dans le corpus utilisé pour former la base lexicale ÉQOL, le mot « symbole » est le plus fréquent au premier cycle du primaire, notamment parce qu'il apparaît dans les manuels de mathématiques. Le deuxième mot le plus fréquent est « olympique », qui apparaît pour la première fois en sixième année, suivi de « olympiades », qui apparaît en quatrième année. Les quatre autres mots, comme « tympan » et « symptômes », sont peu fréquents. Nous constatons donc que la consistance du « ym » est très peu élevée en troisième année.

**Tableau 3.5 Consistance des graphèmes selon la position qu'ils occupent dans les mots – Adapté d'ÉQOL-Infra (Peereman et Stanké, 2022)**

Phonème	Graphème	Exemple	Consistance en position initiale	Consistance en position interne	Consistance en position finale
/ɛ̃/	in	indice	80,808	55,041	21,987
/ɛ̃/	im	impact	14,844	9,323	
/ɛ̃/	en	chien		6,768	31,545
/ɛ̃/	ain	main	4,348	10,152	7,815
/ɛ̃/	ein	plein		5,456	0,486
/ɛ̃/	ym	symbole		12,224	
/ɛ̃/	yn	synthèse		0,967	
/ɛ̃/	aim	faim			0,243
/ɛ̃/	in	coïncidence		0,069	
/ɛ̃/	ens	viens			13,532
/ɛ̃/	ins	lapins			4,724
/ɛ̃/	ains	humains			6,402
/ɛ̃/	ent	tient			9,779
/ɛ̃/	eint	éteint			0,706
/ɛ̃/	eins	freins			0,11
/ɛ̃/	eints	teints			0,066
/ɛ̃/	int	obtint			0,044
/ɛ̃/	aint	saint			2,406
/ɛ̃/	ingt	vingt			0,155

*Note* : Ce tableau tient compte des morphèmes flexionnels et dérivationnels, alors que notre recherche n'en a pas tenu compte pour la cotation des résultats.

Puisque le corpus était de 1200 mots, les élèves n'orthographiaient pas toujours les mêmes mots. Les tableaux 3.6 et 3.7 indiquent le nombre de mots et d'observations<sup>7</sup> orthographiques selon le graphème et sa position dans le mot ainsi que selon le profil des élèves.

<sup>7</sup> Tout au long du présent article, nous employons « observation » pour désigner la réponse comportementale d'un élève effectuant la tâche sur un mot.

**Tableau 3.6 Nombre de mots comportant chacun des graphèmes analysés et nombre d’observations selon le graphème et selon le profil des élèves**

Graphèmes	Nb mots	Nb observations	
		Bons orthographieurs	Faibles orthographieurs
EIN	2	100	30
AIN	2	112	20
EN	14	822	163
IM	4	279	35
IN	34	1949	389
AIN-dérivable	7	393	80
IN-dérivable	10	552	111

**Tableau 3.7 Nombre de mots comportant chacun des graphèmes analysés et nombre d’observations selon le positionnement du graphème dans le mot et selon le profil d’élèves**

Positionnement	Nb mots	Nb observations	
		Bons orthographieurs	Faibles orthographieurs
Initial (ex.: <i>im<u>port</u>ant</i> )	18	1021	167
Interne (ex.: <i>pro<u>vin</u>ce</i> )	17	1022	206
Final (ex.: <i>lut<u>in</u></i> )	38	2164	455

### 3.6 Plan d’analyse

Un traitement statistique de régression logistique a servi à confirmer un lien d’effet (Bissonnette et al., 2010) et a permis de nous adapter à la variabilité naturelle et d’évaluer les effets des multiples variables sur la variable dépendante. Plus précisément, ces analyses de régression logistique multiniveaux ont été réalisées avec la librairie *lme4* pour R (Bates et al., 2015). De telles analyses permettent, entre autres, d’atténuer de possibles biais qui pourraient être attribuables à des

caractéristiques des items et des élèves, caractéristiques pouvant également influencer la réussite de l'orthographe des phonèmes multigraphémiques. En outre, la modélisation multiniveaux demeure fiable lorsque les observations ne sont pas équilibrées entre les groupes ou conditions (El Sanharawi et Naudet, 2013), ce qui est le cas pour la présente étude. Lorsque deux variables sont corrélées, le résultat est un coefficient de corrélation qui indique la taille et la relation entre les variables à partir d'un nombre décimal se situant entre -1,00 et 1,00; si le coefficient se situe près de 0, les variables ne sont pas corrélées; jusqu'à +0,35 et -0,35, la corrélation est faible; entre +0,35 et +0,65 ou -0,35 et -0,65, elle est modérée; entre +0,65 et 1,00 ou -0,65 et -1,00, elle est forte (Gay et al.). Le plan d'analyse est expliqué plus en détail dans la section « traitement statistique » du chapitre 4.

Il n'y avait pas de données manquantes pour les variables utilisées dans le modèle. Les variables continues (longueur, fréquence et consistance du mot) ont été préalablement centrées et normées afin d'être sur la même échelle. Les variables indépendantes qui ont été contrôlées sont les suivantes : âge, sexe, école, classe et statut linguistique des élèves.

Divers outils et critères permettent de confirmer ou d'infirmer les hypothèses de la présente recherche. Ils sont présentés dans le tableau 3.8.

**Tableau 3.8 Critères infirmant ou confirmant les hypothèses, en fonction des outils**

Hypothèses et outils	Critères confirmant ou infirmant les hypothèses
<p><b>Première hypothèse : Les facteurs lexicaux associés à la fréquence, à la longueur (nombre peu élevé de lettres) et à la consistance du mot ont un effet positif sur l'acquisition du phonème multigraphémique /ẽ/ dans les productions orthographiques des élèves québécois bons orthographes de la troisième année du primaire.</b></p>	
<p>La fréquence du mot a un effet positif sur l'acquisition du /ẽ/ chez les bons orthographes (fréquence établie à partir de la base de données ÉQOL)</p>	<p>✓ Une performance moyenne de 75 % et + dans l'orthographe des graphèmes associés au phonème /ẽ/ par les élèves bons orthographes permet de valider tous les critères de l'hypothèse 1</p>

Hypothèses et outils	Critères confirmant ou infirmant les hypothèses
La longueur du mot (nombre peu élevé de lettres) a un effet positif sur l'acquisition du /ɛ̃/ chez les bons orthographes (longueur établie à partir du nombre de lettres des mots)	✓ Idem
La consistance du mot a un effet positif sur l'acquisition du /ɛ̃/ chez les bons orthographes (consistance du mot établie à partir d'ÉQOL)	✓ Idem
<p><b>Deuxième hypothèse : Certains facteurs sous-lexicaux (types de graphèmes, consistance phonographémique, positionnement du graphème et régularités morphologiques dérivationnelles) ont un effet positif et certains autres facteurs sous-lexicaux (contexte d'utilisation et voisinage du graphème) ont un effet négatif sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois bons orthographes de la troisième année du primaire.</b></p>	
Le type de graphème et sa consistance ont un effet positif sur l'acquisition du /ɛ̃/ chez les bons orthographes (critère évalué à partir de la consistance des graphèmes associés au phonème /ɛ̃/ dans ÉQOL-Infra)	✓ Une performance moyenne de 75% et + dans l'orthographe des graphèmes associés au phonème /ɛ̃/ par les élèves bons orthographes permet de valider tous les critères de l'hypothèse 2
Le positionnement du graphème a un effet positif sur l'acquisition du /ɛ̃/ chez les bons orthographes (critère évalué à partir du positionnement initial, interne ou final du graphème associé au phonème /ɛ̃/)	✓ Idem
Les régularités morphologiques dérivationnelles ont un effet	✓ Idem

Hypothèses et outils	Critères confirmant ou infirmant les hypothèses
<p>positif sur l'acquisition du /<math>\tilde{\epsilon}</math>/ chez les bons orthographes (critère évalué dans les mots du corpus se terminant en « in » et ayant une dérivation possible en « i » ainsi que dans les mots se terminant en « ain » ayant une dérivation possible en « a » ou en « ai »)</p>	
<p>Le contexte d'utilisation du graphème a un effet négatif sur l'acquisition du /<math>\tilde{\epsilon}</math>/ chez les bons orthographes (critère évalué dans les mots du corpus dont le « n » se transforme en « m » devant le « b » et le « p »)</p>	<p>✓ Idem</p>
<p>Le voisinage du graphème a un effet négatif sur l'acquisition du /<math>\tilde{\epsilon}</math>/ chez les bons orthographes (critère évalué dans les mots du corpus contenant une semi-voyelle)</p>	<p>✓ Idem</p>
<hr/>	
<p><b>Troisième hypothèse : Les facteurs lexicaux associés à la fréquence, à la longueur (nombre peu élevé de lettres) et à la consistance du mot ont un effet positif sur l'acquisition du phonème multigraphémique /<math>\tilde{\epsilon}</math>/ dans les productions orthographiques des élèves québécois faibles orthographes de la troisième année du primaire</b></p> <hr/>	

Hypothèses et outils	Critères confirmant ou infirmant les hypothèses
<p>La fréquence du mot a un effet positif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs (fréquence établie à partir de la base de données ÉQOL)</p> <p>La longueur du mot (nombre peu élevé de lettres) a un effet positif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs (longueur établie à partir du nombre de lettres des mots)</p>	<p>✓ Une performance moyenne de 75% et + dans l'orthographe des graphèmes associés au phonème /ɛ̃/ par les élèves faibles orthographieurs permet de valider tous les critères de l'hypothèse 3</p> <p>✓ Idem</p>
<p>La consistance du mot a un effet positif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs (consistance du mot établie à partir d'ÉQOL)</p>	<p>✓ Idem</p>
<p><b>Quatrième hypothèse : Certains facteurs sous-lexicaux (types de graphèmes, consistance phonographémique et positionnement du graphème) ont un effet positif et certains autres facteurs sous-lexicaux (voisinage du graphème, contexte d'utilisation du graphème et régularités morphologiques dérivationnelles) ont un effet négatif sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois faibles orthographieurs de la troisième année du primaire.</b></p>	
<p>Le type de graphème et sa consistance ont un effet positif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs (critère évalué à partir de la consistance des graphèmes associés au phonème /ɛ̃/ dans ÉQOL-Infra)</p> <p>Le positionnement du graphème a un effet positif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs (critère évalué à partir du positionnement initial, interne)</p>	<p>✓ Une performance moyenne de 75 % et + dans l'orthographe des graphèmes associés au phonème /ɛ̃/ par les élèves faibles orthographieurs permet de valider tous les critères de l'hypothèse 4</p> <p>✓ Idem</p>

Hypothèses et outils	Critères confirmant ou infirmant les hypothèses
<p>ou final du graphème associé au phonème /ɛ̃/)</p> <p>Les régularités morphologiques dérivationnelles ont un effet négatif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs (critère évalué dans les mots du corpus se terminant en « in » et ayant une dérivation possible en « i » ainsi que dans les mots se terminant en « ain » ayant une dérivation possible en « a » ou en « ai »)</p>	<p>✓ Idem</p>
<p>Le contexte d'utilisation du graphème a un effet négatif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs (critère évalué dans les mots du corpus dont le « n » se transforme en « m » devant le « b » et le « p »)</p>	<p>✓ Idem</p>
<p>Le voisinage du graphème a un effet négatif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs (critère évalué dans les mots du corpus contenant une semi-voyelle)</p>	<p>✓ Idem</p>

### 3.7 Enjeux déontologiques

La recherche de Stanké et ses collaborateurs (2019), qui est à la base de la présente thèse, a été approuvée par le comité d'éthique de recherche en santé de l'Université de Montréal. Le principal moyen pour respecter l'éthique de la recherche de Stanké et ses collaborateurs est le respect de l'anonymat du participant. Aucune information n'est connue sur l'enfant participant, mis à part son

sexe biologique, son âge, son niveau scolaire, son école et le milieu socioéconomique de son école. Une approbation éthique par le comité d'éthique de la recherche de l'UQO a aussi été émise pour la présente recherche doctorale (voir annexe F).

En somme, nous nous sommes assuré que la méthodologie de recherche quantitative retenue permette d'établir les facteurs linguistiques lexicaux et sous-lexicaux qui ont un effet sur l'acquisition du phonème multigraphémique / $\tilde{\epsilon}$ / dans les productions orthographiques des élèves québécois faibles et bons orthographieurs de la troisième année du primaire. Une recherche quantitative de nature hypothéticodéductive a été retenue afin de confirmer ou d'infirmer nos hypothèses de recherche. De plus le type corrélationnel prédictif de la présente recherche a permis de déterminer l'influence des variables les unes sur les autres.

## CHAPITRE 4

### RÉSULTATS

La présente recherche a pour objectif d'identifier les facteurs linguistiques qui favorisent l'acquisition des phonèmes multigraphémiques dans les productions orthographiques des élèves québécois bons et faibles orthographieurs de la troisième année du primaire. Plus précisément, elle cherche à établir les facteurs linguistiques lexicaux et sous-lexicaux qui ont un effet sur l'acquisition du phonème multigraphémique /*ẽ*/ dans les productions orthographiques des élèves québécois faibles et bons orthographieurs de la troisième année du primaire. Les facteurs lexicaux ciblés sont la consistance, la fréquence et la longueur du mot (nombre peu élevé de lettres). Les facteurs sous-lexicaux retenus sont les types de graphèmes, la consistance phonographémique, les régularités graphotactiques (positionnement, contexte d'utilisation et voisinage du graphème) et les régularités morphologiques dérivationnelles. La présentation des résultats correspond aux facteurs lexicaux et sous-lexicaux identifiés.

Le traitement statistique et les statistiques descriptives sont d'abord présentés. Puis, la section « analyses multiniveaux » permet de présenter les variables significatives et de retirer celles qui ne le sont pas. Ensuite, la section « comparaisons multiples » présente les résultats liés aux facteurs sous-lexicaux du positionnement du graphème et de du type de graphème chez les bons et chez les faibles orthographieurs, le type de graphème incluant sa consistance, son contexte d'utilisation, son voisinage et à sa dérivabilité. Cette section présente aussi les résultats liés au seul facteur lexical significatif, soit la longueur du mot (nombre peu élevé de lettres). Enfin, à partir des résultats de la présente recherche doctorale, le tableau des critères et des hypothèses en fonction des outils présenté au chapitre 3 est mis à jour, ce qui permet de faire une synthèse des hypothèses confirmées ou infirmées.

#### 4.1 Traitement statistique

Un traitement statistique a servi à confirmer un lien d'effet. Des analyses de régression logistique multiniveaux ont été réalisées. Ces analyses statistiques ont permis d'éliminer les facteurs non contrôlés. L'isolement de la contribution unique permet de s'assurer qu'aucun autre facteur n'est responsable des effets qui sont observés sur la variable dépendante. C'est là tout l'intérêt de la

régression logistique multiniveaux (aussi appelée mixte ou hiérarchique) pour cette recherche doctorale. En effet, elle permet de contrôler pour plus d'un facteur (les élèves ou le mot, par exemple), elle fonctionne même quand les données ne sont pas balancées et elle permet de faire des contrastes, c'est-à-dire des comparaisons entre des groupes, qui tiennent compte du débalancement des données (El Sanharawi et Naudet, 2013). Cette méthode convient parfaitement puisque les élèves n'ont pas tous vu les mêmes mots.

D'abord, il a fallu faire converger un modèle afin de trouver celui qui pouvait nous permettre de vérifier les hypothèses tout en ayant comme facteurs le profil et la régularité, avec interaction entre les deux. Cela a été fait dans le but de voir si la régularité affecte différemment les élèves selon leur profil et comment elle le fait.

Des analyses multiniveaux ont été conduites pour examiner l'effet des variables d'intérêt sur la réussite de la tâche. Le modèle « nul » comprenait seulement la composante aléatoire du modèle (élève), sans variable explicative; son pseudo-R<sup>2</sup> était d'environ 18 %. Comme montré au tableau 4.1, les modèles « théorique » et « minimal » avaient une performance similaire à l'égard du pseudo-R<sup>2</sup>, soit environ 40 %. Nous avons retenu le modèle « minimal » comme modèle final puisque, tout en étant moins coûteux, il était légèrement plus performant à l'égard des critères d'information d'Akaike et bayésiens (AIC et BIC).

**Tableau 4.1 Comparaison des modèles multiniveaux**

	« Nul »	« Théorique »	« Minimaliste » (Final)
(Intercept)	2,05 (0,06) ***	-1,65 (0,34) ***	-1,44 (0,31) ***
Profil – EN DIFF.		-0,23 (0,70)	-0,87 (0,56)
Graphème – AIN		2,68 (0,37) ***	2,55 (0,35) ***
Graphème – EN		3,83 (0,37) ***	3,58 (0,34) ***
Graphème – IM		1,92 (0,34) ***	1,86 (0,32) ***
Graphème – IN		4,11 (0,30) ***	3,93 (0,28) ***
Graphème – AIN-Dérivable		2,86 (0,35) ***	2,70 (0,33) ***
Graphème – IN-Dérivable		4,69 (0,40) ***	4,42 (0,37) ***
Fréquence		0,03 (0,08)	0,01 (0,06)
Longueur		-0,45 (0,08) ***	-0,49 (0,07) ***
Position – INITIALE		1,79 (0,31) ***	1,70 (0,26) ***
Position – INTERNE		0,78 (0,25) **	0,69 (0,20) ***
Consistance		-0,28 (0,08) **	-0,18 (0,07) **
Fréquence:Longueur		0,02 (0,06)	
Profil:Graphème – AIN		-2,12 (0,92) *	-1,66 (0,86)
Profil:Graphème – EN		-1,57 (0,73) *	-0,88 (0,60)
Profil:Graphème – IM		-1,69 (0,81) *	-1,68 (0,72) *
Profil:Graphème – IN		-0,56 (0,64)	-0,17 (0,60)
Profil:Graphème – AIN-Dérivable		-0,83 (0,73)	-0,35 (0,62)
Profil:Graphème – ING-Dérivable		-0,14 (0,83)	0,58 (0,72)
Profil:Fréquence		-0,17 (0,16)	
Profil:Longueur		-0,17 (0,16)	
Profil:Position – INITIALE		-0,29 (0,58)	
Profil:Position – INTERNE		-0,34 (0,45)	
Profil:Consistance		0,38 (0,16) *	
Profil:Fréquence:Longueur		-0,01 (0,11)	
AIC <sup>a</sup>	3962,04	3305,32	3299,86
BIC <sup>a</sup>	3975,09	3481,47	3430,34
Pseudo-R2	0,18	0,41	0,40
Var : Élève (Intercept)	0,72	0,80	0,77

*Note.* Régression logistique multiniveaux effectuée avec l’algorithme BOBYQA (Powell, 2009). 5035 observations, 509 élèves. AIC : critère d’information d’Akaike; BIC : critère d’information bayésien; AUC : aire sous la courbe ROC. Pseudo-R2 de Nagelkerke. La valeur entre parenthèses est l’erreur standard. \* p < 0,05; \*\* p < 0,01; \*\*\* p < 0,001.

Nous avons testé un modèle « nul » (sans la composante fixe) ainsi que deux modèles qui incluaient l'ensemble des variables d'intérêt. Les analyses multiniveaux ont été effectuées avec la librairie *lme4* pour R (Bates et al., 2015). Leurs équations sont données au tableau 4.2. Le premier modèle s'appuie sur des éléments de théorie linguistique de corpus afin de déterminer les interactions entre les variables. Le second modèle n'inclut que l'interaction entre le profil en orthographe et le graphème, qui est requise pour vérifier la première hypothèse de la présente étude. L'importance relative des variables explicatives a été vérifiée à l'aide du test de Wald implémenté dans la fonction *Anova* de la librairie *car* pour R (Fox et Weisberg, 2019), la statistique  $\chi^2$  produite pour chaque variable étant un indicateur de leur contribution (Thomas et al., 2008). Nous avons retenu le modèle minimaliste à cause de ses critères de performance d'Akaike et bayésiens (AIC et BIC) qui étaient légèrement plus élevés.

**Tableau 4.2 Modèles multivariés testés dans le cadre de la présente étude**

« Nul »	« Théorique »	« Minimaliste »
réussite du graphème ~ (1   élève)	réussite du graphème ~ profil * (graphème + fréquence * longueur + position + consistance) + (1   élève)	réussite du graphème ~ profil * graphème + fréquence + longueur + position + consistance + (1   élève)

*Note.* Pour une explication de la notation employée, voir Bates et al. (2015).

Dans le but d'examiner davantage l'influence du type de graphème et du profil en orthographe sur la réussite, nous avons mené une série de comparaisons *post hoc* (contrastes) par méthode des moyennes marginales. L'implémentation était celle de la librairie *emmeans* pour R (Lenth et al., 2022). Cette méthode s'appuie sur le modèle multiniveaux pour fournir des résultats ajustés tenant compte des composantes fixes et aléatoires du modèle. Nous avons ainsi estimé la probabilité, ajustée par le modèle multiniveaux, de réussir la tâche en fonction du profil et du graphème. Les valeurs *p* associées ont été ajustées par la méthode de Benjamini et Hochberg (1995) afin de réduire la chance de fausses découvertes. Nous avons, en outre, produit un diagramme à boîtes illustrant la réussite du mot (sur l'échelle du logit, tel qu'estimé par le modèle multiniveaux) par groupe et par profil d'élève. Le logit est le ratio des chances (*odds ratio*). Dans ce contexte, il s'agit de la probabilité de réussir le mot divisé par la probabilité de ne pas le réussir exprimée sur une échelle logarithmique.

## 4.2 Statistiques descriptives

L'analyse non ajustée de la proportion de graphèmes réussis, dont le tableau 4.3 fait le sommaire, suggère d'emblée que la réussite des mots a varié selon le profil d'élèves et le type de graphème. Ce tableau précise le taux de réussite orthographique de chacun des graphèmes pour les bons et pour les faibles orthographes. Pour les deux groupes, les graphèmes IN et IN-dérivable sont ceux qui ont présenté le plus haut taux de réussite, alors que le graphème EIN est celui qui a été le moins réussi. Les principales différences observées entre les profils d'élèves ont été pour les mots contenant les graphèmes AIN et IM.

**Tableau 4.3 Statistiques descriptives de la réussite de la tâche selon le type de graphèmes et selon le profil d'élèves**

Graphème	Taux de réussite (non aj.)	
	Bons orthographes	Faibles orthographes
EIN	0,31	0,20
AIN	0,70	0,20
EN	0,88	0,61
IM	0,79	0,26
IN	0,94	0,87
AIN-dérivable	0,81	0,64
IN-dérivable	0,94	0,93

Le tableau 4.4 illustre la proportion non ajustée de réussite orthographique du graphème selon son positionnement et selon le profil d'élèves. Les résultats montrent l'importance de la position du graphème dans le mot pour sa réussite orthographique, notamment pour le positionnement interne où nous avons constaté le plus bas taux de réussite, et ce, pour les deux profils d'élèves.

**Tableau 4.4 Statistiques descriptives de la réussite orthographique du graphème selon son positionnement dans le mot et selon le profil d'élèves**

<b>Positionnement</b>	<b>Taux de réussite (non aj.)</b>	
	<b>Bons orthographieurs</b>	<b>Faibles orthographieurs</b>
Initial	0,93	0,82
Interne	0,84	0,66
Final	0,89	0,74

Une analyse des erreurs a permis d'identifier les graphèmes choisis par les élèves en cas d'erreurs. Les tableaux 4.5 et 4.6 les présentent. Le graphème « in » a été majoritairement choisi par tous les élèves lorsqu'ils ont fait des erreurs, mais dans une plus grande proportion par les faibles orthographieurs. Dans le cas où le graphème cible attendu était le « in », les bons orthographieurs ont eu tendance à choisir le graphème « ain » en cas d'erreur, alors que les faibles orthographieurs ont sélectionné le « en ».

**Tableau 4.5 Graphème choisi par les élèves en cas d'erreur des graphèmes cibles « in » et « ein » selon le profil d'élèves**

Graphème utilisé par les élèves	Pourcentage d'élèves ayant choisi le graphème erroné			
	Graphème attendu : « in »		Graphème attendu : « ein »	
	Élèves bons orthographieurs	Élèves faibles orthographieurs	Élèves bons orthographieurs	Élèves faibles orthographieurs
<b>IN</b>	0	0	68,12	83,33
<b>AIN</b>	29,93	8,62	7,25	0
<b>EN</b>	8,84	20,69	1,45	0
<b>EIN</b>	3,40	6,90	0	0
<b>Inversion</b>	2,04	3,45	1,45	0
<b>É</b>	3,40	8,62	1,45	0
<b>AN</b>	0	1,72	0	0
<b>È</b>	1,36	3,45	0	0
<b>UN</b>	1,36	0	0	0
<b>IM</b>	13,61	3,45	0	0
<b>AUTRE</b>	36,05	43,10	20,29	16,67

**Tableau 4.6 Graphème choisi par les élèves en cas d'erreur dans les graphèmes cibles « ain », « en » et « im » selon le profil d'élèves**

Graphème utilisé par les élèves	Pourcentage d'élèves ayant choisi le graphème erroné					
	Graphème attendu : « ain »		Graphème attendu : « en »		Graphème attendu : « im »	
	Élèves bons orthographieurs	Élèves faibles orthographieurs	Élèves bons orthographieurs	Élèves faibles orthographieurs	Élèves bons orthographieurs	Élèves faibles orthographieurs
<b>IN</b>	68,52	71,11	60,00	71,88	93,33	100
<b>AIN</b>	0	0	13,00	7,81	0	0
<b>EN</b>	4,63	8,89	0	0	0	0
<b>EIN</b>	2,78	0	9,00	1,56	0	0
<b>Inversion</b>	4,63	0	7,00	4,69	0	0
<b>É</b>	0	2,22	0	0	0	0
<b>AN</b>	0	0	4,00	3,13	0	0
<b>AUTRE</b>	19,44	17,78	7,00	10,94	6,67	0

La multiplicité des facteurs qui entrent en jeu a nécessité des contrôles complexes à partir d'analyses multiniveaux. Cela a permis de savoir si les résultats de notre étude sont généralisables. Les résultats de ces analyses sont présentés dans la prochaine section.

### 4.3 Analyses multiniveaux

La portion de la variance expliquée par les variables a d'abord été analysée. Le tableau 4.7 montre les résultats des tests de Wald vérifiant la contribution des variables explicatives du modèle. Le facteur lexical de la fréquence du mot, qui est associé à la première et à la troisième hypothèse, ne contribuait pas au modèle de manière statistiquement significative. La consistance du mot, aussi associé à la première et à la troisième hypothèse, avait un effet significatif, mais sa contribution au modèle était très faible. Les autres variables, de même que l'interaction entre le profil d'élèves et le graphème, contribuaient au modèle de manière statistiquement significative, considérant un seuil de  $p < 0,001$ . Sur la base des statistiques  $\chi^2$ , le type de graphème avait la plus grande contribution, suivi du profil d'élèves. La longueur du mot (nombre peu élevé de lettres) liée aux hypothèses 1 et 3, et le positionnement du graphème lié aux hypothèses 2 et 4, affichaient une contribution d'importance similaire.

**Tableau 4.7 Tests de Wald – modèle final « minimaliste »**

	<i>dl</i>	$\chi^2$	<i>p</i>
Profil de l'élève	1	75,706	<0,001
Type de graphème	6	361,157	<0,001
Fréquence du mot	1	0,038	0,8458
Longueur du mot (nombre peu élevé de lettres)	1	48,576	<0,001
Positionnement du graphème	2	46,525	<0,001
Consistance du mot	1	6,672	0,0098
Interaction profil - graphème	6	23,327	<0,001

*Note.* Degrés de liberté, chi-deux et valeur *p* de tests de Wald.

#### 4.4 Comparaisons multiples entre les bons et les faibles orthographieurs

Nous avons d'abord vérifié les probabilités de bien orthographier les graphèmes selon les profils d'élèves, les bons orthographieurs étant liés à nos deux premières hypothèses et les faibles orthographieurs étant liés aux hypothèses 3 et 4. Le tableau 4.8 présente les probabilités de réussite ajustées par la méthode des moyennes marginales. Les résultats montrent que la différence observée entre les groupes n'est pas statistiquement significative ( $p = 0,14$ ) pour le graphème « ein », qui est d'ailleurs très échoué<sup>8</sup> tant par les bons orthographieurs (environ 34 % de réussite) que par les faibles orthographieurs (environ 18 % de réussite)<sup>9</sup>. Il en est de même pour le graphème « in » dérivable qui, par ailleurs, est très réussi tant par les bons orthographieurs (environ 98 % de réussite) que par les faibles orthographieurs (environ 97 % de réussite). Une grande différence est observée pour le graphème « ain » non dérivable entre les deux groupes. En effet, les faibles orthographieurs ont 2,49 fois plus de chances de l'échouer que leurs pairs. Cette différence est aussi présente dans la réussite du graphème « im », que les faibles orthographieurs ont 3,69 fois plus de chances d'échouer que leurs pairs. La différence entre les deux profils d'élèves est peu élevée pour les graphèmes « en », « in » et « ain » dérivables.

---

<sup>8</sup> Un profil d'élèves est considéré comme ayant réussi la tâche lorsque le taux est de 75 % et plus (Pothier et Pothier, 2004; Stanké et al., 2019).

<sup>9</sup> La comparaison de la réussite des graphèmes a été réalisée grâce à la méthode des intervalles de confiance. Nous considérons alors que la non-superposition d'intervalles indique une différence statistiquement significative.

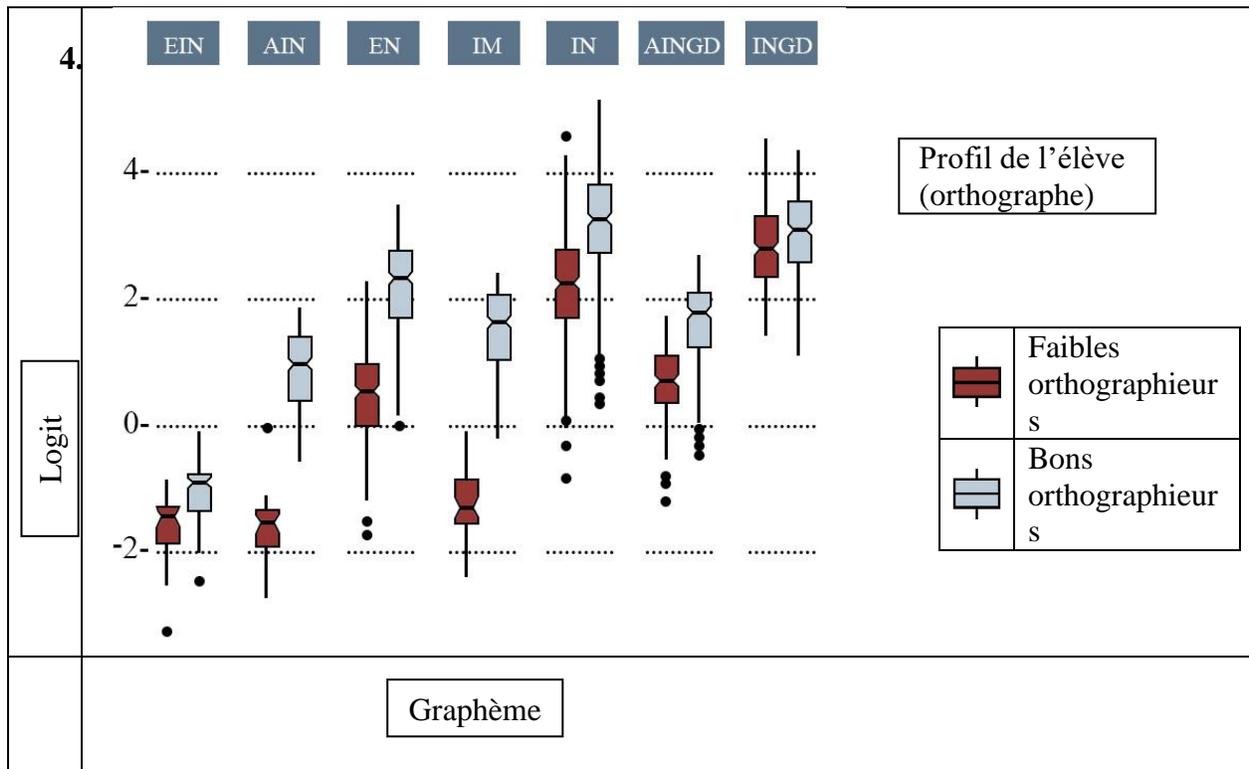
**Tableau 4.8 Contrastes entre les profils d'élèves en orthographe, par graphème**

Graphème	Prob. réponse correcte [IC 95 %]		Ratio de risque [IC 95 %]	<i>p</i>	<i>d</i>
	Bons orthogra- phieus	Faibles orthogra- phieus			
EIN	0,34 [0,24, 0,47]	0,18 [0,07, 0,38]	1,91 [1,10, 2,72]	0,14	- 0,87
AIN	0,87 [0,81, 0,92]	0,35 [0,14, 0,64]	2,49 [1,89, 3,09]	<0,001	- 2,52
EN	0,95 [0,93, 0,96]	0,77 [0,67, 0,84]	1,24 [1,15, 1,33]	<0,001	- 1,75
IM	0,77 [0,70, 0,83]	0,21 [0,10, 0,39]	3,69 [3,04, 4,34]	<0,001	- 2,55
IN	0,96 [0,95, 0,97]	0,90 [0,87, 0,93]	1,07 [1,03, 1,10]	<0,001	- 1,04
AIN – dériv.	0,89 [0,84, 0,92]	0,70 [0,56, 0,80]	1,27 [1,12, 1,42]	<0,001	- 1,22
IN – dériv.	0,98 [0,96, 0,99]	0,97 [0,93, 0,99]	1,01 [0,97, 1,04]	0,53	- 0,29

*Note.* Probabilités obtenues par modélisation multiniveaux. Ratio du risque de non-réussite orthographique. Valeurs *p* calculées par méthode des moyennes marginales, ajustées par méthode de Benjamini-Hochberg. Le *d* indique la taille d'effet *d* de Cohen. Le groupe « faibles orthographes » a été employé comme niveau de référence pour le calcul du ratio de risque et de la taille d'effet.

La méthode de comparaison graphique complète les informations fournies dans le tableau précédent. La figure 4.1 permet de visualiser les contrastes entre les deux profils d'élèves.

**Figure 4.1 Contrastes en orthographe entre les deux profils d'élèves selon les différents graphèmes**



*Note.* La boîte indique les percentiles 25 à 75, la ligne centrale est la valeur médiane. L'encoche indique un intervalle de confiance à 95 % autour de la médiane, la non-superposition de l'encoche indique une différence statistiquement significative.

Le rôle du positionnement du graphème entre les deux profils de difficulté en orthographe a été étudié. Le tableau 4.9 permet de constater que la différence la plus marquée se présente pour le phonème /ɛ̃/ en fin de mot, que les élèves faibles orthographieurs ont environ 1,72 fois plus de chances d'échouer que les élèves bons orthographieurs. En position initiale, le taux de réussite est bon, tant pour les élèves faibles orthographieurs (environ 82 %) que pour les élèves bons orthographieurs (environ 95 %) le graphème étant mieux orthographié par tous les élèves lorsqu'il est en début de mot.

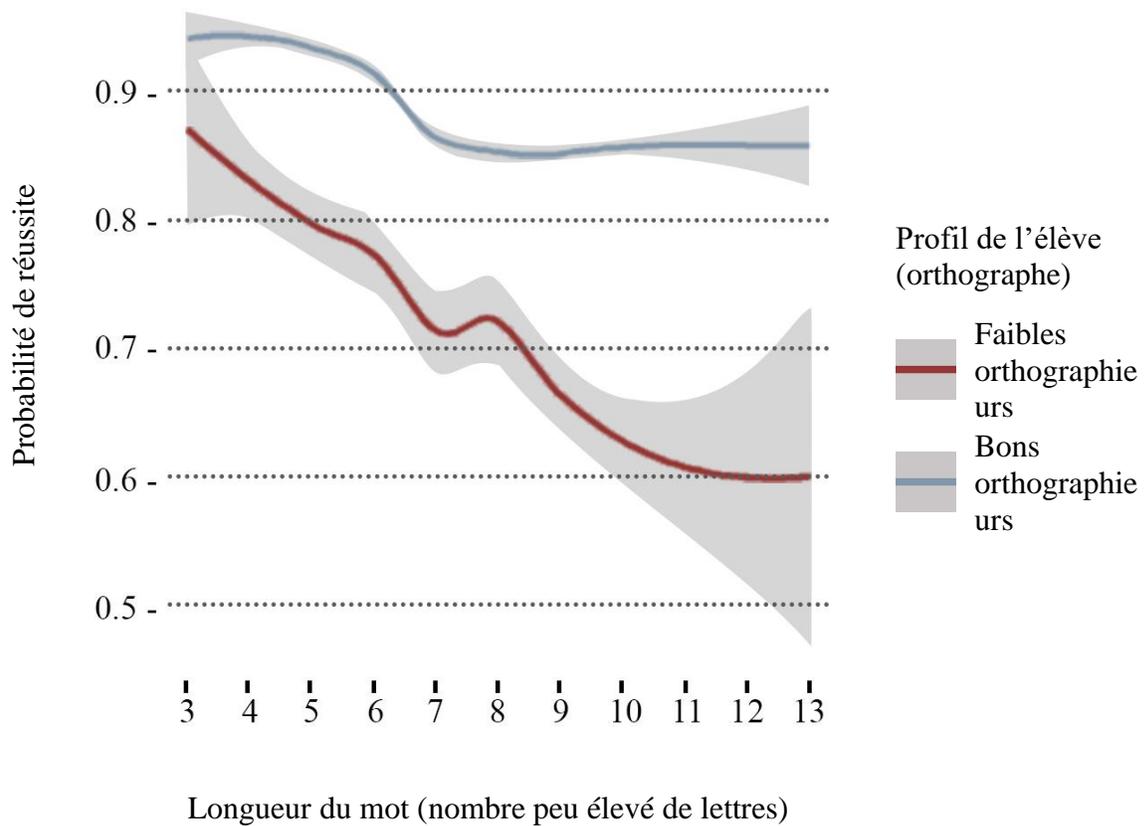
**Tableau 4.9 Contrastes entre les profils de difficulté en orthographe selon le positionnement du graphème**

Positionnement	Prob. réponse correcte [IC 95 %]		Ratio de risque [IC 95 %]
	Bons orthographieurs	Faibles orthographieurs	
Initiale	0,95 [0,93, 0,97]	0,82 [0,74, 0,88]	1,16 [1,09, 1,23]
Interne	0,88 [0,84, 0,91]	0,62 [0,52, 0,72]	1,41 [1,30, 1,52]
Finale	0,78 [0,74, 0,82]	0,45 [0,36, 0,55]	1,72 [1,62, 1,83]

*Note.* Les comparaisons ont été effectuées à partir d'un modèle multiniveaux, par méthode des moyennes marginales. Il n'y avait pas, dans ce modèle, d'interaction entre le positionnement du graphème et le groupe (faibles orthographieurs / bons orthographieurs); la significativité statistique générale pour ces contrastes était de  $p < 0,001$ , avec taille d'effet  $d = 1,46$ .

Enfin, le rôle de la longueur du mot (nombre peu élevé de lettres) entre les profils de difficulté en orthographe a été étudié. La figure 4.2 permet de constater que la longueur des mots affecte peu la réussite des graphèmes associés au phonème / $\tilde{\epsilon}$ / des élèves bons orthographieurs alors que, plus le nombre de lettres est élevé, plus les élèves faibles orthographieurs sont susceptibles de commettre des erreurs.

**Figure 4.2 Contrastes entre les profils de difficulté en orthographe selon la longueur du mot (nombre peu élevé de lettres)**



*Note.* Taux de bonnes réponses ajusté par modèle de régression logistique multiniveaux. Courbes lissées produites avec la librairie *ggplot2* et méthode LOESS (Wickham, 2016).

## 4.2 Synthèse des résultats

Il est maintenant possible de reprendre le tableau des critères et des hypothèses en fonction des outils présenté au chapitre 3 et de faire une synthèse des hypothèses confirmées ou infirmées (tableau 4.10).

**Tableau 4.10 Infirmation ou confirmation des hypothèses**

Hypothèses	Confirmation ou infirmation des hypothèses
<b>Première hypothèse : Les facteurs lexicaux associés à la fréquence, à la longueur et à la consistance du mot ont un effet positif sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois bons orthographes de la troisième année du primaire.</b>	
La fréquence du mot a un effet positif significatif sur l'acquisition du /ɛ̃/ chez les bons orthographes	✓ Infirmée
La longueur du mot (nombre peu élevé de lettres) a un effet positif significatif sur l'acquisition du /ɛ̃/ chez les bons orthographes	✓ Infirmée
La consistance du mot a un effet positif significatif sur l'acquisition du /ɛ̃/ chez les bons orthographes	✓ Infirmée
<b>Deuxième hypothèse : Certains facteurs sous-lexicaux (types de graphèmes, consistance phonographémique, positionnement du graphème et régularités morphologiques dérivationnelles) ont un effet positif et certains autres facteurs sous-lexicaux (contexte d'utilisation et voisinage du graphème) ont un effet négatif sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois bons orthographes de la troisième année du primaire.</b>	
Le type de graphème et sa consistance ont un effet positif significatif sur l'acquisition du /ɛ̃/ chez les bons orthographes	✓ Confirmée
Le positionnement du graphème a un effet positif significatif sur l'acquisition du /ɛ̃/ chez les bons orthographes	✓ Confirmée

Hypothèses	Confirmation ou infirmation des hypothèses
Les régularités morphologiques dérivationnelles ont un effet positif significatif sur l'acquisition du /ɛ̃/ chez les bons orthographieurs	✓ Confirmée
Le contexte d'utilisation du graphème a un effet négatif significatif sur l'acquisition du /ɛ̃/ chez les bons orthographieurs	✓ Infirmée
Le voisinage du graphème a un effet négatif significatif sur l'acquisition du /ɛ̃/ chez les bons orthographieurs (critère évalué dans les mots du corpus contenant une semi-voyelle)	✓ Infirmée
<b>Troisième hypothèse : Les facteurs lexicaux associés à la fréquence, à la longueur et à la consistance du mot ont un effet positif sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois faibles orthographieurs de la troisième année du primaire</b>	
La fréquence du mot a un effet positif significatif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs (fréquence établie à partir de la base de données ÉQOL)	✓ Infirmée
La longueur du mot (nombre peu élevé de lettres) a un effet positif significatif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs (longueur établie à partir du nombre de lettres des mots)	✓ Confirmée
La consistance du mot a un effet positif significatif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs (consistance du mot établie à partir d'ÉQOL)	✓ Infirmée

Hypothèses	Confirmation ou infirmation des hypothèses
<p><b>Quatrième hypothèse : Certains facteurs sous-lexicaux (types de graphèmes, consistance phonographémique et positionnement du graphème) ont un effet positif et certains autres facteurs sous-lexicaux (voisinage du graphème, contexte d'utilisation du graphème et régularités morphologiques dérivationnelles) ont un effet négatif sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois faibles orthographieurs de la troisième année du primaire.</b></p>	
Le type de graphème et sa consistance ont un effet positif significatif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs	✓ Confirmée
Le positionnement du graphème a un effet positif significatif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs	✓ Confirmée
Les régularités morphologiques dérivationnelles ont un effet négatif significatif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs	✓ Infirmée
Le contexte d'utilisation du graphème a un effet négatif significatif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs	✓ Confirmée
Le voisinage du graphème a un effet négatif significatif sur l'acquisition du /ɛ̃/ chez les faibles orthographieurs	✓ Infirmée

Ainsi, la première hypothèse, qui prédisait l'effet positif des facteurs lexicaux chez les bons orthographieurs, est infirmée. La fréquence, la longueur et la consistance du mot n'ont pas d'effet positif significatif sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois bons orthographieurs de la troisième année du primaire. Ces

facteurs ne jouent d'ailleurs aucun rôle négatif non plus. Quant à elle, la troisième hypothèse, qui prédisait l'effet positif significatif de ces mêmes facteurs lexicaux chez les faibles orthographes, elle est partiellement confirmée. La fréquence et la consistance du mot n'ont pas d'effet positif significatif sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois faibles orthographes de la troisième année du primaire et ils n'ont pas d'effet négatif non plus. Toutefois, le facteur lexical de la longueur du mot (nombre peu élevé de lettres) a un effet positif significatif sur l'acquisition du phonème multigraphémique /ɛ̃/ par les faibles orthographes. En effet, les faibles orthographes ne réussissent à obtenir un pourcentage de 75 % et plus de réussite quand ils orthographient les graphèmes associés au phonème multigraphémique /ɛ̃/ que dans les mots de moins de six lettres.

Du côté des facteurs sous-lexicaux, la deuxième hypothèse est partiellement confirmée. Le type de graphème et sa consistance ainsi que son positionnement et sa dérivabilité ont bien un effet positif significatif sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois bons orthographes de la troisième année du primaire. Toutefois, le contexte d'utilisation du graphème et son voisinage n'ont pas l'effet négatif significatif escompté. En effet, le contexte d'utilisation du graphème ne nuit pas à son acquisition et le voisinage du graphème joue même un rôle positif dans son acquisition. Enfin, la quatrième hypothèse est partiellement confirmée, elle aussi. Chez les faibles orthographes, le type de graphème et sa consistance ainsi que son positionnement ont bien un effet positif significatif et le contexte d'utilisation du graphème a bien un effet négatif significatif sur l'acquisition du phonème multigraphémique /ɛ̃/. Toutefois, la dérivabilité du graphème et son voisinage n'ont pas l'effet négatif significatif escompté. En effet, tant la dérivabilité du graphème que le voisinage du graphème ne nuisent pas à son acquisition et jouent même un rôle positif.

Ainsi, le quatrième chapitre a permis de présenter les résultats de la présente recherche doctorale. Le prochain permet d'en faire une interprétation à la lumière des recherches exposées dans le cadre théorique. De plus, il permet de comprendre la portée de ces résultats et de discuter de leur pertinence scientifique et des implications pour la pratique.

## CHAPITRE 5

### DISCUSSION

L'objectif de la présente recherche doctorale était d'identifier les facteurs linguistiques qui favorisent l'acquisition des phonèmes multigraphémiques dans les productions orthographiques des élèves québécois, bons et faibles orthographes de la troisième année du primaire. L'analyse de l'acquisition de l'orthographe des différents graphèmes du phonème multigraphémique /*ẽ*/ dans les productions orthographiques de ces élèves a révélé que plusieurs facteurs linguistiques ont un effet significatif. Dans le présent chapitre, nous discutons de ces facteurs identifiés dans les hypothèses en comparant leur effet chez les élèves bons et faibles orthographes. Tout d'abord, les résultats associés à la première et à la troisième hypothèse sont commentés. Puis, nous discutons des résultats associés à la deuxième et à la quatrième hypothèse. Ensuite, une hiérarchisation des facteurs lexicaux et sous-lexicaux dans l'acquisition du phonème multigraphémique /*ẽ*/ en fonction des deux catégories de participants (les bons orthographes et les faibles orthographes) est proposée. Puis, les implications des nouvelles connaissances pour la pratique sont expliquées. Enfin, les limites de la présente recherche ainsi que les futures recherches qu'elle appelle sont énoncées.

Afin de faire le lien entre les résultats de la présente recherche et la théorie, nous reprenons d'abord le tableau-synthèse des facteurs établis par la recherche ayant un effet sur l'acquisition de l'orthographe lexicale et sur l'acquisition des phonèmes multigraphémiques, présenté au chapitre deux. Nous le mettons à jour à partir de la présente recherche doctorale en nous concentrant uniquement sur l'acquisition du phonème multigraphémique /*ẽ*/ chez les bons et chez les faibles orthographes (tableau 5.1).

**Tableau 5.1 Synthèse des facteurs ayant un effet sur l'acquisition du phonème multigraphémique /ɛ̃/ à partir de la présente recherche doctorale**

Facteurs des recherches antérieures ayant un effet sur l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques	Facteurs de la présente recherche ayant un effet sur l'acquisition du phonème multigraphémique /ɛ̃/ chez les bons orthographieurs	Facteurs de la présente recherche ayant un effet sur l'acquisition du phonème multigraphémique /ɛ̃/ chez les faibles orthographieurs
<b>Fréquence des mots</b>		
Orthographe des mots fréquents acquise plus rapidement (Martinet et al., 2004; Mousty et Leybaert, 1999; Sprenger-Charolles et al., 1998; Stanké et al., 2019)	Prédiction de l'hypothèse 1 infirmée : La fréquence des mots n'a pas d'influence sur l'acquisition du phonème multigraphémique /ɛ̃/	Prédiction de l'hypothèse 3 infirmée : La fréquence des mots n'a pas d'influence sur l'acquisition du phonème multigraphémique /ɛ̃/
<b>Longueur des mots</b>		
Orthographe des mots comptant peu de syllabes, acquise plus rapidement (Stanké et al., 2019)	Prédiction de l'hypothèse 1 infirmée : Le nombre peu élevé de lettres n'a pas d'influence sur l'acquisition du phonème multigraphémique /ɛ̃/	Prédiction de l'hypothèse 3 confirmée : Le nombre peu élevé de lettres a une influence positive sur l'acquisition du phonème multigraphémique /ɛ̃/ lorsque les mots ont moins de six lettres
<b>Consistance des mots</b>		
Le degré de consistance d'une langue joue non seulement un rôle important dans l'acquisition de sa lecture, mais aussi dans l'acquisition de son orthographe (Caravolas 2019; Treiman 2019)	Prédiction de l'hypothèse 1 infirmée : La consistance des mots n'a pas d'influence sur l'acquisition du phonème multigraphémique /ɛ̃/	Prédiction de l'hypothèse 3 infirmée : La consistance des mots n'a pas d'influence sur l'acquisition du phonème multigraphémique /ɛ̃/

Facteurs des recherches antérieures ayant un effet sur l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques	Facteurs de la présente recherche ayant un effet sur l'acquisition du phonème multigraphémique /ɛ̃/ chez les bons orthographes	Facteurs de la présente recherche ayant un effet sur l'acquisition du phonème multigraphémique /ɛ̃/ chez les faibles orthographes
--	--	---

---

**Consistance phonographique**

---

Application de la graphie dominante par les élèves lorsque le lexique orthographique n'est pas assez riche (ex. : utilisation du graphème /o/ plutôt que /eau/) (Mousty et Leybaert, 1999)	Prédiction de l'hypothèse 2 confirmée : Graphie dominante IN majoritairement choisie en cas d'erreurs	Prédiction de l'hypothèse 4 confirmée : Graphie dominante IN majoritairement choisie en cas d'erreurs
Acquisition rapide des graphèmes consistants acontextuels (ex. : /l/ et /ou/) (Mousty et Leybaert, 1999)	Prédiction de l'hypothèse 2 confirmée : Acquisition du graphème non dérivable le plus consistant « in » à 96 % et non-acquisition du graphème le moins consistant « ein » (34 %)	Prédiction de l'hypothèse 4 confirmée : Acquisition du graphème non dérivable le plus consistant « in » à 90 % et non-acquisition du graphème le moins consistant « ein » (18 %)

---

<b>Facteurs des recherches antérieures ayant un effet sur l'acquisition de l'orthographe lexicale et des phonèmes multigraphémiques</b>	<b>Facteurs de la présente recherche ayant un effet sur l'acquisition du phonème multigraphémique /ɛ̃/ chez les bons orthographieurs</b>	<b>Facteurs de la présente recherche ayant un effet sur l'acquisition du phonème multigraphémique /ɛ̃/ chez les faibles orthographieurs</b>
---	--	---

### **Régularités graphotactiques de positionnement**

Sensibilité implicite des jeunes élèves à la position légale des lettres dans l'acquisition de l'orthographe (Pacton et al., 2001)	Prédiction de l'hypothèse 2 confirmée : Les élèves réussissent mieux le phonème multigraphémique /ɛ̃/ en début de mot (95 % au début, vs 78 % à la fin)	Prédiction de l'hypothèse 4 confirmée : Les élèves réussissent mieux le phonème multigraphémique /ɛ̃/ en début de mot (82 % au début vs 45 % à la fin)
Sensibilité implicite des jeunes élèves à l'environnement consonantique et à la position d'une correspondance phonème-graphème (ex. : « eau » impossible en début de mot) (Pacton et al., 2002)		

### **Régularités graphotactiques contextuelles**

Acquisition tardive des graphies consistantes contextuelles (ex. : c devant a, o, u et qu devant e et i) (Mousty et Leybaert, 1999)	Prédiction de l'hypothèse 2 infirmée : La règle contextuelle IM est acquise à 79 %	Prédiction de l'hypothèse 4 confirmée : La règle contextuelle IM n'est acquise qu'à 26 %
Risque accru d'erreurs graphémiques selon le voisinage du graphème, plus précisément en présence d'une semi-voyelle (Pothier et Pothier, 2004)	Prédiction de l'hypothèse 2 infirmée : La semi-voyelle contribue à la réussite du graphème « en » à 95 %	Prédiction de l'hypothèse 4 infirmée : La semi-voyelle contribue à la réussite du graphème « en » à 77 %

---

## Régularités morphologiques

---

Acquisition tardive des stratégies morphologiques (Mousty et Leybaert 1999)	Prédiction de l'hypothèse 2 confirmée : Les bons orthographes sont sensibles aux stratégies morphologiques dès la troisième année du primaire	Prédiction de l'hypothèse 4 infirmée : Les faibles orthographes sont sensibles aux stratégies morphologiques dès la troisième année du primaire
Les faibles orthographes éprouvent davantage de difficulté avec la morphologie dérivationnelle (Bégin et al., 2010)	S.O.	Prédiction de l'hypothèse 4 infirmée : Les faibles orthographes orthographient mieux le graphème « ain » dérivable (70 % de réussite) que le graphème « ain » non dérivable (35 % de réussite)
Graphèmes inconsistants contenus dans les mots dérivés mieux orthographiés que dans les autres mots (Casalis et al., 2011)	Graphème inconsistant « ain » contenu dans les mots dérivables un peu mieux orthographié que dans les mots non dérivables (2 %)	Graphème inconsistant « ain » contenu dans les mots dérivables beaucoup mieux orthographié que dans les mots non dérivables (35 %)

---

### 5.1 Discussion liée à l'hypothèse 1 et à l'hypothèse 3

La première et la troisième hypothèse prédisaient un effet positif de la consistance, de la fréquence et de la longueur du mot (nombre peu élevé de lettres) sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois bons et faibles orthographes de la troisième année du primaire. La présente recherche a montré que le facteur lexical de la longueur du mot (nombre peu élevé de lettres) a un effet sur l'acquisition du phonème multigraphémique /ɛ̃/ chez les faibles orthographes. Toutefois, la consistance du mot n'est pas significative et la fréquence du mot ne joue pas de rôle important dans cette acquisition ni chez les bons ni chez les faibles orthographes.

### **5.1.1 Fréquence du mot**

La fréquence du mot a un effet connu sur l'acquisition de l'orthographe des mots (Fayol, 2017; Martinet et al., 2004; Mousty et Leybaert, 1999; Sprenger-Charolles et al., 1998; Stanké et al., 2019), et ce, dès la première année du primaire (Lété et al., 2008; Martinet et al., 2004). Toutefois, les résultats de la présente recherche doctorale montrent que la fréquence du mot n'a pas d'effet significatif sur l'acquisition des graphèmes associés au phonème multigraphémique /*ɛ̃*/, et ce, autant pour les élèves faibles orthographieurs que pour les élèves bons orthographieurs. Sprenger-Charolles et al. (1998) avaient remarqué que, pour les mots réguliers, les élèves s'en remettaient à leurs connaissances phonologiques de base. Il en va probablement ainsi lorsque les élèves orthographient un phonème multigraphémique, qu'il soit dans un mot fréquent ou non.

### **5.1.2 Consistance du mot**

La consistance du mot a un effet très peu significatif sur l'acquisition du phonème multigraphémique /*ɛ̃*/, tant chez les bons que chez les faibles orthographieurs. En effet, bien que l'inconsistance orthographique ait un effet sur le développement des habiletés en lecture et en orthographe (Caravolas, 2019; Treiman et al., 2019), la consistance totale du mot ne joue pas de rôle significatif dans la capacité à orthographier le phonème multigraphémique /*ɛ̃*/ . Ainsi, le fait qu'un mot soit, dans l'ensemble, consistant ou non a peu d'influence sur la capacité à orthographier un graphème. C'est la consistance du graphème à orthographier qui joue un rôle important. Nous traitons de ce point avec les facteurs sous-lexicaux.

### **5.1.3 Longueur du mot (nombre peu élevé de lettres)**

Les résultats de différentes études avaient déjà montré que les élèves réussissent mieux à orthographier les mots quand ils sont courts (Fayol, 2017; Lété et al., 2008; Martin-Chang et al., 2016). Cela dit, la présente étude montre que la longueur du mot (nombre peu élevé de lettres) n'a un effet positif sur l'acquisition du phonème multigraphémique /*ɛ̃*/ que chez les faibles orthographieurs. En effet, ces derniers voient leur taux de réussite passer sous la barre de 75 % dès que le mot a six lettres et plus. Cette difficulté peut s'expliquer, entre autres, par la multigraphémie, qui demande aux élèves de recourir aux caractéristiques visuelles des mots afin de les garder en mémoire (Daigle et al., 2013), et par la charge cognitive importante qu'exerce l'orthographe phonologique des mots longs, lorsque leur orthographe lexicale n'est pas encore acquise.

Enfin, les mots plus longs sont plus difficiles à stocker dans le lexique orthographique. Ils nécessitent alors de recourir au traitement alphabétique, qui permet d'écrire phonème par phonème ou syllabe par syllabe, ce qui est coûteux en mémoire de travail et peut entraîner une surcharge cognitive (Stanké et al., 2016). Les résultats de la présente étude révèlent que cette surcharge est plus élevée chez les faibles orthographes, chez qui elle entraîne des difficultés à orthographier les graphèmes, qu'ils ont par ailleurs su orthographier dans des mots plus courts, en ayant recours aux régularités graphotactiques ou morphologiques. Cela peut s'expliquer, entre autres, parce que l'empan visuoattentionnel est souvent limité chez les faibles orthographes, ce qui fait en sorte que l'œil ne peut pas traiter plusieurs lettres à la fois (Valdois, 2014). La mémoire de travail, qui a notamment une incidence sur les capacités phonologiques, peut aussi avoir eu une incidence chez les faibles orthographes (Stanké et al., 2016).

#### **5.1.4 Synthèse**

Ainsi, la première hypothèse et la troisième hypothèse prédisaient un effet positif de la fréquence, de la longueur et de la consistance du mot sur l'acquisition du phonème multigraphémique /*ɛ̃*/ dans les productions orthographiques des élèves québécois bons orthographes et faibles orthographes de la troisième année du primaire. L'analyse des résultats mène à constater que seule la longueur des mots joue un rôle sur l'acquisition du phonème multigraphémique /*ɛ̃*/, et ce, uniquement chez les faibles orthographes. Toutefois, l'effet significatif de la consistance et de la fréquence du mot sur l'acquisition des graphèmes associés au phonème /*ɛ̃*/ ne s'observe ni chez les bons ni chez les faibles orthographes. Cela confirme partiellement notre première hypothèse.

#### **5.2 Discussion liée à l'hypothèse 2 et à l'hypothèse 4**

Des constats importants peuvent être faits du côté de notre deuxième et de notre quatrième hypothèse. La deuxième hypothèse prédisait que certains facteurs sous-lexicaux (types de graphèmes, consistance phonographémique, positionnement du graphème et régularités morphologiques dérivationnelles) auraient un effet positif et certains autres facteurs sous-lexicaux (contexte d'utilisation et voisinage du graphème) auraient un effet négatif sur l'acquisition du phonème multigraphémique /*ɛ̃*/ dans les productions orthographiques des élèves québécois bons orthographes de la troisième année du primaire. La quatrième hypothèse prédisait que la plupart des facteurs sous-lexicaux auraient le même effet chez les faibles orthographes. Toutefois, un

effet négatif était prévu pour les régularités morphologiques dérivationnelles, les faibles orthographes éprouvant davantage de difficulté avec la morphologie dérivationnelle (Bégin et al., 2010).

### **5.2.1 Types de graphèmes et consistance phonographémique**

Du côté des types de graphèmes et de leur consistance phonographémique, les résultats d'analyse de la présente recherche confirment les conclusions de recherches antérieures de Mousty et Leybaert (1999). Ces derniers avaient constaté une acquisition rapide des graphèmes consistants acontextuels ainsi qu'une application de la graphie dominante par les élèves lorsque le lexique orthographique n'est pas assez riche.

Tout d'abord, les résultats d'analyse mènent à constater que la graphie dominante « in » est majoritairement choisie en cas d'erreurs, tant par les bons que par les faibles orthographes. L'analyse des graphèmes choisis en cas d'erreurs montre que les faibles orthographes choisissent le graphème le plus fréquent, soit le « in », dans une proportion très élevée, soit de plus de 70 %. Cette tendance confirme le constat de Mousty et Leybaert (1999) voulant que les élèves ayant un lexique orthographique plus pauvre utilisent le graphème le plus fréquent de façon quasi systématique.

De plus, les bons orthographes acquièrent le graphème non dérivable le plus consistant « in » à 96 % alors qu'ils n'acquièrent le graphème le moins consistant « ein » qu'à 34 %. Les faibles orthographes acquièrent le graphème non dérivable le plus consistant « in » à 90 % alors qu'ils n'acquièrent le graphème le moins consistant « ein » qu'à 18 %. Plus précisément, les faibles orthographes ont 90 % de chances de bien orthographier le graphème « in » alors que les bons orthographes ont 96 % de chances d'avoir une réponse correcte. Lorsque le graphème est moins consistant, les faibles orthographes ont beaucoup moins de chances d'obtenir une réponse correcte. Par exemple, ils n'ont que 35 % de chances de bien orthographier le graphème non consistant « ain » alors que les bons orthographes ont 87 % de chances de bien l'orthographier.

Plus précisément, le type de graphèmes (« in », « ain », « ein », « en », « im ») a une incidence importante sur la réussite orthographique des graphèmes pour les élèves faibles et bons orthographes. L'acquisition des graphèmes associés au phonème /ɛ̃/ se fait dans le même ordre,

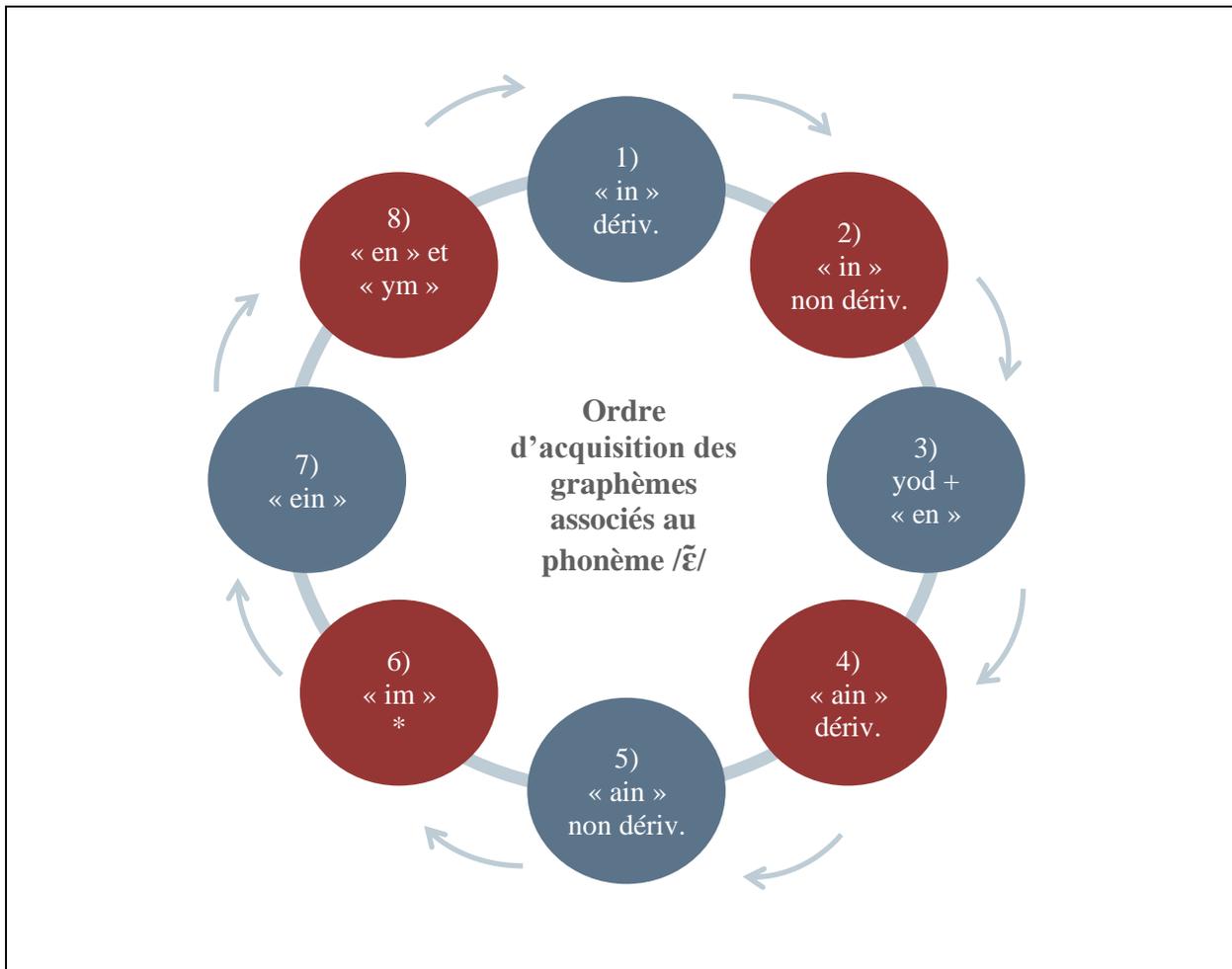
tant pour les élèves faibles orthographes que pour les élèves bons orthographes : premièrement, le « in » dérivable, puis le « in » non dérivable, ensuite, le graphème contextuel « en » précédé du yod (« ien » ou « yen »), puis le « ain » dérivable », le « ain » non dérivable, le « im » et, finalement, le « ein ».

Des différences dans l'acquisition de certains graphèmes existent toutefois entre les profils d'élèves. Plus précisément, les graphèmes très consistants, soit le « in » et le « en » précédé du yod (« ien » et « yen »), sont acquis par tous. Le graphème le plus inconsistant « ein » est, pour sa part, échoué par la majorité des élèves de troisième année du primaire. Une différence importante dans l'acquisition du graphème inconsistant « ain » non dérivable est notée pour les élèves faibles orthographes qui ont 2,49 fois plus de chances d'échouer que les élèves bons orthographes. La différence est encore plus grande dans l'acquisition du graphème contextuel « im » que les élèves faibles orthographes ont 3,69 fois plus de chances d'échouer que les bons orthographes. Ce graphème est acquis pour les élèves bons orthographes, mais pas par leurs pairs. Enfin, le graphème inconsistant « ym » a été échoué à 100 % par les faibles orthographes et le graphème inconsistant « en » non précédé du yod a été échoué à 91,2 % par tous les élèves.

Morin et Montésinos-Gelet (2005) avaient déjà montré que les correspondances consistantes entre les phonèmes et les graphèmes, comme le /ɔ̃/ qui s'écrit presque toujours « on » sauf lorsqu'il est suivi d'un « b » ou d'un « p », peuvent être acquis dès la maternelle. Les résultats de la présente recherche montrent que, dans le cas de correspondances inconsistantes entre les phonèmes et les graphèmes, comme celles du phonème /ɛ̃/, la consistance des graphèmes joue un rôle majeur dans l'acquisition phonographémique. Plus ces graphèmes sont inconsistants, plus ils créent des difficultés orthographiques, et ce, même dans des mots fréquents.

Les précédents résultats pourront jouer un rôle important dans l'enseignement des phonèmes multigraphémiques. À titre d'illustration, la figure 5.1 présente une hiérarchisation de l'acquisition et de l'enseignement des graphèmes associés au phonème /ɛ̃/.

**Figure 5.1** Ordre d'acquisition des graphèmes associés au phonème /*ẽ*/



### **5.2.2 Régularité graphotactique liée au positionnement du graphème**

Pour ce qui est des régularités graphotactiques de positionnement, la présente recherche confirme la sensibilité implicite des jeunes élèves à la position légale des lettres dans l'acquisition de l'orthographe, sensibilité qu'avait remarquée Pacton et al. (2001). En effet, les élèves réussissent mieux le phonème multigraphémique /*ẽ*/ en début de mot. Plus précisément, bien que la position initiale soit un facteur facilitant et que la position finale soit un facteur qui augmente le risque d'erreurs, tant pour les bons que pour les faibles orthographes, les faibles orthographes ont 37 % plus de chances de bien orthographier un graphème associé au phonème multigraphémique /*ẽ*/ en position initiale qu'en position finale, alors que les bons orthographes ont 17 % plus de

chances de bien orthographier un graphème associé au phonème multigraphémique /ɛ̃/ en position initiale qu'en position finale.

Le fait que les graphèmes soient plus faciles à orthographier en position initiale peut s'expliquer par le peu de choix de graphèmes possibles. En effet, seuls le « in », le « im » et le « ain » sont possibles et le graphème le plus consistant est le « in » à 80,8 %. De plus, la régularité graphotactique voulant que le phonème /ɛ̃/ s'orthographie « in » lorsqu'il est positionné au début du mot est un facteur facilitant pour les élèves. Ils ont d'ailleurs si bien acquis cette régularité qu'ils n'ont pas pu réussir la seule exception où le phonème /ɛ̃/ ne s'écrit pas « in » en début de mot. En effet, le graphème exceptionnel « ain » du mot « ainsi » a été échoué à 92,2 % par les bons orthographieurs et à 100 % par les faibles orthographieurs. Ce résultat corrobore ceux des recherches portant sur la sensibilité des élèves aux régularités graphotactiques (Borchardt et al., 2012; Pacton et al., 2014; Pacton, Foulin et al., 2005; Pacton et Casalis, 2006; Sénéchal, 2000).

Lorsque le graphème est en position finale, il est moins réussi par tous les élèves. Cela peut s'expliquer par la présence du graphème « ein » en fin de mot, qui est beaucoup moins consistant que les autres graphèmes (0,49 % de consistance). Il est intéressant de noter que la présence du graphème « ein » en fin de mot donne plus de fil à retordre aux élèves que les nombreux choix qui sont possibles en position interne : « in », « im », « en », « ain », « ein », « ym », « yn » et « in ».

### **5.2.3 Régularité graphotactique liée au contexte d'utilisation du graphème**

Du côté des régularités graphotactiques contextuelles, la présente recherche permet de faire des découvertes importantes. Bien que Mousty et Leybaert (1999) et Catach (2016) avaient constaté une acquisition tardive des graphies consistantes contextuelles, les analyses de la présente recherche suggèrent que les bons orthographieurs de la troisième année du primaire acquièrent la règle contextuelle « im » à 77 %. C'est chez les faibles orthographieurs seulement que le constat des recherches antérieures s'applique. Comme prévu, le contexte d'utilisation n'a pas eu un effet positif chez ces élèves, qui n'arrivent pas à s'adapter au contexte pour modifier des graphèmes au besoin. Ces derniers n'arrivent, en effet, à bien transformer le graphème « in » en « im » selon le contexte que dans 21 % des cas.

Cette difficulté à appliquer les régularités graphotactiques contextuelles par les faibles orthographes peut s'expliquer par une difficulté avec les propriétés visuelles, qui tiennent une place importante dans le modèle de Daigle et al. (2018). En effet, les élèves doivent faire un traitement visuel pour acquérir l'orthographe des mots qui contiennent des règles de positionnement (Daigle et al., 2013). La moitié des erreurs en orthographe lexicale provenant des difficultés avec les propriétés visuelles des mots (Daigle et al., 2013), il n'est pas étonnant que les faibles orthographes peinent à faire le traitement visuel nécessaire pour appliquer la règle contextuelle du « m » devant le « b » et le « p », d'autant que les enseignants font rarement appel à l'enveloppe visuelle des mots dans leur enseignement (Daigle et al., 2015).

#### **5.2.4 Régularité graphotactique liée au voisinage du graphème**

Pour ce qui est de l'effet du voisinage du graphème, il est positif plutôt que négatif pour tous les élèves, mais moins pour un profil d'élèves que pour l'autre. En effet, le yod qui précède le « en » rend ce graphème consistant (« ien » et « yen ») et contribue à sa réussite à 95 % pour les bons orthographes et à 77 % pour les faibles orthographes. En effet, les élèves arrivent très bien à orthographier le graphème « en » parce que les mots dans lesquels il se trouve contiennent une semi-voyelle qui rend le graphème consistant, c'est-à-dire « ien ». Le mot « européens », dans lequel le graphème « en » n'est pas précédé d'une semi-voyelle n'a pour sa part été réussi que par 9,8 % des bons orthographes et par aucun faible orthographe. Le graphème « in » ayant, pour sa part, été très réussi par tous les élèves, il est impossible d'affirmer que la semi-voyelle qui précède le « in » et qui rend ce graphème consistant (« oin ») aide à sa réussite, mais il est possible de constater qu'elle ne nuit pas à sa réussite.

Contrairement à ce qu'ont observé Pothier et Pothier (2004), le risque d'erreurs ne semble donc pas accru en présence d'une semi-voyelle. C'est du moins ce que la présente recherche doctorale a constaté en ce qui concerne l'orthographe des graphèmes associés au phonème / $\tilde{\epsilon}$ / précédé d'une semi-voyelle, et ce, davantage pour les bons orthographes que pour les faibles orthographes.

#### **5.2.5 Régularités morphologiques dérivationnelles**

Enfin, certains constats importants ont été faits du côté des régularités morphologiques. Le graphème « in » ayant été très bien réussi par tous les élèves, qu'il soit dérivable ou non, il est

difficile d'émettre une conclusion. Toutefois, le graphème « ain » ayant été plus échoué, les résultats sont plus significatifs. En effet, les données montrent que les mots dérivables en « ain » (« certAIn » / « certAIne » et « mAin » / « mAnuel », par exemple) sont davantage réussis que les mots qui ne dérivent pas, notamment chez les élèves faibles orthographes. Alors que la recherche montrait que les faibles orthographes éprouvaient davantage de difficulté avec la morphologie dérivationnelle (Bégin et al., 2010), les résultats suggèrent que ces derniers orthographient 35 % mieux les graphèmes inconsistants contenus dans les mots dérivables en « ain » que dans les autres mots non dérivables en « ain » (7 % de plus pour le « in » et 35 % de plus pour le « ain »). Cette conclusion va aussi dans le même sens que les données de la recherche de Casalis et al. (2011), qui avait conclu que les graphèmes inconsistants contenus dans les mots dérivés étaient mieux orthographiés que dans les autres mots. Dans ce cas-ci, toutefois, les graphèmes inconsistants se retrouvent dans des mots dérivables comme « certain » et non dans des mots dérivés comme « certainement », ce qui augmente la difficulté.

Ainsi, contrairement aux conclusions de certains chercheurs qui ont observé que le traitement morphographémique s'acquière plus tardivement (Morin et Montésinos-Gelet, 2005; Mousty et Leybaert, 1999; Seymour, 1997), nos résultats rejoignent ceux de Sénéchal (2000) et de Pacton et ses collaborateurs (2005). En effet, les résultats de notre recherche montrent que les élèves sont sensibles aux régularités morphologiques rapidement, soit dès la troisième année du primaire, et l'effet est encore plus important pour les faibles orthographes. La sensibilité implicite à la morphologie, que Gombert (2003) a remarqué dès le préscolaire, joue donc un rôle important, et ce, malgré les difficultés à faire un traitement explicite des morphogrammes avant que le traitement orthographique soit acquis, difficultés qu'a soulignées Seymour (1997) dans son modèle d'acquisition du français écrit.

### **5.2.6 Synthèse**

Ainsi, la deuxième hypothèse prédisait un effet positif de certains facteurs sous-lexicaux (types de graphèmes, consistance phonographémique, positionnement du graphème et régularités morphologiques dérivationnelles) et un effet négatif de certains autres facteurs sous-lexicaux (contexte d'utilisation et voisinage du graphème) sur l'acquisition du phonème multigraphémique /ɛ̃/ dans les productions orthographiques des élèves québécois bons orthographes de la troisième

année du primaire. La quatrième hypothèse, quant à elle, prédisait un effet positif de certains facteurs sous-lexicaux (types de graphèmes, consistance phonographémique, positionnement du graphème) et un effet négatif de certains autres facteurs sous-lexicaux (voisinage du graphème, contexte d'utilisation du graphème et régularités morphologiques dérivationnelles) sur l'acquisition du phonème multigraphémique / $\tilde{\epsilon}$ / dans les productions orthographiques des élèves québécois faibles orthographes de la troisième année du primaire. L'analyse des résultats suggère que le type de graphème et sa consistance ainsi que le positionnement du graphème ont un effet positif chez tous les élèves. Quant au contexte d'utilisation du graphème, il a un effet négatif chez les faibles orthographes seulement. Le voisinage du graphème n'a pas l'effet négatif escompté ni chez les bons ni chez les faibles orthographes. Enfin, non seulement les régularités morphologiques dérivationnelles ne jouent-elles pas un rôle négatif dans l'acquisition du phonème multigraphémique / $\tilde{\epsilon}$ /, mais elles jouent un rôle positif dans cette acquisition, et ce, chez les bons orthographes, mais surtout chez les faibles orthographes de la troisième année du primaire au Québec.

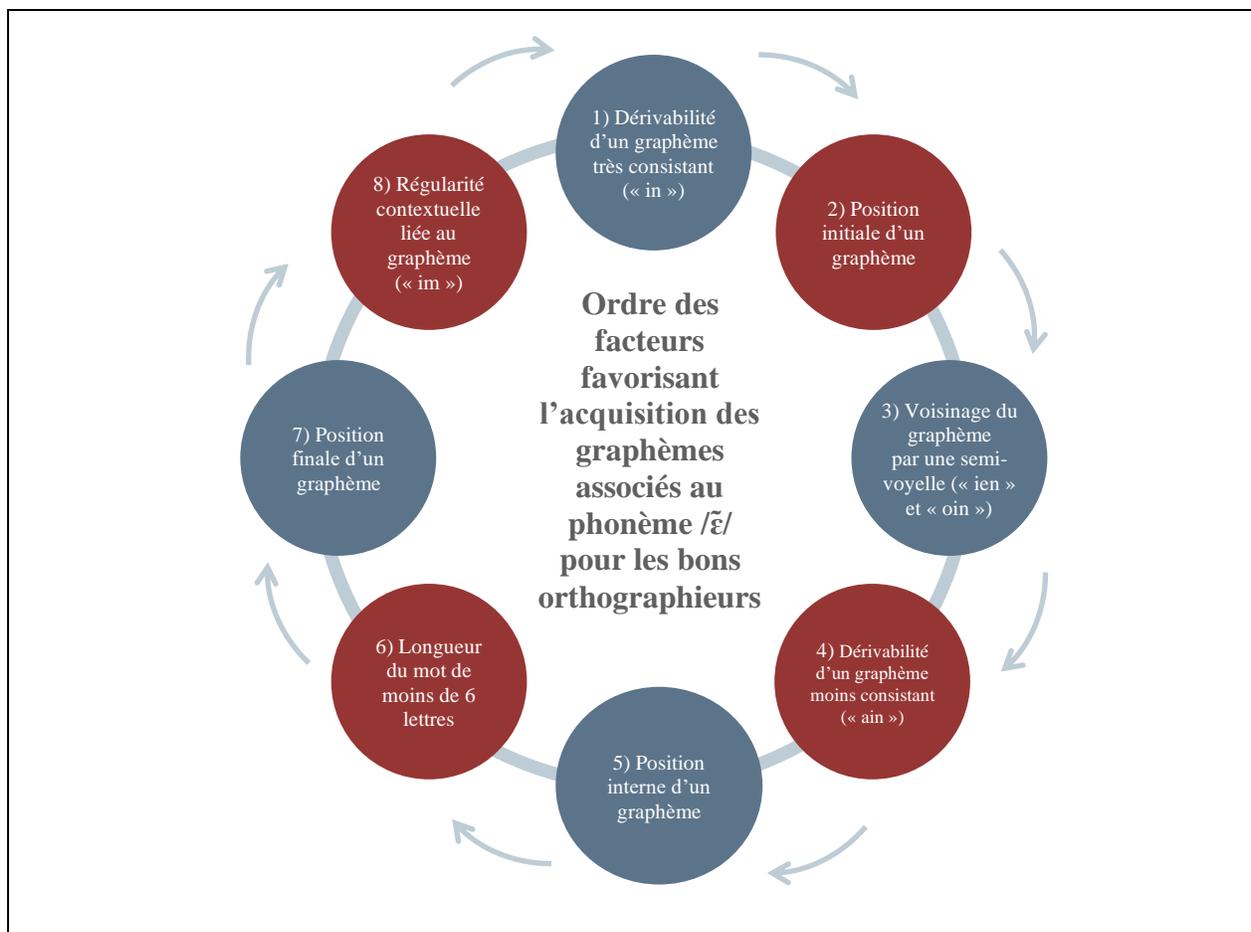
### **5.3 Hiérarchisation des facteurs lexicaux et sous-lexicaux dans l'acquisition du phonème multigraphémique / $\tilde{\epsilon}$ / en fonction des deux catégories de participants : les bons orthographes et les faibles orthographes**

Les facteurs lexicaux et sous-lexicaux qui ont un effet positif sur l'acquisition du phonème multigraphémique / $\tilde{\epsilon}$ / dans les productions orthographiques des élèves québécois de la troisième année du primaire ne sont pas exactement les mêmes pour les élèves bons orthographes que pour les élèves faibles orthographes. De plus, bien que semblables, les facteurs favorisant l'acquisition du phonème multigraphémique / $\tilde{\epsilon}$ / ne se présentent pas exactement dans le même ordre pour les bons orthographes que pour les faibles orthographes. La présente section illustre la hiérarchie de ces facteurs chez les bons et chez les faibles orthographes.

Tout d'abord, la figure 5.2 permet de remarquer que les bons orthographes ont plus de facilité à acquérir les graphèmes associés au phonème multigraphémique / $\tilde{\epsilon}$ / lorsqu'ils sont très consistants et dérivables, soit le « in » (98 % de réussite), qui est un facteur lié aux régularités morphologiques dérivationnelles. Viennent ensuite la position initiale du graphème (95 % de réussite) et le voisinage du graphème par une semi-voyelle (95 % de réussite), qui sont des facteurs liés aux régularités graphotactiques. Le facteur suivant est relié aux régularités morphologiques

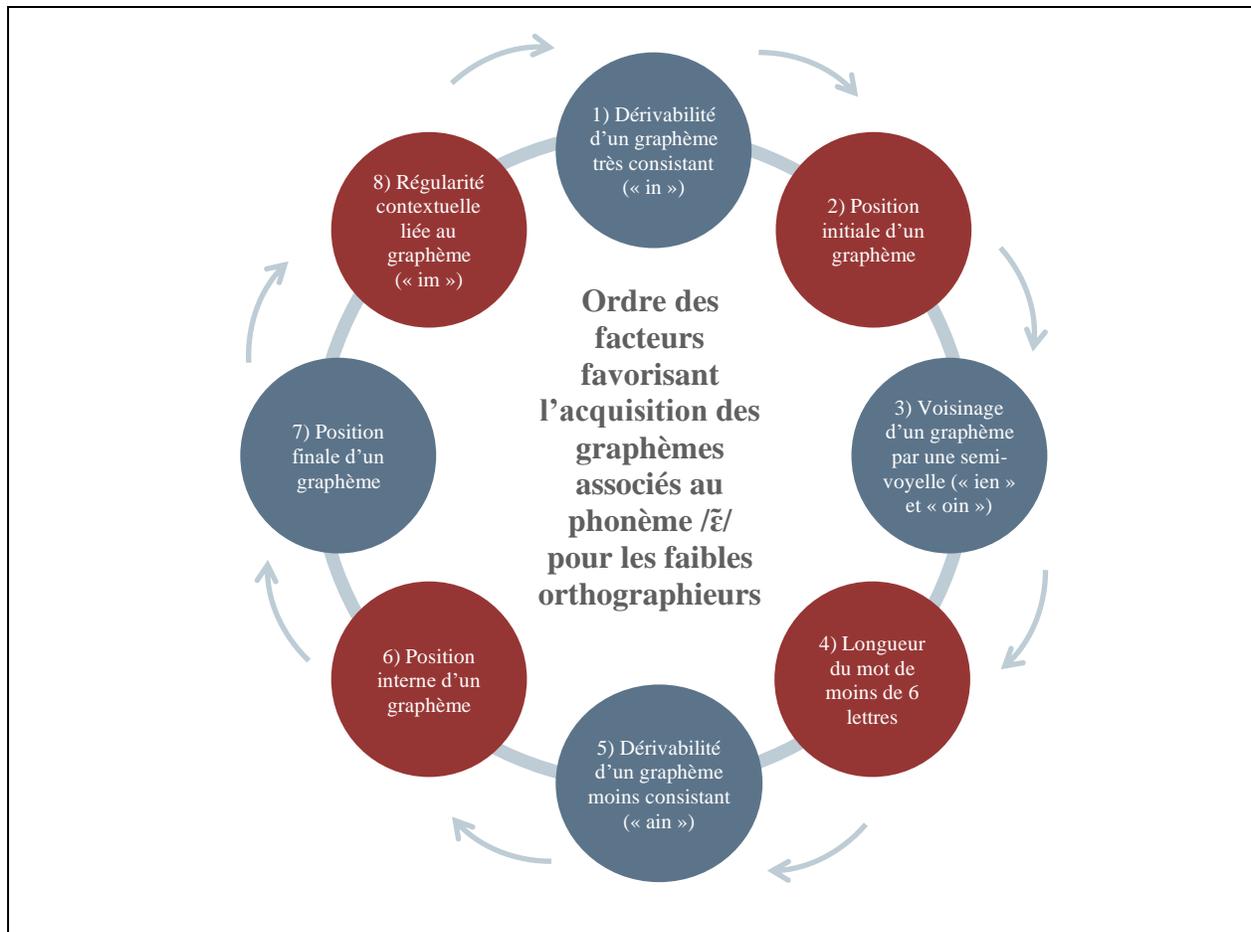
dérivationnelles, mais pour un graphème moins consistant que le « in », soit le « ain » (89 % de réussite). Viennent ensuite la position interne du graphème (88 % de réussite) et le nombre peu élevé de lettres des mots (87 % de réussite pour les mots de six lettres et plus) qui ne causent pas vraiment de difficultés aux bons orthographes. La position finale du phonème multigraphémique (78 % de réussite) et la régularité contextuelle du « im » (77 %) arrivent à la fin, mais sont tout de même acquises puisque leur taux de réussite est de 75 % et plus. Tel que mentionné, un profil d'élèves est considéré comme ayant réussi la tâche lorsque le taux est de 75 % et plus (Pothier et Pothier, 2004; Stanké et al., 2019).

**Figure 5.2** Ordre des facteurs favorisant l'acquisition des graphèmes associés au phonème /ɛ̃/ pour les bons orthographes



Puis, la figure 5.3 présente l'ordre dans lequel les facteurs favorisant l'acquisition du phonème multigraphémique /ɛ̃/ se présentent pour les élèves faibles orthographes. Ils ont, eux aussi, plus de facilité à acquérir les phonèmes multigraphémiques très consistants dérivables, soit le « in » (97 % de réussite), ce qui est un facteur lié aux régularités morphologiques dérivationnelles. Viennent ensuite la position initiale du graphème (82 % de réussite) et le voisinage du graphème par une semi-voyelle (77 % de réussite), qui sont des facteurs liés aux régularités graphotactiques. Le taux de réussite des cinq autres facteurs par les faibles orthographes n'est pas de 75 % et plus, ce qui fait qu'ils ne jouent pas un rôle significatif dans l'acquisition des graphèmes associés au phonème /ɛ̃/. Ils se présentent dans l'ordre suivant : le nombre peu élevé de lettres des mots (72 % de réussite pour les mots six lettres et plus), le facteur relié aux régularités morphologiques dérivationnelles pour un graphème moins consistant que le « in », soit le « ain » (70 % de réussite), la position interne du graphème (62 % de réussite) et la position finale du graphème (45 % de réussite). Tout comme pour les bons orthographes, la régularité contextuelle du « im » arrive à la toute fin, mais avec un pourcentage de réussite nettement plus bas (21 % pour les faibles orthographes contre 77 % pour les bons orthographes). Malgré le fait que les faibles orthographes n'aient pas obtenu un taux de réussite de 75 % et plus pour le graphème dérivable « ain », il faut retenir que la dérivabilité du graphème a joué un rôle majeur pour ces derniers qui, en revanche, n'ont eu un taux de réussite que de 35 % pour le graphème « ain » non dérivable.

**Figure 5.3 Ordre des facteurs favorisant l'acquisition des graphèmes associés au phonème /ɛ̃/ pour les faibles orthographes**



Ainsi, l'analyse des résultats permet de comprendre que les facteurs suivants causent des difficultés aux faibles orthographes, alors que ce n'est pas le cas pour les bons orthographes : le nombre élevé de lettres des mots et la régularité contextuelle liée au phonème multigraphémique « im ». Enfin, la consistance des graphèmes est aidante pour les bons et pour les faibles orthographes et la dérivabilité des graphèmes est particulièrement aidante pour les faibles orthographes lorsque les graphèmes sont moins consistants. Ces résultats ont un effet important pour la pratique.

#### 5.4 Implications des nouvelles connaissances pour la pratique

Les résultats de cette recherche apportent une contribution importante aux connaissances scientifiques, mais aussi aux connaissances liées à la pratique. La jonction entre l'éducation et la linguistique pourrait permettre au personnel enseignant et aux orthopédagogues d'apprendre à mieux manipuler le matériel linguistique et à mieux expliquer les facteurs liés aux phonèmes multigraphémiques en se basant sur des résultats de recherche. Nous en discutons dans cette partie.

Tout d'abord, nous avons constaté que la consistance des graphèmes est très importante pour tous les élèves. Ces derniers utilisent d'ailleurs très majoritairement les graphèmes les plus fréquents quand ils font une erreur; ils écrivent « in » plutôt que « ain », par exemple. De plus, ils réussissent davantage à orthographier les graphèmes consistants qui sont associés au phonème multigraphémique /ɛ̃/. Il serait donc important de ne pas enseigner les graphèmes exceptionnels, comme le « ein », le « ym » ou le « en » non accompagné du yod trop tôt afin que les élèves mémorisent d'abord les graphèmes les plus consistants. Il serait aussi important de s'assurer de présenter un pourcentage de mots correspondant à ces consistances quand vient le temps d'enseigner la multigraphémie. Par exemple, selon la consistance des graphèmes, en fin de mot, il serait préférable de retrouver 27 % de mots en « in », 55 % de mots en « en » et 17 % des mots en « ain », le reste du pourcentage étant occupé par des graphèmes très exceptionnels (Peereman et Stanké, 2022).

De plus, il serait pertinent de limiter les mots comportant un grand nombre de lettres au début de l'acquisition de l'orthographe lexicale afin d'éviter la surcharge cognitive qui engendre des erreurs orthographiques. Plus précisément, pour les élèves du premier cycle du primaire, il faudrait privilégier les mots de moins de six lettres. Ces mots plus courts permettraient d'ailleurs aux faibles orthographieurs d'être plus disponibles pour extraire les régularités associées aux différents graphèmes des phonèmes multigraphémiques, c'est-à-dire les connaissances graphotactiques. Les connaissances graphotactiques ont non seulement un effet sur la mémorisation des graphèmes, mais elles jouent un rôle important dans la mémorisation des mots entiers (Borchardt, 2012). Stanké et ses collaborateurs (2020) ont d'ailleurs montré que l'enseignement explicite des régularités graphotactiques et morphologiques est bénéfique pour tous les élèves, qu'ils soient bons ou faibles orthographieurs.

Plus précisément, il serait important de développer des pistes didactiques pour les élèves présentant des difficultés en orthographe. Plusieurs recherches ont montré que l'enseignement explicite de l'orthographe lexicale a des effets très positifs chez les élèves faibles orthographes (Chapleau et al., 2013; Daigle et al., 2015; Wanzek et al., 2006). Il serait notamment important d'enseigner, de façon explicite, la régularité graphotactique positionnelle du « m » devant le « b » et le « p ». Le traitement articulaire joue un rôle important dans l'apprentissage de la lecture (Desrochers et al., 2009). S'y référer pour enseigner l'orthographe lexicale peut aussi être pertinent. Par exemple, pour enseigner le graphème « im », il est possible de faire prendre conscience aux élèves faibles orthographes que les lèvres se ferment par la suite avec le phonème /b / et le phonème /p/, c'est-à-dire que ce sont des consonnes bilabiales.

Il pourrait aussi être pertinent d'enseigner les régularités morphologiques dérivationnelles dès le premier cycle du primaire. Bien que les habiletés en conscience morphologique soient identifiées comme faibles chez les élèves ayant des difficultés en langage écrit (St-Pierre et al., 2013), les résultats de notre étude montrent, en effet, que les élèves faibles orthographes sont sensibles à la dérivation des mots, et ce, même sans que cela ne leur ait été explicitement enseigné. Il serait donc gagnant de faire comprendre que nombre de mots qui se terminent par le graphème « in » ont une dérivation possible en « i » et que plusieurs mots qui se terminent par le graphème « ain » ont une dérivation possible en « a » ou en « ai ».

Enfin, les résultats de notre étude ont montré que la fréquence des mots ne semble pas avoir d'impact positif sur l'acquisition des phonèmes multigraphémiques. Pour leur enseignement, l'utilisation unique de listes orthographiques basées sur la fréquence des mots et non sur la fréquence des graphèmes, n'est donc pas à privilégier. Présentement, 171 mots contenant le phonème /ɛ̃/ se trouvent dans la liste orthographique à l'usage des enseignantes et des enseignants du Québec et sont présentés par fréquence (ministère de l'Éducation et de l'Enseignement supérieur, 2014) (voir annexes G, H et I). Se référer uniquement à cette liste, qui est basée sur la fréquence des mots et non sur la consistance des graphèmes, pourrait avoir comme effet d'enseigner le mot « princesse » en deuxième année du primaire alors que des mots contenant des graphèmes plus consistants et des régularités graphotactiques, comme « coin » et « foin », ou des régularités morphologiques dérivationnelles, comme « pin », seraient enseignés plus tard, soit respectivement en troisième, en quatrième et en cinquième année du primaire. Il serait aussi important de regrouper

les mots de même famille pour l'enseignement de la régularité graphotactique de positionnement. Par exemple, le mot « invitation », qui est enseigné en troisième année du primaire, le mot « inviter » qui est enseigné en quatrième année du primaire et le mot « invité », qui est enseigné en cinquième année du primaire, pourraient tous être enseignés en troisième année du primaire.

La retombée principale de cette recherche ne permet pas de discuter des listes orthographiques des mots à acquérir au primaire puisque son objectif était lié à l'acquisition des phonèmes multigraphémiques et non à l'acquisition des mots. Toutefois, elle permet de dresser un ordre parmi les principaux phonèmes multigraphémiques à acquérir. À partir des constats liés au phonème multigraphémique /*ɛ̃*/, les enseignants et orthopédagogues peuvent choisir les mots qui correspondent aux mêmes critères pour différents phonèmes. Pour ce, ils peuvent utiliser la liste orthographique à l'usage des enseignantes et des enseignants (ministère de l'Éducation et de l'Enseignement supérieur, 2014) ou ÉQOL (Stanké et al., 2019) au Québec. Les mots d'ÉQOL contenant le phonème multigraphémique /*ɛ̃*/ sont présentés en annexe (voir annexe J). Ailleurs dans la francophonie, ils peuvent utiliser d'autres banques de données.

Ainsi, il est essentiel que, quand vient le temps d'enseigner un phonème multigraphémique, l'accent soit mis sur les graphèmes possibles associés au phonème et non sur les mots en entier. Nous avons, en effet, montré que l'acquisition des mots diffère de l'acquisition des phonèmes multigraphémiques. L'enseignement des mots entiers peut être fait à partir des résultats des recherches antérieures, en se basant notamment sur leurs propriétés phonologiques, visuelles et morphologiques. Ce n'était pas, ici, le but de la présente recherche.

Quant à l'enseignement des phonèmes multigraphémiques, il pourra se faire en portant une attention particulière à certains de ces phonèmes un à la fois. Pour chacun de ces phonèmes multigraphémiques, nous proposons une procédure. Si la fréquence du mot n'a pas d'importance pour l'acquisition des phonèmes multigraphémiques, le nombre de lettres dans le mot a de l'importance. Ainsi, nous proposons de choisir des mots de moins de six lettres pour le premier cycle, de six à neuf lettres pour le deuxième cycle et de dix lettres et plus pour le troisième cycle. Enfin, au premier cycle, les mots devraient contenir le moins de phonèmes multigraphémiques possible afin de ne pas surcharger la mémoire des élèves.

Pour chaque cycle scolaire, deux facteurs associés à un phonème multigraphémique pourraient être présentés à la fois afin que les élèves puissent confronter des facteurs linguistiques dans leur apprentissage.

Bien que le nombre de lettres et de phonèmes multigraphémiques puissent varier, les résultats de notre recherche permettent de dresser une liste de facteurs desquels il serait pertinent de tenir compte pour l'enseignement des phonèmes multigraphémiques. Dans une approche inclusive, tenir compte des facteurs liés aux élèves en difficulté saura être bénéfique pour tous. En intervention orthopédagogique, le nombre de lettres ainsi que le nombre de phonèmes multigraphémiques par mot pourront être réduits au besoin. Ainsi, l'ordre des facteurs pourrait être le suivant pour tous les phonèmes multigraphémiques :

- Le type de graphèmes (partir des plus consistants à partir du tableau 5.1);
- Les graphèmes s'expliquant par leur position dans le mot (consistance de plus de 80 % comme le « in »);
- Les graphèmes s'expliquant par leur voisinage;
- Les graphèmes dérivables associés au phonème multigraphémique (ce critère est en troisième position puisque, bien que le graphème dérivable très consistant « in » ait été mieux réussi, le graphème dérivable moins consistant « ain » est arrivé en quatrième position);
- Les graphèmes contenant des régularités contextuelles (im, g/gu, c/qu, s/ss);
- Les graphèmes plus exceptionnels comme le « ein » et le « ym » (ces graphèmes pourront être présentés à partir du deuxième cycle et acquis par procédés visuo-orthographiques (Daigle et al., 2018).

Afin de bien tenir compte de la consistance des graphèmes associés au phonème multigraphémique /*ẽ*/, il faut prendre certaines données statistiques en considération. En effet, la consistance varie selon la position des graphèmes qui y sont associés. La tableau 5.2 présente la consistance des graphèmes associés au phonème multigraphémique /*ẽ*/ selon la position qu'ils occupent dans les mots.

**Tableau 5.2 Consistance des graphèmes associés au phonème multigraphémique /ɛ̃/ selon la position qu’ils occupent dans les mots – Adapté d’ÉQOL-Infra-Lemme (Peereman et Stanké, 2022)**

Phonème	Graphème	Exemple	Consistance en position initiale	Consistance en position interne	Consistance en position finale
/ɛ̃/	in	indice	<b>87,432</b>	45,349	45,714
/ɛ̃/	en	chien		6,977	29,286
/ɛ̃/	ain	main	0,546	15,116	17,857
/ɛ̃/	im	impact	12,022	9,302	
/ɛ̃/	ein	plein		12,791	1,429
/ɛ̃/	ym	symbole		5,814	
/ɛ̃/	yn	synthèse		3,488	
/ɛ̃/	aim	faim			0,714
/ɛ̃/	in	coïncidence		1,163	
/ɛ̃/	ent	va-et-vient			0,714
/ɛ̃/	eint	éteint			0,714
/ɛ̃/	aint	saint			0,714
/ɛ̃/	ingt	vingt			0,714

*Note* : Les morphèmes flexionnels ne font pas partie d’ÉQOL-Infra-Lemme. Précédé du « o », le graphème « in » associé au phonème /ɛ̃/ a une consistance de 100 %. Précédé du « i » ou du « y », le graphème « en » associé au phonème /ɛ̃/ a aussi une consistance de 100 %.

À partir de ce tableau, nous pourrions :

- 1- Dès le premier cycle, enseigner les trois graphèmes les plus consistants : « in », « en » (précédé du « i ») et « ain »;
- 2- Dès le premier cycle, enseigner la consistance du graphème « in » en début de mot puisque, sauf dans le cas du mot « ainsi », il n’y a aucun mot qui commence par « ain ». Le graphème « in » est employé dans la première syllabe dans une consistance de 87,43 % ;
- 3- Dès le premier cycle, expliquer les graphèmes consistants s’expliquant par le voisinage : « ien » et « oin »;
- 4- Dès le premier cycle, expliquer les graphèmes dérivables associés au phonème multigraphémique : « in » dérivable vers un mot en « in... » (par exemple, « lapin » -

- « lapine ») et « ain » dérivable vers un mot en « a » (par exemple, « main » - « manuel ») ou en « ai » (par exemple, « certain », « certaine »);
- 5- À partir du deuxième cycle, expliquer les graphèmes contenant des régularités contextuelles « im ». Dès ce deuxième cycle, expliquer la dérivation possible du mot « faim » vers « affamé »;
  - 6- Les graphèmes ne s'expliquant pas pourront être présentés à partir du troisième cycle et acquis par procédés visuo-orthographiques;
  - 7- Ces graphèmes plus exceptionnels pourront être présentés à partir du troisième cycle et acquis par procédés visuo-orthographiques : ein, ym, yn, ïn;
  - 8- Si le « in » est suivi d'une lettre muette finale, il sera enseigné au deuxième cycle lorsque la dérivation s'explique et il sera enseigné au troisième cycle de façon visuo-orthographique si la dérivation ne s'explique pas.

Concrètement, à partir de tous les résultats de notre recherche doctorale, des mots d'ÉQOL et des données d'ÉQOL-Infra, nous proposons une gradation des mots contenant le phonème multigraphémique /ɛ̃/ à acquérir du premier au troisième cycle du primaire. Le tableau 5.3 présente ces choix.

**Tableau 5.3 Mots contenant le phonème multigraphémique /ɛ̃/ à acquérir du premier au troisième cycle du primaire à partir des résultats de la présente recherche ainsi que des mots d'ÉQOL et des données d'ÉQOL-Infra**

Premier cycle							
Mot	Phonèmes (API)			Nombre de lettres	Facteur linguistique		
bien	/bjɛ̃/			4			
chien	/ʃjɛ̃/			5			
coin	/kwɛ̃/			4			
fin	/fɛ̃/			3	finale		
<i>incas</i>	/ɛ̃ka/			5			
lien	/ljɛ̃/			4			
loin	/lwɛ̃/			4			
lutin	/lytɛ̃/			5			
main	/mɛ̃/			4	manuel		
marin	/marɛ̃/			5			
matin	/matɛ̃/			5			
pain	/pɛ̃/			4			
rien	/rjɛ̃/			4			
sapin	/sapɛ̃/			5			
train	/trɛ̃/			5			
Légende							
Aucun facteur linguistique	ien	oin	in dérivable	ain dérivable	in ou im toujours en début de mot sauf ainsi	im, aim, ym si b ou p	prin : vient de primus, premier

Deuxième cycle							
Mot	Phonèmes (API)			Nombre de lettres	Facteur linguistique		
afin	/afɛ̃/			4			
besoin	/bɛzwɛ̃/, /bøzwɛ̃/			6			
bientôt	/bjɛ̃to/			7			
cartésien	/kartɛzjɛ̃/			9			
certain	/sɛrtɛ̃/			7			
chemin	/ʃɛmɛ̃/			6			
combien	/kɔ̃bjɛ̃/			7			
dessin	/desɛ̃/			6			
<i>enfin</i>	/ɑ̃fɛ̃/			5			
féminin	/feminɛ̃/			7			
humain	/ymɛ̃/			6			
imparfait	/ɛ̃parfɛ/			9			
important	/ɛ̃pɔ̃rtɑ̃/			9			
indicatif	/ɛ̃dikatif/			9			
indice	/ɛ̃dis/			6			
indiquer	/ɛ̃dike/			9			
industrie	/ɛ̃dystri/			9			
inférieur	/ɛ̃ferjœr/			9			
infinitif	/ɛ̃finitif/			9			
inscrire	/ɛ̃skRiR/			8			
insecte	/ɛ̃sɛktə/, /ɛ̃sɛkt/			7			
intendant	/ɛ̃tɑ̃dɑ̃/			9			
intérieur	/ɛ̃terjœr/			9			
intrus	/ɛ̃tRy/			6			
invention	/ɛ̃vɑ̃sjɔ̃/, /ɛ̃vɑ̃tjɔ̃/			9			
masculin	/maskylɛ̃/			8			
médecin	/mɛdsɛ̃/			7			
moins	/mwɛ̃/			5			
point	/pwɛ̃/			5			
soudain	/sudɛ̃/			7			
terrain	/tɛrɛ̃/, /tɛRɛ̃/			7			
Légende							
Aucun facteur linguistique	ien	oin	in dérivable	ain dérivable	in ou im toujours en début de mot sauf ainsi	im, aim, ym si b ou p	prin : vient de primus, premier

Troisième cycle							
Mot	Phonèmes (API)	Nombre de lettres	Facteurs linguistiques				
ainsi	/ɛ̃si/	5					
amérindien	/amerɛ̃djɛ̃/	10					
atteindre	/atɛ̃drɔ̃/,/atɛ̃dR/	9					
<i>cinq</i>	/sɛ̃k/	4					
cinquante	/sɛ̃kɑ̃t/	9					
demain	/dɛ̃mɛ̃/	6					
européen	/øRɔ̃pẽ/,/øRɔ̃pẽ/	8					
impossible	/ɛ̃pɔ̃siblɔ̃/,/ɛ̃pɔ̃sibl/	10					
information	/ɛ̃fɔ̃Rmasjɔ̃/	11					
intertitre	/ɛ̃tɛ̃RtitRɔ̃/	10					
maintenant	/mɛ̃tnɑ̃/,/mɛ̃tɔ̃nɑ̃/	10					
<i>moyen</i>	/mwajɛ̃/	5					
peinture	/pɛ̃tyR/	8					
prince	/pRɛ̃s/	6					
princesse	/pRɛ̃sɛs/	9					
principal	/pRɛ̃sipal/	9					
principaux	/pRɛ̃sipɔ̃/	10					
printemps	/pRɛ̃tɑ̃/	9					
province	/pRɔ̃vɛ̃s/,/pRɔ̃vɛ̃s/	8					
simple	/sɛ̃plɔ̃/,/sɛ̃pl/	6	im				
singulier	/sɛ̃gylje/	9					
symbole	/sɛ̃bɔ̃l/	7					
vingt	/vɛ̃/	5					
Légende							
Aucun facteur linguistique	ien	oin	in dérivable	ain dérivable	in ou im toujours en début de mot sauf ainsi	im, aim, ym si b ou p	prin : vient de primus, premier

## 5.5 Limites

La présente recherche comporte certaines limites méthodologiques. La première concerne le choix des mots contenant le phonème /ɛ̃/. Puisque notre étude s'inscrit dans le cadre d'une autre recherche, les mots qui ont été choisis ne constituent pas un échantillon parfait des graphèmes associés à ce phonème. Toutefois, cela constitue aussi une force puisque les mots qui ont été donnés en dictée ne contenaient pas que l'inconsistance phonographémique étudiée et que l'attention des élèves n'a donc, en aucun cas, été attirée par le sujet. Plusieurs études menées auprès de patients cérébrolésés ont montré que la production d'un mot écrit pouvait être affectée par les mots qui le précèdent; c'est ce qu'on appelle l'effet d'amorçage (Fayol et Jaffré, 2008). La méthodologie de la présente recherche a donc évité l'effet d'amorçage entre les phonèmes multigraphémiques puisque nous n'avons pas donné deux mots rapprochés contenant le phonème /ɛ̃/ pour que les élèves ne soient pas tentés d'orthographier le phonème de la même façon aux deux endroits. De plus, certains graphèmes ont été testés à partir de peu de mots. C'est le cas du « ein », du « im » et du « ain » non dérivable. Toutefois, si l'on considère que ces graphèmes sont les moins fréquents associés au phonème multigraphémique /ɛ̃/, la proportion de mots retenus est donc correcte.

Par ailleurs, le fait que les élèves n'ont pas tous orthographié les mêmes mots ou le même nombre de mots aurait aussi pu constituer une limite. Du moins, cela a représenté un défi important pour le traitement statistique des données. Cela dit, notre modèle d'analyse multivariée a bien compensé cette limite. Le nombre de 50 élèves par niveau peut aussi sembler insuffisant pour généraliser nos conclusions; il s'agit d'une limite de la présente recherche. Toutefois, des statisticiens qui ont réalisé des recherches importantes ont jugé que ce nombre est suffisant, notamment Pothier et Pothier (2004).

De plus, certains biais de la recherche initiale, qui a permis de récolter les mots étudiés, demeurent. C'est notamment le cas de la provenance des mots testés. En effet, bien que l'équipe de chercheurs ait sollicité toutes les maisons d'édition afin d'avoir accès aux versions Word des ouvrages pour mesurer la fréquence des mots, une seule a accepté, soit ERPI. Le nombre de manuels scolaires de mathématiques, d'univers social et de français est tout de même assez élevé (28) pour que la fréquence des mots établie soit crédible. Il en va de même pour le nombre de livres de littérature jeunesse (24) ainsi que d'extraits de littérature jeunesse (90).

Malgré ces limites, le choix du phonème multigraphémique /*ẽ*/ apporte une excellente validité de construit puisqu'il contient tous les facteurs linguistiques susceptibles d'influencer la réussite d'autres phonèmes multigraphémiques, comme le /*o*/, qui correspond, entre autres, aux graphèmes « o », « au » et « eau ». En effet, le phonème multigraphémique /*ẽ*/ permet de vérifier les facteurs sous-lexicaux liés au type de graphème, à la consistance phonographémique, au positionnement du graphème, aux régularités morphologiques dérivationnelles, aux régularités graphotactiques liées au positionnement, aux graphèmes contextuels et au voisinage du graphème.

Dans le cas de la présente recherche faite dans un milieu naturel, il est difficile de s'assurer que les facteurs linguistiques (variables indépendantes) ont été la seule cause de la réussite du graphème (variable dépendante). Nous n'avons, en effet, pas pu savoir si certains des élèves de l'échantillon retenu avaient eu accès à un enseignement explicite des phonèmes multigraphémiques, ce qui pourrait avoir amélioré leurs résultats. Toutefois, plusieurs recherches confirment que l'enseignement explicite de l'orthographe lexicale est encore peu réalisé dans les écoles, les enseignants étant peu outillés pour mettre en pratique les résultats des recherches en orthographe lexicale (Adoniou, 2013; Daigle et al., 2015; Mansour, 2012). Nous pouvons donc inférer que le pourcentage d'élèves de troisième année du primaire qui auraient pu recevoir un enseignement explicite des divers graphèmes associés au phonème multigraphémique /*ẽ*/ est négligeable.

Enfin, nous n'avons pas vérifié si notre échantillon est comparable à la population; il s'agit d'une limite qui a un effet sur la généralisation des résultats. D'autres études doivent permettre de répondre à cette exigence méthodologique. Jusqu'à ce moment, une prudence s'impose donc quant à la généralisation des résultats à l'ensemble de la population d'élèves québécois de troisième année du primaire. Cependant, le choix de l'échantillon des élèves de troisième année du primaire est judicieux puisque ces élèves ont accès aux traitements logographique, alphabétique, orthographique et morphographique, ce qui permet de faire des liens avec tous les cycles du primaire.

D'ailleurs, concernant l'échantillon du sous-groupe d'élèves faibles orthographieurs, ce dernier représente seulement 10 % du nombre de participants. Cela dit, ce pourcentage est relativement représentatif de la population donnée. En effet, si le pourcentage d'élèves handicapés ou en difficulté d'adaptation ou d'apprentissage (EHDAA) est estimé à 19,1 % au Québec (commission

de la culture et de l'éducation, 2015) et que le trouble des apprentissages spécifiques de la lecture et de l'expression écrite (dyslexie/dysorthographe) représente 80 % des troubles d'apprentissage chez les enfants d'âge scolaire (Neault et Guay, 2006; Shaywitz et al., 1995), nous estimons le pourcentage de faibles orthographieurs à plus ou moins 15 %.

En résumé, malgré certaines limites, nous avons appliqué de hauts critères de signification, de sorte que les conclusions liées aux hypothèses s'appuient sur de solides données d'analyse.

## **5.6 Futures études**

Cette recherche, qui se situe à la jonction des sciences de l'éducation et de la linguistique appliquée, a des incidences importantes, tant sur le plan scientifique que sur les plans pratique et social. Sur le plan scientifique, elle permet d'en savoir davantage sur l'acquisition des phonèmes multigraphémiques chez les élèves québécois faibles et bons orthographieurs. Sur le plan social, des propositions concrètes ressortent des résultats de recherche, tant pour les bons orthographieurs que pour les faibles orthographieurs, favorisant ainsi l'acquisition de l'orthographe lexicale et la réussite scolaire de tous les élèves. Ces nouvelles données québécoises en orthographe lexicale permettront éventuellement au personnel enseignant de faire des choix dans l'enseignement des phonèmes multigraphémiques. En effet, d'un point de vue pratique, ces nouvelles connaissances pourraient ultérieurement permettre de corriger quelques incongruités dans les listes orthographiques actuelles et pourraient éventuellement mener à une première échelle d'acquisition des phonèmes multigraphémiques basée sur une recherche doctorale au Québec, par exemple. Conséquemment, l'apprentissage de l'orthographe lexicale des élèves du primaire serait facilité puisqu'ils pourraient s'approprier l'orthographe de façon mieux réfléchie.

Les études corrélationnelles rassemblent de nombreuses données. Il est donc fréquent qu'elles mènent à d'autres recherches, qui peuvent être de plus grande envergure (Lehrer et al., 2021). C'est le cas de la présente recherche. Plusieurs études restent à faire sur l'acquisition des phonèmes multigraphémiques. Des recherches-développements pourraient permettre de développer des listes graduées d'enseignement des phonèmes multigraphémiques à partir de nos résultats. De plus, des recherches-actions pourraient permettre de les expérimenter en identifiant les bonnes stratégies

d'implantation, entre autres. Enfin, des recherches expérimentales pourraient permettre de préciser les présentes conclusions, sinon, d'aller plus loin.

Enfin, les conclusions de la présente recherche doctorale sont, certes, une motivation à nous questionner sur d'autres éléments. En effet, l'analyse de l'acquisition des graphèmes associés au phonème /*ẽ*/ pourrait servir de modèle à l'acquisition des autres phonèmes multigraphémiques. Ainsi, les enseignants seraient mieux outillés en matière d'enseignement des inconsistances phonographémiques et l'apprentissage de l'orthographe lexicale des élèves du primaire serait facilité. Il serait aussi pertinent d'étudier l'acquisition des phonèmes multigraphémiques des mots concrets et des mots abstraits. Par exemple, parmi les mots retenus contenant le phonème /*ẽ*/, 14 ont un côté figuratif concret : « amérindiens », « chien », « chemin », « dessin », « lutin », « marins », « médecin », « sapin », « incas », « insectes », « prince », « princesse », « coin », « points ». Il pourrait être intéressant d'en tenir compte dans une future recherche, surtout lorsqu'on sait que seulement 4 élèves sur 76 ont fait une erreur dans le graphème « en » du mot « chien », mais que 22 élèves sur 78 ont fait une erreur dans le graphème « en » du mot « lien ». Fayol et Jaffré (2008) confirment d'ailleurs que les jeunes scripteurs tendent à représenter plus facilement, d'abord par des pictogrammes, les éléments concrets comme « chien » ou « chat » que les éléments plus abstraits comme « sur » ou « dans ». Comme nous savons que l'écriture est en fait une série de signes (Saussure, 1985), il est possible de faire un lien entre la représentation de l'abstraction par des pictogrammes et la représentation de l'abstraction par des lettres. Le modèle de Gombert donne d'ailleurs une place importante au traitement sémantique et affirme que les non-mots et les mots inconnus ne peuvent pas être lus par les enfants au début de leur acquisition de l'écrit puisque les mots doivent avoir un sens pour qu'ils arrivent d'abord à les traiter (Demont et Gombert, 2004). Pothier et Pothier (2004) ont aussi observé que les mots ayant une forte signification émotive pourraient être acquis plus rapidement.

## CONCLUSION

Ainsi, la qualité scientifique de cette recherche doctorale repose sur l'arrimage de concepts théoriques sur l'acquisition de l'orthographe lexicale afin de produire un apport scientifique différent. En comparant les données de Pacton, de Sprenger-Charolles, de Pothier et Pothier, de Mousty et Leybaert et de plusieurs autres chercheurs en orthographe lexicale à nos résultats, nous sommes parvenus à mieux comprendre ce qui est à la base des difficultés linguistiques associées à l'acquisition des phonèmes multigraphémiques en orthographe lexicale au Québec. Il s'agit d'une contribution importante, tant du point de vue théorique que du point de vue social. Concrètement, ces nouvelles données québécoises en orthographe lexicale permettront éventuellement au personnel enseignant de faire des choix dans l'enseignement des inconsistances phonographémiques.

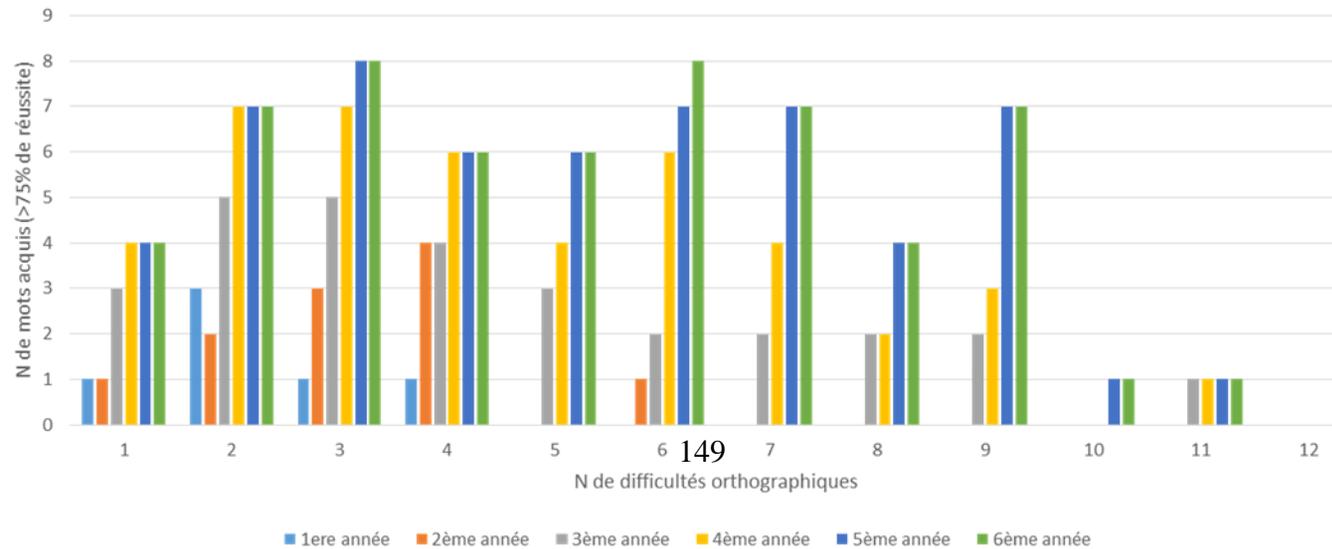
À notre connaissance, la présente recherche est la première à s'être intéressée à l'acquisition des phonèmes multigraphémiques chez les élèves québécois. Les résultats montrent que la fréquence et la consistance des mots ne jouent pas un rôle significatif dans l'acquisition de ces phonèmes. Toutefois, la consistance et le positionnement des graphèmes joue un grand rôle dans leur acquisition. De plus, les résultats suggèrent que les élèves faibles orthographes réussissent beaucoup mieux à orthographier les graphèmes associés aux phonèmes multigraphémiques lorsqu'ils sont dérivables et que les mots sont courts. Toutefois, contrairement aux élèves bons orthographes, ils peinent à intégrer les régularités graphotactiques contextuelles.

L'ordre d'acquisition des graphèmes associés au phonème /*ɛ̃*/, tant pour les élèves bons orthographes que pour les élèves faibles orthographes, pourra guider l'ordre de l'enseignement des autres inconsistances phonographémiques comme le /*o*/, pouvant être représenté par plusieurs graphèmes dont le « o », le « au » et le « eau », le /*ɛ*/ correspondant, entre autres, aux graphèmes « e », « è » et « ai » ainsi que le /*ã*/, correspondant aux graphèmes « an » et « en ». Ces perspectives didactiques et orthodidactiques sont complexes et nécessitent une recherche en collaboration avec le milieu. À cet effet, des recherches-actions devront être éventuellement entreprises.

# ANNEXE A

Extrait du prétest visant à analyser l'effet du nombre de difficultés par mots sur la réussite de la 1<sup>re</sup> à la 6<sup>e</sup> année du primaire

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28					
Mot	Régularité	Phonèmes	rythme	Lettre	phon	% de la graphie de phonème	Cooccurrences	Fréquences	Degré de	Double	Règles	Graphes	Sémi	Lettres	Lettres	Phon	Struct	Régul	Confe	Nomb	Réussite	Freq										
Variables : Difficultés orthographiques																																
Variables : Acquisition - Vent = acquis (75%)																																
certain	ain morpho	/ssu#f/	2	8	3	28%	62%	à venir	3	0	0	1	0	0	0	1	0	0	0	0	0	5	36.07%	-	25.26%	36.25%	48.42%	34.95%	58.00%	0.2887		
humain	ain morpho	/jrn#f/	2	6	3	28%	72.72%	à venir	1	0	0	0	0	1	0	0	0	0	0	0	0	2	4	59.57%	10.72%	45.05%	91.94%	95.40%	96.22%	100.00%	0.1851	
main	ain morpho	/m#f/	1	4	2	28%	59.55%	à venir	1	0	0	0	0	0	0	0	0	0	0	0	0	1	3	97.95%	54.72%	71.63%	90.04%	90.13%	90.00%	100.00%	0.1773	
pain	ain morpho	/p#f/	1	4	2	28%	59.55%	à venir	1	0	0	0	0	0	0	0	0	0	0	0	0	1	3	96.84%	50.98%	69.00%	89.08%	100.00%	97.22%	96.77%	0.0902	
soudain	ain morpho	/su#f/	2	7	4	28%	74.74%	à venir	2	0	0	0	0	0	0	0	0	0	0	0	0	1	3	62.65%	9.26%	31.37%	64.81%	74.60%	92.19%	94.44%	0.1244	
terrain	ain morpho	/tst#f/	2	7	4	28%	33.26%	à venir	2	1	0	0	0	0	0	0	0	0	0	0	0	0	3	59.12%	-	13.46%	68.52%	62.54%	94.38%	94.44%	0.0777	
train	ain morpho	/tst#f/	1	5	3	28%	71.04%	à venir	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	4	94.24%	35.23%	76.98%	83.78%	92.91%	98.40%	95.28%	0.1213
ainsi	ain non spec	/ts#f/	2	5	3	28%	47.61%	à venir	2	0	0	1	0	0	0	0	0	0	0	0	1	0	1	5	39.85%	11.76%	16.88%	10.53%	53.49%	67.14%	68.25%	0.3981
déjà	ain non spec	/dm#f/	2	6	4	28%	73.75%	à venir	1	0	0	0	0	0	0	0	0	0	0	0	0	1	3	73.28%	22.54%	67.39%	79.98%	96.81%	98.00%	0.0591		
maintenant	ain non spec	/mt#f/	3	10	6	28%	60.81%	à venir	2	0	0	0	0	2	0	0	0	0	0	0	0	2	6	49.42%	4.00%	18.81%	32.76%	61.84%	88.38%	95.71%	0.3224	
atteindre	ein non spec	/at#f/	3	9	5	18%	63.29%	à venir	1	1	0	0	0	1	0	0	0	0	0	0	1	0	4	25.59%	-	3.12%	14.28%	38.32%	67.37%	82.54%	0.0933	
peinture	ein non spec	/pt#f/	3	8	5	18%	70.0%	à venir	2	0	0	0	0	1	0	0	0	0	0	0	0	1	2	64.55%	-	44.2%	52.50%	81.00%	91.26%	91.00%	0.028	
européens	en non spec	/e#f/	4	8	6	23%	71.88%	à venir	1	0	0	0	0	0	0	1	1	0	0	0	0	1	2	7	13.90%	145%	2.22%	8.22%	6.45%	38.98%	40.00%	0.1088
américains	ien	/ame#f/	4	11	8	23%	nd	à venir	2	0	0	0	1	0	0	1	0	0	0	0	0	1	4	9	58.91%	8.18%	38.64%	56.10%	73.17%	92.98%	93.33%	0.3483
appartement	ien	/apant#f/	3	10	7	23%	63.83%	à venir	1	1	0	0	1	0	0	0	0	0	0	0	0	2	2	8	25.59%	-	1.92%	35.18%	20.63%	40.63%	51.85%	0.0684
bien	ien	/b#f/	1	4	3	23%	78.40%	à venir	1	0	0	0	1	0	0	0	0	0	0	0	0	1	4	88.22%	42.03%	66.18%	93.55%	98.28%	96.74%	98.13%	1.104	
bien-être	ien	/b#f/	2	7	5	23%	57.59%	à venir	2	0	0	0	1	1	0	0	0	0	0	0	0	1	1	6	61.89%	5.88%	30.21%	54.90%	86.41%	96.86%	96.00%	0.1951
cartésien	ien	/k#f/	3	9	8	23%	nd	à venir	3	0	0	2	1	0	0	0	0	0	0	0	0	0	6	44.72%	7.94%	23.08%	51.43%	60.24%	58.97%	82.49%	0.1213	
claire	ien	/l#f/	1	5	3	23%	78.38%	à venir	1	0	0	0	1	0	0	0	0	0	0	0	0	1	3	94.01%	99.18%	96.92%	97.01%	98.99%	100.00%	98.00%	0.1959	
combien	ien	/b#f/	2	7	5	23%	61.62%	à venir	3	0	1	1	1	0	0	0	0	0	0	0	0	1	1	8	77.60%	1.67%	62.50%	87.98%	89.47%	93.60%	99.07%	1.586
content	ien	/b#f/	2	8	5	23%	61.55%	à venir	3	0	0	1	1	0	0	0	0	0	0	0	0	1	7	52.67%	5.88%	28.40%	44.28%	60.98%	73.79%	75.63%	0.4634	
devenir	ien	/d#f/	2	7	5	23%	73.55%	à venir	1	0	0	0	1	0	0	0	0	0	0	0	0	1	1	5	37.69%	7.58%	21.43%	25.37%	39.34%	70.59%	73.08%	0.2301
ien	ien	/f#f/	1	4	3	23%	78.40%	à venir	1	0	0	0	1	0	0	0	0	0	0	0	0	0	2	5	77.70%	36.07%	69.61%	74.63%	93.22%	97.26%	100.00%	0.2426
moyen	ien	/m#f/	2	5	5	23%	82.02%	à venir	2	0	0	0	2	0	0	0	0	0	0	0	0	0	4	61.56%	-	28.28%	66.67%	81.71%	95.40%	93.22%	0.2674	
obtient	ien	/p#f/	2	7	5	23%	54.81%	à venir	1	0	0	0	1	0	0	0	1	0	0	0	0	1	5	34.07%	-	5.77%	34.55%	24.66%	60.26%	69.81%	0.0995	
rien	ien	/r#f/	1	4	3	23%	78.40%	à venir	1	0	0	0	1	0	0	0	0	0	0	0	0	0	2	5	88.24%	34.07%	82.31%	96.20%	100.00%	100.00%	100.00%	0.3096



**ANNEXE B**

**Extrait du traitement personnel des statistiques sur Excel**

Mot	Fact. env.			Fact. ling.							Fact. intraling.												Performance			Production Orthog. mot et graphème		Réuss. Orthog. semi-voyelle									
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	A	A	A	A	A	AG	AH	AI	AJ	
afin	7	2	sans_diff	7	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99
afin	7	2	sans_diff	7	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99
afin	7	2	sans_diff	7	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99
afin	7	1	sans_diff	7	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	0	affin	INS	99	99
afin	7	1	sans_diff	7	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99
afin	7	1	sans_diff	7	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99
afin	7	2	en_diff	7	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99
afin	7	1	en_diff	7	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99

Mot	Fact. env.	Fact. ind.	Fact. ling.										Fact. intraling.										Performance			Production Orthog. mot et graphème		Réuss. Orthog. semi-voyelle										
			7	9	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99		
afin	7	1	sans_diff	7	9	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99
afin	7	1	sans_diff	7	9	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99
afin	7	1	sans_diff	7	9	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99
afin	7	1	sans_diff	7	9	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99
afin	7	1	en_diff	7	9	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	0	afin	INS	99	99
afin	7	2	sans_diff	7	9	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	0	afin	INS	99	99
afin	7	1	en_diff	7	9	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99
afin	7	1	sans_diff	7	9	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99
afin	7	2	en_diff	7	9	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99
afin	7	1	sans_diff	7	9	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99
afin	7	1	sans_diff	7	9	0,1467953	0,4403859	4	2	48,143	36,3	9,9	9,9	10,7	0	0	1	Final	1	0	0	0	0	0	0	0	0	0	non_deriv	99	IN	1	1	1	afin	INS	99	99

### Légende

- A) mot
- B) milieu socio-économique
- C) sexe
- D) profil en orthographe
- E) tauxreusmot3
- F) fréquence du mot de 1<sup>re</sup> à 3<sup>e</sup> année
- G) fréquence du mot 3<sup>e</sup> année
- H) nombre de lettres dans le mot
- I) nombre de syllabes dans le mot
- J) consistance du mot
- K) consistance totale du graphème
- L) consistance initiale du graphème

- M) consistance interne du graphème
- N) consistance finale du graphème
- O) positionnement initial du graphème
- P) positionnement interne du graphème
- Q) positionnement final du graphème
- R) Positionnement du graphème
- S) in
- T) im
- U) ain
- V) ain
- W) im
- X) en

- Y) semi-voyelle
- Z) graphème dérivable
- AA) graphème dérivable IN
- AB) graphème dérivable AIN
- AC) régularité
- AD) réussite orthographique graphème
- AE) réussite phonologique graphème
- AF) réussite mot
- AG) production mot
- AH) production graphème
- AI) réussite o dans ain
- AJ) réussite i dans ien

## ANNEXE C

### Table des matières de la première analyse statistique

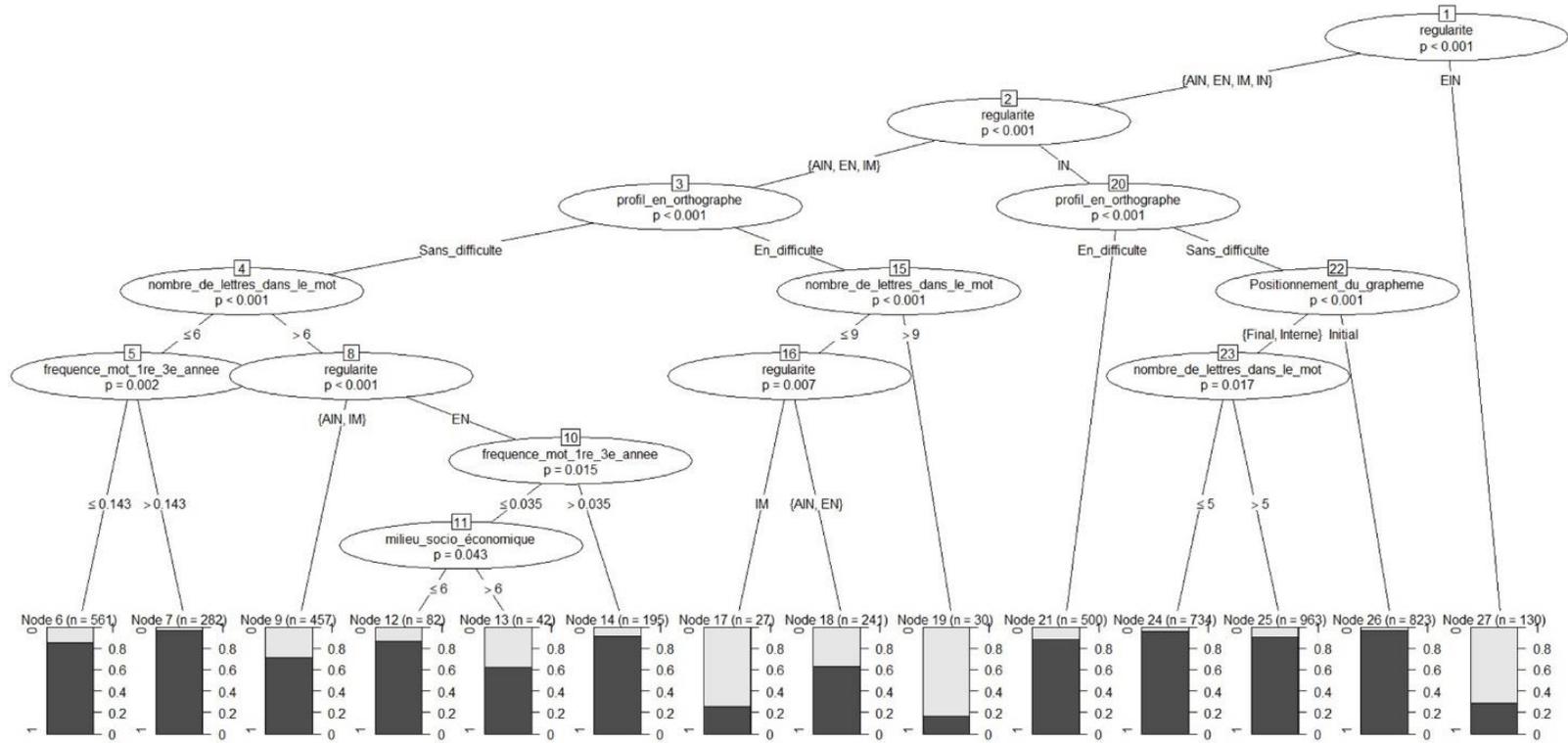
#### Table des matières

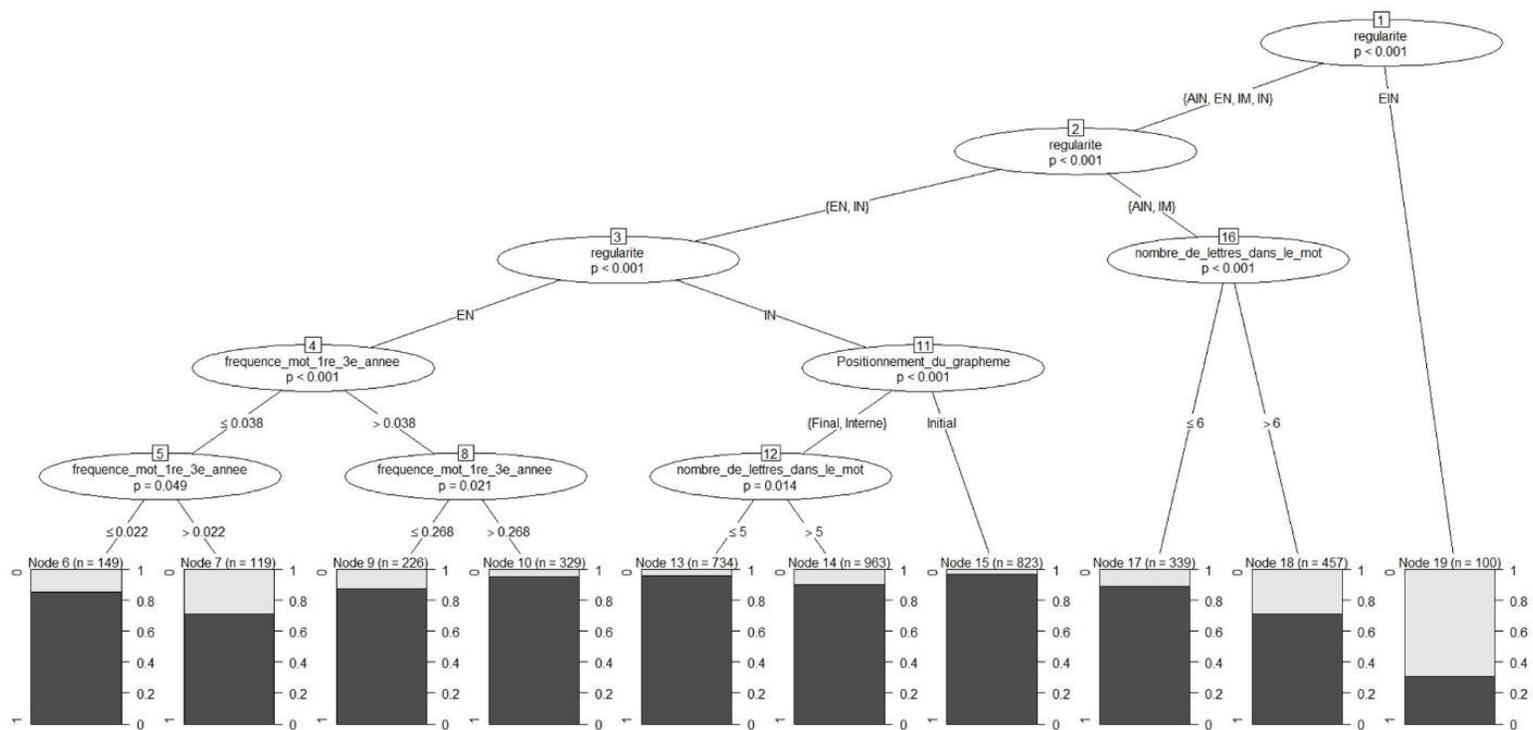
1.	Méthodologie .....	4
2.	Statistiques descriptives .....	4
	Sur les écoles .....	4
	Sur les enfants .....	6
3.	Réussite du mot-cible par l'élève .....	7
	Élèves en difficulté .....	12
	Analyses univariées .....	12
	Sélection de variables .....	13
	Élèves sans difficulté .....	14
	Analyses univariées .....	14
	Sélection de variables .....	15
	Tous les élèves .....	16
	Analyses univariées .....	16
	Sélection de variables .....	17
4.	Réussite orthographique du graphème associé au phonème [e] par l'élève .....	19
	Élèves en difficulté .....	23
	Analyses univariées .....	23
	Sélection de variables .....	24
	Élèves sans difficulté .....	25
	Analyses univariées .....	25
	Sélection de variables .....	26
	Tous les élèves .....	27
	Analyses univariées .....	27
	Sélection de variables .....	28
5.	Réussite orthographique du reste du mot (sauf la CPG [e]) par l'élève .....	30
	Élèves en difficulté .....	34
	Analyses univariées .....	34
	Sélection de variables .....	35
	Élèves sans difficulté .....	36
	Analyses univariées .....	36
	Sélection de variables .....	37
	Tous les élèves .....	38
	Analyses univariées .....	38
	Sélection de variables .....	39
6.	Réussite phonologique du graphème associé au phonème [e] par l'élève .....	41
	Élèves en difficulté .....	45
	Analyses univariées .....	45
	Sélection de variables .....	46
	Élèves sans difficulté .....	47
	Analyses univariées .....	47
	Sélection de variables .....	48
	Tous les élèves .....	49
	Analyses univariées .....	49
	Sélection de variables .....	50
7.	Réussite phonologique du reste du mot (sauf la CPG [e]) par l'élève .....	51
	Élèves en difficulté .....	55
	Analyses univariées .....	55
	Sélection de variables .....	56
	Élèves sans difficulté .....	57
	Analyses univariées .....	57
	Sélection de variables .....	58
	Tous les élèves .....	59
	Analyses univariées .....	59
	Sélection de variables .....	60
8.	Réussite orthographique de la semi-voyelle O dans OIN par l'élève .....	61

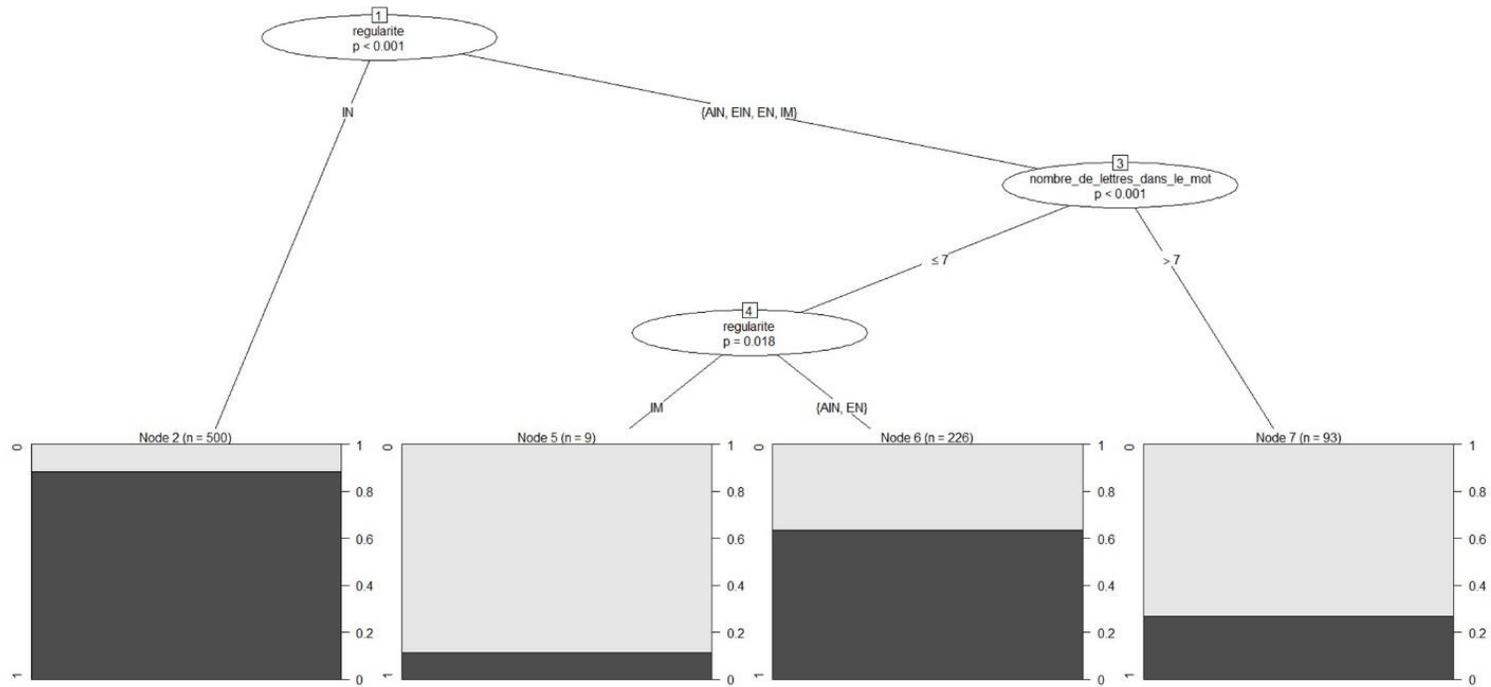
Élèves en difficulté.....	65
Analyses univariées.....	65
Sélection de variables.....	65
Élèves sans difficulté.....	66
Analyses univariées.....	66
Sélection de variables.....	66
Tous les élèves.....	67
Analyses univariées.....	67
Sélection de variables.....	67
9. Réussite orthographique du yod I ou Y dans IEN ou YEN par l'élève....	68
Élèves en difficulté.....	72
Analyses univariées.....	72
Sélection de variables.....	72
Élèves sans difficulté.....	73
Analyses univariées.....	73
Sélection de variables.....	73
Tous les élèves.....	74
Analyses univariées.....	74
Sélection de variables.....	75
10. Réussite orthographique du GT dans VINGT.....	76
Élèves en difficulté.....	80
Analyses univariées.....	80
Sélection de variables.....	80
Élèves sans difficulté.....	81
Analyses univariées.....	81
Sélection de variables.....	81
Tous les élèves.....	81
Analyses univariées.....	81
Sélection de variables.....	81
11. Réussite orthographique du morphème final par l'élève dans le cas de la présence d'un morphème dérivationnel après le graphème associé au phonème [e]	
82	
Élèves en difficulté.....	85
Analyses univariées.....	85
Sélection de variables.....	86
Élèves sans difficulté.....	86
Analyses univariées.....	86
Sélection de variables.....	87
Tous les élèves.....	87
Analyses univariées.....	87
Sélection de variables.....	87

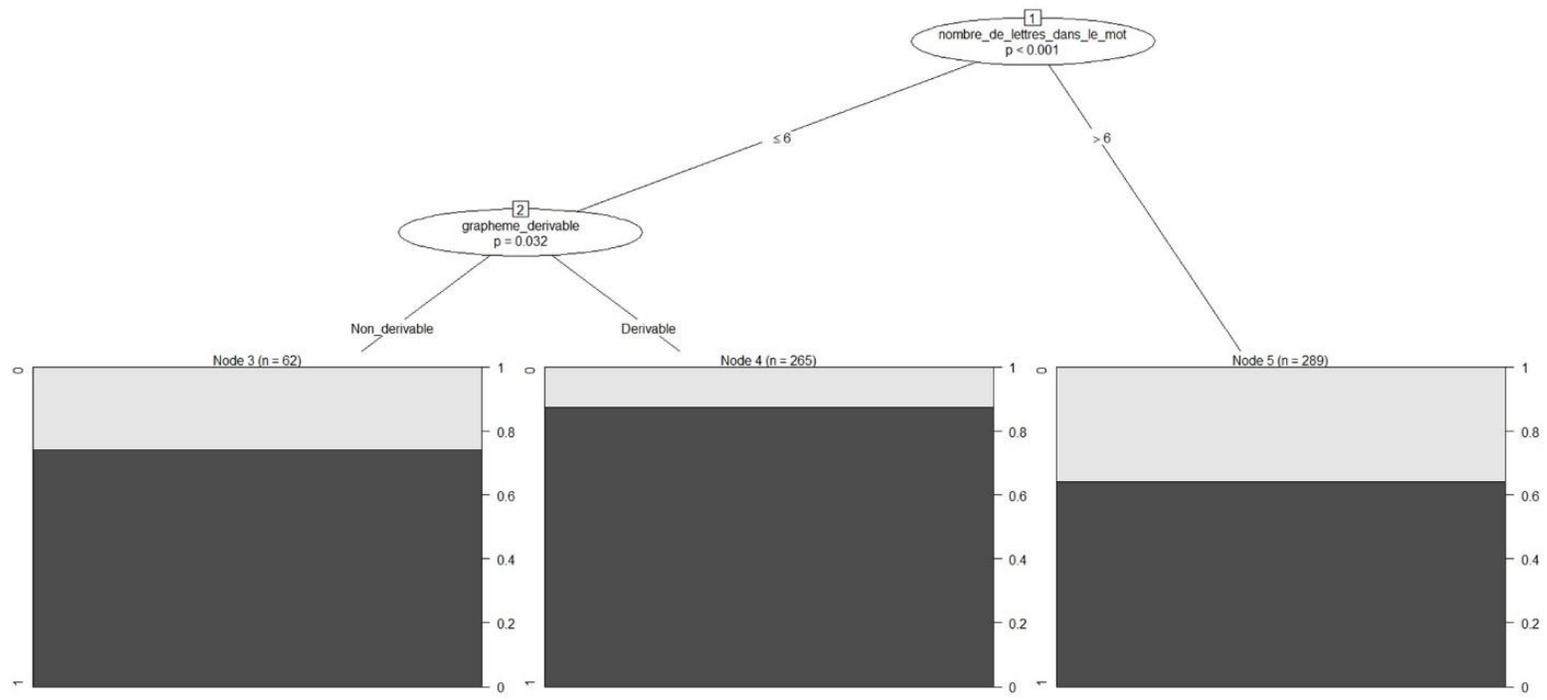
## ANNEXE D

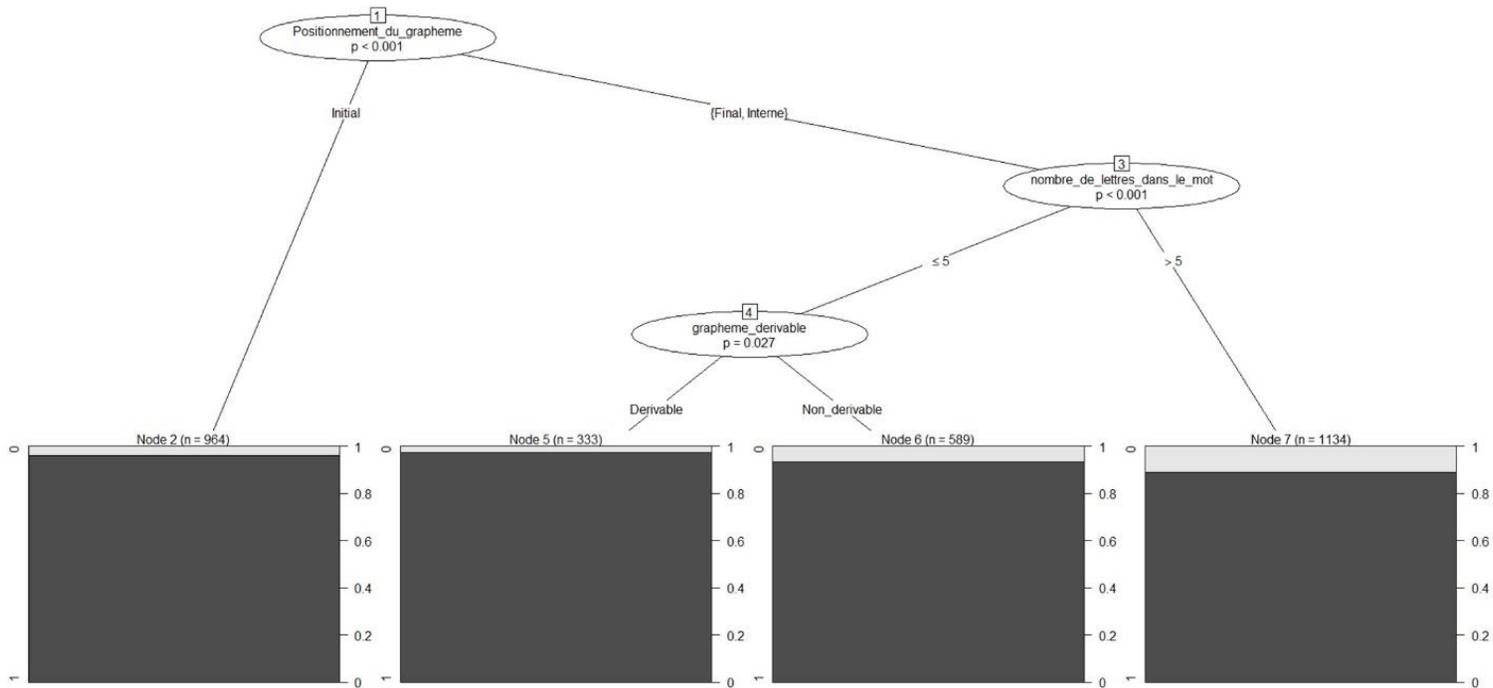
### Arbres produits lors de la deuxième analyse statistique











ANNEXE E

Caractéristiques distributionnelles des correspondances phonographologiques et des unités lexicales dans le lexique de l'école primaire québécoise - Extrait d'ÉQOL-Infra (Peereman et Stanké, 2022)

Frequency and Consistency of Ph-G associations (ver. 1.2)															
				Frequency						Consistency		Consistency		Consistency	
				Initial		Internal		Final		Initial		Internal		Final	
phoneme	grapheme	match	exemple	Type	Token	Type	Token	Type	Token	Type	Token	Type	Token	Type	Token
#	g	g-#	longtemps	0	0	2	183,476	(see note1)	(see note1)			0	0	(see note1)	(see note1)
#	t	t-#	montréalais	0	0	2	18,659	(see note1)	(see note1)			0	0	(see note1)	(see note1)
#	a	a-#	aout	1	46,647	0	0	(see note1)	(see note1)	0	0	0	0	(see note1)	(see note1)
#	w	w-#	clown	0	0	1	80,854	(see note1)	(see note1)			0	0	(see note1)	(see note1)
#	l	l-#	fils	0	0	1	111,952	(see note1)	(see note1)			0	0	(see note1)	(see note1)
%	e	e-%	obtenir	0	0	350	14214,8	9	102457,7			99,715	98,683	100	100

%	on	on-%	monsieur	0	0	1	189,696	0	0			0,285	1,317		
@	en	en-@	entier	104	19457,86	258	8303,116	0	0	66,242	94,688	46,995	45,293		
@	ent	ent-@	vent	0	0	0	0	307	8931,297					56,022	27,259
@	an	an-@	angle	21	677,931	230	7052,985	28	964,034	13,376	3,299	41,894	38,473	5,109	2,942
@	ant	ant-@	important	0	0	0	0	192	6819,757					35,036	20,814
@	em	em-@	emploi	29	373,174	34	2344,769	0	0	18,471	1,816	6,193	12,79		
@	am	am-@	campagne	3	40,427	27	631,286	0	0	1,911	0,197	4,918	3,444		
@	and	and-@	grand	0	0	0	0	5	1666,837					0,912	5,087
@	ang	ang-@	sang	0	0	0	0	3	74,634					0,547	0,228

**ANNEXE F**  
**Approbation éthique**



**Formulaire de demande de renouvellement de l'approbation éthique**

Titre du protocole : **FACTEURS LINGUISTIQUES LIÉS À L'ACQUISITION DES INCONSISTANCES PHONOGRAPHÉMIQUES CHEZ LES ÉLÈVES QUÉBÉCOIS DE LA PREMIÈRE À LA TROISIÈME ANNÉE DU PRIMAIRE**

Numéro(s) de projet : **2021-1503**

Identifiant Nagano : **dutm06**

Chercheur principal (au CER Éval) : **Mélanie Dutablemple**

Date d'approbation du projet par le CER : **2021-02-17**

Formulaire : **F9-10308**

Date de dépôt initial du formulaire :  
**2023-05-25**

Date de dépôt final du formulaire : **2023-05-25**

Statut du formulaire : **Formulaire approuvé**

Suivi du BCER

1. *OBJET: RENOUVELLEMENT DE L'APPROBATION ÉTHIQUE*

2.

*Statut de la demande:*

Demande approuvée

**À la suite du dépôt de votre formulaire de renouvellement, le comité d'éthique de la recherche de l'UQO constate le bon déroulement du projet et vous autorise à poursuivre vos activités de recherche pour une période d'un an.**

*Le renouvellement de votre approbation éthique est valide jusqu'au:*

2024-02-17

**RENOUVELLEMENT ANNUEL: Pour maintenir la validité de votre approbation éthique, vous devez obtenir le renouvellement de votre approbation éthique à l'aide du formulaire F9, et ce avant la date d'échéance. Un rappel automatique vous sera envoyé par courriel quelques semaines avant l'échéance de votre approbation éthique.**

**MODIFICATION: Si des modifications sont apportées à votre projet de recherche, vous devez soumettre les modifications au CER, et ce, AVANT la mise en œuvre de ces modifications en complétant le formulaire F8 - Demande de modification au projet de recherche.**

**FIN DE PROJET: Vous devez remplir le formulaire F10-Rapport final afin d'informer le CER de la fin de votre projet de recherche.**

3.

*La demande a été traitée par :*

Caroline Tardif

**date de traitement:**

2023-05-25

## Section A: Identification

1. **Veillez indiquer le titre complet du projet de recherche.**

*Quel est le titre du projet?*

FACTEURS LINGUISTIQUES LIÉS À L'ACQUISITION DES INCONSISTANCES PHONOGRAPHÉMIQUES CHEZ LES ÉLÈVES QUÉBÉCOIS DE LA PREMIÈRE À LA TROISIÈME ANNÉE DU PRIMAIRE

2. **Veillez indiquer le nom du (de la) chercheur(e) responsable du projet à l'UQO. (L'article 3.1 de la Politique d'éthique de la recherche avec des êtres humains précise qu'un(e) « chercheur(e) » inclut, aux fins des présentes, les professeur(e)s, les étudiant(e)s aux cycles supérieurs ou au premier cycle, ou toute personne impliquée dans les activités de recherche couvertes par la présente Politique.)**

*Qui est le (la) chercheur(e) principal(e) de ce projet à l'UQO?*

Dutemple, Mélanie

3. *En plus du (de la) chercheur(e) principal(e), y a-t-il d'autres personnes dans votre équipe de recherche?*

Oui

Détails cochercheur(e)s	Type de cochercheur
Stanké, Brigitte	Cochercheur(e) - autre université québécoise Établissement - QC Montréal Département Orthophonie

4. **Veillez sélectionner le type de chercheur(e) qui correspond à la situation du (de la) chercheur(e) principal(e).**

*Le (la) chercheur(e) principal(e) est :*

Étudiant(e) de 3e cycle

#### Validation fin de projet

1. **Est-ce que votre projet de recherche est terminé?**

Non

#### Section B: Directeur[s]

1. **Veillez indiquer le nom de votre directeur(trice) de recherche ou des codirecteur(e)s de votre projet. Si un(e) codirecteur(trice) n'est pas professeur(e) de l'UQO, veuillez seulement indiquer son nom ici en l'ajoutant comme contact. Seuls les professeur(e)s de l'UQO peuvent être ajoutés comme utilisateur(trice)s à un projet.**

*Saisir les premières lettres du nom d'abord*

Moreau, André C.

## Section C: Déroulement des travaux

1. **Veillez préciser le statut actuel de la collecte de données en indiquant votre choix ci-dessous.**

*Quel est le statut actuel de la collecte de données?*

- Débutera dans les 12 prochains mois  
 Débutera dans plus d'une année  
 Est en cours  
 Est terminée  
 Le projet n'implique pas de collecte de données, mais plutôt l'utilisation de bases de données impliquant des sujets humains.

2. **Veillez indiquer si des participant(e)s se sont retirés du projet ou si vous avez dû retirer des participant(e)s du projet? Si oui, indiquez pour quelles raisons.**

*Est-ce que des participant(e)s se sont retirés du projet ou avez-vous dû retirer des participant(e)s du projet?*

Non

3. **Veillez indiquer si des participant(e)s ont subi des effets indésirables ou des inconvénients? Si oui, veuillez les décrire et nous indiquer comment il vous a été possible d'y remédier.**

*Est-ce que des participant(e)s ont subi des effets indésirables ou des inconvénients?*

Non

4. **Veillez indiquer si vous avez rencontré des situations où la confidentialité a été compromise? Si oui, dans quelles circonstances et qu'avez-vous pu y faire?**

*Avez-vous rencontré des situations où la confidentialité a été compromise?*

Non

5. **Veillez indiquer si vous avez rencontré d'autres difficultés. Si oui, précisez lesquelles.**

*Avez-vous rencontré d'autres difficultés?*

Non

## Section D: Financement

1. **Veillez indiquer la ou les sources de financement du projet**

- Aucun financement
- FRQNT
- FRQSC
- FRQS
- MITACS
- CRSH
- CRSNG
- IRSC
- Chaire institutionnelle
- Démarrage de projet
- Dépannage
- Contribution institutionnelle (regroupement)
- Équipes
- Fonds de recherche (cours en appoint)
- Fonds de recherche (DFCP)
- Centre de recherche
- Autre ministère ou organisme fédéral
- Autre ministère ou organisme provincial
- Autre

2. **Veillez fournir l'unité budgétaire (si disponible).**

**Section E: Modifications au projet**

1. **Est-ce que votre projet de recherche s'est déroulé comme prévu lors de l'approbation éthique initiale ou en fonction des modifications préalablement apportées et approuvé par le CER?**

Oui

2. **Veillez indiquer si vous envisagez apporter des modifications à votre projet de recherche.**

*Avez-vous l'intention d'apporter des modifications à votre projet de recherche?*

Non

**Section F: Projet sous la responsabilité d'un autre CÉR**

1. **Si votre projet de recherche a fait l'objet d'une évaluation par un autre comité d'éthique que celui de l'UQO. Veillez déposer le document qui démontre que le certificat d'éthique a été renouvelé par l'autre établissement (ex. autres universités, CISSS, etc).**

**Section H: signature du directeur/ codirecteur[s]**

1. **Seuls le (la) directeur(trice) ou les codirecteurs(trices) peuvent signer à cet endroit. LE FORMULAIRE NE DOIT PAS ÊTRE DÉPOSÉ TANT QUE LE (LA) DIRECTEUR(TRICE) DE RECHERCHE N'A PAS SIGNÉ. N'oubliez pas de déposer le formulaire une fois complété.**

*IMPORTANT : Avant de signer et déposer ce formulaire, veuillez vous assurer de bien lire les réponses de l'étudiant(e), car vous partagez la responsabilité du projet avec l'étudiant(e).*

**AVIS AUX ÉTUDIANT(E)S : LE FORMULAIRE NE DOIT PAS ÊTRE DÉPOSÉ TANT QUE LE (LA) DIRECTEUR(TRICE) OU CODIRECTEUR(TRICE) DE RECHERCHE N'A PAS SIGNÉ LE FORMULAIRE. Seul le (la) directeur(trice) ou codirecteur(trice) peut remplir cette section. Si vous signez à la place de votre directeur(trice) OU QUE VOUS DÉPOSÉ LE FORMULAIRE SANS LA SIGNATURE DU (DE LA) DIRECTEUR(TRICE) OU CODIRECTEUR(TRICE) DE RECHERCHE, vous ne ferez que retarder le traitement de votre dossier.**

*Signature électronique du (de la) directeur(trice) ou du (de la) codirecteur(trice) :*

Professeur(e) :  
André C. Moreau  
2023-05-25 13:57

## ANNEXE G

**Mots qui contiennent la prononciation [ɛ̃] dans la liste orthographique à l'usage des enseignantes et des enseignants (ministère de l'Éducation et de l'Enseignement supérieur, 2014) au premier cycle du primaire**

Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :					
Année scolaire 1 <sup>re</sup> et 2 <sup>e</sup>					
Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
certains	[sɛʁtɛ̃]	dét.	m.	pl.	2 <sup>e</sup>
chemin	[ʃəmɛ̃]	n.	m.	s.	1 <sup>re</sup>
chien	[ʃjɛ̃]	n.	m.	s.	1 <sup>re</sup>
cinq	[sɛ̃k]	dét.	m.	pl.	1 <sup>re</sup>
cinq	[sɛ̃k]	dét.	f.	pl.	1 <sup>re</sup>
cinquante	[sɛ̃kɑ̃t]	dét.	m.	pl.	2 <sup>e</sup>
cinquante	[sɛ̃kɑ̃t]	dét.	f.	pl.	2 <sup>e</sup>
cousin	[kuzɛ̃]	n.	m.	s.	2 <sup>e</sup>
enfin	[ɑ̃fɛ̃]	adv.			2 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 1<sup>re</sup> et 2<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
jardin	[ʒardɛ̃]	n.	m.	s.	1 <sup>re</sup>
lapin	[lapɛ̃]	n.	m.	s.	1 <sup>re</sup>
loin	[lwɛ̃]	adv.			2 <sup>e</sup>
main	[mɛ̃]	n.	f.	s.	1 <sup>re</sup>
maintenant	[mɛ̃tənɑ̃]	adv.			2 <sup>e</sup>
matin	[matɛ̃]	n.	m.	s.	1 <sup>re</sup>
moins	[mwɛ̃]	adv.			2 <sup>e</sup>
pain	[pɛ̃]	n.	m.	s.	1 <sup>re</sup>
plein	[plɛ̃]	adj.	m.	s.	1 <sup>re</sup>
prince	[prɛ̃s]	n.	m.	s.	2 <sup>e</sup>
princesse	[prɛ̃ses]	n.	f.	s.	2 <sup>e</sup>
printemps	[prɛ̃tɑ̃]	n.	m.	s. et pl.	2 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 1<sup>re</sup> et 2<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
quinze	[kɛ̃z]	dét.	m.	pl.	2 <sup>e</sup>
quinze	[kɛ̃z]	dét.	f.	pl.	2 <sup>e</sup>
train	[trɛ̃]	n.	m.	s.	1 <sup>re</sup>
vingt	[vɛ̃]	dét.	m.	pl.	2 <sup>e</sup>
vingt	[vɛ̃]	dét.	f.	pl.	2 <sup>e</sup>

## ANNEXE H

**Mots qui contiennent la prononciation [ɛ̃] dans la liste orthographique à l'usage des enseignantes et des enseignants (ministère de l'Éducation et de l'Enseignement supérieur, 2014) au deuxième cycle du primaire**

Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :					
Année scolaire 3 <sup>e</sup> et 4 <sup>e</sup>					
Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
afin de	[afɛ̃dø]	prép.			4 <sup>e</sup>
ainsi	[ɛ̃si]	adv.			4 <sup>e</sup>
ancien	[ɑ̃sjɛ̃]	adj.	m.	s.	4 <sup>e</sup>
au moins	[omwɛ̃]	adv.			4 <sup>e</sup>
bain	[bɛ̃]	n.	m.	s.	3 <sup>e</sup>
besoin	[bøzwɛ̃]	n.	m.	s.	4 <sup>e</sup>
bien	[bjɛ̃]	adv.			3 <sup>e</sup>
bien	[bjɛ̃]	n.	m.	s.	4 <sup>e</sup>
bientôt	[bjɛ̃to]	adv.			3 <sup>e</sup>
certain	[sɛ̃rtɛ̃]	adj.	m.	s.	3 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 3<sup>e</sup> et 4<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
chagrin	[ʃagrɛ̃]	n.	m.	s.	3 <sup>e</sup>
coin	[kwɛ̃]	n.	m.	s.	3 <sup>e</sup>
combien	[kɔ̃bjɛ̃]	adv.			3 <sup>e</sup>
demain	[dəmɛ̃]	adv.			3 <sup>e</sup>
en train de	[ɑ̃trɛ̃də]	prép.			4 <sup>e</sup>
faim	[fɛ̃]	n.	f.	s.	3 <sup>e</sup>
fin	[fɛ̃]	n.	f.	s.	4 <sup>e</sup>
fin de semaine	[fɛ̃dsəmɛ̃n]	n.	f.	s.	4 <sup>e</sup>
gardien	[gardjɛ̃]	n.	m.	s.	4 <sup>e</sup>
important	[ɛ̃pɔ̃rtɑ̃]	adj.	m.	s.	3 <sup>e</sup>
importante	[ɛ̃pɔ̃rtɑ̃t]	adj.	f.	s.	3 <sup>e</sup>
impossible	[ɛ̃pɔ̃sibl]	adj.	m.	s.	4 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 3<sup>e</sup> et 4<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
impossible	[ɛ̃pɔsibl]	adj.	f.	s.	4 <sup>e</sup>
impression	[ɛ̃pʁɛsjɔ̃]	n.	f.	s.	4 <sup>e</sup>
incroyable	[ɛ̃krwajabl]	adj.	m.	s.	4 <sup>e</sup>
incroyable	[ɛ̃krwajabl]	adj.	f.	s.	4 <sup>e</sup>
insister	[ɛ̃siste]	v.			4 <sup>e</sup>
instant	[ɛ̃stɑ̃]	n.	m.	s.	4 <sup>e</sup>
intérieur	[ɛ̃tɛrjœʁ]	adj.	m.	s.	4 <sup>e</sup>
intérieur	[ɛ̃tɛrjœʁ]	n.	m.	s.	4 <sup>e</sup>
intérieure	[ɛ̃tɛrjœʁ]	adj.	f.	s.	4 <sup>e</sup>
intrépide	[ɛ̃tʁɛpid]	adj.	m.	s.	4 <sup>e</sup>
intrépide	[ɛ̃tʁɛpid]	adj.	f.	s.	4 <sup>e</sup>
inventer	[ɛ̃vɑ̃te]	v.			4 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 3<sup>e</sup> et 4<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
invitation	[ɛ̃vitasjɔ̃]	n.	f.	s.	3 <sup>e</sup>
inviter	[ɛ̃vite]	v.			4 <sup>e</sup>
juin	[ʒɥɛ̃], [ʒɥœ̃]	n.	m.	s.	3 <sup>e</sup>
le mien	[ləmjɛ̃]	pron.			4 <sup>e</sup>
lendemain	[lɑ̃dmɛ̃]	n.	m.	s.	3 <sup>e</sup>
le sien	[ləsjɛ̃]	pron.			4 <sup>e</sup>
les miens	[ləmjɛ̃]	pron.			4 <sup>e</sup>
les siens	[ləsjɛ̃]	pron.			4 <sup>e</sup>
les tiens	[lətjɛ̃]	pron.			4 <sup>e</sup>
le tien	[lətjɛ̃]	pron.			4 <sup>e</sup>
magasin	[magazɛ̃]	n.	m.	s.	3 <sup>e</sup>
médecin	[metsɛ̃]	n.	m.	s.	4 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 3<sup>e</sup> et 4<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
médecin	[metsɛ̃]	n.	f.	s.	4 <sup>e</sup>
moyen	[mwajɛ̃]	n.	m.	s.	3 <sup>e</sup>
parrain	[parɛ̃]	n.	m.	s.	4 <sup>e</sup>
patin	[patɛ̃]	n.	m.	s.	3 <sup>e</sup>
pin	[pɛ̃]	n.	m.	s.	3 <sup>e</sup>
plein de	[plɛ̃də]	dét.	m.	s.	4 <sup>e</sup>
plein de	[plɛ̃də]	dét.	f.	s.	4 <sup>e</sup>
point	[pwɛ̃]	n.	m.	s.	3 <sup>e</sup>
pointe	[pwɛ̃t]	n.	f.	s.	4 <sup>e</sup>
poussin	[pusɛ̃]	n.	m.	s.	3 <sup>e</sup>
prochain	[prɔʃɛ̃]	adj.	m.	s.	3 <sup>e</sup>
prochain	[prɔʃɛ̃]	n.	m.	s.	3 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 3<sup>e</sup> et 4<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
refrain	[ʁəfrɛ̃]	n.	m.	s.	4 <sup>e</sup>
rejoindre	[ʁəʒwɛ̃dʁ]	v.			3 <sup>e</sup>
rien	[ʁjɛ̃]	adv.			3 <sup>e</sup>
rien	[ʁjɛ̃]	pron.			3 <sup>e</sup>
sapin	[sapɛ̃]	n.	m.	s.	3 <sup>e</sup>
simple	[sɛ̃pl]	adj.	m.	s.	4 <sup>e</sup>
simple	[sɛ̃pl]	adj.	f.	s.	4 <sup>e</sup>
simplement	[sɛ̃pləmɑ̃]	adv.			4 <sup>e</sup>
singe	[sɛ̃ʒ]	n.	m.	s.	4 <sup>e</sup>
soin	[swɛ̃]	n.	m.	s.	3 <sup>e</sup>
soudain	[sudɛ̃]	adj.	m.	s.	4 <sup>e</sup>
soudain	[sudɛ̃]	adv.			4 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 3<sup>e</sup> et 4<sup>e</sup>

<b>Mot</b>	<b>Phonétique</b>	<b>Classe de mot</b>	<b>Genre</b>	<b>Nombre</b>	<b>Année scolaire</b>
terrain	[tɛʀɛ̃]	n.	m.	s.	3 <sup>e</sup>
vin	[vɛ̃]	n.	m.	s.	4 <sup>e</sup>
voisin	[v wazɛ̃]	n.	m.	s.	3 <sup>e</sup>

## ANNEXE I

**Mots qui contiennent la prononciation [ɛ̃] dans la liste orthographique à l'usage des enseignantes et des enseignants (ministère de l'Éducation et de l'Enseignement supérieur, 2014) au troisième cycle du primaire**

Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :					
Année scolaire 5 <sup>e</sup> et 6 <sup>e</sup>					
Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
afin que	[afɛ̃kə]	conj.			6 <sup>e</sup>
atteindre	[atɛ̃dʁ]	v.			5 <sup>e</sup>
bienvenu	[bjɛ̃vny]	adj.	m.	s.	5 <sup>e</sup>
bienvenue	[bjɛ̃vny]	n.	f.	s.	5 <sup>e</sup>
bienvenue	[bjɛ̃vny]	adj.	f.	s.	5 <sup>e</sup>
brin	[brɛ̃]	n.	m.	s.	6 <sup>e</sup>
ceinture	[sɛ̃tyʁ]	n.	f.	s.	5 <sup>e</sup>
clin d'oeil	[klɛ̃dœj]	n.	m.	s.	5 <sup>e</sup>
convaincre	[kɔ̃vɛ̃kr]	v.			6 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 5<sup>e</sup> et 6<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
copain	[kɔpɛ̃]	n.	m.	s.	5 <sup>e</sup>
craindre	[krɛ̃dr]	v.			5 <sup>e</sup>
crainte	[krɛ̃t]	n.	f.	s.	5 <sup>e</sup>
dessin	[desɛ̃]	n.	m.	s.	5 <sup>e</sup>
destin	[destɛ̃]	n.	m.	s.	5 <sup>e</sup>
écrivain	[ekrivɛ̃]	n.	m.	s.	5 <sup>e</sup>
épingle	[epɛ̃gl]	n.	f.	s.	6 <sup>e</sup>
examen	[ɛgzamɛ̃]	n.	m.	s.	5 <sup>e</sup>
festin	[festɛ̃]	n.	m.	s.	5 <sup>e</sup>
foin	[fwɛ̃]	n.	m.	s.	5 <sup>e</sup>
gamin	[gamɛ̃]	n.	m.	s.	6 <sup>e</sup>
humain	[ymɛ̃]	adj.	m.	s.	6 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 5<sup>e</sup> et 6<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
humain	[ymɛ̃]	n.	m.	s.	6 <sup>e</sup>
impatience	[ɛ̃pasjãs]	n.	f.	s.	5 <sup>e</sup>
importance	[ɛ̃pɔ̃rtãs]	n.	f.	s.	5 <sup>e</sup>
importation	[ɛ̃pɔ̃rtasjɔ̃]	n.	f.	s.	6 <sup>e</sup>
impôt	[ɛ̃po]	n.	m.	s.	6 <sup>e</sup>
incapable	[ɛ̃kapabl]	adj.	m.	s.	6 <sup>e</sup>
incapable	[ɛ̃kapabl]	adj.	f.	s.	6 <sup>e</sup>
incendie	[ɛ̃sãdi]	n.	m.	s.	5 <sup>e</sup>
incident	[ɛ̃sidã]	n.	m.	s.	5 <sup>e</sup>
inconvenient	[ɛ̃kɔ̃venjã]	n.	m.	s.	6 <sup>e</sup>
indépendance	[ɛ̃depãdãs]	n.	f.	s.	6 <sup>e</sup>
indication	[ɛ̃dikasjɔ̃]	n.	f.	s.	6 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 5<sup>e</sup> et 6<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
indice	[ɛ̃dis]	n.	m.	s.	5 <sup>e</sup>
indifférence	[ɛ̃diferãs]	n.	f.	s.	5 <sup>e</sup>
indignation	[ɛ̃diɲasjõ]	n.	f.	s.	5 <sup>e</sup>
individu	[ɛ̃dividy]	n.	m.	s.	5 <sup>e</sup>
industrie	[ɛ̃dystri]	n.	f.	s.	6 <sup>e</sup>
infirmier	[ɛ̃firmje]	n.	m.	s.	5 <sup>e</sup>
infirmière	[ɛ̃firmjɛR]	n.	f.	s.	5 <sup>e</sup>
infirmité	[ɛ̃firmite]	n.	f.	s.	6 <sup>e</sup>
information	[ɛ̃fɔRmasjõ]	n.	f.	s.	5 <sup>e</sup>
injure	[ɛ̃zyR]	n.	f.	s.	6 <sup>e</sup>
injustice	[ɛ̃zystis]	n.	f.	s.	5 <sup>e</sup>
inquiéter	[ɛ̃kjete]	v.			6 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 5<sup>e</sup> et 6<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
inquiétude	[ɛ̃kjetyd]	n.	f.	s.	5 <sup>e</sup>
insecte	[ɛ̃sɛkt]	n.	m.	s.	5 <sup>e</sup>
insolence	[ɛ̃solãs]	n.	f.	s.	5 <sup>e</sup>
inspecteur	[ɛ̃spektœʀ]	n.	m.	s.	6 <sup>e</sup>
inspection	[ɛ̃speksjõ]	n.	f.	s.	6 <sup>e</sup>
inspectrice	[ɛ̃spektris]	n.	f.	s.	6 <sup>e</sup>
inspiration	[ɛ̃spirasjõ]	n.	f.	s.	5 <sup>e</sup>
installation	[ɛ̃stalasjõ]	n.	f.	s.	6 <sup>e</sup>
installer	[ɛ̃stale]	v.			5 <sup>e</sup>
instinct	[ɛ̃stɛ̃]	n.	m.	s.	6 <sup>e</sup>
instruction	[ɛ̃stryksjõ]	n.	f.	s.	5 <sup>e</sup>
instrument	[ɛ̃strymã]	n.	m.	s.	5 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 5<sup>e</sup> et 6<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
intelligence	[ɛ̃tɛlizãs]	n.	f.	s.	5 <sup>e</sup>
intelligent	[ɛ̃tɛlizã]	adj.	m.	s.	5 <sup>e</sup>
intelligente	[ɛ̃tɛlizât]	adj.	f.	s.	5 <sup>e</sup>
intense	[ɛ̃tãs]	adj.	m.	s.	6 <sup>e</sup>
intense	[ɛ̃tãs]	adj.	f.	s.	6 <sup>e</sup>
intention	[ɛ̃tãsjõ]	n.	f.	s.	5 <sup>e</sup>
intéressant	[ɛ̃tɛrɛsã]	adj.	m.	s.	6 <sup>e</sup>
intéressante	[ɛ̃tɛrɛsât]	adj.	f.	s.	6 <sup>e</sup>
intéresser	[ɛ̃tɛrɛsɛ]	v.			5 <sup>e</sup>
intérêt	[ɛ̃tɛrɛ]	n.	m.	s.	5 <sup>e</sup>
intervalle	[ɛ̃tɛrval]	n.	m.	s.	6 <sup>e</sup>
intervention	[ɛ̃tɛrvãsjõ]	n.	f.	s.	5 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 5<sup>e</sup> et 6<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
inventeur	[ɛ̃vɑ̃tœʁ]	n.	m.	s.	6 <sup>e</sup>
inventrice	[ɛ̃vɑ̃tris]	n.	f.	s.	6 <sup>e</sup>
invité	[ɛ̃vite]	n.	m.	s.	5 <sup>e</sup>
invitée	[ɛ̃vite]	n.	f.	s.	5 <sup>e</sup>
lien	[ljɛ̃]	n.	m.	s.	5 <sup>e</sup>
linge	[lɛ̃ʒ]	n.	m.	s.	5 <sup>e</sup>
lointain	[lwɛ̃tɛ̃]	adj.	m.	s.	5 <sup>e</sup>
lointaine	[lwɛ̃tɛ̃n]	adj.	f.	s.	5 <sup>e</sup>
magicien	[maziʒjɛ̃]	n.	m.	s.	5 <sup>e</sup>
mince	[mɛ̃s]	adj.	m.	s.	6 <sup>e</sup>
mince	[mɛ̃s]	adj.	f.	s.	6 <sup>e</sup>
musicien	[myzizjɛ̃]	n.	m.	s.	6 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 5<sup>e</sup> et 6<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
orphelin	[ɔʁfɛlɛ̃]	n.	m.	s.	6 <sup>e</sup>
peindre	[pɛ̃dʁ]	v.			5 <sup>e</sup>
peintre	[pɛ̃tʁ]	n.	m.	s.	5 <sup>e</sup>
peintre	[pɛ̃tʁ]	n.	f.	s.	5 <sup>e</sup>
peinture	[pɛ̃tyʁ]	n.	f.	s.	5 <sup>e</sup>
pinceau	[pɛ̃so]	n.	m.	s.	5 <sup>e</sup>
poing	[pwɛ̃]	n.	m.	s.	5 <sup>e</sup>
principal	[pʁɛ̃sipal]	adj.	m.	s.	6 <sup>e</sup>
principale	[pʁɛ̃sipal]	adj.	f.	s.	6 <sup>e</sup>
principales	[pʁɛ̃sipal]	adj.	f.	pl.	6 <sup>e</sup>
principaux	[pʁɛ̃sipo]	adj.	m.	pl.	6 <sup>e</sup>
saint	[sɛ̃]	adj.	m.	s.	6 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 5<sup>e</sup> et 6<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
saint	[sɛ̃]	n.	m.	s.	6 <sup>e</sup>
sainte	[sɛ̃t]	adj.	f.	s.	6 <sup>e</sup>
sainte	[sɛ̃t]	n.	f.	s.	6 <sup>e</sup>
sein	[sɛ̃]	n.	m.	s.	5 <sup>e</sup>
sympathie	[sɛ̃pati]	n.	f.	s.	6 <sup>e</sup>
sympathique	[sɛ̃patik]	adj.	m.	s.	6 <sup>e</sup>
sympathique	[sɛ̃patik]	adj.	f.	s.	6 <sup>e</sup>
syndicat	[sɛ̃dika]	n.	m.	s.	6 <sup>e</sup>
technicien	[tɛ̃knisjɛ̃]	n.	m.	s.	6 <sup>e</sup>
témoin	[temwɛ̃]	n.	m.	s.	5 <sup>e</sup>
timbre	[tɛ̃br]	n.	m.	s.	5 <sup>e</sup>
vain	[vɛ̃]	adj.	m.	s.	6 <sup>e</sup>

**Les mots qui contiennent la prononciation [ɛ̃] et qui satisfont aux conditions suivantes :**

Année scolaire 5<sup>e</sup> et 6<sup>e</sup>

Mot	Phonétique	Classe de mot	Genre	Nombre	Année scolaire
vainqueur	[vɛ̃kœʀ]	n.	m.	s.	5 <sup>e</sup>

## ANNEXE J

### Mots qui contiennent la prononciation [ɛ̃] dans l'échelle québécoise d'acquisition de l'orthographe lexicale (ÉQOL) (Stanké et al., 2019)

ÉQOL taux de réussite IN							
Mot	Taux de réussite total	Réussite 1 <sup>re</sup> année	Réussite 2 <sup>e</sup> année	Réussite 3 <sup>e</sup> année	Réussite 4 <sup>e</sup> année	Réussite 5 <sup>e</sup> année	Réussite 6 <sup>e</sup> année
afin	82%	69%	75%	79%	81%	96%	90%
ainsi	40%	12%	17%	11%	53%	67%	68%
amérindiens	59%	8%	39%	56%	73%	83%	83%
appartient	26%	0%	2%	35%	21%	41%	52%
atteindre	36%	0%	3%	14%	38%	67%	83%
besoin	69%	15%	43%	70%	90%	95%	98%
bien	88%	42%	86%	94%	98%	97%	98%
bientôt	61%	6%	30%	55%	86%	96%	96%
cartésien	45%	8%	23%	51%	60%	59%	83%
certains	36%	0%	25%	36%	48%	35%	58%
chemin	80%	66%	69%	63%	88%	95%	96%

ÉQOL taux de réussite IN

Mot	Taux de réussite total	Réussite 1 <sup>re</sup> année	Réussite 2 <sup>e</sup> année	Réussite 3 <sup>e</sup> année	Réussite 4 <sup>e</sup> année	Réussite 5 <sup>e</sup> année	Réussite 6 <sup>e</sup> année
chien	94%	76%	96%	97%	98%	100%	98%
cinq	78%	41%	63%	80%	87%	95%	90%
cinquante	48%	7%	42%	38%	52%	78%	78%
coin	72%	20%	55%	94%	89%	92%	92%
combien	78%	2%	63%	88%	89%	94%	99%
contient	53%	6%	28%	44%	61%	80%	79%
demain	73%	23%	67%	78%	92%	97%	98%
dessin	73%	39%	63%	61%	87%	94%	92%
devient	38%	8%	21%	25%	39%	71%	73%
enfin	76%	40%	72%	59%	89%	93%	90%
européens	14%	1%	2%	8%	6%	39%	40%
féminin	81%	64%	78%	78%	86%	83%	97%
fin	96%	83%	96%	97%	99%	99%	100%

ÉQOL taux de réussite IN

Mot	Taux de réussite total	Réussite 1 <sup>re</sup> année	Réussite 2 <sup>e</sup> année	Réussite 3 <sup>e</sup> année	Réussite 4 <sup>e</sup> année	Réussite 5 <sup>e</sup> année	Réussite 6 <sup>e</sup> année
humain	70%	11%	45%	82%	95%	96%	100%
imparfait	56%	0%	12%	75%	81%	85%	100%
important	51%	1%	16%	48%	74%	92%	93%
impossible	52%	5%	21%	50%	75%	84%	92%
incas	70%	63%	54%	84%	75%	83%	55%
indicatif	69%	28%	45%	76%	64%	96%	96%
indices	57%	14%	31%	65%	71%	91%	88%
indique	78%	31%	64%	85%	94%	92%	96%
industrie	49%	13%	23%	45%	71%	83%	88%
inférieur	54%	22%	36%	44%	65%	81%	85%
infinitif	68%	40%	38%	69%	89%	94%	98%
informations	71%	23%	54%	92%	84%	92%	97%
inscrits	52%	8%	23%	47%	73%	86%	94%

ÉQOL taux de réussite IN

Mot	Taux de réussite total	Réussite 1 <sup>re</sup> année	Réussite 2 <sup>e</sup> année	Réussite 3 <sup>e</sup> année	Réussite 4 <sup>e</sup> année	Réussite 5 <sup>e</sup> année	Réussite 6 <sup>e</sup> année
insectes	59%	16%	49%	63%	71%	80%	87%
intendant	50%	1%	20%	50%	84%	80%	79%
intérieur	43%	24%	23%	39%	51%	67%	72%
intertitres	62%	32%	38%	78%	69%	81%	82%
intrus	24%	9%	16%	17%	26%	39%	48%
inventions	64%	6%	22%	84%	73%	88%	94%
lien	78%	36%	70%	75%	93%	97%	100%
loin	77%	18%	71%	73%	86%	96%	94%
lutin	96%	92%	99%	97%	95%	95%	100%
main	88%	55%	82%	98%	98%	90%	100%
maintenant	49%	4%	19%	33%	62%	83%	86%
marins	92%	89%	84%	95%	94%	98%	87%
masculin	84%	44%	78%	87%	91%	94%	98%

ÉQOL taux de réussite IN

Mot	Taux de réussite total	Réussite 1 <sup>re</sup> année	Réussite 2 <sup>e</sup> année	Réussite 3 <sup>e</sup> année	Réussite 4 <sup>e</sup> année	Réussite 5 <sup>e</sup> année	Réussite 6 <sup>e</sup> année
matin	98%	90%	98%	96%	100%	100%	100%
médecin	46%	9%	22%	44%	77%	59%	65%
moins	37%	0%	41%	29%	53%	38%	45%
moyen	62%	0%	28%	67%	82%	95%	93%
obtient	34%	0%	6%	35%	25%	60%	70%
pain	87%	51%	80%	89%	100%	97%	97%
peinture	65%	0%	44%	53%	81%	91%	91%
points	63%	6%	28%	56%	76%	98%	95%
prince	77%	30%	61%	91%	98%	100%	94%
princesse	69%	19%	40%	79%	97%	94%	92%
principal	70%	25%	50%	73%	92%	97%	95%
principaux	47%	0%	2%	60%	53%	73%	98%
printemps	48%	17%	24%	58%	61%	56%	68%

ÉQOL taux de réussite IN

Mot	Taux de réussite total	Réussite 1 <sup>re</sup> année	Réussite 2 <sup>e</sup> année	Réussite 3 <sup>e</sup> année	Réussite 4 <sup>e</sup> année	Réussite 5 <sup>e</sup> année	Réussite 6 <sup>e</sup> année
province	64%	10%	40%	47%	81%	92%	99%
rien	85%	34%	82%	96%	100%	100%	100%
sapin	97%	86%	99%	100%	100%	100%	98%
simple	66%	14%	29%	86%	97%	100%	100%
singulier	78%	27%	67%	85%	94%	94%	96%
soudain	63%	9%	31%	65%	75%	92%	94%
symbole	36%	0%	8%	22%	52%	68%	66%
terrain	59%	0%	13%	69%	83%	84%	94%
train	84%	35%	77%	84%	93%	98%	95%
vient	32%	8%	20%	24%	33%	59%	59%
vingt	54%	6%	23%	50%	79%	82%	78%

## RÉFÉRENCES

- Adoniou, M. (2013). What should teachers know about spelling? *Literacy*, 48(3), 144-154.  
<https://doi.org/10.1111/lit.12017>
- Bates, D., Mächler, M., Bolker, B. et Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67, 1-48.  
<https://doi.org/10.18637/jss.v067.i01>
- 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, 275-297.
- Benjamini, Y. et Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289-300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Benveniste, É. (2006). *Problèmes de linguistique générale. 1*. Gallimard.
- Berger, M.-J. et Desrochers, A. (Éds.). (2011). *L'évaluation de la littératie*. Presses de l'Université d'Ottawa.
- Berthiaume, R. et Daigle, D. (2014). Are dyslexic children sensitive to the morphological structure of words when they read? The case of dyslexic readers of French. *Dyslexia (Chichester, England)*, 20(3), 241-260. <https://doi.org/10.1002/dys.1476>
- Boivin, M.-C. et Pinsonneault, R. (2018). Les erreurs de syntaxe, d'orthographe grammaticale et d'orthographe lexicale des élèves québécois en contexte de production écrite. *Canadian Journal of Applied Linguistics*, 21(1), 43-70. <https://doi.org/10.7202/1050810ar>
- Bonin, P., Collay, S. et Fayol, M. (2008). La consistance orthographique en production verbale écrite : Une brève synthèse. *L'Année psychologique*, 108(03), 517-546.  
<https://doi.org/10.4074/S0003503308003059>

- Borchardt, G. (2012). *L'influence des connaissances graphotactiques sur l'acquisition de l'orthographe lexicale : Étude chez l'enfant d'école élémentaire et chez l'adulte* [Thèse]. Université René Descartes.
- Borchardt, G., Fayol, M. et Pacton, S. (2012). L'influence de la sensibilité aux régularités graphotactiques sur l'apprentissage de l'orthographe de nouveaux mots. *ANAE - Approche Neuropsychologique des Apprentissages chez l'Enfant*, 24(116), 67-73.
- Bosse, M.-L., Valdois, S. et Dompnier, B. (2009). Acquisition du langage écrit et empan visuo-attentionnel : Une étude longitudinale. Dans N. Marc-Breton, A. S. Besse, F. de L. Haye N. B. et E. Bonjour (Éds.), *Approche Cognitive de l'apprentissage de la langue écrite* (p. 132-141). PUR. <https://hal.science/hal-00817793>
- Bousquet, S., Cogis, D., Ducard, D., Massonnet, J. et Jaffre, J.-P. (1999). Acquisition de l'orthographe et modes cognitifs. *Revue française de pédagogie*, 126(1), 23-37. <https://doi.org/10.3406/rfp.1999.1092>
- Brissaud, C. et Cogis, D. (2011a). *Comment enseigner l'orthographe aujourd'hui ?* (Hatier).
- Brissaud, C. et Cogis, D. (2011b). Les fondements de l'enseignement de l'orthographe. Dans C. Brissaud, D. Cogis et J.-P. Jaffré, *Comment enseigner l'orthographe aujourd'hui ?* (Hatier, p. 9-32).
- Caravolas, M. (2019, mars 29). *Prédire le développement des habiletés en lecture et en écriture dans les langues européennes : Effets de la consistance orthographique* [Conférence]. 44e congrès de l'Institut des troubles d'apprentissage, Montréal.
- Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G. et Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science*, 24(8), 1398-1407. <https://doi.org/10.1177/0956797612473122>

- Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavsky, M., Onochie-Quintanilla, E., Salas, N., Schöffelová, M., Defior, S., Mikulajová, M., Seidlová-Málková, G. et Hulme, C. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science*, 23(6), 678-686.  
<https://doi.org/10.1177/0956797611434536>
- Casalis, S., Deacon, S. H. et Pacton, S. (2011). How specific is the connection between morphological awareness and spelling? A study of French children. *Applied Psycholinguistics*, 32(3), 499-511. <https://doi.org/10.1017/S014271641100018X>
- Catach, N. (2011). *L'orthographe* (10 ed). Presses Univ. de France.
- Catach, N. (2016). *L'orthographe française : Traité théorique et pratique avec des travaux d'application et leurs corrigés* (Nouvelle présentation). Armand Colin.
- Chandler, K. (2000). What I Wish I'd Known about Teaching Spelling. *The English Journal*, 89(6), 87. <https://doi.org/10.2307/821268>
- Chapleau, N. (2013). *Effet d'un programme d'intervention orthopédagogique sur la conscience morphologique et la production de mots écrits chez des élèves présentant une difficulté spécifique d'apprentissage de la lecture-écriture*. Université du Québec à Montréal.
- Chapleau, N., Laplante, L. et Brodeur, M. (2013). Rééducation orthopédagogique de l'orthographe lexicale. Portrait d'une élève présentant une dyslexie-dysorthographe. Dans D. Daigle, I. Montésinos-Gelet et A. Plisson, *Orthographe et populations exceptionnelles : Perspectives didactiques* (p. 155-175). Presses de l'Université du Québec.
- Chartrand, S.-G., Cogis, D. et Elalouf, M.-L. (2016). Progression dans l'enseignement de la grammaire. Dans S.-G. Chartrand, *Mieux enseigner la grammaire : Pistes didactiques et activités pour la classe* (p. 81-98). ERPI Pearson.

- Chomsky, N. (2006). *Réflexions sur le langage* (J. Milner, B. Vautherin et P. Fiala, Trad.). Flammarion.
- Cogis, D. (2005a). *Enseigner et apprendre l'orthographe : Nouveaux enjeux, pratiques nouvelles, école-collège*. Delagrave.
- Commission de la culture et de l'éducation. (2015). *Étude des crédits budgétaires 2014-2015—Assemblée nationale du Québec*. Assemblée nationale du Québec.  
<http://www.assnat.qc.ca/fr/travaux-parlementaires/commissions/cce/mandats/Mandat-25329/index.html>
- Daigle, D., Ammar, A., Berthiaume, R., Montésinos-Gelet, I., Ouellet, C. et Prévost, N. (2015). *L'enseignement de l'orthographe lexicale et l'élève en difficulté : Développement et mise à l'essai d'un programme d'entraînement* (p. 174) [Rapport de recherche]. Université de Montréal.
- Daigle, D. et Berthiaume, R. (2021). *L'apprentissage de la lecture et de l'écriture : Décomposer les objets d'enseignement en microtâches pour les rendre accessibles à tous les élèves*. Chenelière.
- Daigle, D., Berthiaume, R., Ruberto, N. et Wolter, J.-A. (2018). Classroom practices in morphological instruction. Dans R. Berthiaume, D. Daigle et A. Desrochers, *Morphological processing and literacy development: Current issues and research* (p. 244-268). Routledge.
- Daigle, D., Costerg, A., Plisson, A., Ruberto, N. et Varin, J. (2016). Spelling Errors in French-speaking Children with Dyslexia: Phonology May Not Provide the Best Evidence. *Dyslexia*, 22(2), 137-157.
- Daigle, D., Montésinos-Gelet, I. et Plisson, A. (2013). *Orthographe et populations exceptionnelles : Perspectives didactiques*. Presses de l'Université du Québec.

- Debout, C. (2012). Méthodologie quantitative et études descriptives simples. *Soins*, 768, 55-60.
- Demont, É. et Gombert, J.-É. (2004). L'apprentissage de la lecture : Évolution des procédures et apprentissage implicite. *Enfance*, 56(3), 245-257.
- Deschamps, L. (2015). Les Français et l'orthographe, ça fait deux. *Huffington Post*.  
[https://www.huffingtonpost.fr/2015/06/11/orthographe-francais-maitrise-langue\\_n\\_7559054.html](https://www.huffingtonpost.fr/2015/06/11/orthographe-francais-maitrise-langue_n_7559054.html)
- Desrochers, A., Kirby, J., Thompson, G. et Fréchette, S. (2009). Le rôle de la conscience phonologique dans l'apprentissage de la lecture. *Revue du Nouvel-Ontario*, 34, 59-82.  
<https://doi.org/10.7202/038720ar>
- Dumais, C., Stanké, B., Moreau, A. et Beaudoin, M. (2014). L'enseignement de l'orthographe lexicale : Réflexion sur les bases de données lexicales. *Les cahiers de l'Association québécoise des professeurs de français (AQPf)*, 4(3), 15-18.
- Écalle, J. et Magnan, A. (2015). *L'apprentissage de la lecture et ses difficultés* (2e éd.). Dunod.
- 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. (1998). *Grapheme-phoneme knowledge is essential for learning to read words in English* (p. x, 403). Lawrence Erlbaum Associates Publishers.
- El Sanharawi, M. et Naudet, F. (2013). Comprendre la régression logistique. *Journal Français d'ophtalmologie*, 36, 710-715. <https://doi.org/10.1016/j.jfo.2013.05.008>
- Fayol, M. (2003). Les difficultés de l'orthographe. *Cerveau et Psycho*, 3, 2-5.
- Fayol, M. (2011). Processus cognitifs et orthographe. Dans C. Brissaud et D. Cogis, *Comment enseigner l'orthographe aujourd'hui ?* (Hatier, p. 86-95).
- Fayol, M. (2017). *L'acquisition de l'écrit* (2e éd. mise à jour). Que-sais-je ?

- Fayol, M. et Jaffré, J.-P. (1999). Note de synthèse : L'acquisition/apprentissage de l'orthographe. *Revue française de pédagogie*, 126(1), 143-170. <https://doi.org/10.3406/rfp.1999.1101>
- Fayol, M. et Jaffré, J.-P. (2008). *Orthographier*. Presses universitaires de France.
- Fayol, M. et Jaffré, J.-P. (2014). *L'orthographe*. Presses universitaires de France.
- Fayol, M. et Jaffré, J.-P. (2016). L'orthographe : Des systèmes aux usages. *Pratiques [En ligne]*, 169-170, Article 169-170. <https://doi.org/10.4000/pratiques.2984>
- Fontaine, M. (2019). *L'effet et l'implantation d'un programme d'intervention orthopédagogique intégrant des aides technologiques sur la production de mots écrits en français d'élèves dysorthographiques âgés de 9-12 ans* [Thèse ou essai doctoral accepté, Université du Québec en Outaouais]. <https://archipel.uqam.ca/13486/>
- Fortin, M.-F. et Gagnon, J. (2022). *Fondements et étapes du processus de recherche : Méthodes quantitatives et qualitatives* (4e édition). Chenelière éducation.
- Fox, J. et Weisberg, S. (2019). *An R companion to applied regression* (Third edition). SAGE.
- Frost, R. et Katz, L. (1992). *Orthography, phonology, morphology, and meaning*. North-Holland. <http://site.ebrary.com/id/10258519>
- Gala, N. et Zock, M. (2013). Ressources lexicales. Contenu, construction, utilisation, évaluation. *Carnets de lecture*, 24. [http://farum.it/lectures/ezone\\_articles.php?art\\_id=344](http://farum.it/lectures/ezone_articles.php?art_id=344)
- Gay, L. R., Mills, G. E. et Airasian, P. W. (2012). *Educational research: Competencies for analysis and applications* (10th ed). Pearson.
- Gombert, J.-E. (2003). Implicit and Explicit Learning to Read: Implication as for Subtypes of Dyslexia. *Current Psychology Letters. Behaviour, Brain & Cognition*, 1(10), Article 10, Vol. 1, 2003. <https://doi.org/10.4000/cpl.202>

- Gombert, J.-É., Bryant, P. et Warrick, N. (1997). Children's Use of Analogy in Learning to Read and Spell. Dans C. A. Perfetti, L. Rieben et M. Fayol, *Learning to Spell: Research, Theory, and Practice Across Languages* (p. 221-235). Routledge.
- Handy, M.-J. (2009). *La différenciation de l'enseignement de l'orthographe lexicale par les entrées sensorielles les gestes mentaux et la métacognition stratégique, auprès des élèves de 3e année du primaire* [Université de Montréal].  
<https://papyrus.bib.umontreal.ca/xmlui/handle/1866/6486>
- Hawken, J. (2009). *Pour un enseignement efficace de la lecture et de l'écriture* (L. Laplante, M. Brodeur, A. Desrochers et G. Jean, Trad.; p. 148). Réseau canadien de recherche sur le langage et l'alphabétisation.
- Hazard, M.-C. (2009). *Consistance orthographique et construction du lexique chez l'enfant d'âge scolaire* [Thèse]. Université de Nice.
- Hébert, M. et Lépine, M. (2013). De l'intérêt de la notion de littératie en francophonie : Un état des lieux en sciences de l'éducation. *Globe*, 16(1), 25-43.
- IBM Corp. (2020). *IBM SPSS Statistics for Windows* (Version 27) [Software]. IBM Corp.  
<https://www.ibm.com/fr-fr/products/spss-statistics>
- Jacquier-Roux, M., Lequette, C., Pouget, G., Valdois, S. et Zorman, M. (2010). *BALE Batterie analytique du langage écrit*. Laboratoire des sciences de l'éducation.
- Jolicoeur, É. (2007). *Rééducation de patrons orthographiques chez trois élèves ayant un trouble de l'orthographe* [Mémoire]. Université du Québec à Montréal.
- Kandel, S. et Spinelli, E. (2010). Processing complex graphemes in handwriting production. *Memory & Cognition*, 38(6), 762-770.

- Karmiloff-Smith, A. (1994). Beyond Modularity : A Developmental Perspective on Cognitive Science. *International Journal of Language & Communication Disorders*, 29(1), 95-105.  
<https://doi.org/10.3109/13682829409041485>
- Karsenti, T. et Savoie-Zajc, L. (2018). *La recherche en éducation : Étapes et approches (4e éd.)* (4e édition). Presses de l'Université de Montréal.
- Labrecque, A.-M. (2011). *Les connaissances et les pratiques des enseignants du 1er cycle du primaire en lien avec l'enseignement et l'apprentissage de la calligraphie*. [Mémoire]. Université de Sherbrooke.
- Lambert, É. et Chesnet, D. (2001). NOVLEX : Une base de données lexicales pour les élèves de primaire. *L'Année psychologique*, 101(2), 277-288.  
<https://doi.org/10.3406/psy.2001.29557>
- Lehrer, J., Bigras, N., Charron, A. et Laurin, I. (2021). *La recherche en éducation à la petite enfance : Origines, méthodes et applications*. PUQ.
- Lenth, R. V., Buerkner, P., Herve, M., Love, J., Miguez, F., Riebl, H. et Singmann, H. (2022). *emmeans : Estimated Marginal Means, aka Least-Squares Means (1.7.5)* [Software].  
<https://CRAN.R-project.org/package=emmeans>
- Lété, B., Peereman, 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. <https://doi.org/10.1016/j.jml.2008.01.001>
- 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*, 36(1), 156-166. <https://doi.org/10.3758/BF03195560>

- Lévesque, J.-Y., Gaté, J.-P., Mansour, M. et Saint-Pierre, A. (2014). Étude de l'orthographe lexicale à domicile et soutien parental. *Nouveaux cahiers de la recherche en éducation*, 16(1), 77-103. <https://doi.org/10.7202/1025764ar>
- Lussier, F., Chevrier, E. et Gascon, L. (2017). Troubles spécifiques des apprentissages. Dans *Neuropsychologie de l'enfant et de l'adolescent : Troubles développementaux et de l'apprentissage* (3<sup>e</sup> éd.). Dunod.  
<http://banq.prenumerique.ca/accueil/isbn/9782100764938>
- Manesse, D. et Cogis, D. (2007). *Orthographe, à qui la faute* (Éditions sociales françaises). ESF.
- Mansour, M. (2012). *Pratiques relatives à l'enseignement de l'orthographe lexicale au premier cycle du primaire* [Masters, Université du Québec à Rimouski].  
<http://semaphore.uqar.ca/id/eprint/963/>
- Martin-Chang, S., Ouellette, G. et Bond, L. (2016). Differential Effects of Context and Feedback on Orthographic Learning: How Good Is Good Enough? *Scientific Studies of Reading*, 21, 1-14. <https://doi.org/10.1080/10888438.2016.1263993>
- Martinet, C., Bosse, M.-L., Valdois, S. et Tainturier, M.-J. (1999). Existe-t-il des stades successifs dans l'acquisition de l'orthographe d'usage ? *Langue française*, 124(1), 58-73.  
<https://doi.org/10.3406/lfr.1999.6306>
- Martinet, C., Valdois, S. et Fayol, M. (2004). Lexical orthographic knowledge develops from the beginning of literacy acquisition. *Cognition*, 91(2), B11-B22.  
<https://doi.org/10.1016/j.cognition.2003.09.002>
- Mauroux, F., Boudreau, C. et Morin, M.-F. (2018). DIDALEX : Élaboration d'un outil francophone d'analyse quantitative et qualitative des performances d'élèves du primaire en orthographe lexicale. *SHS Web of Conferences*, 46, 11002.  
<https://doi.org/10.1051/shsconf/20184611002>

- MELS. (2008). *Mieux soutenir le développement de la compétence à écrire : Rapport du Comité d'experts sur l'apprentissage de l'écriture*. Ministère de l'éducation, du loisir et du sport.
- Ministère de l'Éducation du Loisir et du Sport. (2011). *Programme de formation de l'école québécoise : Progression des apprentissages au primaire, français, langue d'enseignement*. Gouvernement du Québec.
- [http://www.education.gouv.qc.ca/fileadmin/site\\_web/documents/education/jeunes/pfeq/PDA\\_PFEQ\\_francais-langue-enseignement-primaire\\_2011.pdf](http://www.education.gouv.qc.ca/fileadmin/site_web/documents/education/jeunes/pfeq/PDA_PFEQ_francais-langue-enseignement-primaire_2011.pdf)
- Ministère de l'Éducation, du Loisir et du Sport. (2012). *Évaluation du plan d'action pour l'amélioration du français suivi des apprentissages réalisés par les élèves en écriture (2009, 2010) : Deuxième rapport d'étape*. Ministère de l'éducation, du loisir et du sport.
- <http://collections.banq.qc.ca/ark:/52327/2101351>
- Ministère de l'Éducation et de l'Enseignement supérieur. (2014). *Liste orthographique à l'usage des enseignantes et des enseignants*. education.gouv.qc.ca.
- <http://www.education.gouv.qc.ca/references/tx-solrtyperecherchepublicationtx-solrpublicationnouveaute/resultats-de-la-recherche/detail/>
- Ministère de l'Éducation et de l'Enseignement supérieur. (2017). *Politique de la réussite éducative : Le plaisir d'apprendre, la chance de réussir*. éducation.gouv.qc.ca.
- [http://www.education.gouv.qc.ca/fileadmin/site\\_web/documents/PSG/politiques\\_orientations/politique\\_reussite\\_educative\\_10juillet\\_F\\_1.pdf](http://www.education.gouv.qc.ca/fileadmin/site_web/documents/PSG/politiques_orientations/politique_reussite_educative_10juillet_F_1.pdf)
- Montésinos-Gelet, I., Dupin de Saint-André, M. et Charron, A. (2022). *La lecture et l'écriture. Tome 1, fondements et pratiques à l'éducation préscolaire et au 1er cycle du primaire*. Chenelière éducation.
- Moreau, A. C., Herry, Y. et Maltais, C. (2005). *L'éducation inclusive au préscolaire : Accueillir l'enfant ayant des besoins particuliers*. Éditions CEC.

- Moreau, A., Jolicoeur, É., Tremblay, K., Beaulieu, J., Fontaine, M., Brassard, I. et Rhofir, S. (2021). Métasynthèse des recensions antérieures sur les pratiques d'enseignement en littératie d'apprenants ayant une déficience intellectuelle. *Revue internationale de communication et socialisation*, 8(1), 1-15.
- Morin, M.-F. (2009). *Synthèse de connaissances sur l'enseignement de l'écriture à l'école primaire* (p. 82) [Recension]. Université de Sherbrooke. [http://lectureecriture.ca/wp-content/uploads/2014/08/Synth%C3%A8se\\_finale\\_mention-MELS\\_15-mars-2010.pdf](http://lectureecriture.ca/wp-content/uploads/2014/08/Synth%C3%A8se_finale_mention-MELS_15-mars-2010.pdf)
- Morin, M.-F. et Montésinos-Gelet, I. (2005). Les habiletés phonogrammiques en écriture à la maternelle : Comparaison de deux contextes francophones différents France-Québec. *Canadian Journal of Education / Revue canadienne de l'éducation*, 28(3), 508-533. JSTOR. <https://doi.org/10.2307/4126481>
- Mousty, P. et Leybaert, J. (1999). Évaluation des habiletés de lecture et d'orthographe au moyen de BELEC : Données longitudinales auprès d'enfants francophones testés en 2e et 4e années. *Revue européenne de psychologie appliquée*, 49(4), 325-342.
- Mousty, P., Leybaert, J., Alegria, J., Content, A. et Morais, J. (1994). BELEC : une batterie d'évaluation du langage écrit et de ses troubles : Analyse conjointe des habiletés linguistiques et métalinguistiques. Dans J. Grégoire et B. Piérart, *Évaluer les troubles de la lecture. Les nouveaux modèles théoriques et leurs implications diagnostiques* (p. 127-146). De Boeck Supérieur; Cairn.info. <https://www.cairn.info/evaluer-les-troubles-de-la-lecture--9782804119164.htm>
- Neault, I. et Guay, M.-C. (2006). Ces enfants qui souffrent en silence de la dyslexie : Un diagnostic méconnu, souvent tardif et confondu avec le TDAH. *Revue de psychoéducation*, 36(1), 129-147.
- OCDE. (2014). *Regard sur l'éducation 2014* (Les indicateurs de l'OCDE). OCDE.

- OCDE et Statistique Canada. (2012). *La littératie, un atout pour la vie : Nouveaux résultats de l'Enquête sur la littératie et les compétences des adultes*. Statistique Canada, Culture, tourisme et Centre de la statistique de l'éducation.
- Office québécois de la langue française. (s. d.). *Alphabet phonétique international : Liste des symboles*. Consulté 4 juillet 2023, à l'adresse <https://vitrinelinguistique.oqlf.gouv.qc.ca/22137/la-prononciation/notions-de-base-en-phonetique/les-symboles-de-lalphabet-phonetique-international>
- Pacton, S. (2008). L'apprentissage de l'orthographe lexicale du français. Dans A. Desrochers, F. Martineau et Y.-C. Morin, *Orthographe française : Évolution et pratique* (David, p. 327-350).
- Pacton, S. (2018). Apprendre à orthographier les mots. Dans L. Ferrand, *Psychologie cognitive des apprentissages scolaires* (p. 177-187). Dunod; Cairn.info. <https://www.cairn.info/psychologie-cognitive-des-apprentissages-scolaires--9782100775569-p-177.htm>
- Pacton, S. (2019). *Entrevue personnelle réalisée par Mélanie Dutemple : Utilisation des doublets en recherche* [Skype].
- Pacton, S. et Afonso-Jaco, A. (2015). Comment les enfants apprennent-ils l'orthographe des mots ? *Revue française de linguistique appliquée*, XX(2), 51-61. <https://doi.org/10.3917/rfla.202.0051>
- Pacton, S., Fayol, M. et Perruchet, P. (2002). The acquisition of untaught orthographic regularities in French. Dans L. Verhoeven, E. Carsten et P. Reitsma, *Precursors of Functional Literacy* (p. 121-137). John Benjamins Publishing Company. <https://benjamins.com/catalog/swll.11>

- Pacton, S., Fayol, M. et Perruchet, P. (2005). Children's Implicit Learning of Graphotactic and Morphological Regularities. *Child Development*, 76(2), 324-339.  
<https://doi.org/10.1111/j.1467-8624.2005.00848.a.x>
- Pacton, S., Foulin, J.-N. et Fayol, M. (2005). L'apprentissage de l'orthographe lexicale. *Rééducation orthophonique*, 222, 47-68.
- 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. <https://doi.org/10.1037/0096-3445.130.3.401>
- Paret, M.-C. (2010). Le système de l'orthographe française. *Québec français*, 158, 78-80.
- Peereman, R. et Stanké, B. (2022). *ÉQOL-Infra : Caractéristiques distributionnelles des correspondances phono-graphologiques et des unités lexicales dans le lexique de l'école primaire québécoise*. <https://inframorph.github.io/eqolinfra/>
- Pérez, M. (2014a). Proposition de hiérarchisation des 45 graphèmes de base de l'orthographe du français. *SHS Web of Conferences*, 8, 1125-1140.  
<https://doi.org/10.1051/shsconf/20140801178>
- Pérez, M. (2014b). *Quelles variables utiliser pour définir la complexité orthographique des mots?* Colloque international des étudiants chercheurs en didactique des langues et en linguistique, Grenoble.
- Perruchet, P. et Nicolas, S. (1998). L'apprentissage implicite : Un débat théorique. *Psychologie française*, 1(43), 13-25.
- Perruchet, P. et Pacton, S. (2004). Qu'apportent à la pédagogie les travaux de laboratoire sur l'apprentissage implicite? *L'année psychologique*, 104, 121-146.
- Pinard, R., Potvin, P. et Rousseau, R. (2004). Le choix d'une approche méthodologique mixte de recherche en éducation. *Recherches qualitatives*, 24, 58-80.

- Pothier, B. (2011). *Contribution de la linguistique à l'enseignement du français : CLEF*. Presses de l'Université du Québec.
- Pothier, B. et Pothier, P. (2004). *ÉOLE, Échelle d'acquisition en orthographe lexicale pour l'école élémentaire, du CP au CM2*. Retz.
- Powell, M. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. *Undefined*. <https://www.semanticscholar.org/paper/The-BOBYQA-algorithm-for-bound-constrained-without-Powell/0d2edc46f81f9a0b0b62937507ad977b46729f64>
- Préfontaine, R.-R. (1979). *Echelle de langage et d'acquisition orthographique*. Le Sablier.
- Rey, A. (1998). *Dictionnaire historique de la langue française* (Le Robert).
- Saussure, F. de. (1985). *Cours de linguistique générale* (C. Bally, A. Sechehaye et A. Riedlinger, Éds.; Ed. critique). Payot.
- Scarborough, H. S. (1998). Early Identification of Children At Risk for Reading Disabilities. Dans B. K. Shapiro, P. J. Accardo et A. J. Capute, *Specific Reading Disability: A View of the Spectrum* (p. 36). York Press.
- 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. <https://doi.org/10.1037/h0087331>
- Seymour, P. H. K. (1997). Foundations of Orthographic Development. Dans C. A. Perfetti, L. Rieben et M. Fayol, *Learning to Spell: Research, Theory, and Practice Across Languages* (p. 319-337). Routledge.
- Seymour, P. H. K. (2007). Continuity and discontinuity in the development of single-word reading : Theoretical speculations. Dans E. L. Grigorenko et A. J. Naples, *Single-Word Reading Behavioral and Biological Perspectives* (p. 1-24). Psychology Press.

- Seymour, P. H. K., Aro, M. et Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94(2), 143-174.  
<https://doi.org/10.1348/000712603321661859>
- Seymour, P. H. K. et Evans, H. M. (1999). Foundation-Level Dyslexia : Assessment and Treatment. *Journal of Learning Disabilities*, 32(5), 394-405.
- Shaywitz, B. A., Fletcher, J. M. et Shaywitz, S. E. (1995). Defining and Classifying Learning Disabilities and Attention-Deficit/Hyperactivity Disorder. *Journal of Child Neurology*, 10(1), 50-57. <https://doi.org/10.1177/08830738950100S111>
- Sprenger-Charolles, L. (2012). Principaux facteurs expliquant la réussite et l'échec de l'apprentissage de la lecture. *ANAE - Approche Neuropsychologique des Apprentissages chez l'Enfant*, 24(116), 13-20.
- Sprenger-Charolles, L. (2017). Une progression pédagogique construite à partir de statistiques sur l'orthographe du français (d'après Manulex-Morpho) : Pour les lecteurs débutants et atypiques. *ANAE - Approche Neuropsychologique des Apprentissages chez l'Enfant*, 29(148), 247-256.
- Sprenger-Charolles, L. et Colé, P. (2013). *Lecture et dyslexie. Approche cognitive*. Dunod; Cairn.info. <https://www.cairn.info/lecture-et-dyslexie--9782100582921.htm>
- 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), 134-165. <https://doi.org/10.1006/jecp.1997.2422>
- Sprenger-Charolles, Siegel, L. S. et Béchenec, D. (1997). Beginning Reading And Spelling Acquisition In French: A longitudinal Study. Dans C. A. Perfetti, L. Rieben et M. Fayol, *Learning to Spell: Research, Theory, and Practice Across Languages* (p. 339-359). Routledge.

- Stanké, B. (2009). *Facteurs cognitifs liés à l'acquisition du lexique orthographique* [Thèse, Université de Montréal].
- [https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/3632/Stanke\\_Brigitte\\_2009\\_these.pdf](https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/3632/Stanke_Brigitte_2009_these.pdf)
- Stanké, B., Collette, É., Lefebvre, P., Rousseau, N., Schelstraete, M.-A. et Valdois, S. (2016). *Les dyslexies-dysorthographies* (p. xvi, 216 pages). Presses de l'Université du Québec.
- Stanké, B. et Dumais, C. (2017a). An, am, en, em, aon ou aen ? Comment écrire le son /ã/ ? *Vivre le primaire*, 1.
- Stanké, B. et Dumais, C. (2017b). Comment écrire le son /ẽ/ ? *Vivre le primaire*, 36-37.
- Stanké, B. et Flessas, J. (2013). *Outil de dépistage d'élèves à risque de présenter des difficultés d'apprentissage du langage écrit*. Institut de Recherches Psychologiques.
- Stanké, B., Mené, M. L., Rezzonico, S., Moreau, A., Dumais, C., Robidoux, J., Dault, C. et Royle, P. (2019). ÉQOL : Une nouvelle base de données québécoise du lexique scolaire du primaire comportant une échelle d'acquisition de l'orthographe lexicale. *Corpus*, 19, Article 19. <http://journals.openedition.org/corpus/3818>
- Stanké, B., Moreau, A., Dumais, C., Royle, P. et Rezzonico, S. (2020). *Nouvelle approche basée sur un enseignement orthographique favorisant l'apprentissage de la production écrite et de l'orthographe lexicale des élèves faibles orthographes de 3e année du primaire* (2015-LC-187661). Université de Montréal.
- Stanké, B., Rezzonico, S., Moreau, A. C., Robidoux, J. et Royle, P. (2021). Impact de l'enseignement explicite des régularités graphotactiques et morphologiques sur l'apprentissage de l'orthographe lexicale auprès de différents profils d'élèves faibles orthographes de CE2. *ANAE*, 173, 1-11.

- Stanké, B., Rezzonico, S., Moreau, A., Dumais, C. et Royle, P. (2018). *ÉQOL : Guide pédagogique* (19). Université de Montréal.
- St-Pierre, M.-C., Dubé, J.-F. et Croteau, M. (2013). Comment favoriser le développement des habiletés orthographiques auprès d'enfants en difficulté âgés de 7 à 8 ans? Dans D. Daigle, I. Montésinos-Gelet et A. Plisson, *Orthographe et populations exceptionnelles : Perspectives didactiques* (p. 57-80). Presses de l'Université du Québec.
- Thomas, D., Zhu, P., Zumbo, B. et Dutta, S. (2008). On Measuring the Relative Importance of Explanatory Variables in a Logistic Regression. *Journal of Modern Applied Statistical Methods*, 7(1). <https://doi.org/10.22237/jmasm/1209614580>
- Treiman, R. et Kessler, B. (2014). *How Children Learn to Write Words*. Oxford University Press.
- Treiman, R., Kessler, B. et Caravolas, M. (2019). What methods of scoring young children's spelling best predict later spelling performance? Scoring of Young Children's Spelling. *Journal of Research in Reading*, 42(1), 80-96. <https://doi.org/10.1111/1467-9817.12241>
- Tremblay, O., Lefrançois, P. et Lombard, V. (2013). Des listes de mots pour les trois cycles du primaire au Québec. Dans C. Garcia-Debanc, C. Masseron et C. Ronveaux, *Enseigner le lexique*. PU de Namur.
- UNESCO. (2008). *Statistiques internationales sur l'alphabétisme : Examen des concepts, de la méthodologie et des données actuelles*. Institut de statistique de l'UNESCO.
- UNESCO. (2016). *L'éducation pour les peuples et la planète : Créer des avenir durables pour tous | Global Education Monitoring Report*. <https://fr.unesco.org/gem-report/node/1279>
- Valdois, S. (2014). Qu'entendre par déficit visuo-attentionnel en contexte dyslexique ? *A.N.A.E. Approche neuropsychologique des apprentissages chez l'enfant*, 128, 27-35.
- Valdois, S., Guinet, E. et Embs, J.-L. (2017). *Évaluation de l'empan visuo-attentionnel en contexte dyslexique*. Happy Neuron.

- Van der Maren, J.-M. (2003). *La recherche appliquée en pédagogie : Des modèles pour l'enseignement* (De Boeck et Larcier).
- Véronis, J. (1988). From sound to spelling in French: Simulation on a computer. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 8(4), 315-334.
- Wanzek, J., Vaughn, S., Wexler, J., Swanson, E. A., Edmonds, M. et Kim, A.-H. (2006). A synthesis of spelling and reading interventions and their effects on the spelling outcomes of students with LD. *Journal of Learning Disabilities*, 39(6), 528-543.  
<https://doi.org/10.1177/00222194060390060501>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (2<sup>e</sup> éd.). Springer/Sci-Tech/Trade.