



Impacts d'entraînements ciblant les unités morphologiques et orthographiques sur le développement de l'identification des mots et de la production orthographique des enfants avec et sans trouble de la lecture

THÈSE

Soutenue publiquement le 13 septembre 2023

par **Claire FONTAA**

En vue de l'obtention du titre de Docteur en Psychologie

Université de Strasbourg

École doctorale Sciences Humaines et Sociales – Perspectives Européennes

Laboratoire de Psychologie des Cognitions (UR 4440)

Dirigée par :

Elisabeth DEMONT (Professeur des universités, Université de Strasbourg)

Anne-Sophie BESSE (Maître de conférences, Université de Strasbourg)

Rapporteurs :

Séverine CASALIS (Professeur des universités, Université de Lille)

Sébastien PACTON (Professeur des universités, Université de Paris Descartes)

Autres membres du jury :

Eva COMMISSAIRE (Maître de conférences HDR, Université de Strasbourg)

Anila FEJZO (Professeur des universités, Université du Québec à Montréal)

Pascal ZESIGER (Professeur des universités, Université de Genève)

Remerciements

Tout d'abord, je remercie Séverine Casalis, Anila Fejzo, Sébastien Pacton et Pascal Zesiger, de m'avoir fait l'honneur d'accepter de faire partie de mon jury de thèse. Vos publications ont jalonné le parcours de mon doctorat, et je suis réellement honorée de voir mon nom imprimé aux côtés des vôtres.

Je remercie bien sûr mes trois encadrantes, Anne-Sophie, Eva et Elisabeth. Vous avez su avant moi que j'étais capable de terminer cette thèse. Merci pour votre bienveillance, votre gentillesse, votre patience, et surtout, pour l'enseignement que vous m'avez toutes les trois prodigué pendant ces quatre ans. J'espère du fond du cœur m'être montrée à la hauteur de la confiance que vous avez placé en moi.

J'aimerais remercier mes parents, Elisabeth et Vincent, mon cher frère François (le sang), ma (très) belle-sœur Léa, et ma sœur adoptive (chérie), Emma. Merci pour tous ces dimanches passés à nous disputer sur ce que sont une « vraie » thèse et un·e « vrai·e » docteur·e. Vous avez été, à chaque moment passé avec vous, une bouffée d'air frais, et parfois, le souffle qui me manquait pour repartir. À mon neveu Gabriel, que j'aime déjà trop pour pouvoir l'écrire convenablement ici, et qui peut-être me lira un jour.

Je remercie les amis qui ont marché à mes côtés pendant ces quatre années, qu'ils aient été eux-mêmes engagés dans cette folie qu'est le doctorat ou non. Merci Aedena (qui n'est pas du tout hyperactive), Sophie (qui est une mère fantastique), Aurélie (qui craignait sans doute que je l'oublie), Nicolas (qui me doit une partie de poker), Marie S (avec qui il me tarde de boire un nouveau verre) pour être qui vous êtes, tout simplement. Merci à Marie-Ange pour tous ces repas partagés et ces fous-rires, merci à Jad pour avoir toléré avec bienveillance mes imitations, merci à Séléna, Cécile et Bérénice pour avoir œuvré pour préserver la bonne entente qui existe encore aujourd'hui entre les doctorants du LPC. Merci à Estelle d'avoir été une co-thésarde soutenante et bienveillante. Merci à Charlotte, Hiba, Jean-Baptiste, Marie M. et Thomas pour avoir partagé le « bureau du haut » avec moi durant ma dernière année de thèse (ça n'a pas dû être simple tous les jours !). Merci à Marie du Sulisom pour m'avoir redonné le sourire plus d'une fois, merci à Marine pour son écoute lorsque

j'étais vulnérable, merci à Thibaut pour son humour. Merci à Andrei et Anne pour leur humour et leur gentillesse. Merci à Eva L., pour avoir osé, à chaque occasion, me faire des retours pertinents qui m'ont réellement permis de progresser, même si cela signifiait passer pour la méchante. Merci à Anne-Clémence, Carolina, Doha, Eléna, Ianis, Julie, Marie R. et Sophie J.F. pour nos échanges, parfois brefs mais toujours positifs. Bon courage à celles qui commencent ou qui s'accrochent pour continuer, et bonne soutenance à ceux qui finissent ! Merci enfin à tous les membres du LPC que je n'ai pas cités mais qui ont fait de ce labo un lieu chaleureux et ensoleillé.

Merci à l'ensemble vocal féminin Voy'Elles, qui aura été un compagnon de voyage inestimable durant toutes ces années. Merci à Léna et Frédérique, qui par leur intérêt pour mes travaux de recherche, permettront à cette thèse de continuer à vivre après la soutenance. Merci aux grenouilles de CoCyclics, et tout particulièrement à Story et Mosswing pour leurs :chocolat: et leurs :pompom:!

Merci aux étudiant·e·s en orthophonie et en psychologie de l'Université de Strasbourg pour m'avoir aidée dans mes travaux de thèse. J'espère avoir réussi à vous transmettre des choses positives sur la recherche durant le court temps que nous avons passé ensemble. Sachez, en tout cas, que chacun·e d'entre vous m'a énormément appris, à la fois sur mon propre sujet de thèse et sur ce que signifie enseigner.

Merci sans exception à tous les enfants, leurs parents, leurs enseignant·e·s et leurs orthophonistes, qui ont accepté de m'accorder, à moi ou à mes collègues, du temps pour pouvoir mener à bien le travail qui s'est aujourd'hui concrétisé. Si je n'avais trop de respect pour votre anonymat, je dédierais cette thèse à chacun·e d'entre vous.

Mon cher Lolo, je t'ai laissé pour la fin. J'ai souvent pensé à toi comme un quatrième directeur de thèse, mais fidèle à ta nature de caméléon, tu as été bien plus que cela. Tu as été mon co-thésard, mon co-directeur, mon collègue, mon ami, mon amour. Sans toi, ce voyage aurait un goût de larmes et de frustration à l'arrivée. Tu as su me les épargner pour qu'il n'en reste que l'essentiel : l'expérience et l'émotion. Je vole de mes propres ailes, mais c'est toi qui m'as appris à m'en servir. Je t'écris ici ce que j'ai su dès notre premier verre : tu seras pour toujours l'homme de ma vie.

Résumé

La maîtrise de la lecture et de l'écriture constitue un enjeu particulièrement critique pour les individus présentant un trouble de la lecture. Outre les correspondances graphème-phonème, indispensables à l'instauration d'une procédure d'auto-apprentissage, il existe des unités morphologiques et graphotactiques qui concourent au développement de l'identification des mots et à la production orthographique. La présente thèse visait à évaluer l'impact d'entraînements ciblant ces unités sur le développement de l'identification et de la production orthographique d'enfants avec ou sans trouble de la lecture.

Pour cela, nous avons réalisé une méta-analyse sur 40 études indépendantes portant sur les entraînements morphologiques (EM), qui retrouve un impact positif de l'EM pour la conscience morphologique, la production orthographique de morphèmes, la production orthographique de mots et la lecture morphologique (i.e., la lecture de mots présentant au moins deux morphèmes). Nous avons par ailleurs conduit une étude d'entraînement morphologique impliquant 51 enfants scolarisés de la 3^{ème} à la 5^{ème} année d'instruction. Nos résultats indiquent que l'entraînement à la production orthographique de morphèmes, qu'il soit ou non complété d'exercices à l'analyse morphologique, améliore l'identification de mots irréguliers et la production orthographique de mots complexes, comparativement à un entraînement aux mathématiques. Enfin, nous avons conduit une étude d'entraînement orthographique sublexical ciblant les unités phonologiques, morphologiques et graphotactiques, dispensé chez 51 participants, selon un mode soit explicite soit implicite. Des mesures répétées ont permis de dégager différents « profils d'apprenants », et suggèrent une efficacité différenciée en fonction du mode d'administration et du type d'apprenant (avec ou sans trouble de la lecture ; niveau de scolarisation).

Nos résultats confirment l'intérêt de considérer les unités morphologiques et orthographiques pour soutenir l'apprentissage de l'identification des mots et de la production orthographique. Toutefois, en raison de biais méthodologiques, des études ultérieures sont nécessaires pour compléter et confirmer ces résultats.

Mots-clés : entraînement morphologique, entraînement orthographique, identification des mots, production orthographique

Abstract

Achieving reading and writing mastery is a critical issue for individuals with a reading disorder. Besides grapheme-to-phoneme correspondences (CGP) which are essential to the acquisition of a self-teaching procedure, there are morphological and graphotactic units that concur to word identification and spelling development. The present thesis aimed to evaluate the impact of training those units on the development of word identification and spelling in children with or without a reading disorder.

To that end, we conducted a meta-analysis of 40 independent morphological training (MT) studies and found a positive impact of MT on morphological awareness, morpheme spelling, word spelling and morphological reading. We also conducted a MT study involving 51 3rd-to-5th grade schoolchildren. Our results show morpheme spelling training, whether or not it comprises morphological analysis training, seems to benefit irregular word identification and complex word spelling, when compared to a mathematical training. Lastly, we conducted a sublexical orthographic training targeting phonological, morphological and graphotactic units in 51 participants, which we delivered either explicitly or implicitly. Repeated measures analysis seemed to outline different “learning profiles”, suggesting a differential efficiency of the program, depending on delivery mode and on the type of learner (with or without a reading disorder; class level).

All in all, our results appear to confirm that morphological and orthographic units are relevant in supporting the development of word identification and spelling. However, due to methodological bias, further studies are required to complete and confirm such results.

Keywords: morphological training, orthographic training, word identification, spelling.

Table des matières

Remerciements.....	2
Résumé	5
Abstract	6
Table des matières	7
Liste des tableaux	11
Liste des figures	13
Abréviations	15
Présentation de la thèse.....	19
<i>Cadre théorique</i>	24
Chapitre 1 – Le lexique orthographique	25
1.1 Le lexique orthographique.....	25
1.2 Modélisation du lexique orthographique en lecture/écriture	27
1.2.1 Le lexique orthographique dans le processus de lecture experte : le modèle DRC..	27
1.2.2 Le lexique orthographique dans le processus d’écriture experte : le modèle RET..	38
1.3 Le développement du lexique orthographique	44
1.3.1 Le modèle en six stades de Frith (1985)	45
1.3.2 L’hypothèse de l’auto-apprentissage.....	50
1.3.3 Le modèle ST-DRC.....	54
1.4 La théorie de l’Intégration des Multiples Patterns (IMP)	62
Chapitre 2 – Les patterns phonologiques	65
2.1. Connaissance des lettres et conscience phonémique	66
2.1.1 La connaissance des lettres	66
2.1.2 La conscience phonémique.....	67
2.2 Le rôle des patterns phonologiques dans l’identification des mots et la production orthographique	70
2.3 De l’insuffisance des patterns phonologiques	73
Chapitre 3 – Les patterns morphologiques	76
3.1 Morphologie et conscience morphologique.....	76
3.1.1 La morphologie.....	76
3.1.2 La conscience morphologique.....	78

3.2 Le rôle des patterns morphologiques dans l'identification des mots et la production orthographique	81
3.2.1 Modéliser la contribution de la conscience morphologique à la lecture et l'écriture : le Morphological Pathways Framework	81
3.2.2 Morphologie et lecture : identification des morphèmes	83
3.2.3 Morphologie et orthographe : l'invariance morphémique	88
3.3 Limites des patterns morphologiques	90
Chapitre 4 – Les patterns graphotactiques	92
4.1 Définition des patterns graphotactiques	92
4.1.1 Compétences orthographiques sublexicales	92
4.1.2 Connaissance orthographique	96
4.2 Le rôle des patterns graphotactiques dans l'identification des mots et la production orthographique	98
4.2.1 « Apprenabilité » des patterns graphotactiques	98
4.2.2 Étude d'Ise et Schulte-Körne (2010)	100
4.3 Perspectives de recherche concernant les patterns graphotactiques	103
Problématique	104
<i>Contributions scientifiques</i>	108
Chapitre 5 – Vingt-cinq ans d'étude de l'impact des entraînements morphologiques sur le développement du langage écrit : une méta-analyse	109
5.1. Résumé long en français	110
5.1.1 Introduction	110
5.1.2 Méthode	111
5.1.3 Résultats	113
5.1.4 Discussion	116
5.2 Article soumis	118
5.2.1 Abstract	119
5.2.2 Introduction	120
5.2.3 Methods	125

5.2.4 Results	137
5.2.5 Discussion	146
5.2.6 References	160
5.2.7 Tables	176
5.2.8. Figures	196
5.2.9 Appendices	202
Chapitre 6 – Favoriser le développement de l’identification des mots et de la production orthographique par un entraînement morphologique : efficacité, limites et perspectives	212
6.1 Introduction	212
6.2 Étude 1.....	216
6.2.1 Méthode.....	216
6.2.2 Résultats	228
6.2.3 Discussion	236
6.3 Étude 2.....	239
6.3.1 Méthode.....	239
6.3.2 Résultats	241
6.3.3 Discussion	253
6.4 Discussion générale.....	255
Chapitre 7 – Exploration de l’impact d’un entraînement orthographique sublexical sur le développement de la production orthographique et l’identification des mots en français : l’exemple du phonème final /o/.....	262
7.1 Introduction	262
7.2 Méthode.....	268
7.2.1 Participants	268
7.2.2 Protocole d’évaluation.....	271
7.2.3 Protocole d’entraînement.....	276
7.2.4 Design	283

7.2.5 Analyses statistiques	284
7.3 Résultats	287
7.3.1 Résultats sur les mesures pré-/post-tests.....	287
7.3.2 Résultats sur les mesures répétées	294
7.4 Discussion	314
<i>Discussion générale</i>	324
<i>Conclusion</i>	341
<i>Annexes</i>	343
Annexe 1	344
Annexe 2	345
Annexe 3	346
Annexe 4	347
Annexe 5	348
Annexe 6	349
Annexe 7	351
Annexe 8	352
Annexe 9	354
Annexe 10	356
Annexe 11	357
Annexe 12	358
Annexe 13	359
Annexe 14	360
<i>Références bibliographiques</i>	362

Liste des tableaux

Tableau 1. Caractéristiques des groupes EM et EC au pré-test	229
Tableau 2. Moyennes et écarts-types des groupes EM et EC en pré- et post-test sur l'ensemble des variables dépendantes	230
Tableau 3. Effets simples du temps, interactions point de mesure * groupe d'entraînement et contrastes sur l'ensemble des mesures de la DIMO	235
Tableau 4. Caractéristiques des groupes EM, EC-PO et EC-M au pré-test	242
Tableau 5. Moyennes et écarts-types des groupes EM, EC-PO et EC-M sur l'ensemble des variables dépendantes	246
Tableau 6. Analyses de contraste des performances à la DIMO	252
Tableau 7. Moyennes et écarts-types des deux groupes d'entraînement au pré-test sur les variables contrôles	269
Tableau 8. Moyennes, écarts-types et comparaison des scores obtenus à la MEREP à l'issue de la séance 1	276
Tableau 9. Description synthétique des séances et de la prédominance de la dimension phonologique, morphologique ou graphotactique	277
Tableau 10. Moyennes et écarts-types des deux groupes d'entraînement en pré- et post-test sur les variables d'intérêt et sur la variable contrôle	288
Tableau 11. Corrélations partielles entre le gain à l'Alouette-R et les variables pré-test	290
Tableau 12. Corrélations partielles entre le gain à la DIPHO et les variables pré-test	293
Tableau 13. Synthèse de l'évolution du niveau et de la pente entre la phase A et la phase B pour chaque profil de participant.	297
Tableau 14. Synthèse de l'évolution du niveau et de la pente entre la phase A et les trois premiers points de la phase B pour chaque profil de participant.	297

Tableau 15. Indices de non-recouvrement entre les phases A et B chez les sujets de 2ème NL	310
Tableau 16. Indices de non-recouvrement entre les phases A et B chez les sujets de 2ème TL	310
Tableau 17. Indices de non-recouvrement entre les phases A et B chez les sujets de 4ème NL	311
Tableau 18. Indices de non-recouvrement entre les phases A et B chez les sujets de 4ème TL	312
Tableau 19. Indices de non-recouvrement entre les phases A et B chez les sujets de 6ème NL	312
Tableau 20. Indices de non-recouvrement entre les phases A et B chez les sujets de 6ème TL	313

Liste des figures

Figure 1. Modèle à Double Voie en Cascade, adapté et traduit de Coltheart et al. (2001) ____	29
Figure 2. Modèle à Double Voie de Rapp, Epstein et Tainturier, 2002, adapté et traduit ____	40
Figure 3. Modèle en six stades de lecture et d'écriture, adapté et traduit de Frith (1985) ____	46
Figure 4. Modèle d'Auto-Apprentissage à Double Voie en Cascade, adapté et traduit de (Pritchard, 2013) _____	55
Figure 5. Apprentissage orthographique type-based d'un mot régulier d'après Pritchard et al. (2018) _____	57
Figure 6. Apprentissage orthographique type-based d'un mot irrégulier d'après Pritchard et al. (2018). _____	59
Figure 7. Processus de décodage morphologique et d'analyse morphologique au sein du Morphological Pathways Framework (Levesque et al., 2021) _____	82
Figure 8. Évolution des performances des groupes EM et EC sur la lecture de mots irréguliers peu fréquents (BALE) _____	233
Figure 9. Algorithme pour la transcription du phonème /o/ en position finale _____	283
Figure 10. Performances des sujets de 2ème année d'instruction normo-lecteurs aux mesures répétées, en phase A (avant les pointillés) et en phase B (après les pointillés) _____	295
Figure 11. Performances des sujets de 2ème année d'instruction avec un trouble de la lecture aux mesures répétées, en phase A (avant les pointillés) et en phase B (après les pointillés)	299
Figure 12. Performances des sujets de 4ème année d'instruction normo-lecteurs aux mesures répétées, en phase A (avant les pointillés) et en phase B (après les pointillés) _____	301
Figure 13. Performances des sujets de 4ème année d'instruction avec un trouble de la lecture aux mesures répétées, en phase A (avant les pointillés) et en phase B (après les pointillés)	303
Figure 14. Performances des sujets de 6ème année d'instruction normo-lecteurs aux mesures répétées, en phase A (avant les pointillés) et en phase B (après les pointillés) _____	304

Figure 15. Performances des sujets de 6ème année d’instruction avec un trouble de la lecture aux mesures répétées, en phase A (avant les pointillés) et en phase B (après les pointillés) 306

Abréviations

ANOVA = Analysis of Variance [Analyse de variance]

BALE = Batterie Analytique du Langage Ecrit

CDP+ model = Connectionist Dual Process Plus model [modèle Connexionniste au Traitement
Duel +]

CGP = Correspondances graphème-phonème

CM = 1. Conscience morphologique ; 2. Corrects / Mots (indice de précision de l'Alouette)

CP = Conscience phonologique

CPG = Correspondances phonème-graphème

COMOS = Conscience Morphologique – Suffixation

CTL = Mots corrects sur Temps de lecture

DIMO = Dictée Morphologique

DIPHO = Dictée du Phonème final O

DRC model = Dual route cascaded model [Modèle à Double Voie de lecture en Cascade]

DSM-V = Diagnostical Statistical Manual – 5th edition [Manuel diagnostique et statistique –
5ème édition]

EC = Entraînement contrôle

EC-M = Entraînement contrôle – Mathématiques

EC-PO = Entraînement contrôle – Production orthographique

EE = Entraînement explicite

EI = Entraînement implicite

EM = Entraînement morphologique

EOS = Entraînement orthographique sublexical

ES = Effect size [taille d'effet]

EVA = Empan visuo-attentionnel

EVIP = Echelle de Vocabulaire en Images

IMP theory = Integration of Multiple Pattern theory [Théorie de l'Intégration des Patterns Multiples]

IRR = Mots irréguliers

MA = Morphological awareness [Conscience morphologique]

MBD = Multiple Baseline Design [Design en lignes de base multiples]

ME = Mots entraînés

MEA = Modèle à effet aléatoire

MEF = Modèle à effet fixe

MEREP = dictée Mesures répétées

MPC = Matrices Progressives de Raven – Colorisées

MPF = Morphological Patterns Framework [Modèle des Patterns Morphologiques]

MR = Morphological reading [Lecture morphologique]

MS = Morpheme spelling [Production orthographique de morphèmes]

MT = Morphological training [Entraînement morphologique]

NAP index = Non-overlap of All Pairs [Indice de non-recouvrement de toutes les paires]

NAPr = NAP rééchelonné

NE = Non-entraînés

NL = Normo-lecteur

NM B = Non-mots bisyllabiques

ODÉDYS = Outil de Dépistage des Dyslexies

PA = Phonological awareness [Conscience phonologique]

PD = Phonological decoding [Décodage phonologique]

PGC = Phoneme-to-grapheme correspondances [Correspondances phonème-graphème]

POM = Production orthographique de morphèmes

PR = Phonological recoding [Recodage phonologique]

REG S = Mots réguliers simples

RET model = modèle de Rapp, Epsein et Tainturier

RC = Reading comprehension

SCD = Single-Case Design

SLL = Second-language learner [Apprenant d'une langue seconde]

SMD = Standardized mean difference [Différence moyenne standardisée]

ST = Self-Teaching [Auto-Apprentissage]

ST-DRC model = Self-Teaching Dual Route Cascaded model [modèle d'Auto-Apprentissage -
Double Voie de lecture en Cascade]

SMTL = Split-Middle Trend Line

TE = Taille d'effet

TL = Trouble de la lecture

WS = Word spelling [Production orthographique de mots]

Présentation de la thèse

Au XXI^{ème} siècle, la lecture constitue une activité indispensable à la vie en société. Ne pas la maîtriser représente un handicap majeur, limitant considérablement l'accès à l'information, à la culture, aux services publics, à la réalisation des rôles sociaux et à l'épanouissement de soi. Les derniers calculs de prévalence estiment que 7% des français souffrent d'illettrisme (Jonas, 2012). Au vu de la gravité des conséquences en termes de participation sociale et de santé publique, en particulier dans un contexte d'informatisation des services renforcée par la crise du COVID-19, la lutte contre l'illettrisme représente un enjeu essentiel pour lequel la recherche en psycholinguistique demeure essentielle.

L'illettrisme est défini comme la présence de « difficultés graves ou fortes » dans au moins un des trois domaines fondamentaux de l'écrit [i.e., lecture de mots, production de mots écrits, compréhension de texte simple] chez des personnes ayant été scolarisées dans le pays dans lequel elles résident » (Jonas, 2012). Il se distingue ainsi de l'analphabétisme, qui renvoie au fait de n'avoir jamais appris à lire. Les causes de l'illettrisme sont nombreuses et multifactorielles (Vágvölgyi et al., 2016). Dans le cadre de cette thèse, nous avons choisi de nous intéresser à un facteur de risque d'illettrisme d'ordre langagier : les troubles de la lecture.

Environ 6% de la population scolaire en France présenterait un trouble de la lecture (Di Folco et al., 2022). Depuis la 5^{ème} édition du Manuel Diagnostique et Statistique des Troubles Mentaux (DSM-V, American Psychiatric Association, 2013), ils ne sont plus définis en tant qu'entité diagnostique propre mais comme faisant partie des troubles spécifiques des apprentissages. L'atteinte de la lecture et/ou de l'orthographe est ainsi, dans cette classification, considérée comme un symptôme parmi d'autres (e.g., compréhension écrite, raisonnement mathématique...) d'un trouble touchant les facultés d'apprentissage en général. Malgré cette parution véhiculant une prise de position forte quant aux soubassements

théoriques des troubles de la lecture, ces derniers sont encore largement référés aujourd'hui sous le terme de dyslexie développementale. Cette appellation, utilisée depuis le début du XX^e siècle (Orton, 1937) afin de décrire des cas d'incapacité sévère à apprendre à lire en l'absence de toute déficience intellectuelle, neurologique et sensorielle, est toujours en usage dans la littérature¹. Pourtant, son origine cognitive et biologique demeure difficile à cerner (Peterson & Pennington, 2015).

Selon la proposition d'Elliott (2020), quatre acceptions différentes de la dyslexie développementale (ci-après, dyslexie) cohabitent actuellement, tant dans la communauté scientifique que parmi les praticiens. Pour certains (Fletcher et al., 2019; Pennington et al., 2019), la dyslexie doit être comprise comme synonyme de déficits de lecture, survenant en-dehors de toutes conditions plus générales pouvant les expliquer à elles seules (nommément, une déficience intellectuelle ou sensorielle). Cette définition présente l'avantage et l'inconvénient de ne pas s'attacher à la cause neurobiologique du trouble, en ce qu'elle accepte que de multiples combinaisons de facteurs puissent être à l'origine de difficultés importantes en lecture. La deuxième définition suppose quant à elle qu'il existe des différences qualitatives entre ces « mauvais lecteurs » et les individus « véritablement » dyslexiques (Lyon et al., 2003; Miciak & Fletcher, 2020), la plus consensuelle étant une difficulté majeure dans les habiletés phonologiques, et plus particulièrement en conscience phonémique (Carioti et al., 2021; Melby-Lervåg et al., 2012; Saksida et al., 2016), i.e., la capacité d'analyser et de manipuler les phonèmes ou plus petites unités de son de la langue. Ce déficit phonologique ne correspond cependant pas à « l'hypothèse d'un déficit unitaire » (Share, 2021a, p. 8), capable d'expliquer tous les cas de trouble de la lecture (Heim et al., 2008; Valdois et al., 2021). Il paraît donc difficile, à l'heure actuelle, de baser un diagnostic de dyslexie sur la seule présence de ce déficit phonologique. La troisième définition rejoint en

¹ Ainsi que l'attestent les 2810 résultats obtenus sur Google Scholar pour l'année 2022 avec le mot-clé « *developmental dyslexia* » (recherche réalisée le 9 juin 2023).

partie la seconde, en ce qu'elle suppose qu'il existe un facteur, à savoir, la résistance à l'intervention (Elliott & Gibbs, 2008), permettant de poser un diagnostic de dyslexie fiable. Selon cette définition, ne seraient qualifiées de dyslexiques que les personnes ayant épuisé toutes les possibilités d'interventions reconnues comme permettant d'améliorer efficacement les facultés de lecture. Dans cette acception, les personnes dyslexiques ne pourraient accéder à une lecture fonctionnelle, et seules les adaptations environnementales, telles que des aides techniques (e.g., logiciels de lecture et de dictée vocale), sont susceptibles de leur bénéficier. Néanmoins, cette définition est, à l'heure actuelle, difficilement opérationnalisable (détermination de la période de temps minimale durant lequel l'individu doit démontrer son incapacité à progresser ; recensement exhaustif de toutes les interventions potentiellement efficaces à proposer ; etc.) Enfin, la dernière définition correspond à une acception de la dyslexie comme un profil neurodéveloppemental atypique, impliquant un fonctionnement cérébral structurellement différent. Certains partisans de cette conception estiment que l'atteinte de la lecture ne serait que secondaire (Cooke, 2001), arguant « qu'une définition de la dyslexie en termes de performances aux tests de lecture n'aurait aucun espoir d'aboutir au bon diagnostic » (Frith, 1999, p. 197), puisque le trouble neurobiologique sous-jacent ne saurait disparaître avec une amélioration des possibilités de lecture permise par des mesures de remédiation efficaces. D'autres adeptes de cette définition décrivent les personnes dyslexiques comme des individus exceptionnels, dont la façon de fonctionner, bien qu'inadéquate pour l'enseignement ordinaire, implique également des aptitudes remarquables (pour une revue sur cette opinion, voir Kirby, 2019) et indispensables à la société (Taylor & Vestergaard, 2022), du fait principalement d'un avantage créatif par rapport aux personnes non-dyslexiques. Au-delà du fait qu'elle soit peu étayée de preuves (Erbeli et al., 2021; Ramus et al., 2018), cette vision, comme le note Elliott (2020), est notamment exploitée par des organismes de développement personnel (e.g., Davis, 1997), proposant des stages onéreux

de réappropriation du « don de dyslexie » pour toutes les personnes s'estimant concernées par ce problème.

La question de la définition de la dyslexie est donc complexe et repose, en fonction des chercheurs comme des praticiens, sur des soubassements théoriques très différents (Kirby, 2020; Zesiger, 2004). Nous nous sommes concentrés, dans le cadre de cette thèse, sur les troubles de la lecture en général (définition n°1), en incluant aussi bien des lecteurs en difficulté que des enfants diagnostiqués dyslexiques. Par ailleurs, dans une optique de prévention, nous avons également inclus des enfants tout-venants, dont les parents ne rapportaient pas de difficultés particulières dans ce domaine. La remédiation des troubles de la lecture, qu'elle soit mise en place par des enseignants ou des orthophonistes, constitue un élément essentiel de la lutte contre l'illettrisme. Bien que de nombreux travaux aient été répertoriés (Galuschka et al., 2020), les programmes d'intervention en langue française ne représentent qu'une faible part d'entre eux. Par ailleurs, les interventions les plus étayées de preuves se concentrent sur l'aspect « déchiffrage » de la lecture (Galuschka et al., 2020), et non sur l'enrichissement des représentations du lexique orthographique, défini dans le Chapitre 1 de la partie « Cadre théorique » du présent manuscrit. Ces interventions renvoient à des entraînements phonographiques ou phonémiques, que nous détaillons dans le Chapitre 2. Ces derniers, bien qu'essentiels pour permettre l'acquisition du principe alphabétique, n'ont pas démontré leur efficacité pour améliorer la fluence de lecture, i.e., le rapport entre la précision et la vitesse de lecture, ni pour la production orthographique (McArthur et al., 2018).

Cette thèse avait ainsi pour but d'explorer deux interventions sur des unités sublexicales susceptibles de favoriser l'apprentissage orthographique, à la fois en identification des mots et en production orthographique. Elles prenaient la forme d'un entraînement morphologique et d'un entraînement orthographique. Les principes sous-jacents

à ces interventions sont définis respectivement dans les Chapitres 3 et 4 de notre partie « Cadre théorique ». Concernant les entraînements morphologiques, nous présentons tout d'abord, dans le chapitre 5 de notre partie « Contributions scientifiques », une méta-analyse synthétisant les tailles d'effet de 40 études d'entraînement sur différents aspects de la lecture, de la production orthographique et des compétences associées (conscience morphologique, conscience phonologique, vocabulaire). Dans les chapitres 6 et 7 de cette même section, nous présentons la mise en place d'un entraînement morphologique et d'un entraînement orthographique auprès d'enfants scolarisés de la deuxième année de primaire (CE1) à la première année de secondaire (6^{ème}), ainsi que leur impact sur la précision et la fluence de lecture et sur la production orthographique.

Cadre théorique

Chapitre 1 – Le lexique orthographique

Au cours de ce chapitre, nous présenterons le cadre théorique en lien avec le lexique orthographique. Pour cela, nous l’aborderons à travers quatre sous-axes : la réflexion théorique ayant conduit à son élaboration, son rôle dans les modèles de lecture et d’écriture expertes, sa place dans l’acquisition de la lecture et de l’écriture chez l’apprenti-lecteur, et enfin, la façon dont la prise de conscience des régularités de la langue peut contribuer à son enrichissement.

1.1 Le lexique orthographique

Le français est une langue alphabétique, ce qui signifie qu’elle repose sur un système de transcription des plus petites unités de son (ci-après, les phonèmes) via leurs unités graphiques correspondantes, les graphèmes. Les langues alphabétiques varient selon leur consistance, c’est-à-dire selon le « degré d’incertitude associé à l’écriture d’une association phonème-graphème ou à la lecture d’une association graphème-phonème » (Lété, 2008, p. 85). Ainsi, en français, la correspondance phonème-graphème (CPG) /a/ - A est très consistante, puisque que le son /a/ est presque toujours transcrit A ; à l’inverse, la CPG /o/ - O est très inconsistante, puisque /o/ peut se transcrire O aussi bien que EAU, AU, Ô, etc. Ces inconsistances se retrouvent pour le français principalement dans le sens de l’écriture (79.1% des mots monosyllabiques peuvent s’écrire de plusieurs manières différentes) plutôt que dans celui de la lecture (12.4% des mots monosyllabiques peuvent se lire de plusieurs manières différentes ; Ziegler et al., 1996). Autrement dit, s’il y a souvent en français plusieurs façons de transcrire un seul phonème, il n’y a généralement qu’une seule façon de lire un seul graphème (Ziegler, 2018).

Ces inconsistances de correspondances graphème-phonème (CGP), variant en fonction de la langue d’apprentissage (Seymour et al., 2003), ont un impact important sur l’acquisition de la lecture. Une étude de Goswami et al. (1998) a ainsi mis en évidence

une nette différence dans les capacités de lecture de non-mots (i.e., suites de lettres ne correspondant à aucun mot de la langue ; e.g., *JAJE) sans voisin orthographique², entre des enfants anglophones, francophones et hispanophones. À 7 ans, les enfants francophones n'étaient capables de lire correctement qu'environ la moitié des non-mots monosyllabiques présentés (52.9% de bonnes réponses), tandis que les enfants hispanophones de même âge en lisaient sans erreur la quasi-totalité (94.3% de bonnes réponses). Les enfants anglophones, à l'inverse, n'en lisaient correctement qu'une infime partie (11.8% de bonnes réponses). Des trois langues analysées, c'est l'anglais qui possède le plus haut taux d'inconsistance des CGP (30.7% des mots monosyllabiques anglais peuvent se lire de plus d'une manière ; Ziegler et al., 1997), tandis que l'espagnol constitue l'une des langues alphabétiques les plus consistantes (Seymour et al., 2003).

Du fait de l'irrégularité des associations graphème-phonèmes, l'aptitude de lecture dans les langues inconsistantes serait fortement dépendante de l'apprentissage orthographique, i.e., la capacité à apprendre et retenir les formes orthographiques des mots rencontrés (Share, 1999). Ces dernières seraient alors stockées dans un lexique orthographique, sorte de dictionnaire mental permettant de lire et d'orthographier les mots préalablement mémorisés. La présence d'un lexique orthographique chez les lecteurs, qu'ils soient locuteurs d'une langue consistante ou inconsistante (Besner & Smith, 1992; Pagliuca et al., 2008), est démontrée par deux phénomènes, observables dans la lecture à haute voix d'individus ayant été exposés à l'écrit : l'effet de lexicalité et l'effet de fréquence (Forster & Chambers, 1973). L'effet de lexicalité est défini par le fait qu'un mot appartenant à la langue de l'individu (e.g., PLEUR) est lu plus rapidement et précisément qu'un non-mot (e.g., *PLOUR) de même longueur ; l'effet de fréquence est quant à lui défini par le fait qu'un mot possédant une probabilité plus importante d'avoir

² Un voisin orthographique est un mot ayant une différence d'une seule lettre avec un mot cible ; e.g., PAIN et MAIN sont des voisins orthographiques.

déjà été rencontré à plusieurs reprises par l'individu (i.e., un mot fréquent tel que ÊTRE) est lu plus rapidement et plus précisément qu'un mot de même longueur mais rare et donc moins familier (i.e., un mot peu fréquent tel que NÈPE). Ces deux effets témoignent du fait que l'individu se base sur ses expériences de lecture antérieures pour lire les mots qui lui sont présentés. Les tâches de choix pseudo-homophonique (e.g., lequel de ces mots est un mot existant en français : RENTE ou *RANTE ?) attestent également d'un stock mobilisable de représentations orthographiques emmagasinées (Olson et al., 1994), sans lesquelles l'individu ne pourrait décider de la bonne orthographe entre deux propositions homophones.

1.2 Modélisation du lexique orthographique en lecture/écriture

La quasi-totalité des modèles de lecture et d'écriture contemporains incluent une couche correspondant au lexique orthographique. Nous avons choisi de nous concentrer sur le modèle DRC (*Dual Route Cascaded model* ou modèle à Double Voie de lecture en Cascade, Coltheart et al., 2001) pour étudier la modélisation du lexique orthographique en lecture à haute voix, et sur le modèle à double voie de Rapp et al. (2002) pour décrire cette modélisation dans le processus d'écriture. Ces deux modèles en particulier ont été retenus du fait de leurs nombreuses similarités, permettant de réaliser des parallèles entre le processus de lecture et d'écriture : plus précisément, le fait qu'ils s'appuient tous deux sur un processus à double voie, et qu'ils supposent que l'unité de traitement au niveau des activations phonologiques de la voie non- ou sublexicale soit le phonème, et non les sous-composants de la syllabe (attaque, rime).

1.2.1 Le lexique orthographique dans le processus de lecture experte : le modèle

DRC

Le modèle DRC est un modèle computationnel de lecture à haute voix et d'identification des mots, décrit par Coltheart et al. (2001) et établi sur la base du modèle

Interactive Activation and Competition (McClelland & Rumelhart, 1981) et des modèles Logogènes à double voie (Morton, 1969). Bien qu'il soit computationnel, il ne découle pas d'un algorithme d'apprentissage, tels qu'employés dans les modèles *Parallel Distributed Processing* de Seidenberg et McClelland (1989) : la structure du modèle DRC est élaborée par ses concepteurs, sur la base de l'observation des comportements humains. Ce modèle est caractérisé par un fonctionnement « en cascade » (*cascaded*) et non « en seuils » (*thresholded*) : l'activation d'une couche provoque immédiatement l'activation des couches adjacentes, sans dépendre d'un seuil d'activation qui devrait être atteint pour générer cette propagation. Les auteurs du modèle DRC le qualifient de « phonologiquement faible » (*weak phonological model*), en ce qu'il implique que le processus de lecture est toujours à la fois orthographique et phonologique, et non purement phonologique (contrairement à ce qui est supposé dans un modèle phonologiquement fort ou *strong phonological model*). Il convient de préciser que les simulations réalisées par Coltheart et al. (2001) ne comprenaient que des mots de langue anglaise (7881 entrées orthographiques pour 7131 entrées phonologiques³) et des non-mots (7000 entrées orthographiques et phonologiques), tous monosyllabiques⁴, de huit lettres ou moins.

La lecture d'un mot provoque l'activation de deux voies distinctes : la voie non-lexicale et la voie lexicale, elle-même distinguée en voie lexicale asémantique et voie lexicale sémantique (voir Figure 1). Sauf exception mentionnée, les couches adjacentes

³ Cette incongruence est due au nombre plus important en anglais d'hétérophones homographes (e.g., READ, qui peut être lu /ri:d/ ou /red/) par rapport aux homophones hétérographes (e.g., /hɪə/ peut être écrit HERE ou HEAR ; Coltheart et al., 2001).

⁴ Les auteurs ont en effet choisi de se concentrer sur les entrées monosyllabiques du fait de l'impossibilité de se positionner de manière forte sur la manière dont le modèle devrait gérer « l'accentuation et la réduction vocalique » des non-mots polysyllabiques (Coltheart et al., 2001).

communiquent entre elles de manière excitatrice et inhibitrice, à la fois en provenance et en direction de la couche suivante et précédente.

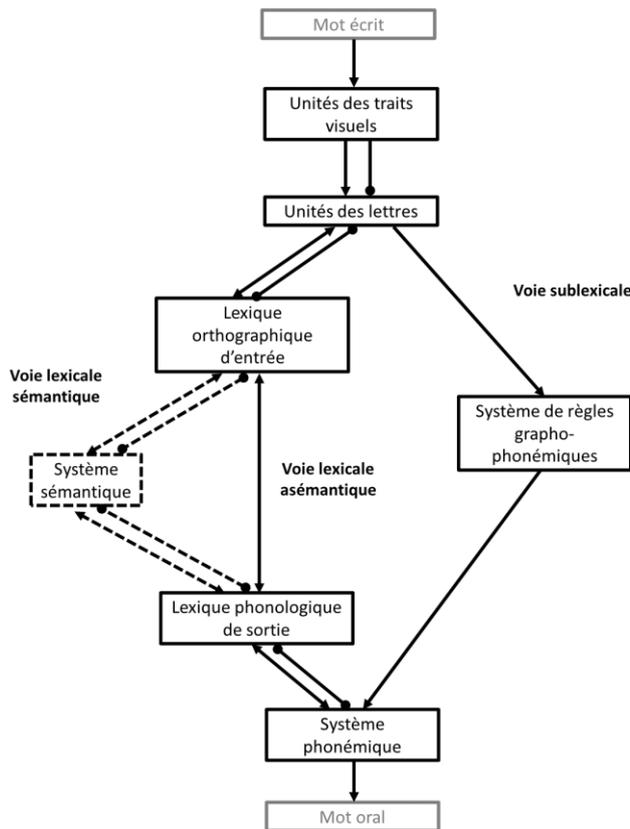


Figure 1. Modèle à Double Voie en Cascade, adapté et traduit de Coltheart et al. (2001)

NB : Les flèches triangulaires indiquent une connexion excitatrice, et les flèches rondes une connexion inhibitrice. Les éléments en pointillés signifient qu'ils n'ont pas été intégrés au modèle lors des simulations.

Les deux voies de lecture ne sont ni indépendantes ni réellement concurrentielles : du fait du fonctionnement en cascade du modèle, tout mot ou non-mot présenté active à des degrés divers les deux voies de lecture et les couches correspondantes. Ce point de vue est légitimé par les études comportementales mettant en évidence l'intervention de représentations lexicales (Andrews, 1997; Laxon et al., 1992; McCann & Besner, 1987) lors de tâches impliquant la lecture de non-mots. En effet, les non-mots pseudohomophones (e.g., *TRAISE), de même que les non-mots possédant de nombreux voisins orthographiques (e.g., *NIRE), sont lus avec plus de précision et plus rapidement

que des non-mots contrôle, tandis que lors de tâches de décision lexicale, ces non-mots sont rejetés moins souvent et moins rapidement que les non-mots contrôlés, suggérant une intervention des représentations lexicales lors du traitement des non-mots.

Trois couches sont communes aux deux voies de lecture : la couche des Unités des traits visuels (*Visual feature units*), la couche des Unités des lettres (*Letter units*) et la couche du Système phonémique (*Phoneme system*). Les couches restantes sont décrites au sein soit de la voie lexicale : Lexique orthographique d'entrée (*Orthographic input lexicon*), Lexique phonologique de sortie (*Phonological output lexicon*), Système sémantique (*Semantic system* ; uniquement dans la voie lexicale sémantique) ; soit de la voie non-lexicale : Système de règle grapho-phonémique (*Grapheme-phoneme rule system*).

Le processus de lecture tel que décrit par le DRC commence avec la couche des Unités des traits visuels⁵. Cette couche est composée de huit sous-emplacements : dans chacun d'entre eux, la lettre du mot présenté active ou non les traits visuels correspondant à cette lettre (Rumelhart & Siple, 1974) : par exemple, la lettre T va activer les unités « trait médial vertical » et « trait supérieur horizontal », et désactiver les traits ne correspondant pas à cette lettre (e.g., « trait inférieur horizontal » comme dans la lettre I). Si l'emplacement ne contient aucune lettre (e.g., quatrième emplacement pour un mot de trois lettres), tous les traits de cet emplacement sont désactivés. Les informations prises dans la couche des Unités des traits visuels se transmettent à la couche des Unités des lettres, et ce de deux façons : les traits visuels activés provoquent l'activation des unités de lettres correspondantes (e.g., le trait médial vertical et le trait supérieur horizontal activent l'unité de lettre « T ») et d'autre part l'inhibition de toutes les lettres restantes. Une fois débutée l'activation de la couche des Unités des lettres, les deux couches inférieures

⁵ Cette notion semble rejoindre la question des *abstract letter identities*, synthétisée dans la publication de Polk et al., 2009, et a été développée originellement par (Bowers, 2000; Coltheart, 1981).

correspondant à l'embranchement des voies lexicales et non-lexicales initient leurs traitements respectifs.

Concernant la voie lexicale⁶, la couche des Unités des lettres provoque dans la couche du Lexique orthographique d'entrée l'activation simultanée de tous les items présentant des similitudes avec le stimulus, c'est-à-dire non seulement l'item cible s'il est présent dans le lexique, mais aussi tous ses voisins orthographiques (Coltheart et al., 1977) : e.g., le stimulus MAIN active l'entrée correspondante ainsi que ses voisins orthographiques tels que PAIN). Parallèlement, les autres entrées orthographiques sont inhibées. Les items les plus fréquents sont activés préférentiellement par rapport aux items peu fréquents (e.g., dans le cas du stimulus MAIN, un mot très fréquent tel que PAIN sera davantage activé qu'un mot plus rare tel que GAIN). En parallèle de ces activations/inhibitions *en provenance* de la couche des Unités des lettres, des activations/inhibitions portées par la couche du Lexique orthographique d'entrée ont également lieu dans la direction opposée, *vers* la couche des Unités des lettres : l'item-cible de même que les voisins orthographiques du Lexique orthographique renforcent les lettres déjà activées dans la couche des Unités des lettres (e.g., l'item-cible MAIN du Lexique orthographique d'entrée renforce les lettres M-A-I-N dans les positions respectives 1, 2, 3 et 4) et participent à l'inhibition des lettres absentes. Ceci contribue à expliquer que les mots soient mieux lus que les non-mots, ces derniers ne bénéficiant pas de cette activation-inhibition rétroactive du Lexique orthographique d'entrée.

Au niveau du Lexique phonologique de sortie, les représentations phonologiques relatives aux entrées orthographiques activées sont récupérées, et activent à leur tour les unités phonémiques correspondantes dans la couche du Système phonémique. Chacun des

⁶ Seule la voie lexicale asémantique sera décrite ici, le fonctionnement de la voie lexicale sémantique n'étant pas à ce jour intégrée au modèle DRC, y compris dans ses versions les plus récentes (e.g., Mousikou et al., 2010).

phonèmes de la langue est donc soit activé soit désactivé, dans chacune des positions pertinentes (e.g., pour MAIN, la représentation phonologique /mɛ̃/ active en première position le phonème /m/ et désactive tous les autres phonèmes de la langue pour cette même position). En retour, les activations ainsi réalisées dans la couche du Système phonémique renforcent les représentations phonologiques activées dans le Lexique phonologique de sortie, et inhibent celles ne correspondant pas à la séquence phonologique du mot.

Concernant la voie non-lexicale également en lien avec la couche des Unités des lettres, bien qu'elle opère en parallèle de la voie lexicale, elle ne s'enclenche pas tout à fait en même temps que cette dernière et procède de manière plus lente. Ce délai est opérationnalisé dans le modèle DRC en imposant un nombre fixe de cycles. Dix cycles entre la couche d'Unités de la lettre et la couche du Lexique orthographique d'entrée doivent en effet avoir lieu avant que la couche du Système de règles grapho-phonémiques (conversion graphème-phonème ou CGP) ne s'applique à la première lettre la plus à gauche du mot. Une fois cette série initiale de cycles réalisée, la première lettre du mot provoque la recherche d'un phonème approprié. Les règles CGP sont interrogées dans un ordre bien précis, privilégiant les graphèmes de plus de trois lettres (e.g., EUIL) et trigraphes (e.g., AIN) ; sont ensuite consultées les graphies contextuelles (e.g., G dont la prononciation varie en fonction de la voyelle suivante) ; après lesquelles viennent les bigraphes (e.g., PH) ; puis les graphèmes d'une seule lettre associées à deux phonèmes (e.g., X) ; et enfin les graphèmes composés d'une seule lettre associés à un seul phonème (e.g., É). Les règles de CGP suivent les associations statistiques les plus fréquentes⁷,

⁷ Comme Pritchard et al. (2012) le soulignent, ceci signifie que le Système de règles grapho-phonémiques du DRC ne tient compte *que* de la prononciation la plus fréquente associée à un graphème (fonctionnement dit « *type count* » par opposition à « *token count* » : d'autres CGP fréquentes (e.g., le graphème anglais OO peut légalement être lu [u:] ou [ʊ]) ne sont donc pas prises en compte.

modulées par des contraintes contextuelles telles qu'évoquées ci-dessus, mais également par des contraintes phonotactiques (e.g., le non-mot TUI ne sera pas décodé en /tyi/ mais en /tqi/) et morphophonémiques (e.g., le non-mot PRUS ne sera pas décodé en /pRys/ mais en /pRy/). Une fois un phonème approprié trouvé, le Système de règles grapho-phonémiques provoque l'activation de l'unité phonémique dans la couche du Système phonémique, complétant celle qu'elle a reçue en provenance du Lexique phonologie de sortie de la voie lexicale. Une deuxième série de cycles (17 dans les simulations) entre la couche d'Unités de la lettre et la couche du Lexique orthographique d'entrée doit avoir lieu avant que le Système de conversion grapho-phonémique ne puisse s'appliquer à la deuxième lettre la plus à gauche du mot. Si la deuxième lettre forme un bigraphe avec la première (e.g. PH), une seule unité phonémique sera activée dans le Système phonémique ; dans le cas contraire, deux unités différentes seront activées. Une nouvelle lettre sera traitée de la même manière tous les 17 cycles par le Système de règles grapho-phonémiques. Ce dernier peut cesser de s'appliquer si la suite de lettres est prononcée via la voie lexicale ; sinon, il traitera chacune des lettres de manière sérielle, l'une après l'autre, jusqu'à arriver à la prononciation du mot oral.

Selon le modèle DRC, le lexique orthographique active donc tous les items emmagasinés présentant des lettres communes avec la cible à partir de la séquence de lettres traitée par la couche des Unités des lettres. Ces items sont activés avec plus ou moins d'intensité, d'abord en fonction de leur similitude, puis de leur fréquence. Du fait de ses afférences avec le Lexique phonologie de sortie et le Système sémantique, la phonologie et le sens des mots présélectionnés par le Lexique orthographique d'entrée sont récupérés. Il convient de préciser que bien que les deux voies (lexicale et non-lexicale) soient impliquées dans la lecture de tous les mots et non-mots, certains types particuliers d'items ne peuvent être lus avec exactitude que si l'une ou l'autre des voies est

entièrement intacte. Plus précisément, les mots irréguliers (i.e., ne respectant pas, au moins partiellement, les règles d'association graphème-phonème édictées plus haut) ne peuvent être lus parfaitement que si la voie lexicale est préservée, c'est-à-dire uniquement si des entrées orthographiques stockées en mémoire sont en mesure de générer une représentation phonologique dans le Lexique phonologique de sortie. Un autre phénomène est constaté pour les non-mots, qui ne peuvent être lus parfaitement que si l'accès au Système de règles grapho-phonémique est préservé. Ainsi, si la voie lexicale est défaillante, la lecture de mots irréguliers mettra en évidence des erreurs de régularisation, c'est-à-dire une prononciation du mot suivant les règles de CGP habituelles, non conforme à la forme phonologique conventionnelle du mot (e.g., FEMME sera lu /fœmə/ et non /fam/ ; profil d'alexie de surface, Coltheart et al., 1983). À l'inverse, si c'est la voie non-lexicale qui est déficitaire, la lecture de non-mots provoquera des erreurs dites de lexicalisation, c'est-à-dire que l'individu remplacera l'item à lire par un vrai mot présentant des similitudes visuelles avec le non-mot cible (e.g., *FIME lu FIRME ; profil d'alexie phonologique, Beauvois & Déruesné, 1979; Temple & Marshall, 1983). De telles atteintes ont pu être correctement simulées via le DRC sur la base de véritables patients alexiques (Coltheart et al., 1996), y compris lorsque les troubles se présentaient à des intensités diverses (Coltheart et al., 2001).

La validité du modèle DRC a été éprouvée par l'intelligence artificielle grâce aux simulations réalisés par ses auteurs, démontrant qu'il était capable de reproduire tous les comportements humains normaux (effets de fréquence et de lexicalité) et pathologiques (alexies de surface et phonologique) observés en tâche de lecture à haute voix et de décision lexicale. Cependant, plusieurs études ont nuancé la validité théorique du modèle DRC. La critique la plus importante concerne son incapacité de simuler le processus d'apprentissage de la lecture, puisqu'il a été modélisé sur le fonctionnement de lecture des

lecteurs experts (*skilled readers*). Il s'agit néanmoins là d'une limite relevée par les auteurs eux-mêmes : « Puisque le modèle DRC n'a pas été développé via un algorithme d'apprentissage de quelque sorte que ce soit (...), il n'a rien à dire sur les processus réels de l'apprentissage de la lecture. » (Coltheart et al., 2001, p. 246), et qui a été adressée dans une refonte récente du modèle (modèle ST-DRC ; Pritchard et al., 2018) que nous traiterons dans la partie 1.3.3. Un autre type de critique, reposant sur une base plus empirique, dénonce le fait que les comportements de lecture obtenus par le modèle DRC s'éloignent parfois grandement du comportement des lecteurs humains. Par exemple, partant du constat que les lecteurs humains experts prononcent souvent de manière irrégulière des non-mots (Andrews & Scarratt, 1998), supposément en se basant sur la prononciation de vrais mots irréguliers (e.g., en anglais, le non-mot *BALF est lu comme le mot irrégulier CALF), Pritchard et al. (2012) ont comparé la lecture de non-mots d'étudiants à celle des modèles DRC et CDP+ (*Connectionist Dual Process Plus*⁸ ; Perry et al., 2007). Bien que le modèle DRC ait obtenu un pourcentage d'accord avec les réponses les plus fréquemment fournies par les étudiants (73.5%) plus important que le CDP+ (12.1%), le modèle DRC, contrairement aux sujets humains et au CDP+, n'a produit aucune erreur de lexicalisation, suggérant qu'il existe une différence importante dans la lecture de non-mots dans le modèle DRC par rapport à la lecture humaine de non-mots. Certains auteurs ont également émis la critique que le DRC ne parvient pas à modéliser correctement les effets d'amorçage phonologique, du fait de la lenteur de la voie non-lexicale (à l'inverse du modèle BIAM ; Diependaele et al., 2010). Une dernière limite du modèle DRC à évoquer concerne l'absence de prise en compte du contexte de liste lors des simulations, qui chez les sujets humains, peut considérablement modifier

⁸ Le modèle CDP+ est élaboré sur la base du modèle DRC mais inclut un algorithme d'apprentissage à sa voie non-lexicale, lui permettant potentiellement de simuler avec plus de réalisme le comportement de lecture des humains face à des non-mots.

leurs performances, notamment concernant les effets de fréquence et de lexicalité (pour une revue, voir Traficante & Burani, 2014), selon que les mots et non-mots sont présentés isolément (contexte pur) ou de manière conjointe (contexte mixte). L'hypothèse sous-jacente (i.e., *route deemphasis hypothesis*) suppose que par rapport à un contexte pur, un contexte mixte privilégiant la présence de mots irréguliers « surstimule » la voie lexicale, conduisant à la commission d'erreurs de lexicalisation sur les non-mots ou à des temps de lecture plus élevés. Un phénomène similaire serait observé sur les mots irréguliers (commission d'erreurs de régularisation et temps de lecture rallongés) lorsqu'ils sont proposés dans un contexte mixte privilégiant la présence de non-mots, du fait d'une « sursimulation » de la voie non-lexicale.

Une limite plus générale, incluant le modèle DRC mais également la plupart des modèles cognitifs de la lecture, concerne la prédominance d'un point de vue anglocentré sur les processus de lecture en général. Selon Share (2008, 2021), l'anglais constitue un cas exceptionnel d'irrégularité parmi les langues alphabétiques, mais représente pourtant la principale – voire la seule – langue considérée lors de l'élaboration des plus influentes théories et modèles de lecture. Les conclusions tirées des recherches basées sur ces derniers pourraient donc être, en réalité, peu généralisables aux autres langues, dont le français. Le modèle DRC a néanmoins été testé pour la langue française dans l'étude de Ziegler et al. (2003). Cette étude met en évidence une différence majeure entre les processus de lecture anglophones et francophones, en l'absence d'un effet de régularité. Ce dernier est défini par un avantage de lecture pour les mots réguliers par rapport aux mots irréguliers, mais uniquement lorsque ces mots sont de faible fréquence, cet effet disparaissant ou s'atténuant lorsque l'on compare des mots réguliers et irréguliers de fréquence élevée (Paap & Noel, 1991), et s'intensifiant lorsque la violation des CGP habituelles survient en début de mot (Coltheart & Rastle, 1999). Cet effet est constaté

pour la langue anglaise et est bien retrouvé lors de la comparaison des mots réguliers peu fréquents versus mots irréguliers peu fréquents dans les simulations de Coltheart et al. (2001). Il s'expliquerait par une incongruence des informations reçues par le Système phonémique en provenance d'une part de la voie lexicale et d'autre part de la voie non-lexicale. En effet, les mots irréguliers peu fréquents bénéficieraient de moins d'activations de la part du Lexique orthographique d'entrée que les mots irréguliers fréquents, laissant davantage de temps à la voie non-lexicale pour opérer et générer des représentations phonologiques contradictoires. Pour le français, néanmoins, l'étude de Ziegler et al. (2003) a montré que chez des lecteurs francophones, l'effet de régularité se traduisait par une lecture plus lente des mots irréguliers par rapport aux réguliers, quelle que soit leur fréquence, répliquant les résultats précédemment obtenus par Content (1991). Ce comportement humain a pu être reproduit dans le modèle en accélérant la mise en œuvre de la voie non-lexicale, ou plus exactement, en supprimant le délai empêchant les deux voies de commencer en même temps à décoder le mot proposé (dix cycles initiaux entre la couche des Unités des lettres et la couche du Lexique orthographique d'entrée).

L'interprétation des auteurs était ainsi qu'en français, les CGP étant plus régulières qu'en anglais, les individus francophones se reposent davantage sur la voie non-lexicale que les individus anglophones. Cette dépendance induit en français un effet de régularité sans interaction avec la fréquence lorsque les informations portées par la voie non-lexicale sont incongruentes avec celles portées par la voie lexicale (i.e., lorsque des mots irréguliers sont proposés).

1.2.2 Le lexique orthographique dans le processus d'écriture experte : le modèle

RET

Si le français et l'anglais diffèrent au niveau de la régularité de leurs CGP, le taux d'irrégularité de leurs CPG, en revanche, paraît très similaire. En effet, comme le soulignent Bonin et al. (2008) :

(...) à l'aide d'une simulation sur ordinateur, Véronis (1988) a estimé que seulement 52,7% des formes graphiques du français pouvaient être prédites à partir de leur forme phonologique (une simulation avait aussi abouti à un score similaire en anglais, Hanna, Hanna, Hodges & Rudolf, 1966). Ainsi, un mot sur deux contient une particularité orthographique qui ne peut donc être levée sur la seule base des correspondances entre unités sonores et unités orthographiques.

Certaines complexités orthographiques apparaissent également typiques à la fois de l'anglais et du français, telles que la présence de lettres muettes déductibles par la dérivation morphologique (e.g., en anglais, DAMN / DAMNATION⁹ ; en français, CHANT / CHANTEUR ; voir partie 3.2.2) ou d'homophones lexicaux (e.g., en anglais, RITE / RIGHT / WRITE ; en français, MER / MÈRE / MAIRE). Par ailleurs, la plupart des modèles de production orthographique élaborés se basent sur des observations aphasiologiques réalisées chez des patients anglophones (Graham et al., 1997) aussi bien que francophones (Beauvois & Dérouesné, 1979). Étant donnés ces éléments, nous considérerons donc que le français et l'anglais possèdent des systèmes orthographiques comparables, justifiant de les étudier conjointement.

Les modèles d'écriture experte (*skilled writing*) sont moins nombreux et moins nettement distingués que les modèles de lecture. Nous avons identifié trois principaux modèles, tous ayant en commun une architecture à double voie : le modèle d'Ellis (1982,

⁹ Les exemples en anglais sont issus d'Ellis (1982).

1988), le modèle de Houghton et Zorzi (1998, 2003) et le modèle de Rapp et collaborateurs (McCloskey et al., 2006; Rapp et al., 2002; Tainturier & Rapp, 2001). Ces trois modèles, bien qu'inscrivant l'écriture dans des cadres théoriques relativement différents, ont également la particularité de s'être enrichis les uns des autres. Nous avons choisi de nous concentrer principalement sur celui de Rapp et collaborateurs car il est celui qui, à notre connaissance, est le plus étayé d'études comportementales et le plus intégratif, incorporant un nombre important de concepts des modèles précédents.

Le modèle à double voie de Rapp et al. (2002 ; ci-après dénommé RET), comme la plupart des modèles d'écriture, a été élaboré sur la base de données comportementales (écriture experte et agraphie¹⁰) et concerne une situation d'écriture sous dictée. Son architecture comprend six couches (voir Figure 2), dont deux sont communes à la voie lexicale et sublexicale : la couche des Phonèmes et la couche du Buffer graphémique. La voie sublexicale comporte une couche de Conversion phonème-graphème située entre les deux couches communes précédemment citées. La voie lexicale est constituée de trois couches spécifiques : la couche du Lexique phonologique d'entrée, constituée de lexèmes phonologiques ; la couche du Niveau sémantique ; et la couche du Lexique orthographique de sortie, constituée de lexèmes orthographiques. Les connexions entre les couches sont toutes « *fed-forward* » (i.e., présence uniquement de connexions excitatrices et dans un seul sens), à l'exception d'une connexion bidirectionnelle entre le Buffer graphémique et le Lexique orthographique de sortie, constituant un feedback graphème-lexème (*grapheme-to-lexeme feedback* ; McCloskey et al., 2006). Contrairement au modèle de Houghton et Zorzi (2003), le modèle RET suppose que l'unité de production orthographique est le graphème et non les sous-composants de la syllabe (attaque, noyau, coda). Cependant, les deux modèles s'accordent à considérer que les voies lexicale et

¹⁰ Ici, le terme « agraphie » ne se réfère pas uniquement à des difficultés concernant le graphisme mais au processus d'écriture dans son ensemble.

sublexicale sont activées en parallèle lors du processus d'écriture. Après une activation du Buffer graphémique issue des deux voies, les représentations abstraites des graphèmes sont, selon la tâche, converties soit en noms de lettres (tâche d'épellation), soit en formes de lettres (tâche d'écriture manuscrite)¹¹.

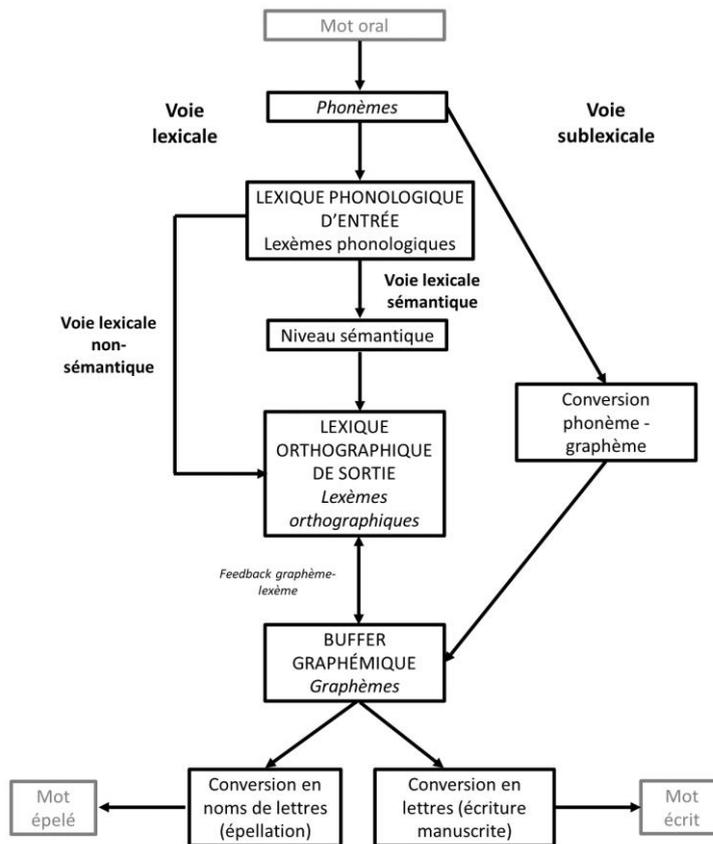


Figure 2. Modèle à Double Voie de Rapp, Epstein et Tainturier, 2002, adapté et traduit

Selon le modèle RET, le processus d'écriture commence avec l'analyse phonémique du mot (niveau des Phonèmes), au terme de laquelle l'embranchement entre les deux voies se forme. Du côté de la voie sublexicale, la Conversion phonème-graphème sélectionne des candidats appropriés pour chaque phonème analysé. La suite de graphèmes adoptées par la voie sublexicale est ainsi activée au niveau du Buffer graphémique. Au niveau de la voie lexicale, en réponse aux phonèmes extraits de

¹¹ En l'état actuel, le modèle n'intègre pas l'écriture au clavier.

l'analyse phonémique, un lexème phonologique semblable en tout (i.e., forme phonologique préalablement encodée) ou partie (i.e., forme phonologique inconnue) est activé dans le Lexique phonologique d'entrée. Cette activation génère au Niveau sémantique¹² l'activation de représentations conceptuelles associées au lexème phonologique, indispensables en cas de présence d'homophones hétérographes (e.g., /mɛR/ qui peut s'orthographier MER, MAIRE ou MÈRE). L'activation du lexème orthographique correspondant à la représentation sémantique pertinente se propage ensuite au niveau du Buffer graphémique, activant les graphèmes correspondants. Ce dernier additionne les activations graphémiques supportées par les voies lexicales et sublexicales ; sorte de « mémoire de travail orthographique », il code l'identité des graphèmes, leur nombre (en cas de lettres doubles) et leur ordre (Rapp et al., 2002). Le feedback graphème-lexème opère ensuite des cycles entre les activations du Lexique orthographique de sortie et du Buffer graphémique, renforçant l'activation du lexème orthographique porté par la voie lexicale (McCloskey et al., 2006), et expliquant l'activation dominante de cette dernière dans le cas de la production orthographique de vrais mots (Folk et al., 2002). Dans le cas de la production de non-mots, la participation de la voie lexicale à la transcription des non-mots serait également impliquée, sous réserve que le non-mot puisse activer un lexème phonologique au niveau du Lexique phonologique d'entrée (i.e., qu'une telle représentation soit préalablement stockée et suffisamment ressemblante) ; cette dernière susciterait une activation (quoique présumément faible puisque n'étant pas identique au non-mot énoncé) du ou des lexèmes

¹² Il est à noter qu'une voie lexicale non-sémantique, ne sollicitant pas le Niveau sémantique (et ne permettant donc pas la distinction entre différents homophones lexicaux), a été récemment intégrée au modèle RET (Hepner et al., 2017), afin de tenir compte des cas d'agraphies asémantiques décrits dans la littérature (Bub et al., 1985; Funnell, 1983; Sartori et al., 1987; Schwartz et al., 1980).

orthographiques associés, activant *in fine* les unités graphémiques correspondants dans le buffer graphémique (Rapp et al., 2002).

L'hypothèse d'une participation simultanée des voies lexicales et sublexicales a été émise en raison de phénomènes spécifiques à la situation d'écriture. Typiquement, les productions témoignant de représentations lexicales partielles (*partial lexical knowledge* ; Rapp et al., 2002; Tainturier & Rapp, 2001) mettent en évidence chez un individu le recours à un lexème orthographique pour une partie du mot seulement (e.g., CHRYSANTHÈME transcrit CHRYSANTÈME), signifiant la contribution pour la transcription d'un même mot à la fois de la voie lexicale (transcription de /kRi/ et de /zã/) et de la voie sublexicale (transcription de /tem/). Ces erreurs phonologiquement plausibles (Rapp et al., 2002), attribuées à des « irrptions » de la voie sublexicale, seraient surtout constatées lors d'incongruence entre les lexèmes orthographiques et les CPG plus fréquentes (Bonin et al., 2008), non seulement dans le cas de mots irréguliers mais aussi pour certains mots réguliers en lecture. Ces derniers peuvent en effet présenter des CGP consistantes, mais qui correspondent à des CPG inconsistantes, susceptibles d'interférer avec la production de la bonne orthographe du mot : par exemple, pour le mot CIRQUE, les CGP sont sans ambiguïté ; néanmoins, ce mot possède des phonèmes pouvant être transcrits par plusieurs graphèmes différents (e.g., /s/ peut être transcrit S au lieu de C ; /k/ être transcrit C ou K au lieu de QUE). De même, une action concomitante des deux voies est observable chez les scripteurs experts amenés à écrire des non-mots sous dictée. Ces derniers semblent recourir spontanément à des *patterns* orthographiques rencontrés dans des vrais mots (e.g., /no/ écrit NEAU comme dans BEAU, PEAU, etc. ; Campbell, 1983). Dans l'écriture experte, la conversion phono-graphémique semble soumise à des règles plus nombreuses que celles édictées dans le sens graphème-phonème. En plus des règles relativement stables, telles que les CPG extrêmement consistantes (e.g., le son /a/ est

toujours transcrit A, sauf dans les mots irréguliers FEMME, SOLENNEL, MOELLE, POÊLE, et dans les adverbes dérivés d'adjectifs terminant en ENT comme ÉVIDENT/ÉVIDEMMENT) et les graphies contextuelles (e.g., G ne peut pas transcrire /ʒ/ lorsqu'il est positionné avant les voyelles A, O et U, à moins de le faire suivre d'un E comme dans PIGEON), d'autres interrogations se posent. En effet, lorsqu'ils sont en situation de dictée de non-mots, les individus choisissent-ils le graphème le plus fréquemment associé au phonème cible (e.g., /fiz/ écrit *FISE car /f/ est plus souvent transcrit F en français que PH), ou celui possédant le plus haut degré de consistance graphème-phonème (e.g., /kyf/ écrit *KUFE car K ne peut jamais se lire autrement que /k/ en français, contrairement à C) ? Tiennent-ils compte du nombre de voisins orthographiques (e.g., /zas/ écrit *ZASSE en référence à PASSE, TASSE, NASSE, LASSE, etc. plutôt que *ZACE en référence à FACE, LACE et RACE) ou de la consistance rimique (e.g., /deʒ/ écrit *DEIGE comme dans BEIGE et NEIGE) ? Il est intéressant de souligner que ces différentes potentielles influences sont à la fois d'ordre lexical (e.g., nombre de voisins orthographiques du non-mot proposé) et sublexical (e.g., fréquence d'association phonème-graphème). Les résultats des différentes études menées par Perry et Ziegler (2004) ont mis en évidence que le premier facteur déterminant l'orthographe des non-mots chez les scripteurs experts était la fréquence d'association phonème-graphème, suivie par la consistance rimique (ces deux facteurs pouvant interagir entre eux). Concernant la fréquence d'association phonème-graphème, cette dernière apparaît elle-même modulée par plusieurs facteurs (Perry et al., 2002) : la position dans le mot (e.g., /f/ écrit F en position médiane comme dans *NUFT mais FF en position finale comme dans *NUFF), la présence de certaines consonnes (e.g., /u:/ écrit UE après un /j/ mais OO dans les autres cas), le caractère court ou long de la voyelle (e.g., /neif/ écrit

*NAFE et non *NAFFE ; /nʌf/ écrit *NUFF et non *NUF) et le statut morphologique (e.g., /z/ écrit S et non Z dans *NUGS comme pour marquer le pluriel).

Bien que de nombreuses études comportementales menées chez des sujets sains et agraphiques étayent les différentes composantes du modèle RET, il n'a pas bénéficié, contrairement au modèle de Houghton et Zorzi (2003), de simulations computationnelles. Il n'est donc pas possible de savoir si son fonctionnement reproduirait fidèlement le comportement de scripteurs humains. Par ailleurs, contrairement au modèle développé par Ellis (1982, 1988), le modèle RET constitue davantage un modèle d'orthographe qu'un modèle d'écriture, puisque les processus périphériques (i.e., la réalisation motrice des graphèmes, qu'elle implique un outil scripteur ou un clavier) ne sont pas développés au sein du modèle. Enfin, en dépit du fait que les auteurs soulignent les ressemblances entre les processus mis en œuvre lors de la lecture et de l'écriture (Tainturier & Rapp, 2001), elles ne statuent pas, tout comme Coltheart et al. (2001), sur la question de modules réellement communs aux deux activités. Plus précisément, la question d'un même lexique orthographique duquel serait récupérées les représentations impliquées dans la lecture et dans l'écriture est abordée mais n'est pas intégrée à l'architecture du modèle proposée. Pourtant, cette hypothèse d'un lexique orthographique commun, à la base de l'apprentissage orthographique, est largement admise dans les modèles développementaux de la lecture / écriture, que nous allons décrire dans les paragraphes suivants.

1.3 Le développement du lexique orthographique

Nous aborderons deux conceptions de l'apprentissage orthographique par des apprentis lecteurs/scripteurs, illustrant chacune l'indissociabilité et la constante interaction des processus de lecture et d'écriture dans le développement du langage écrit : le modèle en six stades de Frith (1985) suggérant une transition de l'utilisation de la voie non-lexicale/sublexicale vers la voie lexicale, et l'hypothèse de l'auto-apprentissage

développée initialement par Firth (1972) puis par Share et collaborateurs (Jorm & Share, 1983; Share, 1995, 1999), postulant la non-linéarité du recours à ces deux voies. Enfin, nous aborderons comment cette dernière peut être intégrée au modèle DRC (Pritchard et al., 2018).

1.3.1 Le modèle en six stades de Frith (1985)

Le modèle d'acquisition de la lecture et de l'écriture de Frith (1985 ; Figure 3) stipule que cette acquisition est marquée par une alternance de trois types de stratégies : logographique, alphabétique et orthographique (voir infra), inspirées et adaptées des travaux de (Marsh et al., 1977). L'utilisation de ces stratégies se succède dans un ordre strict, et chaque stratégie se base sur la précédente. Le modèle de Frith suppose également que les processus de lecture et d'écriture se développent de manière concomitante mais « déphasée » : à la première étape de chaque stade (a), la nature des processus de lecture et d'écriture relève d'une stratégie différente, avant de se resynchroniser dans une seconde étape (b), dans laquelle les stratégies redeviennent similaires. Chaque stade est ainsi marqué par la domination de l'un des processus (lecture ou écriture) qui initie le changement de stratégie, tandis que l'autre stagne dans la stratégie précédente. La stratégie des stades 1 et 3 est donc, en première étape, portée par la lecture. À l'inverse, la stratégie du stade 2 est en première étape mise en œuvre par le processus d'écriture. Le degré de maîtrise de chaque stratégie est signifié par un nombre allant de 1 à 3 ; l'auteure suggère que le processus dominant de chaque stade (lecture ou écriture) doit d'abord obtenir un degré de maîtrise équivalent au niveau 2 avant que l'autre processus ne puisse être investi de cette stratégie (i.e, l'enfant doit atteindre suffisamment de maîtrise en lecture logographique avant d'être capable d'écrire des mots en utilisant cette stratégie).

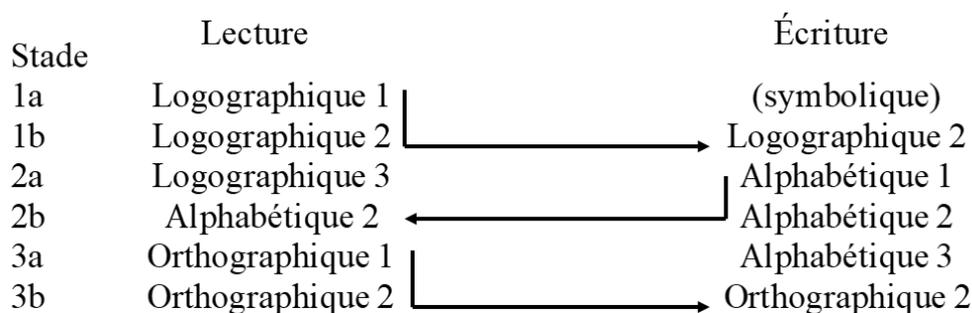


Figure 3. Modèle en six stades de lecture et d'écriture, adapté et traduit de Frith (1985)

Le stade 1 renvoie à un processus de lecture selon une stratégie logographique. Cette dernière consiste pour l'enfant à lire un nombre restreint de mots (qui peuvent davantage être considérés comme des « images graphiques » à ce stade), en s'appuyant sur certaines propriétés visuelles saillantes (e.g., le X du prénom ALEX), dont l'apprentissage est facilité par des indices issus du contexte dans lequel le mot est rencontré. Elle s'appuie sur les facultés symboliques du jeune enfant, grâce auxquelles il s'approprie le concept de mot. Aux premiers niveaux de la stratégie logographique, il est supposé que l'enfant ne tient pas compte de l'ordre ni de l'identité des lettres, mis à part les indices saillants repérés, qu'il sera par la suite capable de reproduire lorsqu'il souhaitera transcrire le mot (e.g., l'enfant écrit X pour signifier ALEX). L'auteure suppose que plus l'enfant est capable de se fier à des indices intrinsèques au mot plutôt qu'à des indices contextuels issus de l'épisode dans lequel le mot a été rencontré, plus il est capable de le reconnaître en dépit de modifications situationnelles. Il est également présumé qu'un stimulus qui présenterait les indices saillants d'un mot pourrait être faussement reconnu comme ce dernier (e.g., MAX sera lu ALEX car il contient également X). Même si un nombre parfois important d'items peut potentiellement faire partie du vocabulaire visuel de l'enfant (à noter toutefois d'importantes variabilités inter-individuelles ; Francis, 1982), ce dernier est incapable de les mettre à profit pour décoder des mots inconnus (e.g., il ne peut pas lire ALEXA à partir d'ALEX). Bien que le nombre d'items du vocabulaire visuel soit

le plus souvent limité, l'auteure postule que ce n'est pas ce qui motive le changement vers la stratégie alphabétique. Selon elle, il s'agit plutôt l'acquisition de la conscience phonémique (Gleitman & Rozin, 1977; Liberman et al., 1977), i.e., « l'accès conscient au niveau phonémique du flot de la parole ainsi qu'une certaine faculté de manipuler cognitivement les représentations à ce niveau » (Stanovich, 1986, p. 362), ajoutée à la pratique de l'écriture (e.g., copie du prénom) qui engendre une première appréhension de la séquence, essentielle au stade 2.

L'écriture selon une stratégie alphabétique domine donc le stade 2, et est définie comme l'application systématique des CPG, permise par une dose minimale et indispensable d'enseignement explicite (Christensen & Bowey, 2005). La lecture, quant à elle, est encore marquée par une stratégie logographique à la première étape du stade 2 (voir Figure 3). Au commencement de la stratégie alphabétique, la mise en œuvre des CPG est rudimentaire : chaque phonème est simplement converti en graphème dans un rapport d'un pour un (e.g., /fløR/ transcrit FLER, /fRit/ transcrit FRIT, /møsjø/ transcrit MESIE), ce qui est problématique dans une langue opaque mais relativement fonctionnel dans une langue très transparente. Du fait de la non-convergence des stratégies à ce stade, l'enfant ne peut souvent pas relire ce qu'il a écrit (Read, 1971). Quand la stratégie alphabétique se développe et ses éléments plus complexes (graphies contextuelles notamment) sont intégrés, un nombre important de mots réguliers aux CPG très consistantes (i.e., FRITE, FLEUR) peuvent être écrits sans faute. Les mots aux GPG inconsistantes sont écrits avec les graphèmes les plus fréquents (e.g., /bẽ/ transcrit BIN), et les mots irréguliers sont régularisés (e.g., /møsjø/ transcrit MEUSSIEU), donnant lieu à des erreurs phonologiquement plausibles. À l'étape 2b, la lecture par CGP est possible, mais, selon l'auteure, uniquement pour celles que l'enfant est également capable d'écrire (Bradley & Bryant, 1979; Bryant & Bradley, 1980). Ce principe lui permet donc de

décoder des mots nouveaux, même si sa lecture n'est pas systématiquement correcte (e.g., BANC lu /bãk/).

Après que l'enfant ait été capable d'appliquer la plupart des règles de conversion grapho-phonémique en situation de lecture, une nouvelle stratégie émerge, marquant l'entrée dans le stade 3 : la stratégie orthographique. Cette dernière se définit comme l'utilisation d'unités plus larges que le graphème, ne nécessitant pas d'analyse phonologique. Par exemple, les graphèmes de MAMAN ne sont pas lus isolément, l'un après l'autre (M/A/M/AN), mais par groupements de syllabes fréquemment rencontrées (MA/MAN), procédé fluidifiant nettement la lecture. L'auteure postule que l'enfant se constitue, de par ses expériences répétées face à l'écrit, une sorte de « syllabaire », qui « idéalement coïncide avec des morphèmes » (Frith, 1985), i.e. plus petite unité de sens d'un mot (voir partie 3.1.1). Ce syllabaire permettrait de lire virtuellement tous les mots. À la première étape du stade 3, l'enfant a néanmoins encore recours à la stratégie alphabétique pour écrire. Une fois que les représentations orthographiques sont suffisamment stables et définies (étape 3b), il peut s'en servir également pour orthographier des parties de mots. L'auteure note qu'à ce moment du développement, les productions orthographiques de l'enfant seront marquées par des représentations lexicales partielles (e.g., /møsjø/ transcrit MONSIEU, /fam/ transcrit FAMME) pouvant parfois être non-phonologiquement plausibles, ce qui peut sembler constituer une régression, puisque l'enfant préserve d'abord les unités orthographiques ou morphémiques avant les unités phonologiques. Des mots inconnus sont ainsi transcrits par analogie à des mots familiers (e.g., /boko/ transcrit *BEAUCAUP au lieu de BOCAUX), de même que les non-mots (e.g., /bRafo/ transcrit *BRASCHAUD par analogie avec BRAS et CHAUD). En d'autres termes, le stade 3 marque le début de l'apprentissage orthographique et de la constitution chez l'enfant d'un lexique orthographique. L'auteure souligne le probable rôle de la

conscience morphologique (i.e., faculté d'identifier et de manipuler les morphèmes de la langue, Carlisle, 1995 ; voir partie 3.1.2) dans l'avènement de ce stade. Elle note que ce dernier ne constitue *a priori* pas l'ultime stade de développement de la lecture (i.e., présence supposée d'un quatrième stade correspondant à l'acquisition du système d'écriture dans son ensemble) mais que l'arrêt d'un individu au stade 3 ne crée pas de comportement de lecture pathologique, contrairement aux deux précédents. Frith (1985) suppose en effet que les troubles de la lecture (en l'occurrence, les dyslexies développementales) peuvent être expliqués par une interruption du développement soit au stade 1 (lecture uniquement logographique ; profil de dyslexie phonologique), soit au stade 2 (lecture logographique et alphabétique, mais sans accès à des unités plus larges ; profil de dyslexie de surface). Une absence d'acquisition de la stratégie logographique est également envisagée (i.e., aucune stratégie de lecture), mais ne correspondrait selon l'auteure qu'à des cas de déficience intellectuelle sévère, et non de dyslexie à proprement parler.

Les premières observations de comportements des apprentis-lecteurs semblent confirmer l'existence de stratégies de lecture logographique (Ehri & Wilce, 1985; Neumann et al., 2012; Torrey, 1979), alphabétique (Condry et al., 1979; Pick et al., 1978) et orthographique (McCaughy et al., 1980; Snowling & Frith, 1981), ainsi que la forte interaction entre les processus de lecture et d'écriture (Ehri & Wilce, 1987; Uhry & Shepherd, 1993). De même, Frith (1985) argumente que les troubles de la lecture semblent pouvoir être décrits en termes d'absence de développement de l'une ou l'autre de ces stratégies, puisque certains enfants semblent capables d'appliquer une stratégie alphabétique mais ne jamais réussir à atteindre l'efficacité permise par une stratégie orthographique, tandis que d'autres n'accèdent jamais à une stratégie alphabétique, ne reconnaissant que quelques mots via la stratégie logographique. Néanmoins, certains

principes inhérents au modèle de Frith sont contestés, tel que l'hypothèse que chaque stratégie se baserait sur les compétences développées au sein de la stratégie précédente (Stuart & Coltheart, 1988). Par exemple, contrairement à ce que le modèle infère, le passage à une stratégie alphabétique ne paraît pas dépendre de l'acquisition d'une stratégie logographique (cette dernière n'étant par ailleurs pas corrélée au niveau de lecture ultérieur ; Budoff & Quinlan, 1964; Firth, 1972; Jorm, 1977, 1981; Masonheimer et al., 1984; Rozin et al., 1971) mais plutôt d'un apprentissage explicite des CPG en lien avec les facultés de conscience phonémique (Scott & Ehri, 1990). La conception de Frith paraît également suggérer qu'un individu ayant commencé à lire via la stratégie orthographique abandonnerait immédiatement et entièrement la stratégie alphabétique pour l'ensemble des mots rencontrés, puisque ces deux stratégies ne peuvent pas, en théorie, co-exister pour la lecture. Ce point de vue implique qu'il existerait une « borne » dans le développement à partir de laquelle cette évolution opérerait. Or, plusieurs travaux (Barron & Baron, 1977; Condry et al., 1979; Kimura & Bryant, 1983) mettent en évidence une lecture rapide et non-alphabétique de quelques mots de très haute fréquence dès la première ou la deuxième année de primaire. En résumé, les interrogations suscitées par la proposition de Frith ne concernent pas tant la nature ou l'existence des stratégies décrites que la pertinence du principe de fonctionnement en stades appliqué à l'acquisition de la lecture (Content & Zesiger, 2005).

1.3.2 L'hypothèse de l'auto-apprentissage

Il s'agit du principal point d'achoppement de l'hypothèse de l'auto-apprentissage introduite par Firth (1972) et développée par Share et collaborateurs (Jorm & Share, 1983; Share, 1999, 1995) : l'utilisation des stratégies de lecture serait *item-dépendante* et non *stade-dépendante*. La lecture via un recodage phonologique (i.e., récupération de la forme phonologique du mot via ses composants sublexicaux ; correspond à la stratégie

alphabétique) ou la reconnaissance du mot sans recodage phonologique (correspond à la stratégie orthographique)¹³ serait fonction de la familiarité du mot à décoder, et donc directement de sa fréquence : les mots bénéficiant d'une fréquence élevée auraient toutes les chances d'avoir été précédemment rencontrés par l'apprenti-lecteur, qui en aurait déjà extrait des représentations orthographiques, lui permettant de les reconnaître visuellement. En revanche, les mots peu fréquents ayant moins de probabilité d'avoir déjà été rencontrés seraient, pour la plupart, lus via un recodage phonologique. Share (1995) ajoute que même lorsqu'il est rudimentaire, ce recodage constitue la pierre angulaire de l'apprentissage orthographique et ainsi du développement de la lecture/écriture : chaque expérience réussie d'identification d'un nouveau mot par recodage phonologique constituerait une « opportunité d'auto-apprentissage » pouvant créer une entrée dans le lexique orthographique. À chaque nouvelle rencontre avec ce mot, la représentation orthographique associée serait ainsi renforcée, jusqu'à ce que le mot puisse être lu sans passer par un recodage phonologique, et suffisamment spécifié pour être écrit sans erreur. Dans le même temps, d'autres informations sur le système d'écriture en général seraient extraites de cette rencontre, permettant à l'individu de sophistication ses règles de CGP. Par exemple, un individu qui rencontrerait le mot ROSE à l'écrit pour la première fois pourrait, à tort, faire correspondre le graphème S au phonème /s/, ignorant les règles contextuelles associées au S (i.e., lorsqu'il est compris entre deux voyelles, S se lit /z/) ; néanmoins, grâce au contexte dans lequel ce mot est rencontré, l'individu déduira qu'il ne s'agit pas du mot ROSSE mais bien du mot ROSE, conduisant à une opportunité d'auto-apprentissage. Il extraira alors d'une part une trace de l'orthographe de ce mot dans son

¹³ La stratégie logographique n'est pas considérée par Share (1995) comme une stratégie de lecture en tant que telle, mais plutôt comme de la « pré-lecture ».

lexique orthographique et d'autre part les prémices de la règle de prononciation du S entre deux voyelles.

Suivant cette logique, un auto-apprentissage réussi dépendra également des capacités de conscience phonémique de l'apprenti-lecteur. En effet, comme le souligne Share (1995), un sujet qui ne posséderait pas une conscience phonémique, même rudimentaire, serait incapable de se corriger lorsqu'il décoderait /Rɔs/ au lieu de /Roz/, puisque les similitudes dans la structure phonologique de ces deux mots ne seraient pas perçues (Stuart, 1990). Comme évoqué précédemment, le contexte jouerait également un rôle-clef dans la résolution des ambiguïtés de décodage (e.g., un jardin de ROSSES versus un jardin de ROSES ; Goswami, 1990; Nicholson & Hill, 1985; Pring & Snowling, 1986). Ce rôle serait d'autant plus essentiel pour les mots irréguliers ou possédant des CGP inconsistantes (Gough & Walsh, 1991; Wang et al., 2011), en ce qu'il permettrait de compenser un décodage partiel (Share, 1995). Par exemple, un mot irrégulier tel que FEMME possède néanmoins des CGP pour la plupart consistantes, i.e., F, MM et E ; l'appui sur le contexte permet ainsi de déduire que la CGP inconsistante doit être lue /a/ et non pas /œ/. Share (1995) note par ailleurs qu'en anglais, ce sont principalement les voyelles qui sont porteuses des inconsistantes résultant en des mots irréguliers (Bryson & Werker, 1989; Zinna et al., 1986), tandis que les consonnes restent le plus souvent consistantes ; or, ce sont les consonnes qui sont les plus décisives dans l'identification d'un mot (Lee et al., 2001). Ainsi que le remarquent Pritchard et al. (2018), d'autres facteurs que le contexte pourraient également être suggérés comme conditionnant la réussite du décodage, tels que la richesse du vocabulaire de l'individu, i.e., le fait de préalablement posséder une entrée dans son lexique phonologique et/ou dans son système sémantique pour le mot FEMME (Tunmer & Chapman, 2012).

Dans une étude publiée en 1999, Share a testé l'hypothèse d'apprentissage orthographique par auto-apprentissage dans une série d'expériences menées en Hébreu voyellisé (système de transcription considéré comme transparent ; Ravid, 1996; Schiff, 2003) chez des enfants scolarisés en deuxième année de primaire. Dans une première expérience, 40 enfants devaient lire des textes dans lesquels un non-mot apparaissait soit six fois, soit quatre fois. Les évaluations réalisées trois jours plus tard montrent qu'après seulement quatre expositions, les enfants étaient capables au-delà d'une probabilité relevant du hasard, de reconnaître (tâche de choix homophonique), de lire à haute voix, et d'orthographier les non-mots de l'ensemble des textes, démontrant une préférence pour l'orthographe présentée par rapport à une orthographe homophone. Afin de déterminer si ces résultats pouvaient être attribuées à la simple exposition visuelle aux cibles, une deuxième expérience a été réalisée, dans laquelle 20 autres participants de même niveau de classe étaient exposés six fois à des non-mots identiques à ceux de l'expérience 1, cette fois au cours d'une tâche de décision lexicale. Dans cette expérience, la consigne était de déterminer si la suite de lettres présentées était un mot ou non, tout en vocalisant un autre non-mot sans rapport avec les stimuli, dans le but d'empêcher la subvocalisation et ainsi diminuer les capacités de recodage phonologique des participants. Les résultats des évaluations pré-tests montrent une chute significative des performances par rapport à l'expérience 1, suggérant le rôle essentiel du recodage phonologique lors de l'apprentissage orthographique. Une troisième expérience, quasi-identique à l'expérience 2, a ensuite été réalisée afin de déterminer si cette absence d'effet d'exposition visuelle pouvait être attribuée à une durée d'exposition trop brève aux stimuli par rapport à celle de l'expérience 1. Dans cette troisième expérience, au lieu de vocaliser un autre non-mot sans rapport, 20 autres participants de deuxième année de primaire devaient lire à haute voix la cible présentée avant de décider s'il s'agissait d'un mot ou non, afin de s'assurer

qu'ils réalisaient effectivement un décodage phonologique du non-mot présenté. Les résultats obtenus étaient cette fois comparables à ceux de l'expérience 1.

Cette série d'expériences, de même que de nombreuses autres études menées (Bowey & Muller, 2005; Cunningham, 2006; Cunningham et al., 2002; de Jong et al., 2009; de Jong & Share, 2007; Kyte & Johnson, 2006) soutiennent ainsi l'hypothèse que l'apprentissage orthographique est médié par le recodage phonologique, y compris pour la langue française (Bosse et al., 2015). Cette conception du développement de la lecture permet également d'envisager une réelle continuité entre « l'apprentie-lecture » et la lecture experte. En effet, si l'apprentissage orthographique est item-dépendant, il n'y a pas de différence *qualitative* fondamentale (inhérente à la vision stadiste du développement de la lecture) entre ces deux situations, puisque chaque lecteur évolue sur un continuum où seules peuvent être remarquées des différences *quantitatives* au niveau de la maturité des deux voies de lecture, i.e., la maîtrise des règles de CGP pour la voie non-lexicale et la richesse du lexique orthographique pour la voie lexicale (la seconde étant conditionnée par le bon fonctionnement de la première).

1.3.3 Le modèle ST-DRC

Compte tenu de cette importante perméabilité entre l'hypothèse de l'auto-apprentissage et les processus de lecture experte du modèle DRC (Coltheart et al., 2001), Pritchard et collaborateurs (Pritchard et al., 2018, 2016) ont proposé une modification de ce modèle intégrant une possibilité d'auto-apprentissage (ST ou *self-teaching*) : le modèle ST-DRC (Pritchard et al., 2018), ou « modèle d'auto-apprentissage à double voie en cascade de lecture à haute voix et d'apprentissage orthographique ». Afin d'incorporer l'hypothèse de l'auto-apprentissage dans le fonctionnement du modèle de lecture à double

voie, l'architecture du modèle DRC¹⁴ a été reprise et modifiée (voir Figure 4). Outre des changements dans l'appellation des différentes couches, un Système sémantique rudimentaire a été intégré afin de permettre au contexte de jouer son rôle lors du recodage phonologique. Par rapport à la version précédemment présentée du DRC, il inclut également des seuils d'activation à certaines étapes du processus, et les entrées des Lexiques orthographique et phonologique sont considérées comme des nœuds dépendants de ces seuils d'activations.

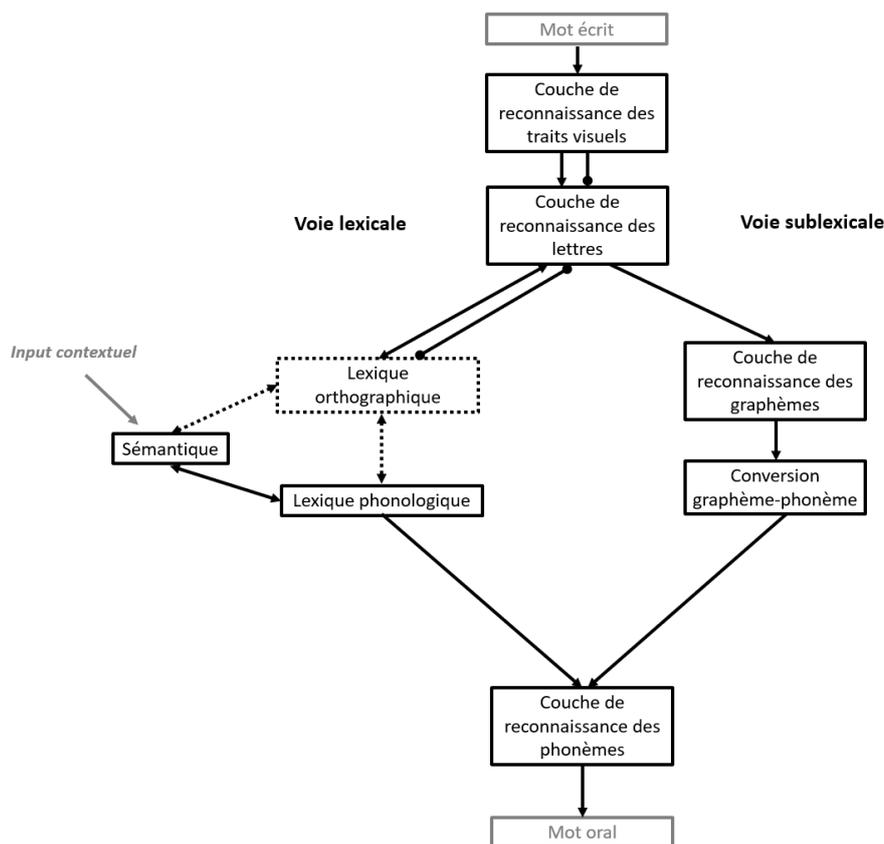


Figure 4. Modèle d'Auto-Apprentissage à Double Voie en Cascade, adapté et traduit de (Pritchard, 2013)

NB : Les flèches triangulaires indiquent une connexion excitatrice, et les flèches rondes une connexion inhibitrice. Les éléments en pointillés renvoient à ce qui doit être appris par le modèle.

¹⁴ Il est à noter que les auteurs utilisent la version 1.2.1 du DRC, présentant quelques différences avec la version décrite dans Coltheart et al. (2001) au niveau du fonctionnement de la voie non-lexicale (Mousikou et al., 2010) ; ces dernières n'étant pas critiques pour la compréhension du propos, nous ne les détaillerons pas.

Le modèle ST-DRC permet de rendre compte de manière computationnelle de deux types d'auto-apprentissage : l'apprentissage orthographique *type-based* (i.e., création d'une toute nouvelle entrée dans le lexique orthographique) et l'apprentissage orthographique *token-based* (i.e., renforcement d'une entrée orthographique déjà existante). L'apprentissage *type-based* a lieu lorsque le mot est rencontré pour la première fois. Puisque le mot n'a jamais été rencontré, aucune activation au niveau du Lexique orthographique ne peut avoir lieu (les éventuels voisins orthographiques pourront bénéficier d'une activation, mais celle-ci est négligeable du fait des liens inhibiteurs entre le Lexique orthographique et les unités des lettres). L'activation des unités phonémiques dans la Couche de reconnaissance des phonèmes (i.e., Système phonémique dans le modèle DRC de 2001) sera donc réalisée via la voie sublexicale. Si le mot fait partie du vocabulaire du sujet et qu'il est régulier, la suite de phonèmes activera (selon un seuil d'activation au paramétrage déterminé par le modèle) le nœud correspondant dans le Lexique phonologique, accomplissant un recodage phonologique réussi. Un nouveau nœud (i.e., une nouvelle entrée) est donc créé dans le Lexique orthographique ; du fait de sa récente création, le modèle considérera ce mot comme ayant une fréquence très faible (valeur de 1), correspondant au fait qu'il n'ait été rencontré qu'une seule fois. L'apprentissage *type-based* d'un mot régulier est illustré dans la Figure 5.

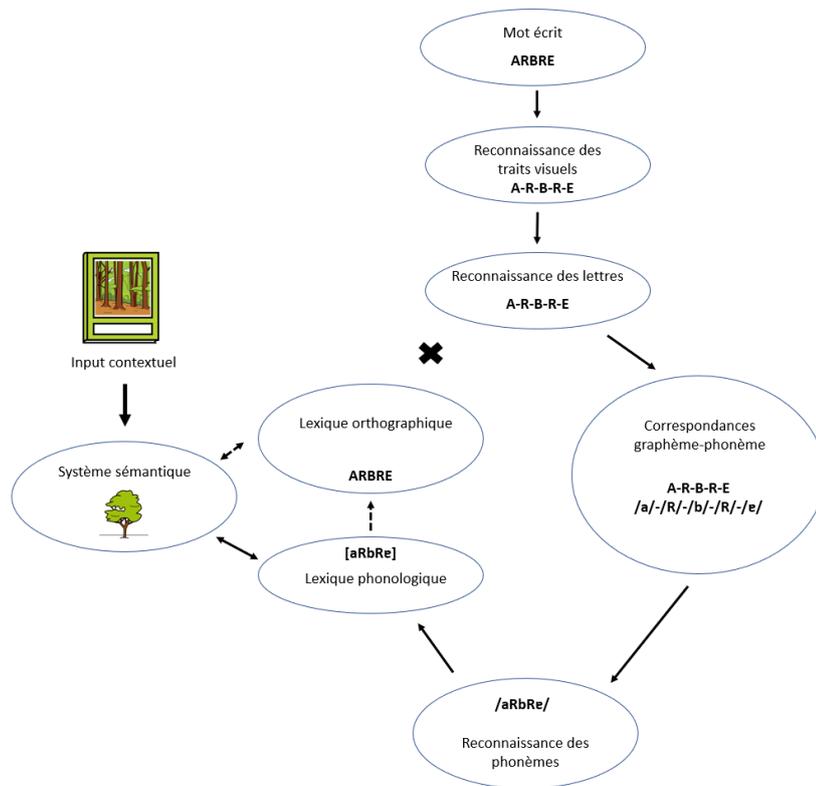


Figure 5. Apprentissage orthographique type-based d'un mot régulier d'après Pritchard et al. (2018)

NB : Les pictogrammes sont issus du site ARASAAC (<https://arasaac.org>) et créés par Sergio Palao ; elles sont la propriété du Gouvernement d'Aragon, Espagne (BY-NC-SA)

L'apprentissage *token-based* a lieu lorsqu'un nœud a été créé dans le Lexique orthographique suite à un apprentissage *type-based* : le mot est alors lu de manière similaire au modèle DRC (i.e., via une double activation en cascade de la voie lexicale et de la voie sublexicale), à la différence que l'acte de lecture provoque, après la production de la prononciation du mot, une modification de la fréquence attribuée au nœud du Lexique orthographique, cette dernière augmentant à chaque nouvelle rencontre¹⁵. Ultiment, après suffisamment d'événements d'auto-apprentissage, le modèle ST-DRC se comporte donc de manière identique au modèle DRC.

¹⁵ Dans le modèle ST-DRC, ce phénomène d'augmentation progressive de la fréquence du mot dans le lexique orthographique est simplifié par l'implémentation d'un coefficient multiplicateur correspondant à la fréquence d'occurrence du mot dans la langue.

Dans le cas de l'apprentissage *type-based* d'un mot irrégulier, la voie sublexicale ne parviendra pas à elle seule à activer l'entrée correspondante dans le Lexique phonologique, du fait de l'incongruence entre la suite de phonèmes de la Couche de reconnaissance des phonèmes et celle contenue dans le Lexique phonologique. C'est alors le Système sémantique qui fournit l'activation nécessaire manquante pour atteindre le seuil d'activation du nœud dans le Lexique phonologique. Toutefois, le nœud correspondant au concept du mot dans le Système sémantique doit avoir lui-même été préalablement activé par le contexte. L'apprentissage orthographique *type-based* d'un mot irrégulier est illustré dans la Figure 6. La participation du contexte dans l'activation du nœud phonologique nécessaire à un recodage phonologique réussi n'est cependant pas exclusive aux mots irréguliers et participe à la lecture de tous les mots (cf. Figure 5). Néanmoins, sans cette dernière, le recodage phonologique des mots irréguliers échoue, tandis que dans le cas des mots réguliers, elle n'est pas indispensable. Ceci est cohérent avec les études comportementales mettant en évidence l'absence d'effet du contexte dans la lecture des mots réguliers (Nation et al., 2007; Wang et al., 2011). Dans le cadre des simulations réalisées par le modèle, la participation du contexte apparaît également critique dans le cas des potentiophones (i.e., mot irrégulier dont la prononciation régularisée correspond à un autre mot existant, e.g., FILS) et des homophones (e.g., MAIRE/MÈRE). Bien que nous n'ayons pas connaissance d'études ayant exploré ces mots en particulier chez des apprentis-lecteurs, il est cohérent d'estimer qu'ils représentent une difficulté pour eux également, et que le contexte permettrait de résoudre l'ambiguïté posée par le recodage phonologique.

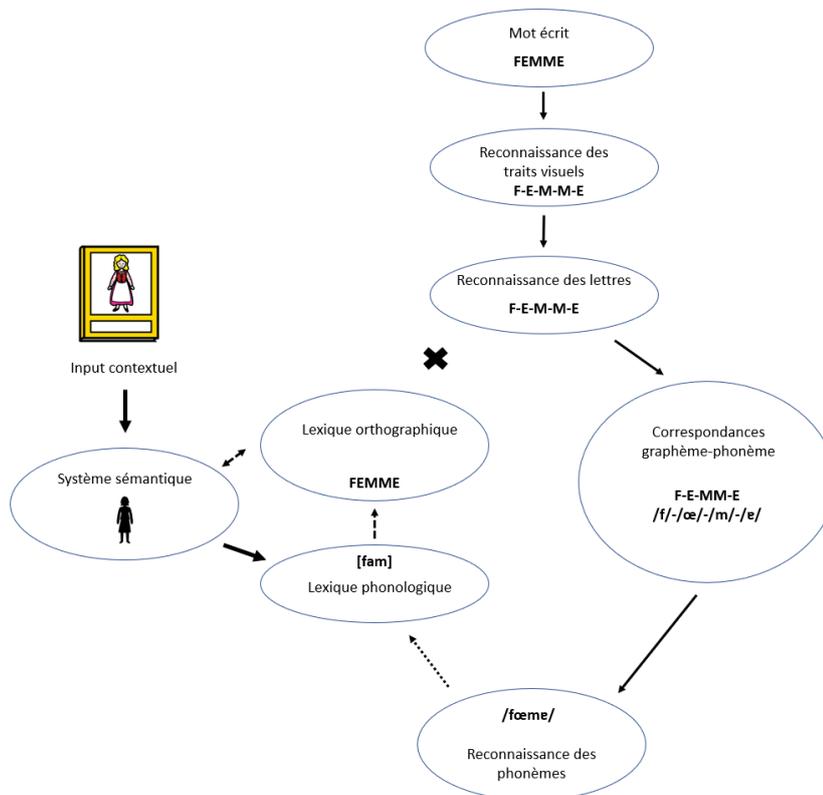


Figure 6. Apprentissage orthographique *type-based* d'un mot irrégulier d'après Pritchard et al. (2018).

NB : Les pictogrammes sont issus du site ARASAAC (<https://arasaac.org>) et créés par Sergio Palao ; elles sont la propriété du Gouvernement d'Aragon, Espagne (BY-NC-SA)

Dans les apprentissages *type-based* et *token-based* décrits ci-dessus, l'apprenant simulé par le modèle dispose néanmoins de règles de CGP complètes et d'un vocabulaire extensif, en ce qu'il possède déjà les nœuds phonologiques et sémantiques correspondant aux mots qui font l'objet d'un apprentissage orthographique. Cette situation ne reflète pas celle d'un appreni-lecteur typique, dont la voie sublexicale n'est pas parfaite, et dont le vocabulaire s'enrichit en partie avec ses lectures (Stanovich, 1986). Si Pritchard et al. (2018) n'ont pas modélisé les liens entre l'acquisition du vocabulaire et l'apprentissage orthographique, ils ont cependant réalisé une simulation correspondant au cas d'un appreni-lecteur présentant des CGP appauvries, i.e., dont l'unité graphémique maximale était réduite soit à deux lettres (e.g., OIN lu /wan/), soit à une lettre (e.g., OIN lu /o/-i/-/n/). Les auteurs faisaient également varier l'appui sur le contexte, soit en réduisant ou en

augmentant l'activation sémantique, soit en rendant l'input contextuel plus ou moins ambigu. Les résultats mettent en évidence le rôle essentiel de l'interaction entre contexte et recodage phonologique postulée par Share (1995) : lorsqu'une contribution minimale du contexte est présente (i.e., contexte ni trop ambigu, ni trop prépondérant), les CGP, même réduites à une lettre, permettent de lire plus de la moitié des mots réguliers et irréguliers présentés. Ce taux de réussite monte à 65.6% pour les deux types de mots lorsque la simulation autorise les digraphes.

Ainsi, le modèle ST-DRC paraît correctement simuler les prédictions portées par l'hypothèse d'auto-apprentissage et propose un cadre conceptuel valide d'un point de vue computationnel pour l'apprentissage de la lecture, au moins pour la langue anglaise. Les hypothèses de Share (1995, 1999) concernant la production orthographique ne sont, en revanche, nullement abordées dans le modèle ST-DRC, puisque le modèle DRC sur lequel il se base ne prévoit pas la situation de transcription. La capacité à orthographier correctement un mot est souvent considérée comme découlant directement de la représentation orthographique encodée via l'auto-apprentissage, et est, à cette occasion, fréquemment évaluée dans les expériences d'auto-apprentissage (Bosse et al., 2015; Cunningham, 2006; Cunningham et al., 2002; de Jong & Share, 2007). L'étude de Carrillo et al. (2013) a notamment permis de tester cette hypothèse, en comparant les performances d'apprenants natifs hispanophones et francophones en production orthographique de mots fréquents et peu fréquents, sélectionnés pour présenter des CPG d'une inconsistance comparable (i.e., nécessitant, dans une langue comme dans une autre, le recours à des représentations orthographiques préalablement engrammées). L'espagnol présentant, de manière générale, des CPG très consistantes, il devrait, selon la théorie de l'auto-apprentissage, permettre à ses locuteurs de recourir facilement au recodage phonologique et ainsi d'acquérir des représentations orthographiques plus précocement qu'une langue

aux CGP inconsistantes telle que le français. De fait, dans l'étude de Carrillo et al. (2013), les apprenants hispanophones montraient, dès le début de la 2^{ème} année d'instruction, des performances plus élevées pour la transcription de mots fréquents par rapport à celle de mots peu fréquents, contrairement aux apprenants francophones, qui ne montraient un tel effet qu'à la fin de cette même année d'instruction. Ce résultat semble ainsi confirmer que les performances en production orthographique de mots inconsistants sont d'autant plus élevées que le système de transcription possède des CGP consistantes, permettant un recodage phonologique plus fiable – ce qui correspond aux prédictions de la théorie de l'auto-apprentissage.

En conclusion, selon les partisans de la théorie de l'auto-apprentissage, l'apprentissage orthographique serait attribuable à trois facteurs essentiels : la capacité à acquérir un recodage phonologique fiable ; la capacité à s'appuyer sur des informations contextuelles, phonologiques et sémantiques préalablement engrammées pour identifier un candidat approprié en cas de décodage partiel ; et enfin, de s'exposer suffisamment à l'écrit pour multiplier les opportunités d'auto-apprentissage. Néanmoins, cette position ne s'intéresse qu'à une situation d'identification des mots ou de transcription relativement automatique. Elle implique que les apprenants ne se baseraient que sur des représentations orthographiques spécifiques à chaque mot pour l'orthographe, et qu'il serait impossible pour eux de transcrire correctement un mot qu'ils n'auraient pas préalablement rencontré en situation de lecture. Par ailleurs, dans cette conception, aucune place n'est laissée aux stratégies explicites utilisées (à bon ou à mauvais escient) par les individus lorsqu'ils sont confrontés à une difficulté pour identifier ou écrire un mot, puisqu'elles nécessitent des traitements conscients qui ne sont pas abordés dans ces modèles.

Plusieurs auteurs ont souligné la nécessité de tenir compte des facteurs énoncés ci-dessus lors de l'apprentissage orthographique, à la fois pour décrire au mieux le

développement des capacités de lecture et d'orthographe des apprenants, et pour pouvoir orienter judicieusement les pratiques d'enseignement. Cette conception, dénommée « Théories des répertoires » par Templeton (2020), regroupe un ensemble de travaux (e.g., Théorie de la Triple Forme des Mots écrits, Daffern et al., 2015; Théorie des *Overlapping Waves*, Kwong & Varnhagen, 2005 ; Théorie de l'Intégration de Multiples Patterns, Treiman & Kessler, 2014) ayant tous en commun le fait de considérer que l'apprenant pourrait et *devrait* avoir recours aux différentes dimensions linguistiques d'un même mot pour faciliter son apprentissage orthographique. Nous choisirons de nous concentrer sur l'une de ces théories développées par Treiman et collaborateurs : la théorie de l'Intégration des Multiples Patterns (IMP).

1.4 La théorie de l'Intégration des Multiples Patterns (IMP)

La théorie de l'IMP, comme la plupart des théories des répertoires, se base sur le constat que les systèmes de transcription des langues alphabétiques présentent des régularités sublexicales (ci-après, des patterns), accessibles à la fois à un apprentissage incident (Pacton et al., 2005) et à une instruction explicite (Singh et al., 2021). Ces patterns constituent des représentations orthographiques générales, applicables à un grand nombre de mots sur la base d'une approche déterministe (e.g., le mot TASSE *doit* prendre deux S sinon il se lirait /taz/) et/ou statistique (i.e., le mot TASSE *a plus de chance* de s'écrire avec deux S qu'avec un C), et existent en parallèle des représentations orthographiques spécifiques (Castles & Nation, 2006) abordées jusqu'à présent.

Selon Treiman et collaborateurs (Treiman, 2017b; Treiman & Kessler, 2014), ces régularités sont de deux classes : la première traite de la forme extérieure de l'écrit, i.e., les patterns graphotactiques ; la seconde concerne les liens entre l'écrit et le langage, i.e., les patterns phonologiques et morphologiques. La théorie de l'IMP postule que la capacité à coordonner ces différents patterns pour un même mot facilite l'apprentissage

orthographique en la rendant moins arbitraire. À l'inverse, lorsque l'un de ces patterns n'est pas applicable pour un mot, la mémorisation est rendue plus ardue. Par exemple, la graphie EAU dans le mot AGNEAU respecte les différents patterns cités : sur le plan phonologique, EAU est une graphie acceptable pour transcrire le phonème [o] ; EAU correspond à un morphème portant la signification de diminutif ; le phonème final [o] est le plus souvent transcrit EAU en français dans les mots bisyllabiques. En revanche, dans le mot LANDAU, la graphie AU ne respecte pas les patterns graphotactiques (AU est une graphie peu fréquente pour transcrire le son /o/ en fin de mot) et morphologiques (AU ne correspond pas à un quelconque morphème), rendant la mémorisation de cette partie du mot plus complexe. Ce phénomène est attesté dans l'étude de Kemp et Bryant (2003) auprès d'enfants anglais scolarisés de la première à la troisième année de primaire. Dans cette étude, les enfants devaient écrire sous dictée des cibles (mots appariés sur leur fréquence, leur longueur et leur complexité phonétique) présentés au sein de phrases. Ces dernières étaient pour un premier tiers des pluriels surdéterminés (i.e., l'usage du S pluriel était suggéré à la fois par la morphosyntaxe et par les régularités graphotactiques de la langue, e.g., FIBS), et pour un second tiers des pluriels déterminés uniquement sur la base de la morphosyntaxe (i.e., l'usage du S pluriel, bien que correct, ne correspond pas à un pattern graphotactique fréquent, e.g., FLEAS). Le dernier tiers était des mots qui n'étaient pas des pluriels mais qui comprenaient également le phonème final /z/ (e.g., FIZZ) afin de ne pas contraindre l'utilisation du S sur l'ensemble des items. Les résultats obtenus montrent que les enfants avaient significativement plus de facilité à écrire S pour les pluriels surdéterminés par rapport aux pluriels qui n'étaient déterminés que morphologiquement.

Dans les chapitres suivants, nous aborderons les différents patterns (phonologiques, morphologiques et graphotactiques) évoqués par la théorie de l'IMP.

Nous tenterons d'illustrer la contribution dans l'identification des mots et la production orthographique de chacun d'entre eux, et de discuter de l'adéquation de cette approche pour remédier aux difficultés relatives à ces acquisitions.

Chapitre 2 – Les patterns phonologiques

Selon la théorie de l'IMP (Treiman, 2017a; Treiman & Kessler, 2014), les patterns phonologiques font partie de la classe des patterns transcrivant un aspect du langage oral, en l'occurrence la phonologie. Ils renvoient à la maîtrise des associations grapho-phonémiques qui limitent le choix de lettres utilisables lors d'une transcription. Ainsi, dans le mot /rido/, bien qu'une multitude de graphies différentes puisse être envisagée pour transcrire le phonème /o/ (e.g., O, EAU, AU, OT, AUD, etc.), l'influence des patterns phonologiques devrait interdire au scripteur d'envisager des graphèmes qui ne sont pas conventionnellement associés à ce phonème, tels que ON ou AI. Les patterns phonologiques concernent également l'orthographe d'un phonème en fonction de son contexte phonologique (Treiman, 2018). Par exemple, en français, le phonème /ɛ/ possède plusieurs transcriptions possibles (E, È, Ê, AI, EI, ET) ; néanmoins, son orthographe dans un mot est en partie prévisible par son environnement phonémique. Par exemple, en français, l'ajout d'un accent grave ou circonflexe sur le E dans des mots tels que RESTE ou FLEXION n'est pas nécessaire, puisqu'il n'existe pas de séquence ES ou EX dans laquelle E peut être prononcé /œ/. Ces règles de transcription, évidentes pour un expert, requièrent pourtant l'acquisition et la coordination de différentes compétences.

En premier lieu, l'intégration des patterns phonologiques lors de l'apprentissage orthographique n'est possible que si l'apprenant bénéficie d'un enseignement explicite (Ehri et al., 2001) et relativement structuré (i.e., commencer par les associations graphème-phonème les plus consistantes plutôt que par les exceptions ; Sprenger-Charolles, 2017) des relations conventionnelles entre phonologie et graphie. Au-delà de l'instruction formelle, d'autres compétences supplémentaires apparaissent néanmoins sous-tendre la capacité à acquérir ces CGP. Nous évoquerons successivement les trois habiletés les plus consensuellement reconnues comme participant de cette acquisition : la connaissance des

lettres et la conscience phonémique. Notons que ces compétences, bien que détaillées séparément, interagissent très probablement entre elles au sein du développement (de Jong & Olson, 2004; Mann & Foy, 2003).

2.1. Connaissance des lettres et conscience phonémique

2.1.1 La connaissance des lettres

La connaissance des lettres renvoie à un ensemble de savoirs concernant les lettres d'un alphabet, comprenant leur identité visuelle, la connaissance de leur nom et de leurs sons (Foulin, 2005). L'identification des lettres concerne la capacité à reconnaître les lettres comme des signes graphiques propres et à les discriminer entre elles. Dès l'âge de 4 ans, les enfants seraient capables de reconnaître les lettres de leur alphabet par rapport à d'autres symboles ou caractères issus d'un autre système de transcription (Lavine, 1977; Levy et al., 2006). La discrimination des lettres entre elles, cependant, nécessiterait une période plus longue d'apprentissage, d'autant plus en ce qui concerne les lettres miroirs (i.e., b/d, p/q, etc.) en raison de la capacité du système visuel à identifier immédiatement un stimulus indépendamment de son orientation droite-gauche (*mirror invariance* ; Biederman & Cooper, 2009; Rollenhagen & Olson, 2000). La connaissance du nom des lettres constitue une autre compétence chez les pré-lecteurs, ce savoir leur permettant de déduire le son des lettres dont ils ont appris le nom, du fait de leur acrophonie (e.g., le nom attribué à la lettre D commence par le son /d/). Quant à la connaissance du son des lettres, elle constitue véritablement la première étape vers l'acquisition des CGP (Byrne & Fielding-Barnsley, 1989) et donc des patterns phonologiques, mais qui devra s'affiner pour intégrer notamment la notion de graphème (e.g., CH ne correspond pas à la somme des sons des lettres C et H) et de règles contextuelles (e.g., C se lit /s/ devant E, I et Y mais /k/ devant A, O et U). La connaissance du son et du nom des lettres renvoie ainsi à des *constructs* très proches, mais dont la capacité à prédire le développement de la lecture/écriture évolue en fonction de l'âge de l'enfant. En

effet, comme le note Foulin (2005), en maternelle, la connaissance des lettres présente une meilleure variance que celle des sons, et prédit donc de manière plus fiable l'acquisition ultérieure de la lecture ; néanmoins, la première plafonne à l'entrée à l'école élémentaire, rendant alors la seconde plus pertinente à mesurer.

2.1.2 La conscience phonémique

En somme, l'intégration des patterns phonologiques suppose dans un premier temps que l'individu maîtrise les lettres de l'alphabet pour pouvoir abstraire le son conventionnellement attribué à la lettre. Cette association, bien qu'imparfaite car ne tenant pas compte des complexités du système de transcription, représente une première étape vers l'acquisition des CGP/CPG. Néanmoins, ce savoir ne peut être appliqué si la structure phonologique de la cible n'est pas traitée comme telle. L'apprenant doit en effet être capable, pour lire comme pour écrire, de percevoir les composants de la chaîne parlée afin de les mettre en relation avec le système de transcription sous-tendu par les lettres. Cette faculté d'analyse renvoie à la conscience phonologique, et peut concerner différentes unités de la parole : les syllabes (conscience syllabique), les rimes (conscience rimique) ou les phonèmes (conscience phonémique). Parmi ces trois unités, c'est celle du phonème qui paraît la plus corrélée à l'acquisition de la lecture (Melby-Lervåg et al., 2012) dans les langues alphabétiques, ce qui s'explique aisément lorsque l'on considère que ces systèmes de transcription reposent sur le graphème, qui est le pendant graphique du phonème.

Le phonème est cependant souvent considéré comme une unité artificielle et difficile à percevoir dans le continuum de la parole (Liberman et al., 1967), expliquant les résultats d'études suggérant que les pré-lecteurs n'y ont pas accès avant l'apprentissage formel de la lecture (Lonigan et al., 1998). Il existerait ainsi une relation circulaire entre conscience phonémique et apprentissage des CGP, l'une favorisant l'autre et inversement (Goswami & Bryant, 2016). En revanche, selon Gallet et al. (2020), la syllabe orale correspond à « l'unité

phonologique prélexicale la plus précocement accessible par les enfants prélecteurs francophones (Maïonchi-Pino et al., 2010, 2012) ». Ainsi, si la syllabe constitue une unité phonologique facilement manipulable, renforcer la conscience syllabique permettrait de créer une sorte de « pont » permettant le développement de la conscience phonémique, facilitant *in fine* l'acquisition des CGP (hypothèse du Pont Syllabique, Doignon-Camus & Zagar, 2014). L'intérêt de passer par la conscience syllabique pour favoriser le développement de la conscience phonémique n'est néanmoins pas établi avec certitude (Ukrainetz et al., 2011). Tout d'abord, contrairement à la conscience phonémique, la conscience syllabique ne constitue qu'un prédicteur modeste (Engen & Høien, 2002; Høien et al., 1995) du développement de la lecture. Ensuite, bien qu'il soit généralement admis que les enfants n'aient pas accès au phonème avant leur première année d'instruction élémentaire, ce constat n'est pas systématiquement retrouvé (Chaney, 1992; Maclean et al., 1987). En effet, la conscience phonémique des enfants pré-lecteurs pourrait émerger tout aussi précocement que la conscience syllabique : cela dépendrait de la formulation des consignes, les contraintes mnésiques et cognitives pouvant hautement varier en fonction de la tâche employée (Anthony et al., 2003; Kenner et al., 2017). En revanche, les entraînements à la conscience syllabique n'ont pas rapporté d'impact significatif, ni sur la conscience phonémique (Ukrainetz, 2008) ni sur l'acquisition de la lecture/écriture (Halls, 2021).

Le rôle de la conscience phonémique est en revanche essentiel dans le développement des CGP/CPG. Des confusions dans les représentations des phonèmes peuvent en effet mener à des erreurs non-phonologiquement plausibles, traduisant des confusions entre des sons proches (e.g., FRITE transcrit VRITE) ou des difficultés à traiter les différents phonèmes d'un même groupe consonantique (e.g., BRAS transcrit BAS ; Bruck & Treiman, 1990). Bien que la conscience phonémique bénéficie de l'apprentissage des CGP (Goswami & Bryant, 2016), la première joue ainsi un rôle causal dans l'application des secondes. Ce fait est

illustré via les études d'entraînement, dont l'intérêt est ainsi double : d'un point de vue fondamental, elles visent à tester l'hypothèse d'une relation de causalité entre conscience phonémique et acquisition des CGP/CPG. D'un point de vue plus appliqué, leur objectif est d'instruire les méthodes d'enseignement et/ou de remédiation auprès d'apprenants en difficulté de lecture, ces dernières pouvant alors se baser sur des pratiques éprouvées scientifiquement (*evidence-based practice*).

L'efficacité des entraînements phonémiques (i.e., à la conscience phonémique) pour favoriser l'apprentissage précoce de la lecture et de l'orthographe est connue depuis longtemps. La méta-analyse de Bus et van IJzendoorn (1999) menée sur 36 études indépendantes explorant l'impact d'interventions phonémiques met en évidence une taille d'effet importante ($d = 0.85$) pour la lecture simplifiée¹⁶ et une taille d'effet plus modérée pour la lecture de mots ($d = 0.34$). Un effet plus faible mais également positif a été retrouvé pour la production orthographique ($d = 0.25$). La méta-analyse plus récente de Suggate (2016), menée sur 71 études indépendantes, retrouve cet effet positif de l'entraînement phonémique (bien qu'il soit moins important ; $d = 0.40$) pour ce qui est dénommé « prélecture » (tâches de conscience phonémique, lecture de lettres, de non-mots, etc.). Un effet positif est également retrouvé pour la lecture de texte ($d = 0.32$) et pour la production orthographique de mots ($d = 0.33$). Enfin, une étude très récente de Wolff & Gustafsson (2022), menée sur 364 enfants de maternelle (138 dans le groupe entraîné à la conscience phonémique ; 226 dans le groupe contrôle chez qui des compétences non-phonologiques, telles que la géométrie ou des connaissances sur le monde, étaient travaillées), a mis en évidence que seulement deux sessions de six semaines d'entraînement phonologique (une par an jusqu'à l'apprentissage formel de la lecture) permettaient d'obtenir des gains significatifs

¹⁶ Les auteurs n'explicitent pas ce qu'ils entendent par cette formulation. Étant donné que cette variable dépendante est opposée à la lecture de vrais mots, nous inférons qu'elle renvoie à la lecture de syllabes ou de non-mots.

($d = 0.27$; $p < .001$) sur la lecture en première année, et sur la production orthographique ($d = 0.28$; $p < .01$) en deuxième année de primaire. Il est intéressant de préciser que lors de ces entraînements, les habiletés phonologiques étaient travaillées sans aucun support écrit.

L'ensemble de ces études témoigne du rôle prédictif de la conscience phonémique pour l'apprentissage de la lecture et de l'écriture. Nous avons néanmoins supposé que cet effet était obtenu via l'acquisition des CGP/CPG, pour lesquelles la conscience phonémique représente un précurseur indispensable. Il est ainsi nécessaire d'examiner plus directement le lien entre la capacité à appliquer les règles phono-graphémiques et l'apprentissage orthographique via des mesures de lecture et de production orthographique de mots.

2.2 Le rôle des patterns phonologiques dans l'identification des mots et la production orthographique

Si l'efficacité des entraînements phonographiques pour le développement de la lecture et de l'écriture a été régulièrement réaffirmée ces vingt dernières années (e.g., Castles et al., 2018; Ehri et al., 2001; Suggate, 2010), certains auteurs ont alerté sur l'existence de biais dans les revues réalisées (Torgerson et al., 2019), ainsi que sur les fréquentes imprécisions dans les interprétations basées sur ces dernières (Bowers, 2020). Il convient donc de déterminer d'une part sur quels domaines précis un effet bénéfique des entraînements phonographiques peut être attendu, et d'autre part dans quelle mesure cet effet est observé, c'est-à-dire quels sont les facteurs favorisant leur efficacité (e.g., âge des apprenants, intensité et contenu de l'entraînement).

Les méta-analyses publiées ces cinq dernières années tendent à se concentrer sur les apprenants en difficulté de lecture (e.g., Galuschka et al., 2020; McArthur et al., 2018; Wanzek et al., 2018). À notre connaissance, la plus récente synthèse incluant des enfants tout-venants correspond à la méta-analyse réalisée par Suggate (2016), précédemment citée. Dans cette dernière, concernant l'impact des entraînements phonographiques sur la lecture,

les tailles d'effet trouvées paraissent étonnamment moins importantes que pour les entraînements phonémiques ($d = 0.26$ par rapport à 0.32 pour les entraînements phonémiques pour la lecture de textes ; $d = 0.32$ pour la prélecture par rapport à 0.40 pour les entraînements phonémiques). L'impact sur la production orthographique semble en revanche légèrement plus élevée ($d = 0.38$ par rapport à 0.33 pour les entraînements phonémiques). L'auteur rappelle cependant que la méta-analyse réalisée ne permet pas de déterminer s'il existe une véritable différence de tailles d'effet entre celles obtenues pour les entraînements phonémiques et celles obtenues pour les entraînements phonographiques. Il suggère par ailleurs que les entraînements phonémiques ont pu suffire à faciliter l'apprentissage des CGP et donc de la lecture, initiant un effet Mathieu (Pfost et al., 2014; Stanovich, 1986) au final équivalent à un entraînement phonographique.

Les méta-analyses suivantes, portant uniquement sur les entraînements phonographiques dispensées auprès d'enfants en difficulté de lecture, ont néanmoins retrouvé des tailles d'effet plus conséquentes. Par exemple, la méta-analyse de McArthur et al. (2018) réalisée sur 14 études soigneusement sélectionnées pour la qualité de leur *design* retrouve un impact important de l'entraînement phonographique pour toutes les mesures de précision de lecture, qu'il s'agisse de mots réguliers et irréguliers ($d = 0.51$) ou de non-mots ($d = 0.67$). En revanche, les auteurs ne retrouvent pas d'effet significatif sur la production orthographique de mots, phénomène attribué à un nombre d'études incluses pour ces variables très faible ($k \leq 5$). La méta-analyse de Galuschka et al. (2020), réalisée sur huit études, retrouve quant à elle un impact similaire des interventions phonographiques sur la lecture ($d = 0.62$) et l'orthographe ($d = 0.68$). Il convient néanmoins de préciser que dans cette synthèse, les études incluses devaient avoir pour but l'amélioration de la production orthographique. Les interventions accordaient donc, pour une majorité d'entre elles, une place importante à

l'apprentissage orthographique, au-delà des seules associations phono-graphémiques (e.g., Blachman et al., 2014).

En complément des tailles d'effet moyennes pour chaque variable d'intérêt, McArthur et al. (2018) ont réalisé des analyses de sous-groupes rassemblant les différentes études selon les caractéristiques de leur *design* (e.g., contenu ou durée de l'entraînement). Ces analyses, bien que post-hocs (i.e., les catégories sont réalisées après que les résultats des études soient obtenus) et non-expérimentales (i.e., il est impossible de cibler un paramètre du *design* tout en maintenant les autres constants), sont utiles pour identifier de potentiels facteurs conditionnant l'efficacité d'une intervention. Ainsi, une taille d'effet significative a été obtenue pour les études proposant moins de deux heures d'entraînement par semaine ($d = 0.54$), pendant moins de 3 mois ($d = 0.61$). L'intervention phonographique était efficace lorsqu'elle était associée à un entraînement à la reconnaissance rapide des mots (*sight word training* ; $d = 0.73$), qu'elle était dispensée en petits groupes ($d = 0.33$) et de manière non-informatisée ($d = 0.70$). Toutefois, il est important de garder à l'esprit que, du fait de la catégorisation des études en sous-groupes, le nombre d'études par catégorie était très faible (entre 2 et 9 études), ce qui d'une part limite les possibilités de généralisation de ces résultats, et qui d'autre part ne permet pas de conclure sur le fait que l'efficacité de ces études soit bien due à la caractéristique du *design* ciblée.

En somme, il apparaît nettement, grâce aux méta-analyses les plus récentes réalisées à la fois sur l'ensemble des apprenants et sur des apprenants en difficulté, que les entraînements phonographiques conduisent à des résultats positifs pour la lecture (Galuschka et al., 2020; McArthur et al., 2018; Suggate, 2016). En revanche, les analyses menées sur l'orthographe montrent des résultats plus mitigés. De plus, la supériorité des entraînements phonographiques par rapport aux entraînements phonémiques sans support écrit ne paraît pas démontrée (Suggate, 2016), contrairement à ce que les résultats de méta-analyses plus

anciennes semblaient avancer (e.g., Ehri et al., 2001). Enfin, bien que les analyses de sous-groupes suggèrent que les interventions phonographiques sont plus efficaces selon certaines caractéristiques, telles qu'une administration relativement brève et peu intense en petits groupes, le manque d'études ne permet pas de l'affirmer formellement.

2.3 De l'insuffisance des patterns phonologiques

Si les méta-analyses confirment ainsi le rôle essentiel des CGP et de la conscience phonémique dans le développement de la lecture, en permettant notamment l'accès à un recodage phonologique et donc *in fine* à l'apprentissage orthographique, comme discuté précédemment (cf. §1.3.2), elles n'ont pas permis de démontrer avec certitude un impact sur la production orthographique via la seule intégration des patterns phonologiques. En d'autres termes, il est possible que les systèmes de transcription de l'anglais et du français soient trop opaques pour que les apprentis-scripteurs de ces langues puissent s'appuyer de manière fiable sur les patterns phonologiques, au-delà de certains mots très consistants (e.g., FRITE, POISSON). Cette interprétation est confortée par deux observations : premièrement, les apprentis-scripteurs de langue transparente (e.g., l'italien) commettent considérablement moins d'erreurs d'orthographe au même âge que ceux de langue opaque (e.g., l'anglais). Par exemple, dans l'étude de Marinelli et al. (2015), des enfants italiens de 9 à 10 ans obtiennent un pourcentage d'erreurs d'environ 10% sur la production orthographique des mots réguliers comme irréguliers, tandis que des enfants anglais de même âge commettent environ 30 (pour les mots réguliers) à 40% (pour les mots irréguliers) d'erreurs. Par ailleurs, chez les enfants italiens, l'effet de fréquence signant la présence d'un lexique orthographique et l'intervention de la voie lexicale (cf. §1.1.2 et §1.2.2) est majoritairement absent, tandis qu'il est systématiquement présent chez les enfants anglais. Ces résultats suggèrent ainsi que la consistance des CPG des langues transparentes telles que l'italien permettent à ses apprenants de s'appuyer en grande partie sur les patterns orthographiques pour orthographier

correctement près de 90% des mots, et ce dès l'âge de 9 ans. En revanche, pour les apprenants d'une langue opaque telle que l'anglais ou le français, d'autres patterns doivent être pris en compte, du fait de l'inconsistance des CPG.

Un autre argument en faveur de l'insuffisance des patterns phonologiques réside dans l'existence d'un seuil de développement au-delà duquel les entraînements phonographiques semblent perdre leur efficacité sur l'orthographe. En effet, la méta-analyse d'Ehri et al. (2001) met en évidence que, si l'impact des entraînements phonographiques sur la production orthographique est très important pour les enfants de maternelle ($d = 0.97$), il diminue sensiblement chez les enfants de première année de primaire ($d = 0.52$) pour devenir faible ($d = 0.14$) et non significatif, une fois passé ce niveau de classe. Ainsi, les patterns phonologiques joueraient principalement un rôle au début de l'apprentissage de l'orthographe, permettant aux apprenants de conserver, dans la transcription de chaque mot, un « squelette phonologique » (i.e., représentation phonémique minimale, e.g., PINTUR pour PEINTURE ; Bahr et al., 2012; Bourassa & Treiman, 2003), sur lequel viendront se greffer par la suite des représentations lexicales. Cet argument est étayé par les études d'analyse d'erreurs en français et en anglais, mettant en évidence une diminution progressive des erreurs non-phonologiquement plausibles avec l'âge, qui ne représentent plus que 5 à 15% des erreurs totales commises en cinquième année de primaire (Bahr et al., 2012; Hazard et al., 2020).

La dimension phono-graphémique, incarnée par les patterns phonologiques tels que décrits par Treiman et collaborateurs (Treiman, 2017a, 2017b; Treiman & Kessler, 2014), apparaît donc essentielle mais insuffisante dans les langues opaques. Certains auteurs argumentent même que, dès le début de l'apprentissage des CGP, d'autres dimensions linguistiques devraient être enseignées pour donner un sens aux nombreuses irrégularités rencontrées par les élèves (Bowers & Bowers, 2017). Dans la théorie de l'IMP, cette

dimension correspond aux patterns morphologiques, que nous abordons dans le chapitre suivant.

Chapitre 3 – Les patterns morphologiques

Parallèlement aux patterns phonologiques, les patterns morphologiques transcrivent également des informations d'ordre linguistique (Treiman, 2017b; Treiman & Kessler, 2014). Cette influence est particulièrement importante en français, qui, d'après l'analyse lexicologique de Rey-Debove (1984), comporterait près de 80% (27 390 entrées sur les 34 290 du *Robert Méthodique*) de mots complexes, c'est-à-dire comprenant au moins deux morphèmes. Ces mots plurimorphémiques sont, d'après l'auteure, constitués à 69% environ de morphèmes communs. Ainsi, il serait bien plus efficace de considérer l'unité du morphème plutôt que du mot entier durant l'apprentissage orthographique, les patterns morphologiques renvoyant à la part de l'orthographe d'un mot qui pourrait être déduite via la connaissance préalable des morphèmes qu'elle contient. Cette prise en compte des patterns morphologiques en lecture et en orthographe nécessiterait d'avoir développé une certaine sensibilité à la morphologie du langage.

3.1 Morphologie et conscience morphologique

3.1.1 La morphologie

Huot (2010) définit la morphologie comme l'étude linguistique descriptive de la forme des mots et de la signification associée à celle-ci. Nous nous focaliserons uniquement sur la morphologie morphématique, qui considère le morphème comme la base de son analyse (Fradin, 2003). Le morphème est défini comme la plus petite unité de sens identifiable au sein d'un mot. L'étude des morphèmes peut être réalisée soit du point de vue diachronique (i.e., historique) soit synchronique (i.e., contemporaine). En morphologie synchronique, le lien entre le mot et ses morphèmes doit pouvoir être évident sans recours à l'étymologie : ainsi, bien que le mot PLAFOND soit historiquement composé des morphèmes PLAT et FOND, ce mot est considéré comme monomorphémique d'un point de vue synchronique, tandis qu'il est bimorphémique d'un point de vue diachronique (Huot, 2010).

La morphologie synchronique s'intéresse à deux phénomènes : la flexion et la formation des mots (Booij, 2015).

La flexion ou morphologie flexionnelle est très proche du domaine de la morphosyntaxe (Huot, 2010) et est assimilée à un sous-système de la grammaire (Booij, 2015). La flexion d'un mot s'accomplit selon deux axes : la flexion inhérente (choix du locuteur d'une certaine forme verbale plutôt qu'une autre) et la flexion contextuelle (contrainte pour la forme verbale de s'accorder en genre et en nombre avec le sujet). Par exemple, dans la phrase « Je voudrais du café », la flexion du verbe VOULOIR possède une flexion inhérente (choix du conditionnel par le locuteur plutôt que le présent) et une flexion contextuelle (première personne du singulier). La structure d'un mot fléchi comporte ainsi typiquement deux sortes d'éléments (Booij, 2015) : une base lexicale ou lexème (i.e., morphème appartenant au lexique d'une langue, e.g. VOULOIR) et un suffixe flexionnel (e.g., -AIS). La base se réalise sous la forme d'un radical (e.g., VOUDR-) qui possède souvent des allomorphes (i.e., des formes différentes d'un même morphème, e.g., VEU-, VOUL-, etc.).

La formation des mots englobe plusieurs phénomènes distincts, dont le plus fréquent en français est la dérivation (Clark, 1998). La morphologie dérivationnelle fonctionne de manière similaire à la morphologie flexionnelle en ce qu'elle emploie également des bases lexicales et des suffixes. En morphologie dérivationnelle, les affixes permettent de créer de nouveaux lexèmes (e.g., COIFFER+EUR - COIFFEUR). Ceci n'est pas le cas des suffixes flexionnels : même s'ils résultent en des formes orthographiques différentes pour un même mot (e.g., COIFFEZ, COIFFONS, etc.), ces dernières ne sont pas considérées comme des entrées lexicales distinctes de la base. La dérivation emploie également des préfixes qui ont la particularité de changer le sens mais pas la nature du mot qu'ils complètent (e.g., LIRE - RELIRE : le mot reste un verbe même après adjonction du préfixe RE), contrairement aux

suffixes qui ont la possibilité d'en changer à la fois la nature et le sens (e.g., CROIRE - CROYABLE).

En plus de la dérivation, le français a également recours à la composition pour former de nouvelles entrées lexicales. Celle-ci procède par la combinaison de deux (e.g., CHOU-FLEUR) voire trois bases (e.g., RADIOTÉLÉSCOPE) entre elles. Celle-ci procède soit par accollement simple (e.g., PORTEMANTEAU), avec un ou plusieurs tirets (e.g., REZ-DE-CHAUSSÉE) ou par l'emploi de prépositions (e.g., POMME DE TERRE, SAC À MAIN). Cette combinaison est dite populaire lorsque les bases sont des mots français, et savante lorsqu'elles sont d'origine grecque ou latine (e.g., PTÉRODACTYLE). Notons que cette combinaison savante illustre la proximité entre le processus de composition et celui de dérivation : par exemple, dans le mot composé GASTRONOMIE, la base NOMIE pourrait être fonctionnellement considérée comme un suffixe, au même titre que ERIE dans MENUISERIE (Huot, 2010).

En français, les morphèmes sont donc incarnés par des bases (réalisées par des radicaux allomorphes) et des affixes (préfixes et suffixes). Les premières sont considérées comme des morphèmes « libres », en ce qu'elles correspondent à des lexèmes pouvant exister sans aucun affixe (e.g., PIE, RENARD, CROCODILE). À l'inverse, les affixes, qu'ils soient de nature flexionnelle (e.g., -EZ) ou dérivationnelle (e.g., -TION), constituent des morphèmes liés, ce qui signifie qu'ils ne peuvent exister comme des entrées lexicales propres.

3.1.2 La conscience morphologique

La conscience morphologique (CM) est définie comme la capacité d'identifier et de manipuler les morphèmes (Carlisle, 1995), et peut concerner des opérations mentales très variées. Plusieurs auteurs ont proposé une typologie d'épreuves permettant de la mesurer chez les apprenants (e.g., Apel et al., 2013; Apel & Henbest, 2016; Berthiaume et al., 2010; Goodwin et al., 2021). À partir de 27 études investiguant le développement de la CM en lien

avec la lecture, Berthiaume et collaborateurs ont répertorié dix types de tâches morphologiques, variant singulièrement dans le degré de conscientisation nécessaire à leur exécution. Par exemple, dans cette typologie, la lecture à haute voix de mots complexes est considérée comme une tâche morphologique. Cette habileté repose probablement sur des processus moins cognitivement contrôlés qu'une tâche de jugement d'intrus (e.g., lequel de ces mots est un intrus : REDONNER, REVOLER, REGARDER ?). Les auteurs notent qu'il est difficile de déterminer avec certitude le caractère implicite ou explicite de l'ensemble des tâches répertoriées, tant au niveau des processus cognitifs mis en œuvre pour leur réalisation que de leur caractère enseigné ou non dans un cursus scolaire classique. Par ailleurs, la réalisation même de certaines tâches semble hautement dépendante des connaissances lexicales préalables du sujet, telles que la tâche de jugement d'analogie : en effet, dans cette tâche, une paire de mots est proposée (e.g., POISSON-POISSONNIER), puis le sujet est invité à compléter la paire d'un troisième mot, sur l'exemple du couple précédemment présenté (e.g., FROMAGE-[FROMAGER]). Dans cette tâche, non seulement le sujet doit être capable d'extraire le suffixe cible dans l'exemple, mais il doit aussi pouvoir récupérer en mémoire un mot existant afin de compléter la paire de manière adéquate (dans l'exemple ci-dessus, un apprenant risquerait de proposer *FROMAGIER ou LAITIER s'il n'a jamais rencontré le mot FROMAGER auparavant).

En fonction des tâches, la CM se développerait plus ou moins précocement chez les apprenants. Dans son étude princeps, Berko (1958) a exploré, chez 56 enfants américains âgés de 4 à 7 ans, les capacités de CM sur des non-mots (*Wug test*), via des tâches de production flexionnelle (S pluriel, possessif et de la troisième personne du singulier ; présent de l'indicatif et progressif ; participe passé), dérivationnelle (adjectifs dérivatif, comparatif, superlatif et diminutif ; agentif) et compositionnelle (composition d'un mot à partir de deux mots proposés). Les réponses des enfants étaient comparées à celles de 12 étudiants ; les

réponses des premiers étaient considérées justes si elles concordait avec celles des seconds. L'analyse des réponses montre que les tâches de morphologie dérivationnelle étaient les moins réussies, avec seulement 11% de bonnes réponses. La composition était également difficile, quoique légèrement mieux réussie (18%). À l'inverse, de bons résultats étaient généralement obtenus pour la morphologie flexionnelle : par exemple, la plupart des tâches impliquant la formation de pluriels étaient réussies pour plus de 50% des enfants, avec sur ces items une amélioration significative entre la maternelle et la première année de primaire (les scores de réussite devenant proches de 100%). Les autres tâches de morphologie flexionnelle donnaient lieu à des résultats plus variables mais qui étaient, là aussi, en général réussies pour plus de la moitié des enfants.

Afin de préciser cette apparente différence de développement entre la CM flexionnelle et la CM dérivationnelle/compositionnelle, Selby (1972) a fait passer le *Wug test* à des enfants écossais anglophones âgés de 3 à 15 ans (150 enfants de maternelle et 100 enfants pour les autres groupes d'âges, totalisant ainsi près de 1 000 sujets), dont les performances ont été comparées à celles de 50 adultes (25 étudiants et 25 puériculteurs et enseignants de maternelle). Les résultats montrent que les compétences en CM flexionnelles sont acquises pour la moitié des participants aux alentours de l'âge de 3 ans (S pluriel et présent progressif), 6 ans (S possessif) et 8 ans (présent de l'indicatif et participe passé), générant des scores plafonds à l'âge de 10-12 ans. Les tâches de CM dérivationnelle, en revanche, présentent un tout autre type de développement : les scores agrégés de production de mots dérivés et composés ne donnent lieu à des performances avoisinant les 50% qu'à partir de 13 ans, ce score moyen se maintenant à 15 ans sans atteindre de plafond. Quant au score agrégé de production d'adjectifs superlatifs et comparatifs, le taux de réussite reste en-dessous de 20% jusqu'à l'âge de 10 ans, et n'atteint 42% de réussite qu'entre 13 et 15 ans.

Casalis & Louis-Alexandre (2000) suggèrent que cette importante différence dans l'acquisition de la CM flexionnelle versus dérivationnelle s'explique par le caractère prédictif de la flexion, par opposition à la dérivation. En effet, les suffixes flexionnels constituent un ensemble limité, produisant des combinaisons hautement régulières (e.g., CHANTER - CHANT+ONS ; PARTIR - PART+ONS, etc.), tandis que les affixes dérivationnels, plus nombreux et moins productifs, se combinent parfois de manière imprévisible aux bases (e.g., CHEVEUX - CHEVELURE mais POIL - PELAGE ; IN+CROYABLE mais A+SOCIAL). Ainsi, l'acquisition de la CM dérivationnelle dépendrait hautement des mots étudiés, de même que du système de transcription de la langue (Casalis et al., 2015). Ainsi, les apprenants de langue française semblent s'emparer plus tôt de la dérivation que les apprenants de langue anglaise, présumément du fait que ce phénomène de formation de mots soit plus prégnant en français qu'en anglais (Duncan et al., 2009) De plus, une relation phonologiquement et/ou orthographiquement opaque entre la base et le dérivé pourrait significativement limiter les capacités de CM (Carlisle, 1988) : par exemple, le morphème SANG dans SAIGNER est plus difficile à percevoir que le morphème PLEUR dans PLEURER. Parallèlement à cette opacité orthographique et phonologique, l'opacité sémantique entre la base et le dérivé (Marslen-Wilson et al., 1994) semblerait également jouer un rôle dans le traitement morphologique (e.g., le morphème LUNE dans LUNETTE serait plus difficile à percevoir que le morphème FILLE dans FILLETTE).

3.2 Le rôle des patterns morphologiques dans l'identification des mots et la production orthographique

3.2.1 Modéliser la contribution de la conscience morphologique à la lecture et l'écriture : le Morphological Pathways Framework

La contribution de la conscience morphologique (CM) aux processus de lecture et d'écriture ont été formalisés par Levesque et collaborateurs dans leur modèle *Morphological*

Pathways Framework (MPF ; Levesque et al., 2021). Le MPF étant constitué de nombreux niveaux et sous-niveaux, nous en présentons une version simplifiée dans la Figure 7, dans laquelle nous nous concentrons sur la triade Processus d'Identification des Mots (*Word Identification Processes*), Système Orthographique (*Orthographic System*) et Système Linguistique (*Linguistic System*).

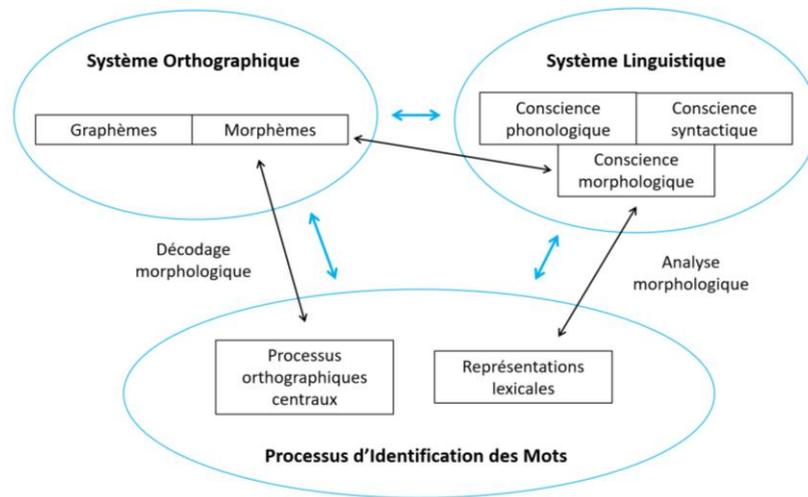


Figure 7. *Processus de décodage morphologique et d'analyse morphologique au sein du Morphological Pathways Framework (Levesque et al., 2021)*

NB. Le processus liant les morphèmes du Système Orthographique à la Conscience morphologique du Système Linguistique n'a pas de nom dans la version actuelle du modèle.

L'architecture du MPF constitue une proposition théorique se basant sur l'ensemble de la littérature ayant exploré les liens entre la conscience morphologique et les processus de lecture et d'écriture, à la fois chez les lecteurs experts et les apprentis-lecteurs. Elle n'a pas été testée via l'intelligence artificielle et s'applique, d'après ses auteurs, uniquement à la langue anglaise. Toutefois, le MPF présente un intérêt pour le français, en ce que ces deux langues sont toutes deux constituées à environ 80% de mots plurimorphémiques (Nagy & Anderson, 1984; Rey-Debove, 1984).

Le MPF introduit deux processus par lesquels l'unité morphémique contribue à l'identification des mots écrits : le décodage morphologique et l'analyse morphologique.

Le décodage morphologique correspond à l'identification visuo-orthographique des morphèmes durant le processus d'identification des mots écrits, permettant de les lire plus précisément et plus rapidement (nous détaillons les études l'ayant mis en évidence dans la partie 3.2.2). L'analyse morphologique renvoie quant à elle au processus réflexif permettant d'accéder à la signification d'un mot complexe sur la base de ses constituants morphémiques. Selon les auteurs du MPF, elle représenterait le lien entre les propriétés morphologiques d'un mot et la conscience morphologique du locuteur. Elle s'apparenterait, selon eux, à un procédé de décomposition morpho-sémantique. Levesque et collaborateurs ont proposé des tâches visant à l'isoler des processus de décodage morphologique et de conscience morphologique (Levesque et al., 2017, 2019) ; toutefois, nous n'avons pas connaissance d'autres études ayant repris cette distinction et retrouvons le plus souvent ces termes utilisés de manière interchangeable.

3.2.2 Morphologie et lecture : identification des morphèmes

Si l'analyse morphologique semble, à l'heure actuelle, se différencier difficilement du procédé de conscience morphologique en lui-même, le « décodage morphologique », en revanche, est étayé par une littérature riche et ancienne. Deux types de littérature se côtoient au sein de l'étude des processus morphologiques dans l'identification des mots écrits : les traitements morphologiques précoces et automatiques (classiquement évalués via un paradigme d'amorçage), et l'impact de la structure morphologique sur la précision et la fluence de lecture (évalué par des tâches de lecture à haute voix de mots ou non-mots).

3.2.2.1 Traitements morphologiques automatiques

Concernant les traitements automatiques impliqués lors de la lecture des mots complexes, plusieurs hypothèses ont été formulées (Marelli et al., 2020). Historiquement, il a tout d'abord été question de contraster les théories dites « *full-parsing* », telles que le modèle de Taft et Forster (1975) qui suggèrent que tous les mots sont décomposés en morphèmes,

aux modèles « *full-listing* ». Ces derniers, au contraire, supposent que tous les mots, qu'ils soient simples ou complexes, sont traités dans leur entièreté (Butterworth, 1983). Par la suite, les modèles connexionnistes ont envisagé que les effets morphologiques n'émergeraient que du fait des ressemblances orthographiques, phonologiques et sémantiques entre les mots morphologiquement liés (e.g., JARDIN, JARDINIER, JARDINAGE ; Seidenberg & Gonnerman, 2000). Une position intermédiaire, similaire à celle du modèle DRC (voir § 1.2.1) a également été envisagée. Selon cette dernière (e.g., Burani & Caramazza, 1987), le traitement des mots complexes pourrait s'effectuer selon deux voies en fonction de certains facteurs tels que la fréquence ou la transparence morphologique : un découpage morphémique et un accès direct aux représentations lexicales (Caramazza et al., 1988; Frauenfelder & Schreuder, 1992; Schreuder & Baayen, 1995). Les deux processus s'appliqueraient de manière parallèle, la voie « directe » étant plus efficace pour décoder les mots de haute fréquence, tandis que le procédé de découpage morphémique le serait davantage pour les mots inédits et sémantiques transparents (Schreuder & Baayen, 1995).

Par ailleurs, les aspects précis de cette décomposition morphémique n'ont pas clairement été identifiés (Marelli et al., 2020). Nommément, les morphèmes étant définis comme une forme orthographique liée à une unité de signification, la question de la prédominance des aspects morpho-orthographiques versus morpho-sémantiques dans le traitement morphologique précoce se pose. Elle est notamment critique pour établir leurs rôles respectifs au sein du processus d'identification des mots complexes : l'aspect « signification » du morphème y contribue-t-il, ou le morphème ne constitue-t-il qu'une unité orthographique pure, au même titre que le graphème ou la syllabe ? L'hypothèse « *form-then-meaning* » (e.g., Rastle et al., 2004) suppose une primauté de l'aspect orthographique aux premiers stades de l'identification des mots, des effets sémantiques n'étant retrouvés que dans un second temps (e.g., Beyersmann et al., 2016; Kazanina et al., 2008; Longtin et al.,

2003; Rastle et al., 2000, 2004). À l'inverse, l'hypothèse « *form-with-meaning* » (Diependaele et al., 2005, 2009) considère que les aspects sémantiques du morphème ont cours *durant* la décomposition morphémique (e.g., Feldman et al., 2012; Kazanina, 2011; Pyykkönen & Järviö, 2012). Des résultats empiriques cohérents avec ces deux hypothèses ayant été retrouvés via le paradigme d'amorçage, la question de l'influence des aspects sémantiques du morphème dans les traitements morphologiques demeure, encore aujourd'hui, sujette à débat (Beyersmann et al., 2016; Hasenäcker et al., 2016; Heathcote et al., 2018; Marelli et al., 2020).

3.2.2.2 *Impact de la morphologie dans la lecture à haute voix*

De manière complémentaire à ce champ de recherche, les études comparant la lecture de mots simples bisyllabiques (i.e., monomorphémiques, e.g., COULEUR) versus complexes bimorphémiques (e.g., DANSEUR), ainsi que les études comparant les non-mots simples (e.g., VANDIAR) à des non-mots morphologiquement construits (i.e., adjonction d'un affixe existant à une base existante ou inventée, e.g., VENDIER) mettent en évidence l'effet facilitateur d'une structure morphologique sur les capacités de lecture. De fait, lorsque les mots complexes et simples présentés étaient appariés sur des caractéristiques telles que leur fréquence lexicale, la fréquence de leur base et de leurs bigrammes/trigrammes, leur nombre de lettres et la taille de leur voisinage orthographique, les participants lisaient plus rapidement les mots comprenant plusieurs morphèmes que les mots n'en comprenant qu'un seul (Carlisle & Stone, 2005; Laxon et al., 1992; Quémart et al., 2012). De manière similaire, les non-mots qui contiennent de vrais morphèmes sont lus plus vite et avec davantage de précision (Burani et al., 2008; Colé et al., 2012; Traficante et al., 2011) que les non-mots contenant des suites de lettres appariées sur leurs caractéristiques phonologiques (e.g, phonème initial, structure syllabique) et orthographiques (e.g, taille du voisinage orthographique, fréquence des bigrammes). Il est intéressant de noter que la capacité des apprenants à lire les mots

complexes paraît liée à leurs compétences en CM (Carlisle, 2000; Kearns et al., 2016; Nagy et al., 2006), alimentant l'hypothèse d'un rôle causal de la CM sur les habiletés de lecture via un décodage morphologique (Levesque et al., 2021).

Les études longitudinales menées chez des apprenants de primaire, cependant, ne confirment pas systématiquement l'idée que la CM prédise positivement la précision ou la fluence de lecture des mots complexes (Deacon & Kirby, 2004; Manolitsis et al., 2017) lorsque d'autres facteurs tels que le QI et le décodage phonologique sont pris en compte. Plusieurs d'entre elles ont tout de même trouvé un effet prédicteur modeste, nonobstant significatif, de la CM sur le développement de la précision et de la fluence de lecture (Casalis & Louis-Alexandre, 2000; Desrochers et al., 2018; Kirby et al., 2012), y compris après avoir contrôlé les variables susnommées. Il paraît pertinent de noter qu'à ces occasions, aucun test utilisé pour mesurer ces variables d'intérêt n'était conçu pour évaluer particulièrement la lecture des mots complexes : ils en incluaient peu, ou bien de manière à refléter correctement leur proportion dans un texte réel. Ceci suggère que le bénéfice de la CM sur les mots complexes pourrait être suffisamment important pour améliorer globalement la lecture de n'importe quel corpus, du moment que ce dernier présente assez de mots plurimorphémiques ou pseudo-affixés (e.g., COULEUR dans lequel la suite de lettres EUR serait traité comme un suffixe).

De même que les entraînements phonologiques et phonographiques ont pu éprouver le caractère causal de la relation entre conscience phonologique/CGP et les habiletés de lecture, des études d'entraînement ont également investigué l'impact d'une intervention morphologique sur le développement de ces habiletés. Les méta-analyses réalisées par Goodwin & Ahn (2010, 2013) retrouvent une taille d'effet moyenne positive, modérée et significative sur la précision (nommée *decoding* par les auteures) mais pas sur la fluence de lecture. Néanmoins, ces résultats ne distinguant pas le caractère simple ou complexe des

mots, ni le caractère entraîné ou non-entraîné des stimuli, un apparent effet positif de la CM sur la précision de lecture a pu être porté par les mots complexes travaillés lors des entraînements. Ainsi, des études d'entraînement récentes ont investigué l'impact d'une intervention morphologique sur différents types de mots, et remettent en question les conclusions portées par ces méta-analyses. En effet, dans ces études, lorsqu'un effet positif d'une intervention morphologique sur l'identification des mots est observé, ce dernier concerne les tâches impliquant uniquement des mots (Colenbrander et al., 2021) ou non-mots (Jöbstl et al., 2021) complexes. Dans ces études, les résultats sur les mesures post-entraînement incluant des mots à la fois complexes et simples ne suggèrent pas d'effet significatif de l'intervention, que ce soit sur la précision ou la fluence de lecture. Il convient également de préciser que la plupart des études d'entraînement se centrant spécifiquement sur la CM par rapport à des études plus généralement métalinguistiques (i.e., incluant des entraînements phonographiques et/ou phonologiques au même titre que des entraînements morphologiques ; e.g., Kirk & Gillon, 2009) n'obtiennent de résultat significatif pour aucune des mesures de lecture proposées comparativement à une intervention contrôle (e.g., Abbott & Berninger, 1999; Apel & Diehm, 2014; Berninger et al., 2008; Georgiou et al., 2021). Nous réalisons une analyse plus complète de la question de l'impact des interventions morphologiques sur les capacités d'identification des mots, ainsi que sur les différents facteurs pouvant concourir à son efficacité (type d'apprenants et d'intervention contrôle, durée de l'intervention, etc.) dans le chapitre 5 du présent manuscrit.

En résumé, il apparaît que la présence de morphèmes dans les mots ou non-mots conduise à de meilleurs scores en lecture par rapport à des mots ou non-mots simples de même fréquence et aux propriétés orthographiques et phonologiques similaires (e.g., Quémart et al., 2012). Cette facilitation semble bien dépendre des capacités de CM (Colenbrander et al., 2021; Jöbstl et al., 2021), comme le suggéraient les études non-interventionnelles étudiant

la corrélation entre le développement des capacités de lecture et de CM (e.g., Kearns et al., 2016). Toutefois, la portée de la CM, et par conséquent des patterns morphologiques, semble cantonné, du moins pour la lecture, aux mots complexes.

3.2.3 Morphologie et orthographe : l'invariance morphémique

En ce qui concerne la production orthographique, de nombreuses observations suggèrent que les patterns morphologiques pourraient jouer un rôle essentiel dans ce processus, via la CM. Les apprenants semblent en effet capables de déduire l'orthographe de mots inconnus dès la deuxième année de primaire en se référant à la famille morphologique du mot (Pacton et al., 2013, 2018) et en identifiant les morphèmes le composant, qui tendent à toujours être écrits de la même manière en dépit d'éventuelles modifications de prononciations (i.e., invariance morphémique ; Bowers & Bowers, 2017; Burton et al., 2021; Landerl & Reitsma, 2005). Par ailleurs, dans les langues inconsistantes, et plus particulièrement en français, la prise en compte des patterns morphologiques permet de régulariser les CPG en position finale (Peereman et al., 2013), cette dernière correspondant à la fois à la principale localisation des inconsistances (e.g., le phonème [o] ne peut s'orthographier que O, EAU, Ô ou AU en milieu de mot, tandis qu'il peut être transcrit ÔT, OT, AUD, AUX, etc. en fin de mot), et l'endroit où l'information morphologique tend à être la plus fréquemment transcrite.

Les études longitudinales conduites à la fois dans des langues inconsistantes (Deacon et al. 2009) et consistantes (Grigorakis & Manolitsis, 2021) suggèrent que les capacités en CM des apprenants prédisent le développement des compétences en production orthographique, même après régression des variables contrôles appropriées (e.g., QI, décodage phonologique). Les mesures utilisées portaient à la fois sur des mots complexes et simples, suggérant que le bénéfice de la CM ne serait pas restreint aux mots plurimorphémiques. Cependant, de manière similaire à ce qui a été observé pour

l'identification des mots (cf. § 3.2.1), cet impact positif n'est pas systématiquement retrouvé (Pittas & Nunes, 2014).

Les méta-analyses de Goodwin et Ahn (2010, 2013) retrouvent un effet positif significatif des interventions morphologiques sur la production orthographique, de taille modérée-faible. Cependant, la taille d'effet moyenne calculée porte aussi bien sur la capacité de production orthographiques de mots entiers simples et complexes (e.g., Elbro & Arnbak, 1996), de morphèmes (e.g., Devonshire & Fluck, 2010), que de non-mots simples (e.g., Berninger et al., 2008), ce qui ne permet pas de déterminer si la CM bénéficie au processus d'écriture indirectement via l'amélioration du décodage (telle que suggérée par les études longitudinales), ou directement via l'invariance morphémique. Au vu des résultats des études plus récentes, la deuxième hypothèse semble devoir être privilégiée. En effet, la totalité des études à avoir, à notre connaissance, exploré les capacités de transcription phonologique suite à un entraînement morphologique, n'ont pas réussi à mettre en évidence une amélioration à ce niveau, comparativement à une intervention contrôle (Colenbrander et al., 2021; Taha & Saiegh-Haddad, 2016; Y. Wang & McBride, 2016). En revanche, un effet bénéfique de l'intervention sur la production orthographique des morphèmes est quasi-systématiquement retrouvé (Apel & Diehm, 2014 ; Casalis et al., 2017 ; Jöbstl et al., 2021 ; Zhang & Li, 2016 ; mais voir Colenbrander et al., 2021; Tsesmeli, 2017). Une exploration plus fine de la contribution de la CM au développement de la production orthographique est également proposée dans le chapitre 5 du présent manuscrit (méta-analyse de l'efficacité des entraînements morphologiques sur le développement de la littératie), et nous explorons, via le chapitre 6 (étude d'entraînement morphologique), la manière dont les effets d'une intervention portant sur l'orthographe des suffixes peut se généraliser à la production de mots complexes non travaillés.

En conclusion, contrairement à ce que laissaient supposer les études longitudinales (Casalis & Louis-Alexandre, 2000; Deacon et al., 2009; Desrochers et al., 2018; Grigorakis & Manolitsis, 2021; Kirby et al., 2012), le bénéfice de la CM pour la lecture et l'écriture semblent cantonnés aux mots complexes. Bien que ces derniers concernent, en théorie, une large majorité des mots de la langue française (environ 80% ; Rey-Debove, 1984), les études d'entraînement évaluant l'impact d'une intervention morphologique sur le développement de la lecture et de la production orthographique rapportent des tailles d'effet étonnamment assez modestes (tailles d'effet faibles à modérées ; effet réduit voire absent en-dehors des corpus présentant exclusivement des mots complexes).

3.3 Limites des patterns morphologiques

Ainsi, bien que les patterns morphologiques constituent, à l'image des patterns phonologiques, un aspect central du système de transcription des langues alphabétiques inconsistantes, leur utilisation comporte des limites. Même si la prise de conscience de la structure morphologique des mots écrits (i.e., conscience morphologique) peut conduire à de meilleures performances, notamment en identification des mots et en production orthographique (Fejzo et al., 2018), ce bénéfice est d'une part probablement restreint aux mots complexes, et d'autre part d'une taille d'effet peu importante (bien que significative). Il est également important de souligner que les régularités morphologiques, en particulier en ce qui concerne la dérivation, présentent elles-mêmes des inconsistances. En outre, la surgénéralisation des patterns morphologiques peut mener à des erreurs, plus particulièrement en production orthographique (e.g., RETARDER - RETARD mais CAUCHEMARDER - CAUCHEMAR ; Pacton & Afonso Jaco, 2015).

L'effet relativement modeste d'une intervention morphologique sur la lecture/écriture pourrait résider dans une dimension qui n'a, à notre connaissance, jamais été prise en compte dans les entraînements morphologiques : les contraintes graphotactiques de la langue. Tout

comme l'application des patterns phonologiques, l'application des patterns morphologiques repose sur l'acquisition de règles (i.e., enseignées telles qu'elles dans le cadre d'une instruction formelle) ou de régularités (i.e., intégrées incidemment par les apprenants eux-mêmes au contact de l'écrit), sous la forme d'affirmations telles que « les mots finissant par la rime /*et*/ sont transcrits -ETTE s'ils correspondent à un diminutif, e.g., VOITURETTE ». Elles sont supposément extraites par les apprenants afin d'être réinvesties dans n'importe quelle situation dans laquelle les règles/régularités peuvent s'appliquer (Manza & Reber, 1997). Toutefois, Pacton et al. (2005) ont démontré que le recours aux patterns morphologiques était conditionné par les propriétés distributionnelles du système de transcription. En effet, lorsque l'on demandait à des élèves de deuxième, troisième et cinquième année de primaire de transcrire des non-mots définis comme étant des diminutifs finissant par la rime /*et*/ (e.g., « une petite /*tRav*/ est une /*tRavet*/ »), l'emploi du suffixe -ETTE diminuait significativement si la consonne qui précédait la rime correspondait à une occurrence inhabituelle dans la langue (e.g., en français, une rime en /*et*/ n'est presque jamais transcrite ETTE si elle suit la lettre T) comme dans le non-mot /*tRatet*/. Ainsi, l'apprentissage orthographique des apprenants serait moins déterminé par leur capacité à extraire et/ou à appliquer des règles ou régularités que par leur sensibilité aux propriétés statistiques de la langue. Dans la théorie de l'IMP, cette dimension correspond aux patterns graphotactiques, que nous détaillons dans le chapitre suivant.

Chapitre 4 – Les patterns graphotactiques

Dans la théorie de l'IMP, les patterns graphotactiques sont définis comme transcrivant non pas une dimension linguistique, mais plutôt la « forme externe de l'écriture » (Treiman, 2017b, p. 273). Ils renvoient aux aspects purement distributionnels et probabilistes du système de transcription (e.g., les consonnes simples sont plus fréquentes que les consonnes doubles ; certaines consonnes ne doublent pas, certaines doublent moins que d'autres ; Danjon & Pacton, 2009).

4.1 Définition des patterns graphotactiques

4.1.1 Compétences orthographiques sublexicales

Lors de la formalisation de l'hypothèse de l'auto-apprentissage, Share (1995) évoquait déjà un possible « rôle secondaire » (p. 170) de compétences visuo-orthographiques dans l'apprentissage orthographique. Ces dernières, nommées compétences orthographiques (*orthographic processing skills*) par Stanovich et West (1989), renvoient à la « capacité de former, stocker et accéder à des représentations orthographiques » (p. 404) et participaient, après prise en compte des capacités de CGP, à une part significative de la variance dans la capacité de lecture des mots irréguliers ainsi qu'évaluée dans leur étude 2. Elles sont typiquement évaluées par une tâche de choix entre deux orthographes homophones d'un mot existant (e.g., ROOM – *RUME ; Olson et al., 1985). Néanmoins, ainsi que l'illustrent Castles et Nation (2006), une telle tâche se distingue difficilement du processus d'identification des mots lui-même, puisqu'elle implique un traitement lexical du mot, et donc une certaine forme de lecture par reconnaissance. Il est donc malaisé de considérer ces compétences orthographiques lexicales comme un prédicteur de l'apprentissage orthographique sans entrer dans une relation circulaire avec la lecture, du fait de la tâche habituellement retenue pour les évaluer.

Un autre sous-composant des compétences orthographiques, complémentaire à la première et d'ordre sublexical, a été parallèlement décrit. Il renvoie à la sensibilité aux régularités graphotactiques de la langue, mesurée par des tâches de choix entre deux non-mots, homophones ou non, se distinguant par la légalité (e.g., *EPPI – *EJJI ; Cassar & Treiman, 1997) ou la fréquence (e.g., *PUIRE – *PUIRRE ; Commissaire & Besse, 2019) de leurs séquences de lettres. Selon Treiman (2017a), l'intégration de ces régularités, à l'origine de la notion des patterns graphotactiques décrits dans l'IMP, commencerait à émerger avant même l'apprentissage du principe alphabétique. En effet, dès l'âge de 3 ans, 50% des enfants américains testés sont capables de tenir compte dans une certaine mesure des conventions du système de transcription, en ce qu'ils refusent d'admettre comme mots des propositions telles que HHHH, du fait de l'absence de variations dans les symboles utilisés (Ganopole, 1987). Par ailleurs, après seulement quelques mois d'instruction formelle en lecture, les apprentis-lecteurs de première année de primaire montrent une nette préférence pour des non-mots présentant des séquences légales versus illégales (Cassar & Treiman, 1997), et des séquences fréquentes versus peu fréquentes (Pacton et al., 2001).

Contrairement aux compétences orthographiques lexicales précédemment abordées, la sensibilité aux patterns graphotactiques est plus consensuellement considérée comme un processus distinct de celui de lecture et paraît constituer une mesure plus robuste des compétences orthographiques en général (Castles & Nation, 2006). Il est intéressant de noter que les compétences orthographiques lexicales et sublexicales (i.e., la sensibilité aux patterns graphotactiques) ne sont pas systématiquement corrélées. En effet, bien que des analyses factorielles suggèrent qu'elles sont sous-tendues par un facteur commun (Cunningham et al., 2001; S. H. Deacon et al., 2019; Hagiliassis et al., 2006), les coefficients de corrélation obtenus entre leurs mesures respectives pouvaient, selon les études de réplication, apparaître modestes (e.g., Commissaire et al., 2011; Conrad et al., 2013) voire non-significatives

(Connors et al., 2011). Commissaire et Besse (2019) ont évoqué la possibilité que de tels résultats contradictoires puissent découler de la construction des mesures de compétences orthographiques. En effet, il apparaît essentiel de considérer, au sein même d'une tâche de choix orthographique lexicale (e.g., choix entre FRAISE et *FREIZE) la fréquence des patterns sublexicaux qui constituent les propositions. Le cas échéant, une proposition pourrait être préférée du fait de la sensibilité du sujet aux patterns graphotactiques de sa langue, plutôt que par référence à une représentation orthographique spécifique stockée en mémoire (e.g., FRAISE pourrait être choisi plutôt que *FREIZE uniquement parce que le pattern -AISE est plus fréquent en français que -EIZE). De sorte, dans l'étude de Commissaire et Besse (2019), lorsque la tâche implique des cibles dont l'orthographe correcte présente des patterns sublexicaux peu fréquents, face à des leurres aux patterns sublexicaux plus fréquents (e.g., TREIZE – *TRAISE ; Lex OP1-LF), les résultats à cette tâche ne corrèlent pas significativement avec les tâches évaluant les compétences orthographiques sublexicales. À l'inverse, une corrélation entre compétences orthographiques lexicale et sublexicale est bien observée lorsque les cibles proposées dans la tâche lexicale possèdent des orthographe aux patterns sublexicaux plus fréquents que ceux présentés par les leurres (e.g., FRAISE – *FREIZE ; Lex OP1-HF) ou lorsque les patterns sublexicaux sont de même fréquence entre la cible et le leurre (e.g., HANTE – *HENTE ; Lex OP 2).

De manière similaire, la construction des tâches évaluant la sensibilité aux patterns graphotactiques pourrait distinguer les items présentant des séquences variant selon leur légalité (e.g., *ROUVE – *ROUVVE ; tâche SOP1 de Commissaire et Besse, 2019) et ceux variant selon leur fréquence (e.g., *TOIRE – *TOIRRE ; tâche SOP2 de Commissaire et Besse, 2019), puisqu'ils requièrent probablement différents degrés de sensibilité ou d'exposition au fonctionnement du système de transcription. Ainsi, dans l'étude de Commissaire et Besse (2019), une corrélation plus importante avec la tâche Lex OP 2 a été

retrouvée pour la tâche SOP 2 ($r = .36, p < .001$) par rapport à la SOP 1 ($r = .27, p < .01$), suggérant que la sensibilité à la fréquence des patterns graphotactiques serait plus en lien avec les compétences orthographiques lexicales que la sensibilité à leur légalité.

La question de la relation entre sensibilité aux patterns graphotactiques et lecture/écriture est également fondamentale à aborder, afin de déterminer dans quelle mesure ces patterns sont pertinents à considérer du point de vue de l'apprentissage orthographique. L'étude de Deacon et al. (2019) retrouve des corrélations significatives et modérées voire fortes entre cette sensibilité et l'identification des mots (coefficients de corrélation allant de .53 à .76). La même chose est constatée pour la production orthographique dans l'étude de Conrad et al. (2013). Ces deux études paraissent donc répliquer les premiers résultats de Cunningham et al. (2001) qui suggéraient qu'il existait une corrélation significative entre compétences orthographiques sublexicales et lecture/écriture.

Concernant le sens de cette relation, il est communément admis que les compétences orthographiques lexicales et sublexicales s'acquièrent via le fait de lire (Cunningham & Stanovich, 1990; Deacon et al., 2012; Leslie & Shannon, 1981) ; toutefois, ceci n'a pas, à notre connaissance, été testé expérimentalement, par exemple en contrôlant l'exposition à la lecture de pré-lecteurs sur une période donnée. L'analyse factorielle confirmatoire (*confirmatory factor analysis*, CFA) réalisée par Deacon et al. (2019), sur la base de données longitudinales, suggère que les facultés d'identification des mots et les compétences orthographiques lexicales et sublexicales sont sous-tendues par un même facteur commun, qui se distingue des capacités d'apprentissage orthographique. Dans le modèle retenu, le facteur « identification des mots/compétences orthographiques » prédit *et* est prédit par les capacités d'apprentissage orthographique ($\beta = .62, p < .001$). L'identification des mots et les compétences orthographiques ultérieures (T2) sont déterminées à la fois par le niveau de compétences initial (T1) en « identification des mots/compétences orthographiques » ($\beta =$

.91, $p < .001$), et par les capacités initiales (T1) d'apprentissage orthographique ($\beta = .10$, $p < .001$). Les capacités d'apprentissage orthographique ultérieures (T2), toutefois, semblent uniquement prédites par les capacités d'apprentissage orthographique initiales ($\beta = .77$, $p < .001$). En somme, le modèle basé sur les analyses de Deacon et al. (2019) confirme qu'il existe des liens très forts entre lecture et compétences orthographiques, puisqu'elles apparaissent sous-tendues par un même facteur ; cependant, le modèle n'offre pas de support à l'hypothèse que l'apprentissage orthographique soit influencé par la sensibilité aux patterns graphotactiques. Afin de déterminer si l'intégration de ces patterns pourrait constituer, comme Treiman (2017a) le suggère, un facteur d'importance dans l'apprentissage orthographique, au même titre que les patterns phonologiques et morphologiques, nous décrivons les quelques études d'entraînement existantes dans la partie 4.2.

4.1.2 Connaissance orthographique

Dans la littérature, la notion de compétence orthographique côtoie l'appellation « conscience orthographique » (*orthographic awareness* ; Berninger et al., 2010, p. 142), par analogie à la conscience phonologique et morphologique. La définition proposée par Berninger et al. (2010) de cette dernière correspond à la capacité de mémoriser la séquence de lettres d'un mot présenté visuellement. Dans les tâches évaluant cette conscience orthographique, le sujet doit déterminer si un nouvel item est identique à un item précédent, si l'item présenté comprenait certaines lettres ou non, ainsi que l'ordre dans lequel s'enchaînaient ces lettres (Berninger, 1987). L'auteure précise que cette définition englobe également les processus impliqués dans les tâches de choix entre deux orthographes possibles (e.g., RENTE – *RANTE) ainsi que les « connaissances orthotactiques des régularités de l'écriture, y compris les séquences de lettres probables et les positions de lettres spécifiques » (p. 143). Ainsi, dans cette acception, la conscience orthographique rejoint presque exactement la notion de compétence orthographique abordée dans le point 4.1.1, et les études

citées par Berninger et al. (2010) pour décrire les liens entre conscience orthographique et lecture/écriture sont identiques. Il convient toutefois de noter que les partisans de cette appellation ont rarement recours à l'aspect sublexical de la conscience orthographique lorsqu'ils conçoivent des entraînements fondés sur cette définition. Les programmes d'intervention proposés se rapportent davantage à des instructions de type « mot entier » (e.g., Berninger et al., 2008), dans lesquelles c'est en réalité la mémorisation de la représentation orthographique lexicale dans son ensemble qui est ciblée. De ce fait, afin d'éviter des confusions d'interprétation, nous avons préféré nous centrer sur les protocoles compatibles avec la description réalisée par Treiman et collaborateurs (Treiman, 2017a, 2017b; Treiman & Kessler, 2014), qui situe clairement les patterns graphotactiques à un niveau sublexical.

La notion de « conscience orthographique » (parfois également nommée « connaissance orthographique » ; Apel, 2011) est toutefois essentielle à aborder, en raison de la confusion dont elle est à l'origine dans la littérature. En effet, la nature extrêmement polysémique du terme « orthographe » et sa capacité à englober à la fois les compétences orthographiques lexicales et sublexicales conduisent à considérer comme « orthographique » tout entraînement ne pouvant être qualifié purement de phonologique ou morphologique, mais empruntant soit à ces deux dimensions linguistiques (e.g., Darch et al., 2006), soit se référant à un protocole exhaustif, traitant aussi bien des CGP que du vocabulaire, de la compréhension écrite et de la fluence (e.g., Calhoun et al., 2010; Given et al., 2008). C'est ainsi que dans leur méta-analyse, Galuschka et al. (2020) catégorisent comme orthographiques les entraînements réalisés dans les études citées, quand bien même ils donnaient la définition suivante de la connaissance orthographique (*orthographic knowledge* ; p. 3) :

Compréhension du système de règle orthographique permettant l'écriture correcte selon les règles et les patterns du langage écrit. Les interventions favorisant les compétences orthographiques se concentrent principalement sur les régularités graphotactiques et phono-orthographiques. Les enfants doivent apprendre les contraintes et les règles orthographiques qui gouvernent les positions, combinaisons et prononciations résultantes des lettres au sein des mots.

Le recours au terme générique « orthographique » rend ainsi très difficile l'identification d'études d'entraînement centrées sur l'impact des patterns graphotactiques, et non sur l'orthographe dans ses dimensions phonologiques et morphologiques.

4.2 Le rôle des patterns graphotactiques dans l'identification des mots et la production orthographique

4.2.1 « Apprenabilité » des patterns graphotactiques

Les patterns graphotactiques étant, dans la théorie de l'IMP, les seuls patterns à ne pas transcrire une dimension langagière, il est légitime de se demander si de telles régularités sont à même de faire l'objet d'une verbalisation pouvant être investie comme stratégie d'apprentissage orthographique. Cette question a été explorée via les expériences 1 et 3 de l'étude de Singh et al. (2021). Dans cette étude, 70 apprenants anglais de première année de primaire (35 participants dans chaque expérience) étaient exposés à un apprentissage graphotactique qui était soit implicite, soit explicite. Dans les deux modalités, les non-mots étaient les mêmes et possédaient une structure consonne-voyelle-consonne. La première consonne était toujours D, G, M ou R. En fonction de l'identité de la voyelle, la seconde consonne pouvait donner lieu à un doublement. Ainsi, quand le non-mot comprenait la voyelle U, la consonne qui suivait (F, L, S ou T) n'était jamais doublée. À l'inverse, quand la voyelle du non-mot était E, la consonne qui suivait était systématiquement doublée (FF, LL, SS ou TT). Selon ces règles, DUF était donc un non-mot légal, tandis que DUFF ne l'était

pas¹⁷. Dans l'expérience 1, les enfants étaient ainsi exposés à plusieurs non-mots suivant ces patterns, mais ces derniers ne leur étaient pas explicités. Dans l'expérience 2, avant de commencer l'apprentissage graphotactique, les enfants étaient au contraire avertis des patterns de formation des non-mots comme édictés précédemment. À l'issue de l'exposition aux non-mots, les enfants devaient réaliser deux types de tâches visant à déterminer si un apprentissage graphotactique avait eu lieu. La première tâche consistait en une épreuve de choix de la bonne voyelle : un squelette consonantique leur était présenté et ils devaient déterminer, entre E et U, quelle voyelle convenait d'être ajoutée pour composer un non-mot acceptable selon ce qu'ils avaient vu précédemment (e.g., S_FF [E]) ; les non-mots ainsi formés correspondaient à des items non-entraînés. La seconde tâche consistait en une épreuve de jugement de légalité : les enfants devaient déterminer si des non-mots n'ayant pas été présentés durant la phase d'apprentissage respectaient le pattern habituel. Cette étude a ainsi révélé que le fait d'être conscient des règles de formation des non-mots avant d'y être exposé conduisait à des performances plus élevées, et ce pour la tâche de choix ($z = 3.83 ; p < .001$) comme pour la tâche de jugement ($z = 2.19 ; p < .03$).

L'étude de Singh et al. (2021) suggère donc que les apprenants seraient capables, dès la première année de primaire, de tirer profit d'une explicitation des patterns graphotactiques. Toutefois, la question de l'intérêt d'une telle approche dans le cadre de l'apprentissage orthographique de vrais mots reste à traiter. À notre connaissance, il n'existe qu'une seule étude d'entraînement ayant fait l'objet d'une publication internationale à ce sujet : l'étude d'Ise et Schulte-Körne (2010). D'autres études publiées en langue allemande semblent lui préexister (Faber, 2005; Schulte-Körne et al., 2001, 2003), mais leur contenu exact est inaccessible aux non-germanophones. Stanké et al. (2020) rapportent également le résultat d'un protocole d'entraînement en partie basé sur des régularités graphotactiques et réalisé

¹⁷ Ces règles étaient contrebalancées à travers les participants de chaque expérience.

auprès d'enfants bons et faibles orthographes de troisième année de primaire. Toutefois, en raison du manque important d'informations se rapportant aux participants (attrition ; mode d'allocation des sujets aux groupes ; nombre de sujets faibles/bons orthographes ; nombre de sujets dans les groupes intervention/contrôle) ainsi qu'aux analyses statistiques conduites, il ne nous est pas possible de décrire cette recherche dans le cadre de la présente thèse. Ainsi, du fait de la paucité de la littérature concernant les entraînements graphotactiques, nous avons choisi de nous intéresser particulièrement à l'étude d'Ise et Schulte-Körne (2010), dans la sous-section dédiée ci-dessous.

4.2.2 Étude d'Ise et Schulte-Körne (2010)

Cette étude décrit l'impact d'une intervention visant l'intégration de certains patterns graphotactiques de la langue allemande, proposée à des apprenants porteurs d'un diagnostic de dyslexie. Deux plans pré-test/post-test ont été réalisés : dans le premier plan (étude 1), 10 participants recevaient 15 séances d'entraînement individuelles hebdomadaires de 60 minutes, tandis que 4 autres participants constituaient un groupe sans intervention (il s'agissait d'enfants dont les parents avaient pris contact tardivement avec l'équipe de recherche, et qui ont été intégrés au groupe « traitement » de l'étude 2). Tous les participants étaient scolarisés en 5^{ème} année de primaire. Le second plan (étude 2) avait recours à un *design* traitement (n = 13) versus liste d'attente (n = 14) ; les participants diagnostiqués dyslexiques étaient scolarisés soit en 5^{ème} soit en 6^{ème} année d'instruction. Le nombre de séances était réduit à 12, mais leur contenu et intensité restaient identiques. Dans les deux plans pré-test/post-test évoqués, l'entraînement ciblait la transcription des voyelles courtes et longues selon les régularités graphotactiques de l'allemand, exprimées sous la forme d'un algorithme décisionnel. Par exemple, l'une des « règles » ainsi travaillées pouvait prendre la forme suivante : « si une voyelle courte est suivie par seulement une seule consonne au sein d'un même morphème, alors la consonne doit être doublée lors de la transcription (e.g.,

MANN [homme] » ; p. 22). Cet algorithme décisionnel était présenté aux apprenants sous une forme graphique évoquant un plan de métro. Avant que celui-ci ne soit abordé, les notions sous-jacentes (voyelle longue/courte, morphème, etc.) étaient préalablement travaillées.

L'impact de l'entraînement sur la lecture n'a été évalué que dans l'étude 1. L'épreuve utilisée consistait en un test standardisé (ELFE 1-6 ; Lenhard & Schneider, 2006) calculant un score global sur la base des scores combinés des subtests d'identification de mots et de compréhension de phrases et de textes. Étant donné le faible nombre de participants inclus dans le groupe contrôle dans l'étude 1 ($n = 4$), les auteurs ont choisi de réaliser un test t pour échantillons appariés uniquement pour le groupe d'intervention. Ce dernier met en évidence une amélioration entre le pré et le post-test ($t(9) = 3.13$; $p = .006$). Toutefois, du fait de l'impossibilité de réaliser des analyses statistiques sur les performances pré-test/post-test du groupe contrôle, il n'est pas possible de distinguer ce gain d'un simple effet développemental. Par ailleurs, en l'absence de détails concernant les subtests de l'épreuve d'habileté de lecture, une potentielle amélioration à celle-ci à la suite de l'entraînement graphotactique est, en tout état de cause, difficile à interpréter.

La production orthographique était évaluée dans les deux études, avant et après l'entraînement. L'épreuve consistait également en un test standardisé (RST ; Grund, 2003), sous la forme d'une dictée de 60 mots (seule l'orthographe correcte du mot entier était comptabilisée comme juste). Afin de constituer des échantillons plus nombreux, les auteurs ont décidé de combiner les groupes d'intervention/traitement et contrôle/liste d'attente¹⁸ des études 1 et 2, portant ainsi le nombre de sujets respectifs à 19 et 18 dans les analyses. Une ANOVA à mesures répétées a ensuite été réalisée, révélant un effet du temps ($F(1,35) = 44.35$, $p < .001$) et une interaction significative entre temps et groupe ($F(1,35) = 5.13$, $p =$

.03), suggérant que le groupe d'intervention a davantage progressé que le groupe contrôle entre le pré-test et le post-test. Il convient néanmoins de rappeler que le groupe contrôle consistait en une absence d'intervention, et non en un contrôle actif, ce qui ne permet pas de s'affranchir d'un potentiel effet Hawthorne ou placebo.

Dans l'étude 2, afin de s'assurer de l'effet spécifique de l'entraînement graphotactique, la connaissance des régularités travaillées était évaluée via un questionnaire (huit QCM et trois questions ouvertes). Une ANOVA à mesures répétées a mis en évidence un effet du temps ($F(1,23) = 5.68, p = .03$), indiquant que l'ensemble des participants ont obtenu des performances plus importantes lors du post-test. L'interaction entre le temps et le groupe d'entraînement a atteint de justesse le seuil de significativité ($F(1,23) = 4.37, p = .048$) ; le groupe « traitement » semble ainsi avoir davantage progressé dans l'acquisition des patterns graphotactiques par rapport au groupe « liste d'attente ». Ce résultat confirme que les patterns graphotactiques semblent pouvoir faire l'objet d'un apprentissage explicite. Toutefois, les auteurs précisent qu'il n'existait pas de corrélation significative entre l'amélioration au test de production orthographique et l'amélioration au questionnaire.

Plusieurs conclusions essentielles peuvent être tirées de l'étude d'Ise et Schulte-Körne (2010). Premièrement, bien que l'impact sur la lecture soit difficile à établir pour des raisons méthodologiques, cette étude suggère que la production orthographique, évaluée par un test standardisé d'orthographe sous dictée, bénéficie bien d'une intervention graphotactique. Ces résultats sont cependant modérés par le faible nombre de participants, limitant à la fois les interprétations des analyses paramétriques réalisées et l'extrapolation des résultats à la population cible. De plus, l'absence d'une part de groupe contrôle actif, et d'autre part de corrélation entre l'amélioration en dictée et à un questionnaire évaluant la connaissance des règles travaillées, met en doute le fait que les progrès constatés soient réellement spécifiques

¹⁸ Les 4 sujets du groupe contrôle de l'étude 1, qui ont fait partie du groupe « traitement » de l'étude 2, ont donc

à l'entraînement. En effet, le simple fait d'exposer des participants à une intervention en langage écrit a pu bénéficier au développement de leur production orthographique.

4.3 Perspectives de recherche concernant les patterns graphotactiques

En somme, l'étude d'Ise et Schulte-Körne (2010) suggère qu'une intervention ciblant l'intégration de patterns graphotactiques, tels que décrits par Treiman et collaborateurs dans la théorie de l'IMP (Treiman & Kessler, 2014), offre un complément intéressant aux approches phonologique et morphologique. Il existe toutefois des limitations importantes à cette conclusion. Tout d'abord, cette étude d'entraînement est la seule à notre connaissance à avoir été publiée sur le sujet, restreignant ces résultats positifs aux seules caractéristiques du cas présenté (langue allemande, enfants dyslexiques de 5^{ème} et 6^{ème} année de primaire, etc.), phénomène amplifié par le faible nombre de participants inclus. Concernant le gain en lecture, bien que le groupe d'intervention ait progressé entre le pré-test et le post-test, ce dernier n'est pas distinguable d'un effet développemental, et il n'est pas possible, du fait du choix de la mesure, d'estimer quelles composantes (précision ; fluence ; compréhension) ont réellement été améliorées. Enfin, l'absence de groupe contrôle actif et de corrélation entre gains en production orthographique / gains en connaissance des régularités graphotactiques n'a pas permis de s'assurer d'un effet spécifique de l'entraînement.

Ainsi, avec une seule étude d'entraînement explorant l'impact des patterns graphotactiques sur l'apprentissage orthographique, de nombreuses interrogations demeurent encore à traiter. Nous avons cherché, via l'étude présentée dans le chapitre 7 du présent manuscrit, à contribuer à la compréhension de la prise en compte de la dimension graphotactique dans la mise en œuvre d'un entraînement orthographique sublexical.

été ici uniquement comptabilisés dans le groupe « liste d'attente ».

Problématique

L'identification rapide et précise des mots, tout comme la production orthographique, sont sous-tendues par le bon développement du lexique orthographique (Burt & Tate, 2002; Reitsma, 1983). Ce dernier s'accomplit via l'apprentissage orthographique qui est, selon Share (1995, 1999), item-dépendant et essentiellement permis par l'acquisition de représentations orthographiques spécifiques, via un recodage phonologique. Toutefois, cette conception ne tient pas compte des régularités caractérisant le système de transcription, ni de la capacité des apprenants à les investir pour décoder, transcrire et encoder des mots inconnus. Ces régularités ont été formalisées dans la théorie de l'Intégration des Multiples Patterns (IMP, Treiman & Kessler, 2014) et passées en revue dans les chapitres 2 à 4 du présent manuscrit. Cet état des lieux a permis à la fois de justifier de l'intérêt, pour les apprenants, de prêter attention au fonctionnement de leur système de transcription, mais également de souligner les aspects qui gagneraient à être davantage explorés par le biais d'études d'entraînement, en particulier concernant les patterns morphologiques et graphotactiques.

La morphologie représente un aspect essentiel des systèmes de transcription opaques tels que le français et l'anglais, et concerne près de 80% des mots de leur lexique (Nagy & Anderson, 1984; Rey-Debove, 1984). Sa prise en compte permet notamment, en français, de régulariser singulièrement les inconsistances observables en fins de mots (Peereman et al., 2013), et la capacité d'identification des mots des sujets semble bénéficier significativement de la présence de morphèmes (e.g., Quémart et al., 2012), en lien avec leur faculté de les identifier et de les manipuler volontairement (i.e., conscience morphologique ; e.g., Kearns et al., 2016). Si l'intervention morphologique est dorénavant considérée comme efficace pour de nombreux aspects de la littératie, dont la précision, la fluence de lecture et la production orthographique (Goodwin & Ahn, 2010, 2013), certains éléments tendent à nuancer ce

constat. Premièrement, les synthèses se basent sur des études d'entraînement dans lesquelles la conscience morphologique ne constituait pas le principal « ingrédient actif » de l'intervention ; par ailleurs, les variables dépendantes choisies ne permettaient pas de déterminer si les progrès constatés concernaient uniquement les mots (morphologiquement) complexes ou s'ils pouvaient se généraliser à des mots simples. Enfin, les modalités de l'efficacité de l'intervention morphologique, notamment en termes d'opacité du système de transcription, demeurent pour la plupart inconnues.

Le chapitre 5, consistant en une méta-analyse des études d'entraînement morphologique publiées entre 1996 et 2021, visait à répondre à ces interrogations. Nous avons pour cela synthétisé les 173 tailles d'effets de 40 études indépendantes proposant une intervention centrée sur la morphologie, à l'attention d'enfants scolarisés de la maternelle jusqu'à la terminale, pouvant présenter divers profils (tout-venants ; en difficulté d'apprentissage ; dyslexiques ; apprenants d'une langue seconde), mais sans comorbidité sensorielle, intellectuelle ou neurologique. Nous avons calculé une taille d'effet moyenne pour onze variables dépendantes : la conscience morphologique, la précision de lecture, la fluence de lecture, la lecture morphologique (lecture de mots ou non-mots complexes uniquement), le décodage phonologique (lecture de non-mots), la compréhension écrite, la production orthographique de mots, la production orthographique de morphèmes, le recodage phonologique (transcription de non-mots), le vocabulaire et la conscience phonologique. Nous avons également exploré, via une analyse de sous-groupes, les variables modératrices se rapportant au *design* des études (type d'entraînement contrôle, unité d'intervention, type d'instructeur, durée de l'entraînement), aux caractéristiques des participants (type d'apprenants, niveau de classe), à la langue d'instruction (système de transcription, opacité du système de transcription) et aux caractéristiques de l'entraînement morphologique (type de morphologie enseignée, type d'exercice morphologique, informatisation de l'entraînement).

Cette méta-analyse faisant l'objet d'une soumission pour publication scientifique, nous en proposons un résumé long en français ainsi que le manuscrit en anglais.

Le chapitre 6 renvoie à une étude d'entraînement visant à déterminer, au sein d'un même *design*, l'impact d'une intervention morphologique sur l'identification des mots et sur la production orthographique. Les questions posées et les stimuli choisis pour y répondre étaient néanmoins distincts. Ainsi, concernant l'identification des mots, nous souhaitions préciser les potentiels gains d'un entraînement morphologique à des tests standardisés de lecture, qu'ils concernent des mots irréguliers et réguliers (de fréquence lexicale faible ou élevée), et de non-mots (aux patterns sublexicaux de haute versus basse fréquence). Cette investigation visait à estimer si, en-dehors de mesures n'impliquant que des mots complexes, la conscience morphologique pouvait avoir un impact sur l'identification des mots.

Concernant la production orthographique, la question portait sur le potentiel de généralisation des patterns morphologiques (score de production orthographique des suffixes entraînés au sein de mots non-entraînés), évaluée par une tâche expérimentale. L'entraînement a été proposé en petits groupes, sur le temps scolaire, à 27 élèves scolarisés de la 3^{ème} à la 5^{ème} année de primaire, de divers niveaux de lecture. Il se composait de deux parties : une partie de manipulation consciente de suffixes via des jeux, et une partie de transfert orthographique, dans laquelle les enfants produisaient les mots entraînés sous dictée avec *feedback* correctif immédiat. Il était, dans un premier temps, comparé à un entraînement aux mathématiques complété d'une session de dictée portant sur les mêmes mots que l'entraînement morphologique (n = 24). Dans un second temps, nous avons également inclus un groupe contrôle entraîné aux mathématiques seul (n = 17), afin d'exclure un effet placebo ou Hawthorne.

Afin de compléter et d'enrichir les investigations ciblant la morphologie, nous avons évalué l'impact sur la production orthographique et l'identification des mots (fluence et

précision) d'un entraînement orthographique sublexical, traitant à la fois de patterns phonologiques, morphologiques et graphotactiques. Du fait du caractère inédit d'un tel entraînement, nous avons souhaité explorer un nombre important de paramètres relatifs à sa mise en œuvre. **Le chapitre 7** présente ainsi une étude d'un entraînement métalinguistique ciblant des unités orthographiques sublexicales, proposé à 54 enfants scolarisés en 2^{ème}, 4^{ème} ou 6^{ème} année d'instruction. Les enfants recrutés présentaient soit des difficultés de lecture significatives avec ou sans diagnostic formel de dyslexie (n = 13), soit possédaient des capacités dans la norme (n = 41). Nous avons choisi de nous centrer sur une régularité en particulier : la transcription du phonème /o/ en fin de mot (EAU – O – AUX, etc.).

L'entraînement était proposé selon deux modalités : explicitation forte suivant les principes de l'enseignement explicite (Hughes et al., 2017) avec proposition d'un algorithme décisionnel sur le modèle de celui d'Ise et Schulte-Körne, 2010 (n = 27) ; versus explicitation faible, dite « implicite », sans présentation de l'algorithme (n = 27). Le *design* employé était double : d'une part, avons eu recours à une méthodologie classique de type pré-test/post-test évaluant les progrès en production orthographique (dictée de mots entraînés et non-entraînés) et l'identification des mots (test de l'Alouette-R) ; et d'autre part, à une méthodologie de type *Single-Case Design* (SCD ; Lane & Gast, 2014) explorant, pour 45 de nos sujets, leur trajectoire développementale en production orthographique (dictées de mots non-entraînés) suite à l'introduction de l'entraînement.

Contributions scientifiques

**Chapitre 5 – Vingt-cinq ans d'étude de l'impact des entraînements
morphologiques sur le développement du langage écrit : une méta-analyse**

5.1. Résumé long en français

5.1.1 Introduction

La conscience morphologique (CM) i.e., la capacité d'identifier et de manipuler les morphèmes (Carlisle et al., 2010) ou plus petites unités de sens dans les mots, est aujourd'hui considérée comme un élément essentiel contribuant au développement de la littératie. Des études longitudinales ont mis en évidence la valeur prédictive de la CM sur la précision (Desrochers et al., 2018; Diamanti et al., 2017; Kirby et al., 2012), la fluence et compréhension de lecture (Casalis & Louis-Alexandre, 2000; Deacon & Kirby, 2004; Kirby et al., 2012; Levesque et al., 2019; Manolitsis et al., 2017), le décodage phonologique (Diamanti et al., 2017; Kirby et al., 2012), la production orthographique (Desrochers et al., 2018; Diamanti et al., 2017) et le vocabulaire (Han et al., 2022; Sparks & Deacon, 2015).

Plusieurs revues (Carlisle et al., 2010; Ke & Zhang, 2021; Reed, 2008) et méta-analyses (Bowers et al., 2010; Galuschka et al., 2020; Goodwin & Ahn, 2010, 2013; Rueda-Sanchez & Lopez-Bastida, 2016) ont identifié les entraînements morphologiques (EM) comme constituant une intervention bénéfique pour de nombreux aspects de la littératie. Toutefois, il est important de noter que les effets positifs décrits ont été obtenus sur un corpus d'études dans lesquelles l'EM n'était pas nécessairement l'élément principal de l'intervention (e.g., Kirk & Gillon, 2009). Elle pouvait en effet par exemple être proposée en parallèle d'activités de conscience phonologique ou d'entraînements phonographiques, qui constituent tous deux des interventions efficaces pour le développement de la littératie (Ehri et al., 2001; Galuschka et al., 2014; Torgerson et al., 2019; Weiser & Mathes, 2011), ce qui limite la compréhension de l'effet indépendant de l'EM sur l'ensemble des variables investiguées. Il est également pertinent de signaler que, dans les revues et méta-analyses citées plus haut, les variables dépendantes choisies ne permettaient pas de différencier l'effet de l'EM sur les mots simples et sur les mots complexes, notamment en ce qui concerne la lecture et la

production orthographique. Il convient par ailleurs d'évoquer le fait que les EM menés dans des langues sémitiques (Dallaseh-Khatib et al., 2014; Ravid & Geiger, 2009; Taha & Saiegh-Haddad, 2016; Vaknin-Nusbaum & Raveh, 2019) ou non-alphabétiques (Chow et al., 2008; Packard et al., 2006; Wang & McBride, 2016; Wu et al., 2009; Zhang & Li, 2016) ont régulièrement été exclues des efforts de synthèses réalisés jusqu'à présent, en raison du fonctionnement de leur morphologie, éloigné des langues alphabétiques européennes telles que l'anglais. Ces études, combinées à celles conduites dans des langues alphabétiques européennes, permettraient pourtant de mieux comprendre l'impact différencié de l'EM en fonction des caractéristiques des langues, notamment la modalité du système de transcription (alphabétique ou non-alphabétique) ainsi que son opacité (transparente, intermédiaire ou opaque). Enfin, l'impact de certains aspects de l'EM en lui-même, tels que le type de morphologie enseignée (lexicale, flexionnelle ou mixte), d'exercices proposés (modalité d'enseignement explicite ou implicite ; avec ou sans support écrit ; avec ou sans contexte phrastique ; avec ou sans référence à l'étymologie...), ou l'informatisation de l'entraînement, demeurent à explorer.

L'objectif de cette méta-analyse était donc de répondre à ces questions, d'une part en calculant une taille d'effet (TE) moyenne pour onze variables en lien avec la littératie (cf. 5.1.2 Méthode) ; et d'autre part en réalisant des analyses de sous-groupes permettant d'estimer l'impact de différentes variables modératrices sur l'efficacité de l'EM, concernant les caractéristiques du *design*, des participants, de la langue d'instruction ou de l'EM.

5.1.2 Méthode

Nous avons inclus 40 études indépendantes répondant aux critères de sélection suivants : être publiée entre 1996 et 2021 ; être rédigée en anglais, français ou espagnol ; avoir comme objectif d'améliorer un aspect de la littératie par le biais d'un EM dans lequel l'intervention morphologique représente au moins les 2/3 de l'entraînement ; présenter un

design pré/post-test comprenant au moins un groupe contrôle et figurant la moyenne (brute), l'écart-type et le nombre de participants nécessaires au calcul de l'indice de TE ; inclure des participants scolarisés de la maternelle (*preschool*) à la douzième année d'instruction sans condition particulière excepté celle d'être un apprenant de langue seconde ou de présenter un retard / trouble de la lecture. Nous avons privilégié les mesures issues de tests standardisés par rapport à des tâches expérimentales, ainsi que celles réalisées sur des items non-entraînés par rapport à celles réalisées sur des items entraînés. Lorsque plusieurs mesures étaient disponibles pour une même variable, nous choisissons la mesure possédant le meilleur indice de fiabilité inter-item. Des TE moyennes ont été calculées pour les variables de CM, de précision de lecture, de fluence de lecture, de lecture morphologique (i.e., identification de mots ou non-mots complexes), de décodage phonologique (i.e., lecture de pseudomots), de compréhension écrite, de production orthographique de mots, de production orthographique de morphèmes, de recodage phonologique (i.e., transcription de pseudomots), de vocabulaire et de conscience phonologique. Elles ont été calculées sur la base d'une différence moyenne standardisée (*standardized mean difference*, SMD) selon les recommandations Morris (2008) pour les *designs* pré/post-tests. Nous avons calculé la TE moyenne à la fois selon un modèle à effet fixe (i.e., ne prenant pas en compte la variabilité inter-études) et un modèle à effet aléatoire.

Les caractéristiques des participants, du *design* et de l'EM ont été rassemblées en catégories afin de réaliser des analyses de sous-groupes. Nous avons établi 4 grandes catégories de variables regroupant les caractéristiques relatives au *design*, aux participants, à la langue d'instruction ou à l'EM. La catégorie « caractéristiques du *design* » comprenait le type de groupe contrôle (modalités : instruction usuelle ; entraînement phonologique / phonographique ; entraînement au vocabulaire ; entraînement orthographique ; entraînement à la lecture ; autre traitement alternatif linguistique ; autre traitement alternatif non-

linguistique), la taille d'administration de l'intervention (modalités : individuelle ; groupe inférieur à 12 ; groupe de 12 ou plus), le type d'instructeur (modalités : chercheur ; non-chercheur) et la durée de l'entraînement (modalités : moins de 10 heures ; entre 10 et 15 heures ; 15 heures ou plus). La catégorie « caractéristiques des participants » comprenait le type d'apprenants (modalités : apprenants ordinaires ; apprenants de langue seconde ; apprenant à risque de présenter un trouble de la lecture/écriture ; apprenant avec un trouble de la lecture/écriture) et le niveau de classe (modalités : maternelle/première année d'instruction ; deuxième/cinquième année d'instruction ; sixième/neuvième année d'instruction ; dixième/douzième année d'instruction). La catégorie « caractéristiques de la langue d'instruction » comprenait le type de système de transcription (modalités : alphabétique ; non-alphabétique) et l'opacité du système de transcription (modalités : transparent ; intermédiaire ; opaque). Enfin, la catégorie « caractéristiques de l'EM » comprenait le type de morphologie (modalités : lexicale ; flexionnelle ; mixte), le type d'exercice (modalités : enseignement implicite ; entraînement à la CM ; entraînement à la CM avec support écrit ; entraînement à la CM avec support contextuel ; entraînement exhaustif), et l'informatisation de l'entraînement (modalités : informatisé ; non-informatisé). Les analyses de sous-groupes se basaient sur le calcul de l'hétérogénéité (Q de Cochran). Si la p -valeur du $Q_{between}$ (i.e., hétérogénéité au sein d'un même sous-groupe) était significative, nous considérons qu'il existait des différences de TE dans les modalités du sous-groupe analysé.

5.1.3 Résultats

À partir des 40 études incluses, nous avons obtenu 173 TE indépendantes sur l'ensemble de nos variables investiguées. Cinq de nos 11 variables ont obtenu des TE significatives, qu'elles soient positives (CM, production orthographique de morphèmes, production orthographique de mots, lecture morphologique) ou négatives (conscience

phonologique). Des analyses de sous-groupes ont été conduites sur les variables présentant une hétérogénéité inter-études significative ou tendancielle, i.e., la CM, la compréhension écrite (CE), la production orthographique de morphèmes (POM), le vocabulaire et la conscience phonologique (CP).

La TE moyenne pour la CM comprenait 35 tailles d'effet, dont un *outlier* (SMD = 5.34) issu de l'étude de Harris et al. (2011). Après exclusion de cette valeur aberrante, la TE moyenne pour la CM était de 0.34 ($p < .001$) pour le modèle à effet fixe (MEF) et de 0.48 ($p < .001$) pour le modèle à effet aléatoire (MEA), suggérant un effet positif, faible à modéré et significatif de l'EM. Les analyses de sous-groupes ont révélé 5 modérateurs significatifs pour la CM. Le premier de ces modérateurs concerne le type de groupe contrôle ($Q_{between}(6) = 35.67, p < .001$), impliquant que l'EM était significativement plus efficace pour améliorer la CM qu'un entraînement orthographique ($k_{TE} = 1$; SMD = 1.05), qu'un entraînement au vocabulaire ($k_{TE} = 5$; SMD = 0.73) ou qu'une instruction usuelle ($k_{TE} = 11$; SMD = 0.56) ; les autres types d'entraînement donnant lieu à une TE non-significatives de la part de l'EM (entraînement phonologique/phonographique ; entraînement à la lecture ; autre traitement alternatif linguistique ; autre traitement alternatif non-linguistique). Le modérateur concernant la taille d'administration du groupe d'entraînement ($Q_{between}(2) = 22.82, p < .001$) suggérait que l'EM était efficace pour améliorer la CM lorsqu'il était dispensé dans des groupes, qu'ils soient supérieurs ($k_{TE} = 18, SMD = 0.54$) ou inférieurs ($k_{TE} = 13, SMD = 0.53$) à 12 participants ; la TE concernant l'administration individuelle n'étant pas significative. Les modérateurs relatifs au système de transcription de la langue ($Q_{between}(1) = 8.11, p < .01$) et à l'opacité de la langue ($Q_{between}(2) = 15.93, p < .001$) a révélé que l'EM était efficace pour les langues alphabétiques ($k_{TE} = 31, SMD = 0.53$) opaques ($k_{TE} = 28, SMD = 0.64$) ; les TE pour les langues non-alphabétiques, ainsi que pour les langues intermédiaires ou transparentes n'étant pas significatives. Enfin, le dernier modérateur significatif pour la

CM concernait l'informatisation de l'entraînement ($Q_{between}(1) = 18.70, p < .001$), la TE étant significative uniquement pour les entraînements non-informatisés ($k_{TE} = 32, SMD = 0.52$).

La TE moyenne pour la POM comprenait 14 TE et était de 0.38 ($p < .001$) pour le MEF et de 0.41 ($p < .001$) pour le MEA, suggérant un effet positif, faible à modéré et significatif de l'EM. Le seul modérateur retenu pour la POM était marginalement significatif et concernait le type de groupe contrôle, suggérant que l'EM était plus efficace qu'un entraînement orthographique ($k_{TE} = 3, SMD = 0.62$) et qu'une instruction usuelle ($k_{TE} = 7, SMD = 0.45$) pour améliorer la POM, les autres types d'entraînement donnant lieu à une TE non-significatives de la part de l'EM (entraînement au vocabulaire ; entraînement à la lecture) ou n'étant pas représentés pour cette variable (entraînement phonologique/phonographique ; autre traitement alternatif linguistique ; autre traitement alternatif non-linguistique).

La TE moyenne pour la production orthographique de mots comprenait 28 TE et était de 0.19 ($p < .001$) pour le MEF et de 0.22 ($p < .001$) pour le MEA, suggérant un effet positif, faible et significatif de l'EM. La variance et l'hétérogénéité inter-études n'étant pas significatives, nous n'avons pas réalisé d'analyses de sous-groupes pour cette variable.

La TE moyenne pour la lecture morphologique comprenait 9 TE et était de 0.15 ($p < .01$), à la fois pour le MEF et pour le MEA, suggérant un effet positif, faible et significatif de l'EM. La variance et l'hétérogénéité inter-études n'étant pas significatives, nous n'avons pas réalisé d'analyses de sous-groupes pour cette variable.

La TE moyenne pour la CP comprenait 15 TE et était de -0.15 ($p < .05$) pour le MEF et de -0.09 ($p > .05$) pour le MEA, suggérant un effet négatif, faible et significatif de l'EM, uniquement lorsque la variabilité inter-études n'est pas prise en compte. Un seul modérateur significatif a été identifié pour la CP, et concernant l'opacité de la langue d'intervention ($Q_{between}(1) = 11.04, p < .001$), suggérant que l'EM était moins efficace que le groupe contrôle pour améliorer la CP pour les langues transparentes ($k_{TE} = 1, SMD = -0.72$), les langues

opaques donnant lieu à une TE non-significative et les langues intermédiaires n'étant pas représentées pour cette variable.

Les variables de CE et de vocabulaire n'ont pas donné lieu à une TE moyenne significative. Toutefois, la variabilité et l'hétérogénéité inter-études étant significatives pour ces dernières, des analyses de sous-groupes ont été réalisées. Ces dernières ont révélé une efficacité significative de l'EM sur la CE en fonction du type de morphologie entraîné ($Q_{between}(1) = 4.40, p < .05$) et du type d'exercice ($Q_{between}(4) = 15.24, p < .01$), i.e., lorsque l'EM abordait à la fois la morphologie lexicale et flexionnelle ($k_{TE} = 10, SMD = 0.31$), et lorsque les mots étaient présentés en contexte ($k_{TE} = 1, SMD = 1.44$). Concernant le vocabulaire, l'analyse de sous-groupes a révélé que le niveau de classe constituait un facteur modérateur significatif ($Q_{between}(2) = 9.39, p < .01$), une TE significative était obtenue pour cette variable uniquement lorsque l'EM était proposé à des élèves de 6^{ème}/9^{ème} année d'instruction ($k_{TE} = 2, SMD = 0.35$).

Les variables de précision de lecture, de fluence de lecture, de décodage phonologique et de recodage phonologique ont toutes donné lieu à des TE moyennes, des variance et des hétérogénéités inter-études non-significatives.

5.1.4 Discussion

Notre méta-analyse visant à étudier l'effet spécifique de la morphologie sur le développement de la morphologie a mis en évidence de nombreuses TE significatives sur diverses composantes de la littératie. Tout d'abord, l'impact sur la CM, bien qu'attendu, semble être modéré par plusieurs facteurs, à savoir la nature du groupe contrôle, la taille du groupe d'entraînement, le système de transcription, l'opacité de la langue et l'informatisation de l'entraînement. L'impact de l'EM sur l'identification des mots et la production orthographique semble dépendre de la complexité morphologique du matériel étudié. En effet, une TE significative a été retrouvée pour la variable de lecture morphologique (i.e., de

mots ou non-mots complexes) mais pas pour les variables de précision et de fluence de lecture ni de décodage phonologique. De même, une TE plus importante a été obtenue pour la POM, comparativement à la production orthographique de mots. Concernant la POM, l'analyse de sous-groupes suggère que l'EM serait plus efficace pour l'améliorer qu'un entraînement orthographique et qu'une instruction usuelle. En ce qui concerne les variables de CE et de vocabulaire, une efficacité de l'EM pourrait être retrouvée dans certaines conditions. Pour la CE, l'EM semblait efficace uniquement lorsqu'il abordait à la fois la morphologie lexicale et flexionnelle, ou lorsque les exercices proposés veillaient à présenter les mots au sein d'un contexte (phrase ou texte) ; et pour le vocabulaire, lorsque les apprenants étaient scolarisés de la 6^{ème} à la 9^{ème} année d'instruction. Quant à l'effet adverse de l'EM sur la CP (retrouvé uniquement lorsque la variabilité inter-études n'est pas prise en compte), celui-ci semblait porté par une étude investiguant l'impact de l'EM sur une langue transparente (Lyster, 2002).

En dépit de ces résultats, il convient de garder à l'esprit des limites intrinsèques à notre étude, en particulier concernant le recours aux analyses de sous-groupes, ces dernières étant par définition non-expérimentales, menées sur des catégories post-hocs figurant un faible nombre d'études. Par ailleurs, les TE moyennes calculées ne portaient que sur les pré-tests et post-tests immédiats ; ainsi, notre méta-analyse ne permet pas de conclure sur l'impact à long terme de l'EM sur le développement de la littératie.

Les analyses menées ont permis d'identifier des lacunes dans la littérature, notamment le faible nombre d'EM mené sur des apprenants de langue seconde, dans des langues non-alphabétiques et dont le contenu se focalisait sur la morphologie flexionnelle seule. Par ailleurs, nous n'avons étudié qu'une seule étude investiguant l'impact de l'EM sur la production écrite, et aucune ciblant les apprenants scolarisés de la 10^{ème} à la 12^{ème} année d'instruction.

5.2 Article soumis

Twenty-five years of studying the impact of morphological training on written language

learning: a meta-analysis.

Fontaa, C.^a, Commissaire, E.^a, Landré, L.^b. & Besse, A.-S.^a

^a Laboratoire de Psychologie des Cognitions, Université de Strasbourg

^b Laboratoire Icube, Université de Strasbourg

5.2.1 Abstract

Morphological training (MT) is widely considered as an effective way to improve both reading and spelling, as well as skills concurring to written language acquisition, such as morphological awareness, vocabulary and phonological awareness. However, its positive impact found in previous systematic reviews and meta-analyses was inferred from studies in which morphology was not always the main focus of intervention. Additionally, it is unclear whether these effects, especially regarding word identification and spelling, are restricted to morphologically complex words, or if they could benefit to more general reading and writing situations. In order to address these caveats, we conducted a meta-analysis of 40 individual MT studies published between 1996 and 2021, totalling 173 effect sizes (ES). Significant positive ES were obtained for the outcomes of morphological awareness, morpheme spelling, word spelling and morphological reading. A significant negative ES was found for the phonological awareness outcome. No significant ES was found for the remaining outcomes (reading accuracy and fluency, phonological decoding, reading comprehension, phonological recoding and vocabulary). Subgroup analyses found eight significant factors moderating the impact of MT (comparison group, intervention unit, grade, language transcription system, language opacity, type of morphology taught, type of morphological exercise and computerization of training). However, they were not consistently found across all of our outcome variables. Overall, MT was found to yield a significant impact on written language including reading and spelling, although more notably so for complex words. Future research is needed to further specify this impact in non-alphabetic languages and in high schoolers.

Keywords: morphological training, meta-analysis, morphological awareness, written language

5.2.2 Introduction

Morphological awareness (MA) is defined as the ability to identify and manipulate morphemes (Carlisle et al., 2010) which are the smallest units of meaning in words. For instance, the word *seller* contains two morphemes, distinguishable in form as well as in meaning: the base word *sell* and the suffix *-er*. Morphemes may play a syntactic role by signalling grammatical functions (e.g., *says*, *washed*, *dogs*). They also concur to the creation of new words via processes of derivation (e.g., a game – a *gamer*) and compounding (e.g., a *video-game*). Over the years, MA has increasingly been considered as a key aspect facilitating written language acquisition. Its positive impacts include the ability to decode words with precision, also known as reading accuracy (Desrochers et al., 2018; Diamanti et al., 2017; Kirby et al., 2012). They also concern more proficient aspects of reading, such as the ability to read at an appropriate speed and accuracy (i.e., reading fluency) and to extract meaning during reading (i.e., reading comprehension; Casalis & Louis-Alexandre, 2000; Deacon & Kirby, 2004; Kirby et al., 2012; Levesque et al., 2019; Manolitsis et al., 2017). A positive impact has also been noted on spelling (Desrochers et al., 2018; Diamanti et al., 2017), vocabulary (Han et al., 2022; Sparks & Deacon, 2015) and, rather counterintuitively, on the ability to decode nonsense words using only letter-to-sound mappings, referred to as phonological decoding (Diamanti et al., 2017; Kirby et al., 2012). Different underlying mechanisms may account for these benefits. For instance, accurate identification of morphemes could speed up word reading, as assessed by longitudinal studies examining the predictive value of MA on reading outcomes (Casalis & Louis-Alexandre, 2000; Desrochers et al., 2018; Kirby et al., 2012). MA also seems to support higher-level dimensions such as reading comprehension through the knowledge of a morpheme's meaning and use (Deacon et al., 2017; Levesque et al., 2019). Word spelling could also benefit from MA, as being able to identify a morpheme in a word (whether it is a base word or an affix) provides reliable orthographic indications, (e.g.,

spelling “*played*” versus “*playd*” (Pacton & Deacon, 2008). The effects on word decoding and spelling might be especially critical for opaque languages (Lee et al., 2022), as they are defined as languages containing many letter-to-sound inconsistencies (Seymour et al., 2003), which are often motivated from a morphological perspective (e.g., the inconsistent “tw” sequence in “two” is justified by its word family, such sequence appearing in “twenty”, “twins”, “twice”, etc.; Bowers & Bowers, 2017).

Considering these possible mechanisms, educational interventions drawing on these morphological skills have been designed, further referred to as morphological training (MT) studies. Reviews and meta-analyses conducted on these studies have so far unanimously concluded to the positive impact of this intervention on written language, both in skilled and less-skilled learners (Bowers et al., 2010; Carlisle et al., 2010; Galuschka et al., 2020; Goodwin & Ahn, 2010, 2013; Ke & Zhang, 2021; Reed, 2008; Rueda-Sanchez & Lopez-Bastida, 2016). Regarding meta-analyses more specifically, significant and positive effects were obtained for outcomes of MA and phonological awareness (Bowers et al., 2010; Goodwin & Ahn, 2010, 2013), which is the ability to analyze the sounds of a language (Høien et al., 1995). A positive impact on reading was also reported, whether it related to word identification accuracy (Goodwin & Ahn, 2013) or reading comprehension (Goodwin & Ahn, 2010; Rueda-Sanchez & Lopez-Bastida, 2016). However, MT repeatedly failed to demonstrate efficacy on reading fluency (Goodwin & Ahn, 2010, 2013). By contrast, spelling seemed to systematically benefit from MT (Galuschka et al., 2020; Goodwin & Ahn, 2010, 2013; Rueda-Sanchez & Lopez-Bastida, 2016), as did (Bowers et al., 2010; Goodwin & Ahn, 2010, 2013; Rueda-Sanchez & Lopez-Bastida, 2016).

Given the substantial heterogeneity between studies, and in order to more thoroughly understand the mechanisms underlying the effects of MT on these several outcomes, subgroup analyses have commonly been conducted. Some of these concern rather methodologi-

cal aspects of the training. For instance, in Goodwin & Ahn's first analysis (2010) that included studies carried out in several different languages, best mean effect sizes (ES) were obtained when training duration lasted at least 10 hours and for students with a speech or language impairment. In their 2013 analysis, focusing only on studies conducted in English, the only significant moderator was participants' grade, mean ES being largest for preschool/early elementary students. Interestingly, while Goodwin & Ahn (2013) found no significant differences between administering MT in isolation or as part of a diverse linguistic program, other reviews showed that studies focusing on MT alone obtained less significant (Goodwin & Ahn, 2010) and weaker (Bowers et al., 2010) general training effects. It was therefore inferred that MT is most efficient when it is integrated to other linguistic instruction elements. Critically, other very powerful training components to which MT was integrated, such as phonological training (Ehri et al., 2001; Galuschka et al., 2020) or phonics instruction (Galuschka et al., 2014; Torgerson et al., 2019; Weiser & Mathes, 2011), could have at least partially explained the training positive effects on their own. Consequently, we still do not know how strong the independent impact of MT is on written language learning, which calls for a re-examination of the effects of MT when its content is grounded (almost) solely on morphological dimensions.

It is also unclear to what extent MT could benefit word identification and spelling beyond complex words, i.e., words containing at least two morphemes (e.g. "seller"), as opposed to simple words (i.e., words containing only one morpheme; e.g., "flower"). Notably, spelling has been considered as a unitary variable in previously published meta-analyses. Differentiating between spelling of morphemes and spelling of whole words might give a more precise idea of the impact of MT on spelling skills. Similarly, reviews have explored decoding, fluency and occasionally phonological decoding individually (Goodwin & Ahn, 2010), but have not distinguished between reading of simple versus complex words. Alt-

though the latter supposedly represent around 80% of encountered words, at least in English and French (Nagy & Anderson, 1984; Rey-Debove, 1984), the impact of MT on word identification might be overestimated if reading of complex words is not assessed separately from reading of more diverse words. On the other hand, an effect that would generalise to all words would be a strong argument in favour of this type of training – the underlying processes of which still need to be clarified.

Additionally, a few questions identified by previous works remain unaddressed, one of the most compelling being whether MT is similarly efficient depending on language properties. Namely, non-alphabetic languages (i.e., languages which do not resort to an alphabet to transcribe written language), as well as Semitic languages, have been constantly excluded from most reviews and meta-analyses, due to differences in their morphology. Yet, in non-alphabetic languages such as Chinese (Ruan et al., 2018; Wei et al., 2014) and Semitic languages such as Arabic (Schiff & Saiegh-Haddad, 2018) and Hebrew (Vaknin-Nusbaum & Sarid, 2021), MA has also been identified as a crucial aspect of written language learning. We argue that, given the increasing number of linguistically varied MT studies, these languages should no longer be excluded from reviewing efforts. On the contrary, the inclusion of non-alphabetic languages is critical to determining whether the way a language is transcribed (i.e., language transcription system) represent a significant moderator to the impact of MT on reading and spelling. Additionally, the inclusion of Semitic languages may contribute to understanding the moderating effect of opacity, or inconsistencies in letter-to-sound mappings. It is unclear whether MT is as useful for languages such as Greek, German or Norwegian, which have almost a one-to-one match between letter and sound and are considered as transparent or intermediate languages (Seymour et al., 2003), as it is for English or French (Bowers & Bowers, 2017; Peereman et al., 2013), which are considered opaque. Semitic languages

are critical to this question, as they may alternatively be considered as opaque or transparent, depending on whether or not vowels are transcribed (Taha, 2016).

On a complementary note, investigating the extent to which the content of the MT protocol may moderate training efficiency has been mentioned as an important moderator to investigate, in order to further guide clinicians and researchers' practices (Carlisle et al., 2010). Relevant training characteristics may include the type of morphology taught during training (derivational, compositional or inflectional), the type of morphology exercise, and the computerization of training.

The present study

The aim of the present study was to systematically and quantitatively review findings published in the last twenty-five years in the field of MT. Noteworthy, there has been an increasing number of newly published articles investigating the efficiency of MT on various outcomes, especially between 2010 and 2020. Even though the last published meta-analysis exploring the impact of MT on written language skills is recent (Galuschka et al., 2020), it included only two studies published after 2010. We assume that a ten-year span arguably represents a timely moment to update our knowledge about MT.

More specifically, this meta-analysis aims to investigate the impact of MT on several written language outcomes as well as outcomes concurring to its acquisition: MA, reading accuracy, reading fluency, morphological reading, phonological decoding, reading comprehension, word spelling, morpheme spelling, vocabulary and phonological awareness. Importantly, we set out to include only studies featuring a substantial amount of morphological instruction (at least two-thirds of the training program; see our eligibility criteria in the Methods section for more details) in order to focus on the specific impact of MT per se. An additional goal was to determine whether the potential effect of MT varied according to the properties of the items, i.e. whether or not they involved complex words. We also investigated

the significance of moderating factors in terms of study design (comparison group, intervention unit, instructor, training duration), participants (grade level, type of students), language of instruction (alphabetic or non-alphabetic, degree of letter-to-sound opacity) and morphological instruction characteristics (lexical or grammatical morphology, type of morphological exercises, computerization of training). Although some of these factors have already been recognized as significantly modulating the effects of MT, such as training duration, type of students and grade, their effect has been shown to greatly vary from one meta-analysis to another. Moderating variables regarding language characteristics and training content remain to be explored.

5.2.3 Methods

5.2.3.1 Eligibility criteria

Before conducting our literature search, 10 eligibility criteria were established. First, studies would only be included in our review if they had been published between 1996 and 2021. Studies prior (e.g., Henry, 1988) or subsequent to this 25-year period were not included. Second, studies could be redacted in either English, French or Spanish, as the authors were proficient in all three languages. Third, they had to be designed to improve at least one language skill through lexical (derivational or compositional) and/or inflectional morphological instruction. Studies in which morphological instruction represented less than two thirds of the intervention program were not included (e.g., Kirk & Gillon, 2009) to ensure any training effect would be due to MT. This was assessed by proportionating the time or number of sessions allocated to morphological exercises out of the total duration of the training period. However, when morphological instruction represented less than two thirds of the intervention program, the study could still be included if the control group shared the exact same type and amount of non-morphological intervention (e.g., Abbott & Berninger, 1999; Wolter & Dilworth, 2014). Fourth, studies had to provide a substantial description of the training protocol

that would allow for the characterization of the type of morphological exercises. If studies mentioned a MT program but without any detail regarding its content, they were not included (e.g., Hurry et al., 2005). Fifth, they had to feature a pretest-posttest design disclosing raw values for means, standard deviations and number of participants in each group in order to calculate ES. Case studies and n-of-one trials (e.g., Davidson & O'Connor, 2019), as well as studies using adjusted means (e.g., Kieffer & Lesaux, 2012; Nunes et al., 2003), were therefore not included. However, when raw means and/or standard deviations were not available, we contacted the study's authors to try and obtain them. Sixth, studies had to feature at least one behavioural outcome measure; studies which only featured neuroimaging measures were not included (e.g., Richards et al., 2005). Seventh, they had to feature a control training condition, whether a waiting group, regular treatment, or an alternative treatment, in order to make sure training effects could not be explained by a simple retest effect or by a developmental effect. Studies without a control group (e.g., Apel et al., 2013), or in which the control group was not of the same nature as the experimental group (e.g., students of a different age group, as in Crosson et al., 2019) were not included. Eighth, they had to focus on schoolchildren from preschool to grade 12, whether typical achievers, second language learners (SLL), undifferentiated learners, at-risk of reading disability, struggling readers or diagnosed with a learning disability. Training studies carried out on adult subjects were not included (e.g., Amirjalili et al., 2018). To make sure populations would be relatively comparable, we also chose not to include studies focusing on children with comorbidities such as language (e.g., Good et al., 2015), intelligence, sensory (e.g., Bow et al., 2004), neurological or attentional impairments. Ninth, studies had to have been published in a peer-reviewed journal, in the proceedings of a colloquium, or to have been featured in a dissertation manuscript. Lastly, studies had to present original data in order to avoid potential duplicates (e.g., Lyster et al., 2016).

5.2.3.2 Selection process

Using online databases, we looked up keywords appearing in registered studies' title and/or abstract that targeted morphological training, by entering either these two terms or the closest synonyms for each. Although some training programs may have included morphological components without specifically referring to morphology in their title or abstract (e.g., by using terms such as “word families”, “etymology”, etc.), we decided that these studies were less likely to have made morphology the main element of their instruction content. Consequently, our search queries were, in English (for all databases except “these.fr”, “Archipel” and “Dialnet”): *morph** OR *morphem** AND *instruction** OR *intervention** OR *training*; in French (only for the “these.fr” and “Archipel” databases): *morph** OR *morphèm** AND *instruction** OR *intervention** OR *entraînement*; in Spanish (only for the “Dialnet” database): *morfol** OR *morfem** AND *instrucción** OR *intervención** OR *entrenamiento*.

Between December 2020 and December 2021, 6 databases were searched for eligible records: ERIC, Jstor, Pubmed, Sage Journals, ScienceDirect and Web of Science. We considered that these 6 databases offered a representative panel of the databases in the fields of education, educational psychology and rehabilitation research. Similarly to Goodwin and Ahn (2010, 2013), we chose not to focus solely on peer-reviewed works in order to try and avoid publication bias. We therefore additionally searched the E-Thesis Online Service (EThOS) database, as well as the Open Access Thesis and Dissertation (OATD) database, for defended and online-accessible doctoral dissertations. We also explored non-English sources in order to increase the potential number of studies that could be included in our meta-analysis.

French and Spanish being the other two languages in which we were proficient enough to conduct our search, we reviewed the French national university dissertation database “theses.fr”, the University of Quebec in Montreal’s “Archipel” dissertation database, and the

Spanish dissertation database “Dialnet”. During this time period, we also manually searched citations from previous reviews.

After the removal of duplicates, 5,101 individual articles and dissertations were screened for eligibility based on their title, either via database search or citation search. Titles that displayed keywords clearly indicative of ineligibility (e.g., medical interventions, review papers, etc.) were removed. In case of doubt, records were saved for the next step of the selection process. Ensuing title screening, 152 records were selected. We then carefully read their abstract looking for clear ineligibility indicators identical to title screening and identified 116 eligible publications. We then read the full text of those remaining records and systematically excluded studies that did not meet our 10 eligibility criteria. Seventy-seven studies were excluded, either because they did not feature pre-posttest measures as previously defined ($n = 24$), because they were not an intervention study ($n = 8$), because morphology was not the main component of the intervention program ($n = 28$), because their type of participants was invalid ($n = 9$) or because their control group was invalid ($n = 8$). Our final sample was therefore comprised of 39 publications and 40 studies, due to the publication by Devonshire and Fluck (2010) featuring two independent studies. Included studies are displayed in Table 1.

5.2.3.3 Coding of studies

Outcome variables were coded for each study following categories established by Goodwin & Ahn (2013) and are summarized in Table 1. We differentiated reading outcomes and spelling outcomes into several distinct categories for each type of variable. For reading outcomes, we differentiated between reading accuracy, reading fluency, morphological reading, phonological decoding and reading comprehension. For spelling outcomes, we differentiated between word spelling, morpheme spelling and phonological recoding. We also added a writing outcome variable to account for recent studies aiming to improve writing skills

through morphological training (e.g., Allen & Lembke, 2020). Post-test outcome measures used in the analysis were always the ones that first followed the end of the intervention. Therefore, no follow-up or mid-test measures were included in our analysis. When there were several grade or learner subgroups within one intervention group that the authors had not combined, we created different ES for each of those subgroups.

5.2.3.4 Defining the outcome variables

Correspondences between the outcome variables used in the present meta-analysis and the task used in the included publications are detailed in Tables S1 and S2 (online only).

Morphological awareness. We considered a task to be a morphological awareness (MA) task when the participant had to purposefully identify or manipulate morphemes in order to answer a question or solve a problem given by the examiner. The task could be proposed in the oral or written modality, but tasks measuring the spelling of morphologically complex words (i.e., words that contain at least two morphemes, e.g., seller) or morphemes were always considered to be spelling tasks, even if mentioned as an MA task in the original publication.

Reading accuracy. We considered a task to be a reading accuracy task when its purpose was to measure the participant's capacity to correctly read real words out loud, regardless of reading time. Consequently, we only included reading scores that did not feature a time measurement.

Reading fluency. We considered a task to be a reading fluency task when its purpose was to assess both the participant's reading speed and accuracy of real words or sentences. Typically, there were expressed as the number of correctly read words read aloud within a time limit. Although reading comprehension of short sentences may be considered an indication of reading fluency ability, they were always coded as reading comprehension tasks within this meta-analysis. Regarding the reading accuracy and fluency outcomes, we chose to

only include tasks featuring words that were not exclusively complex. Reading lists of words that only featured complex words were coded as morphological reading tasks.

Morphological reading. We considered a task to be a morphological reading (MR) task when it exclusively featured complex words (e.g., “seller”) or complex nonwords (e.g., “deller”). When both were available, we preferred nonwords because we considered that they were more likely to be read through morpheme decoding than real words. We also preferred fluency over accuracy indexes, as we considered that fluency index was a more precise indication of reading proficiency.

Phonological decoding. We considered a task to be a phonological decoding (PD) task when the participant had to apply grapheme-to-phoneme (i.e., letter-to-sound) conversion. For alphabetic languages, we only included simple nonword reading tasks, and for non-alphabetic languages, we only included pinyin reading tasks. Fluency scores were preferred whenever available. Tasks involving complex nonwords were included in the MR outcome.

Reading comprehension. We considered a task to be a reading comprehension (RC) task when the participant had to demonstrate comprehension of a sentence or text. RC could be assessed through selecting a correct picture, multiple choice quiz, true-or-false quiz, or rephrasing or synthesizing a given written material. Oral comprehension tasks were not considered to be RC tasks.

Word spelling. We considered a task to be a word spelling (WS) task when the participant had to spell dictated words. This category only included spelling measures of whole words as opposed to word parts (e.g., spelling of base words, affixes, etc.), which were coded as morpheme spelling tasks.

Morpheme spelling. We considered a task to be a morpheme spelling (MS) task when the spelling score was based on the spelling of individual morphemes as opposed to the

whole word. Spelling of complex nonwords were included in this category, as well as spelling scores of word parts (i.e., different spelling scores for base words and affixes).

Phonological recoding. We considered a task to be a phonological recoding (PR) task when the participant had to apply phoneme-to-grapheme (i.e., sound to letter) conversion, i.e., spelling non-words. Spelling tasks of complex nonwords were coded as MS tasks. Regarding spelling tasks in general, none of our included studies featured a spelling time for their measures. Therefore, all of our spelling outcomes tackled on spelling accuracy exclusively.

Writing. We considered as a writing task the production of a written sentence or text, evaluated in terms of quality or quantity of generated concepts or ideas. If the production was evaluated in terms of correctly spelled words, we considered it to be a WS task.

Vocabulary. We considered a task to be a vocabulary task when the participant had to demonstrate his or her knowledge of the meaning of a word orally or through writing, either by picture naming/selection, or by producing/selecting a definition for the presented word. Tasks in which the participant had to determine the meaning of a written sentence was considered a RC task. Tasks in which the participant had to determine the meaning of a target morpheme was considered a MA task.

Phonological awareness. Lastly, we considered a task to be a phonological awareness (PA) task when the participant had to purposefully identify or manipulate phonemes, syllables or rimes in order to answer a question or solve a problem given by the examiner. Only oral tasks were included in this category.

Studies occasionally assessed the same construct through the administration of several different measures. For instance, an initial-phoneme deletion task and a syllable counting task could both be used to measure the participants' PA skill (Lyster, 2002). We decided not to include all of the different measures of a same construct under the form of an averaged or

weighted score, as we considered that these measures could not always be merged together from a theoretical perspective. For example, although phoneme deletion and syllable identification are both proxies of PA, they were found to only moderately correlate with each other (Høien et al., 1995). In order to choose from all those several possible measures, we established a priori preferences. Firstly, for all outcome variables, alike Galuschka et al. (2020), we preferred standardized measures whenever they were available. A task was categorized as experimental if it was created by the authors, or if it was truncated, adapted or modified from an existing standardized test or subtest. Furthermore, we preferred tasks that had the best inter-item reliability. In the absence of such coefficient, we chose the one with the best inter-rater reliability. When reliability coefficients were not provided, and if only experimental measures were available, we preferred tests measuring performance on untrained material rather than on trained material. For the MA variable in particular, if reliability coefficients were not available, non-word material was preferred to word material, in order to minimize a potential lexical bias (i.e., a morphological task may be succeeded because the participant knows the word's meaning and not because he or she resorted to morphological analysis).

5.2.3.5 Labelling of categorical variables

Categorical variables were also coded in order to conduct subgroup analyses and are summarized in Table 1. Details regarding categorical variable coding are available in Tables S1 and S2 (online only). Again, we mostly mirrored the categories established by Goodwin and Ahn (2013). Regarding the *study design characteristics* category, we explored four variables:

- **comparison group:** regular treatment (waiting group or “business as usual”), phonology/phonics training, vocabulary training, orthographic training, reading-based training, other alternative linguistic treatment (e.g., study skills training) and alternative non-linguistic treatment (e.g., mathematics training). Type of alterna-

tive training was determined based on the objective described by the authors in the original article. When the article featured both an alternative treatment group and a regular treatment group, we always preferred the alternative treatment group in order to avoid a potential placebo or Hawthorne effect.

- **intervention unit:** individual, small group (less than twelve students), large group (twelve students or more).
- **instructor:** researcher or non-researcher (e.g., the child’s parent or usual teacher). If the professional who delivered the instruction was a certified teacher but only intervened in the school for the purpose of the study, we labelled them as “researcher”.
- **duration:** less than 10 hours of training (10 hours excluded), between 10 and 15 hours of training (15 hours excluded), 15 hours or more.

Regarding the *participant characteristics* category, we explored two variables:

- **type of students:** typical learners (which included undifferentiated learners as well as children specifically selected to be “average readers”), second language learner (i.e., multilingual students), at-risk of learning disability (including both students experiencing vulnerable schooling conditions impairing their ability to learn and students displaying significant weaknesses in reading/spelling but without any formal diagnosis of learning disability) and learning disabled (students formally diagnosed with specific learning disability in reading/spelling by a trained clinician following the diagnosis criteria in use in their own country or state).
- **grade:** preschool to grade 1 (or “beginning readers”), grade 2 to grade 5 (or “young readers”), grade 6 to grade 9 (or “confirmed readers”), grade 10 to grade

12 (or “advanced readers”) and broad grade (i.e., grades of participants crossed the listed ranges).

Thirdly, regarding the *language of instruction characteristics* category, we explored two variables:

- **language transcription system:** alphabetic and non-alphabetic. Regarding consonantal alphabet languages such as Hebrew and Arabic, we considered these languages to be “alphabetic”.
- **language opacity:** opaque, intermediary and transparent. These subtypes reuse previously published classifications for European languages (Landerl et al., 2013; Reis et al., 2020) inspired by Seymour et al. (2003). Opacity was not coded for non-alphabetic languages. Regarding Hebrew and Arabic, we considered these languages to be “opaque” when students were presented with regular transcription (i.e., no transcription of vowels) and “transparent” when the vowelized transcription was used (Taha, 2016); however, none of the included studies used vowelized transcription to our knowledge. The opacity of our included studies is indicated in Table 1 and is detailed in Table S1 (online only).

Lastly, regarding the *morphological instruction characteristics* category, we explored three variables:

- **type of morphology:** lexical (derivational or compositional morphology), grammatical (inflectional morphology) or both.
- **type of morphological exercise:** implicit learning (i.e., passive exposition to complex words), MA (i.e., exercises involving production, identification, segmentation or fusion of morphemes), MA with an emphasis on spelling (i.e., exercises focusing on orthographic properties of morphemes), MA with words presented in context (i.e., complex words were systematically presented within a sentence or a

text), and thorough training (a combination of the three explicit morphological exercises)¹⁹

- **computerization of the training program:** yes (computerized) or no (non-computerized).

Our coding of morphological exercises was based on the description of the training protocol made by the study's authors and is displayed in Table S3 (online only). Definitions and examples for exercises of each type are available in Table S4 (online only).

5.2.3.6 Statistical analyses

We conducted our statistical analyses with the free and open-source software RStudio. Our entire script is available upon request. First of all, publication bias was assessed through Egger's test (Egger et al., 1997) conducted for each outcome measure, using the function *metabias*, which is part of the *meta* package (Balduzzi et al., 2019). For the first part of the meta-analysis, we computed ES for each outcome measure. To that end, we used the function *metacont* which is also part of the *meta* package, following the steps described by Schwarzer et al. (2015). One of the possible ES indexes that can be used in *metacont* is Hedges' *g*, also known as Standardised Mean Difference (*SMD*), which we modified to follow the recommendations by Morris (2008) on pre-posttest designs, similarly to Goodwin and Ahn (2010, 2013). We therefore used a modified *SMD* – known as d_{ppc2} – for each outcome measure, within a single study. The used formula was the following:

$$d_{ppc2} = c_P \frac{MDe - MDc}{pooledSD_{pretest}}$$

where *MDe* is the difference between the post-test and the pre-test performance of the experimental group and *MDc* is the difference between the post-test and the pre-test perfor-

¹⁹ Those training categories were established after reading the entire corpus of studies, so that at least 3 studies could be represented in each category.

mance of the control group. c_P is the bias adjustment obtained with the following formula (Morris, 2008):

$$c_P = 1 - \frac{3}{4(n_t + n_c - 2) - 1}$$

where n_t is the number of participants in the treatment group and n_c is the number of the participants in the control group.

SMD are expressed in terms of *Z*-scores and are usually interpreted similarly to a Cohen's *d* (Lakens, 2013).²⁰ Mean *SMD* may be calculated either under the fixed-effects model or the random-effects model, producing two potentially different mean ES for each outcome. Both models weigh studies by taking into account intra-study variability in the computation of the mean ES. However, the fixed-effects model does not take into account inter-study variability,²¹ while the random-effects model does. This difference results in the fixed-effects model being generally less subject to bias due to under-powered studies, although ignoring between-study differences. It therefore does not quantify heterogeneity due to methodological differences in studies, which may be helpful when trying to determine the comparability of studies within the same research field or in determining sources of effect moderation (as in subgroup analyses). Considering both models therefore provide interesting perspectives about the impact of MT on written language skills, we decided to conduct our analyses under both models and to reflect on potential differences in mean ES.

For the second part of the meta-analysis, we conducted subgroup analyses, again using the function *metacont*. Subgroup analyses are used to determine whether significant heterogeneity for an outcome measure (or Q_{total}) may be explained by testing for moderators'

²⁰ Cohen (1988) suggested the following interpretations for *d*: 0.20-0.30 = small effect sizes, 0.50 = medium effect size and 0.80 = large effect size.

²¹ τ^2 was calculated using the DerSimonian-Laird estimator, with the assumption of separate between-study variances.

heterogeneity (or $Q_{between}$), i.e., differences in the moderators' modalities (Schwarzer et al., 2015).

5.2.4 Results

A total of 39 articles published between 1996 and 2021 were included in the meta-analysis, producing 40 independent studies. Thirty-three of these studies were published in peer-reviewed journals, 1 was published in the proceedings of a colloquium (Fejzo et al., 2019) and 5 were dissertation manuscripts (Bowers, 2006; Colchester, 2010; Frank, 2008; Kolne, 2019; Rassel, 2020). Only one study (Allen & Lembke, 2020) was found to feature a writing task, therefore we could not compute an averaged ES for the writing outcome. In total, 173 individual ES were obtained.

5.2.4.1 Publication bias

Funnel plots for all outcome measures are available in Figure S1 (online only). Egger's test was conducted for each outcome variable in order to statistically estimate publication bias. Risk of publication bias was significant for the MA ($t(32) = 2.33, p = .03$)²² and the MR ($t(7) = -2.45, p = .04$) outcomes. It was marginally significant for the RC ($t(14) = 1.86, p = .08$) and the WS ($t(26) = 1.42, p = .17$) outcomes. Egger's test was not significant for the remaining outcomes: reading accuracy ($t(11) = -0.15, p = .88$), reading fluency ($t(11) = -0.28, p = .79$), PD ($t(9) = -1.34, p = .21$), MS ($t(12) = 0.68, p = .51$), PR ($t(7) = 1.06, p = .32$), vocabulary ($t(13) = -0.72, p = .48$) and PA ($t(13) = 0.77, p = .45$).

In summary, 4 out of our 11 outcome variables were at risk of a publication bias: MA, MR, RC and WS. Although we specifically attempted to include dissertation theses and proceedings of colloquiums in order to avoid publication bias, we therefore recommend keeping this risk in mind when interpreting results for these four outcomes.

²² This calculation excludes the outlier study by Harris et al. (2011).

5.2.4.2 Meta-analysis

Mean *SMD*, number of ES, confidence intervals and study heterogeneity for our outcome variables are summarized in Table 2. Forest plots for each outcome variable are displayed in Figures S2 to S6 (online only).

Morphological awareness (MA)

A total of 35 ES were computed for the MA outcome. All included studies taken together, the MT group totalled 1,655 participants, while the control group totalled 1,578. Mean *SMD* for MA was noticeably different depending on the computational model: 0.39 [0.32; 0.46] for the fixed-effects model (small-to-moderate significant) and 0.61 [0.36; 0.86] for random-effects model (moderate-to-strong, significant). This suggests that important study variations may have impacted the computed mean *SMD* in the random-effects model. Notably, the random-effects model may have been less robust to the presence of an outlier ES: 5.34 [4.63; 6.05], obtained by Harris et al. (2011). Between-study variance ($\tau^2 = 0.48, p < .01$) and heterogeneity ($Q_{total}(34) = 361.20, p < .001$) were both significant. When the study by Harris et al. (2011) was removed, there were 1,586 participants in the MT group and 1,505 in the control group. Mean *SMD* was 0.34 [0.27; 0.41] (small-to-moderate, significant) and 0.48 [0.30; 0.66] (moderate, significant), respectively for the fixed-effects model and the random-effects model. Between-study variance ($\tau^2 = 0.21, p < .01$) and heterogeneity ($Q_{total}(33) = 173.07, p < .001$) remained both significant.

Reading accuracy

A total of 8 ES were computed for the reading accuracy outcome. All included studies taken together, the MT group totalled 245 participants, while the control group totalled 244. Mean *SMD* was 0.17 [-0.01; 0.35] for the fixed-effects model and 0.18 [-0.07; 0.43] for the random-effects model (very small, not significant). Between-study variance ($\tau^2 = 0.05, p = .11$) and heterogeneity ($Q_{total}(7) = 11.82, p = .11$) were both not significant.

Reading fluency

A total of 13 ES were computed for the reading fluency outcome. All included studies taken together, the MT group totalled 472 participants, while the control group totalled 450. Mean *SMD* was 0.04 [-0.09; 0.17] both for the fixed-effects model and the random-effects model, which equals to a very small, not significant ES. Between-study variance ($\tau^2 = 0.00$, $p = 1$) and heterogeneity ($Q_{total}(12) = 2.66$, $p = 1$) were both not significant.

Morphological reading (MR)

A total of 9 ES were computed for the MR outcome. All included studies taken together, the MT group totalled 710 participants, while the control group totalled 658. Mean *SMD* was 0.15 [0.04; 0.25] both for the fixed-effects model and the random-effects model (very small, significant). Between-study variance ($\tau^2 = 0.00$, $p = 1$) and heterogeneity ($Q_{total}(8) = 0.91$, $p = 1$) were both not significant.

Phonological decoding (PD)

A total of 11 ES were computed for the PD outcome. All included studies taken together, the MT group totalled 426 participants, while the control group totalled 407. Mean *SMD* was 0.01 [-0.13; 0.14] both for the fixed-effects model and the random-effects model (negligible, not significant). Between-study variance ($\tau^2 = 0.00$, $p = .58$) and heterogeneity ($Q_{total}(10) = 8.48$, $p = .58$) were both not significant.

Reading comprehension (RC)

A total of 16 ES were computed for the RC outcome. All included studies taken together, the MT group totalled 638 participants, while the control group totalled 609. Mean *SMD* was 0.10 [-0.02; 0.21] for the fixed-effects model and 0.16 [-0.03; 0.35] for the random-effects model (very small, not significant). Between-study variance ($\tau^2 = 0.08$, $p < .01$) and heterogeneity ($Q_{total}(15) = 36.16$, $p < .01$) were both significant.

Word spelling (WS)

A total of 28 ES were computed for the WS outcome. All included studies taken together, the MT group totalled 1,155 participants, while the control group totalled 1,139. The mean *SMD* was 0.19 [0.11; 0.28] in the fixed-effects model and 0.22 [0.11; 0.33] in the random-effects model (small, significant). Between-study variance ($\tau^2 = 0.01$, $p = .18$) and heterogeneity ($Q_{total}(27) = 33.59$, $p = .18$) were both not significant.

Morpheme spelling (MS)

A total of 14 ES were computed for the MS outcome. All included studies taken together, the MT group totalled 503 participants, while the control group totalled 507. The mean *SMD* was 0.38 [0.25; 0.50] in the fixed-effects model and 0.41 [0.24; 0.59] in the random-effects model (small-to-moderate, significant). Between-study variance ($\tau^2 = 0.04$, $p = .08$) and heterogeneity ($Q_{total}(13) = 20.81$, $p = .08$) were both marginally significant.

Phonological recoding (PR)

A total of 9 ES were computed for the PR outcome. All included studies taken together, the MT group totalled 274 participants, while the control group totalled 267. The mean *SMD* was -0.03 [-0.20; 0.14] in the fixed-effects model and 0.01 [-0.21; 0.19] in the random-effects model (negligible, not significant). Between-study variance ($\tau^2 = 0.02$, $p = .29$) and heterogeneity ($Q_{total}(8) = 9.60$, $p = .29$) were both not significant.

Vocabulary

A total of 15 ES were computed for the vocabulary outcome. All included studies taken together, the MT group totalled 1,055 participants, while the control group totalled 1,013. The mean *SMD* was 0.04 [-0.04; 0.13] in the fixed-effects model and 0.03 [-0.10; 0.15] in the random-effects model (negligible, not significant). Between-study variance ($\tau^2 = 0.02$, $p = .06$) and heterogeneity ($Q_{total}(14) = 23.24$, $p = .06$) were both marginally significant.

Phonological awareness (PA)

A total of 15 ES were computed for the PA outcome. All included studies taken together, the MT group totalled 440 participants, while the control group totalled 408. Mean *SMD* was -0.15 [-0.29; -0.01] (very small, significant) for the fixed-effects model, indicating an adverse effect of MT on this outcome. For the random-effects model, however, mean *SMD* was -0.09 [-0.37; 0.18] (very small, not significant) suggesting that this adverse effect does not hold when inter-study variability is considered. Between-study variance ($\tau^2 = 0.19$, $p < .01$) and heterogeneity ($Q_{total}(14) = 49.01$, $p < .001$) were both significant.

Conclusion on the effect of morphological training on written language outcomes

To sum up, significant and positive ES were obtained for MA (0.34 to 0.48, small-to-moderate), MR (0.15, very small), WS (0.19 to 0.22, small) and MS (0.38 to 0.41, moderate). One significant and adverse ES was obtained for PA (-0.15, very small; only significant when inter-study variability was ignored). No significant ES were obtained for reading accuracy, reading fluency, PD, RC, PR and vocabulary, with or without considering between-study variability.

5.2.4.3 Subgroup analyses

In order to try and identify sources of effect moderation, we conducted subgroup analyses on the 5 literacy outcomes that showed significant or marginally significant p-values for between-study variance (τ^2) and between-study heterogeneity (Q_{total}): morphological awareness²³ (MA), reading comprehension (RC), morpheme spelling (MS), vocabulary and phonological awareness (PA). Detailed subgroup analyses for each outcome measure are available in Appendices S1 to S5 (online only).

Study design characteristics.

Comparison group.

²³ We excluded the outlier study by Harris et al. (2011)

Comparison group was a significant predictor for MA ($Q_{between}(6) = 35.67, p < .001$) suggesting its modalities obtained significantly different ES. The strongest ES was obtained for studies which compared MT to orthographic training (SMD = 1.05 [0.54; 1.55], though only one study was featured under this modality), after which came vocabulary training ($k = 5, SMD = 0.73 [0.41; 1.06], \tau^2 = .07, I^2 = 55\%, Q_{within} = 8.88$) and regular treatment or instruction ($k = 11, SMD = 0.56 [0.31; 0.80], \tau^2 = .12, I^2 = 73.3\%, Q_{within} = 37.46$). MT was not found to yield significant ES on MA when compared to other comparison groups (phonology/phonics training, reading-based training, other alternative linguistic treatments, alternative non-linguistic treatments). Comparison group was a marginally significant predictor for MS ($Q_{between}(3) = 7.12, p = .07$): strongest ES was obtained for studies which compared MT to orthographic training ($k = 3, SMD = 0.62 [0.20; 1.05], \tau^2 = .07, I^2 = 51.8\%, Q_{within} = 4.15$), after which came regular treatment or instruction ($k = 7, SMD = 0.45 [0.25; 0.65], \tau^2 = .01, I^2 = 15.6\%, Q_{within} = 7.11$). MT was not found to yield significant ES on MS when compared to other available comparison groups (reading-based training or vocabulary training). No studies were featured under the “phonology/phonics training”, “other alternative linguistic treatment” and “alternative non-linguistic treatment” modalities. Comparison group was not a significant predictor for the remaining outcome variables (RC, vocabulary and PA).

Intervention unit.

Intervention unit was a significant predictor for MA ($Q_{between}(2) = 22.82, p < .001$) suggesting its modalities obtained significantly different ES. The strongest ES was obtained for studies which delivered MT in large groups ($k = 13, SMD = 0.54 [0.33; 0.75], \tau^2 = .09, I^2 = 66.8\%, Q_{within} = 36.15$), very closely followed by studies which delivered MT in small groups ($k = 18, SMD = 0.53 [0.20; 0.86], \tau^2 = .39, I^2 = 83.5\%, Q_{within} = 102.82$). MT was not found to yield significant ES on MA when delivered individually. Intervention unit was not a significant predictor for the remaining outcome variables (RC, MS, vocabulary and PA).

Instructor.

The type of instructor was not a significant predictor for any of the outcome variables.

Duration.

Duration was not a significant predictor for any of the outcome variables.

Participant characteristics.***Type of students.***

The type of was not a significant predictor for any of the outcome variables.

Grade.

Grade was a significant predictor for vocabulary ($Q_{between}(2) = 9.39, p < .01$) suggesting its modalities obtained significantly different ES. A significant ES was only obtained for studies which targeted students from grade 6 to grade 9 ($k = 2, SMD = 0.35$ [0.11; 0.58], $\tau^2 = .00, I^2 = 0.0\%, Q_{within} = 0.07$), indicating that MT did not obtain significant ES on vocabulary when it targeted other grade ranges (preschool to grade 1; grade 2 to grade 5). No studies were featured under the “grade 10 to grade 12” or the “broad grade” modalities. Grade was not a significant predictor for the remaining outcomes variables (MA, RC, MS and PA).

Language of instruction characteristics.***Language transcription system.***

Language transcription system was a significant predictor for MA ($Q_{between}(1) = 8.11, p < .01$) suggesting its modalities obtained significantly different ES. A significant ES was only obtained when it was conducted in alphabetical languages ($k = 31, SMD = 0.53$ [0.33; 0.73], $\tau^2 = .23, I^2 = 81.9\%, Q_{within} = 166.01$), indicating that MT did not obtain significant ES on MA when it was conducted in non-alphabetical languages. Language transcription system was not a significant predictor for the remaining outcomes variables (RC, MS, vocabulary and PA).

Language opacity.

Non-alphabetical languages were excluded from subgroup analyses of language opacity. Language opacity was a significant predictor for MA ($Q_{between(2)} = 15.93, p < .001$) suggesting its modalities obtained significantly different ES. A significant ES was only obtained when it was conducted in opaque languages ($k = 28, SMD = 0.64 [0.48; 0.80], \tau^2 = .10, I^2 = 58.4\%, Q_{within} = 64.96$), indicating that MT did not obtain significant ES on MA when it was conducted in intermediate or transparent languages. Regarding PA, language opacity was also a significant predictor ($Q_{between(1)} = 11.04, p < .001$). A significant, negative ES was only obtained when it was conducted in a transparent language ($SMD = -0.72 [-1.02; -0.43]$, though only one study was featured under this modality), indicating that for transparent languages, MT might have an adverse effect on PA. No studies were featured under the “intermediate” modality. Language opacity was not a significant predictor for the remaining outcomes variables (RC, MS and vocabulary).

Morphological training characteristics.

Type of morphology.

Type of morphology was a significant predictor for RC ($Q_{between(1)} = 4.40, p < .05$). A significant ES was only obtained when it targeted both types of morphology ($k = 10, SMD = 0.31 [0.02; 0.59], \tau^2 = .13, I^2 = 66.9\%, Q_{within} = 27.19$). MT did not obtain significant ES on RC when it targeted lexical morphology alone. Furthermore, no studies were featured under the “grammatical morphology” modality. Type of morphology was not a significant predictor for the remaining outcomes variables (MA, MS, vocabulary and PA).

Type of morphological exercises.

Type of morphological exercise was a significant predictor for RC ($Q_{between(4)} = 15.24, p < .01$). A significant ES was only obtained for MA+context ($SMD = 1.44 [0.74; 2.15]$, though only one study was featured under this modality), indicating that MT did not

obtain significant ES on RC when the content of the MT focused on other types of exercises (MA alone, MA+spelling, thorough training, implicit learning). Type of morphological exercise was not a significant predictor for the remaining outcomes variables (MA, MS, vocabulary and PA).

Computerization of training.

Lastly, computerization of training was a significant predictor for MA ($Q_{between}(I) = 18.70$, $p < .001$). A significant ES was only obtained when MT was not computerized ($k = 32$, $SMD = 0.52$ [0.33; 0.71], $\tau^2 = .22$, $I^2 = 78.1\%$, $Q_{within} = 141.31$), indicating that MT did not obtain significant ES on MA when training was delivered using some sort of computerization. Computerization of training was not a significant predictor for the remaining outcomes variables (RC, MS, vocabulary and PA).

Conclusion on subgroup analyses.

In conclusion, significant predictors were identified for 5 outcome variables. For morphological awareness, morphological training (MT) obtained significant effect sizes (ES) when it was compared to orthographic training ($SMD = 1.05$, though only one study was featured under this modality), vocabulary training ($SMD = 0.73$) and regular treatment or instruction ($SMD = 0.56$); when the intervention unit was a large group ($SMD = 0.54$) or a small group ($SMD = 0.53$); when the language transcription system was alphabetical ($SMD = 0.53$); when language opacity was opaque ($SMD = 0.64$); and lastly when training was not computerized ($SMD = 0.52$). For reading comprehension, MT obtained significant ES when it targeted both types of morphology ($SMD = 0.31$) rather than just lexical morphology, and when morphological exercises focused on MA+context ($SMD = 1.44$, though only one study was featured under this modality). For morpheme spelling, MT obtained significant ES when it was compared to orthographic training ($SMD = 0.62$) and regular treatment or instruction ($SMD = 0.45$), while no significant ES were obtained when compared to other types of

control trainings. For phonological awareness, MT was less effective than a control training when language opacity was transparent (SMD = -0.72, though only one study was featured under this modality). Lastly, for vocabulary, MT obtained significant ES when it was delivered to students in grade 6 to grade 9 (SMD = 0.35) rather than younger students.

5.2.5 Discussion

The aim of this meta-analysis was to examine the effect of MT as the main element of instruction on outcomes of written language, and to try to determine potential sources of effect moderation, especially ones that had not yet been investigated in previous meta-analyses (Goodwin & Ahn, 2010, 2013; Rueda-Sanchez & Lopez-Bastida, 2016). We included 40 independent studies totalling 173 different ES. Overall, MT resulted in moderate and positive ES for MA and MS. Small and positive ES were obtained for MR and WS. An adverse effect of MT was obtained on PA, only when inter-study variability was ignored. No significant ES were obtained for reading accuracy, reading fluency, PD, RC, PR and vocabulary. Eight significant predictors were also identified: comparison group, intervention unit, grade, language transcription system, language opacity, type of morphology, type of morphological exercise and computerization of training. However, those effect moderators were not consistently significant for all of our investigated variables. Our findings are discussed in more detail in the sections below.

5.2.5.1 Effects of MT on morphological awareness

Firstly, we found MT to have a small-to-moderate impact on MA once an outlier study (Harris et al., 2011) had been removed. Both Bowers et al. (2010) and Goodwin and Ahn (2010; 2013) had obtained ES of a similar value for their respective MA outcomes. Our subgroup analyses suggest that MT is significantly better at improving MA than orthographic training (though only one study by Casalis et al., 2017 was featured under that category), vocabulary training and regular treatment, in that order. However, it is unclear whether MA is better than

other types of control training such as phonological training or reading-based training. This lack of clear superiority may be explained by the close relationships between MA and PA, as well as between MA and reading. PA intervention has been found to result in positive gains for MA, to a degree (Casalis & Colé, 2009). To our knowledge, the effect of reading-based interventions on MA has not yet been investigated. However, there is evidence that reading development positively predicts growth in MA skills (Deacon et al., 2013; Kruk & Bergman, 2013). We also found that MT allowed for better improvement on MA when it was delivered in large or small groups (the difference between the two was extremely minimal) rather than individually, which appears rather counterintuitive, as practical syntheses tend to advocate delivering instruction in an individual manner (Galuschka et al., 2020; Galuschka & Schulte-Körne, 2016). However, a number of reviews on various types of trainings have similarly failed to obtain a significant advantage for individual sessions compared to group delivery (e.g., Goodwin & Ahn, 2010, 2013; Lou et al., 2001; Marulis & Neuman, 2013). Regarding the development of MA in particular, it would appear that some form of collective learning allows for better results, possibly because the student is exposed to a wider number of morphological productions than what can be provided in a one-on-one session. The nonsignificant effect of MT on MA in the case of individual sessions may also result from the existence of a confounding factor: training computerization. Indeed, computerized trainings were always administered individually in included studies for the MA outcome. Although Galuschka et al. (2020) have argued that computer-based instruction should be considered as a form of intervention unit, different from regular individual sessions, we chose to consider it as a type of moderator of its own (computerized training: yes or no) to investigate its specific independent influence on MT. Our rationale was that computerized training sessions would not necessarily always be individual (e.g., in the study by Colchester, 2010, training was displayed via a screen to a whole class). We found, similarly to Galuschka et al. (2020), that a

physical delivery should be preferred to a computerized one. However, as mentioned before, individual sessions also appeared to yield a weaker ES on MA than group sessions. Therefore, the superiority of non-computerized training could also be due to the fact that those trainings provided collective learning.

Lastly, we found MT to be most effective on MA for alphabetic languages with an opaque transcription system, a moderator that had not been addressed so far. This latter finding appears rather expected from a linguistic perspective, as opaque languages require learners to be aware of morphemic relationships between words in order to make sense of grapheme-to-phoneme inconsistencies: for instance, the silent g in “sign” is voiced in words pertaining to its morphological family, such as “signal” or “signature” (Bowers & Bowers, 2017). Although MA is predictive of written language learning in transparent and intermediate languages (Diamanti et al., 2017; Haase & Steinbrink, 2022), learners of these languages may benefit less from MT than learners of opaque languages. In addition, although the effect of MT on MA was mostly observed in alphabetic languages, and was nonsignificant for non-alphabetic ones, this issue remains to be explored further due to the small number of studies we were able to feature under the “non-alphabetic languages” category. We advocate for future research to focus on determining differential benefits of MT depending on the language transcription system.

5.2.5.2 Effect of MT on word identification

Our meta-analysis found no significant ES of MT for reading accuracy and fluency, when focusing on a diverse array of words. This result adds to the list of contradictory results from the various meta-analyses already published. While Rueda-Sanchez and Lopez-Batisda (2016) found moderate, yet significant impacts of MT on general word identification skills, null effects were reported by Galuschka et al. (2020). Examining reading skills in more detail, Goodwin and Ahn found positive impacts of MT on reading *accuracy* (labelled as “de-

coding” in their study) in their 2013 analysis, but not in their 2010 one. Nevertheless, neither their 2010 or 2013 analyses found that MT could have a positive impact on reading *fluency*. Overall, this piece of data suggests that there is still poor convincing evidence for a benefit of MT on general word identification. However, while previous meta-analyses did not consider complex words separately from other measures of word identification, we decided to compute another separate ES for tasks that specifically selected complex words or non-words, i.e., MR. Here, a small yet positive and significant ES was obtained, suggesting that MT may only impact reading of complex words specifically, possibly through morphological decoding (Levesque et al., 2021). The fact that MT only had a significant impact on the reading of complex words, which did not generalize to all types of words, is an important theoretical finding in the field. Indeed, it was assumed that MA could benefit to word reading beyond the decoding of complex words because longitudinal studies found significant contribution of MA to standardized measures of reading accuracy and fluency (Desrochers et al., 2018; Kirby et al., 2012) that were not designed to feature a majority of derived words. Further, a few of those longitudinal studies found positive impacts of MA on the reading accuracy of simple nonwords (Diamanti et al., 2017; Kirby et al., 2012), suggesting an effect on phonological decoding. This pattern was however not replicated in our analysis, since the effect of MT was nonsignificant on phonological decoding (PD). Considering none of our word identification outcomes yielded significant heterogeneity, it was not possible to further investigate moderating factors that may have provided leads of interpretation. All in all, our results suggest that MT may have a small impact on word identification skills, but only in words that are complex.

5.2.5.3 *Effects of MT on reading comprehension*

Regarding RC, no significant ES was found in our analysis. Again, this challenges findings from longitudinal studies that almost systematically found MA to predict the devel-

opment of RC (Casalis & Louis-Alexandre, 2000; Deacon & Kirby, 2004; Kirby et al., 2012; Levesque et al., 2019; Manolitsis et al., 2017). This benefit was however not as systematically replicated by meta-analyses of training studies. In Rueda-Sanchez and Lopez-Batisda (2016), the impact of MT on RC appeared moderate, while in Goodwin and Ahn's 2010 analysis and in Bowers et al. (2010), this impact was much smaller. No significant impact of MT was found in Goodwin and Ahn's 2013 analysis. Our subgroup analyses suggest that this variability may be due to different study moderators. Firstly, RC seems to benefit best from MT when the latter tackled both on lexical and grammatical morphology rather than just lexical morphology. This makes sense from a reading comprehension perspective, as both syntactic information (impacted by inflectional morphology) and semantic information (impacted by derivational and compositional morphology) concur to the elaboration of a mental representation (Perfetti, 2007). In order to benefit to RC development, MT should therefore emphasize both lexical and grammatical aspects of morphology. Secondly, MT appears to yield better effects on RC when MA exercises refer to complex words in the context of a sentence or a text rather than isolated. RC instruction is indeed more effective when words are presented among their natural, written context, allowing for word meaning derivation (Fukking & de Glopper, 1998), though in the case of MT within this meta-analysis, this effect was driven by only one study by Vaknin-Nusbaum and Raveh (2019).

5.2.5.4 Effects of MT on spelling

Opaque languages often become more consistent when switching from graphemic to morphemic units, especially in the spelling direction (Bowers & Bowers, 2017; Peereman et al., 2013): for instance, both “ee” and “ea” may transcribe the /i:/ sound in English, but it will always be transcribed “ea” in the morpheme “sea”, such as in seaweed, seagull, etc. Therefore, the practical relevance of teaching students to master this shift of perspective is an especially important matter. In the present meta-analysis, we found a significant impact of MT on

WS, as in previous works on the topic (Galuschka et al., 2020; Goodwin & Ahn, 2010, 2013; Rueda-Sanchez & Lopez-Bastida, 2016). Our ES could, however, be qualified as small. Although this is consistent with the findings obtained by Goodwin and Ahn in their 2010 analysis, the range of effects was generally higher in most other meta-analyses (Bowers et al., 2010; Goodwin & Ahn, 2013; Rueda-Sanchez & Lopez-Bastida, 2016). Several explanations may account for this different pattern of result, but they could not be further investigated through moderator analysis, as measures of heterogeneity were not significant for this outcome variable. One possible explanation could lie in the languages involved in the analysis. For instance, unlike Goodwin and Ahn's (2013) work, our own analysis as well as their 2010 publication included intermediate languages. This difference in ES might therefore reflect the fact that the impact of MT on spelling is stronger (hence, better captured) when focusing on opaque languages such as English (as we found for MA). Moreover, much stronger effects on spelling have been reported in meta-analyses whose authors purposefully aimed to review the efficiency of spelling interventions, i.e., interventions which were designed around spelling improvement (Galuschka et al., 2020). This differs from other analyses (including our own) that included studies which did not necessarily aim to teach students connections between MA and spelling, even if the impact of MT on spelling was subsequently evaluated.

Importantly, our meta-analysis also included additional specific measures of spelling, allowing us to better understand underlying mechanisms of a MT benefit. On the one hand, we report a lack of MT effect on PR, as in Goodwin & Ahn's (2010) analysis, suggesting that MT is not efficient in improving pure phoneme-to-grapheme conversion (PGC) skills, specifically required in simple nonword spelling. On the other hand, we report a significant, moderate effect of MT on MS (e.g., spelling a suffix such as "-ful"), which according to our subgroup analyses, was modulated by the type of control training. We found that MT was better at improving MS than orthographic training and regular treatment, but not better than read-

ing-based training or vocabulary training. A possible interpretation could be that those two types of control training also focus on morphemes, or at the very least word meaning. This could imply that when aiming to improve MS, meaning is as important – if not more important – as form.

Taken together, these results suggest that MT is primarily effective in improving orthographic representations of individual morphemes, rather than more general aspects of spelling, such as PGC or orthographic skills. However, it should be reminded that – as in all other outcome variables – we preferentially included untrained words, either through standardized or experimental measures. Thus, even though the impact of MT on WS was small, it also indicates that generalization to the spelling of unknown words is possible.

5.2.5.5 Effects of MT on vocabulary

Our analysis found no significant effect of MT on vocabulary, a result that differs markedly from Rueda-Sanchez and Lopez-Batisda (2016) as well as Goodwin and Ahn (2010; 2013) who did report a significant benefit. However, our subgroup analyses suggest that while no significant ES was found for students in preschool to grade 1, or in grade 2 to grade 5, MT may be efficient to improve vocabulary in students who are at least in grade 6. Although this observation may be limited in its generalizability, as we could only include two studies (Crosson et al., 2021; Harris et al., 2011) that investigated the impact of MT on vocabulary in grade 6 to grade 9 students, this hypothesis is worth exploring. Research efforts in the field seem to focus on a specific grade level (Davidson & O'Connor, 2019; Goodwin, 2016; Kieffer & Lesaux, 2012; Lesaux et al., 2010). The finding that MT may be most effective in improving vocabulary for children in grades higher than grade 6 in comparison to younger children is therefore unprecedented. This advantage would imply that middle schoolers are better disposed at using connections between words in a semantic purpose than elementary students. There is evidence that the proportion of complex words increases with

grade (Nagy & Anderson, 1984) and that middle schoolers benefit from MT directed at understanding the meaning of those words (Lesaux et al., 2010). However, to further support the hypothesis that middle schoolers and potentially high schoolers are more likely to develop vocabulary through MT than younger children, a comparison between training gains in both populations is needed.

A possible explanation for the discrepancy regarding vocabulary between our results and those of previous meta-analysis lies in fact that we preferred to include untrained words for our outcome measure. The positive impact of MT on vocabulary found by Goodwin and Ahn (2013) has been questioned in a reanalysis article by McQuillan (2020). The latter recalculated an ES for vocabulary on studies included by Goodwin and Ahn by considering the impact of training on *transfer* words, that is, using untrained words only. In doing so, McQuillan found a negligible, non-significant mean ES of MT on vocabulary, in line with our result. He concluded that while MT may partly help students recall word meaning of *trained* material, it does not appear to be appropriate as a generalization strategy to infer the meaning of unknown words. He therefore advised practitioners to use other effective treatments instead of MT, such as “free voluntary reading” (Krashen, 2004). Note that, surprisingly, our subgroup analyses did not reveal that any training program, including reading-based treatments, was more effective than MT in improving vocabulary, suggesting the need for further research to uncover appropriate intervention programs for vocabulary development.

5.2.5.6 Effects of MT on phonological awareness

A fairly surprising result in our analysis was the negative impact of MT on the PA outcome. PA was alternatively the most (Goodwin & Ahn, 2010) and the second most (Goodwin & Ahn, 2013) positively impacted outcome in Goodwin and Ahn’s successive analyses. This finding suggests that MT, by training MA in children, also trained processes that are common to PA: most likely, structured analysis of oral language. Our results do not

corroborate this hypothesis since MT appeared to be generally less efficient than control training in improving PA. We explain this major difference by considering the nature of the control groups. Indeed, out of our 15 included studies that evaluated the impact of MT on PA, 11 of them (73%) featured a control group that consisted in a phonological and/or phonics training. For these studies, PA was not necessarily expected by the authors to improve in the experimental group, but more likely in the control group. Although the moderating effect of control training was not significant, it is rather reasonable to suggest that this adverse effect of MT on PA is related to the content of the control training, since higher benefits in PA should be observed after specific training in phonology rather than morphology. In Goodwin and Ahn (2010; 2013), it appears that only two studies featured a phonological and/or phonics control group (Berninger et al., 2003; Filippini, 2007), which may explain why MT was more successful in improving PA in their analyses than it was in ours. Concomitantly, our computation of an averaged ES for PA included a study conducted in Norwegian, which is a transparent language (Lyster, 2002). This study was given the most weight in our analysis for PA, both in the fixed-effects and in the random-effects model. Goodwin and Ahn's reviews only focused on opaque and – to a lesser degree – on intermediate languages, supporting the idea that the origin of this difference may also be nested there. This hypothesis is comforted by our subgroup analyses, which revealed a significant, negative impact of MT for transparent languages, albeit being carried by the one study by Lyster (2002). Lastly, it should be noted that this adverse ES for PA was found in the fixed-effect model, but not in the random-effect one, which accounts for inter-study variability. An alternative explanation would therefore be that this adverse mean ES is simply a statistical artefact, resulting from the weight given to the study by Lyster (2002). Taken together, we do not believe that our results challenge the hypothesis that MT could support the development of PA. Rather, it simply sug-

gests that MT is likely not better at improving PA than a PA or phonics training, especially in transparent languages.

5.2.5.7 Summary of the impact of MT on written language subskills

All in all, our results suggest that MT is most effective for improving morpheme processing: the largest ES was obtained for MA; MR was the only word identification task to yield a significant ES; and MS obtained a larger ES than WS. Morphemes thus appear to be the unit to benefit the most from MT, which stands to reason, considering the nature of the instruction that was provided. Moreover, tasks requiring either whole-word processing such as reading fluency and accuracy or phonological processing (PA, phonological decoding and recoding) did not appear to benefit from MT. Subgroup analyses revealed that vocabulary and RC could result in significant ES, but only under particular conditions (respectively, for students in grades 6 to 9, and when MT presented morphemes in context and covered both lexical and grammatical morphology). Unlike previous meta-analyses and systematic reviews, our work therefore suggests that MT yields rather “morpheme-specific” training effects.

5.2.5.8 Absence of duration, instructor and type of student effects

Several moderating effects related to intervention design and population variables were not observed in the present analysis. Firstly, we found no better results of MT intervention on any written language skills by studies with longer duration times, against all expectations. As a matter of fact, a lack of training duration effect was also reported by Rueda-Sanchez and Lopez-Bastida (2016) and by Goodwin and Ahn in their 2013 analysis. However, in their 2010 one, Goodwin and Ahn did find that training effects were significantly larger for studies that administered at least 10 hours of training. This difference appears to relate to individual studies published after 2010 with small training durations obtaining large training effects. For instance, in their study 1, Devonshire and Fluck (2010) delivered only 30 minutes of morphological training to students from a large panel of different grades (kindergarten to

grade 5) and obtained a significant and moderate ES on word spelling. This would suggest that for MT, it is not so much the amount of time devoted to training that matters as much as other parameters such as the content of training itself.

The type of instructor seemed to have no moderating effect on the impact of MT within our meta-analysis. This is rather in line with the findings of Goodwin and Ahn (2013), who similarly showed no significant advantage of intervention being delivered by researchers over non-researchers or vice-versa. This consensual finding, also obtained by Rueda-Sanchez and Lopez-Bastida (2016), is encouraging for MT, as it suggests high research-to-practice transfer. Finally, the type of students also did not predict training effects, suggesting that all students, no matter their specificities, potentially benefit from MT. This was also the conclusion of Rueda-Sanchez and Lopez-Bastida (2016) and of Goodwin and Ahn (2010). Note that in the analysis by Bowers et al. (2010), “less able students” seemed to generally benefit more from MT on all measures. However, no significance values were calculated, thus limiting interpretation.

5.2.5.9 Limitations and perspectives

Likewise, a number of limitations within our study should be kept in mind. Firstly, from both a theoretical and methodological perspective, we opted for a research approach that mostly considers morphemes or words as the sole outcome measures. While these units of measurement may be considered proxies of written language, they are necessarily limited in their ability to reflect the actual proficiency of students in their own language. Secondly, the use of subgroup analyses presents a few caveats. The purpose of this type of analysis is usually to determine if modalities of the established categorical variables can moderate training effects. However, due to their post-hoc, i.e., non-experimental nature, it is impossible to keep all other categorical variables constant while trying to isolate one. Thereby, a seemingly significant predictor may not entirely or solely reflect the concept behind its attributed label.

Additionally, subgroups often lack representativity as only a few studies are typically featured under each predictor's modalities, meaning that an apparently significant modality may further reflect the case of a few, isolated studies rather than offer generalization perspectives. The conclusions of our subgroup analyses should therefore be considered cautiously in light of those two major limitations. The exclusive use of immediate post-tests evaluations in determining training gains also limits our analysis. A previous meta-analysis has calculated both immediate and delayed training effects of phonics training on reading and spelling outcomes (Ehri et al., 2001); however, meta-analyses synthesizing the results of training studies usually only include measures of performance right after the end of the training period. This is due to the fact that follow-up evaluations are not systematically conducted in individual studies for practical reasons, and that the durations between the end of the training and the subsequent post-tests may vary drastically from one study to another, ranging from one month (e.g., Bar-Kochva et al., 2020) to several years (e.g., Lyster et al., 2016) after the end of the training period. Whether intervention benefits may be maintained overtime is nonetheless of critical importance, and an increasing number of training studies now strive to include follow-up measures to their design. Some significant differences between experimental and control groups may even only be observed on those occasions (Torres et al., 2021). Therefore, future syntheses on the effect of MT on written language-related skills should try and consider those additional post-test measures.

Our work also points to additional future directions for research on MT. Although we computed average ES for eleven major outcomes in relation to written language, and conducted subgroup analyses on five of them, a few questions are still left unanswered, due to gaps in the literature. Firstly, we were not able to compute an average ES for the writing outcome, as we could only identify the one study by Allen and Lembke (2020). Functional written expression skills, however, could be considered as the ultimate goal targeted by spelling

interventions, in the same way that reading comprehension represents the underlying finality of word identification instruction programs. Following this logic, the evaluation of writing skills should be more frequently integrated to MT designs. Secondly, among all 40 included studies, none explored MT in students from grade 10 to 12. As we previously argued, middle and high school students may benefit particularly from MT due to the higher proportions of complex words in their curriculum (Nagy & Anderson, 1984); however, to date, research efforts on older participants appear to have focused almost exclusively on early middle schoolers (e.g., Bar-Kochva & Hasselhorn, 2017; Lesaux et al., 2010; Taha & Saiegh-Haddad, 2016) or young adults (e.g., Amirjalili et al., 2018). Considering the encouraging results obtained on both of these populations, it would be reasonable to believe that morphological instruction programs purposefully designed for high schoolers could result in similarly positive gains. Thirdly, although inflectional morphology was widely represented in our included studies, it was always in association with either derivational or compositional morphology (i.e., lexical morphology), except for the one study by Colchester (2010). Inflectional and derivational morphologies seem to play different roles in learning to read (Casalis & Louis-Alexandre, 2000) and have different development trajectories (Selby, 1972). However, while exclusive lexical morphology training studies were numerous in our study sample, the same could not be said about inflectional morphology. Exploring the respective impacts of both types of morphology training would be extremely informative for designing future MT protocols in function of student needs. Lastly, only three included studies focused on non-alphabetic languages (Chow et al., 2008; Wang & McBride, 2016; Wu et al., 2009), despite a growing interest for MA development in these languages in relation to written language learning (Han et al., 2022). Our conclusions regarding these languages in particular were therefore limited in interpretation. We hope that future research will seek to fill these gaps in the near future.

5.2.5.10 Conclusion

When focusing specifically on morphological awareness, MT yields significant and positive results for morphological awareness, morphological reading, word spelling and morpheme spelling. Namely, the effect of morphological training on reading and spelling appears greater for complex words. Further research is still needed in order to better understand the impact of MT for non-alphabetic languages, for students in grade 10 or higher and of inflectional morphology training in general.

5.2.6 References

*Studies included in the meta-analysis are signalled by an asterisk.

- *Abbott, S. P., & Berninger, V. W. (1999). It's never too late to remediate : Teaching word recognition to students with reading disabilities in grades 4–7. *Annals of Dyslexia*, 49(1), 221-250. <https://doi.org/10.1007/s11881-999-0025-x>
- *Allen, A. A., & Lembke, E. S. (2020). The Effect of a Morphological Awareness Intervention on Early Writing Outcomes. *Learning Disability Quarterly*.
<https://doi.org/10.1177/0731948720912414>
- Amirjalili, F., Jabbari, A. A., & Rezai, M. J. (2018). The effect of explicit instruction on derivational morphological awareness amongst Iranian EFL learners. *Linguistic Research*, 35, 47-82. <https://doi.org/10.17250/KHISLI.35..201809.003>
- Apel, K., Brimo, D., Diehm, E., & Apel, L. (2013). Morphological awareness intervention with kindergartners and first- and second-grade students from low socioeconomic status homes : A feasibility study. *Language, Speech, and Hearing Services in Schools*, 44(2), 161-173. [https://doi.org/10.1044/0161-1461\(2012/12-0042\)](https://doi.org/10.1044/0161-1461(2012/12-0042))
- *Apel, K., & Diehm, E. (2014). Morphological Awareness Intervention With Kindergarteners and First and Second Grade Students From Low SES Homes : A Small Efficacy Study. *Journal of Learning Disabilities*, 47(1), 65-75.
<https://doi.org/10.1177/0022219413509964>
- Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R : A practical tutorial. *Evidence Based Mental Health*, 22(4), 153-160.
<https://doi.org/10.1136/ebmental-2019-300117>
- *Bar-Kochva, I., & Hasselhorn, M. (2017). The Training of Morphological Decomposition in Word Processing and Its Effects on Literacy Skills. *Frontiers in Psychology*, 8, 1583.
<https://doi.org/10.3389/fpsyg.2017.01583>

- *Bar-Kochva, I., Korinth, S. P., & Hasselhorn, M. (2020). Effects of a morpheme-based training procedure on the literacy skills of readers with a reading disability. *Applied Psycholinguistics*, *41*(5), 1061-1082. <https://doi.org/10.1017/S0142716420000120>
- Berninger, V. W., Nagy, W. E., Carlisle, J., Thomson, J., Hoffer, D., & Abbott, S. (2003). Effective treatment for children with dyslexia in grades 4-6 : Behavioral and brain evidence. In *Preventing and remediating reading difficulties : Bringing science to scale* (B. R. Foorman, p. 381-417). Baltimore: York Press.
- *Berninger, V. W., Winn, W. D., Stock, P., Abbott, R. D., Eschen, K., Lin, S.-J. (Cindy), Garcia, N., Anderson-Youngstrom, M., Murphy, H., Lovitt, D., Trivedi, P., Jones, J., Amtmann, D., & Nagy, W. (2008). Tier 3 specialized writing instruction for students with dyslexia. *Reading and Writing*, *21*(1), 95-129. <https://doi.org/10.1007/s11145-007-9066-x>
- Bow, C. P., Blamey, P. J., Paatsch, L. E., & Sarant, J. Z. (2004). The effects of phonological and morphological training on speech perception scores and grammatical judgments in deaf and hard-of-hearing children. *Journal of Deaf Studies and Deaf Education*, *9*(3), 305-314. <https://doi.org/10.1093/deafed/enh032>
- *Bowers. (2006). *Adding transparency to morphologically opaque words through instruction* [Master's dissertation in Education]. Queen's University.
- Bowers, J. S., & Bowers, P. N. (2017). Beyond Phonics : The Case for Teaching Children the Logic of the English Spelling System. *Educational Psychologist*, *52*(2), 124-141. <https://doi.org/10.1080/00461520.2017.1288571>
- Bowers, Kirby, J. R., & Deacon, S. H. (2010). The Effects of Morphological Instruction on Literacy Skills : A Systematic Review of the Literature. *Review of Educational Research*, *80*(2), 144-179. <https://doi.org/10.3102/0034654309359353>

- *Brimo, D. (2016). Evaluating the Effectiveness of a Morphological Awareness Intervention : A Pilot Study. *Communication Disorders Quarterly*, 38(1), 35-45.
<https://doi.org/10.1177/1525740115604592>
- Carlisle, J. F., McBride-Chang, C., Nagy, W., & Nunes, T. (2010). Effects of Instruction in Morphological Awareness on Literacy Achievement : An Integrative Review. *Reading Research Quarterly*, 45(4), 464-487. <https://doi.org/10.1598/RRQ.45.4.5>
- *Casalis, S., & Colé, P. (2009). On the relationship between morphological and phonological awareness : Effects of training in kindergarten and in first-grade reading. *First Language*, 29(1), 113-142. <https://doi.org/10.1177/0142723708097484>
- Casalis, S., & Louis-Alexandre, M.-F. (2000). Morphological analysis, phonological analysis and learning to read French : A longitudinal study. *Reading and Writing*, 12(3), 303-335. <https://doi.org/10.1023/A:1008177205648>
- *Casalis, S., Pacton, S., Lefevre, F., & Fayol, M. (2017). Morphological training in spelling : Immediate and long-term effects of an interventional study in French third graders. *Learning and Instruction*, 53, 89-98.
<https://doi.org/10.1016/j.learninstruc.2017.07.009>
- *Chow, B. W.-Y., McBride-Chang, C., Cheung, H., & Chow, C. S.-L. (2008). Dialogic reading and morphology training in Chinese children : Effects on language and literacy. *Developmental Psychology*, 44(1), 233-244. <https://doi.org/10.1037/0012-1649.44.1.233>
- *Colchester, E. J. R. (2010). *Can we improve children's spelling ability by teaching morphemes through text reading ? An intervention study exploring the relationship between morphological awareness and literacy* [Doctoral dissertation in Professional Educational, Child and Adolescent Psychology]. Institute of Education, University of London.

- *Colenbrander, D., Parsons, L., Bowers, J. S., & Davis, C. J. (2021). Assessing the Effectiveness of Structured Word Inquiry for Students in Grades 3 and 5 With Reading and Spelling Difficulties : A Randomized Controlled Trial. *Reading Research Quarterly*, rrq.399. <https://doi.org/10.1002/rrq.399>
- *Crosson, A. C., McKeown, M. G., Lei, P., Zhao, H., Li, X., Patrick, K., Brown, K., & Shen, Y. (2021). Morphological analysis skill and academic vocabulary knowledge are malleable through intervention and may contribute to reading comprehension for multilingual adolescents. *Journal of Research in Reading*, 44(1), 154-174. <https://doi.org/10.1111/1467-9817.12323>
- Crosson, A. C., McKeown, M. G., Moore, D. W., & Ye, F. (2019). Extending the bounds of morphology instruction : Teaching Latin roots facilitates academic word learning for English Learner adolescents. *Reading and Writing*, 32(3), 689-727. <https://doi.org/10.1007/s11145-018-9885-y>
- *Dallashah-Khatib, R., Ibrahim, R., & Karni, A. (2014). Longitudinal Data on the Relations of Morphological and Phonological Training to Reading Acquisition in First Grade : The Case of Arabic Language. *Psychology*, 05(08), 918-940. <https://doi.org/10.4236/psych.2014.58103>
- Davidson, S. J., & O'Connor, R. E. (2019). An intervention using morphology to derive word meanings for English language learners. *Journal of Applied Behavior Analysis*, 52(2), 394-407. <https://doi.org/10.1002/jaba.539>
- Deacon, S. H., Benere, J., & Pasquarella, A. (2013). Reciprocal relationship : Children's morphological awareness and their reading accuracy across grades 2 to 3. *Developmental Psychology*, 49(6), 1113-1126. <https://doi.org/10.1037/a0029474>

- Deacon, S. H., & Kirby, J. R. (2004). Morphological awareness : Just “more phonological”? The roles of morphological and phonological awareness in reading development. *Applied Psycholinguistics*, 25(2), 223-238. <https://doi.org/10.1017/S0142716404001110>
- Deacon, S. H., Tong, X., & Francis, K. (2017). The relationship of morphological analysis and morphological decoding to reading comprehension : Morphological Analysis and Morphological Decoding. *Journal of Research in Reading*, 40(1), 1-16. <https://doi.org/10.1111/1467-9817.12056>
- Desrochers, A., Manolitsis, G., Gaudreau, P., & Georgiou, G. (2018). Early contribution of morphological awareness to literacy skills across languages varying in orthographic consistency. *Reading and Writing*, 31(8), 1695-1719. <https://doi.org/10.1007/s11145-017-9772-y>
- *Devonshire, V., & Fluck, M. (2010). Spelling development : Fine-tuning strategy-use and capitalising on the connections between words. *Learning and Instruction*, 20(5), 361-371. <https://doi.org/10.1016/j.learninstruc.2009.02.025>
- Diamanti, V., Mouzaki, A., Ralli, A., Antoniou, F., Papaioannou, S., & Protopapas, A. (2017). Preschool Phonological and Morphological Awareness As Longitudinal Predictors of Early Reading and Spelling Development in Greek. *Frontiers in Psychology*, 8, 2039. <https://doi.org/10.3389/fpsyg.2017.02039>
- Egger, M., Smith, G. D., & Phillips, A. N. (1997). Meta-analysis : Principles and procedures. *BMJ (Clinical Research Ed.)*, 315(7121), 1533-1537. <https://doi.org/10.1136/bmj.315.7121.1533>
- Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic Awareness Instruction Helps Children Learn to Read : Evidence From the National Reading Panel’s Meta-Analysis. *Reading Research Quarterly*, 36(3), 250-287. <https://doi.org/10.1598/RRQ.36.3.2>

- *Elbro, C., & Arnbak, E. (1996). The role of morpheme recognition and morphological awareness in dyslexia. *Annals of Dyslexia*, 46(1), 209-240.
<https://doi.org/10.1007/BF02648177>
- *Fejzo, A., Saidane, R., Whissell-Turner, K., & Chapleau, N. (2019, juillet 19). *Effects of morphological awareness training on the development of morphological strategy use in polymorphemic word spelling among 4th grade French-speaking students* [Poster]. 26th annual meeting of Society of Scientific Studies of Reading (SSSR), Toronto, Canada.
- Filippini, A. L. (2007). *Effects of a vocabulary-added instructional intervention for at-risk English learners : Is efficient reading instruction more effective?* [Unpublished doctoral dissertation]. University of California, Santa Barbara.
- *Frank, A. R. (2008). *The effect of instruction in orthographic conventions and morphological features on the reading fluency and comprehension skills of high-school freshmen* [Doctoral dissertation in Education, University of San Francisco].
<https://repository.usfca.edu/diss/263>
- Fukkink, R. G., & de Glopper, K. (1998). Effects of Instruction in Deriving Word Meaning from Context : A Meta-Analysis. *Review of Educational Research*, 68(4), 450-469.
<https://doi.org/10.2307/1170735>
- Galuschka, K., Goergen, R., Kalmar, J., Haberstroh, S., Schmalz, X., & Schulte-Koerne, G. (2020). Effectiveness of spelling interventions for learners with dyslexia : A meta-analysis and systematic review. *Educational Psychologist*, 55(1), 1-20.
<https://doi.org/10.1080/00461520.2019.1659794>
- Galuschka, K., & Schulte-Körne, G. (2016). The Diagnosis and Treatment of Reading and/or Spelling Disorders in Children and Adolescents. *Deutsches Arzteblatt International*, 113(16), 279-286. <https://doi.org/10.3238/arztebl.2016.0279>

- Galuschka, K., Von Wirth, E., Krick, K., & Schulte-Körne, G. (2014). Effectiveness of Treatment Approaches for Children and Adolescents with Reading Disabilities : A Meta-Analysis of Randomized Controlled Trials. *PloS one*, *9*, e89900. <https://doi.org/10.1371/journal.pone.0089900>
- Gebauer, D., Fink, A., Kargl, R., Reishofer, G., Koschutnig, K., Purgstaller, C., Fazekas, F., & Enzinger, C. (2012). Differences in brain function and changes with intervention in children with poor spelling and reading abilities. *PloS One*, *7*(5), e38201. <https://doi.org/10.1371/journal.pone.0038201>
- *Gellert, A. S., Arnbak, E., Wischmann, S., & Elbro, C. (2021). Morphological Intervention for Students With Limited Vocabulary Knowledge : Short- and Long-Term Transfer Effects. *Reading Research Quarterly*, *56*(3), 583-601. <https://doi.org/10.1002/rrq.325>
- *Georgiou, G. K., Martinez, D., Vieira, A. P. A., & Guo, K. (2021). Is orthographic knowledge a strength or a weakness in individuals with dyslexia? Evidence from a meta-analysis. *Annals of Dyslexia*, *71*(1), 5-27. <https://doi.org/10.1007/s11881-021-00220-6>
- Good, J. E., Lance, D. M., & Rainey, J. (2015). The Effects of Morphological Awareness Training on Reading, Spelling, and Vocabulary Skills. *Communication Disorders Quarterly*, *36*(3), 142-151. <https://doi.org/10.1177/1525740114548917>
- Goodwin, A. P. (2016). Effectiveness of word solving : Integrating morphological problem-solving within comprehension instruction for middle school students. *Reading and Writing*, *29*(1), 91-116. <https://doi.org/10.1007/s11145-015-9581-0>
- Goodwin, & Ahn. (2010). A meta-analysis of morphological interventions : Effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia*, *60*(2), 183-208. <https://doi.org/10.1007/s11881-010-0041-x>

- Goodwin, & Ahn. (2013). A Meta-Analysis of Morphological Interventions in English : Effects on Literacy Outcomes for School-Age Children. *Scientific Studies of Reading*, 17(4), 257-285. <https://doi.org/10.1080/10888438.2012.689791>
- Haase, A., & Steinbrink, C. (2022). Associations between morphological awareness and literacy skills in German primary school children : The roles of grade level, phonological processing and vocabulary. *Reading and Writing*. <https://doi.org/10.1007/s11145-021-10247-1>
- Han, B., Koh, P. W., Zhang, S., Joshi, R. M., & Li, H. (2022). The relative contributions of facets of morphological awareness to vocabulary development in Chinese : A longitudinal study in grades one to three. *Contemporary Educational Psychology*, 69, 102063. <https://doi.org/10.1016/j.cedpsych.2022.102063>
- *Harris, M. L., Schumaker, J. B., & Deshler, D. D. (2011). The Effects of Strategic Morphological Analysis Instruction on the Vocabulary Performance of Secondary Students with and without Disabilities. *Learning Disability Quarterly*, 34(1), 17-33. <https://doi.org/10.1177/073194871103400102>
- Henry, M. K. (1988). Beyond phonics : Integrated decoding and spelling instruction based on word origin and structure. *Annals of Dyslexia*, 38(1), 258-275. <https://doi.org/10.1007/BF02648260>
- Høien, T., Lundberg, I., Stanovich, K. E., & Bjaalid, I. K. (1995). Components of phonological awareness. *Reading and writing*, 7(2), 171-188.
- Hurry, J., Nunes, T., Bryant, P., Pretzlik, U., Parker, M., Curno, T., & Midgley, L. (2005). Transforming research on morphology into teacher practice. *Research Papers in Education*, 20, 187-206. <https://doi.org/10.1080/02671520500078291>
- *Jöbstl, V., Kargl, R., Prattes, A. E., Beyersmann, E., & Landerl, K. (2021). Effects of a Morpheme-based Spelling Intervention Challenging Previous Results. *International*

Electronic Journal of Elementary Education, 13(5), 651-671.

<https://doi.org/10.26822/iejee.2021.219>

Ke, S. (Echo), & Zhang, D. (2021). Morphological instruction and reading development in young L2 readers : A scoping review of causal relationships. *Studies in Second Language Learning and Teaching*, 11(3), 331-350.

<https://doi.org/10.14746/sslts.2021.11.3.2>

Kieffer, M. J., & Lesaux, N. K. (2012). Effects of Academic Language Instruction on Relational and Syntactic Aspects of Morphological Awareness for Sixth Graders from Linguistically Diverse Backgrounds. *The Elementary School Journal*, 112(3),

519-545. JSTOR. <https://doi.org/10.1086/663299>

Kirby, Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012).

Children's morphological awareness and reading ability. *Reading and Writing*, 25(2),

389-410. <https://doi.org/10.1007/s11145-010-9276-5>

Kirk, C., & Gillon, G. T. (2009). Integrated morphological awareness intervention as a tool for improving literacy. *Language, Speech, and Hearing Services in Schools*, 40(3),

341-351. [https://doi.org/10.1044/0161-1461\(2008/08-0009\)](https://doi.org/10.1044/0161-1461(2008/08-0009))

*Kolne, K. (2019). *Disambiguating the benefits of morphology and vocabulary instruction for improving spelling among English-speaking children* [Doctoral dissertation in Communication Sciences and Disorders]. School of Communication Sciences and Disorders, McGill University.

Krashen, S. (2004). *The Power of Reading : Insights from the Research* (2d edition). Heinemann.

Kruk, R. S., & Bergman, K. (2013). The reciprocal relations between morphological processes and reading. *Journal of Experimental Child Psychology*, 114(1), 10-34.

<https://doi.org/10.1016/j.jecp.2012.09.014>

- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science : A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*.
<https://doi.org/10.3389/fpsyg.2013.00863>
- Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P. H. T., Lohvansuu, K., O'Donovan, M., Williams, J., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., Tóth, D., Honbolygó, F., Csépe, V., Bogliotti, C., Iannuzzi, S., Chaix, Y., Démonet, J.-F., ... Schulte-Körne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity : Cross-linguistic predictors of dyslexia. *Journal of Child Psychology and Psychiatry, 54*(6), 686-694.
<https://doi.org/10.1111/jcpp.12029>
- Lee, J. won, Wolters, A., & Grace Kim, Y.-S. (2022). The Relations of Morphological Awareness with Language and Literacy Skills Vary Depending on Orthographic Depth and Nature of Morphological Awareness. *Review of Educational Research, 00346543221123816*. <https://doi.org/10.3102/00346543221123816>
- Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The Effectiveness and Ease of Implementation of an Academic Vocabulary Intervention for Linguistically Diverse Students in Urban Middle Schools. *Reading Research Quarterly, 45*(2), 196-228. <https://doi.org/10.1598/RRQ.45.2.3>
- Levesque, K. C., Breadmore, H. L., & Deacon, S. H. (2021). How morphology impacts reading and spelling : Advancing the role of morphology in models of literacy development. *Journal of Research in Reading, 44*(1), 10-26. <https://doi.org/10.1111/1467-9817.12313>
- Levesque, K. C., Kieffer, M. J., & Deacon, S. H. (2019). Inferring Meaning From Meaningful Parts : The Contributions of Morphological Skills to the Development of Children's

- Reading Comprehension. *Reading Research Quarterly*, 54(1), 63-80.
<https://doi.org/10.1002/rrq.219>
- Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small Group and Individual Learning with Technology : A Meta-Analysis. *Review of Educational Research*, 71(3), 449-521.
<https://doi.org/10.3102/00346543071003449>
- *Lyster, S.-A. H. (2002). The effects of morphological versus phonological awareness training in kindergarten on reading development. *Reading and Writing*, 15(3), 261-294.
<https://doi.org/10.1023/A:1015272516220>
- Lyster, S.-A. H., Lervåg, A. O., & Hulme, C. (2016). Preschool morphological training produces long-term improvements in reading comprehension. *Reading and Writing*, 29, 1269-1288. <https://doi.org/10.1007/s11145-016-9636-x>
- Manolitsis, G., Grigorakis, I., & Georgiou, G. K. (2017). The longitudinal contribution of early morphological awareness skills to reading fluency and comprehension in Greek. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01793>
- Marulis, L. M., & Neuman, S. B. (2013). How Vocabulary Interventions Affect Young Children at Risk : A Meta-Analytic Review. *Journal of Research on Educational Effectiveness*, 6(3), 223-262. <https://doi.org/10.1080/19345747.2012.755591>
- McQuillan, J. (2020). *The Effects of Morphological Training on Vocabulary Knowledge : A Reanalysis of Goodwin & Ahn*. 19-38.
- Morris, S. B. (2008). Estimating Effect Sizes From Pretest-Posttest-Control Group Designs. *Organizational Research Methods*, 11(2), 364-386.
<https://doi.org/10.1177/1094428106291059>
- Nagy, & Anderson, R. C. (1984). How Many Words Are There in Printed School English? *Reading Research Quarterly*, 19(3), 304. <https://doi.org/10.2307/747823>

- Nunes, T., Bryant, P., & Olsson, J. (2003). Learning Morphological and Phonological Spelling Rules : An Intervention Study. *Scientific Studies of Reading - SCI STUD READ*, 7, 289-307. https://doi.org/10.1207/S1532799XSSR0703_6
- Pacton, S., & Deacon, S. H. (2008). The timing and mechanisms of children's use of morphological information in spelling : A review of evidence from English and French. *Cognitive Development*, 23(3), 339-359. <https://doi.org/10.1016/j.cogdev.2007.09.004>
- Peereman, R., Sprenger-Charolles, L., & Messaoud-Galusi, S. (2013). The contribution of morphology to the consistency of spelling-to-sound relations : A quantitative analysis based on French elementary school readers. *L'Année Psychologique*, Vol. 113(1), 3-33.
- Perfetti, C. A. (2007). Reading Ability : Lexical Quality to Comprehension. *Scientific Studies of Reading*, 11(4), 357-383.
- *Rassel, A. (2020). *Contribution de la conscience morphologique à l'apprentissage de la lecture : Études en milieu défavorisé* [Doctoral dissertation in Psychology, University of Lille 3]. <http://www.theses.fr/2020LIL3H023>
- *Ravid, D., & Geiger, V. (2009). Promoting morphological awareness in Hebrew-speaking grade-schoolers : An intervention study using linguistic humor. *First Language*, 29(1), 81-112. <https://doi.org/10.1177/0142723708097483>
- Reed, D. K. (2008). A Synthesis of Morphology Interventions and Effects on Reading Outcomes for Students in Grades K–12. *Learning Disabilities Research & Practice*, 23(1), 36-49. <https://doi.org/10.1111/j.1540-5826.2007.00261.x>
- Reis, A., Araújo, S., Morais, I. S., & Faísca, L. (2020). Reading and reading-related skills in adults with dyslexia from different orthographic systems : A review and meta-analysis. *Annals of Dyslexia*, 70(3), 339-368. <https://doi.org/10.1007/s11881-020-00205-x>

- Rey-Debove, J. (1984). Le domaine de la morphologie lexicale. In *Cahiers de lexicologie* (p. 3-19).
- Richards, T., Berninger, V., Nagy, W., Parsons, A., Field, K., & Richards, A. (2005). Brain activation during language task contrasts in children with and without dyslexia : Inferring mapping processes and assessing response to spelling instruction. *Educational and Child Psychology, 22*, 62-80.
- Ruan, Y., Georgiou, G. K., Song, S., Li, Y., & Shu, H. (2018). Does writing system influence the associations between phonological awareness, morphological awareness, and reading? A meta-analysis. *Journal of Educational Psychology, 110*(2), 180-202.
<https://doi.org/10.1037/edu0000216>
- Rueda-Sanchez, M. I., & Lopez-Bastida, P. (2016). Effects of morphological awareness training on reading, writing and comprehension : Meta-analysis. *Anales De Psicologia, 32*(1), 60-71. <https://doi.org/10.6018/analesps.32.1.196261>
- Schiff, R., & Saiegh-Haddad, E. (2018). Development and Relationships Between Phonological Awareness, Morphological Awareness and Word Reading in Spoken and Standard Arabic. *Frontiers in Psychology, 9*.
<https://www.frontiersin.org/article/10.3389/fpsyg.2018.00356>
- Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). *Meta-analysis with R*. Springer.
- Selby, S. (1972). The development of morphological rules in children. *British Journal of Educational Psychology, 42*(3), 293-299. <https://doi.org/10.1111/j.2044-8279.1972.tb00722.x>
- Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology (London, England: 1953), 94*(Pt 2), 143-174. <https://doi.org/10.1348/000712603321661859>

- Sparks, E., & Deacon, H. (2015). Morphological awareness and vocabulary acquisition : A longitudinal examination of their relationship in English-speaking children. *Applied Psycholinguistics*, *36*, 299-321. <https://doi.org/10.1017/S0142716413000246>
- Taha, H. (2016). Deep and shallow in Arabic orthography : New evidence from reading performance of elementary school native Arab readers. *Writing Systems Research*, *8*. <https://doi.org/10.1080/17586801.2015.1114910>
- *Taha, H., & Saiegh-Haddad, E. (2016). The Role of Phonological versus Morphological Skills in the Development of Arabic Spelling : An Intervention Study. *Journal of Psycholinguistic Research*, *45*(3), 507-535. <https://doi.org/10.1007/s10936-015-9362-6>
- Torgerson, C., Brooks, G., Gascoine, L., & Higgins, S. (2019). Phonics : Reading policy and the evidence of effectiveness from a systematic ‘tertiary’ review. *Research Papers in Education*, *34*(2), 208-238. <https://doi.org/10.1080/02671522.2017.1420816>
- *Torkildsen, J. von K., Bratlie, S. S., Kristensen, J. K., Gustafsson, J.-E., Lyster, S.-A. H., Snow, C., Hulme, C., Mononen, R.-M., Næss, K.-A. B., López-Pedersen, A., Wie, O. B., & Hagtvet, B. (2021). App-based morphological training produces lasting effects on word knowledge in primary school children : A randomized controlled trial. *Journal of Educational Psychology*. <https://doi.org/10.1037/edu0000688>
- Torres, A. R., Mota, N. B., Adamy, N., Naschold, A., Lima, T. Z., Copelli, M., Weissheimer, J., Pegado, F., & Ribeiro, S. (2021). Selective Inhibition of Mirror Invariance for Letters Consolidated by Sleep Doubles Reading Fluency. *Current Biology*, *31*(4), 742-752.e8. <https://doi.org/10.1016/j.cub.2020.11.031>
- *Tsesmeli, S. N. (2017). Spelling and Meaning of Compounds in the Early School Years through Classroom Games : An Intervention Study. *Frontiers in Psychology*, *8*, 2071. <https://doi.org/10.3389/fpsyg.2017.02071>

- *Vaknin-Nusbaum, V., & Raveh, M. (2019). Cultivating morphological awareness improves reading skills in fifth-grade Hebrew readers. *Journal of Educational Research*, *112*(3), 357-366. <https://doi.org/10.1080/00220671.2018.1528541>
- Vaknin-Nusbaum, V., & Sarid, M. (2021). The role of morphological awareness in the development of reading comprehension in Hebrew-speaking second-graders. *Reading and Writing*, *34*(10), 2645-2671. <https://doi.org/10.1007/s11145-021-10162-5>
- *Wang, Y., & McBride, C. (2016). Beyond Copying : A Comparison of Multi-Component Interventions on Chinese Early Literacy Skills. *International Journal of Behavioral Development*, *41*(3), 380-389. <https://doi.org/10.1177/0165025416637212>
- Wei, T.-Q., Bi, H.-Y., Chen, B.-G., Liu, Y., Weng, X.-C., & Wydell, T. N. (2014). Developmental changes in the role of different metalinguistic awareness skills in Chinese reading acquisition from preschool to third grade. *PloS One*, *9*(5), e96240. <https://doi.org/10.1371/journal.pone.0096240>
- Weiser, B., & Mathes, P. (2011). Using Encoding Instruction to Improve the Reading and Spelling Performances of Elementary Students At Risk for Literacy Difficulties : A Best-Evidence Synthesis. *Review of Educational Research*, *81*(2), 170-200. <https://doi.org/10.3102/0034654310396719>
- *Weiss, S., Grabner, R. H., Kargl, R., Purgstaller, C., & Fink, A. (2010). Behavioral and neurophysiological effects of morphological awareness training on spelling and reading. *Reading and Writing*, *23*(6), 645-671. <https://doi.org/10.1007/s11145-009-9177-7>
- *Wolter, J. A., & Dilworth, V. (2014). The Effects of a Multilinguistic Morphological Awareness Approach for Improving Language and Literacy. *Journal of Learning Disabilities*, *47*(1), 76-85. <https://doi.org/10.1177/0022219413509972>
- *Wu, X., Anderson, R., li, W., Wu, X., Li, H., Zhang, J., Zheng, Q., Zhu, J., Shu, H., Jiang, W., Chen, X., Wang, Q., Yin, L., He, Y., Packard, J., & Gaffney, J. (2009). Morpho-

logical Awareness and Chinese Children's Literacy Development : An Intervention Study. *Scientific Studies of Reading*, 13, 26-52.

<https://doi.org/10.1080/10888430802631734>

*Zhang, D. (2016). Morphology in Malay–English biliteracy acquisition : An intervention study. *International Journal of Bilingual Education and Bilingualism*, 19(5), 546-562.

<https://doi.org/10.1080/13670050.2015.1026873>

*Zhang, D., & Li, L. (2016). Morphological Teaching and Singaporean Children's English Word Learning. In R. Elaine Silver & W. D. Bokhorst-Heng (Éds.), *Quadilingual Education in Singapore* (p. 199-218). Springer Singapore.

https://doi.org/10.1007/978-981-287-967-7_12

*Zoski, J. L., & Erickson, K. A. (2017). Multicomponent Linguistic Awareness Intervention for At-Risk Kindergarteners. *Communication Disorders Quarterly*, 38(3), 161-171.

<https://doi.org/10.1177/1525740116660817>

5.2.7 Tables

Table 1

Summary of Included Studies and their Associated Variables

Included studies	Outcome variables	Categorical variables
1. Abbott & Berninger, 1999	RA, RF, MR, PD, RC, WS, PA	ALT; Individual; RD; ≥ 15 h; At-risk; G4-7 (BR); O; Lex; MA; Not computerized
2. Allen & Lembke, 2020	WS, W	RT; Large group; RD; 10h-15h; At-risk; G2-3 (YR); O; Both; TT; Not computerized
3. Apel & Diehm, 2014	MA, RF, PD, RC, MS	RT; Small group; RD; 10h-15h; At-risk; a Kindergarten (BR), b. G1 (BR), c. G2 (YR); O; Both; TT; Not computerized
4. Bar-Kochva & Hasselhorn, 2017	RF, MR, RC, WS	ALT; Individual; RD; < 10 h; SLL; G5-6 (BG); ITM; Both; IL; Computerized
5. Bar-Kochva et al. 2020	RF, MR, RC, WS	ALT; Individual; RD; < 10 h; At-risk; G5-6 (BG); ITM; Both; IL; Computerized
6. Berninger et al., 2008 (study 1)	PD, PR	OT; Small group; RD; ≥ 15 h; LD; a. G4-6 (BG), b. G7-9 (CR); O; Lex; MAS; Not computerized
7. Bowers, 2006	MA, RA, WS	RT; Large group; RD; ≥ 15 h; TL; G4-5 (YR); O; Lex; TT; Not computerized
8. Brimo, 2016	MA, MS	RBT; Small group; RD; 10h-15h; LD; G3 (YR); O; Both; TT; Not computerized
9. Casalis & Colé, 2009 (study 1)	MA, V, PA	PT; Small group; RD; < 10 h; TL; K (BR); O; Both; MA; Not computerized
10. Casalis et al., 2017	MA, WS, MS, V	OT; Large group; NRD; < 10 h; TL; G3 (YR); O; Lex; TT; Not computerized
11. Chow et al., 2008	MA, RA, V	RBT; Individual; NRD; < 10 h; TL; K (BR); NA; Lex; MAC; Not computerized
12. Colchester, 2010	MS	OT; Large group; RD; < 10 h; TL; G1-2 (BG); O; Gram; MAS; Computerized
13. Colenbrander et al., 2021	RF, MR, PD, RC, WS, MS, PR, V	RBT; Small group; NRD; ≥ 15 h; At-risk; G3-5 (YR); O; Lex; TT; Not computerized
14. Crosson et al., 2021	MA, V	RT; Large group; NRD; ≥ 15 h; SLL; G7-8 (CR); O; Lex; TT; Not computerized
15. Dallahsheh-Khatib et al., 2014	MA, V, PA	PT; Small group; RD; ≥ 15 h; TL; K (BR); O; Both; MA; Not computerized
16. Devonshire & Fluck, 2010 (study 1)	WS	ALT; Small group; RD; < 10 h; TL; K to G5 (BG); O; Lex; MA; Not computerized
17. Devonshire & Fluck, 2010 (study 2)	WS, MS	OT; Large group; RD; < 10 h; TL; G2-3 (YR); O; Lex; TT; Not computerized
18. Elbro & Arnbak, 1996 (study 3)	MA, RA, PD, RC, WS, V, PA	RT; Small group; NRD; < 10 h; LD; G4-5 (YR); O; Both; MA; Not computerized
19. Fejzo et al., 2019	MA, WS, MS	RT; Large group; RD; ≥ 15 h; TL; G4 (YR); O; Lex; MAS; Not computerized
20. Frank, 2008	RF, RC, WS	RBT; Large group; NRD; 10h-15h; TL; G9 (CR); O; Both; TT; Not computerized
21. Gebauer et al., 2012	RC, WS	RT; Individual; RD; 10h-15h; TL; BG; ITM; Lex; TT; Not computerized
22. Gellert et al., 2021	MA, RC, V	VT; Small group; RD; 10h-15h; At-risk; G5 (YR); O; Both; TT; Not computerized
23. Georgiou et al., 2021	MA, RA, PD	PT; Individual; RD; 10h-15h; At-risk; G3 (YR); O; Lex; TT; Not computerized
24. Harris et al., 2011	MA, V	VT; Large group; RD; < 10 h; TL; G9 (CR); O; Lex; MAS; Not computerized

Table 1 (continued)

Included studies	Outcome variables	Categorical variables
25. Jöbstl et al., 2021	MA, RF, MR, PD, RC, WS, MS	RT; Large group; RD; ≥15h; TL; G2 (YR); O; Both; MAS; Computerized
26. Kolne, 2019	MA, WS, MS, V	VT; Large group; RD; 10h-15h; TL; a. G3 (YR), b. G5 (YR); O; Lex; MAS; Not computerized
27. Lyster, 2002	MA, V, PA	PT; Small group; NRD; 10h-15h; TL; Preschool (BR); TR; Both; MA; Not computerized
28. Rassel, 2020	MA, RA, RF, MR, PD, RC, V, PA	VT; Large group; NRD; 10h-15h; At-risk; a. G2 (YR), b. G3 (YR); O; Lex; TT; Not computerized
29. Ravid & Geiger, 2009	MA	ALT; Large group; RD; 10h-15h; TL; G4 (YR); O; Both; MA; Not computerized
30. Taha & Saiegh-Haddad, 2016	MA, WS, PA	PT; Small group; RD; ≥15h; a. At-risk, b. TL; a. G2 (YR); b. G4 (YR); c. G6 (CR); O; Both; MAS; Not computerized
31. Torkildsen et al., 2021	MA, MR, WS, V	ANT; Individual; NRD; < 10h; TL; G2 (YR); TR; Both; TT; Computerized
32. Tsesmeli, 2017	WS	RT; Large group; RD; 10h-15h; TL; a. G1 (BR), b. G2 (YR); TR; Lex; MAS; Not computerized
33. Vaknin-Nusbaum & Raveh, 2019	MA, RF, RC	RBT; Small group; NRD; ≥15h; At-risk; G5 (YR); O; Both; MAC; Not computerized
34. Wang & McBride, 2016	MA, MR, WS, PA	PT; Small group; RD; < 10h; TL; Kindergarten (BR); NA; Lex; TT; Not computerized
35. Weiss et al., 2010	RC, WS	RT; Individual; RD; ≥15h; TL; G3 to 9 (BG); ITM; Lex; TT; Computerized
36. Wolter & Dilworth, 2014	RA, PD, RC, WS	OT; Small group; RD; 10h-15h; At-risk; G2 (YR); O; Both; TT; Not computerized
37. Wu et al., 2009	MA	RT; Large group; NRD; duration unknown; TL; G2 (YR); NA; Lex; MA; Not computerized
38. Zhang, 2016	MA, MR	RT; Large group; NRD; < 10h; SLL; G4 (YR); O; Lex; MAS; Not computerized
39. Zhang & Li, 2016	MA, MS	RT; Large group; NRD; < 10h; SLL; G4 (YR); O; Lex; MAS; Not computerized
40. Zoski & Erickson, 2017	MA, RF, WS, PA	PT; Small group; RD; 10h-15h; At-risk; K (BR); O; Both; MAC; Not computerized

Note. ALT = Alternative linguistic treatment. ANT = Alternative non-linguistic treatment. At-risk = At-risk of learning disability. Both = Lexical and grammatical morphology. BG = Broad grade. BR = Beginning readers. G = Grade. Gram = Grammatical morphology. IL = Implicit learning. ITM = Intermediate language. Lex = Lexical morphology. LD = Learning disabled. MA = Morphological awareness. MAC = Morphological awareness with words presented in context. MAS = Morphological awareness with an emphasis on spelling. MR = Morphological reading. MS = Morpheme spelling. NA = Non-alphabetic language. NRD = Non-researcher delivered. O = Opaque language. OT = Orthographic training. PA = Phonological awareness. PD = Phonological decoding. PR = Phonological recoding. PT = Phonology/phonics training. RA = Reading accuracy. RBT = Reading-based treatment. RC = Reading comprehension. RD = Researcher delivered. RF = Reading fluency. RT = Regular treatment. SLL = Second language learner. TL = Typical learner. TR = Transparent language. TT = Thorough training. V = Vocabulary. VT = Vocabulary training. W = Writing. WS = Word spelling. YR = Young readers.

Table 2*Effects of Morphological Training on Written Language Outcomes*

Outcome	<i>k</i>	Fixed-effects model			Random-effects model			τ^2	Q_{total}
		Mean <i>SMD</i>	95% CI		Mean <i>SMD</i>	95% CI			
			LL	UL		LL	UL		
MA ^a	34	0.34***	0.27	0.41	0.48***	0.30	0.66	0.21**	173.07***
RA	13	0.11	-0.02	0.24	0.10	-0.06	0.26	0.03	17.54
RF	13	0.04	-0.09	0.17	0.04	-0.09	0.17	0.00	2.66
MR	9	0.15**	0.04	0.25	0.15**	0.04	0.25	0.00	0.91
PD	11	0.01	-0.13	0.14	0.01	-0.13	0.14	0.00	8.48
RC	16	0.10	-0.02	0.21	0.16	-0.03	0.35	0.08**	36.16**
WS	28	0.19***	0.11	0.28	0.22***	0.11	0.33	0.01	33.59
MS	14	0.38***	0.25	0.50	0.41***	0.24	0.59	0.04 [†]	20.81 [†]
PR	9	-0.03	-0.20	0.14	0.01	0.21	0.19	0.02	9.60
V	15	0.04	-0.04	0.13	0.03	-0.10	0.15	0.02 [†]	23.24 [†]
PA	15	-0.15*	-0.29	-0.01	-0.09	-0.37	0.18	0.19**	49.01***

Note. CI = Confidence interval. LL = Lower limit. MA = Morphological awareness. MR = Morphological reading. MS = Morpheme spelling. PA = Phonological awareness. PD = Phonological decoding. PR = Phonological recoding. RA = Reading accuracy. RC = Reading comprehension. RF = Reading fluency. *SMD* = Standardized Mean Difference. UL = Upper limit. V = Vocabulary. WS = Word spelling.

^a The presented result excludes the outlier study by Harris et al. (2011)

[†] $p < .09$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table S1

Included Studies

Study	Opacity	Comparison group	Intervention unit	Instructor	Duration	Type of student	Grade	Type of morphology	Outcome variables	Corresponding tasks in publication
1. <u>Abbott & Berninger (1999)</u>	Opaque (English)	Alternative linguistic training (study skills training)	Individual	Researcher	1h/week for 16 weeks =16h	At-risk of LD	4-7	L (derivation)	RA _{STD} ; RF _{STD} ; MR _E ; PD _{STD} ; RC _{STD} ; WS _{STD} ; PA _{STD}	WJ-R: word identification; TOWRE: sight word efficiency; Taught words; WJ-R: word attack; WJ-R: passage comprehension; WIAT: spelling; CTOPP
2. <u>Allen & Lembke (2020)</u>	Opaque (English)	Regular treatment	Small group	Researcher	112.5m/week for 5 weeks =9.5h	At-risk of LD	2-3	Both	WS _{STD} ; W _{STD}	WJ-3-ACH: spelling; WJ-3-ACH: writing
3. <u>Apel & Diehm (2014)</u>	Opaque (English)	Regular treatment	Small group	Researcher	100m/week for 8 weeks =13.3h	At-risk of LD	a. = K b. = 1 c. = 2	Both	MA _E ; RF _{STD} ; PD _{STD} ; RC _{STD} ; MS _E	Rehit ¹ ; TOWRE: sight word efficiency; TOWRE: phonemic decoding efficiency; TOSREC;

¹ Although the Affix identification task had a slightly better inter-item reliability (.92 vs .90), we chose the Rehit task because it was undertaken by all three grade subgroups.

4. <u>Bar-Kochva & Hasselhorn (2017)</u>	Intermediate (German)	Alternative linguistic treatment (same words segmented in a non-morphological way)	Individual	Researcher	12 sessions of 15-25m =4h	SLL	5-6	Both	RF _{STD} ; MR _E ; RC _{STD} ; WS _{STD}	Spelling multimorphemic words SLRT-2: word reading; Reading of untrained words; ELFE 1-6; DRT-5/RST 6-7: spelling
5. <u>Bar-Kochva et al. (2020)</u>	Intermediate (German)	Alternative linguistic treatment (same words segmented in a non-morphological way)	Individual	Researcher	12 sessions of 15-20m =3.5h	At-risk of LD	5-6	Both	RF _{STD} ; MR _E ; RC _{STD} ; WS _E	SLRT-2: word reading; Reading of untrained words; ELFE 1-6; Spelling of untrained items
6. <u>Berninger et al. (2008): study 1</u>	Opaque (English)	Orthographic training	Small group	Researcher	2h/day for 14 days =28h	Learning disabled	a.4-6 b.7-9	L (derivation)	PD _{STD} ; PR _{STD}	TOWRE: phonemic decoding efficiency; WJ-3: spelling sounds
7. <u>Bowers (2006)</u>	Opaque (English)	Regular treatment	Large group	Researcher	20 lessons of 50m =16.6h	Typical learners	4-5	L (derivation)	MA _E ; RA _{STD} ; WS _{STD}	Morphological choice; WJ-R: word identification; WRAT-3: spelling
8. <u>Brimo (2016)</u>	Opaque (English)	Reading-based training	Small group	Researcher	75m/week for 10 weeks = 12.5h	Learning disabled	3	Both	MA _E ; MS _E	Affix identification task; Spelling multimorphemic words
9. <u>Casalis & Colé (2009): study 1</u>	Opaque (French)	Phonology/phonics training (phonological training)	Small group	Researcher	12 sessions of 30m = 6h	Typical learners	K	Both	MA _E ; V _{STD} ; PA _E	Derivation in sentence context; Picture vocabulary test; Phoneme deletion

10. Casalis et al. (2017)	Opaque (French)	Orthographic training	Large group	Non-researcher (teacher)	24 sessions of 20m = 8h	Typical learners	3	L (derivation)	MA _E ; WS _{STD} ; MS _E ; V _{STD}	Definition; L2MA-2: Corbeau dictation test; Base extraction; PPVT
11. Chow et al. (2008)	Not applicable (Cantonese)	Reading-based training (Dialogic reading)	Individual	Non-researcher (parent)	40m/week for 12 weeks = 8h	Typical learners	K	L (compounding)	MA _E ; RA _E ; V _E	Morphological construction; List of characters adapted from the HKT-SpLD; Hong Kong Cantonese Receptive Vocabulary Test + PPVT
12. Colchester (2010)	Opaque (English)	Orthographic training	Large group	Researcher	20m/week for 5 weeks = 1,67h	Typical learners	1-2	G	MS _E	Pseudoword tasks (total score)
13. Colenbrander et al. (2021)	Opaque (English)	Reading-based training (Motivated Reading)	Small group	Non-researcher (teaching assistant)	1h/week for 24 weeks = 24h	At-risk of learning disabilities	3-5	L (derivation)	RF _{STD} ; MR _E ; PD _{STD} ; RC _{STD} ; WS _E ; MS _{STD} ; PR _{STD} ; V _E	TOWRE: sight word efficiency; Generalization nonwords; TOWRE: phonemic decoding efficiency; NGRT; Morphological spelling task: untrained words; MoSTn; DiSTn; Multiple-choice vocabulary task
14. Crosson et al. (2021)	Opaque (English)	Regular treatment	Large group	Non-researcher (teacher)	80 lessons of 15m = 20h	SLL	7-8	L (derivation)	MA _E ; V _{STD}	Morphological analysis task;

15. Dallashah-Khatib et al. (2014)	Opaque (Arabic)	Phonology/phonics training (phonological training)	Small group	Researcher	112.5m /week for 8 weeks = 15h	Typical learners	K	Both	MA _E ; VE; PA _E	Evaluation of Academic Vocabulary Combined morphology score; Adaptation of the PPVT; Combined phonology score Experimental spelling test
16. Devonshire & Fluck (2010): study 1	Opaque (English)	Alternative linguistic treatment (sharing of spelling experiences and strategies)	Small group	Researcher	1 session of 30m = 0.5h	Typical learners	K to G5	L (derivation)	WS _E	
17. Devonshire & Fluck (2010): study 2	Opaque (English)	Orthographic training (Nelson Spelling Scheme)	Large group	Researcher	9 sessions of 35m = 5.25h	Typical learners	2-3	L (derivation)	WS _{STD} ; MS _E	Schonell spelling test; Correctly-spelled morphemes
18. Elbro & Arnbak (1996): study 3; Arnbak & Elbro (2000)	Opaque (Danish)	Regular treatment	Small group	Non-researcher (teacher)	33 sessions of 15m = 8.25h	Learning disabled	4-5	Both	MA _E ; RA _E ; PDE; RC _{STD} ; WS _E ; V _{STD} ; PA _E	Morphological analogy; Word reading (total number correct); Nonword reading; SL60: passage comprehension; Word spelling (total number correct); PPVT; Phonological discrimination Summed scores of the three morphological tasks; Word spelling;
19. Fejzo et al., (2019)	Opaque (French)	Regular treatment	Large group	Researcher	55m/week for 20 weeks = 18.3h	Typical learners	4	L (derivation)	MA _E ; WS _E ; MS _E	

20. Frank (2008)	Opaque (English)	Reading-based training (independent reading)	Large group	Non-researcher (teacher)	33 sessions of 20m = 11h	Typical learners	9	Both	RF _{STD} ; RC _{STD} ; WS _{STD}	Morpheme spelling STAR: passage reading; NDRT: passage reading; Spelling Feature Assessment
21. Gebauer et al. (2012)	Intermediate (German)	Regular treatment	Individual	Researcher	2h/week for 5 weeks = 10h	Typical learners	N/A ²	L (derivation)	RC _{STD} ; WS _{STD}	ELFE 1-6; HSP: word spelling
22. Gellert et al., (2021)	Opaque (Danish)	Vocabulary training (context-based training)	Small group	Researcher	24 lessons of 30m = 12h	At-risk of learning disabilities	5	Both	MA _E ; RC _{STD} ; V _{STD}	Definition of proximal transfer words; Text reading; PPVT
23. Georgiou et al. (2021)	Opaque (English)	Phonology/ phonics training (Simplicity)	Individual	Researcher	12h	At-risk of learning disabilities	3	L (derivation)	MA _{STD} ; RA _{STD} ; PD _{STD}	MRT; WRAT-4: word reading; Decoding factor: combination of several standardized nonword reading tests
24. Harris et al. (2011)	Opaque (English)	Vocabulary training (LNCS strategy)	Large group	Researcher	7.5h	Typical learners	9	L (derivation)	MA _E ; V _E	Morphological analysis test; Word knowledge test
25. Jöbstl et al., (2021)	Intermediate (German)	Regular treatment	Large group	Researcher	210m/week for 8 weeks = 28h	Typical learners	2	Both	MA _E ; RF _{STD} ; MR _E ; PD _{STD} ;	Morphological awareness;

² Grade of participants is not mentioned by the author. However, ages ranged from 10 to 15 and participants did not suffer from any reading deficit, so it was assumed that they came from a broad range of grades.

									RC _{STD} ; WS _{STD} ; MS _E	SLRT-2: word reading; Suffixed nonwords ³ ; SLRT-2: nonword reading; SLS 1-4; SLRT-2: spelling; Morphological stems
26. <u>Kolne (2019)</u>	Opaque (English)	Vocabulary training	Large group	Researcher	1h/day for 10 days = 10h	Typical learners	a. 3 b. 5	L (derivation)	MA _E ; WSE; MS _E ; V _E	Morphological production task; Taught words spelling (word type 1); Stem spelling (word types 2 and 4); Definitions task (untaught items)
27. <u>Lyster (2002)</u>	Transparent (Norwegian)	Phonology/ phonics training (phonological training)	Small group	Non-researcher (teacher)	15-20 sessions of 30-40m = 10.2h	Typical learners	P	Both	MA _E ; V _E ; PA _E	Word compounds; Homophones; Syllable counting
28. <u>Rassel (2020)</u>	Opaque (French)	Vocabulary training	Large group	Non-researcher (teacher)	14 sessions of 45m = 10.5h	At-risk of learning disabilities	a. 2 b. 3	L (derivation)	MA _{STD} ; RA _{STD} ; RF _{STD} ; MR _E ; PD _E ; RC _{STD} ; V _E ; PA _{STD}	MORPHOTE: production task; Alouette-Revised accuracy index; BALE: text reading; Derived words;

³ Suffixed nonwords were preferred over prefixed nonwords because included studies focused overall more on suffixes than on prefixes, thus this measure was considered more generalizable.

29. Ravid & Geiger (2009)	Opaque (Hebrew)	Alternative linguistic treatment (non-verbal humour training)	Large group	Researcher	16 lessons of 45m = 12h	Typical learners	4	Both	MA _E	Pseudoword reading fluency; TéCoPé; Total vocabulary score; EVALEC: initial phoneme deletion (CCV syllable structure) Derivation of pseudowords ⁴
30. Taha & Saiegh-Haddad (2016)	Opaque (Arabic)	Phonology/phonics training (phonological training)	Small group	Researcher	90m/week for 24 weeks = 36h	1. At-risk of learning disabilities 2. Typical learners	a. 2 b. 4 c. 6	Both	MA _E ; WS _E ; PA _E	Morphological awareness; Word spelling production; Phonological awareness Receptive unexposed; Word reading fluency unexposed; Spelling accuracy unexposed; WISC-4: vocabulary Spelling pseudo-compounds
31. Torkildsen et al. (2021)	Transparent (Norwegian)	Alternative non-linguistic treatment (math training)	Individual	Non-researcher (teacher)	38 sessions of 13m = 8.2h	Typical learners	2	Both	MA _E ; MR _E ; WS _E ; V _{STD}	Receptive unexposed; Word reading fluency unexposed; Spelling accuracy unexposed; WISC-4: vocabulary Spelling pseudo-compounds
32. Tsesmeli (2017)	Transparent (Greek)	Regular treatment	Large group	Researcher	5 sessions of 2h = 10h	Typical learners	a. 1 b. 2	L (compounding)	WS _E	Spelling pseudo-compounds

⁴ We chose to use the root scores, as we considered that it was a more language-generalizable concept than rhyme or pattern scores.

33. <u>Vaknin-Nusbaum & Raveh (2019)</u>	Opaque (Hebrew)	Reading-based training (reading comprehension training)	Small group	Non-researcher (teacher)	24 sessions of 45m = 18h	At-risk of learning disabilities	5	Both	MA _{STD} ; RF _{STD} ; RC _{STD}	Alef-Taf: pseudoword production; Alef-Taf: number of words read in one minute; Elul: text reading comprehension
34. <u>Wang & McBride (2016)</u>	Not applicable (Mandarin Chinese)	Phonology/phonics training (copy + pinyin training)	Small group	Researcher	16 sessions of 30m = 8h	Typical learners	K	L (compounding)	MA _E ; MR _E ; WS _E ; PA _E	Morphological awareness; Chinese word reading (trained set); Chinese word writing; Phonological awareness
35. <u>Weiss et al., (2010)</u>	Intermediate (German)	Regular treatment	Individual	Researcher	30-40h	Typical learners	3 to 9	L (derivation)	RC _{STD} ; WS _{STD}	SLS 1-4/5-8: sentence reading; HSP: word spelling
36. <u>Wolter & Dilworth (2014)</u>	Opaque (English)	Orthographic training	Small group	Researcher	90m/day for 9 days = 13.5h	At-risk of learning disabilities	2	Both	RA _{STD} ; PD _{STD} ; RC _{STD} ; WS _{STD}	WRMT-R: word identification; WRMT-R: word attack; WRMT-R: passage comprehension; TWS-4: spelling
37. <u>Wu et al., (2009)</u>	Not applicable (Mandarin Chinese)	Regular treatment	Large group	Non-researcher (teacher)	N/A	Typical learners	2	L (compounding)	MA _E	Radical meaning judgement
38. <u>Zhang (2016)</u>	Opaque (English)	Regular treatment	Large group	Non-researcher (teacher)	400m = 6.6h	Second-language learners	4	L (derivation)	MA _E ; MR _E	English morphological relatedness; English word reading fluency

39. <u>Zhang & Li (2016)</u>	Opaque (English)	Regular treatment	Large group	Non-researcher (teacher)	400m = 6.6h	Second-language learners	4	L (derivation)	MA _E ; MS _E	Morphological relatedness; Morphological production
40. <u>Zoski & Erickson (2017); Zoski (2015)</u>	Opaque (English)	Phonology/ phonics training (phonological/ letter knowledge training)	Small group	Researcher	2h/week for 6 weeks = 12h	At-risk of learning disabilities	K	Both	MA _E ; RF _{STD} ; WS _{STD} ; PA _{STD} ;	Adapted from the Kindergarten Oral Morphological Production Task; TOWRE: composite score of sight word efficiency and phonetic decoding efficiency; The single word morphological spelling test; CTOPP: composite score of elision and blending words subtests

Table S2*Abbreviations of Outcome Variables and Tests*

Abbrevia- tion	Description
BALE	Analytical battery of reading and writing
CTOPP	Comprehension Test of Phonological Processing
DiSTn	Diagnostic Spelling Test – non words
DRT5	Diagnostic Spelling Test for Fifth Graders
ELFE 1/6	A Reading Comprehension Test for First to Sixth Graders
EVALEC	Computerized battery for the diagnostic assessment of specific reading disabilities
F _E	Experimental reading fluency task
F _{STD}	Standardized reading fluency task
G	Grammatical (inflectional) morphology
HKT-SpLD	Hong Kong Test of Specific Learning Difficulties in reading and writing
HSP	Hamburger-Schreibprobe
K	Kindergarten
L	Lexical (derivational and/or compositional) morphology
L2MA-2	Battery of oral language, written language, memory, attention – Second edition
MA _E	Experimental morphological awareness task
MA _{STD}	Standardized morphological awareness task
MOR- PHOTE	A French battery assessing morphological awareness for primary school children
MoSTn	Morphological Spelling Test – non-words
MR _E	Experimental morphological reading task
MR _{STD}	Standardized morphological reading task.

MRT	Morphological Relatedness Test
MS _E	Experimental morpheme spelling task
MS _{STD}	Standardized morpheme spelling task
NDRT	Nelson-Denny Reading Test
NGRT	New group reading test
P	Preschool
PA _E	Experimental phonological awareness task
PA _{STD}	Standardized phonological awareness task
PD _E	Experimental phonological de- coding task
PD _{STD}	Standardized phonological de- coding task
PPVT	Peabody Picture Vocabulary Test
PR _E	Experimental phonological re- coding task
PR _{STD}	Standardized phonological re- coding task
RA _E	Experimental reading accuracy task
RA _{STD}	Standardized reading accuracy task
RC _E	Experimental reading compre- hension task
RC _{STD}	Standardized reading compre- hension task
RST 6/7	Spelling test for 6th and 7th graders
SL60	The sentence reading tests
SLRT2	Reading and Spelling Tests
SLS 1/4	Salzburger-Lese-Screening 1-4
SLS 5/8	Salzburger-Lese-Screening 5-8
TeCoPé	Test of Sentence Comprehen- sion
TOSREC	Test of Silent Reading Com- prehension
TOWRE	Test of Word Reading Efficien- cy
TWS-4	Test of Written Spelling – Fourth edition
V _E	Experimental vocabulary task
V _{STD}	Standardized vocabulary task
WIAT	Weschler Individual Achieve-

	ment Test
WISC-4	Weschler Intelligence Scale for Children - Fourth edition
WJ-3	Woodcock-Johnson – Third edition
WJ-3-ACH	Woodcock-Johnson tests of achievement – Third edition
WJ-R	Woodcock Johnson-Revised
W _{STD}	Standardized writing task
WRAT-3	Wide Range Achievement Test – Third Edition
WRAT-4	Wide Range Achievement Test - Fourth Edition
WRMT-R	Woodcock Reading Mastery Test-Revised
W _{SE}	Experimental word spelling task
W _{STD}	Standardized word spelling task

Table S3*Types of Morphological Exercises: Coding of Studies*

	Explicit learning				Implicit learning	Type of MA exercise	Computerized
	Morphological awareness exercises	Instruction on word origins	Links to spelling	Links to context			
1. Abbott & Berninger, 1999	X	X				MA	No
2. Allen & Lembke, 2020	X		X	X		Thorough training	No
3. Apel & Diehm, 2014	X		X	X		Thorough training	No
4. Bar-Kochva et al., 2020					X	Implicit learning	Yes
5. Bar-Kochva & Hasselhorn, 2017					X	Implicit learning	Yes
6. Berninger et al., 2008 (study 1)	X		X			MA+spelling	No
7. Bowers, 2006	X	X	X			Thorough training	No
8. Brimo, 2015	X		X	X		Thorough training	No
9. Casalis & Colé, 2009 (study 1)	X					MA	No
10. Casalis et al., 2017	X		X	X		Thorough training	No
11. Chow et al., 2008	X			X		MA+context	No
12. Colchester, 2010	X		X			MA+spelling	Yes

13. Colenbrander et al., 2021	X	X	X		Thorough training	No
14. Crosson et al., 2021	X	X		X	Thorough training	No
15. Dallahsheh-Khatib et al., 2014	X				MA	No
16. Devonshire & Fluck, 2010 (study 1)	X				MA	No
17. Devonshire & Fluck, 2010 (study 2)	X	X	X		Thorough training	No
18. Elbro & Arnbak, 1996 (study 3)	X				MA	No
19. Fejzo et al., 2019	X		X		MA+spelling	No
20. Frank, 2008	X	X	X		Thorough training	No
21. Gebauer et al., 2012	X		X	X	Thorough training	No
22. Gellert et al., 2021	X		X	X	Thorough training	No
23. Georgiou et al., 2021	X	X	X		Thorough training	No
24. Harris et al., 2011	X		X		MA+spelling	No
25. Jöbstl et al., 2021	X		X		MA+spelling	Yes
26. Kolne, 2020	X		X		MA+spelling	No
27. Lyster, 2002	X				MA	No
28. Rassel, 2020	X		X	X	Thorough training	No
29. Ravid & Geiger, 2009	X				MA	No

30. Taha & Saiegh-Haddad, 2015	X		X		MA+spelling	No
31. Torkildsen et al. 2021	X		X	X	Thorough training	Yes
32. Tsesmeli, 2017	X		X		MA+spelling	No
33. Vaknin-Nusbaum & Raveh, 2019	X			X	MA+context	No
34. Wang & McBride, 2016	X	X	X		Thorough training	No
35. Weiss et al., 2010	X		X	X	Thorough training	Yes
36. Wolter & Dilworth, 2013	X		X	X	Thorough training	No
37. Wu et al., 2009	X	X			MA	No
38. Zhang, 2016	X		X		MA+spelling	No
39. Zhang & Li, 2016	X		X		MA+spelling	No
40. Zoski & Erickson, 2017	X			X	MA+context	No

Table S4*Definitions and Examples of Types of Morphological Exercises*

Type of exercise	Definition	Examples
Morphological awareness	<ul style="list-style-type: none"> - Conscious identification, deletion or manipulation of morphemes; - Instruction on meaning of morphemes (e.g., agentive role of the suffix -er in English; “the base carries the core meaning of a word”; etc). 	<p>Word sorting (e.g., Apel & Diehm, 2014); identification of target morpheme in a word; segmentation of a word into morphemes (e.g., Berninger et al., 2008); producing words sharing a common morpheme; word analysis (e.g., Gellert et al., 2021); choosing the appropriate suffix out of several propositions; ordering morphemes to create words; counting morphemes (e.g., Weiss et al., 2010) e.g., What word is left when you take away “door” in <i>doorstep</i>?</p>
Instruction on word origins	<p>Lessons focusing on the etymology of words, such as Latin/Greek roots (e.g., <i>equi-</i>, <i>phon-</i>, etc), or subcharacter units instruction (origins of radicals, shape-to-meaning connections of ideographic characters, etc.)</p>	<p>Shape-to-meaning instruction of pictograms (e.g., Wang & McBride, 2016); instruction on twin bases (e.g., Bowers & Kirby, 2010) e.g., The twin bases <i>duce/duct</i> both mean “to lead, to build” as in <i>product</i>, <i>produce</i>, etc</p>
Links to spellings	<p>Learned morphemes are systematically written by the students or spelling irregularities are explicitly studied or remarked on.</p>	<p>Spelling algorithms (e.g., Bowers & Kirby, 2010); connect morphological analysis to spelling (e.g., Casalis et al., 2017) e.g., “Drop the final <i>e</i>” rule (i.e., when a suffix starts with a vowel and the base ends with a silent <i>e</i>, the silent <i>e</i> is dropped as in <i>forgive</i> - <i>forgiving</i>)</p>

Links to context	Students are prompted to produce a morphologically complex word in the context of a sentence or text or to derive its meaning from context.	“Say it another way” or morphological closes (e.g., Apel & Diehm, 2014); “identify the correct word or morpheme to fit a discourse” (e.g., Torkildsen et al., 2021) e.g., Someone who <i>reads</i> is a ... (<i>reader</i>)
Implicit learning	Implicit emphasis on morphological structure of words without any prompt of conscious manipulation.	Exposure to written morphological segmentation (e.g., Bar-Kochva & Hasselhorn, 2017; Bar-Kochva et al., 2020) e.g., <i>Un#break#able</i>

5.2.8. Figures

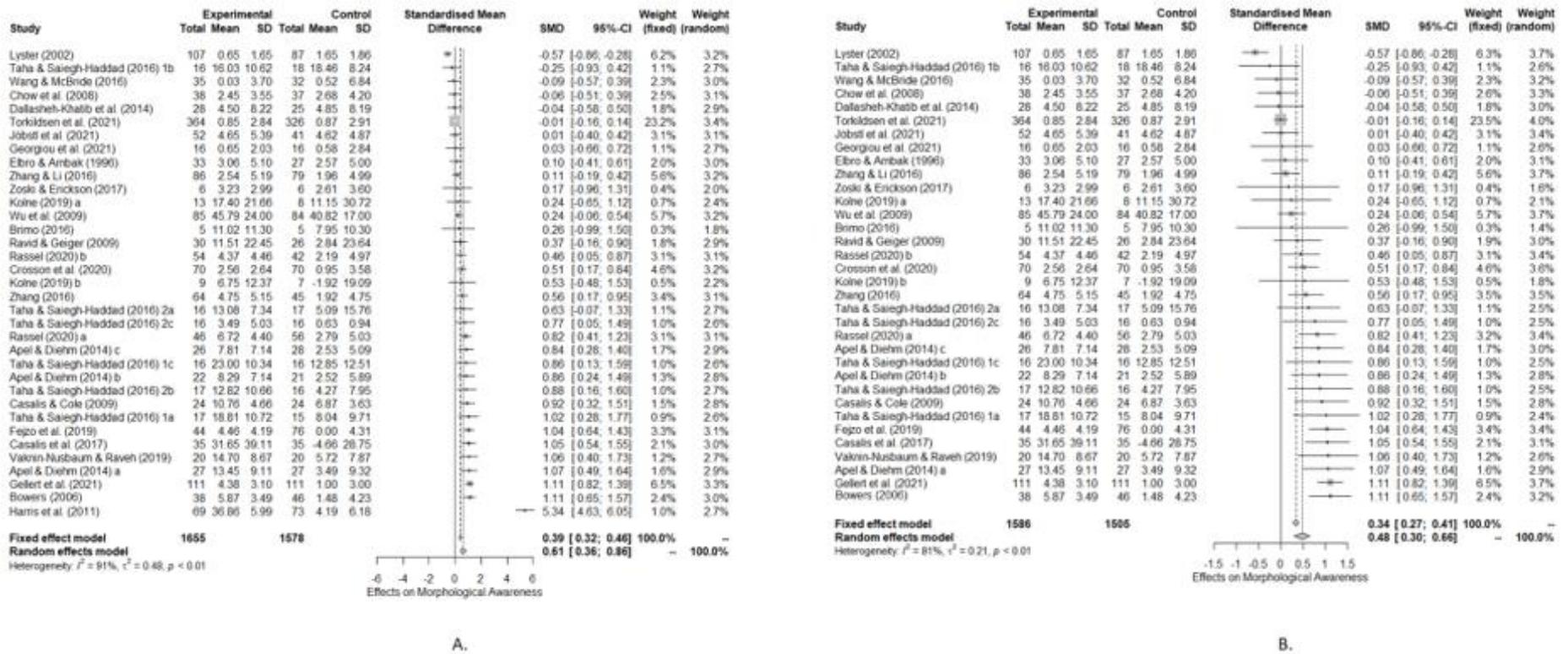


Figure S1
Funnel Plots for all of our Outcome Variables

* This graph excludes the outlier study by Harris et al. (2011)

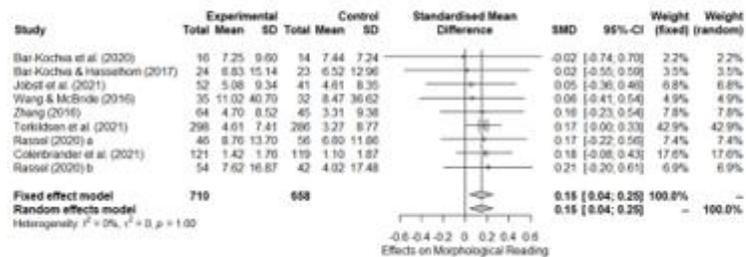
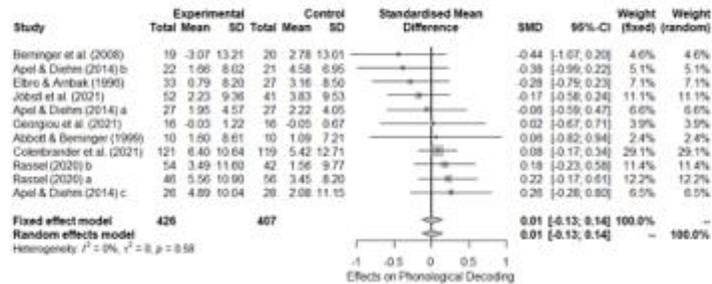
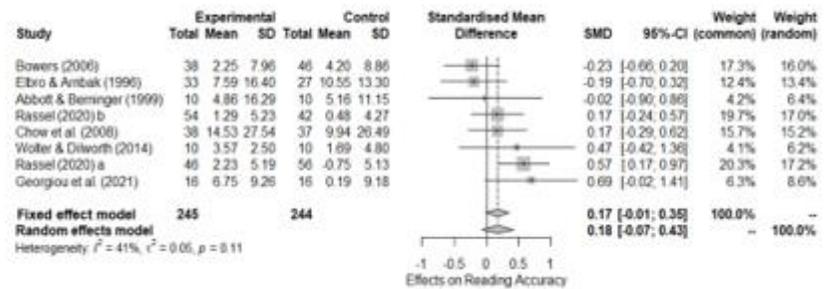
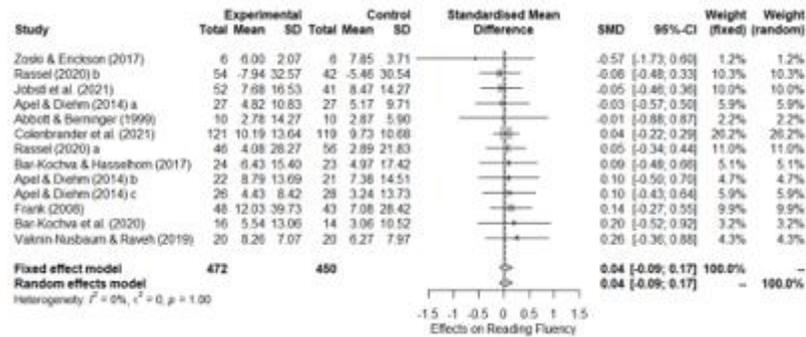


Figure S2

Forest Plots for the Morphological Awareness Outcome

A. Including the study by Harris et al. (2011)

B. Excluding the study by Harris et al. (2011)

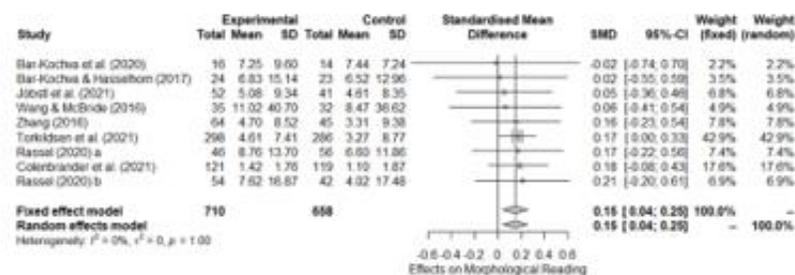
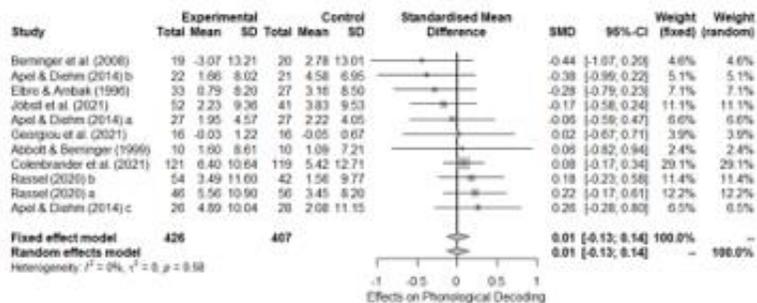
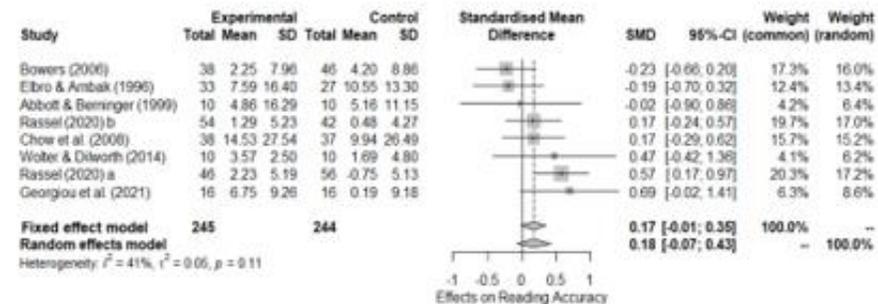
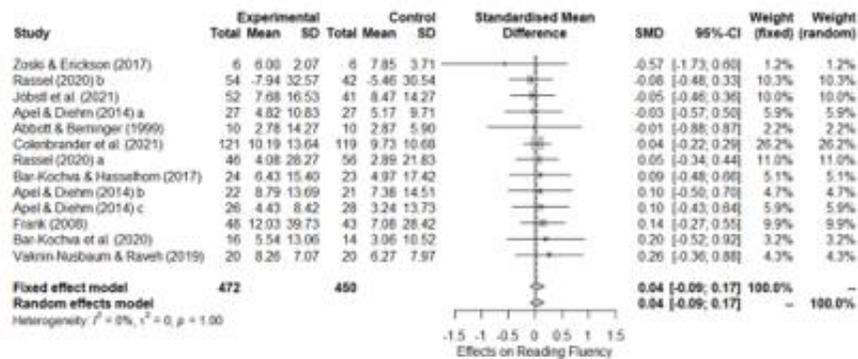


Figure S3

Forest Plots for the Reading Accuracy, Reading Fluency, Morphological Reading and Phonological Decoding Outcomes

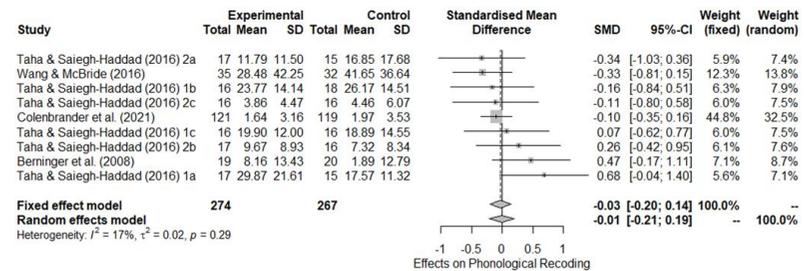
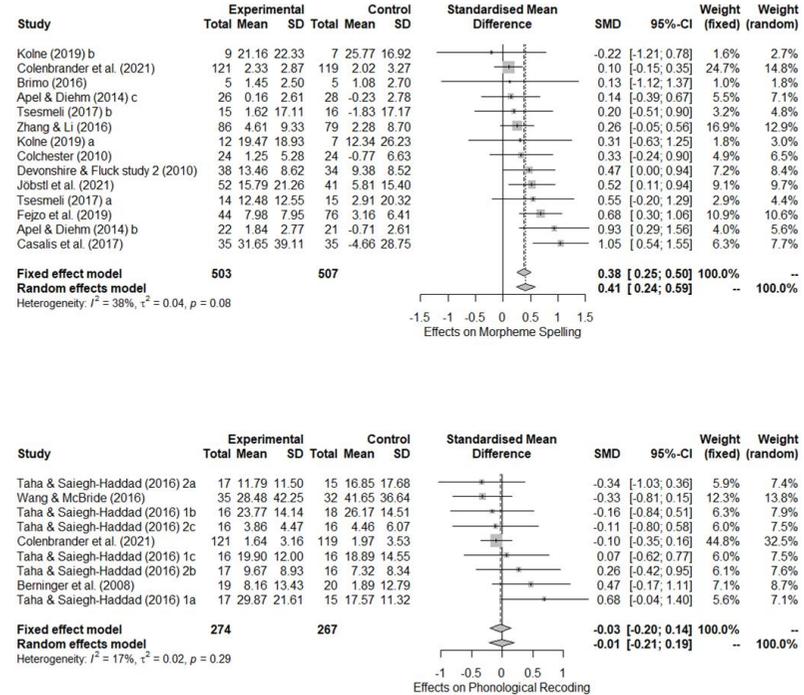
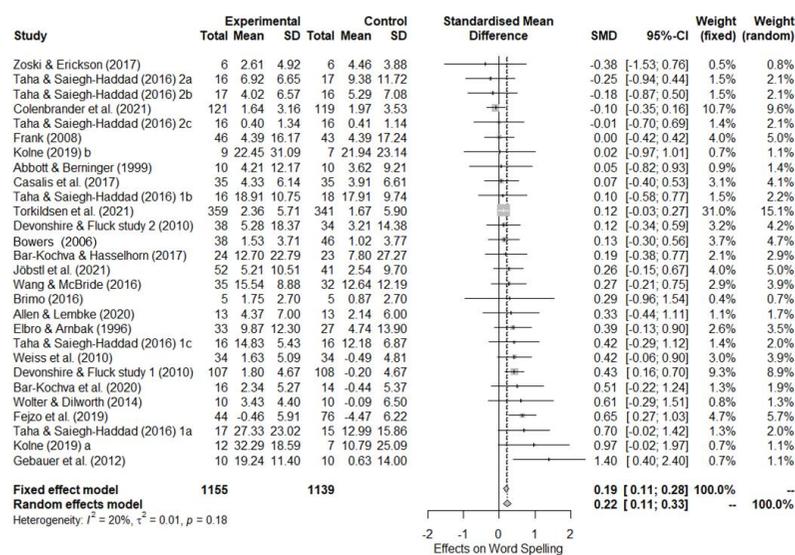


Figure S4

Forest Plots for the Word Spelling, Morpheme Spelling and Phonological Recoding Outcomes

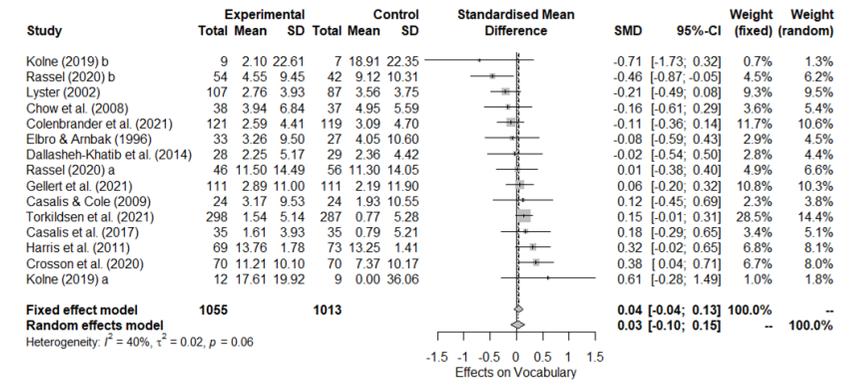
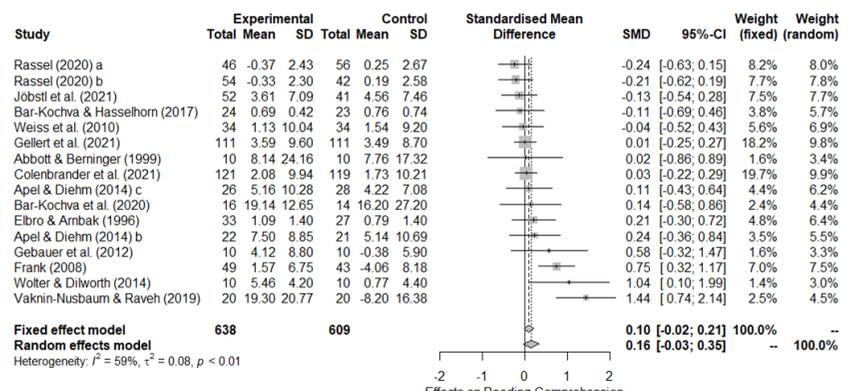


Figure S5
Forest Plots for the Reading Comprehension and Vocabulary Outcomes

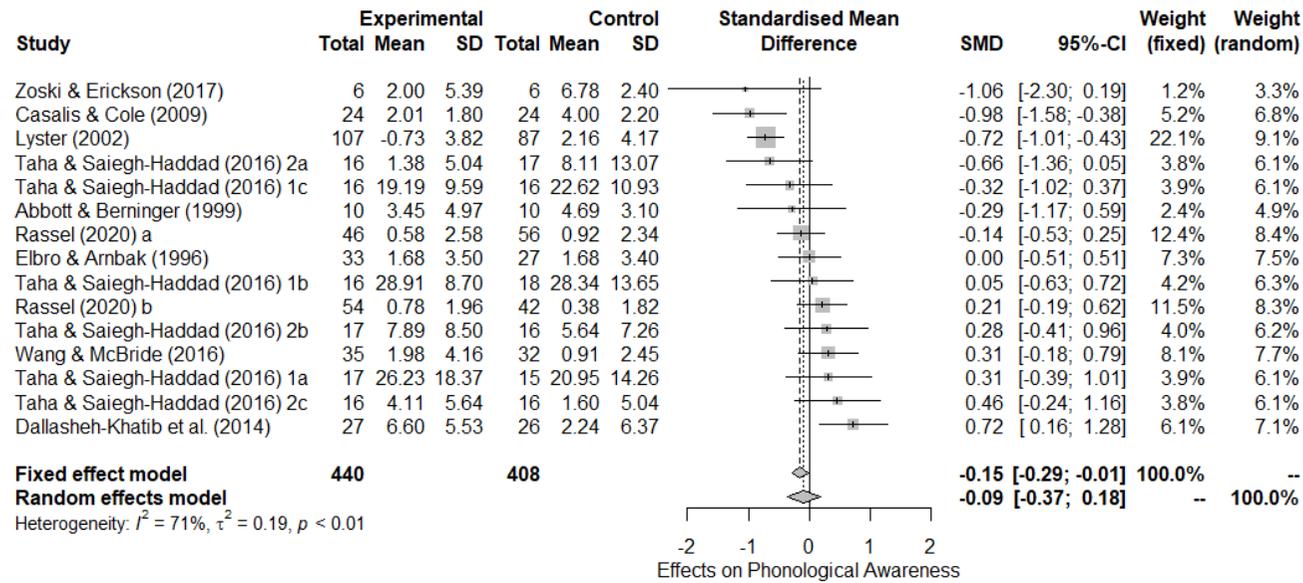


Figure S6
Forest Plot for the Phonological Awareness Outcome

5.2.9 Appendices

Appendix S1

Subgroup Analyses: Morphological Awareness

$Q_{total}(33) = 173.07^{***}, k = 34^1$								
	k	SMD	τ^2	τ	I^2	95% CI		Q_{within}
						LL	UL	
Comparison group, $Q_{between}(6) = 35.67^{***}$								
Orthographic training	1	1.05	--	--	--	0.54	1.55	--
Phonology/ phonics training	12	0.33	0.32	0.57	78.0%	-0.04	0.70	50.09
Reading-based training	3	0.42	0.35	0.60	73.1%	-0.40	1.22	7.44
Vocabulary training	5	0.73	0.07	0.26	55.0%	0.41	1.06	8.88
Other alternative linguistic treatment	1	0.37	--	--	--	-0.16	0.90	--
Alternative non-linguistic treatment	1	-0.01	--	--	--	-0.16	0.14	--
Regular treatment or instruction	11	0.56	0.12	0.35	73.3%	0.31	0.80	37.46
Intervention unit, $Q_{between}(2) = 22.82^{***}$								
Individual	3	-0.01	0.00	0.00	0.0%	-0.15	0.13	0.06
Small group	18	0.53	0.39	0.63	83.5%	0.20	0.86	102.82
Large group	13	0.54	0.09	0.30	66.8%	0.33	0.75	36.15
Instructor, $Q_{between}(1) = 2.49$								
Researcher	22	0.59	0.16	0.40	65.5%	0.37	0.80	60.96
Non-researcher	12	0.32	0.16	0.40	84.4%	0.07	0.57	70.34
Duration, $Q_{between}(3) = 4.37$								
Less than 10h	8	0.29	0.10	0.32	75.8%	0.02	0.56	28.98
Between 10h and 15h	13	0.49	0.39	0.62	85.4%	0.11	0.88	82.20
15h or more	12	0.62	0.15	0.39	67.4%	0.34	0.90	33.76
N/A	1	0.24	--	--	--	-0.07	0.54	--
Type of students, $Q_{between}(3) = 7.18$								
Typical learners	17	0.38	0.22	0.47	83.3%	0.12	0.64	95.73
Second-language learners	3	0.38	0.04	0.19	53.3%	0.09	0.67	4.29
At-risk of learning disability	12	0.72	0.09	0.30	54.6%	0.48	0.96	24.21
Learning disabled	2	0.12	0.00	0.00	0.0%	-0.35	0.59	0.05

Grade, $Q_{between}(2) = 2.26$								
P-G1	9	0.25	0.27	0.52	82.9%	-0.14	0.64	46.66
G2-G5	22	0.55	0.20	0.44	80.4%	0.33	0.77	107.12
G6-G9	3	0.60	0.00	0.00	0.0%	0.32	0.88	1.01
G10-G12	0	--	--	--	--	--	--	--
Broad grade	0	--	--	--	--	--	--	--
Language transcription system, $Q_{between}(1) = 8.11^{**}$								
Alphabetic	31	0.53	0.23	0.48	81.9%	0.33	0.73	166.01
Non-alphabetic	3	0.10	0.00	0.00	0.0%	-0.13	0.32	1.87
Language opacity, $Q_{between}(3) = 24.76^{***}$								
Opaque	28	0.64	0.10	0.31	58.4%	0.48	0.80	64.96
Intermediate	1	0.01	--	--	--	-0.40	0.42	--
Transparent	2	-0.28	0.15	0.38	91.4%	-0.83	0.28	11.67
Type of morphology, $Q_{between}(1) = 0.00$								
Lexical	14	0.48	0.11	0.33	70.0%	0.26	0.70	43.31
Grammatical	0	--	--	--	--	--	--	--
Both	20	0.49	0.29	0.54	84.5%	0.21	0.76	122.24
Type of morphological exercise, $Q_{between}(3) = 3.55$								
MA	6	0.14	0.22	0.47	82.5%	-0.28	0.56	28.64
MA+spelling	13	0.90	1.24	1.11	94.1%	0.26	1.53	202.05
MA+context	3	0.39	0.34	0.58	73.2%	-0.39	1.18	7.47
Thorough training	12	0.51	0.12	0.35	61.2%	0.25	0.78	28.36
Implicit learning	0	--	--	--	--	--	--	--
Computerized training, $Q_{between}(1) = 18.70^{***}$								
Yes	2	-0.01	0.00	0.00	0.0%	-0.15	0.13	0.00
No	32	0.52	0.22	0.46	78.1%	0.33	0.71	141.31

¹ This subgroup analysis excludes the outlier study by Harris et al. (2011).

** = $p < .01$. *** = $p < .001$.

Appendix S2

Subgroup Analyses: Reading Comprehension

$Q_{total}(15) = 36.16^{**}$, $k = 16$								
	k	SMD	τ^2	τ	I^2	95% CI		Q_{within}
						LL	UL	
Comparison group, $Q_{between}(4) = 9.20$								
Orthographic training	1	1.05	--	--	--	0.10	1.99	--
Phonology/ phonics training	0	--	--	--	--	--	--	--
Reading-based training	3	0.69	0.39	0.62	89.2%	-0.07	1.44	18.56
Vocabulary training	3	-0.10	0.00	0.00	0.0%	-0.29	0.09	1.45
Other alternative linguistic treatment	3	-0.01	0.00	0.00	0.0%	-0.41	0.39	0.29
Alternative non-linguistic treatment	0	--	--	--	--	--	--	--
Regular treatment or instruction	6	0.07	0.00	0.00	0.0%	-0.14	0.29	2.94
Intervention unit, $Q_{between}(2) = 1.89$								
Individual	5	0.04	0.00	0.00	0.0%	-0.25	0.33	1.83
Small group	7	0.31	0.10	0.32	67.4%	0.01	0.62	18.40
Large group	4	0.04	0.17	0.41	79.7%	-0.41	0.49	14.76
Instructor, $Q_{between}(1) = 1.01$								
Researcher	10	0.05	0.00	0.00	0.0%	-0.11	0.21	7.30
Non-researcher	6	0.27	0.20	0.45	82.3%	-0.13	0.68	28.26
Duration, $Q_{between}(2) = 0.29$								
Less than 10h	3	0.08	0.00	0.00	0.0%	-0.25	0.42	0.71
Between 10h and 15h	8	0.19	0.10	0.32	64.3%	-0.10	0.48	19.63
15h or more	5	0.21	0.16	0.40	74.7%	-0.22	0.63	15.79
Type of students, $Q_{between}(3) = 1.05$								
Typical learners	4	0.26	0.16	0.40	71.5%	-0.22	0.73	10.53
Second-language learners	1	-0.11	--	--	--	-0.69	0.46	--
At-risk of learning disability	10	0.15	0.08	0.29	61.7%	-0.10	0.40	23.50
Learning disabled	1	0.21	--	--	--	-0.30	0.72	--
Grade, $Q_{between}(3) = 7.98$								
P-G1	1	0.24	--	--	--	-0.36	0.84	--

G2-G5	9	0.13	0.09	0.29	66.9%	-0.12	0.37	24.17
G6-G9	1	0.76	--	--	--	0.33	1.18	--
G10-G12	0	--	--	--	--	--	--	--
Broad grade	5	0.04	0.00	0.00	0.0%	-0.25	0.33	1.83
Language transcription system								
Alphabetic	16	0.16	0.08	0.28	58.5%	-0.03	0.35	36.16
Non-alphabetic	0	--	--	--	--	--	--	--
Language opacity, $Q_{between}(1) = 1.96$								
Opaque	11	0.23	0.11	0.33	69.5%	-0.02	0.48	32.83
Intermediate	5	-0.02	0.00	0.00	0.0%	-0.27	0.23	2.26
Transparent	0	--	--	--	--	--	--	--
Type of morphology, $Q_{between}(1) = 4.40^*$								
Lexical	6	-0.05	0.00	0.00	0.0%	-0.22	0.12	3.81
Grammatical	0	--	--	--	--	--	--	--
Both	10	0.31	0.13	0.36	66.9%	0.02	0.59	27.19
Type of morphological exercise, $Q_{between}(4) = 15.24^{**}$								
MA	2	0.16	0.00	0.00	0.0%	-0.28	0.60	0.14
MA+spelling	1	-0.13	--	--	--	-0.54	0.28	--
MA+context	1	1.44	--	--	--	0.74	2.15	--
Thorough training	10	0.13	0.06	0.24	55.3%	-0.08	0.34	20.12
Implicit learning	2	-0.02	0.00	0.00	0.0%	-0.46	0.43	0.29
Computerized training, $Q_{between}(1) = 1.96$								
Yes	5	-0.02	0.00	0.00	0.0%	-0.27	0.23	2.26
No	11	0.23	0.11	0.33	69.5%	-0.02	0.48	32.83

* = $p < .05$. ** = $p < .01$.

Appendix S3

Subgroup Analyses: Morpheme Spelling

$Q_{total}(13) = 20.81, k = 14, p = .08$								
	k	SMD	τ^2	τ	I^2	95% CI		Q_{within}
						LL	UL	
Comparison group, $Q_{between}(3) = 7.12, p = .07$								
Orthographic training	3	0.62	0.07	0.27	51.8%	0.20	1.05	4.15
Phonology/ phonics training	0	--	--	--	--	--	--	--
Reading-based training	2	0.10	0.00	0.00	0.0%	-0.14	0.35	0.00
Vocabulary training	2	0.06	0.00	0.00	0.0%	-0.62	0.74	0.58
Other alternative linguistic treatment	0	--	--	--	--	--	--	--
Alternative non-linguistic treatment	0	--	--	--	--	--	--	--
Regular treatment or instruction	7	0.45	0.01	0.10	15.6%	0.25	0.65	7.11
Intervention unit, $Q_{between}(1) = 0.74$								
Individual	0	--	--	--	--	--	--	--
Small group	4	0.29	0.07	0.26	47.7%	-0.09	0.67	5.74
Large group	10	0.48	0.01	0.12	18.0%	0.30	0.66	10.98
Instructor, $Q_{between}(1) = 0.03$								
Researcher	11	0.46	0.00	0.00	0.0%	0.29	0.64	7.81
Non-researcher	3	0.42	0.13	0.36	81.7%	-0.04	0.87	10.93
Duration, $Q_{between}(2) = 0.46$								
Less than 10h	4	0.50	0.07	0.27	57.9%	0.16	0.85	7.13
Between 10h and 15h	7	0.35	0.00	0.00	0.0%	0.07	0.63	5.62
15h or more	3	0.41	0.08	0.29	72.5%	0.03	0.79	7.28
Type of students, $Q_{between}(3) = 3.04$								
Typical learners	9	0.54	0.00	0.07	5.0%	0.36	0.73	8.42
Second-language learners	1	0.26	--	--	--	-0.05	0.56	--
At-risk of learning disability	3	0.32	0.10	0.32	65.1%	-0.13	0.77	5.72
Learning disabled	1	0.13	--	--	--	-1.12	1.37	--
Grade, $Q_{between}(2) = 2.25$								
P-G1	2	0.77	0.00	0.00	0.0%	0.29	1.25	0.59

G2-G5	11	0.38	0.04	0.20	42.9%	0.18	0.58	17.50
G6-G9	0	--	--	--	--	--	--	--
G10-G12	0	--	--	--	--	--	--	--
Broad grade	1	0.33	--	--	--	-0.24	0.90	--
Language transcription system								
Alphabetic	14	0.41	0.04	0.19	37.5%	0.24	0.59	20.81
Non-alphabetic	0	--	--	--	--	--	--	--
Language opacity, $Q_{between}(2) = 0.29$								
Opaque	11	0.41	0.06	0.24	49.6%	0.20	0.63	19.83
Intermediate	1	0.52	--	--	--	0.11	0.94	--
Transparent	2	0.36	0.00	0.00	0.0%	-0.15	0.87	0.45
Type of morphology, $Q_{between}(2) = 0.21$								
Lexical	9	0.41	0.06	0.24	51.2%	0.17	0.64	16.38
Grammatical	1	0.33	--	--	--	-0.24	0.90	--
Both	4	0.48	0.03	0.16	21.6%	0.14	0.81	3.83
Type of morphological exercise, $Q_{between}(1) = 0.12$								
MA	0	--	--	--	--	--	--	--
MA+spelling	8	0.40	0.00	0.00	0.0%	0.23	0.58	5.33
MA+context	0	--	--	--	--	--	--	--
Thorough training	6	0.47	0.13	0.36	67.3%	0.11	0.84	15.31
Implicit learning	0	--	--	--	--	--	--	--
Computerized training, $Q_{between}(1) = 0.05$								
Yes	2	0.46	0.00	0.00	0.00	0.12	0.79	0.29
No	12	0.41	0.05	0.23	45.7%	0.21	0.62	20.26

Appendix S4

Subgroup Analyses: Vocabulary

$Q_{total}(14) = 23.24, k = 15, p = .06$								
	k	SMD	τ^2	τ	I^2	95% CI		Q_{within}
						LL	UL	
Comparison group, $Q_{between}(5) = 6.34$								
Orthographic training	1	0.18	--	--	--	-0.29	0.65	--
Phonology/ phonics training	3	-0.12	0.00	0.00	0.0%	-0.35	0.11	1.19
Reading-based training	2	-0.12	0.00	0.00	0.0%	-0.34	0.10	0.04
Vocabulary training	6	0.00	0.06	0.25	58.4%	-0.28	0.28	12.03
Other alternative linguistic treatment	0	--	--	--	--	--	--	--
Alternative non-linguistic treatment								
Alternative non-linguistic treatment	1	0.15	--	--	--	-0.02	0.31	--
Regular treatment or instruction								
Regular treatment or instruction	2	0.19	0.06	0.24	53.3%	-0.25	0.63	2.14
Intervention unit, $Q_{between}(2) = 1.38$								
Individual	2	0.07	0.02	0.13	36.0%	-0.19	0.33	1.56
Small group	6	-0.06	0.00	0.00	0.0%	-0.20	0.07	2.39
Large group	7	0.09	0.08	0.28	60.8%	-0.20	0.37	15.32
Instructor, $Q_{between}(1) = 1.31$								
Researcher	6	0.13	0.01	0.07	8.3%	-0.06	0.32	5.45
Non-researcher	9	-0.02	0.03	0.17	52.0%	-0.18	0.14	16.65
Duration, $Q_{between}(2) = 3.41$								
Less than 10h	6	0.14	0.00	0.00	0.0%	0.01	0.26	3.49
Between 10h and 15h	6	-0.11	0.03	0.19	44.8%	-0.35	0.12	9.05
15h or more	3	0.08	0.05	0.23	61.9%	-0.25	0.41	5.25
Type of students, $Q_{between}(3) = 5.95$								
Typical learners	9	0.06	0.02	0.13	30.5%	-0.10	0.22	11.51
Second-language learners	1	0.38	--	--	--	0.04	0.71	--
At-risk of learning disability	4	-0.09	0.01	0.12	36.1%	-0.29	0.10	4.70

Learning disabled	1	-0.08	--	--	--	-0.59	0.43	--
Grade, $Q_{between}(2) = 9.39^{**}$								
P-G1	4	-0.13	0.00	0.00	0.0%	-0.33	0.08	1.21
G2-G5	9	-0.01	0.02	0.14	38.0%	-0.16	0.15	12.89
G6-G9	2	0.35	0.00	0.00	0.0%	0.11	0.58	0.07
G10-G12	0	--	--	--	--	--	--	--
Broad grade	0	--	--	--	--	--	--	--
Language transcription system, $Q_{between}(1) = 0.66$								
Alphabetic	14	0.04	0.02	0.15	42.1%	-0.09	0.17	22.44
Non-alphabetic	1	-0.16	--	--	--	-0.61	0.29	--
Language opacity, $Q_{between}(2) = 0.74$								
Opaque	12	0.05	0.03	0.16	38.6%	-0.11	0.20	17.92
Intermediate	0	--	--	--	--	--	--	--
Transparent	2	-0.01	0.05	0.22	77.8%	-0.35	0.33	4.50
Type of morphology, $Q_{between}(1) = 0.03$								
Lexical	9	0.03	0.06	0.23	56.3%	-0.18	0.24	18.31
Grammatical	0	--	--	--	--	--	--	--
Both	6	0.05	0.00	0.00	0.0%	-0.06	0.17	4.90
Type of morphological exercise, $Q_{between}(3) = 2.06$								
MA	4	-0.11	0.00	0.00	0.0%	-0.32	0.10	1.21
MA+spelling	3	0.17	0.14	0.38	51.1%	-0.42	0.76	4.09
MA+context	1	-0.16	--	--	--	-0.61	0.29	--
Thorough training	7	0.04	0.02	0.16	53.0%	-0.12	0.21	12.77
Implicit learning	0	--	--	--	--	--	--	--
Computerized training, $Q_{between}(1) = 1.71$								
Yes	1	0.15	--	--	--	-0.01	0.31	--
No	14	0.01	0.02	0.16	38.3%	-0.13	0.14	21.07

** = $p < .01$.

Appendix S5

Subgroup Analyses: Phonological Awareness

$Q_{total}(14) = 49.01^{***}$, $k = 15$								
	k	SMD	τ^2	τ	I^2	95% CI		Q_{within}
						LL	UL	
Comparison group, $Q_{between}(1) = 0.66$								
Orthographic training	0	--	--	--	--	--	--	--
Phonology/ phonics training	11	-0.12	0.31	0.55	77.6%	-0.50	0.26	44.71
Reading-based training	0	--	--	--	--	--	--	--
Vocabulary training	2	0.03	0.02	0.15	34.0%	-0.31	0.38	1.51
Other alternative linguistic treatment	1	-0.29	--	--	--	-1.17	0.59	--
Alternative non-linguistic treatment	0	--	--	--	--	--	--	--
Regular treatment or instruction	1	0.00	--	--	--	-0.51	0.51	--
Intervention unit, $Q_{between}(2) = 0.60$								
Individual	1	-0.29	--	--	--	-1.17	0.59	--
Small group	12	-0.11	0.27	0.52	75.8%	-0.46	0.24	45.40
Large group	2	0.03	0.02	0.15	34.0%	-0.31	0.38	1.51
Instructor, $Q_{between}(1) = 0.20$								
Researcher	11	-0.05	0.22	0.47	64.9%	-0.40	0.30	28.51
Non-researcher	4	-0.18	0.17	0.41	81.4%	-0.63	0.27	16.17
Duration, $Q_{between}(2) = 2.04$								
Less than 10h	3	-0.20	0.33	0.57	81.8%	-0.92	0.51	11.01
Between 10h and 15h	4	-0.33	0.21	0.46	81.3%	-0.85	0.20	16.03
15h or more	8	0.10	0.11	0.32	45.9%	-0.23	0.43	12.94
Type of students, $Q_{between}(2) = 0.07$								
Typical learners	7	-0.10	0.41	0.64	85.0%	-0.62	0.43	39.90
Second-language learners	0	--	--	--	--	--	--	--
At-risk of learning disability	7	-0.03	0.00	0.07	4.4%	-0.26	0.20	6.28
Learning disabled	1	0.00	--	--	--	-0.51	0.51	--
Grade, $Q_{between}(3) = 1.12$								
P-G1	5	-0.29	0.50	0.71	87.9%	-0.98	0.39	32.97
G2-G5	7	0.02	0.00	0.06	4.4%	-0.19	0.22	6.28

G6-G9	2	0.07	0.18	0.42	58.3%	-0.70	0.83	2.40
G10-G12	0	--	--	--	--	--	--	--
Broad grade	1	-0.29	--	--	--	-1.17	0.59	--
Language transcription system, $Q_{between}(1) = 2.28$								
Alphabetic	14	-0.13	0.20	0.44	71.3%	-0.41	0.16	45.28
Non-alphabetic	1	0.31	--	--	--	-0.18	0.79	--
Language opacity, $Q_{between}(2) = 17.25^{***}$								
Opaque	13	-0.05	0.13	0.36	57.8%	-0.32	0.21	28.46
Intermediate	0	--	--	--	--	--	--	--
Transparent	1	-0.72	--	--	--	-1.02	-0.43	--
Type of morphology, $Q_{between}(1) = 1.00$								
Lexical	4	0.07	0.00	0.05	4.5%	-0.17	0.31	3.14
Grammatical	0	--	--	--	--	--	--	--
Both	11	-0.15	0.27	0.52	75.4%	-0.52	0.22	40.58
Type of morphological exercise, $Q_{between}(3) = 4.08$								
MA	5	-0.26	0.38	0.62	84.8%	-0.86	0.34	26.24
MA+spelling	6	0.02	0.06	0.23	30.4%	-0.32	0.36	7.18
MA+context	1	-1.10	--	--	--	-2.30	0.19	--
Thorough training	3	0.10	0.01	0.10	18.5%	-0.17	0.37	2.45
Implicit learning	0	--	--	--	--	--	--	--
Computerized training								
Yes	0	--	--	--	--	--	--	--
No	15	-0.09	0.19	0.44	71.4%	-0.37	0.18	49.01

*** = $p < .001$.

Chapitre 6 – Favoriser le développement de l’identification des mots et de la production orthographique par un entraînement morphologique : efficacité,

limites et perspectives

6.1 Introduction

La conscience morphologique (CM), définie comme la capacité de manipuler les morphèmes (i.e., plus petites unités de sens) au sein des mots (Carlisle, 1995), est largement considérée comme contribuant au développement de la littératie des enfants d’âge scolaire (Goodwin & Ahn, 2010, 2013). Si son intérêt pour la production orthographique est bien connu et intégré aux curricula de primaire et de secondaire en France (Bulletin officiel spécial n°11 du 26 novembre 2015 : Programmes d’enseignement de l’école élémentaire et du collège, 2015), la compréhension de son impact sur l’identification des mots peut paraître moins intuitive. En effet, ainsi que le souligne Rastle (2019), l’importance de la morphologie dans le processus de lecture a pu être négligée du fait que « les modèles théoriques d’acquisition de la lecture ont eu tendance à se concentrer sur des mots monosyllabiques, et bien souvent, monomorphémiques (...) [et] également parce que la recherche dans ce champ s’est concentrée sur de jeunes enfants aux premières étapes de leur apprentissage de la lecture » (p. 45).

Un ensemble de travaux suggère pourtant que, de manière complémentaire à la conscience phonologique (capacité à identifier et manipuler les sons d’une langue ; Carlisle, 1995), la CM contribue directement à l’identification des mots, via ce que nous nommerons décodage morphologique en référence au *Morphological Pathways Framework* (Levesque et al., 2021). Partant du constat que les mots et non-mots complexes sont lus plus précisément et plus rapidement que les mots et non-mots simples (cf. § 3.2.1), il a en effet été suggéré que le

fait d'être capable d'identifier ces morphèmes au sein des mots faciliterait la lecture des mots plurimorphémiques. C'est dans ce but que certaines études d'entraînement ont été conduites (e.g., [Ardanouy et al., 2023](#); [Bar-Kochva et al., 2020](#); [Jöbstl et al., 2021](#)). Ainsi que nous l'avons évoqué dans le chapitre 5 du présent manuscrit, l'instruction morphologique peut comprendre des activités diverses de CM, telles que la segmentation, la fusion ou la génération de familles morphologiques. Une synthèse des résultats des études réalisées, présentée dans le cadre de notre méta-analyse (cf. chapitre précédent), paraît ainsi confirmer qu'un impact positif sur l'identification des mots et non-mots complexes, quoique faible, est effectivement constaté à l'issue d'une intervention morphologique.

Dans notre méta-analyse, les gains obtenus pour l'identification des mots se limitaient à la lecture de ces mots complexes, et ne se généralisaient pas à des mesures comprenant à la fois des mots simples et des mots complexes, qu'elles ciblent la précision ou la fluence de lecture. Ce pattern se retrouve dans la totalité des études à avoir, à notre connaissance, investigué l'impact d'un entraînement morphologique sur la fluence ou la précision de lecture ([Abbott & Berninger, 1999](#); [Apel & Diehm, 2014](#); [Arnbak & Elbro, 2000](#); [Bar-Kochva et al., 2020](#); [Bar-Kochva & Hasselhorn, 2017](#); [Colenbrander et al., 2021](#); [Frank, 2008](#); [Georgiou et al., 2021](#); [Jöbstl et al., 2021](#); [Vaknin-Nusbaum & Raveh, 2019](#); [Wolter & Dilworth, 2014](#); [Zoski & Erickson, 2017](#)), à l'exception de l'étude de [Rassel \(2020\)](#) menée dans le cadre de sa thèse de doctorat. Cette étude a retrouvé des résultats positifs d'un entraînement morphologique sur la précision et la fluence de lecture de texte chez des enfants de 2^{ème} année de primaire présentant un retard de lecture et issus d'un milieu socio-économique très défavorisé. Cet effet dépendait de la nature du groupe contrôle auquel l'intervention morphologique était comparée : ainsi, l'effet sur la précision de lecture était obtenu comparativement à un groupe entraîné au vocabulaire mais pas à un groupe « sans

entraînement ». À l'inverse, l'effet sur la fluence était obtenu comparativement au groupe « sans entraînement », mais pas au groupe entraîné au vocabulaire.

Ainsi, bien que cette étude offre des perspectives encourageantes et relativement inédites, suggérant un impact positif de l'entraînement morphologique sur la lecture, ils demandent à être répliqués en raison du manque de clarté du pattern des résultats. C'est dans cette perspective que nous avons, dans la présente étude, évalué l'impact d'une intervention morphologique sur la lecture des mots de la Batterie Analytique du Langage Écrit (Jacquier-Roux et al., 2010).

Outre ce premier objectif, notre étude visait également à préciser la manière dont un entraînement morphologique peut favoriser le développement de la production orthographique de mots complexes. Plus spécifiquement, nous voulions déterminer dans quelle mesure une intervention morphologique proposant des exercices de manipulation morphémique explicites était plus efficace qu'un simple entraînement à la production orthographique de morphèmes. Parmi les études ayant exploré l'impact d'un entraînement morphologique sur l'orthographe, le groupe de comparaison consistait le plus souvent soit en une instruction usuelle (e.g., Apel & Diehm, 2014; Arnbak & Elbro, 2000; Jöbstl et al., 2021) soit en une liste d'attente (Gebauer et al., 2012), soit en un entraînement alternatif (e.g., Brimo, 2016; Devonshire & Fluck, 2010; Torkildsen et al., 2021) qui n'exposait pas les participants témoins à la forme écrite des morphèmes évalués en pré- et post-test (voir toutefois Casalis et al., 2017; Colenbrander et al., 2021). Les programmes d'intervention morphologique visant l'amélioration de la production orthographique semblent d'autant plus efficaces lorsqu'ils incluent des activités de production orthographique de morphèmes, plutôt que lorsqu'ils présentent ces derniers de manière exclusivement orale (Nunes et al., 2003). Néanmoins, le simple fait d'entraîner des sujets à écrire correctement des cibles conduit à une amélioration

des performances orthographiques sur celles-ci (e.g., Graham, 1999). Ainsi, il nous semblait essentiel que notre entraînement contrôle présente également cet « entraînement orthographique ». Sans cela, il ne nous serait pas possible de déterminer avec certitude qu'une amélioration sur les mesures de production orthographique soit réellement due aux activités de manipulation de morphèmes, et non plus généralement à la production orthographique de ces derniers. Par ailleurs, dans sa recherche de 2016 sur l'étude des liens entre CM et production orthographique de morphèmes chez les élèves de 3^{ème} et de 4^{ème} primaire, Fejzo a trouvé une corrélation positive et significative entre le niveau de CM (plus particulièrement, la dérivation de non-mots complexes) et les capacités d'orthographe des morphèmes. Notons que la CM n'apportait une contribution significative qu'à l'orthographe des suffixes et des morphogrammes (i.e., lettres muettes déductibles sur la base d'un mot dérivé de la même base), et non des bases et des préfixes (Fejzo, 2016). Ces résultats sont cohérents avec l'hypothèse que la CM présente un intérêt majeur pour pallier les inconsistances graphème-phonème en fin de mots (Peereman et al., 2013). Nous avons donc choisi, dans notre étude, d'explorer principalement l'impact d'un entraînement morphologique sur la production de mots suffixés.

Ainsi, notre étude avait deux grands objectifs : d'une part, nous souhaitions déterminer si un entraînement morphologique pouvait résulter en une amélioration dans l'identification des mots d'une tâche standardisée de lecture, et d'autre part, nous souhaitions déterminer si la manipulation explicite de morphèmes permettait d'obtenir des gains en production orthographique de mots suffixés. Dans les deux cas, nous souhaitions estimer l'impact d'une telle intervention, comparée à un entraînement à la seule production orthographique des mots entraînés.

6.2 Étude 1

6.2.1 Méthode

6.2.1.1 Participants

Nous avons recruté nos participants par l'intermédiaire de directeurs et directrices d'écoles primaires du département du Bas-Rhin. Après avoir été informés des tenants et aboutissants de la présente étude, les enseignants de chaque école prenaient contact avec la doctorante s'ils souhaitaient que leurs élèves y participent. Une lettre d'information concernant l'étude et un formulaire de consentement éclairé à destination des élèves et de leur famille étaient ensuite transmis aux classes concernées. Préalablement, nous avons sollicité et obtenu l'autorisation du rectorat ainsi que du Comité d'Éthique pour la Recherche de l'Université de Strasbourg. Soixante-deux élèves de 3^{ème} (n = 12), 4^{ème} (n = 20) et 5^{ème} (n = 30) années de primaire nous ont retourné les formulaires de consentement éclairé et ont donc pu passer les évaluations pré-tests, réalisées en février 2021.

La présente étude visant à estimer l'intérêt d'une intervention morphologique pour la pratique courante, nous n'avons pas exclu les élèves qui présentaient un trouble des apprentissages. En revanche, concernant les élèves de 5^{ème} année de primaire, nous avons exclu une élève qui présentait un handicap visuel, ainsi qu'un élève avec un diagnostic de déficience intellectuelle, considérant que le matériel que nous avons développé risquait de poser des problèmes d'accessibilité pour ces élèves. Une autre élève de 5^{ème} année de primaire ainsi qu'un élève de 4^{ème} année de primaire faisaient partie d'un dispositif ULIS-DYS. Bien qu'intégrés au protocole au commencement de l'étude, ils ont dû en être exclus en raison des consignes sanitaires qui préconisaient de ne pas mélanger les élèves de classes différentes sur le temps de classe. L'échantillon d'élèves ayant bénéficié d'une intervention était donc constitué de 58 enfants : 12 élèves de 3^{ème} année de primaire, 19 élèves de 4^{ème} année de

primaire et 27 élèves de 5^{ème} année de primaire. Avant de commencer l'entraînement, chaque élève était aléatoirement attribué soit au groupe d'intervention morphologique (n = 30), soit au groupe d'intervention contrôle (n = 28), de sorte que chaque classe présente un nombre équivalent d'élèves dans chaque groupe. Néanmoins, en raison de la situation sanitaire liée au COVID-19, l'entraînement a été retardé et n'a pu débuter qu'en mai 2021, ce qui a également repoussé le moment des passations post-tests. Ces dernières n'ayant pu avoir lieu que quelques jours avant la fin de l'année scolaire, 2 élèves du groupe contrôle de 5^{ème} année de primaire étaient déjà partis en vacances sans nous en avertir. Par ailleurs, 5 élèves de 5^{ème} année de primaire (3 du groupe d'entraînement morphologique et 2 du groupe d'entraînement contrôle) présentaient un taux d'absentéisme important aux séances d'entraînement (supérieur ou égal à 40%) ; il a donc été décidé de les exclure également.

Notre échantillon final était donc composé de 51 élèves, dont 12 scolarisés en 3^{ème} année de primaire, 19 en 4^{ème} année de primaire et 20 en 5^{ème} année de primaire, constituant soit le groupe d'entraînement morphologique (n = 27) soit le groupe d'entraînement contrôle (n = 24).

6.2.1.2 Matériel d'évaluation

Nous avons mesuré les performances de nos sujets au pré-test afin de répondre à deux objectifs : premièrement, s'assurer de l'équivalence de nos deux groupes sur un ensemble de variables contrôles ; ensuite, de s'assurer qu'une éventuelle différence de performance sur l'une ou l'autre de nos variables d'intérêt en post-test ne puisse être attribuée à des différences déjà présentes entre les groupes avant l'entraînement. Pour l'ensemble de ces mesures, étant donné que certaines d'entre elles étaient expérimentales ou ne disposaient pas de normes adaptées à la classe d'âge de nos sujets, nous avons choisi de ne considérer que le score brut

et, pour les mesures d'identification des mots, de ne retenir qu'un indice de fluence pour chaque type de mot, mêlant la vitesse de réalisation et le nombre de mots correctement lus.

Variables contrôles. Les variables contrôles renvoient à des épreuves mesurant diverses capacités susceptibles d'influer sur l'efficacité de l'intervention proposée. Elles ne sont proposées qu'en pré-test. Les capacités de raisonnement non-verbal étaient évaluées via les Matrices Progressives Colorées (Raven et al., 1990), pour lesquelles le score maximal est de 36. Le vocabulaire était estimé via la forme B du test d'Échelle de Vocabulaire en Images Peabody (Dunn et al., 1993). Le score maximal est de 170²⁴. Les capacités de mémoire à court terme et de travail verbales étaient respectivement évaluées via l'empan de chiffres endroit et envers de la batterie Outil de Dépistage des Dyslexies (ODÉDYS ; Jacquier-Roux et al., 2005), pour lesquels le score maximal est de 8. La conscience phonémique était évaluée via le subtest CCV de l'épreuve de suppression phonémique du test informatisé Evalec (Sprenger-Charolles et al., 2005) : une syllabe de structure consonne-consonne-voyelle était présentée oralement à l'enfant qui devait supprimer mentalement le premier phonème et énoncer la syllabe résultante (e.g., *FLA – LA) ; le score maximal est de 12. Les capacités générales d'identification des mots étaient évaluées via l'indice CTL (Mots corrects sur Temps de lecture) du test révisé de l'Alouette (Lefavrais, 2005), estimant, sur la base d'un texte sans signification, le nombre de mots que le sujet est capable de décoder en 3 minutes. Les capacités générales de production orthographiques étaient évaluées par la dictée de l'ODÉDYS comprenant trois listes de 10 mots (10 mots réguliers consistants, 10 mots irréguliers et 10 non-mots bisyllabiques). Pour les deux premières listes, un point était attribué

²⁴ La réalisation de l'EVIP nécessite de prendre en compte trois paramètres : la base (plus haute série de 8 items consécutifs correctement réussis), le plafond (6 erreurs commises au sein d'une suite de 8 items) et le nombre d'erreurs produites entre la base et le plafond. Le score brut est ainsi obtenu en additionnant le nombre d'erreurs commises entre la base et l'item plafond, puis en soustrayant ce nombre au numéro de l'item plafond (dernier item réussi dans le plafond).

pour chaque mot parfaitement transcrit. Pour les non-mots, la production devait être une transcription phonologiquement plausible du mot dicté (e.g., pour le non-mot /pulã/, les variantes *POULEN, *POULAN, *POULENT, *POULANT étaient toutes acceptées, mais pas *POULON ni *PULAN). Les erreurs d'inversion sur des lettres isolées n'étaient pas comptabilisées si la lettre ainsi obtenue ne correspondait à aucune autre lettre existante (e.g., la lettre S écrite en miroir dans le non-mot /savet/ n'était pas considérée comme une erreur). En revanche, elles étaient considérées comme créant des productions non phonologiquement plausibles si elles touchaient une lettre dont l'inversion correspond à une autre lettre de l'alphabet (e.g., la transcription inversée de la lettre D dans le non-mot /datwaR/ était considérée comme aboutissant à une production non phonologiquement plausible, i.e., *BATOIR – /batwaR/).

Variables d'intérêt. Les variables d'intérêt renvoient à des compétences sur lesquelles un effet de l'entraînement morphologique est attendu, et sont donc évaluées à la fois en pré-test et en post-test.

L'impact de l'entraînement morphologique sur la conscience morphologique était évalué par la tâche expérimentale « Conscience Morphologique – Suffixation » (COMOS). Dans cette tâche proposée uniquement à l'oral, le sujet doit trouver le dérivé suffixé d'une base à partir d'une définition (e.g., « Comment appellerait-on une petite fille ? [Une fillette] »). Les dérivés à produire sont composés d'une base entraînée ou non-entraînée (7 entraînées et 7 non-entraînées) et d'un suffixe entraîné parmi les 7 suivants : -ER, -IER, -ERIE, -URE, -AGE, -ETTE et -EUR, de sorte que la tâche COMOS était ainsi composée de 7 mots entraînés (i.e., FROMAGER, FILLETTE, ÉPICIER, MAQUILLEUR, COIFFURE, JARDINAGE, CRÈMERIE) et de 7 mots non-entraînés (i.e., HORLOGER, FOURCHETTE, CAISSIER, BALAYEUR, BLESSURE, PATINAGE, LINGERIE) dont les suffixes ont

toutefois été abordés en entraînement. Les items entraînés et non-entraînés étaient appariés via la base de données Manulex (Lété et al., 2004; Peereman et al., 2007), sur la fréquence (occurrence pour un million de mots) ainsi que sur la longueur (nombre de lettres) de leur dérivé et de leur base ; le détail de cet appariement est présenté en Annexe 1. La fiabilité inter-item était de .50 pour les mots entraînés et de .60 pour les mots non-entraînés. Afin d'augmenter cette fiabilité, nous avons exclu l'item MAQUILLEUR ainsi que son équivalent non-entraîné BALAYEUR, résultant en une fiabilité inter-item de .51 pour les mots entraînés et de .65 pour les mots non-entraînés. La faible fiabilité inter-item de la tâche COMOS pour les mots entraînés sera discutée dans la partie 7.3 (Discussion générale) du présent chapitre.

Nous souhaitons déterminer si un entraînement morphologique pouvait faciliter le développement de l'identification des mots sur des mesures standardisées. Six listes de 20 items étaient ainsi proposées, toutes issues de la Batterie Analytique du Langage Écrit (Jacquier-Roux et al., 2010) : 20 mots réguliers fréquents, 20 mots irréguliers fréquents, 20 non-mots au pattern sublexical fréquent, 20 mots réguliers peu fréquents, 20 mots irréguliers peu fréquents et 20 non-mots au pattern sublexical peu fréquent. En nous appuyant sur la base de données Morpho-Lex (Mailhot et al., 2020)²⁵, nous avons déterminé la proportion de mots complexes dans chaque liste de mots réguliers et irréguliers. Cette proportion était de 4/20 (20%) pour les mots réguliers fréquents, de 3/20 (15%) pour les mots irréguliers fréquents, de 5/20 (25%) pour les mots réguliers peu fréquents et de 5/20 (25%) pour les irréguliers peu fréquents. Les non-mots au pattern sublexical fréquent présentaient 15 patterns orthographiques (75%) constituant un suffixe selon cette même base de données. Cette

²⁵ Notons que le mot irrégulier fréquent SECONDE, le mot régulier fréquent FUIITE et les mots réguliers peu fréquents DORADE et AVANIE n'étaient pas répertoriés dans la base de données. Le premier pouvant être compris comme le féminin de l'adjectif SECONDE, nous avons de fait considéré qu'il s'agissait d'un mot fléchi, donc complexe. Concernant les mots restants, le *Dictionnaire de l'Académie Française* (Académie française, 1932) considère qu'il s'agit de réfections latines et n'identifie qu'un morphème dans chacun de ces mots.

proportion était de 12 (60%) pour les non-mots au pattern sublexical peu fréquent. Sur l'ensemble des items de la BALE, un seul mot présentait un suffixe travaillé lors de l'entraînement (i.e., le suffixe -EUR du mot irrégulier peu fréquent COMPTEUR), et aucune base n'était entraînée. Quant aux non-mots au pattern sublexical fréquent, nous relevons 3 items en contenant (*TAUBAGE, *ABINDEUR, *ONTAGE). Les non-mots au pattern sublexical peu fréquent n'en contiennent aucun. Concernant la cotation de cette tâche, pour les mots, seule la prononciation usuelle permettait d'obtenir un point. Pour les non-mots, la prononciation devait correspondre aux règles de CGP du français (e.g., pour le non-mot *COGINTE, seule la prononciation /koʒɛ̃t/ était acceptée, et non les variantes /kogɛ̃t/, /kozint/, etc.) Pour chaque sujet, nous avons obtenu un indice de lecture en divisant le score obtenu sur 20 par le temps de lecture (exprimé en secondes), et ce pour chaque liste. Chaque indice a ensuite été multiplié par 100 pour une meilleure lisibilité. Ainsi, plus la performance donnait lieu à un score élevé et à un temps de lecture réduit, plus l'indice qui en résultait était élevé.

Enfin, la production orthographique était évaluée via la tâche expérimentale « Dictée Morphologique » (DIMO). Cette dictée comprend 16 mots dérivés, constitués à la fois d'une base (8 entraînées et 8 non-entraînées) et de l'un des 5 suffixes entraînés suivants : -IER, -ERIE, -AGE, -EUR, -ETTE. Ainsi, la DIMO est composée de 8 mots entraînés (COIFFEUR, BIJOUTIER, COLLAGE, ÉPICERIE, JARDINERIE, CRÉMIER, FILLETTE, POISSONNERIE) et de 8 mots non-entraînés (ROMANCIER, LAVAGE, DANSEUR, BOTTIER, PARFUMERIE, CACHETTE, GENDARMERIE, LAITERIE) dont les suffixes ont toutefois été abordés lors de l'entraînement. Les items des listes des mots entraînés (ME) et non-entraînés (NE) étaient appariés via la base de données Manulex (Lété et al., 2004; Peereman et al., 2007) sur leur fréquence (occurrence pour un million de mots), leur longueur

(nombre de lettres, de syllabes, de phonèmes et de graphèmes), la consistance de leur CPG et la fréquence de leur CGP (*token count*) ; le détail de cet appariement est présenté en Annexe 2. La DIMO a été cotée de plusieurs manières : dans le sous-score « mot entier », un point est attribué pour chaque mot parfaitement transcrit ; dans le sous-score « bases », un point est attribué pour chaque base correctement transcrite ; dans le sous-score « suffixes », un point est attribué pour chaque suffixe correctement transcrit²⁶. La fiabilité inter-item était acceptable pour la quasi-totalité des sous-scores de la DIMO : concernant les ME, elle était de .82 pour les mots entiers, de .71 pour les bases et de .77 pour les suffixes. Concernant les NE, elle était de .73 pour les mots entiers, de .52 pour les bases et de .82 pour les suffixes. La faible fiabilité inter-item du score « bases » des NE sera considérée dans la partie 7.3 (Discussion générale) du présent chapitre.

6.2.1.3 Matériel d'entraînement

Entraînement morphologique. Le matériel d'entraînement était à la fois oral et écrit et conçu pour être proposé en supplément de la pratique courante d'un·e enseignant·e ou d'un·e orthophoniste. Dans cette optique, il prenait la forme de 5 jeux différents ciblant le processus de suffixation via différentes activités de conscience morphologique : les familles de mot, l'assemblage, la segmentation et l'identification des morphèmes. Des exemples des jeux utilisés sont présentés en Annexe 3. Le premier jeu était inspiré du jeu des 7 familles, dans lequel chaque base (e.g., POISSON) représente le « chef » de famille et chaque mot dérivé (e.g., POISSONNIER, POISSONNERIE) constitue ses autres membres. Durant la première session, la notion de famille de mots était ainsi présentée aux élèves et nous attirions leur

²⁶ En raison d'irrégularités propres à la morphologie dérivationnelle française, pour un même mot, un point obtenu à la fois aux sous-scores « base » et « suffixe » ne signifie pas systématiquement qu'un point a été obtenu au sous-score « mot entier » : par exemple, dans la production erronée *ROMANSIER, la base ROMAN et le suffixe IER sont tous deux bien orthographiés, mais la transcription du mot entier est incorrecte, totalisant 1 point pour le sous-score base, 1 point pour le sous-score suffixe, mais 0 point pour le sous-score « mot entier » .

attention sur le fait que les mots dérivés d'une même famille partagent la même base. Dans la seconde session, nous réutilisons le même jeu mais avec une règle différente : les enfants devaient créer de nouvelles « familles » en classant les mots selon leurs affixes plutôt que selon leurs bases (e.g., ils devaient classer ensemble les mots POISSONNERIE, BIJOUTERIE, FROMAGERIE, CRÈMERIE et JARDINERIE, qui partagent tous le suffixe -ERIE exprimant la notion de lieu). Les deuxième, troisième et cinquième jeux ont été inspirés des travaux de deux étudiantes orthophonistes (Truer & Martray, 2018) et adaptés à cette étude. Le deuxième jeu, qui visait l'assemblage morphémique, était constitué de cartes rappelant les pièces d'un jeu de domino : chaque carte comportait une partie base et une partie suffixe (e.g., ETTE/JARDIN) de sorte que les cartes devaient être assemblées afin de former des mots dérivés. Lors de la troisième session, les enfants devaient utiliser les cartes pour construire de manière collective la plus grande chaîne de mots possible, tandis que dans la quatrième session, ils devaient jouer avec les cartes à la manière d'un jeu de dominos (i.e., les joueurs devaient, chacun leur tour, essayer de se débarrasser de leurs cartes en formant un mot existant avec une carte déjà posée ; ils devaient piocher une nouvelle carte s'ils ne pouvaient pas jouer). Nous attirions l'attention des enfants sur les changements orthographiques et phonologiques de la base causés par la suffixation (e.g., l'assemblage de la base FROMAGE et du suffixe IER nécessite de supprimer le I pour créer le mot FROMAGER). Le troisième jeu était présenté dans les sessions cinq et six et visait la segmentation morphémique. Dans la cinquième session, des mots suffixés aux lettres évidées étaient présentés sur une feuille de papier et les enfants devaient colorier les suffixes communs d'une même couleur. À nouveau, nous attirions l'attention des enfants sur les changements orthographiques et phonologiques causés par la suffixation (FROMAGER et ÉPICIER partagent un suffixe pouvant être considéré comme identique du point de vue du sens, en dépit de leurs différences de

transcription et de prononciation). Dans la sixième session, les mêmes mots aux suffixes colorés étaient présentés sur des cartes individuelles, et les enfants s'en servaient pour jouer au Mistigri : deux mots partageant le même suffixe étaient considérés comme une paire devant être retirée du jeu aussitôt qu'elle apparaissait dans la main d'un joueur, jusqu'à ce qu'un seul mot suffixé puisse être isolé, constituant le « mistigri » et entraînant la défaite du dernier joueur à l'avoir dans sa main. Le quatrième jeu était inspiré du jeu du Dobble® (Zygomatic - Board Game Studio, 2019) : une quarantaine de cartes étaient distribuées entre les différents joueurs, chacune comprenant 5 mots suffixés. Les joueurs devaient rapidement identifier un morphème commun entre les mots dérivés de leur propre carte et ceux de la carte posée au milieu de la table : lorsqu'un morphème (base ou suffixe) commun était identifié, le joueur devait l'énoncer à haute voix et poser sa carte sur le tas du milieu, forçant l'ensemble des joueurs à trouver un nouveau morphème commun. Cette façon de jouer correspondait à la septième session. Dans la huitième session, le jeu était utilisé à la manière d'un bingo : les élèves piochaient au hasard l'une des 40 cartes, puis l'adulte prodiguant les entraînements faisait de même. Cette dernière énonçait ensuite l'un après l'autre les mots sur sa carte : si les élèves identifiaient, sur leur propre carte, un morphème commun (base ou suffixe) avec les mots prononcés, ils devaient recouvrir le mot concerné d'un jeton, jusqu'à ce que l'un des joueurs obtienne un « bingo » (i.e., tous ses mots étaient recouverts d'un jeton). Enfin, le cinquième jeu était un labyrinthe constitué de mots suffixés. Dans la neuvième session, les élèves devaient trouver individuellement le chemin permettant à leur pion de sortir du labyrinthe, en évoluant case par case d'un morphème commun à un autre. Dans la dixième et ultime session, le même labyrinthe était présenté sous la forme d'un jeu de *Serpents et Échelles* : chaque joueur faisait progresser son pion sur le plateau selon le résultat d'un jet de dé. Ils pouvaient encore avancer si la case adjacente présentait un morphème commun avec la

case sur laquelle ils avaient atterri. À la fin de chacune des dix sessions, les enfants effectuaient une dictée collective de dix des mots suffixés utilisés au cours du jeu. Cette dictée avait pour but de favoriser les liens entre les morphèmes identifiés via la manipulation morphologique et leurs équivalents orthographiques, partant du constat qu'un entraînement proposant des exercices métalinguistiques est d'autant plus efficace lorsqu'il est associé à des activités de production orthographique (Nunes et al., 2003). Chaque mot suffixé entraîné était dicté à quatre reprises sur l'ensemble des 10 sessions (voir Annexe 4 pour la composition de ces dictées). Les erreurs commises par les enfants étaient immédiatement corrigées afin qu'ils ne mémorisent pas des orthographes incorrectes.

Entraînement contrôle. Le matériel d'entraînement contrôle (voir Annexe 5 pour des exemples) a été élaboré sur la base du matériel d'entraînement morphologique afin que les facteurs motivationnels et que les processus cognitifs généraux mis en œuvre soient les plus similaires possibles entre les deux groupes, tout en ne sollicitant pas d'activité métalinguistique. Nous avons pour cela choisi de créer un matériel d'entraînement mathématique, de sorte que les familles de mots suffixés de l'entraînement morphologique étaient remplacées par des additions simples, souvent sollicitées lors du calcul mental, i.e. les complétions à 10 (e.g., 4+6), à 11 (e.g., 5+6), à 12 (e.g., 5+7), à 13 (e.g., 6+7), à 14 (e.g., 5+9), à 15 (e.g., 8+7) et à 16 (e.g., 7+9). Étant donné que nous travaillions le lien entre morphème oral et écrit dans l'entraînement morphologique via des dictées avec feedback correct immédiat, et que ces dernières constituent un entraînement efficace en lui-même (e.g., Graham, 1999), nous avons choisi de proposer les mêmes dictées dans l'entraînement contrôle, afin de nous assurer qu'un potentiel effet bénéfique de l'entraînement morphologique était bien dû aux activités de conscience morphologique et non à ce seul exercice de production orthographique. Les modalités et contenus des dictées étaient

identiques à ceux proposés dans l'entraînement morphologique, mais aucune analyse morphologique n'était fournie aux élèves, et les adultes veillaient à ne pas attirer leur attention sur les morphèmes communs entre les différents items.

6.2.1.4 Procédure

Les participants étaient évalués et entraînés au sein de leur école durant le temps scolaire, au moment le plus commode pour les enseignants, partant du principe que, dans une perspective écologique où cet entraînement morphologique serait intégré à leur pratique courante, il serait, de fait, proposé aux élèves sur ces créneaux plutôt qu'à tout autre moment. À cet effet, une salle calme et bien éclairée était mise à disposition par l'école. Les évaluations pré-tests et post-tests étaient réalisées par la doctorante et par des étudiant·e·s en orthophonie et en psychologie du développement. Les évaluations pré-test avaient lieu au courant du mois de février et étaient réalisées de manière individuelle, dans des sessions durant entre 45 minutes et une heure en fonction de l'âge et des facilités de chaque participant. Seules les Matrices Progressives Colorées étaient administrées collectivement, dans la salle de classe habituelle des enfants, par le moyen d'un cahier individuel. Les évaluations post-test étaient réalisées juste après la fin de l'entraînement, entre fin juin et début juillet. Elles étaient entièrement individuelles et duraient entre 30 et 45 minutes par élève. Les entraînements commençaient une semaine après la fin des sessions pré-tests et consistaient en 10 sessions à raison de deux sessions par semaine, chacune durant approximativement 45 minutes (30 minutes d'entraînement et 15 minutes de dictée), résultant environ en 7.5 heures d'entraînement au total. L'entraînement était prodigué en groupes de 5 à 6 élèves. L'adulte délivrant les entraînements était soit la doctorante, soit une psychologue du développement, soit une étudiante en dernière année d'études d'orthophonie, toutes formées à l'entraînement morphologique. Chaque adulte délivrait entièrement les dix sessions

d'entraînement pour un même groupe, et afin d'exclure des biais liés aux différences de style d'enseignement, chaque adulte entraînait à la fois un groupe d'intervention morphologique et un groupe d'intervention contrôle au sein d'une même classe. De plus, il leur était demandé de baser leur enseignement sur des scénarii prérédigés, et ce pour chaque session d'entraînement ; un exemple de ces scénarii est disponible en Annexe 6. Les adultes qui dispensaient des entraînements ne participaient pas à l'évaluation post-test des classes des élèves qu'elles avaient entraînés. Ainsi, tous les examinateurs post-tests étaient aveugles quant au groupe d'intervention du participant qu'ils ou elles évaluaient.

6.2.1.5 Analyses statistiques

Des analyses de variance (ANOVA) à mesures répétées ont été utilisées pour déterminer la significativité des interactions entre le point de mesure (pré-test et post-test) et le groupe d'entraînement (morphologique et contrôle). Nous avons eu recours à un modèle de somme des carrés de type III (Langsrud, 2003) et avons réalisé des comparaisons planifiées (analyses de contraste) entre les moyennes pré-test et post-test de chaque groupe, ainsi qu'entre les moyennes post-tests des deux groupes d'entraînement. Étant donné que ces contrastes correspondaient aux hypothèses que nous souhaitions tester, nous avons décidé de toujours regarder les comparaisons à titre indicatif, y compris lorsque l'interaction n'était pas significative (Jin, 2021) ; toutefois, en raison du risque accru d'erreur de type 2, nous avons tâché de considérer les analyses effectuées avec prudence lorsque ce cas se présentait.

Une p-valeur inférieure ou égale à .05 était considérée comme statistiquement significative, et celles comprises entre .06 et .10 étaient considérées comme tendanciellement significatives. Les tailles d'effet de l'entraînement sont indiquées pour les ANOVA via l'éta carré partiel (η^2_p) dont les valeurs peuvent s'interpréter comme suit (Richardson, 2011) : .01 :

effet faible ; .06 : effet médium ; .14 : effet fort. Les ANOVA, tout comme les analyses de contraste, ont été réalisées avec le logiciel JASP (JASP Team, 2023).

6.2.2 Résultats

Au moment du pré-test, aucune différence significative entre nos deux groupes d'entraînement n'était constatée sur les mesures de mémoire à court terme verbale, de vocabulaire passif, de raisonnement non-verbal, de conscience phonémique et morphologique, ni sur les mesures de lecture et de production orthographique (voir Tableau 1).

Tableau 1. Caractéristiques des groupes EM et EC au pré-test

		EM (n = 27)			EC (n = 24)		
Nombre de participants en fonction de l'année d'instruction		3 ^{ème} (n = 6)	4 ^{ème} (n = 9)	5 ^{ème} (n = 12)	3 ^{ème} (n = 6)	4 ^{ème} (n = 10)	5 ^{ème} (n = 8)
Âge (années)		8.91 (0.26)	9.78 (0.53)	10.9 (0.39)	9.02 (0.61)	9.58 (0.38)	10.8 (0.32)
Habiletés cognitives générales							
MPC (/36)		28.5 (6.09)	29.8 (4.29)	32.5 (2.75)	29.2 (5.89)	31.1 (3.52)	31.9 (2.42)
EVIP (/170)		103 (30.7)	115 (23.7)	127 (17.5)	110 (27.2)	117 (17.0)	123 (21.4)
Empan endroit (/8)		5.00 (0.89)	5.00 (0.71)	5.17 (0.72)	5.33 (1.63)	4.90 (0.88)	5.88 (1.25)
Empan envers (/8)		3.17 (0.41)	3.44 (0.53)	4.00 (0.74)	3.50 (0.55)	3.30 (0.82)	3.88 (0.64)
Conscience phonologique							
Evalec	Items corrects (/12)	7.83 (3.31)	7.89 (2.93)	9.75 (2.05)	9.17 (2.04)	9.40 (1.26)	8.25 (1.04)
	Temps de réponse items corrects (ms)	3856 (1210)	4929 (3155)	3895 (2514)	4715 (3386)	3639 (1496)	2987 (906)
Identification des mots							
Alouette-R	CTL (mots lus en 3 mn)	222 (57.8)	200 (102)	313 (145)	257 (126)	232 (72.6)	346 (113)
Production orthographique							
ODÉDYS	REG S (/10)	8.50 (1.87)	7.44 (1.94)	8.83 (1.53)	8.00 (2.28)	9.20 (0.92)	8.38 (1.92)
	IRR (/10)	5.17 (2.32)	4.46 (2.46)	7.33 (2.19)	6.83 (2.32)	5.40 (2.07)	6.38 (3.16)
	NM B (/10)	8.17 (1.17)	7.67 (2.18)	9.00 (1.28)	7.50 (2.35)	8.20 (2.30)	9.00 (1.31)

CTL = Indice de fluence de lecture. EC = Entraînement Contrôle. EM = Entraînement Morphologique. EVIP = Échelle de Vocabulaire en Images Peabody. IRR = Mots irréguliers. MPC = Matrices Progressives Colorées. NM B = Non-mots bisyllabiques. ODÉDYS = Outil de Dépistage des Dyslexies. REG S = Mots réguliers simples.

Les résultats obtenus sur les différentes variables d'intérêt en pré- et post-test sont présentés dans le Tableau 2.

Tableau 2. Moyennes et écarts-types des groupes EM et EC en pré- et post-test sur l'ensemble des variables dépendantes

		Pré-test		Post-test		
		EM	EC	EM	EC	
COMOS	Mots entraînés (/6)	3.58 (1.5)	3.79 (1.5)	5.19 (0.83)	4.21 (1.35)	
	Mots non-entraînés (/6)	3.81 (1.72)	3.67 (1.66)	4.37 (1.15)	3.92 (1.38)	
DIMO	Mots entraînés (/8)	Mot entier	3.70 (2.28)	4.25 (2.59)	6.63 (1.39)	6.79 (1.86)
		Bases	5.04 (2.10)	5.46 (2.15)	7.26 (0.81)	7.17 (1.27)
		Suffixes	5.63 (2.19)	6.13 (2.09)	7.44 (0.97)	7.42 (1.25)
	Mots non-entraînés (/8)	Mot entier	3.93 (2.32)	4.54 (2.26)	4.96 (2.03)	5.46 (1.82)
		Bases	5.19 (1.57)	5.83 (1.40)	5.56 (1.69)	6.08 (1.64)
		Suffixes	5.78 (2.19)	5.67 (2.55)	7.04 (1.16)	7.00 (1.59)
BALE	Mots irréguliers fréquents	88.9 (50.2)	93.7 (53.8)	100.4 (52.9)	105.7 (51.3)	
	Mots réguliers fréquents	104.9 (49.9)	118.7 (47.9)	107.9 (44.1)	122.7 (44.8)	
	Non-mots au pattern sublexical fréquent	60.4 (27)	69.5 (31.6)	65.1 (29.1)	72.7 (30.2)	
	Mots irréguliers peu fréquents	54.6 (45.9)	64.7 (48.1)	61.2 (47.3)	63.1 (43.2)	
	Mots réguliers peu fréquents	74.1 (43.6)	90.1 (49.4)	78.3 (45.4)	89.6 (43.5)	
	Non-mots au pattern sublexical peu fréquent	54.7 (34.8)	54.9 (20.8)	57.0 (31.1)	57.2 (21.4)	

BALE = Batterie Analytique du Langage Écrit. COMOS = Conscience Morphologique – Suffixation. DIMO = Dictée Morphologique. EC = Entraînement Contrôle. EM = Entraînement Morphologique.

6.2.2.1 EM versus EC : Conscience Morphologique – Suffixation (COMOS)

L'impact de l'intervention morphologique sur les capacités de conscience morphologique est évalué via la tâche COMOS. Le groupe entraîné à la morphologie et le groupe contrôle semblent avoir augmenté leurs performances entre le pré-test et le post-test, à la fois sur les mots entraînés ($F(1,48) = 28.97, p < .001, \eta^2_p = .376$)²⁷ et les mots non-entraînés ($F(1,48) = 5.52, p = .02, \eta^2_p = .106$).

²⁷ Le sujet manquant correspond à un sujet absent lors du pré-test dans le groupe EM.

Concernant les items entraînés, l'interaction entre le point de mesure et le groupe d'entraînement était significative ($F(1,48) = 9.81, p < .01, \eta^2_p = .170$). Les analyses de contrastes ont révélé que le groupe entraîné à la morphologie a progressé entre le pré-test et le post-test ($t(48) = -6.15, p < .001$), contrairement au groupe contrôle ($t(48) = -1.56, p = .13$). Les résultats obtenus au post-test montrent une supériorité significative du groupe entraîné à la morphologie par rapport au groupe contrôle ($t(48) = -2.53, p = .01$).

Concernant les items non-entraînés, l'interaction entre le point de mesure et le groupe d'entraînement n'était pas significative ($F(1,48) < 1, p = .44, \eta^2_p = .013$). Nous notons une différence significative pour le groupe entraîné à la morphologie entre le pré-test et le post-test ($t(48) = -2.26, p = .03$), contrairement au groupe contrôle ($t(48) = -1.09, p = .28$). En revanche, les performances entre le groupe expérimental et le groupe contrôle ne sont pas significativement différentes en post-test ($t(48) = -0.93, p = .36$).

6.2.2.2 EM versus EC : identification des mots

L'impact de l'entraînement morphologique sur le développement de l'identification des mots est évalué via les listes de lecture de la BALE.

Sur les mots irréguliers fréquents, nous notons un effet significatif du point de mesure ($F(1,48) = 13.5, p < .001, \eta^2_p = .220$)²⁸ mais l'interaction entre le point de mesure et le groupe d'entraînement n'est pas significative pour ces items ($F(1,48) < 1, p = .99, \eta^2_p = .000$). Les analyses de contraste montrent en effet une différence significative entre le pré-test et le post-test, à la fois pour le groupe expérimental ($t(48) = -2.64, p = .01$) et pour le groupe contrôle ($t(48) = -2.56, p = .01$), mais ces deux groupes ne se distinguent pas significativement en post-test ($t(48) = 0.36, p = .72$).

²⁸ Les sujets manquants correspondent à des oublis de chronométrage lors de la passation des tâches en pré-test ou en post-test, n'ayant pas permis de calculer un indice de fluence pour 3 sujets du groupe entraînés à la morphologie (1 en pré-test et 2 en post-test) et pour 1 sujet du groupe contrôle (1 en post-test).

Concernant les mots réguliers fréquents, nous ne notons pas d'effet du point de mesure ($F(1,48) = 1.28, p = .26, \eta^2_p = .026$) ni d'interaction significative entre le point de mesure et le groupe d'entraînement ($F(1,48) < 1, p = .97, \eta^2_p = .000$). Ainsi, nous ne retrouvons pas de différence significative entre le pré-test et le post-test pour les groupes expérimental ($t(48) = -0.84, p = .40$) et contrôle ($t(48) = -0.76, p = .45$). La différence entre ces deux groupes au post-test n'est pas significative ($t(48) = 1.11, p = .27$).

Sur les non-mots au pattern sublexical fréquent, un effet tendanciel du temps est noté ($F(1,49) = 3.05, p = .09, \eta^2_p = .060$) mais nous n'avons pas retrouvé d'interaction significative entre le point de mesure et le groupe d'entraînement pour cette variable ($F(1,49) < 1, p = .74, \eta^2_p = .002$). Nous ne notons pas de différence significative entre le pré-test et le post-test pour les groupes expérimental ($t(49) = -1.52, p = .14$) et contrôle ($t(49) = -0.97, p = .34$), ni entre ces deux groupes au post-test ($t(49) = 0.93, p = .36$).

Il n'y avait pas d'effet du point de mesure pour les mots irréguliers peu fréquents ($F(1,48) = 1.01, p = .32, \eta^2_p = .021$). En revanche, une interaction significative entre le point de mesure et le groupe d'entraînement a été retrouvée ($F(1,48) = 4.19, p = .046, \eta^2_p = .080$). Les analyses de contraste ont révélé une différence significative entre le pré-test et le post-test pour le groupe entraîné à la morphologie ($t(48) = -2.25, p = .03$), tandis le groupe contrôle montre une baisse non-significative des performances entre le pré-test et le post-test ($t(48) = 0.77, p = .48$; voir Figure 8). En revanche, nous n'observons pas de différence significative au post-test entre le groupe entraîné à la morphologie et le groupe contrôle ($t(48) = 1.14, p = .89$).

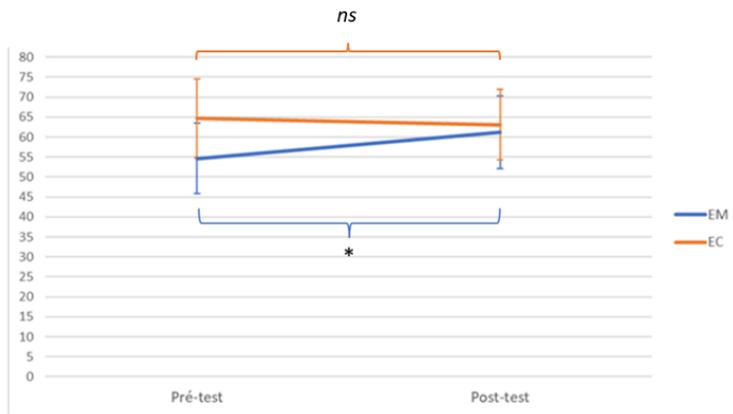


Figure 8. Évolution des performances des groupes EM et EC sur la lecture de mots irréguliers peu fréquents (BALE)

EC = Groupe contrôle. EM = Groupe entraîné à la morphologie. *ns* = Non significatif.

* = $p < .05$.

NB : Les barres d'erreurs représentent l'erreur standard.

Il n'y avait pas d'effet du point de mesure pour les mots réguliers peu fréquents ($F(1,48) < 1, p = .35, \eta^2_p = .018$), ni d'interaction significative entre le point de mesure et le groupe d'entraînement ($F(1,48) = 1.31, p = .26, \eta^2_p = .027$). Ainsi, nous ne notons pas de différence significative entre le pré-test et le post-test pour les groupes expérimental ($t(48) = -1.51, p = .14$) et contrôle ($t(48) = 0.14, p = .89$), ni entre ces deux groupes au post-test ($t(48) = 0.79, p = .44$).

Un pattern de résultats identique est observé pour les non-mots au pattern sublexical peu fréquent. Nous ne notons pas d'effet du temps ($F(1,49) = 1.37, p = .25, \eta^2_p = .027$), ni d'interaction entre le point de mesure et le groupe d'entraînement ($F(1,49) < 1, p = .98, \eta^2_p = .000$). Ainsi, nous ne notons pas de différence significative entre le pré-test et le post-test pour les groupes expérimental ($t(49) = -0.87, p = .39$) et contrôle ($t(49) = -0.78, p = .44$), ni entre ces deux groupes au post-test ($t(49) = 0.01, p = .99$).

6.2.2.3 EM versus EC : Dictée Morphologique (DIMO)

L'impact de l'entraînement morphologique sur la production orthographique est évalué via la DIMO. Ainsi que le détaille le Tableau 3, nos deux groupes d'entraînement semblent avoir progressé sur l'ensemble des scores entre le pré-test et le post-test ; toutefois, aucune interaction significative n'a été trouvée entre le point de mesure et le groupe d'entraînement, pour aucune de nos mesures. Ainsi, les analyses de contraste mettent en évidence des différences significatives entre le pré-test et le post-test pour l'ensemble des mesures, à la fois pour le groupe EM et le groupe EC, sauf pour le score « bases » des mots non-entraînés. Pour ce score en particulier, le groupe EM montre une amélioration tendancielle entre le pré-test et le post-test ($t(49) = - 1.76, p = .09$), contrairement au groupe EC ($t(49) = - 3.50, p = .27$).

Nous ne notons aucune différence significative entre le groupe EM et le groupe EC sur l'ensemble des mesures post-test de la DIMO (cf. Tableau 3).

Tableau 3. Effets simples du temps, interactions point de mesure * groupe d'entraînement et contrastes sur l'ensemble des mesures de la DIMO

		ANOVA à mesures répétées		
		Effet du point de mesure	Interaction point de mesure * groupe d'entraînement	
Mots entraînés	Mot entier	$F(1,49) = 133.12, p < .001, \eta^2_p = .731$	$F(1,49) < 1, p = .42, \eta^2_p = .013$	
	Base	$F(1,49) = 64.28, p < .001, \eta^2_p = .567$	$F(1,49) = 1.10, p = .30, \eta^2_p = .022$	
	Suffixe	$F(1,49) = 40.31, p < .001, \eta^2_p = .451$	$F(1,49) = 1.14, p = .29, \eta^2_p = .023$	
Mots non-entraînés	Mot entier	$F(1,49) = 29.41, p < .001, \eta^2_p = .375$	$F(1,49) < 1, p = .74, \eta^2_p = .002$	
	Base	$F(1,49) = 4.08, p = .049, \eta^2_p = .077$	$F(1,49) < 1, p = .70, \eta^2_p = .003$	
	Suffixe	$F(1,49) = 28.96, p < .001, \eta^2_p = .371$	$F(1,49) < 1, p = .88, \eta^2_p = .000$	
		Contrastes		
		EM pré-test versus EM post-test	EC pré-test versus EC post-test	EC post-test versus EM post-test
Mots entraînés	Mot entier	$t(49) = -9.00, p < .001$	$t(49) = -7.37, p < .001$	$t(49) = 0.28, p = .78$
	Base	$t(49) = -6.61, p < .001$	$t(49) = -4.79, p < .001$	$t(49) = 0.20, p = .84$
	Suffixe	$t(49) = -5.41, p < .001$	$t(49) = -3.63, p < .001$	$t(49) = -0.06, p = .95$
Mots non-entraînés	Mot entier	$t(49) = -4.20, p < .001$	$t(49) = -3.50, p = .001$	$t(49) = 0.83, p = .41$
	Base	$t(49) = -1.76, p = .09$	$t(49) = -3.50, p = .27$	$t(49) = 1.18, p = .24$
	Suffixe	$t(49) = -3.81, p < .001$	$t(49) = -3.80, p < .001$	$t(49) = -0.07, p = .95$

6.2.3 Discussion

Notre étude visait à déterminer dans quelle mesure un entraînement morphologique (EM) de 45 minutes proposé deux fois par semaine sur le temps scolaire pouvait favoriser le développement de l'identification de mots écrits et de la production orthographique.

L'objectif était d'une part d'évaluer la capacité d'un tel entraînement à traduire son efficacité via des mesures standardisées de lecture, et d'autre part de déterminer si la manipulation morphémique explicite présentait un avantage pour la production orthographique, par rapport à entraînement contrôle (EC) de production orthographique des morphèmes sans activité réflexive sur ces derniers.

En premier lieu, nous notons une interaction significative entre le groupe d'entraînement et le point de mesure sur les mots entraînés de notre tâche expérimentale de conscience morphologique (COMOS), en faveur du groupe EM. La manipulation explicite de morphèmes semble avoir aidé les participants de ce groupe à approfondir leur compréhension des mots travaillés par rapport à la seule activité de transcription, puisqu'en comparaison, les participants du groupe EC ont obtenu des performances similaires entre le pré-test et le post-test. De plus, au post-test, le groupe EM obtient des performances significativement supérieures à celles du groupe EC sur ces mots entraînés. En ce qui concerne les mots non-entraînés, le groupe EM a également significativement progressé entre les deux points de mesure, ce qui n'est pas le cas du groupe EC. Toutefois, les performances obtenues par ces deux groupes ne sont pas significativement différentes au post-test. Ainsi, les participants du groupe EM semblent avoir, dans une certaine mesure, transféré leurs connaissances sur les morphèmes travaillés en entraînement à des mots suffixés non-entraînés. Dans l'ensemble, les résultats obtenus sur la tâche COMOS suggèrent ainsi que la simple transcription orthographique des morphèmes seule ne suffit pas à accéder à une analyse morphémique. Celle-ci semble devoir faire l'objet d'un travail explicite et spécifique.

En ce qui concerne l'identification des mots, nous observons une amélioration de nos deux groupes sur les mots irréguliers *fréquents* entre les deux points de mesure, qui n'est donc pas distinguable d'un simple effet développemental ou placebo / Hawthorne. Sur les mots irréguliers *peu fréquents*, le groupe EM semble avoir progressé de manière significative entre le pré-test et le post-test, contrairement au groupe EC. Toutefois, les performances aux post-tests entre ces deux groupes n'étaient pas significatives. Rappelons que cette dernière liste présentait un pourcentage de mots complexes de 25% (contre 15% pour les mots irréguliers *fréquents*), et qu'elle présentait un suffixe travaillé en séance (COMPTEUR), tandis que la liste de mots irréguliers *fréquents* n'en comptait aucun. Ceci semble suggérer que des gains en fluence de lecture suite à un EM peuvent être observés dès lors que la proportion de mots complexes atteint les 25%, y compris lorsque la majorité d'entre eux ne comprennent pas de morphèmes ayant fait l'objet d'un entraînement. Remarquons toutefois qu'aucun progrès, ni de la part du groupe EM, ni de la part du groupe EC, n'est noté sur les autres types de mots, bien qu'ils comprennent des proportions de mots complexes similaires : e.g., les mots réguliers peu fréquents comportaient également 25% de mots complexes, et les non-mots au pattern sublexical fréquent ainsi que les non-mots au pattern sublexical peu fréquent présentaient des proportions d'affixes qui étaient respectivement de 75% (dont 3 vus en entraînement) et de 60%. Le bénéfice d'un EM pour la fluence de lecture semble donc, dans le cadre de notre étude, se restreindre aux 1) *mots*, 2) *irréguliers*, 3) *peu fréquents* et 4) présentant une proportion d'au moins 25% de mots complexes. Ceci suggère que l'analyse morphologique, par rapport à la simple activité de transcription, permettrait aux participants de recourir à une stratégie de lecture – potentiellement, le décodage morphologique (Levesque et al., 2021) – leur permettant de lire de manière plus fluente lorsqu'ils sont confrontés à des mots présentant les quatre caractéristiques susnommées.

Rappelons toutefois qu'il s'agit là d'une hypothèse avancée sur la base d'une seule mesure, pour laquelle l'effet d'interaction était très proche du seuil de non-significativité ($p = .046$). Outre cela, la faible taille d'échantillon de nos deux groupes d'entraînement contribue également à nous exposer à un risque d'artefact statistique. Cette hypothèse doit donc être considérée avec la plus grande prudence, d'autant qu'à notre connaissance, il s'agit de la première étude ayant mis en évidence un bénéfice d'un EM pour les mots irréguliers peu fréquents ne comportant que 25% de mots complexes.

Concernant la production orthographique (DIMO), nos deux groupes d'entraînement semblent avoir progressé de manière similaire sur l'ensemble des scores, excepté sur le score « bases » des mots non-entraînés, seul score ne comportant aucun morphème travaillé en entraînement. Pour ce score particulier, seul le groupe EM montre une amélioration tendancielle entre le pré-test et le post-test, bien que ses performances ne se différencient pas significativement du groupe EC en post-test. Ceci pourrait suggérer que l'analyse morphologique a permis aux participants du groupe EM de réaliser des liens entre les mots à écrire et des mots de même famille morphologique dont ils connaissaient l'orthographe, bien qu'ils n'aient pas été vus en entraînement, ce qui pourrait suggérer une forme de transfert.

Excepté pour cette épreuve, nos deux entraînements (avec et sans analyse morphologique) semblent avoir permis à nos participants d'améliorer leurs performances à la DIMO. Les deux entraînements semblent toutefois avoir été d'une efficacité comparable, l'analyse morphologique n'ayant pas constitué d'avantage par rapport à la seule transcription des morphèmes entraînés. Il convient de rappeler qu'un entraînement à la transcription sans erreur constitue une intervention efficace en elle-même pour améliorer la production orthographique de ces mots (e.g., Joseph et al., 2012). Ainsi, il est possible que l'entraînement à la production orthographique de morphèmes prodigué à notre groupe contrôle aient donné lieu à des effets suffisamment importants pour masquer l'effet de l'entraînement

morphologique. Une autre interprétation demeure également possible : l'effet bénéfique constaté pourrait en réalité ne refléter qu'un effet placebo ou Hawthorne. Puisque nous n'avions pas de groupe non-entraîné, il ne nous est pas possible d'exclure cette hypothèse.

Dans une seconde étude, nous avons donc souhaité recruter un nouveau groupe contrôle, cette fois uniquement entraîné aux mathématiques, et qui ne comporterait donc pas d'élément susceptible d'impacter le développement de la littératie. L'objectif de l'étude 2 était ainsi de déterminer si des élèves de même niveau de classe que ceux de notre étude 1, évalués sur une période de temps similaire, et à qui l'on aurait prodigué de l'attention à la même intensité, montreraient le même profil d'évolution que nos groupes d'entraînement de l'étude 1. Plus précisément, nous souhaitions déterminer si entraîner à la production orthographique avec et sans analyse morphologique donnait bien lieu à une amélioration similaire pour l'identification des mots irréguliers fréquents et pour la transcription de morphèmes, ou si ces améliorations étaient en réalité attribuables à un simple effet placebo / Hawthorne.

6.3 Étude 2

6.3.1 Méthode

6.3.1.1 Participants

Nous avons donc recruté des participants supplémentaires pour constituer notre deuxième groupe contrôle, que nous appelons EC-M (entraînement contrôle – maths seules). Afin de limiter les différences inter-groupes avec les groupes EM (entraînement morphologique) et EC-PO (entraînement contrôle – production orthographique), nous avons tenté de recontacter les écoles qui avaient donné leur accord pour l'étude 1. Malheureusement, ce protocole constituant un à-côté relativement chronophage pour les enseignants dans un contexte de nombreux retards scolaires exacerbés par la pandémie du COVID-19, elles n'ont pas souhaité y participer deux années consécutives. Nous avons donc sollicité une autre école, dans la ville de Schiltigheim (en périphérie de Strasbourg). Après avoir eu connaissance des

modalités de l'entraînement proposé et des évaluations qui en découlaient, nous avons obtenu l'accord de cette école pour y réaliser notre recrutement. Étant donné que dans l'étude 2, le groupe EC-M consistait en des activités uniquement mathématiques (voir § 6.3.1.4), la lettre d'information a été modifiée en conséquence et les enseignants étaient informés du contenu des entraînements. Neuf élèves de 5^{ème} année et 13 élèves de 4^{ème} année de primaire ont ainsi retourné leurs formulaires de consentement éclairé et ont passé les évaluations pré-test, réalisées en mars 2022.

Parmi les élèves évalués, deux ne parlaient pas couramment français et ont donc été exclus (1 en 5^{ème} année et 1 en 4^{ème} année de primaire), puisque nous avons considéré qu'ils risquaient de constituer des profils atypiques lors de l'analyse des résultats. Étant donné que nous avons essayé de reproduire autant que possible la période dans laquelle s'est déroulé l'entraînement de l'étude 1, nous avons été confrontés au même problème de départs en vacances anticipés : parmi les élèves de 4^{ème} année de primaire, 3 élèves n'ont pas passé les post-tests pour cette raison. Au final, notre groupe EC-M était donc constitué de 17 élèves, dont 8 en 5^{ème} année et 9 en 4^{ème} année de primaire.

6.3.1.2 Matériel d'évaluation

Les tests constituant les variables contrôles (pré-test uniquement) et les variables d'intérêt (pré- et post-test) étaient identiques à ceux décrits dans l'étude 1 (cf. § 6.2.1.3).

6.3.1.3 Matériel d'entraînement

L'intervention proposée au groupe EC-M reprenait le matériel utilisé avec le groupe EC-PO (cf. § 6.2.1.3), à la différence qu'aucune dictée n'était proposée à l'issue des sessions. Le groupe EC-M ne bénéficiait donc que d'un entraînement au calcul mental et n'était pas exposé aux mots considérés comme entraînés dans les conditions EM et EC-PO.

6.3.1.4 Procédure

Le protocole d'évaluation pré- et post-test était identique à celui décrit pour l'étude 1 (cf. § 6.2.1.4). Les élèves étaient évalués en pré-test au courant du mois de mars 2022 par la doctorante et des étudiantes en psychologie du développement, et en post-test au courant du mois de juin 2022, juste après la fin de la période d'entraînement. Les entraînements ont été menés par la doctorante, qui n'a donc pas pris part aux évaluations post-test. Ces dernières ont été réalisées par des étudiantes en orthophonie formées à la passation de tests. Étant donné que l'intervention EC-M ne comprenait pas de moment dédié à la production orthographique de mots suffixés, les sessions duraient en moyenne moins longtemps que celles des groupes EM et EC-PO (approximativement 30 minutes).

6.3.1.5 Analyses statistiques

Les analyses réalisées étaient identiques à celles menées dans l'étude 1 (cf. § 6.2.1.5).

6.3.2 Résultats

Après analyse des données pré-test, il est apparu que les participants du groupe EC-M se différencient significativement des groupes EM et EC-PO sur le raisonnement non-verbal et les scores à la tâche de COMOS. Nous avons donc décidé de sélectionner, parmi les groupes EM et EC-PO de l'étude 1, les participants dont les profils de performance au pré-test étaient les plus similaires à ceux des participants de l'étude 2 sur un maximum de variables. Nous avons procédé à un appariement trois à trois, de sorte que chaque groupe d'entraînement soit constitué de 17 sujets (voir Tableau 4). Après appariement, il n'y avait plus de différence significative entre nos groupes sur aucune des mesures en pré-test.

Tableau 4. Caractéristiques des groupes EM, EC-PO et EC-M au pré-test

		EM (n = 17)			EC-PO (n = 17)			EC-M (n = 17)			
		3 ^{ème} (n = 4)	4 ^{ème} (n = 7)	5 ^{ème} (n = 6)	3 ^{ème} (n = 4)	4 ^{ème} (n = 7)	5 ^{ème} (n = 6)	3 ^{ème} (n = 0)	4 ^{ème} (n = 9)	5 ^{ème} (n = 8)	
Nombre de participants en fonction de l'année d'instruction											
Âge (années)		9.04 (0.17)	9.75 (0.60)	10.9 (0.45)	9.10 (0.75)	9.66 (0.34)	10.8 (0.36)	-	9.78 (0.29)	10.8 (0.32)	
Habilités cognitives générales											
MPC (/36)		28.0 (7.53)	28.9 (4.45)	32.7 (1.63)	27.8 (5.68)	30.0 (3.90)	31.3 (2.42)	-	29.6 (3.81)	24.3 (4.95)	
EVIP (/170)		97.5 (33.7)	113 (26.1)	122 (18.2)	109 (31.1)	113 (16.2)	118 (22.9)	-	112 (29.1)	116 (18.8)	
Empan endroit (/8)		4.75 (0.96)	5.00 (0.82)	5.00 (0.89)	5.25 (2.06)	5.00 (0.82)	5.50 (1.05)	-	5.22 (0.83)	5.13 (0.99)	
Empan envers (/8)		3.25 (0.50)	3.43 (0.54)	3.83 (0.75)	3.25 (0.50)	3.29 (0.49)	3.83 (0.41)	-	3.33 (1.22)	3.63 (0.74)	
Conscience phonologique											
Evalec		Items corrects (/12)	8.00 (4.08)	7.71 (2.69)	8.67 (2.42)	8.75 (2.36)	9.00 (1.15)	8.00 (1.10)	-	8.78 (2.73)	9.00 (2.07)
		Temps de réaction items corrects (ms)	3794 (1353)	5382 (3488)	4112 (1717)	5666 (3912)	3627 (1817)	3055 (1037)	-	3529 (1483)	3207 (1198)
Identification des mots											
Alouette-R		CTL (mots lus en 3 mn)	209 (39.0)	217 (109)	283 (168)	200 (114)	206 (73.1)	333 (129)	-	284 (111)	258 (149)

Production orthographique										
ODÉDYS	REG S (/10)	7.75 (1.89)	7.57 (1.90)	8.17 (1.83)	7.00 (2.16)	8.86 (0.90)	8.00 (2.10)	-	8.00 (2.65)	7.38 (3.16)
	IRR (/10)	4.00 (1.41)	4.29 (2.63)	7.00 (2.37)	5.75 (2.06)	5.00 (1.73)	5.50 (3.21)	-	4.89 (2.98)	5.63 (2.13)
	NM B (/10)	7.75 (1.26)	7.86 (2.19)	8.33 (1.51)	7.50 (3.00)	7.57 (2.51)	8.67 (1.37)	-	7.89 (2.09)	7.00 (2.33)

CTL = Indice de fluence de lecture. EC = Entraînement Contrôle. EM = Entraînement Morphologique. EVIP = Échelle de Vocabulaire en Images Peabody. IRR = Mots irréguliers. MPC = Matrices Progressives Colorées. NM B = Non-mots bisyllabiques. ODÉDYS = Outil de Dépistage des Dyslexies. REG S = Mots réguliers simples.

6.3.2.1 Analyses préalables : comparaison des patterns de résultats des groupes EM et EC-PO entre l'étude 1 et 2

Étant donné que nous avons supprimé des participants dans les groupes EM et EC-PO, nous souhaitons vérifier si les patterns de résultats pour les différentes variables d'intérêt demeuraient semblables à ceux de l'étude 1 pour les 34 participants sélectionnés, et ce afin de s'assurer que les interrogations sur la base desquelles nous avons mené cette seconde étude (cf. 6.2.3) étaient applicables à ce nouvel échantillon.

Les patterns de résultats obtenus via les ANOVA menées sur les groupes EM et EC-PO étaient globalement similaires entre l'étude 1 ($n = 51$) et l'étude 2 ($n = 34$). Nous notons toutefois trois points de divergence. En premier lieu, concernant l'identification des mots (BALE), les résultats obtenus à l'étude 1 sur les non-mots au pattern sublexical fréquent montraient un effet tendanciel du point de mesure ($F(1,49) = 3.05, p = .09, \eta^2_p = .060$). Cet effet n'était plus significatif après sélection des participants pour l'étude 2 ($F(1,32) = 2.64, p = .11, \eta^2_p = .076$). Toutefois, nous notons que la p-valeur demeure dans le même ordre de grandeur. Une autre différence est notée concernant les mots irréguliers peu fréquents : les résultats de l'étude 1, menée sur les 51 participants, montraient une interaction significative entre le point de mesure et le groupe d'entraînement ($F(1,48) = 4.19, p = .046, \eta^2_p = .080$). Après sélection des participants pour l'étude 2, cette interaction n'est plus significative ($F(1,31) = 2.68, p = .11, \eta^2_p = .080$). À nouveau, cependant, les p-valeurs pour ces interactions demeuraient proches du seuil de tendance que nous avons préalablement établi ($[.05 ; .10]$). Enfin, concernant la production orthographique (DIMO), un effet du point de mesure était constaté dans l'étude 1 pour le score « bases » des mots non-entraînés ($F(1,49) = 4.08, p = .049, \eta^2_p = .077$). Cet effet n'est pas retrouvé après sélection des participants pour l'étude 2 ($F(1,32) = 1.11, p = .30, \eta^2_p = .034$). Ainsi, excepté pour l'absence d'évolution du score « bases » des mots non-entraînés de la DIMO, les résultats obtenus par les 51

participants des groupes EM et EC-PO à l'étude 1 et par les 34 participants issus de ce même échantillon et sélectionnés pour l'étude 2 semblent similaires.

Concernant les analyses de contraste, seules trois différences dans les seuils de significativité sont notées : au score « mots entraînés » de la COMOS, les 24 sujets du groupe EC présentaient une amélioration non-significative de leurs performances entre le pré-test et le post-test ($t(48) = -1.56, p = .13$). Après suppression des 7 participants pour constituer l'échantillon EC-PO de l'étude 2, cette amélioration atteint le seuil de tendance ($t(31) = -1.75, p = .09$). Nous notons toutefois que ces deux p-valeurs restent dans le même ordre de grandeur. De manière similaire, le score « mots non-entraînés » ne donnait pas lieu à une amélioration entre le pré-test et le post-test dans le groupe EC lorsque les 24 sujets du groupe étaient considérés ($t(48) = -1.09, p = .28$). Après retrait des 7 participants pour former le groupe EC-PO, cette amélioration atteint le seuil de tendance ($t(31) = -1.70, p = .10$). La dernière différence concerne l'indice de fluence pour les mots irréguliers peu fréquents : les 27 participants du groupe EM obtenaient une amélioration significative entre le pré-test et le post-test ($t(48) = -2.25, p = .03$). Après suppression des 10 participants pour constituer l'échantillon EM de l'étude 2, cette différence n'est plus significative ($t(31) = -1.25, p = .22$).

En résumé, après suppression des participants dans les groupes EM et EC de l'étude 1, les patterns de résultats restent globalement similaires. Nous notons toutefois que les sujets retenus pour former le groupe EC-PO semblent donc avoir tendancielllement bénéficié de l'entraînement à la production orthographique pour améliorer leurs performances en conscience morphologique. Toutefois, nous n'avons pas constaté d'amélioration en identification des mots sur les mots irréguliers peu fréquents. L'hypothèse qu'un EM pourrait bénéficier à une liste de lecture constituée à 25% de mots complexes ne semble donc pas tenir pour le groupe EM nouvellement formé. Les performances obtenues par nos trois groupes sur l'ensemble des variables dépendantes sont présentées dans le Tableau 5.

Tableau 5. Moyennes et écarts-types des groupes EM, EC-PO et EC-M sur l'ensemble des variables dépendantes

		Pré-test			Post-test			
		EM n = 17	EC-PO n = 17	EC-M n = 17	EM n = 17	EC-PO n = 17	EC-M n = 17	
COMOS	Mots entraînés (/6)	3.25 (1.61)	3.41 (1.28)	2.94 (1.20)	5.00 (0.94)	4.00 (1.41)	3.59 (1.54)	
	Mots non-entraînés (/6)	3.50 (1.79)	3.29 (1.49)	3.18 (1.38)	4.18 (1.13)	3.76 (1.48)	3.88 (1.27)	
DIMO	Mots entraînés (/8)	Mot entier	2.94 (2.19)	3.12 (2.18)	3.06 (2.22)	6.35 (1.46)	6.35 (2.06)	3.65 (2.40)
		Bases	4.47 (2.10)	4.59 (1.94)	4.47 (2.43)	7.24 (0.83)	6.88 (1.41)	5.12 (2.15)
		Suffixes	4.88 (2.32)	5.47 (2.15)	4.76 (2.70)	7.18 (1.13)	7.18 (1.42)	5.41 (2.55)
	Mots non-entraînés (/8)	Mots entier	3.18 (2.24)	3.73 (2.22)	3.18 (2.13)	4.24 (1.89)	5.00 (1.87)	3.47 (2.12)
		Bases	4.82 (1.67)	5.47 (1.50)	5.06 (1.68)	5.06 (1.56)	5.65 (1.69)	4.88 (1.80)
		Suffixes	4.94 (2.33)	4.82 (2.58)	4.65 (2.23)	6.65 (1.22)	6.76 (1.79)	5.06 (2.54)
	BALE	Mots irréguliers fréquents	80.0 (53.9)	81.5 (49.3)	92.6 (50.4)	89.7 (58.7)	93.4 (50.8)	89.0 (44.3)
		Mots réguliers fréquents	102 (55.0)	108 (45.1)	111 (50.4)	101 (49.9)	113 (43.4)	108 (46.4)
Non-mots au pattern sublexical fréquent		58.6 (26.6)	60.1 (30.8)	62.6 (32.0)	64.1 (30.7)	62.9 (29.6)	59.2 (30.0)	
Mots irréguliers peu fréquents		48.4 (47.9)	51.6 (47.7)	62.1 (42.2)	52.9 (46.4)	47.9 (39.3)	58.4 (43.0)	
Mots réguliers peu fréquents		69.5 (45.8)	76.4 (44.6)	86.7 (33.8)	74.6 (45.1)	78.9 (43.6)	74.7 (38.4)	
Non-mots au pattern sublexical peu fréquent		52.3 (36.6)	46.8 (19.0)	55.0 (27.5)	55.7 (29.7)	51.1 (19.9)	53.9 (32.1)	

BALE = Batterie Analytique du Langage Écrit. COMOS = Conscience Morphologique – Suffixation. DIMO = Dictée Morphologique. EC-M = Entraînement Contrôle – Maths seules. EC-PO = Entraînement Contrôle – Production Orthographique. EM = Entraînement Morphologique.

6.3.2.2 EM versus EC-PO versus EC-M : Conscience Morphologique – Suffixation

L'impact de l'intervention morphologique sur les capacités de conscience

morphologique est évalué via la tâche COMOS. Dans l'ensemble, nos sujets semblent avoir progressé sur les items entraînés ($F(2,47) = 24.92, p < .001, \eta^2_p = .346$)²⁹ entre le pré- et le post-test. L'interaction entre le point de mesure et le groupe d'entraînement était significative ($F(2,47) = 3.28, p = .046, \eta^2_p = .122$). Les analyses de contraste révèlent que le groupe EM a significativement augmenté ses performances entre le pré-test et le post-test ($t(47) = -4.89, p < .001$). Cette augmentation n'était que tendanciellement significative pour les groupe EC-PO ($t(47) = -1.76, p = .09$) et EC-M ($t(47) = -1.93, p = .06$). Au post-test, le groupe EM obtient des performances significativement supérieures au groupe EC-M ($t(47) = -2.87, p = .005$). Il obtient des performances tendanciellement supérieures au groupe EC-PO ($t(47) = -2.00, p = .05$). Les deux groupes contrôle ne se distinguent pas significativement au post-test ($t(47) = -0.89, p = .38$).

Nous notons également un effet du point de mesure pour les mots non-entraînés de la tâche COMOS ($F(2,47) = 10.15, p = .003, \eta^2_p = .178$). Nous ne notons pas d'interaction significative entre le point de mesure et le groupe d'entraînement ($F(2,47) < 1, p = .87, \eta^2_p = .006$). Concernant les analyses de contraste, de manière surprenante, seul le groupe EC-M a significativement progressé entre le pré-test et le post-test ($t(47) = -2.26, p = .03$). Cette augmentation n'est que tendancielle pour le groupe EM ($t(47) = -1.75, p = .09$) et n'atteint pas le seuil de significativité pour le groupe EC-PO ($t(47) = -1.51, p = .14$). Au post-test, on ne note aucune différence significative entre les groupes, qu'il s'agisse de la comparaison des groupes EM versus EC-M ($t(47) = -0.36, p = .72$), EM versus EC-PO ($t(47) = -0.60, p = .55$) ou EC-M versus EC-PO ($t(47) = 0.24, p = .81$).

²⁹ La donnée manquante sur la tâche COMOS est due à l'absence d'un sujet du groupe EM lors du pré-test.

6.3.2.3 EM versus EC-PO versus EC-M : identification des mots

L'ensemble de nos sujets semblent avoir progressé de manière significative entre le pré-test et le post-test sur les mots irréguliers fréquents ($F(2,47) = 5.05, p = .03, \eta^2_p = .097$)³⁰ ; une interaction tendancielle entre le point de mesure et le groupe d'entraînement est par ailleurs relevée pour cette variable ($F(2,47) = 3.13, p = .053, \eta^2_p = .118$). Les analyses de contraste montrent que le groupe EM ($t(47) = -2.19, p = .03$) et le groupe EC-PO ($t(47) = -2.42, p = .02$) ont tous deux significativement progressé entre le pré-test et le post-test, tandis que le groupe EC-M a diminué ses performances de manière non significative ($t(47) = 0.74, p = .46$). Toutefois, au post-test, nous ne constatons aucune différence significative entre nos groupes, qu'elle concerne les comparaisons EM versus EC-M ($t(47) = -0.04, p = .97$), EM versus EC-PO ($t(47) = 0.20, p = .84$) ou EC-PO versus EC-M ($t(47) = -0.25, p = .81$).

Concernant les mots réguliers fréquents, il n'y avait pas d'effet du point de mesure ($F(2,47) < 1, p = .68, \eta^2_p = .004$)³¹ ni d'interaction entre le point de mesure et le groupe d'entraînement ($F(2,47) < 1, p = .59, \eta^2_p = .022$). Les analyses de contraste montrent que ni le groupe EM ($t(47) = -0.39, p = .70$), ni le groupe EC-M ($t(47) = 0.56, p = .58$), ni le groupe EC-PO ($t(47) = -0.88, p = .38$) n'ont amélioré leurs performances de manière significative entre le pré-test et le post-test. Par conséquent, aucune différence significative n'est notée au post-test, qu'elle concerne les comparaisons EM versus EC-M ($t(47) = 0.39, p = .70$), EM versus EC-PO ($t(47) = 0.69, p = .50$) ou EC-PO versus EC-M ($t(47) = -0.30, p = .77$).

De manière similaire, nous ne notons pas d'effet du point de mesure pour les non-mots au pattern sublexical fréquent ($F(2,48) < 1, p = .48, \eta^2_p = .011$), ni d'interaction entre le point de mesure et le groupe d'entraînement ($F(2,48) = 1.35, p = .27, \eta^2_p = .053$). À nouveau, les analyses de contraste ne retrouvent aucune amélioration significative entre le pré-test et le

³⁰ La donnée manquante sur cette variable est due à l'absence d'un sujet du groupe EM lors du post-test.

³¹ La donnée manquante sur cette variable est due à l'absence d'enregistrement de la performance d'un sujet du groupe EM lors du post-test.

post-test, que ce soit pour le groupe EM ($t(48) = -1.41, p = .17$), EC-M ($t(48) = -0.87, p = .39$), et EC-PO ($t(48) = -0.70, p = .49$). De même, aucune différence significative n'est notée au post-test, qu'elle concerne les comparaisons EM versus EC-M ($t(48) = -0.48, p = .64$), EM versus EC-PO ($t(48) = -0.12, p = .90$) ou EC-PO versus EC-M ($t(48) = -0.36, p = .72$).

Il n'y avait pas d'effet du point de mesure pour les mots irréguliers peu fréquents ($F(2,45) < 1, p = .60, \eta^2_p = .006$)³², ni d'interaction entre le point de mesure et le groupe d'entraînement ($F(2,45) = 1.75, p = .19, \eta^2_p = .072$). Là encore, les analyses de contraste ne retrouvent aucune amélioration significative entre le pré-test et le post-test, que ce soit pour le groupe EM ($t(45) = -1.37, p = .18$), EC-M ($t(45) = -0.74, p = .46$), et EC-PO ($t(45) = 1.17, p = .25$). Aucune différence significative n'est notée au post-test, qu'elle concerne les comparaisons EM versus EC-M ($t(45) = 0.59, p = .56$), EM versus EC-PO ($t(45) = -0.32, p = .75$) ou EC-PO versus EC-M ($t(45) = 0.88, p = .38$).

Concernant les mots réguliers peu fréquents, nous ne retrouvons pas d'effet du point de mesure ($F(2,47) < 1, p = .99, \eta^2_p = .000$)³³. Toutefois, une interaction tendancielle est notée entre le point de mesure et le groupe d'entraînement ($F(2,47) = 3.09, p = .055, \eta^2_p = .116$). Les analyses de contraste ne retrouvent pas d'amélioration significative entre le pré-test et le post-test pour les groupes EM ($t(47) = -1.34, p = .19$) et EC-PO ($t(47) = -0.67, p = .51$). En ce qui concerne le groupe EC-M, nous retrouvons une diminution tendancielle significative des performances entre les deux points de mesure ($t(47) = 1.98, p = .053$). Au post-test, nous ne notons pas de différence significative entre nos groupes, que cela concerne les comparaisons EM versus EC-M ($t(47) = 0.30, p = .77$), EM versus EC-PO ($t(47) = 0.31, p = .76$) ou EC-PO versus EC-M ($t(47) = -0.00, p = 1$).

³² Les deux données manquantes sur cette mesure sont dues à l'absence d'enregistrement de la performance d'un sujet du groupe EC-M au pré-test, et d'un autre sujet de ce groupe, ainsi que du groupe EC-PO, lors du post-test.

³³ La donnée manquante sur cette variable est due à l'absence d'enregistrement de la performance d'un sujet du groupe EC-M lors du post-test.

Enfin, concernant les non-mots au pattern sublexical peu fréquent, nous notons un effet tendanciel du point de mesure ($F(2,47) = 2.86, p = .10, \eta^2_p = .057$)³⁴. Il n'y avait pas d'interaction entre le point de mesure et le groupe d'entraînement ($F(2,47) < 1, p = .88, \eta^2_p = .006$). En dépit de l'effet tendanciel relevé sur le point de mesure, les analyses de contraste ne révèlent pas de différence significative entre le pré-test et le post-test, que ce soit pour le groupe EM ($t(47) = -1.04, p = .30$), EC-M ($t(47) = -0.58, p = .57$), et EC-PO ($t(47) = -1.32, p = .19$). Au post-test, aucune différence significative entre nos groupes n'est relevée, que cela concerne la comparaison EM versus EC-M ($t(47) = 0.12, p = .90$), EM versus EC-PO ($t(47) = -0.49, p = .63$) ou EC-PO versus EC-M ($t(47) = 0.60, p = .55$).

6.3.2.4 EM versus EC-PO versus EC-M : Dictée Morphologique

Un effet du point de mesure est retrouvé pour la production orthographique des mots entraînés, à la fois sur les scores « mots entiers » ($F(2,48) = 135.23, p < .001, \eta^2_p = .738$), « bases » ($F(2,48) = 68.86, p < .001, \eta^2_p = .578$) et « suffixes » ($F(2,48) = 44.18, p < .001, \eta^2_p = .479$). Une interaction significative entre le point de mesure et le groupe d'entraînement est notée pour tous les mots entraînés, aussi bien pour le score « mots entiers » ($F(2,48) = 19.39, p < .001, \eta^2_p = .447$), « bases » ($F(2,48) = 7.50, p = .001, \eta^2_p = .238$) que « suffixes » ($F(2,48) = 4.28, p = .02, \eta^2_p = .151$).

Un effet du point de mesure est également noté pour la production orthographique de mots non-entraînés, pour les scores « mots entiers » ($F(2,48) = 19.48, p < .001, \eta^2_p = .289$) et « suffixes » ($F(2,48) = 32.33, p < .001, \eta^2_p = .402$). Le score « bases » ($F(2,48) < 1, p = .64, \eta^2_p = .005$) relatifs aux mots non-entraînés n'a en revanche pas donné lieu à une amélioration significative entre le pré- et le post-test. Une interaction significative entre le point de mesure et le groupe d'entraînement est notée pour le score « suffixes » ($F(2,48) = 3.99, p = .03, \eta^2_p =$

³⁴ La donnée manquante sur cette variable est due à l'absence d'enregistrement de la performance d'un sujet du groupe EC-M lors du post-test.

.143). Nous ne relevons pas d'interaction significative entre le point de mesure et le groupe d'entraînement pour les scores « mots entiers » ($F(2,48) = 2.18, p = .12, \eta^2_p = .083$) et « bases » ($F(2,48) < 1, p = .56, \eta^2_p = .024$).

L'ensemble des analyses de contraste réalisées sur les résultats à la DIMO sont présentées dans le Tableau 6. Concernant les mots entraînés, le groupe EM tout comme le groupe EC-PO ont significativement amélioré leurs performances entre le pré- et le post-test sur les scores « mots entiers », « bases » et « suffixes », à l'inverse du groupe EC-M. Ce pattern de résultats est identique pour les scores « mots entiers » et « suffixes » des mots non-entraînés. Concernant le score « bases » des mots non-entraînés, cependant, aucun des groupes ne semble avoir progressé entre le pré- et le post-test. Par rapport au groupe EC-M, le groupe EM et le groupe EC-PO ont obtenu en post-test des performances significativement supérieures sur les trois scores relatifs aux mots entraînés, ainsi qu'au score « suffixes » des mots non-entraînés. Concernant le score « mots entiers » des mots non-entraînés, seul le groupe EC-PO a obtenu une performance supérieure au groupe EC-M au post-test. En revanche, ni le groupe EM ni le groupe EC-PO n'ont obtenu de score significativement supérieur au groupe EC-M sur le score « bases » des mots non-entraînés. Enfin, les groupes EM et EC-PO ne se différencient de manière significative sur aucune des mesures analysées.

Tableau 6. Analyses de contraste des performances à la DIMO

	Différences pré- et post-test									Différences entre les groupes d'entraînement au post-test								
	EM			EC-PO			EC-M			EM vs EC-M			EM vs EC-PO			EC-PO vs EC-M		
	<i>t</i>	<i>ddl</i>	<i>p</i>	<i>t</i>	<i>ddl</i>	<i>p</i>	<i>t</i>	<i>ddl</i>	<i>p</i>	<i>t</i>	<i>ddl</i>	<i>p</i>	<i>t</i>	<i>ddl</i>	<i>p</i>	<i>t</i>	<i>ddl</i>	<i>p</i>
ME mots entiers	- 9.50	48	< .001	- 9.01	48	< .001	- 1.64	48	.11	- 3.75	48	< .001	0.00	48	1	- 3.75	48	< .001
ME bases	- 6.81	48	< .001	- 5.65	48	< .001	- 1.59	48	.12	- 3.28	48	.002	- 0.55	48	.59	- 2.73	48	.008
ME suffixes	- 5.68	48	< .001	- 4.23	48	< .001	- 1.60	48	.12	- 2.42	48	.02	0.00	48	1	- 2.42	48	.02
NE mots entiers	- 3.13	48	.003	- 3.75	48	< .001	- 0.87	48	.39	- 1.07	48	.29	1.07	48	.29	- 2.14	48	.04
NE bases	- 0.82	48	.42	- 0.61	48	.54	0.61	48	.54	- 0.31	48	.76	1.04	48	.30	- 1.35	48	.18
NE suffixes	-4.14	48	< .001	-4.71	48	< .001	-1.00	48	.32	- 2.14	48	.04	0.16	48	.88	- 2.29	48	.03

DIMO = Dictée Morphologique. EC-M = Entraînement contrôle – maths seules. EC-PO = Entraînement contrôle – production orthographique. EM = Entraînement morphologique. ME = Mots entraînés. NE = Mots non-entraînés.

6.3.3 Discussion

L'objectif de l'étude 2 était de compléter les conclusions de l'étude 1 en incluant dans l'analyse un second groupe contrôle, cette fois uniquement entraîné aux mathématiques (EC-M). Ce dernier nous permettait de déterminer si les différences pré/post tests observées à la fois dans les groupes EM et EC-PO étaient bel et bien dues aux interventions, ou si elles n'étaient pas distinguables d'un effet placebo ou Hawthorne. En raison de la non-équivalence au pré-test du groupe EC-M et des précédents groupes EM et EC-PO sur certaines variables contrôle (raisonnement non-verbal) et d'intérêt (conscience morphologique), nous avons choisi de réaliser un appariement trois à trois, afin de minimiser les différences pré-test et de pouvoir réaliser les analyses sur des participants au profil le plus similaire possible. La suppression de participants dans les groupes EM et EC-PO a légèrement modifié les patterns des résultats des ANOVA menées sur ces groupes lors de l'étude 1. En particulier, nous notons une disparition de l'effet d'interaction précédemment obtenu pour la lecture des mots irréguliers peu fréquents et pour la transcription de bases non-entraînées.

Concernant la lecture de mots irréguliers fréquents et la production orthographique, l'ajout de ce deuxième groupe contrôle semble confirmer qu'un entraînement à la transcription de mots complexes sans analyse morphémique explicite suffit à générer un gain sur ces tâches. Pour la production orthographique, ceci se vérifie, que l'on évalue le bénéfice sur les mots entiers, leurs bases ou leurs suffixes entraînés (y compris, pour ces derniers, dans des mots non-entraînés), et ce de manière aussi efficace qu'un entraînement morphologique proposant des exercices de conscience morphologique en plus de la production orthographique de ces morphèmes (EM). Il semblerait même qu'il puisse être « plus efficace » qu'un EM dans une certaine mesure, puisque seul le groupe EC-PO a obtenu des performances supérieures au groupe EC-M au score « mots entiers » des mots non-entraînés. Une interprétation de ce résultat assez surprenant pourrait être qu'une intervention courte

(environ 15 minutes) ciblant spécifiquement la production orthographique engendre de meilleurs résultats pour cette même production orthographique lorsqu'elle n'est pas combinée à un autre entraînement langagier (analyse morphémique explicite). La dictée survenant après 30 minutes d'entraînement métalinguistique, les participants du groupe EM avaient potentiellement épuisé leurs ressources cognitives et attentionnelles, ne leur permettant pas d'engrammer aussi bien les représentations orthographiques travaillées que le groupe EC-PO. Une autre hypothèse est que les sujets du groupe EC-PO de l'étude 2 ont pu développer de manière implicite des compétences morphologiques comparables à celles du groupe EM, et qu'ils aient été davantage capables de les mobiliser lors d'une tâche de production orthographique. Cette interprétation se base sur l'augmentation tendancielle des performances du groupe EC-PO aux mots entraînés de la COMOS entre le pré-test et le post-test. Il est en effet possible que cette exposition à elle seule ait suffi à engager une réflexion sur la signification de ces morphèmes chez nos sujets. Les effets d'une instruction morphologique implicite ont été étudiés sur le développement de l'orthographe (Bar-Kochva et al., 2020; Bar-Kochva & Hasselhorn, 2017; Burton et al., 2021) mais pas, à notre connaissance, sur celui de la conscience morphologique. Il serait intéressant de déterminer dans quelle mesure l'exposition à du matériel morphologiquement construit sans pression environnementale particulière permet réellement de susciter un raisonnement métalinguistique chez des apprenants. Il convient toutefois de rappeler que ces interprétations sont réalisées sur des échantillons au faible nombre de participants. Rappelons également que les résultats auraient pu être différents sans l'exclusion des 10 participants nécessaire à la constitution du groupe EM de l'étude 2, puisqu'on note, entre l'étude 1 et l'étude 2, une disparition de la progression entre le pré-test et le post-test sur la transcription des bases non-entraînées, que nous avons précédemment interprété comme un possible transfert permis par l'analyse morphémique. Les sujets du groupe EM sélectionnés pour l'étude 2 semblent, en moyenne, avoir moins bénéficié

de l'entraînement à l'analyse morphémique que lorsque le groupe était considéré dans son ensemble (cf. étude 1).

Un dernier résultat surprenant obtenu lors de cette deuxième étude, et qui semble devoir être adressé, concerne le fait que seul le groupe EC-M, le seul groupe à n'avoir pas été exposé aux morphèmes durant leur entraînement, a progressé entre le pré-test et le post-test sur les mots non-entraînés de la COMOS. Il nous semble que l'explication doive être cherchée non pas dans le contenu de l'entraînement mathématique, mais dans un potentiel facteur confondant. En effet, les participants du groupe EC-M étaient issus d'une seule école, et différente de celles ayant permis le recrutement des participants des groupes EM et EC-PO. Il est donc par exemple possible que ces élèves aient bénéficié d'une instruction morphologique dispensée par leurs enseignants, en parallèle de leur entraînement mathématique proposé dans le cadre de cette étude. Il aurait également été opportun de contrôler l'enseignement morphologique reçu par les élèves en parallèle de l'entraînement prodigué, aussi bien dans l'étude 1 que l'étude 2. Si l'absence de contrôle de ce biais nous paraissait tolérable dans l'étude 1, puisqu'il était « réparti » équitablement à travers nos groupes (chacun comportant un nombre équivalent d'élèves issus d'une même classe), il apparaît particulièrement critique pour interpréter correctement les résultats de l'étude 2.

6.4 Discussion générale

La présente étude visait à contribuer à la compréhension de l'impact d'un entraînement à la conscience morphologique (EM) sous forme de jeu intégré au temps scolaire (deux fois 45 minutes par semaine) sur le développement de l'identification des mots écrits et de la production orthographique. Plus particulièrement, concernant l'identification des mots, nous avons souhaité déterminer si un EM permettait d'améliorer la fluence de lecture dans le cadre d'un test standardisé, ne présentant pas une majorité de mots complexes. Le but était d'évaluer dans quelle mesure l'EM peut constituer une intervention efficace pour

améliorer l'identification des mots en général, les études longitudinales et les études d'entraînement ne semblant pas aboutir à la même conclusion (voir § 3.2.2.2), les premières notant un lien entre les compétences morphologiques et les compétences de lecture ultérieures, tandis que les secondes ne semblent retrouver un effet de la conscience morphologique que sur les mots complexes (cf. Chapitre 5). Concernant la production orthographique, nous souhaitons estimer si le fait de stimuler la manipulation explicite de morphèmes, en plus d'entraîner à la production orthographique de ces derniers, représentait un avantage pour la transcription de mots suffixés, par rapport à un entraînement contrôle qui n'aborderait ces mots que via leur transcription sous dictée. Partant du constat que les entraînements morphologiques les plus efficaces pour le développement de la production orthographique incluent le plus souvent des activités de transcription de morphèmes (e.g., Nunes et al., 2003), nous souhaitons déterminer si cette seule activité de transcription était suffisante à améliorer les performances en production orthographique, ou si des exercices de manipulation consciente de morphèmes participaient de l'efficacité d'une telle intervention.

Pour ce faire, nous avons dans un premier temps (étude 1) comparé les performances en conscience morphologique, en identification des mots et en production orthographique en pré- et post-test d'un groupe EM (n = 27) à un groupe contrôle entraîné au calcul mental et à la transcription de morphèmes (EC-PO ; n = 24). Plusieurs résultats semblaient se dégager : tout d'abord, seul le groupe EM semblait avoir progressé sur l'épreuve de conscience morphologique des mots entraînés, sur l'identification de mots irréguliers peu fréquents et sur la production orthographique de bases non-entraînées, suggérant que l'analyse morphémique a pu constituer pour les participants du groupe EM une stratégie facilitant la réalisation de ces tâches, à l'opposé du groupe EC-PO. Ces résultats étant tout à fait inédits à notre connaissance, mais conclus sur la base d'échantillons au faible nombre de participants, ils demandent toutefois à être confirmés par des études ultérieures. De manière surprenante, en

revanche, nos deux groupes semblaient avoir tous deux progressé de manière similaire sur l'identification des mots irréguliers fréquents et sur le reste des tâches de production orthographique de mots suffixés. L'entraînement à la transcription sans erreur constituant une intervention efficace pour la production orthographique (Wanzek et al., 2006; Williams et al., 2017), il n'était pas possible de déterminer si ces progrès observés pour les deux groupes entre le pré-test et le post-test étaient dus à un simple effet Hawthorne / placebo, ou si les deux entraînements pouvaient être considérés comme efficaces.

Afin d'approfondir ces résultats, nous avons donc recruté dans un second temps (étude 2) un groupe contrôle supplémentaire, cette fois uniquement entraîné aux mathématiques (EC-M). Malheureusement, les écoles dans lesquelles nous avons effectué les entraînements de l'étude 1 n'ayant pas souhaité renouveler leur participation, nous avons dû effectuer le recrutement de l'étude 2 dans une autre école, et avons obtenu un groupe de participants se distinguant significativement des groupes EM et EC-PO sur des mesures pré-test critiques (conscience morphologique et raisonnement non-verbal). De fait, nous avons choisi de réaliser un appariement trois à trois sur nos trois groupes (17 participants dans chaque groupe), afin de ne conserver que des participants au profil similaire. Cette sélection a toutefois remis en cause les conclusions de l'étude 1, puisque nous n'observons plus de progrès entre le pré-test et le post-test sur l'identification des mots irréguliers peu fréquents, ni sur la transcription de bases non-entraînées de la part des sujets du nouveau groupe EM, et ce avant même d'inclure le groupe EC-M. Par ailleurs, les sujets sélectionnés du groupe EC-PO semblent améliorer leurs performances en conscience morphologique entre le pré-test et le post-test.

L'étude 2 semble confirmer que les entraînements EM et EC-PO sont tous deux bien efficaces pour améliorer l'identifications de mots irréguliers fréquents et la production orthographique de mots suffixés, puisque ces deux groupes ont significativement progressé

entre le pré-test et le post-test, contrairement au groupe EC-M. Toutefois, les groupes EM et EC-PO ne se distinguent en post-test sur aucune de nos variables d'intérêt, suggérant que l'analyse morphémique explicite n'apporte pas d'avantage particulier par rapport à l'entraînement à la production orthographique de morphèmes, excepté, de manière tendancielle, pour les items entraînés de la tâche de conscience morphologique.

Cette conclusion complète les résultats obtenus par Nunes et al. (2003), qui ont évalué l'efficacité d'entraînements morphologique et phonologique, avec ou sans activité de transfert écrit. Dans leur étude, il ressort que la pratique de l'écrit génère toujours un bénéfice supérieur à l'entraînement métalinguistique seul, quel qu'il soit. Ceci s'explique probablement par le fait que l'activité de transcription contraint le sujet à traiter l'ensemble des unités graphémiques d'un mot. De fait, la représentation orthographique issue d'un tel exercice est plus précise que celle qui pourrait être atteinte par seule exposition à une forme écrite (Shahar-Yames & Share, 2008). Il est même possible qu'elle soit si prégnante qu'elle ait engendré, dans le cadre de notre étude, un plafonnement des scores à la DIMO aussi bien dans le groupe EM que EC-PO. Il est possible que la mesure utilisée n'ait pas été assez sensible pour capturer une potentielle différence de performance entre ces deux groupes : les deux entraînements apparaissent donc donner lieu à des gains équivalents. Il est important de signaler que la dictée d'entraînement abordait des items évalués lors de la DIMO. Ces items sont par ailleurs les seuls à avoir généré des améliorations (i.e., scores « mots entiers », « bases » et « suffixes » des mots entraînés ; score « suffixes » des mots non-entraînés). Les performances très similaires des groupes EM et EC-PO à la DIMO pourraient ainsi également s'expliquer par un effet de *testing* partagé (Rowland, 2014) : le fait d'avoir régulièrement présenté ces items dans une situation proche de celle de l'évaluation finale a pu permettre aux deux groupes de réaliser des progrès importants sur la transcription de ces mots entraînés. Leur performances auraient ainsi peut-être été différentes si la situation d'évaluation avait

d'avantage différé de la situation d'enseignement. Par exemple, le fait d'évaluer la qualité de la transcription orthographique via la production d'un texte libre aurait davantage pu permettre de mettre en évidence un effet de la manipulation consciente de morphèmes sur les facultés de production orthographique (Wolter & Dilworth, 2014). Il paraît en effet essentiel d'examiner de quelle manière une intervention morphologique peut permettre aux apprenants d'avoir recours à la bonne orthographe d'un mot, dans une situation plus écologique que la simple transcription sous dictée.

En ce qui concerne les améliorations entre le pré-test et le post-test obtenues sur l'identification des mots irréguliers fréquents par les groupes entraînés à la production orthographique de morphèmes (i.e., EM et EC-PO) mais pas par le groupe EC-M, ce résultat est à notre connaissance inédit. Une hypothèse pourrait être que la prise de conscience de la forme orthographique des unités morphémiques, même uniquement de manière implicite dans le cas du groupe EC-PO, ait incité les participants de ces deux groupes à faire davantage appel à leur système sémantique lors du décodage du fait de la capacité de l'unité morphémique à receler à la fois une forme et un sens. Les deux groupes ont pu ainsi améliorer leur précision de lecture, le système sémantique permettant, selon le modèle ST-DRC, de « compenser » les irrégularités grapho-phonémiques en activant des représentations phonologiques compatibles (cf. 1.3.3). Une autre hypothèse pourrait être que l'activité de production orthographique elle-même a permis ces progrès en identification des mots irréguliers fréquents. Ainsi que nous l'avons évoqué dans le paragraphe précédent, la transcription oblige à traiter chaque élément de la séquence phono-graphémique. L'entraînement à la transcription avec feedback correctif immédiat a pu permettre de renforcer les correspondances phonème-graphème (CPG) et graphème-phonème (CGP) de nos participants, conduisant à une amélioration en lecture. En effet, bien que les progrès soient constatés sur des mots irréguliers, et donc inconsistants du point de vue de ces correspondances, tout mot irrégulier comporte une certaine proportion de

graphèmes consistants (Share, 1995). Par exemple, dans le mot FEMME, bien que le E ne suive pas les règles de CGP habituelles, les graphèmes restants sont tout à fait consistants. Ainsi, tout travail de transcription renforçant les CPG (et donc, dans une certaine mesure, les CGP), est susceptible de permettre une amélioration en identification des mots, y compris lorsque ceux-ci sont inconsistants. Pour finir sur les résultats concernant l'identification des mots, il est intéressant de s'interroger sur la raison pour laquelle seuls les mots irréguliers fréquents semblent avoir bénéficié des entraînements EM et EC-PO. En effet, cette liste ne comportant que 15% de mots complexes et aucun morphème travaillé en entraînement, il paraît surprenant qu'elle soit la seule à avoir apparemment bénéficié d'un travail sur les morphèmes. Que l'on privilégie la première ou la seconde hypothèse avancée, il nous semble que deux éléments de réponse sont essentiels à considérer : 1) ces mots étaient suffisamment fréquents pour qu'il existe chez nos participants une représentation phonologique ; 2) ces mots étaient d'un niveau de difficulté adéquat pour être sensible à une amélioration pré/post-test, les autres listes de mots étant potentiellement trop « faciles » (e.g., mots réguliers fréquents), ou au contraire trop « difficiles » (e.g., mots irréguliers peu fréquents).

L'ensemble des interprétations exposées ci-dessus doit cependant tenir compte des limites présentées par notre étude. Nous avons soulevé les biais méthodologiques évidents dans les discussions des études 1 (absence de contrôle d'un effet Hawthorne / placebo) et 2 (introduction de facteurs confondants relatifs au groupe EC-M ; absence de contrôle de l'instruction morphologique en-dehors des entraînements), ainsi que ceux communs aux deux études (petits échantillons exacerbant les potentielles variations inter-individuelles). Outre ces biais, il est important de rappeler que les considérations relatives à la liste des mots entraînés de la tâche de COMOS et du score « bases » des mots non-entraînés de la DIMO sont limités par la faible fiabilité inter-item de ces mesures ($\alpha \approx .50$). Celle-ci a en effet pu impacter la mesure des performances de nos sujets, tant dans le sens d'une surestimation que d'une sous-

estimation du score réel (Tavakol & Dennick, 2011). Par ailleurs, bien que nous ayons fourni des scripts à l'ensemble des adultes chargées de dispenser les entraînements, nous n'avons pas évalué formellement la fidélité du traitement, par exemple en leur demandant de cocher les cases d'une *checklist* récapitulant les étapes durant la session d'entraînement, ou en filmant les sessions réalisées afin de vérifier que les étapes du script étaient bien respectées. Il n'a donc pas été possible d'estimer la fidélité du traitement pour chaque expérimentatrice, et ainsi de déterminer si nos résultats ont souffert d'une potentielle variabilité dans la manière d'administrer l'intervention. Enfin, la portée de nos résultats est par ailleurs limitée par l'absence de post-test réalisé à distance de l'entraînement (*follow-up*), bornant nos conclusions à la période suivant immédiatement la fin des entraînements.

En conclusion, les résultats de notre étude sur l'impact d'un entraînement morphologique ludique intégré à la pratique courante suggèrent que la transcription de morphèmes permet d'obtenir des gains sur la conscience morphologique, sur l'identification des mots et sur la production orthographique. Toutefois, nous n'avons pas pu mettre en évidence l'intérêt ajouté d'exercer à la manipulation explicite des morphèmes. Enfin, nos résultats pâtissent de nos faibles échantillons et de plusieurs biais méthodologiques, limitant la fiabilité de ces conclusions. Il nous semble essentiel que cette recherche soit répliquée en adressant ces biais, car les résultats relevés apparaissent d'une part critiques dans le cadre d'applications fondamentales, et d'autre part tout à fait inédites, notamment en ce qui concerne l'identification des mots.

Chapitre 7 – Exploration de l’impact d’un entraînement orthographique sublexical sur le développement de la production orthographique et l’identification des mots en français : l’exemple du phonème final /o/

7.1 Introduction

Selon un rapport de la Direction de l’Évaluation, de la Prospective et de la Performance (DEPP, 2016), le niveau en orthographe des élèves français en classe élémentaire connaît une baisse soutenue depuis 1987. Bien que l’essentiel des erreurs commises lors de la dictée d’évaluation proposée concerne l’orthographe grammaticale, les erreurs lexicales augmentent notablement entre 1987 et 2015. Par exemple, le mot « chemin », correctement orthographié par 92.2% des enfants de CM2 en 1987, ne l’est que par 79.9% des élèves de ce même niveau de classe en 2015, confirmant une première baisse des performances, notée en 2007 (85.9% de réussite sur ce mot). Bien que la diminution des performances en orthographe des élèves français en dictée de texte soit connue et régulièrement évoquée (Manesse et al., 2007), il demeure difficile d’en identifier les causes et par conséquent de l’adresser de manière ciblée. Il paraît ainsi nécessaire, afin de lutter contre les conséquences négatives associées à un mauvais niveau d’orthographe, de proposer aux élèves des stratégies innovantes et efficaces, en particulier pour ceux à risque de présenter un trouble de la lecture.

La méta-analyse de Galuschka et al. (2020) propose une synthèse statistique de l’effet de différentes interventions ciblant la production orthographique de sujets avec un trouble de l’orthographe (i.e., dont la performance aux tests d’orthographe est au moins inférieure au 25^{ème} percentile). Parmi les entraînements identifiés comme efficaces, les entraînements dits « orthographiques » étaient considérés distinctement des entraînements phonologiques et morphologiques, et concernaient les interventions proposant une « instruction systématique de l’application de règles d’orthographe graphotactiques ou ortho-phonologiques » (p. 8). Nous

avons choisi, dans le cadre du présent chapitre, de nommer ces interventions « entraînements orthographiques sublexicaux » (EOS), puisqu'ils ne concernent non pas l'orthographe des mots dans leur entièreté, mais bien les unités sublexicales pouvant faire l'objet de règles ou régularités applicables à un grand nombre d'items. Ainsi que nous l'avons évoqué dans le point 4.1.2, après examen des six études d'entraînement incluses dans la méta-analyse de Galuschka et al. (2020) sous cette catégorie, il apparaît que seules trois études, dont une seule rédigée en anglais, correspondaient réellement à un entraînement orthographique tel que défini par les auteurs eux-mêmes. Cette dernière, menée par Ise et Schulte-Körne (2010), présentée dans le point 4.2.2, visait à déterminer l'impact d'un tel entraînement sur la lecture et l'orthographe d'apprenants de 5^{ème} et de 6^{ème} année d'instruction. Il consistait en 12 à 15 séances hebdomadaires de 60 minutes ayant pour objectif l'apprentissage d'un algorithme décisionnel. Celui-ci ciblait l'application de régularités graphotactiques ou grapho-phonologiques à la production orthographique de mots, e.g., « si une voyelle courte est suivie par seulement une seule consonne au sein d'un même morphème, alors la consonne doit être doublée lors de la transcription (e.g., MANN [homme] » ; p. 22). Un gain entre le pré- et le post-test a été attribué à cette intervention, mais l'absence de groupe contrôle actif ne permettait pas de distinguer un potentiel effet de l'intervention d'un effet développemental en ce qui concernait la mesure de lecture, et d'un effet placebo ou Hawthorne concernant la mesure de production orthographique. Ainsi, bien que l'EOS ait été mentionné comme efficace à la fois pour l'identification des mots ($g' = 0.19 [0.09 ; 0.29]$) et pour la transcription ($g' = 0.67 [0.05 ; 1.28]$) dans la méta-analyse de Galuschka et al. (2020), l'effet de ce type d'intervention sur le développement de la littératie demeure en grande partie à explorer.

La compréhension de l'impact de l'EOS sur la littératie est rendue ardue par la difficulté à élaborer un entraînement portant sur des unités orthographiques qui ne soit ni purement graphophonologique (cf. § 2.2), ni purement morphologique. Pourtant, ainsi que

nous l'avons rappelé dans les chapitres 2 et 3 du présent manuscrit, seuls 50% des mots peuvent s'écrire par simple application des correspondances phonème-graphème (Véronis, 1988), et environ 20% des mots de la langue française sont monomorphémiques (Rey-Debove, 1984) et ne peuvent donc s'écrire en référence à d'autres morphèmes (excepté les mots monomorphémiques présentant des lettres muettes morphologiquement motivées, e.g., BLANC, CHANT, etc.). Bien que minoritaires sur l'ensemble du corpus, les mots à la fois simples et inconsistants sont souvent très fréquents (e.g., OISEAU ; Lété et al., 2004; Peereman et al., 2007), et donc essentiels à maîtriser. Il paraît ainsi pertinent de réfléchir à des régularités susceptibles de compléter les stratégies graphophonologiques et morphologiques, plus répandues et plus étudiées. Le phonème final /o/, nommément, est particulièrement inconsistant en français, puisqu'il peut s'orthographier de 12 façons différentes : O (RADIO), OT (HARICOT), ÔT (AUSSITÔT), OS (CHAOS), OP (GALOP), OC (ACCROC), AU (PRÉAU), AUX (ANIMAUX), AUD (CRAPAUD), AUT (DÉFAUT), EAU (CORBEAU). Concernant cette terminaison précise, ni un entraînement morphologique, ni un entraînement graphophonologique seuls n'apparaissent suffisants pour déduire la graphie qu'il convient d'employer en toute situation. Nous basant sur la Théorie de l'Intégration des Multiples Patterns (Treiman, 2017b), qui souligne l'importance d'enseigner conjointement les patterns phonologiques, morphologiques et graphotactiques, nous avons mis au point un EOS ciblant la terminaison /o/. Ce dernier, détaillé dans le point 7.2.3, intègre à la fois une dimension phonologique (i.e., identification du phonème final ; segmentation syllabique), morphologique (i.e., lettres muettes à valeur morphologique ; règle de morphologie flexionnelle ; éléments d'étymologie) et graphotactique (i.e., notion de graphie légale versus illégale ; notion de graphie fréquente versus peu fréquente). Nous avons synthétisé l'ensemble de ces stratégies métalinguistiques en un algorithme décisionnel, en nous inspirant des travaux d'Ise et Schulte-Körne (2010 ; voir § 7.2.3.3).

À notre connaissance, peu d'études d'entraînement se sont penchées sur l'intérêt d'un EOS empruntant à l'ensemble des dimensions susnommées. L'une d'entre elle, réalisée par Cordewener et al. (2016), a comparé l'efficacité d'une stratégie orthographique basée sur des « principes phonologiques, morphologiques et orthographiques » (p. 50) à une approche d'auto-correction, chez des enfants de troisième année d'instruction. Les sujets ayant appris à appliquer les stratégies métalinguistiques au cours de 4 sessions d'entraînement ont amélioré de manière significative leurs performances sur des mots non-entraînés entre le pré-test et le post-test, et ce de manière significativement plus importante qu'un groupe de sujets s'étant exercé à écrire des mots sans feedback correctif sur la même période de temps. Le groupe ayant appliqué une stratégie d'auto-correction a également amélioré leurs performances de manière significative, mais le gain n'était pas significativement distinct de celui réalisé par les sujets sans feedback correctif.

En dépit de ce premier résultat, de nombreux paramètres restent à déterminer afin de correctement caractériser l'impact à attendre d'un tel entraînement. Il convient notamment d'établir le profil de participants pouvant bénéficier d'une telle intervention. En effet, il n'est pas rare que des entraînements reconnus efficaces pour un public ciblé possèdent une portée limitée pour d'autres. En outre, la présence ou non de troubles de la lecture conditionne l'efficacité d'interventions phonémiques sur la production orthographique, de même que l'âge auquel ces entraînements sont proposés (Ehri et al., 2001). En second lieu, les caractéristiques de l'entraînement, et plus précisément, un mode d'instruction explicite (Hughes et al., 2017) ou non, paraît en moduler l'efficacité. Les approches non-explicites sont établies comme moins (voir peu) efficaces pour l'acquisition de certains apprentissages fondamentaux, tels que la lecture et les mathématiques (Bissonnette et al., 2010). Pour la production orthographique, cependant, des gains significatifs ont été constatés avec une approche non-explicite (Arfé et al., 2018; Kemper et al., 2012). Dans l'étude de Kemper et al. (2012), qui

visait à comparer l'impact d'une approche explicite versus implicite sur l'acquisition d'une règle d'orthographe de morphologie flexionnelle auprès d'enfants avec ou sans trouble de l'orthographe, des résultats différents ont ainsi été obtenus en fonction du profil d'apprenant. En effet, bien que seul l'enseignement explicite ait permis aux sujets de généraliser les règles orthographiques abordées à des mots non-entraînés, il est intéressant de noter que, lorsque les mots entraînés et non-entraînés étaient considérés ensemble, l'instruction explicite n'était supérieure à l'instruction implicite que chez le groupe d'enfants sans trouble de l'orthographe. Le groupe présentant un trouble de l'orthographe, aux performances orthographiques similaires aux enfants sans trouble mais plus âgés d'un ou deux ans, ne semblait ainsi pas bénéficier d'une explicitation de la règle ciblée. Les deux groupes d'enfants se distinguant à la fois du point de vue de leur niveau de lecture et de leur âge moyen, l'absence de supériorité de l'enseignement explicite dans le groupe présentant un trouble de l'orthographe pourrait être attribuée soit à leurs difficultés linguistiques, soit à leur âge plus élevé. Étant donné le faible nombre d'études investiguant l'impact d'un EOS sur le développement de la production orthographique, nous avons choisi de recourir à un panel diversifié d'apprenants, présentant ou non un trouble de la lecture, et appartenant à trois niveaux de classe différents (2^{ème}, 4^{ème} ou 6^{ème} année d'instruction), et d'en explorer l'efficacité à la fois selon une modalité explicite et implicite.

Il nous paraît important, lorsqu'il est question d'opposer un mode d'apprentissage explicite versus implicite dans le domaine de l'orthographe, d'établir une distinction entre règle et régularité. Dans leur étude de 2005, Pacton et collaborateurs proposaient de considérer comme règle, en prenant l'exemple de la transcription des diminutifs /o/ (-EAU) et /ɛt/ (-ETTE), une convention d'écriture qui fait l'objet d'une définition formalisée (e.g., « une petite *TRAVE est une *TRAVETTE. ») La notion de régularité, quant à elle, était abordée en tant que convention d'écriture que les usagers d'une langue écrite appliquent, mais qu'ils

ne verbalisent pas consciemment (e.g., la terminaison -EAU est souvent retrouvée après les consonnes R, T ou V mais jamais après K, F et G). Dans la présente étude, la distinction entre règle et régularité est directement mise en lien avec la modalité d'apprentissage. Ainsi, dans le groupe entraîné explicitement, la convention d'écriture travaillée était systématiquement présentée comme une règle, i.e., suivant une définition formalisée, tandis que pour le groupe entraîné implicitement, les différentes conventions d'écriture étaient simplement mises en avant en tant que régularité récurrente du système de transcription (voir § 7.2.3.2).

En complément du design pré-/post-test fréquemment employé pour évaluer quantitativement les effets d'une intervention sur le développement de la littératie, nous avons également eu recours à une approche plus qualitative, nous permettant de décrire en détail l'impact individuel de l'EOS sur nos sujets : le *Single-Case Design* ou SCD (Lobo et al., 2017). Cette approche intra-sujet se base sur l'étude de l'évolution de la performance d'un participant au cours du temps, via la réalisation de mesures répétées, avant et pendant l'intervention. Bien que principalement employé en sciences de l'éducation (Maggin et al., 2018) et de la réadaptation (Krasny-Pacini & Evans, 2018), ce *design* présente également un intérêt pour le domaine de la psychologie du développement (Cohen et al., 2014), en ce qu'il permet de décrire précisément la trajectoire développementale d'un individu, sans la sacrifier à celle du groupe.

L'objectif de cette étude était ainsi d'explorer dans quelle mesure un entraînement orthographique sublexical (EOS) portant sur le phonème final /o/ présente un intérêt pour le développement de la production orthographique et de l'identification des mots. Nous avons exploré cette efficacité chez des enfants de trois niveaux de classe différents (2^{ème}, 4^{ème} ou 6^{ème} année d'instruction), avec ou sans trouble de la lecture. De plus, l'intervention était dispensée soit selon une modalité soit explicite, soit implicite. Concernant la production orthographique plus particulièrement, l'efficacité de l'entraînement était doublement évaluée,

à la fois selon une approche quantitative traditionnelle pré-/post-test, et selon une approche plus qualitative en *Single-Case Design*.

7.2 Méthode

7.2.1 Participants

Les participants ont été recrutés via une affiche transmise à des médecins pédiatres et des directeurs d'école du département du Bas-Rhin, qui l'ont ensuite diffusée auprès de leurs patients et de leurs élèves. Cette procédure de recrutement a été privilégiée dans le contexte de la pandémie de COVID-19, où nous ne savions pas si les écoles pouvaient fermer à nouveau leurs portes, et étant donné que nous ne souhaitions pas surcharger les établissements déjà sollicités pour l'étude d'entraînement morphologique. L'annonce contenait l'adresse e-mail et le numéro de téléphone de la doctorante. Les parents intéressés pouvaient ainsi prendre contact avec elle afin d'inscrire leur enfant au protocole de recherche. Via l'affiche, les parents étaient informés que nous recherchions des enfants scolarisés en 2^{ème}, 4^{ème} ou 6^{ème} année d'instruction, avec ou sans diagnostic de dyslexie. En fonction des préférences et des disponibilités des parents, un premier rendez-vous était fixé soit à leur domicile, soit dans les locaux de faculté de psychologie. Au cours de cette entrevue, une lettre d'information, ainsi qu'un formulaire de consentement éclairé, étaient présentés à l'enfant et à ses parents. Ces derniers remplissaient également un questionnaire de renseignements visant à préciser la présence ou l'absence de critères d'exclusion incompatibles avec les modalités de l'entraînement (i.e., déficit sensoriel, intellectuel ou neurologique ; maîtrise insuffisante du français). Le cas échéant, la doctorante ou une étudiante en orthophonie formée à la passation des tests réalisait l'évaluation pré-test (cf. § 7.2.2.).

Après avoir passé l'intégralité des épreuves du pré-test, le participant était assigné aléatoirement soit au groupe d'entraînement explicite soit au groupe d'entraînement implicite (cf. § 7.2.3). Toutefois, si le sujet présentait un score parfait ou quasi parfait (25 ou plus) au

test expérimental de production orthographique, il n'était pas inclus dans le protocole. Nous avons recruté des participants pendant deux années consécutives, sur la période s'étirant entre septembre et mars, à la fois en 2020-2021 et en 2021-2022. Nos participants ayant été recrutés « au fil de l'eau », les moments d'évaluation pré-test, d'entraînement et d'évaluation post-test étaient différents pour chaque sujet. Afin de nous assurer que les deux groupes d'entraînement ne différaient pas significativement sur l'ensemble des variables contrôles, à mesure que de nouveaux participants étaient recrutés, nous réalisions des ANOVA à un facteur sur les données pré-test, et attribuions ainsi les sujets nouvellement évalués à l'un ou l'autre groupe d'entraînement, de sorte à maintenir cet équilibre au fur et à mesure des inclusions.

Nous avons utilisé les normes relatives aux deux indices CM et CTL du test de l'Alouette-R (Lefavrais, 2005) pour déterminer si un enfant pouvait être considéré comme présentant un trouble de la lecture. Parmi les 54 enfants ayant réalisé l'intégralité de l'entraînement, 13 (dont 4 avec un diagnostic formel de dyslexie) présentaient un score significativement en-dessous du seuil (score Z inférieur ou égal à -1.5 écarts-types) correspondant à leur niveau de classe. Vingt-sept enfants ont été intégrés au groupe entraîné explicitement (EE) et 27 autres au groupe entraînement implicitement (EI). Quarante-cinq des 54 participants ont réalisé des mesures répétées, dont 23 dans le groupe EE et 22 dans le groupe EI. Les caractéristiques des deux groupes d'entraînement au pré-test sont détaillées dans le Tableau 7.

Tableau 7. Moyennes et écarts-types des deux groupes d'entraînement au pré-test sur les variables contrôles

	EE (n = 27)			EI (n = 27)		
	2 ^{ème} (n = 14)	4 ^{ème} (n = 9)	6 ^{ème} (n = 4)	2 ^{ème} (n = 12)	4 ^{ème} (n = 10)	6 ^{ème} (n = 5)
Âge (années)	7.25 (0.30)	9.29 (0.41)	11.2 (0.25)	7.20 (0.26)	9.23 (0.18)	11.3 (0.22)
Diagnostic de dyslexie (n)	0	2	0	0	0	2
Trouble de la lecture (n)	3	4	1	1	4	3
Habiletés cognitives générales						
MPC (/36)	27.1 (4.39)	29.4 (3.47)	33.0 (1.15)	27.6 (3.40)	30.9 (3.14)	30.8 (2.77)

Empan endroit (/8)		5.00 (0.88)	5.22 (0.97)	6.00 (0.82)	4.67 (0.99)	5.50 (0.71)	6.40 (1.14)
Empan envers (/8)		3.57 (1.02)	3.56 (1.01)	4.75 (1.50)	3.42 (0.90)	3.60 (1.17)	5.00 (2.00)
Identification des mots							
Alouette-R	CM (%)	83.8 (7.92)	91.6 (4.99)	94.4 (3.00)	85.4 (5.86)	88.5 (7.54)	92.2 (6.22)
	CTL (mots lus en 3 mn)	134 (59)	250 (112)	363 (142)	153 (36)	246 (85)	353 (129)
Production orthographique							
ODÉDYS	REG S (/10)	5.21 (2.61)	8.33 (1.87)	9.00 (0.82)	5.58 (1.62)	8.40 (2.22)	9.20 (0.84)
	IRR (/10)	0.79 (1.05)	4.67 (2.92)	6.50 (1.73)	0.58 (0.67)	4.60 (2.80)	7.80 (1.79)
	NM B (/10)	7.43 (1.28)	8.67 (1.22)	9.25 (0.96)	6.92 (1.98)	7.90 (1.60)	9.40 (0.55)
Vocabulaire							
EVIP (/170)		105 (17.4)	125 (14.3)	133 (10.9)	106 (9.75)	122 (13.2)	142 (10.6)
Conscience phonémique							
Evalec	Items corrects (%)	69.6 (20.0)	74.1 (12.8)	83.3 (13.6)	66.0 (16.8)	71.7 (12.6)	75.0 (20.4)
	Temps de réponse items corrects (ms)	5099 (1731)	3570 (905)	2593 (549)	4232 (993)	3597 (1611)	3037 (591)
Empan visuo-attentionnel							
Indice global (/100)		65.0 (18.2)	75.0 (13.4)	85.0 (13.1)	70.9 (12.4)	78.7 (11.0)	78.5 (7.42)
Indice partiel (/50)		30.7 (9.03)	38.8 (4.79)	44.0 (6.06)	29.2 (9.13)	37.3 (4.58)	35.8 (4.35)
Compétences orthographiques							
Choix orthographique	Sublex OP1 (/28)	21.0 (3.37)	25.6 (2.35)	25.8 (1.71)	21.3 (7.00)	24.7 (3.53)	26.2 (1.92)
	Sublex OP2 (/20)	12.9 (2.96)	15.1 (2.85)	16.8 (2.36)	13.3 (3.06)	15.1 (2.64)	16.8 (1.92)
	Lex OP1-HF (/26)	15.2 (3.04)	18.9 (3.95)	20.8 (3.59)	14.7 (2.87)	20.1 (3.63)	21.0 (3.00)
	Lex OP1-BF (/26)	13.2 (2.61)	16.6 (3.54)	19.8 (1.50)	13.0 (3.22)	17.2 (3.16)	20.4 (3.91)
	Lex OP 2 (/50)	27.9 (3.87)	37.4 (4.88)	40.5 (1.73)	28.3 (3.52)	37.7 (7.29)	40.2 (7.19)

CM = Indice de précision de lecture. CTL = Indice de fluence de lecture. EE = Groupe entraîné explicitement. EI = Groupe entraîné implicitement. EVIP = Échelle de Vocabulaire en Images Peabody. IRR = Mots irréguliers. Lex OP 1-BF = Patterns orthographiques lexicaux de basse fréquence. Lex OP 1-HF = Patterns orthographiques lexicaux de haute fréquence. Lex OP 2 = Patterns orthographiques lexicaux de même fréquence. MPC = Matrices Progressives Colorées. NM B = Non-mots bisyllabiques. ODÉDYS = Outil de Dépistage des Dyslexies. REG S = Mots réguliers simples. Sublex OP 1 = Patterns orthographiques sublexicaux légaux versus illégaux. Sublex OP 2 = Patterns orthographiques sublexicaux fréquents versus peu fréquents.

Bien que 54 participants aient réalisé entièrement l'entraînement proposé, 3 d'entre eux (2 apprenants de 6^{ème} année d'instruction et 1 participante de 4^{ème} année d'instruction)

n'ont pas donné suite aux sollicitations de prise de rendez-vous pour réaliser les évaluations post-tests. Les ANOVA à mesures répétées n'ont donc été menées que sur 25 participants du groupe EE et 26 du groupe EI.

7.2.2 Protocole d'évaluation

7.2.2.1 Procédure

Les évaluations pré- et post-tests ont été réalisées au domicile des participants ou dans les locaux de la faculté de psychologie, selon la préférence et les disponibilités des parents, toujours dans une pièce calme et bien éclairée. Les pré-tests étaient réalisés soit par la doctorante, soit par une étudiante en orthophonie formée à la passation des tests. Étant donné que la doctorante administrait l'intégralité des entraînements, les post-tests ont tous été réalisés soit par des étudiantes en orthophonie, soit par des étudiants en psychologie. Les passations pré-tests duraient en moyenne 1h pour les enfants scolarisés en 6^{ème} année d'instruction (i.e., 1^{ère} année de collège ou 6^{ème}), 1h30 pour les enfants scolarisés en 4^{ème} année d'instruction (i.e., 4^{ème} année de primaire ou CM1) et 2h pour les enfants scolarisés en 2^{ème} année d'instruction (i.e., 2^{ème} année de primaire ou CE1). Pour ces derniers en particulier, mais également pour tous les enfants en grande difficulté de lecture, la passation pouvait être séquencée en plusieurs sessions afin de ne pas saturer les capacités de concentration des enfants.

7.2.2.2 Variables contrôles

Nos groupes EE et EI étaient appariés sur des variables considérées comme pouvant impacter la réussite de nos participants à l'entraînement. Certaines d'entre elles sont dénommées variables « contrôles », notamment les Matrices Progressives Colorées (Raven et al., 1990) évaluant le raisonnement non-verbal, la forme A de l'Échelle de Vocabulaire en Images Peabody (EVIP, Dunn et al., 1993) évaluant le vocabulaire réceptif, et les empan de

chiffres endroit et envers de l'ODÉDYS (Jacquier-Roux et al., 2005) évaluant respectivement la mémoire à court terme et la mémoire de travail verbales. L'appariement sur ces variables visait à s'assurer que les participants de nos deux groupes d'entraînement ne se distinguaient pas significativement au niveau de leur fonctionnement cognitif général.

7.2.2.2 Variables langagières

D'autres variables faisaient également l'objet d'un appariement lors du pré-test, et étaient considérées comme susceptibles de conditionner la réussite à l'entraînement en raison de leur lien avec le développement de la littératie. Ces variables étaient dénommées variables « langagières. » La tâche de conscience phonémique de l'Evalec (Sprenger-Charolles et al., 2005), identique à celle utilisée dans le cadre de l'étude d'entraînement morphologique (cf. § 6.2.1.2), consistait en la suppression du phonème initial d'un non-mot monosyllabique de structure CCV (e.g., /gRu/ - /Ru/). L'empan visuo-attentionnel était évalué par les tâches informatisées de reports global et partiel (Bosse et al., 2007). Dans la tâche de report global, après l'apparition d'une croix de fixation pendant 500 ms, le sujet voit apparaître sur l'écran une suite de 5 consonnes majuscules (e.g., P S F H T) parmi 10 possibles (B P T F L M D S R H) pendant 200 ms. Le sujet doit, après leur disparition, dénommer l'ensemble des lettres perçues. Chaque lettre apparaissant 20 fois dans 5 positions différentes, cette tâche donne lieu à un score sur 100. Dans la seconde tâche, après la disparition des 5 lettres, un trait vertical apparaît en-dessous de l'une des 5 positions de lettres pendant 50 ms ; la consigne est alors de dénommer uniquement la lettre correspondant à cette position. L'épreuve de dictée de l'ODÉDYS évalue les compétences de production orthographique générales via la transcription de mots réguliers (e.g., FRITE), de mots irréguliers (e.g., MONSIEUR) et de non-mots bisyllabiques (e.g., *POULEN). Enfin, deux tâches de choix orthographique élaborées par Commissaire et Besse (2019) évaluaient respectivement les compétences orthographiques sublexicales et lexicales. La première comprenait les subtests Sublex OP 1

(choix entre deux non-mots homophones se distinguant par la légalité de leurs patterns sublexicaux ; e.g., *ROUVE – *ROUVVE) et Sublex OP 2 (choix entre deux non-mots homophones se distinguant par la fréquence de leurs patterns sublexicaux ; e.g., *DUIRE – *DUIRRE). La seconde comprenait les subtests Lex OP 1-HF (choix entre un mot au pattern sublexical fréquent et un leurre homophone au pattern sublexical peu fréquent ; e.g., FRAISE – *FREIZE), Lex OP 1-BF (choix entre un mot au pattern sublexical peu fréquent et un leurre homophone au pattern sublexical fréquent ; e.g., TREIZE – *TRAISE), et Lex OP 2 (choix entre un mot et un leurre homophone au pattern sublexical de même fréquence ; e.g., GRAINE – *GREINE).

Afin d'estimer l'effet spécifique de l'intervention sur le développement des compétences orthographiques, nous avons également proposé la tâche de choix orthographique Lex OP 2 en post-test. L'entraînement n'ayant pas porté sur les mots présentés dans cette tâche, une amélioration à cette dernière entre ces deux points de mesure constituerait un argument en faveur d'un effet non spécifique de l'intervention (i.e., effet test / re-test, Hawthorne ou placebo), ou d'un simple effet développemental.

7.2.2.3 Variables d'intérêt

L'efficacité de l'entraînement a été évaluée via deux types de variables d'intérêt : les mesures pré- et post-tests, proposées avant le commencement de l'entraînement et à la fin de ce dernier ; et les mesures répétées, proposées à chaque séance, tout au long de l'entraînement.

Mesures pré- et post-tests. Nous inscrivant en grande partie dans la continuité de l'étude d'Ise et Schulte-Körne (2010), nous avons choisi d'évaluer l'impact de l'entraînement sur l'identification des mots, via les indices CM et CTL de l'Alouette-R (Lefavrais, 2005). Bien que ce texte ne soit pas en majorité constitué de mots présentant une terminaison en /o/, nous avons souhaité évaluer l'effet d'un entraînement orthographique sublexical (EOS) sur la

précision et la fluence de lecture. En effet, le caractère méta-linguistique de notre EOS, qui abordait des notions aussi bien phonologique, morphologique que graphotactique, était susceptible de bénéficier aux capacités d'identification des mots de nos participants, d'une part en induisant une prise de conscience sur l'écrit de manière générale, et d'autre part du fait des activités proposées, qui engageaient leurs compétences de lecture et de production orthographique (cf. 7.2.3.2). L'indice CM (Corrects / Mots) est obtenu en divisant le nombre de mots correctement lus par le nombre de mots lus au total, puis en multipliant ce résultat par 100, résultant en un pourcentage de mots correctement lus. L'indice CTL (Corrects / Temps de Lecture) est obtenu en divisant le nombre de mots correctement lus par le temps de lecture exprimé en secondes, puis en multipliant ce résultat par 180. Le test consistant en la lecture d'un texte asémantique de 265 mots, l'indice CTL estime le nombre de mots qui auraient pu être lus par l'individu dans le temps imparti (180 secondes) si le test comprenait davantage de mots, permettant ainsi d'éviter une saturation du score à 265 mots pour les meilleurs lecteurs.

L'impact de l'entraînement sur la production orthographique était évalué via la tâche expérimentale Dictée du Phonème final O (DIPHO), consistant en une épreuve de production orthographique sous dictée de 22 mots finissant tous par le phonème O et dont la transcription de la graphie correspondante obéissait à une « règle » travaillée durant l'entraînement. Douze des mots de la DIPHO étaient abordés au cours de l'entraînement pour illustrer ces règles aux participants, tandis que les 10 autres n'étaient pas travaillés mais présentaient également des graphies finales à l'orthographe déductible en suivant ces mêmes « règles ». Deux sous-scores ont ainsi été élaborés, selon que les mots étaient entraînés (ME) ou non-entraînés (NE). Les ME étaient composés des 12 items vus en entraînement ($\alpha = .68$) et les NE des 10 items non travaillés ($\alpha = .76$). La liste complète des items de la DIPHO est présentée en Annexe 7. Un point était accordé par terminaison correctement transcrite, sans égard pour le reste du mot.

Au pré-test, il n'existait pas de différence significative entre nos deux groupes d'entraînement sur aucune des variables d'intérêt.

Mesures répétées. Quarante-cinq³⁵ des 54 participants ont également été évalués à la fin de chaque séance d'entraînement via une dictée de mots non-entraînés dénommée MEREP (dictée Mesures Répétées). Trois versions (A, B, C) de cette même dictée étaient présentées en alternance à chaque séance, dans un ordre aléatoirement défini pour chaque sujet, afin de limiter un effet d'apprentissage trop important. Ce sont les scores obtenus à cette dictée qui ont permis de réaliser les analyses en *Single-Case Design* (SCD) détaillées dans le point 7.2.5.2. La MEREP était constituée de 10 items, propres à chaque variante A, B ou C : 3 mots « EAU », 3 mots « AUX », 2 mots « O apocope » et 2 mots « O mot d'emprunt ». Un point était attribué pour chaque terminaison correctement transcrite, sans égard pour le reste du mot ; le score maximum pour chaque dictée réalisée était donc de 10. Il est à noter qu'aucun des mots de la MEREP n'était commun avec ceux de la DIPHO.

Chaque variante de la MEREP était appariée via la base de données Manulex-infra (Peereman et al., 2007) sur la fréquence d'occurrence ($\chi^2(2) = 0.15, p = .93$), le nombre de lettres ($\chi^2(2) = .24, p = .89$), le nombre de syllabes ($\chi^2(2) = 1.52, p = .47$), le nombre de phonèmes ($\chi^2(2) = 1.68, p = .43$), le nombre de graphèmes³⁶ ($\chi^2(2) = 1.30, p = .52$), la fréquence des associations graphème-phonème *token count* ($\chi^2(2) = 0.24, p = .89$), la consistance des associations graphème-phonème *token count* ($\chi^2(2) = 0.81, p = .67$) et la consistance des associations phonème-graphème *token count* ($\chi^2(2) = 0.84, p = .67$). Le détail des items de la MEREP est présenté dans l'Annexe 8. À l'issue de la première séance

³⁵ Neuf des 54 participants, tous scolarisés en 2^{ème} année d'instruction, étaient trop fatigables pour se voir proposer la dictée MEREP en fin de séance.

³⁶ Bien que le graphème soit souvent envisagé comme la contrepartie écrite du phonème (le nombre de graphèmes d'un mot devant ainsi être identique à son nombre de phonèmes), plusieurs auteurs ont souligné de nombreuses exceptions à cette règle (Catach, 1979; Lockwood, 2000). Dans le cadre des mots de la MEREP, ces exceptions concernaient la graphie AUX considérée dans Manulex-infra comme donnant lieu à deux graphèmes (AU+X), la graphie OI considérée comme correspondant à deux phonèmes (/w/ + /a/), et la graphie EN de FLAMENCO, considérée comme un seul graphème donnant lieu à deux phonèmes (/ε/ + /n/).

(premier point de la phase A), après réalisation d'une ANOVA unidirectionnelle sur rangs de Kruskal-Wallis, il n'existait pas de différence significative de performances entre les participants aux trois variantes de la MEREP pour les enfants de 2^{ème} ($\chi^2(2) = 2.85, p = .24$), 4^{ème} ($\chi^2(2) = 2.64, p = .27$) et 6^{ème} année d'instruction ($\chi^2(2) = 5.14, p = .08$) ; toutefois, la valeur du p de la comparaison pair à pair Dwass, Steel, Critchlow et Fligner entre la dictée B et C chez les enfants de 6^{ème} année d'instruction était proche du seuil de tendance ($W = 2.76, p = .12$), suggérant que la dictée C était en moyenne marginalement mieux réussie que la dictée B (cf. Tableau 8). Cette différence non-significative présente à l'issue de la première séance de la phase A sera évoquée lors de la Discussion (§ 7.3).

Tableau 8. Moyennes, écarts-types et comparaison des scores obtenus à la MEREP à l'issue de la séance 1

Variante proposée lors de la séance 1	Moyenne (ET)		
	2 ^{ème}	4 ^{ème}	6 ^{ème}
Dictée A	2.14 (1.07) n = 7	5.67 (1.03) n = 6	7.00 (1.41) n = 2
Dictée B	2.86 (1.57) n = 7	4.50 (2.51) n = 6	4.00 (2.83) n = 2
Dictée C	3.67 (1.53) n = 3	5.57 (2.23) n = 7	8.60 (1.14) n = 5
Comparaisons pair à pair DCSF			
Dictée A			
	2 ^{ème}	4 ^{ème}	6 ^{ème}
versus dictée B	$W = 1.67$ $p = .47$	$W = -2.07$ $p = .31$	$W = -1.73$ $p = .44$
versus dictée C	$W = 2.18$ $p = .27$	$W = -0.94$ $p = .78$	$W = 1.95$ $p = .35$
B			
	2 ^{ème}	4 ^{ème}	6 ^{ème}
versus dictée C	$W = .99$ $p = .76$	$W = 1.65$ $p = .47$	$W = 2.76$ $p = .12$

DCSF = Dwass, Steel, Critchlow et Fligner.

7.2.3 Protocole d'entraînement

7.2.3.1 Procédure

Notre entraînement orthographique sublexical portait sur la transcription du phonème final /o/ et consistait en 10 sessions d'entraînement de 30 minutes chacune, à raison de deux séances par semaine pendant 5 semaines, pour une durée totale d'entraînement d'environ 5

heures par enfant. Il intègre des éléments phonologiques, morphologiques et graphotactiques (voir Tableau 9). Les séances étaient individuelles et entièrement menées par la doctorante, au domicile des participants ou dans les locaux de la faculté de psychologie, en fonction des préférences des parents de l'enfant. De même que pour l'entraînement morphologique présenté dans le précédent chapitre, toutes les séances étaient scénarisées afin de permettre à la doctorante de délivrer un entraînement le plus similaire possible à tous les participants. Un exemple de scénario est présenté dans l'Annexe 9.

Tableau 9. Description synthétique des séances et de la prédominance de la dimension phonologique, morphologique ou graphotactique

Numéro de séance	Compétence centrale travaillée durant la séance	Prédominance
1	- Identification du phonème final /o/ - Listing des graphies possibles pour le phonème final /o/	Phonologique / Graphotactique
2	Comptage et segmentation syllabique	Phonologique
3	Inférence de la lettre muette par la morphologie	Morphologique
4	Règle de morphologie flexionnelle (AUX)	Morphologique
5	Règle de morphologie dérivationnelle (O apocope)	Morphologique
6	Règle d'orthographe étymologique (O mot d'emprunt)	Morphologique
7	Terminaison fréquente (EAU)	Graphotactique
8	Terminaisons rares (OS, OT, AU, AUD, AUT)	Graphotactique
9	Application de l'algorithme	/
10	Application de l'algorithme	/

NB : Les trois premières séances (grisées) sont considérées comme des prérequis et constituent, dans l'analyse SCD, les trois points de mesure de la phase de ligne de base.

Les participants suivaient deux variantes du même entraînement, détaillé dans le paragraphe suivant. Dans l'entraînement dit « explicite » (EE), l'objectif de chaque séance était annoncé et les régularités orthographiques travaillées étaient présentées comme des règles à retenir et à utiliser pour pouvoir correctement écrire la terminaison /o/ à la fin des mots. Dans l'entraînement dit « implicite » (EI), bien que l'essentiel du matériel soit commun avec le groupe EE, les exercices à réaliser n'étaient jamais explicitement mis en lien avec le travail d'une compétence ou d'une règle, et l'enfant devait seul en comprendre la raison sous-jacente, sans indication de la part de la doctorante. Ces deux groupes d'entraînement se basaient ainsi sur des logiques d'apprentissage différentes. Dans l'EE, nous suivions les

principes de l'enseignement explicite, selon lesquels une approche présentant une structure et des objectifs clairs permet au plus grand nombre d'élèves d'apprendre efficacement (Hughes et al., 2017). Dans l'EI, le postulat était que les unités orthographiques ciblées par le matériel d'entraînement seraient perçues comme des régularités plutôt que comme des « règles ». L'apprentissage pour les participants se ferait par simple exposition à un matériel sélectionné pour ses propriétés linguistiques, et serait alors incident, c'est-à-dire qu'il ne ferait pas (ou dans une moindre mesure) l'objet d'une verbalisation par l'enfant lui-même.

7.2.3.2 Contenu des séances

Les trois premières séances d'entraînement visaient à travailler les prérequis indispensables à la compréhension de l'algorithme et ne portaient pas sur les régularités orthographiques en elles-mêmes. La première séance se centrait sur deux compétences : l'identification auditivo-verbale du phonème /o/ à la fin des mots et l'identification visuo-orthographique de la graphie finale utilisée pour transcrire ce phonème. Dans ce but, une liste de mots finissant par le phonème /o/ était présentée à l'enfant. Après avoir colorié la terminaison correspondant à ce phonème pour l'ensemble des mots présentés, la liste était retirée et l'enfant devait réécrire de mémoire les différentes terminaisons vues dans l'exercice. Cette séance avait ainsi également pour but d'amener l'enfant à distinguer les graphies légales de ce phonème final en français (e.g., -OS ou -OT sont légales tandis que -OE ne l'est pas). Au cours de cette première séance, l'EE et l'EI se différenciaient essentiellement par le matériel utilisé, les participants de l'EE ayant la possibilité de colorier la terminaison en /o/ des mots de la liste en ayant sous les yeux un inventaire des graphies existantes pour transcrire ce phonème en fin de mot (cf. Annexe 10), contrairement aux participants de l'EI.

La deuxième séance abordait la notion de syllabe, indispensable pour la première étape de l'algorithme décisionnel (i.e., « Est-ce que c'est un mot avec au moins 2 syllabes ? »). Après avoir donné une brève définition de la syllabe en tant qu'unité naturelle

d'articulation (i.e., un morceau de mot qu'on arrive à dire en une seule fois), l'enfant était invité à déterminer le nombre de syllabes dans son prénom puis dans celui de la doctorante. Ensuite, trois étiquettes comportant respectivement les nombres 1, 2 et 3 étaient placées devant l'enfant. Les mots travaillés lors de la séance précédente étaient également présentés sous forme d'étiquettes, et l'enfant devait les placer sous le chiffre correspondant au nombre de syllabes dans le mot. Lorsque tous les mots étaient correctement classés, ils étaient retirés de la vue de l'enfant et la doctorante piochait deux mots d'une, 2 et 3 syllabes, que l'enfant devait ensuite réécrire sous dictée. Les participants du groupe EE étaient invités à recompter le nombre de syllabes du mot avant de l'écrire, tandis qu'aucune instruction supplémentaire n'était donnée aux participants du groupe EI.

La troisième séance se concentrait sur les mots monosyllabiques comportant le phonème final /o/. Il était expliqué aux participants du groupe EE que les mots d'une seule syllabe constituaient une exception aux règles d'orthographe qui seraient travaillées dans les séances suivantes, tandis que cette information n'était pas divulguée aux participants du groupe EI. Deux listes de 4 mots monosyllabiques étaient présentées aux participants, regroupés selon qu'ils présentaient une lettre muette inférable par la morphologie (e.g., GROS) ou non (e.g., TÔT). Il était demandé à l'enfant de colorier la terminaison en /o/ de ces mots. Aux participants du groupe EE, il était expliqué que, puisque ces mots ne respectaient pas les règles d'orthographe, il n'y avait d'autre choix que de retenir par cœur l'orthographe du mot. La possibilité de recourir à la famille morphologique du mot était alors proposée comme stratégie à l'enfant, qui restait libre de s'en servir ou non. Les deux listes étaient ensuite cachées, puis l'on demandait à l'enfant de réécrire sous dictée les 8 mots précédemment vus. Les participants du groupe EI ne se voyaient proposer aucune stratégie particulière et devaient simplement transcrire les mots dictés après avoir colorié leur

terminaison. Dans les deux groupes d'entraînement, un feedback correctif immédiat était fourni si la production était incorrecte.

Les séances 4 à 7 travaillaient chacune une graphie différente du phonème final /o/ : les graphies AUX, O (dans le cas des mots apocopes, puis dans le cas des mots empruntés à une langue étrangère), et EAU. Dans la séance 4, une liste de mots terminant par la graphie -AUX était présentée, au regard d'une liste de ces mêmes mots au singulier et portant donc la terminaison -AL ou -AIL. Il était demandé à l'enfant de relier correctement les mots au pluriel à leur équivalent singulier. Puis, les deux listes étaient cachées et l'enfant devait réécrire les mots au pluriel sous dictée. Aux participants de l'EE, il était ainsi expliqué la seconde étape de l'algorithme, correspondant à la première règle orthographique : « si le phonème /o/ à la fin d'un mot peut être remplacé par /al/ ou /aj/ à l'oral, alors ce mot *doit* s'écrire -AUX ». Aux participants de l'EI, aucune règle n'était présentée, et l'exercice consistait simplement à relier les singuliers et les pluriels, puis à réécrire correctement les pluriels sous dictée.

Les séances 5 et 6 abordaient pour les participants de l'EE la troisième étape de l'algorithme, visant à déterminer si le mot cible doit s'écrire avec un -O ou non, selon qu'il obéisse à l'une des deux règles vues individuellement. Dans la séance 5, des mots faisant l'objet d'apocopes en français (e.g., VÉLOCIPÈDE) étaient présentés à l'enfant, qui devait essayer de déterminer « le plus petit mot dans le grand mot ». Les mots apocopes étaient ensuite dictés. Aux participants de l'EE, il était ainsi explicité la deuxième règle orthographique : « s'il existe un mot plus grand que le mot finissant par le phonème /o/, alors ce mot *doit* s'écrire -O ». Aux participants de l'EI, aucune règle n'était présentée, et ils devaient simplement écrire sous dictée les apocopes des mots présentés dans la première partie de l'exercice.

La séance 6 commençait avec la présentation d'une feuille comportant 7 images de mots étrangers (cf. Annexe 11), présentés comme tels aux participants du groupe EE. La

deuxième règle de l'étape 3 de l'algorithme était ainsi introduite : « si tu sais que le mot finissant par le phonème /o/ est un mot étranger, alors il *doit* s'écrire -O ». Aux participants du groupe EI, il était simplement expliqué que ces mots étaient étrangers et qu'ils s'écrivaient avec la terminaison -O, mais sans présenter ce fait comme une règle. Dans les deux groupes d'entraînement, les participants devaient ensuite dénommer les mots illustrés et les transcrire.

La séance 7 introduisait la dernière terminaison du phonème /o/ dont la transcription correspondait à une « règle » de notre algorithme. Aux participants de l'EE, les trois règles précédemment vues étaient rappelées, avant de présenter la dernière : « si le mot finissant par le phonème /o/ n'entre dans aucun des cas de figure précédents, alors il *doit* s'écrire -EAU dans la majorité des cas ». Il était ensuite précisé que la terminaison -EAU étant la plus fréquente en français pour les mots d'au moins deux syllabes finissant par le phonème /o/, si le mot à écrire n'était ni un pluriel d'un mot en -al ou -aj, ni un mot « raccourci », ni un mot étranger, alors il était très probable que ce mot s'écrive -EAU. L'exercice de la séance consistait à illustrer cette règle en proposant à l'enfant 8 « devinettes » de mots ayant en commun le fait de comporter 2 syllabes et de finir par -EAU (cf. Annexe 12). Le participant devait lire la devinette et essayer d'inférer de quel mot il s'agissait sur la base des indices présentés. Pour les participants de l'EI, la séance commençait avec la lecture des devinettes et aucune règle n'était abordée. Après le jeu de devinettes, l'enfant était invité à transcrire les 8 mots sous dictée. Dans le cas du groupe EE, les participants avaient la consigne d'essayer d'appliquer les règles précédemment vues sur chacun des mots, afin de voir qu'elles ne fonctionnaient pas, avant d'écrire la terminaison -EAU.

La huitième séance abordait le cas de mots présentant une terminaison moins fréquente que -EAU, mais n'entrant pas dans les cas de figure des règles présentées dans les séances 4 à 6. Aux participants de l'EE, il était expliqué que certains mots dits « exception » *devraient* s'écrire -EAU, mais possédaient encore une autre transcription, et que ces mots

n'obéissant pas aux règles abordées, il n'y avait d'autre choix que de retenir l'orthographe de ces mots par cœur. Pour les participants de l'EI, la séance commençait directement sur la seconde partie de la séance, sans qu'aucune explication particulière ne soit fournie : le jeu de famille des exceptions. Trois mots comportant la même graphie « exception » (-OT, -OS, -AU-, -AUD et -AUT) étaient présentés à l'enfant sous forme de cartes (cf. Annexe 13), puis ces cartes étaient utilisées dans le cadre d'un jeu sur la base du jeu des 7 familles. Pendant le jeu, il était précisé aux participants de l'EE qu'il leur serait demandé de réécrire un mot de chaque famille, afin de les inciter à prêter attention à la terminaison de chaque mot. Les participants de l'EI n'étaient pas prévenus et devaient simplement jouer au jeu. À l'issue du jeu, cinq mots (un pour chaque graphie « exception ») étaient dictés à l'enfant, qui devait les retranscrire.

Les séances 9 et 10 étaient dédiées, dans l'entraînement EE, à l'utilisation de l'algorithme (cf. Figure 9). Ce dernier était tout d'abord présenté comme la synthèse de toutes les règles abordées jusqu'à présent et chaque étape était lue avec l'enfant, puis la doctorante dictait 12 mots abordés dans les séances précédentes et l'enfant était invité à initier une verbalisation suivant les étapes de l'algorithme afin de déduire la terminaison qu'il convenait d'utiliser pour chaque mot, sous la supervision de la doctorante. Concernant les participants de l'EI, il leur était simplement annoncé qu'ils devaient écrire sous dictée des mots qu'ils avaient déjà écrits auparavant, lors des séances précédentes. Ainsi, les participants de l'EI ne recevaient pas d'indication sur la manière dont il fallait écrire la terminaison des mots dictés ; toutefois, un feedback correctif immédiat leur était dispensé lorsque la production était incorrecte.

Quelle terminaison choisir pour le son [o] à la fin d'un mot ?

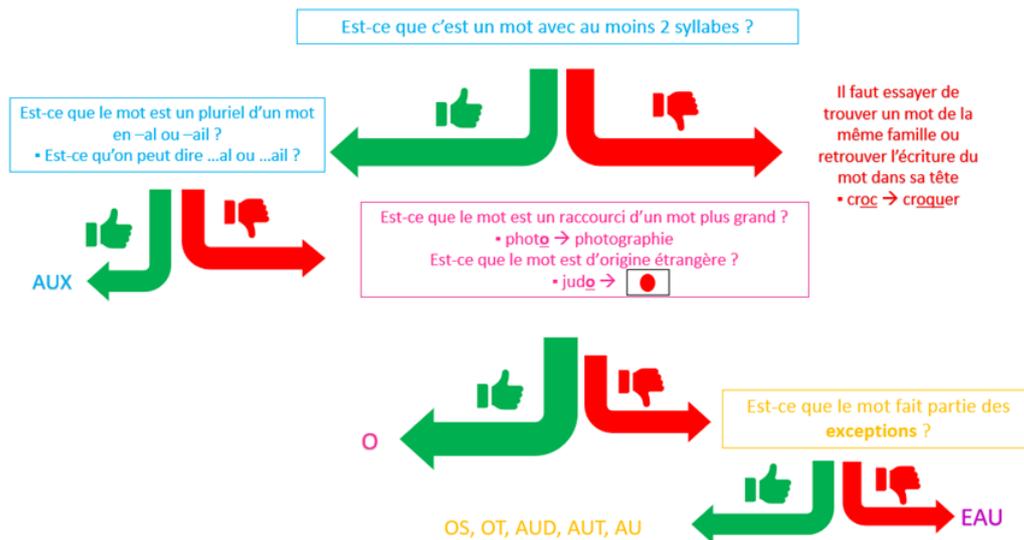


Figure 9. Algorithmes pour la transcription du phonème /o/ en position finale

7.2.4 Design

L'efficacité de l'entraînement et de ses modalités était évaluée selon deux *designs* différents et complémentaires. Nous avons à la fois eu recours à un *design* « classique » en pré-/post-test, semblable à celui utilisé lors de la précédente étude, présentée dans le chapitre 6, ainsi qu'à un *design* en lignes de bases multiples (*Multiple Baseline Design*, MBD). Ce dernier nous permettait d'estimer les progrès de chaque participant sur nos mesures répétées, avant (phase A) et pendant (phase B) l'acquisition des « règles » de transcription de la terminaison /o/.

Parmi les *Single-Case Design* (SCD), le *design* MBD est le plus fréquemment utilisé dans les études d'entraînement en psychologie ou en sciences de l'éducation, car il est le seul qui permette de tenir compte de l'irréversibilité de l'apprentissage (Petersen-Brown et al., 2012). Ce *design* nécessite la répétition d'un effet de l'intervention à travers plusieurs lignes de base ou phases A. Dans le cadre de cette étude, nous avons choisi de tester la répétition de cet effet à travers nos participants (Kazdin, 2011). Notre intervention s'étalait sur 10 séances, les trois premières séances ciblant le travail de prérequis à l'acquisition de l'algorithme (cf. §

7.2.3). Ainsi, pour chacun de nos sujets, notre phase A était constituée de 3 points de mesure (i.e., les trois premières séances), et notre phase B de 7 points de mesure (i.e., les sept séances restantes), excepté pour les sujets 25 (EE), 32 (EI) et 34 (EE), pour lesquels la phase B était respectivement constituée de 5, 5 et 3 points de mesure, suite à la perte des feuilles comprenant les points de mesure manquants.

Dans les MBD, le début de l'intervention doit être aléatorisé et échelonné pour chaque nouvelle ligne de base (Kratochwill et al., 2021), afin de minimiser des effets non-spécifiques à l'intervention ; par exemple, sans cet échelonnage, il ne serait pas possible de déterminer si une amélioration qui aurait lieu au terme d'un nombre fixe de séances est dû à l'intervention ou à un développement « naturel », inhérent à la compétence évaluée, ou bien encore à un effet placebo différé. Toutefois, un tel échelonnage n'a pas pu être réalisé dans le cadre de notre étude, du fait de notre recrutement « au fil de l'eau » et de notre *design* mixte qui, vis-à-vis du pré-/post-test, se prêtait difficilement à une variation de la durée de la ligne de base. Le commencement de chaque phase d'intervention était défini par avance et correspondait toujours à la 4^{ème} séance avec le participant.

7.2.5 Analyses statistiques

7.2.5.1 Mesures pré-/post-test

Concernant les mesures pré- et post-test, nous avons eu recours à des analyses de variance (ANOVA) à mesures répétées, ainsi qu'à des analyses post-hocs avec correction de Tukey. Elles visaient à déterminer la significativité des interactions entre le point de mesure (pré-test et post-test) et le groupe d'entraînement (explicite et implicite). Dans le cas de la DIPHO, nous avons également inclus le paramètre entraîné ou non-entraîné des stimuli proposés en tant que variable intra supplémentaire dans l'ANOVA.

Étant donné que nous n'avions pas de groupe contrôle nous permettant de contrôler un potentiel effet développemental, nous avons choisi d'intégrer en tant que covariable le gain de

nos participants entre le pré- et le post-test à la tâche de choix orthographique Lex OP 2. Bien qu'il ne soit pas possible de considérer a priori que la performance à cette tâche évolue de la même manière que sur nos variables d'intérêt, notre objectif était ainsi de neutraliser une part de la variance attribuable à une progression due au temps.

Dans une perspective exploratoire, nous avons également réalisé des corrélations partielles (contrôlant l'âge développemental en mois) entre le gain entre le pré- et le post-test de nos variables d'intérêt (indices CM et CTL de l'Alouette, production orthographique de mots entraînés et non-entraînés), et nos variables contrôle suivantes : empan visuo-attentionnel (empan global et partiel), performances à la dictée de l'ODÉDYS (mots réguliers simples, mots irréguliers, non-mots) et tâches de choix orthographique (Sublex OP 1, Sublex OP 2, Lex OP 1-HF, Lex OP 1-BF, Lex OP 2). Ceci a été réalisé afin de d'estimer de potentiels prédicteurs de l'efficacité de l'entraînement, qu'il soit explicite ou implicite. Nous avons réalisé ces analyses pour les participants scolarisés en 2^{ème} et en 4^{ème} année d'instruction. En raison du faible nombre de sujets scolarisés en 6^{ème} année d'instruction, nous n'avons pas mené d'analyses sur ce sous-groupe d'élèves.

Une p-valeur inférieure ou égale à .05 obtenue sur les analyses de variance et sur les corrélations était considérée comme statistiquement significative, et celles comprises entre .06 et .10 étaient considérées comme tendanciellement significatives. Concernant les ANOVA à mesures répétées, les tailles d'effet de l'entraînement sont indiquées par l'éta carré partiel (η^2_p) dont les valeurs peuvent s'interpréter comme suit (Richardson, 2011) : .01 : effet faible ; .06 : effet médium ; .14 : effet fort. Concernant les coefficients de corrélation, nous avons suivi les recommandations de Mukaka (2012) pour l'interprétation des valeurs de r (le signe précédant la valeur détermine si la corrélation est positive ou négative) : entre .00 et .30 : corrélation négligeable ; entre .30 et .50 : corrélation faible ; entre .50 et .70 : corrélation

modérée ; entre .70 et .90 : corrélation élevée ; entre .90 et 1 : corrélation très élevée. Toutes les analyses ont été réalisées avec le logiciel Jamovi (The jamovi project, 2021).

7.2.5.2 Mesures répétées

Concernant les mesures répétées, les analyses réalisées étaient de trois types. Tout d'abord, ainsi qu'il en est d'usage pour les SCD (Kratochwill et al., 2010), nous avons préalablement réalisé une analyse visuelle descriptive en nous basant sur le niveau, la tendance et l'immédiateté de l'effet de l'intervention. Le niveau était évalué par le calcul de la médiane et la comparaison de cette valeur entre les phases A et B. La tendance de chaque phase était inférée par le tracé d'une *split-middle trend line* (SMTL ; Kazdin, 2011; White & Haring, 1980) : chaque phase est séparée en deux moitiés (pour un nombre impair de mesures, ce qui est notre cas, la donnée « du milieu » est ignorée), et la médiane est déterminée pour chacune de ces moitiés qui sont ensuite reliées par une ligne. L'introduction de l'intervention étant censée accélérer l'amélioration des performances d'un sujet, nous avons comparé, pour chaque phase, la pente (via le coefficient directeur) en phase A et celle en phase B. Enfin, l'immédiateté de l'effet de l'intervention était évaluée en établissant le niveau et la tendance de la phase A par rapport aux trois premiers points de la phase B. Sur la base de ces trois indicateurs, nous avons ensuite procédé à une analyse du pattern de progression pour l'ensemble des participants.

Suite à ces analyses visuelles descriptives, nous avons eu recours à un calcul de taille d'effet appropriée pour les SCD, via un indice de non-recouvrement appelé *Non-overlap of All Pairs* ou NAP (Parker & Vannest, 2009). Le calcul de cet indice non-paramétrique consiste à comparer chaque point de la phase A à chaque point de la phase B, afin de déterminer, pour chaque paire, l'existence d'un recouvrement (*overlap*), d'un non-recouvrement (*non-overlap*) ou d'un *ex-aequo* (*tie*). Un point est attribué pour chaque *overlap*, 0.5 point pour chaque *tie* et 0 point pour chaque *non-overlap*. L'ensemble de ces

« scores » pour chaque paire est ensuite additionné, puis soustrait au nombre total de paires possibles. Ce nombre représente ainsi le nombre de points en phase B ne recouvrant pas les points en phase A. Afin d'obtenir un pourcentage, ce nombre est enfin divisé par le nombre total de pair possible et multiplié par 100. Ainsi défini, le calcul du NAP est quasiment identique au calcul du U_L (i.e., le nombre de points d'une phase dont la valeur excède celle des points de l'autre phase) d'un test *Rank-Sum* de Wilcoxon, à la différence que dans le NAP, le U_L est ramené à la somme du nombre total de paires possibles (Parker & Vannest, 2009). De sorte, il est possible d'estimer pour cet indice une p-valeur, pour laquelle un seuil de significativité inférieur à .05, et un seuil de tendance compris entre .05 et .10, ont été retenus.

Sur la base de ce calcul, toute valeur de NAP inférieure à 50% doit ainsi évoquer un effet adverse. Une taille d'effet positive de l'intervention étant donc forcément comprise entre 50 et 100%, un ré-échelonnage de cette valeur a été proposé (Parker & Vannest, 2009) afin d'obtenir pour cet indice une valeur comprise entre 0 et 100% : une valeur de NAP entre 0 et 31% suggère un effet faible ; une valeur comprise entre 32 et 84% suggère un effet médium ; et une valeur comprise entre 85 et 100% indique un effet fort.

Le niveau (médiane) et la tendance (SMTL) ont été calculés à main levée. Le calcul du NAP a été réalisé avec le logiciel R (R Core Team, 2022), et plus précisément avec le *package* « scan » (Wilbert & Lueke, 2022).

7.3 Résultats

7.3.1 Résultats sur les mesures pré-/post-tests

Les performances en pré- et post-test de nos deux groupes entraînés aux variables de lecture, de production orthographique et de choix orthographique sont présentées dans le Tableau 10. Nous décrivons ci-dessous le résultat des analyses des variance (avec le gain à la

tâche de choix orthographique en covariable) pour les performances à l'Alouette-R et pour les performances à la DIPHO.

Tableau 10. Moyennes et écarts-types des deux groupes d'entraînement en pré- et post-test sur les variables d'intérêt et sur la variable contrôle

		Pré-test		Post-test	
		EE (n = 25)	EI (n = 26)	EE (n = 25)	EI (n = 26)
Alouette-R	CM (%)	87.9 (8.06)	87.7 (6.93)	91.7 (7.89)	92.2 (4.78)
	CTL (nombre de mots correctement lus en 3 mn)	203 (127)	220 (106)	230 (137)	234 (85)
DIPHO	ME (/12)	6.48 (2.22)	6.27 (2.16)	9.12 (2.30)	8.58 (2.34)
	NE (/10)	4.68 (2.10)	4.73 (2.32)	6.56 (2.57)	6.35 (2.40)
Lex OP 2 (/50)		32.4 (6.58)	33.9 (7.81)	33.2 (6.39)	35.3 (6.99)

NB : En raison de données manquantes au post-test pour 2 sujets du groupe EE et 1 sujet du groupe EI, les données correspondantes ont également été retirées des pré-tests.

CM = Indice de précision. CTL = Indice de fluence. DIPHO = Dictée Phonème final O. EE = Groupe entraîné explicitement. EI = Groupe entraîné implicitement. Lex OP 2 = Tâche de choix orthographique au pattern sublexical de même fréquence.

7.3.1.1 Alouette-R : ANOVA à mesures répétées

Un effet du point de mesure ($F(2,47) = 40.4, p < .001, \eta^2_p = .462$) semble indiquer que nos deux groupes ont progressé de manière significative entre le pré-test et le post-test sur l'indice CM ($t(47) = -6.87, p_{Tukey} < .001$). Un effet du point de mesure est également retrouvé pour l'indice CTL ($F(2,47) = 7.29, p = .01, \eta^2_p = .134$) en faveur d'une amélioration générale avec le temps pour nos deux groupes ($t(47) = -2.98, p_{Tukey} = .005$). En revanche, nous n'avons pas trouvé d'interaction significative entre le point de mesure et le groupe d'entraînement, ni pour l'indice CM ($F(2,47) < 1, p = .70, \eta^2_p = .003$), ni pour l'indice CTL ($F(2,47) = 1.21, p = .28, \eta^2_p = .025$), suggérant qu'aucun de nos groupes d'entraînement n'a davantage progressé que l'autre. Par ailleurs, nous n'avons pas retrouvé d'interaction significative entre le point de mesure et la covariable « gain à la tâche Lex OP 2 », que ce soit pour l'indice CM ($F(2,47) = 1.15, p = .29, \eta^2_p = .024$) comme pour l'indice CTL ($F(2,47) < 1, p = .51, \eta^2_p = .009$).

7.3.1.2 Alouette-R : corrélations entre le gain et les variables contrôle.

Le Tableau 11 présente le résultat des corrélations partielles réalisées entre le gain aux indices de précision et de fluence de l'Alouette-R et certaines variables contrôle susceptibles de conditionner la réussite à l'entraînement : empan visuo-attentionnel (empan global et partiel), performances à la dictée de l'ODÉDYS (mots réguliers simples, mots irréguliers, non-mots) et tâches de choix orthographique (Sublex OP 1, Sublex OP 2, Lex OP 1-HF, Lex OP 1-BF, Lex OP 2). Concernant les sujets de 2^{ème} année d'instruction, la quasi-totalité des corrélations réalisées donne lieu à des coefficients non-significatifs, excepté pour une corrélation positive faible entre le gain sur l'indice CM et la tâche de choix orthographique Lex OP 1 BF ($r = .44, p = .03$). Cette tâche donne également lieu à un coefficient de corrélation positif faible et tendanciellement significatif entre cette dernière et le gain sur l'indice CTL ($r = .43 p = .08$), pour les sujets de 4^{ème} année d'instruction. Le reste des coefficients de corrélation obtenus pour cette classe d'âge sont non-significatifs.

Tableau 11. Corrélations partielles entre le gain à l'Alouette-R et les variables pré-test

2^{ème} année d'instruction												
	Gain CM	Gain CTL	EVA-G	EVA-P	REG	IRR	NM	SOP1	SOP2	LOP1 HF	LOP1 BF	LOP2
Gain CM	-	-	.04	-.13	-.02	-.14	.17	-.01	.17	-.22	.44*	-.26
Gain CTL	-	-	.13	.03	-.21	-.02	.21	-.05	-.08	.09	.11	-.32
4^{ème} année d'instruction												
	Gain CM	Gain CTL	EVA-G	EVA-P	REG	IRR	NM	SOP1	SOP2	LOP1 HF	LOP1 BF	LOP2
Gain CM	-	-	.00	.39	-.25	-.33	-.30	-.23	-.06	-.08	-.14	-.37
Gain CTL	-	-	.12	-.06	.13	.21	.04	-.22	-.20	.24	.43 ^t	.18

^t = $p < .10$ - * = $p < .05$

NB : Les corrélations présentées contrôlent l'âge des participants en mois.

EVA-G = Empan visuo-attentionnel global. EVA-P = Empan visuo-attentionnel partiel. IRR = Dictée de mots irréguliers. NM = Dictée de non-mots. REG = Dictée de mots réguliers.

7.3.1.3 DIPHO : ANOVA à mesures répétées

Les scores entraînés (ME) et non-entraînés (NE) de la DIPHO ont été convertis en pourcentages avant de réaliser l'ANOVA à mesures répétées sur ces derniers. De même que pour les ANOVA réalisées sur l'Alouette, nous avons intégré le gain à la tâche Lex OP 2 à l'analyse en tant que covariable.

Nous avons retrouvé un effet du point de mesure ($F(2,47) = 44.5, p < .001, \eta^2_p = .486$) : d'une manière générale, nos participants semblent avoir progressé entre le pré- et le post-test à la DIPHO ($t(47) = -7.10, p_{Tukey} < .001$). Toutefois, nous ne retrouvons pas d'effet d'interaction entre le point de mesure et le groupe d'entraînement ($F(2,47) < 1, p = .54, \eta^2_p = .008$), suggérant qu'aucun de nos groupes d'entraînement n'a davantage progressé que l'autre. De plus, nous ne retrouvons pas d'interaction entre le point de mesure et le gain à la tâche Lex OP 2 ($F(2,47) < 1, p = .48, \eta^2_p = .011$). De la même manière, bien qu'un effet simple du type de mots soit retrouvé ($F(2,47) = 21.7, p < .001, \eta^2_p = .316$), indiquant, selon les analyses post-hocs, de meilleures performances en général sur les ME que sur les NE ($t(47) = 5.29, p_{Tukey} < .001$), les interactions entre le type de mots et le point de mesure ($F(2,47) = 2.64, p = .11, \eta^2_p = .053$), ainsi qu'entre le type de mots et le groupe d'entraînement ($F(2,47) < 1, p = .38, \eta^2_p = .017$) n'étaient pas significatives. En revanche, l'interaction entre le type de mots et le gain à la tâche Lex OP 2 était tendancielle ($F(2,47) = 2.93, p = .09, \eta^2_p = .059$) ; afin d'explorer de quelle manière ce gain impactait les scores aux ME et aux NE, nous avons réalisé des régressions linéaires entre le gain à la tâche Lex OP 2 et le gain à ces deux scores. Toutefois, le gain à la tâche Lex OP 2 n'expliquait significativement ni le gain sur les ME ($B = 0.01, SE B = 0.09, \beta = 0.01 ; t = 0.08, p = .94 ; 95\% CI = [-0.17, 0.18]$), ni sur les NE ($B = 0.09, SE B = 0.08, \beta = 0.15 ; t = 1.04, p = .31 ; 95\% CI = [-0.08, 0.26]$).

La triple interaction point de mesure * groupe d'entraînement * type de mots n'était pas significative ($F(2,47) < 1, p = .62, \eta^2_p = .005$), de même que la triple interaction point de

mesure * groupe d'entraînement * gain à la tâche Lex OP 2 ($F(2,47) = 2.43, p = .13, \eta^2_p = .049$).

7.3.1.4 DIPHO : corrélations entre le gain et les variables contrôle

Les corrélations partielles entre le gain en production orthographique sur les scores ME et NE de la DIPHO et les variables pré-test sont présentées dans le Tableau 12.

Concernant les sujets de 2^{ème} année d'instruction, aucun coefficient de corrélation significatif n'est retrouvé, ni pour les ME, ni pour les NE. Concernant les sujets de 4^{ème} année d'instruction, en revanche, plusieurs corrélations négatives, qu'elles soient significatives ou tendanciellemment significatives, sont retrouvées entre le gain à la DIPHO et les variables contrôle. Ainsi, un coefficient de corrélation négatif, faible et tendanciellemment significatif est obtenu entre le gain sur les ME et l'empan visuo-attentionnel partiel ($r = -.46, p = .06$). Un coefficient de corrélation négatif, faible et tendanciellemment significatif est obtenu entre le gain sur les ME et le score à la dictée de mots réguliers ($r = -.44, p = .08$) ; il est négatif, modéré et significatif pour le score à la dictée de mots irréguliers ($r = -.52, p = .03$).

Concernant les tâches de choix orthographique, nous retrouvons une corrélation négative, modérée et significative entre le gain sur les ME et la tâche Sublex OP 2 ($r = -.57, p = .02$), et la tâche Lex OP 1 – HF ($r = -.58, p = .01$). Une corrélation négative, faible et tendanciellemment significative est retrouvée entre le gain sur les ME et la tâche Lex OP 2 ($r = -.45, p = .07$).

Concernant les NE, une corrélation négative, faible et tendanciellemment significative est obtenue entre ce score et l'empan visuo-attentionnel partiel ($r = -.44, p = .07$). Des corrélations négatives, moyennes et significatives sont obtenues entre le score sur les NE et la performance à la dictée de mots réguliers ($r = -.54, p = .02$) et irréguliers ($r = -.52, p = .03$), ainsi qu'aux tâches de choix orthographiques Lex OP 1 – HF ($r = -.60, p = .01$) et Lex OP 2 ($r = -.58, p = .01$).

Tableau 12. Corrélations partielles entre le gain à la DIPHO et les variables pré-test

2^{ème} année d'instruction												
	Gain ME	Gain NE	EVA-G	EVA-P	REG	IRR	NM	SOP1	SOP2	LOP1 HF	LOP1 BF	LOP2
Gain ME	-	-	-.08	-.27	.05	.29	-.22	.10	.11	.15	.23	.01
Gain NE	-	-	.03	-.14	-.07	.18	-.10	-.01	.12	.20	.30	-.04
4^{ème} année d'instruction												
	Gain ME	Gain NE	EVA-G	EVA-P	REG	IRR	NM	SOP1	SOP2	LOP1 HF	LOP1 BF	LOP2
Gain ME	-	-	-.17	-.46 ^t	-.44 ^t	-.52*	-.40	-.40	-.57*	-.58*	-.33	-.45 ^t
Gain NE	-	-	-.04	-.44 ^t	-.54*	-.52*	-.27	-.36	-.35	-.60*	-.16	-.58*

^t = $p < .10$ - * = $p < .05$

NB : Les corrélations présentées contrôlent l'âge des participants en mois.

EVA-G = Empan visuo-attentionnel global. EVA-P = Empan visuo-attentionnel partiel. IRR = Dictée de mots irréguliers. NM = Dictée de non-mots. REG = Dictée de mots réguliers

7.3.1.5 Lex OP 2

Afin d'avoir une indication du progrès de nos participants uniquement dû au temps en l'absence de groupe contrôle, nous avons également réalisé une ANOVA à mesures répétées sur le score obtenu à la tâche de choix orthographique Lex OP 2. Un effet tendanciel du point de mesure est retrouvé ($F(1,48) = 3.43, p = .07, \eta^2_p = .067$), suggérant que les participants de nos deux groupes d'entraînement ont marginalement progressé à cette tâche entre le pré- et le post-test. Nous ne constatons pas d'interaction entre le point de mesure et le groupe d'entraînement ($F(1,48) < 1, p = .34, \eta^2_p = .019$) et n'avons donc pas mené d'analyse post-hoc sur cette variable.

7.3.2 Résultats sur les mesures répétées

7.3.2.1 Analyses visuelles

Le tableau en Annexe 14 détaille les valeurs de niveau et de pente pour chaque participant entre la phase A et B. Le Tableau 13 synthétise l'évolution de ces indicateurs pour chaque profil de participant. Le Tableau 14 réévalue ces indicateurs en ne tenant compte que des trois premiers points en phase B, afin d'estimer l'immédiateté de l'intervention.

2^e NL. Les graphiques correspondant aux mesures répétées en phase A et en phase B de nos sujets de 2^{ème} année d'instruction normo-lecteurs (2^e NL) sont présentées dans la Figure 10a pour ceux ayant suivi l'EE (n = 8) et dans la Figure 10b pour ceux ayant suivi l'EI (n = 6).

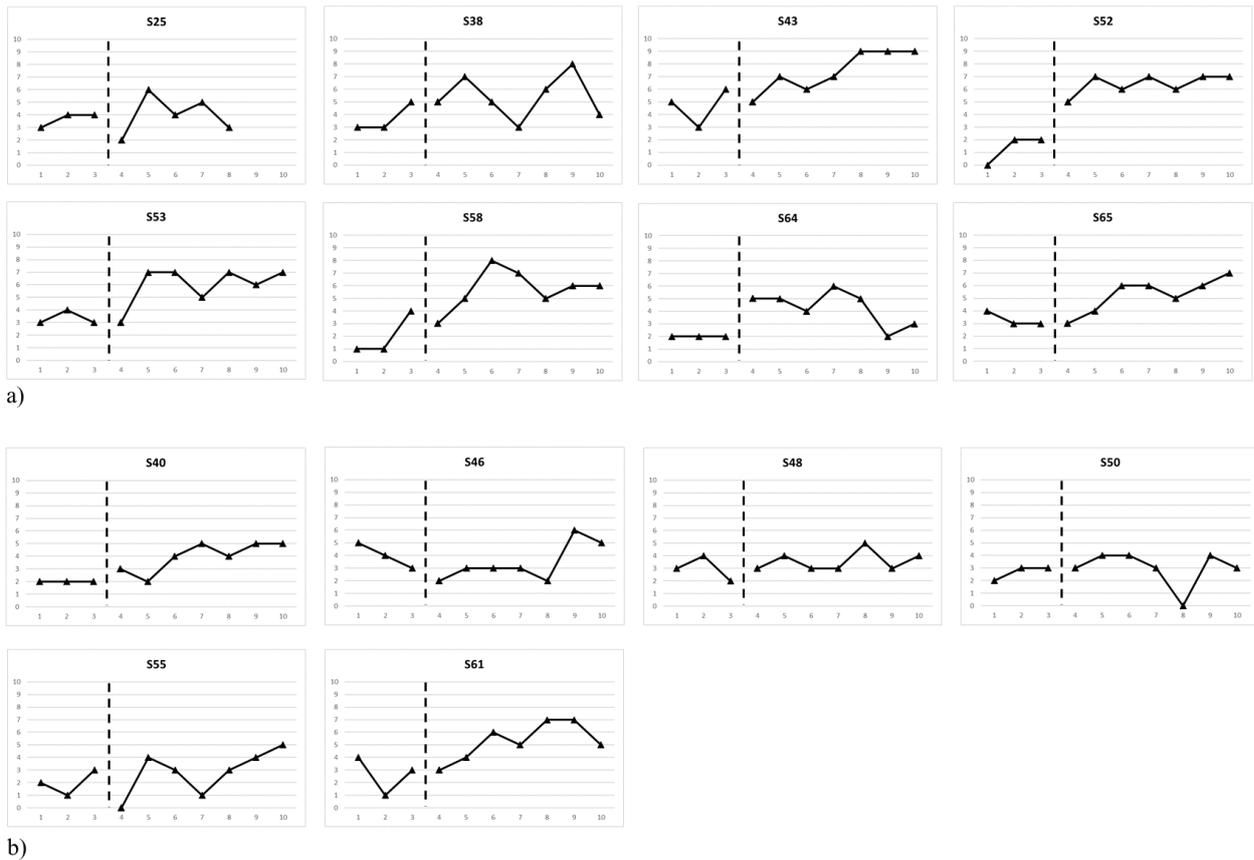


Figure 10. Performances des sujets de 2ème année d'instruction normo-lecteurs aux mesures répétées, en phase A (avant les pointillés) et en phase B (après les pointillés)

- a) Entraînement explicite
- b) Entraînement implicite

Tous nos sujets ayant suivi l'EE ont amélioré leur niveau en phase B par rapport à la phase A, excepté pour le sujet 25, pour lequel le niveau en phase A et B est demeuré le même ; il est à noter que ce sujet est également le seul à ne présenter que 5 points de mesure en phase B. En comparaison, seuls 3 sujets ayant suivi l'EI ont vu l'augmentation de leur niveau en phase B ; 2 ont maintenu le même niveau qu'en phase A ; et 1 (sujet 46) a diminué de niveau. Concernant les sujets ayant suivi l'EE, ce dernier n'a donné lieu à une accélération de la progression des performances en phase B que dans 2 cas. La progression est demeurée identique à celle en phase A pour un sujet (sujet 53), et celle des 5 autres a diminué. En revanche, parmi les sujets ayant suivi l'EI, l'intervention a donné lieu à une accélération de la progression des performances en phase B pour 4 d'entre eux et à une stagnation pour 2

d'entre eux. La comparaison de la médiane et de la pente en phase A par rapport à la médiane et à la pente des trois premiers points de la phase B montre, pour le groupe EE, une amélioration rapide du niveau de performance pour la quasi-totalité d'entre eux (7/8). Les pentes, bien qu'étant pour la plupart positives en phase B (6/8), ne traduisent une accélération de la progression des performances par rapport à la phase A que pour 3 sujets, suggérant un plafonnement rapide des performances autour du niveau ainsi atteint. Concernant le groupe EI, on note également une amélioration du niveau entre la phase A et les trois premiers points de la phase B pour la plupart des participants (4/6), 1 présentant une stagnation (sujet 48) et 1 une régression (sujet 46). La quasi-totalité des sujets (5/6) présente une pente positive sur les trois premiers points de la phase B, et elle est, également pour la quasi-totalité d'entre eux (5/6), d'une valeur plus importante que celle en phase A (le sujet 48 présente une pente nulle sur les trois premiers points en phase B tandis qu'il présentait une pente négative en phase A).

Tableau 13. Synthèse de l'évolution du niveau et de la pente entre la phase A et la phase B pour chaque profil de participant.

	Niveau (médiane)			Pente (SMTL)		
	A = B	A > B	B > A	A = B	A > B	B > A
2^{ème} NL						
EE (n = 8)	1/8	0/8	7/8	1/8	5/8	2/8
EI (n = 6)	2/6	1/6	3/6	0/6	2/6	4/6
2^{ème} TL						
EE (n = 2)	0/2	0/2	2/2	0/2	1/2	1/2
EI (n = 1)	1/1	0/1	0/1	0/1	0/1	1/1
4^{ème} NL						
EE (n = 5)	1/5	0/5	4/5	1/5	2/5	2/5
EI (n = 6)	1/6	0/6	5/6	1/6	1/6	4/6
4^{ème} TL						
EE (n = 4)	0/4	0/4	4/4	0/4	2/4	2/4
EI (n = 4)	2/4	0/4	2/4	0/4	0/4	4/4
6^{ème} NL						
EE (n = 3)	1/3	0/3	2/3	0/3	2/3	1/3
EI (n = 2)	1/2	0/2	1/2	0/2	1/2	1/2
6^{ème} TL						
EE (n = 1)	0/1	1/1	0/1	0/1	1/1	0/1
EI (n = 3)	0/3	0/3	3/3	0/3	2/3	1/3

EE = Groupe entraîné explicitement. EI = Groupe entraîné implicitement. NL = Normo-lecteurs. SMTL = *split-middle trend line*. TL = Trouble de la lecture.

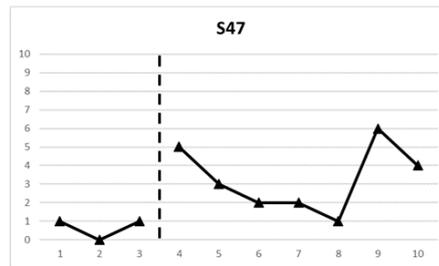
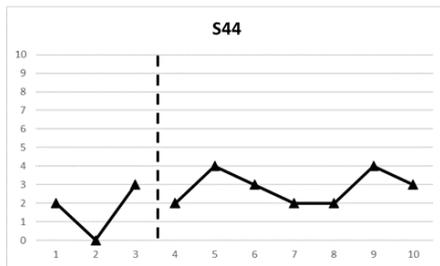
Tableau 14. Synthèse de l'évolution du niveau et de la pente entre la phase A et les trois premiers points de la phase B pour chaque profil de participant.

	Niveau (médiane)			Pente (SMTL)		
	A = B	A > B	B > A	A = B	A > B	B > A
2^{ème} NL						
EE (n = 8)	1/8	0/8	7/8	2/8	3/8	3/8
EI (n = 6)	0/6	1/6	5/6	1/6	0/6	5/6
2^{ème} TL						
EE (n = 2)	0/2	0/2	2/2	1/2	1/2	0/2
EI (n = 1)	0/1	0/1	1/1	0/1	0/1	1/1
4^{ème} NL						
EE (n = 5)	2/5	0/5	3/5	1/5	3/5	1/5
EI (n = 6)	1/6	0/6	5/6	2/6	2/6	2/6
4^{ème} TL						
EE (n = 4)	1/4	0/4	3/4	1/4	2/4	1/4
EI (n = 4)	2/4	0/4	2/4	1/4	1/4	2/4
6^{ème} NL						
EE (n = 3)	1/3	1/3	1/3	1/3	0/3	2/3
EI (n = 2)	1/2	0/2	1/2	0/2	1/2	1/2
6^{ème} TL						
EE (n = 1)	0/1	1/1	0/1	0/1	1/1	0/1
EI (n = 3)	1/3	0/3	2/3	0/3	1/3	2/3

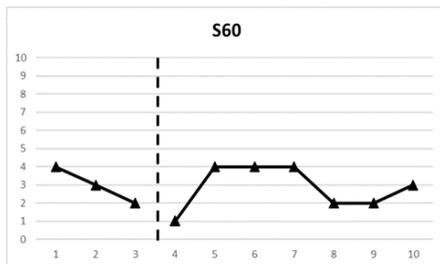
EE = Groupe entraîné explicitement. EI = Groupe entraîné implicitement. NL = Normo-lecteurs. SMTL = *split-middle trend line*. TL = Trouble de la lecture.

En somme, pour les 2^e NL, bien que l'EE ait davantage que l'EI donné lieu à une amélioration du niveau entre la phase A et la phase B, il semblerait que l'EE n'ait pas été capable d'impulser une accélération de la progression des performances, contrairement à l'EI. Ceci pourrait notamment s'expliquer par le fait que la plupart des participants (4/6) auxquels l'EI a été administré présentaient, en phase A, une pente de progression négative ou plate, tandis que les participants de l'EE présentaient déjà, pour la majorité d'entre eux (5/8), une pente positive en phase A. Par ailleurs, l'observation des courbes en phase B laisse entrevoir, pour l'EE comme pour l'EI, une tendance au plafonnement autour du niveau. Toutefois, ce plafond est plus élevé pour l'EE. De plus, l'effet de l'intervention paraît pratiquement immédiat dans l'EE. Ainsi, même si la pente paraît augmenter pour l'EI contrairement à l'EE, ceci pourrait s'expliquer à la fois par une tendance à la diminution progressive des performances en phase A dans l'EI, et par une stabilisation de ces dernières dans les deux groupes d'entraînement en phase B, mais qui fait suite, pour l'EE, à une augmentation quasi-immédiate des scores.

2^e TL. Les graphiques correspondant aux mesures répétées en phase A et en phase B de nos sujets de 2^{ème} année d'instruction avec un trouble de la lecture (2^e TL) sont présentées dans la Figure 11a pour ceux ayant suivi l'EE (n = 2) et dans la Figure 11b pour celui ayant suivi l'EI.



a)



b)

Figure 11. Performances des sujets de 2^{ème} année d’instruction avec un trouble de la lecture aux mesures répétées, en phase A (avant les pointillés) et en phase B (après les pointillés)

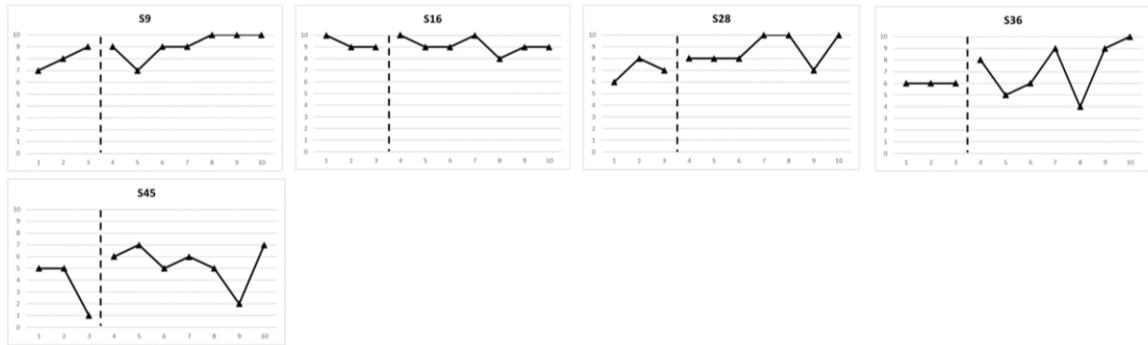
- a) Entraînement explicite
- b) Entraînement implicite

Les sujets 44 et 47 (groupe EE) ont légèrement augmenté leur niveau en phase B (3 ; 3) par rapport à la phase A (2 ; 1), tandis qu’il est resté le même (3) pour le sujet 60 ayant suivi l’EI. Des deux sujets ayant suivi l’EE, seul le sujet 47 présente une accélération de la progression de ses performances en phase B par rapport à la phase A. Le sujet 60, quant à lui, présente une pente négative en phase B (coefficient directeur : -0.5), mais qui constitue nonobstant une amélioration par rapport à la phase A (coefficient directeur : -1). Lorsque seuls les trois premiers points de la phase B sont considérés, toutefois, le pattern de résultats est différent pour tous nos sujets de 2[°] TL. En effet, le sujet 60 semble améliorer son niveau sur cette période, comme en témoigne sa médiane qui passe de 3 en phase A à 4 en phase B. Il présente par ailleurs une pente positive sur ces trois premiers points, tandis qu’elle était négative en phase A. Les deux sujets ayant suivi l’EE présentent l’augmentation de niveau

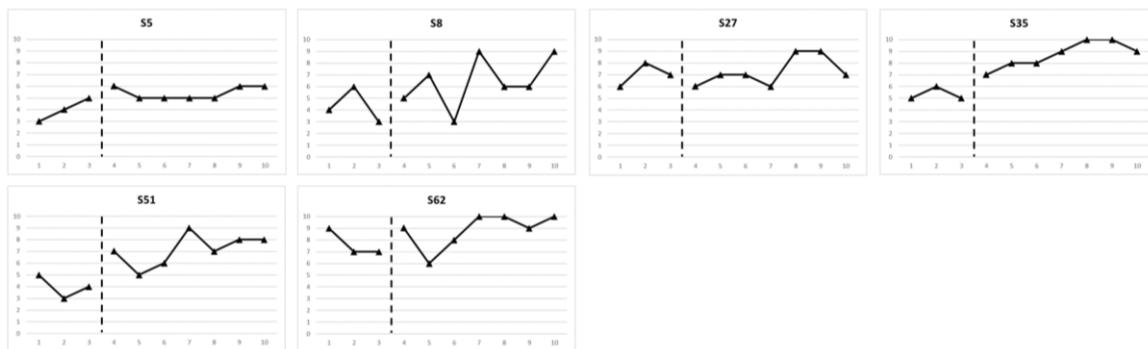
décrite dès ces trois points en phase B, mais les pentes sur ces derniers sont soit similaire à celle de la phase A (sujet 44), soit négative (sujet 47).

En conclusion, les performances de nos sujets 2^e TL aux deux types d'entraînement sont globalement moins bons que ceux obtenus par les 2^e NL. L'augmentation de niveau, bien qu'étant retrouvée pour les deux sujets ayant suivi l'EE, est notamment peu importante et donne rapidement lieu à un plafonnement. Le seul sujet de 2^e NL ayant suivi l'EI présente une accélération de ses performances, et ce dès les trois premiers points de mesure de la phase B. Toutefois, son niveau sur l'ensemble des points reste le même en phase A par rapport à la phase B.

4^e NL. Les graphiques correspondant aux mesures répétées en phase A et en phase B de nos sujets de 4^eme année d'instruction normo-lecteurs (4^e NL) sont présentées dans la Figure 12a pour ceux ayant suivi l'EE (n = 5) et dans la Figure 12b pour ceux ayant suivi l'EI (n = 6).



a)



b)

Figure 12. Performances des sujets de 4ème année d’instruction normo-lecteurs aux mesures répétées, en phase A (avant les pointillés) et en phase B (après les pointillés)

- a) Entraînement explicite
- b) Entraînement implicite

Une augmentation de niveau est constatée pour 4 des 5 sujets du groupe EE et pour 5 des 6 sujets du groupe EI. Le seul sujet du groupe EE à ne pas s’être amélioré sur ce point (sujet 16) a en réalité maintenu son niveau initial, puisqu’il obtenait déjà des scores maximums en phase A. Quant au sujet du groupe EI (sujet 27), ceci semble être dû à un effet retardé de l’intervention, visible seulement à partir de la séance 8. Une nouvelle tendance au plafonnement est retrouvée pour les participants ayant bénéficié de l’EE, ne se traduisant par une accélération de la progression des performances de la pente en phase B par rapport à la phase A que dans 2 cas sur les 5. Pour les participants du groupe EI, en revanche, une accélération de la progression des performances en phase B est constatée pour la plupart des sujets (4/6). Dans l’ensemble, les patterns présentés par les apprenants de cette tranche d’âge, qu’ils aient suivi l’EE paraissent de deux types : les sujets 9, 16 et 28, qui présentaient dès la

phase A un bon niveau à la dictée, semblent plafonner rapidement au maximum. Les sujets 36 et 45, quant à eux, présentent une variabilité marquée dans leurs performances, avec parfois une chute brutale des scores. Les sujets ayant suivi l'EI, semblent, quant à eux, tous bénéficier d'une certaine amélioration de leurs performances à la suite de l'introduction de l'intervention, y compris pour le sujet 5, qui maintient de manière consistante ses performances autour d'une même valeur plus élevée qu'en phase A, et pour le sujet 8, qui montre lui aussi une variabilité importante dans les scores à la dictée. Lorsque seuls les trois premiers points de la phase B sont considérés, les valeurs associées au niveau et la médiane changent peu entre la phase A et la phase B, suggérant un effet rapide de l'intervention pour la majorité des participants, pour les deux types d'intervention.

Pour résumer, dans le cas des 4^e NL, l'EE comme l'EI sont associés à des patterns suggérant une amélioration rapide des performances pour la quasi-totalité des participants. Toutefois, seul l'EI permet une augmentation à la fois du niveau et de la pente en phase A par rapport à la phase B, et ce de manière consistante pour l'ensemble des sujets, bien que cela se traduise par un pattern de progression propre à chaque participant.

4^e TL. Les graphiques correspondant aux mesures répétées en phase A et en phase B de nos sujets de 4^{ème} année d'instruction avec un trouble de lecture (4^e TL) sont présentées dans la Figure 13a pour ceux ayant suivi l'EE (n = 4) et dans la Figure 13b pour ceux ayant suivi l'EI (n = 4).

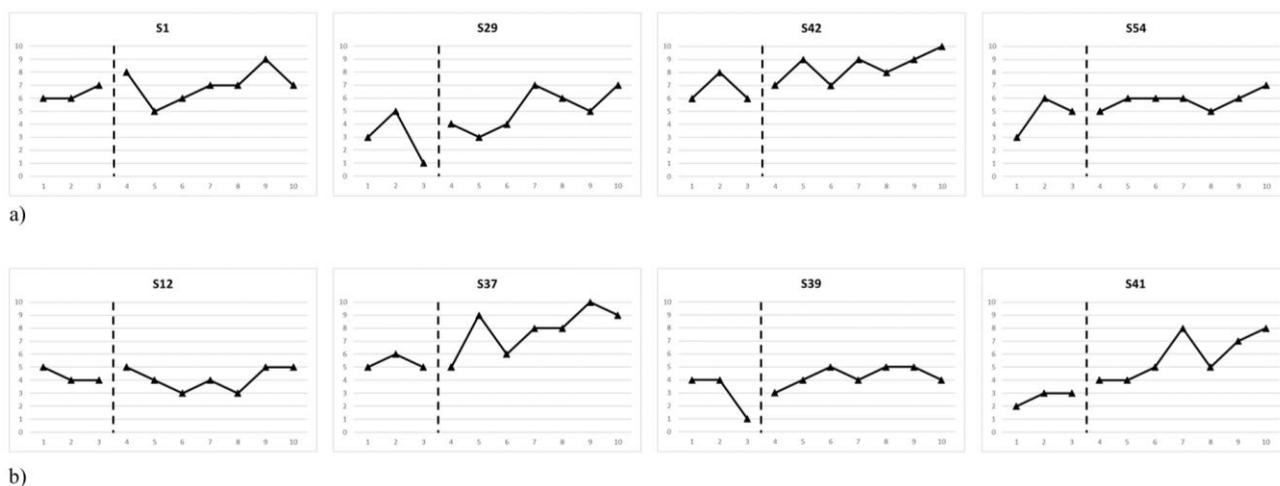


Figure 13. Performances des sujets de 4^{ème} année d'instruction avec un trouble de la lecture aux mesures répétées, en phase A (avant les pointillés) et en phase B (après les pointillés)

- a) Entraînement explicite
- b) Entraînement implicite

Tous les sujets de 4^e TL ayant reçu l'EE ont augmenté leur niveau entre la phase A et B ; en comparaison, ceci n'est constaté que pour la moitié des participants de l'EI. En revanche, tous les sujets de l'EI ont amélioré en phase B la pente de progression constatée en phase A. Les participants de l'EE, eux, présentent à nouveau un profil plus hétérogène sur ce point : l'on constate une accélération de la progression des performances dans 2 cas sur les 4, et un ralentissement dans les 2 autres cas. L'observation des performances sur les 3 premiers points de la phase B révèle, dans le cas de l'EE, un effet rapide de l'intervention pour 3 des 4 participants. En revanche, dans le cas de l'EI, un délai est observé, notamment concernant la pente de progression, qui ne traduit une amélioration des performances qu'à partir de la seconde moitié de la phase B.

En conclusion, et comme constaté sur les profils précédents, les participants de 4^e TL ayant suivi l'EE semblent avoir davantage amélioré leur niveau entre la phase A et B que ceux ayant suivi l'EI. Toutefois, seul l'EI semble avoir donné lieu à une véritable accélération de la progression des performances entre ces deux phases. Lorsque ces deux indices sont considérés, les deux types d'entraînement semblent avoir permis aux participants d'augmenter

leurs scores en dictée entre la phase A et la phase B. Dans le cas de l'EI, toutefois, on note un léger délai à l'apparition de cet effet.

6^e NL. Les graphiques correspondant aux mesures répétées en phase A et en phase B de nos sujets de 6^{ème} année d'instruction normo-lecteurs (6^e NL) sont présentées dans la Figure 14a pour ceux ayant suivi l'EE (n = 3) et dans la Figure 14b pour ceux ayant suivi l'EI (n = 2).

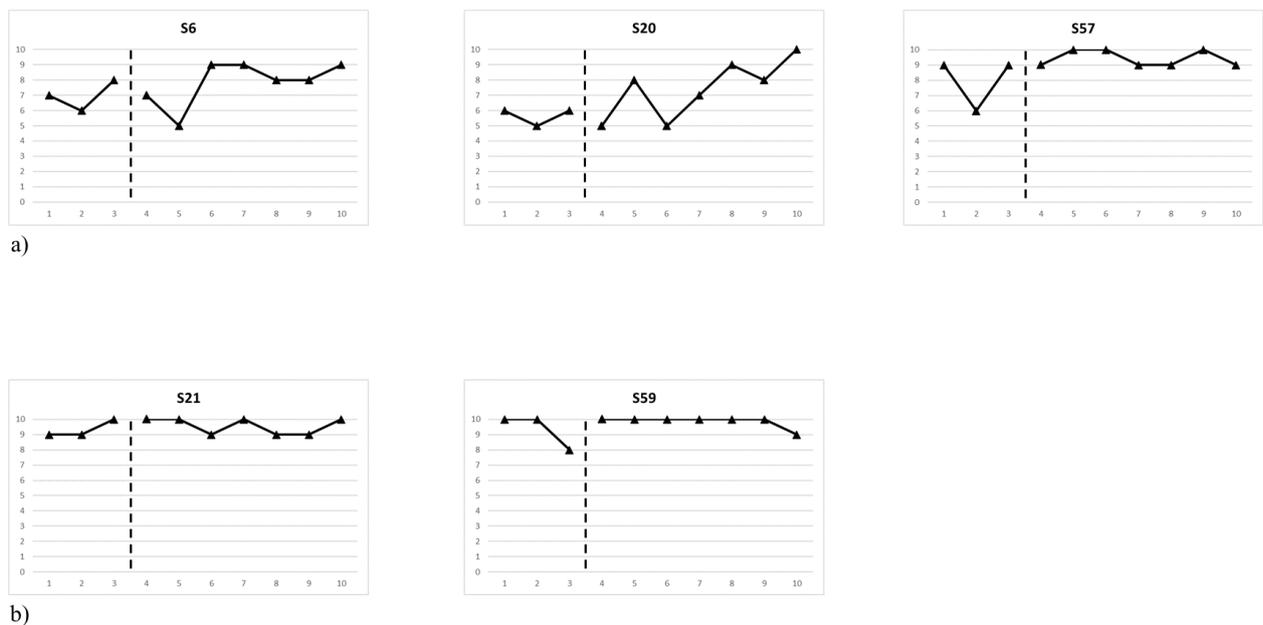


Figure 14. Performances des sujets de 6^{ème} année d'instruction normo-lecteurs aux mesures répétées, en phase A (avant les pointillés) et en phase B (après les pointillés)

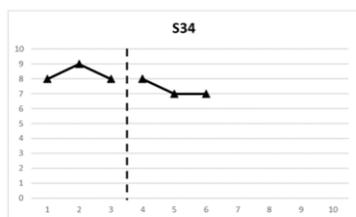
- a) Entraînement explicite
- b) Entraînement implicite

Les deux sujets ayant suivi l'EI montrent un plafonnement aux valeurs maximales (10/10) ou quasi-maximales (9/10) dès la phase A. À l'exception d'un score nettement plus bas lors de la 2^{ème} séance de la phase A (6/10) et qui semble assimilable à un *outlier*, le sujet 57 ayant suivi l'EE paraît montrer un pattern similaire. Ainsi, nous suggérons de ne prêter attention qu'aux sujets 6 et 20, ayant tous deux suivi l'EE, et qui ne montrent pas un tel plafonnement ; le pattern des participants restants ne pouvant être analysé au-delà du constat d'une saturation de leurs performances dès la phase A.

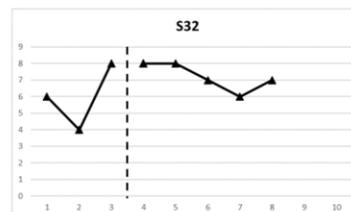
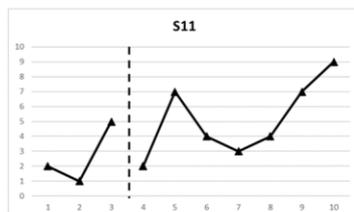
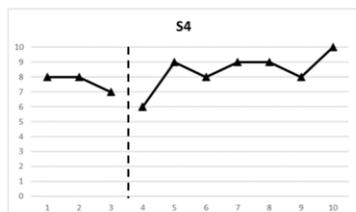
Ces deux participants présentent une augmentation de niveau entre la phase A et la phase B, mais seul le sujet 20 présente une accélération de la progression de ses performances entre ces deux phases. Pour ces deux sujets, les améliorations observées ne se retrouvent pas dès les premiers points de la phase B, suggérant un délai à l'apparition de l'efficacité de l'entraînement, avec des valeurs restant proches de celles de la phase A jusqu'à la seconde moitié de l'entraînement.

En conclusion, 3 de nos 5 participants présentant un profil 6^e NL ont obtenu des performances plafonnant autour du score maximal dès la phase A. L'analyse visuelle n'a donc été conduite que sur les deux participants à n'avoir pas atteint ces scores maximums avant l'introduction de l'intervention, et qui ont tous deux suivi l'EE. Dans le cas de ces deux sujets, l'intervention semble bien avoir donné lieu à une amélioration du niveau de performances, mais seulement à partir de la seconde moitié de la phase B.

6^e TL. Les graphiques correspondant aux mesures répétées en phase A et en phase B de nos sujets de 6^{ème} année d'instruction avec un trouble de la lecture (6^e TL) sont présentées dans la Figure 15a pour celui ayant suivi l'EE (n = 1) et dans la Figure 15b pour ceux ayant suivi l'EI (n = 3).



a)



b)

Figure 15. Performances des sujets de 6^{ème} année d’instruction avec un trouble de la lecture aux mesures répétées, en phase A (avant les pointillés) et en phase B (après les pointillés)

a) *Entraînement explicite*

b) *Entraînement implicite*

L’unique sujet présentant un profil 6^e TL ayant suivi l’EE ne disposant que de trois mesures en phase B, l’analyse visuelle le concernant ne portera que sur la première moitié de l’intervention. Cette dernière ne semble ainsi avoir généré, chez ce participant, aucune amélioration, tant en termes de niveau que de pente de progression. L’effet de l’intervention n’ayant pas été immédiat chez les sujets 6^e NL ayant suivi l’EE (cf. paragraphe précédent), cette absence d’amélioration pourrait être directement attribuable, chez ce sujet, à l’absence de points de mesure pour la seconde moitié de l’EE.

Trois sujets de 6^e TL ont suivi l’EI, et chacun semble présenter un pattern distinct de progression ; tous ont néanmoins amélioré leur niveau entre la phase A et B. Le sujet 5, bien que présentant déjà des performances relativement élevées en phase A, semble avoir amélioré ses scores à la dictée en phase B, ainsi que l’atteste sa pente de progression, qui est passée de négative à positive entre la phase A et la phase B (et ce dès les trois premiers points de cette dernière). Son augmentation de niveau était toutefois soumise à un délai, puisqu’aucune différence entre la phase A et les trois premiers points de la phase B n’est constatée.

Concernant le sujet 11, l'introduction de l'intervention a également donné lieu à un pattern particulier : après une chute momentanée des performances sur les 3 premières séances d'intervention (nuancée par un score relativement élevé à la séance 5), une amélioration notable des performances est retrouvée pour les deux dernières séances. Le sujet 11 a ainsi augmenté de niveau entre la phase A et la phase B ; toutefois, cette évolution positive ne se retrouve pas sur les pentes de progression en phase B, qui sont moins élevées qu'en phase A, tant sur l'ensemble des points que lorsque seuls les trois premiers sont considérés. Enfin, le sujet 32 (qui ne bénéficiait, contrairement aux deux précédents, que de 5 points de mesure en phase B), présente une amélioration de niveau rapide (présente dès les trois premiers points de la phase B), mais son pattern de progression après introduction de l'intervention semble indiquer une baisse progressive des performances, ainsi que le suggère le coefficient directeur négatif de sa pente, déjà observé sur les trois premiers points de cette phase.

Résumé des analyses visuelles et conclusion.

L'analyse visuelle descriptive des mesures répétées visait à dégager différents patterns de progression en fonction des profils d'apprenant et du type d'intervention. De manière générale, l'EE paraît associé à une augmentation rapide du niveau entre la phase A et la phase B, excepté pour les 6^e NL, chez lesquels cette évolution est soumise à un délai. Le seul sujet de 6^e TL à avoir suivi l'EE n'ayant que trois points en phase B, il ne nous est pas possible de déterminer si ce pattern se serait également retrouvé chez lui. Néanmoins, l'absence de toute amélioration après introduction de l'intervention dans son cas pourrait être attribuable à cet effet retardé, constaté chez les 6^e NL.

L'élément caractéristique de l'EI semble être, contrairement à l'EE, une amélioration systématique et immédiate de la pente de progression entre la phase A et la phase B, excepté pour les sujets de 6^e année d'instruction. Les 2 sujets 6^e NL à avoir suivi l'EI ayant obtenu des scores maximaux dès la phase A, leurs performances ne se prêtaient pas à une analyse

descriptive au-delà de ce simple constat. En revanche, les 3 sujets de 6^e TL ont montré un pattern se distinguant de celui observé pour les autres profils de participants : premièrement, tous ont montré une amélioration de niveau entre la phase A et B (ce qui n'était retrouvé, dans le cas de l'EI, que pour les apprenants de 4^e NL). En dépit de ce point de convergence, les courbes de progression en phase B se démarquaient nettement : l'un des participants montrait une amélioration progressive mais lente ; l'un montrait une chute initiale de ses performances avant d'améliorer ses scores sur les deux derniers points de mesure ; et le dernier, en dépit d'un niveau plus élevé en phase B, montrait une tendance progressive à la diminution de ses scores au fil des séances.

Les apprenants de 4^e NL ayant reçu l'EI présentaient également des profils particuliers. Dans l'ensemble, ils ont rapidement amélioré à la fois leur niveau et leur pente de progression entre la phase A et B. De manière intéressante, cet état de fait était retrouvé pour la totalité des sujets, y compris pour les sujets présentant une courbe de progression particulière, comme par exemple un plafonnement autour d'une valeur plus élevée qu'en phase A, ou une variabilité importante des performances tendant toutefois vers l'amélioration. Quant aux sujets de 4^e NL ayant suivi l'EE, bien que montrant tous soit une évolution positive de leur niveau en phase B, soit une saturation de la mesure dès la phase A, ils semblaient présenter un pattern de progression moins lisible.

En conclusion, sur la base de l'analyse visuelle descriptive, l'EE comme l'EI semblent induire une amélioration aux scores de la dictée MEREP pour la plupart des participants, que celle-ci se traduise par une augmentation du niveau ou de la pente de progression en phase B. Cet effet est observé dès la 2^{ème} année d'instruction, mais de manière moins prononcée pour les sujets présentant un trouble de la lecture. La fenêtre d'efficacité semble également plus difficile à déterminer pour les participants de 6^{ème} année d'instruction normo-lecteurs, qui

avaient tendance à saturer leurs scores dès la phase A dans le cas de ceux ayant suivi l'EI, ou à rencontrer un certain délai dans le cas de ceux ayant suivi l'EE.

Ainsi, concernant l'EE, les profils d'apprenant semblant avoir le plus profité de l'intervention dispensée paraissent être les sujets de 2^{ème} année d'instruction normo-lecteurs, et dans une moindre mesure, les sujets de 4^{ème} année d'instruction avec un trouble de la lecture. Concernant l'EI, le profil le plus indiqué concernerait principalement les sujets de 4^{ème} année d'instruction normo-lecteurs.

7.3.2.2 *Indice de non-recouvrement (NAP)*

Afin de compléter ces analyses visuelles, nous avons calculé un indice de *Non-overlap of All Pairs* qui permet d'estimer la taille d'effet de l'intervention ainsi que sa significativité.

Les NAP obtenus pour les sujets 2^e NL sont présentés dans le Tableau 15. Des 8 sujets ayant suivi l'EE, 6 ont obtenu des NAP significatifs. Parmi ces derniers, la valeur rééchelonnée était comprise, pour la moitié d'entre eux, entre 76 et 81%, indiquant un effet médium de l'intervention, et pour les 3 autres, entre 86 et 100%, indiquant un effet fort. Un des participants présentait un NAP rééchelonné à 62% (effet médium) dont la p-valeur était tendancielle ($p = .08$). En ce qui concerne l'EI, 2 des 6 sujets 2^e NL ont obtenu un NAP significatif, respectivement médium (NAPr = 81% ; $p = .03$) et fort (NAPr = 86% ; $p = .02$).

Tableau 15. Indices de non-recouvrement entre les phases A et B chez les sujets de 2^{ème} NL

	Paires	Overlaps	Ties	Non-overlaps	NAP	NAPr	U_L	p
EE								
S25	15	5	3	7	57%	13%	6.5	.44
S38	21	2	4	15	81%	62%	4.0	.08
S43	21	1	2	18	90%	81%	2.0	.03
S52	21	0	0	21	100%	100%	0.0	.01
S53	21	1	2	18	90%	81%	2.0	.03
S58	21	1	0	20	95%	90%	1.0	.02
S64	21	0	3	18	93%	86%	1.5	.02
S65	21	1	3	17	85%	76%	2.5	.04
EI								
S40	21	0	3	18	93%	86%	1.4	.02
S46	21	12	4	5	33%	-33%*	14	.83
S48	21	4	6	11	67%	33%	7.0	.23
S50	21	3	6	12	71%	43%	6.0	.16
S55	21	5	3	13	69%	38%	6.5	.21
S61	21	1	2	18	90%	81%	2.0	.03

* La valeur négative du NAPr indique un effet adverse de l'intervention.

EE = Groupe entraîné explicitement. EI = Groupe entraîné implicitement. NAP = *Non-overlap of All Pairs*. NAPr = *Non-overlap of All Pairs* ré-échelonné.

Les NAP obtenus par les sujets 2^e TL sont présentés dans le Tableau 16. Des deux sujets ayant suivi l'EE, seul le sujet 47 a obtenu un NAP rééchelonné significatif à 90% ($p = .02$). Le sujet 60, unique participant 2^e TL du groupe EI, a obtenu un NAPr non-significatif (-4.8% ; $p = .59$), suggérant un effet adverse de l'intervention.

Tableau 16. Indices de non-recouvrement entre les phases A et B chez les sujets de 2^{ème} TL

	Paires	Overlaps	Ties	Non-overlaps	NAP	NAPr	U_L	p
EE								
S44	21	3	5	13	74%	48%	5.5	.14
S47	21	0	2	19	95%	90%	1.0	.02
EI								
S60	21	8	6	7	48%	-4.8%*	11	.59

* La valeur négative du NAPr indique un effet adverse de l'intervention.

EE = Groupe entraîné explicitement. EI = Groupe entraîné implicitement. NAP = *Non-overlap of All Pairs*. NAPr = *Non-overlap of All Pairs* ré-échelonné.

Les NAP obtenus par les sujets 4^e NL sont présentés dans le Tableau 17. Parmi les sujets ayant reçu l'EE, aucun n'a obtenu un NAPr significatif. Trois sujets, toutefois, ont obtenu de manière tendancielle un NAPr médium compris entre 62 et 71%. Concernant les sujets du groupe EI, 3 d'entre eux ont présenté un NAPr significatif. Pour 2 d'entre eux, il

s'agissait d'un effet fort (95.2% ; $p = .01 - 100%$; $p = .01$), et pour le sujet 5, d'un effet médium (81% ; $p = .02$). Le sujet 8 a également obtenu, de manière tendancielle, un NAPr médium (79% ; $p = .10$).

Tableau 17. Indices de non-recouvrement entre les phases A et B chez les sujets de 4ème NL

	Paires	Overlaps	Ties	Non-overlaps	NAP	NAPr	U_L	p
EE								
S9	21	2	4	15	81%	62%	4.0	.08
S16	21	7	10	4	43%	-14%*	12	.70
S28	21	1	4	16	86%	71%	3.0	.05
S36	21	6	3	12	64%	29%	7.5	.28
S45	21	2	4	15	81%	62%	4.0	.08
EI								
S5	21	0	4	17	90%	81%	2.0	.02
S8	21	3	3	15	79%	57.1%	4.5	.10
S27	21	7	5	9	55%	9.5%	9.5	.45
S35	21	0	0	21	100%	100%	0.0	.01
S51	21	0	1	20	98%	95.2%	0.5	.01
S62	21	4	2	15	76%	52.4%	5.0	.12

* La valeur négative du NAPr indique un effet adverse de l'intervention.

EE = Groupe entraîné explicitement. EI = Groupe entraîné implicitement. NAP = *Non-overlap of All Pairs*. NAPr = *Non-overlap of All Pairs* ré-échelonné.

Les NAP obtenus par les sujets 4e TL sont présentés dans le Tableau 18. Parmi les 4 sujets du groupe EE, seul le sujet 42 présente un NAPr significatif, de taille d'effet médium (NAPr = 76% ; $p = .04$). Un NAPr tendanciel médium (62% ; $p = .08$) est toutefois obtenu par le sujet 29. Concernant le groupe EI, deux participants obtiennent des NAPr significatifs, témoignant, chez le sujet 37, d'une taille d'effet médium (NAPr = 76.2% ; $p = .04$), et chez le sujet 41, d'une taille d'effet forte (NAPr = 100% ; $p = .01$).

Tableau 18. Indices de non-recouvrement entre les phases A et B chez les sujets de 4^{ème} TL

	Paires	Overlaps	Ties	Non-overlaps	NAP	NAPr	U_L	p
EE								
S1	21	4	5	12	69%	38%	6.5	.20
S29	21	3	2	16	81%	62%	4.0	.08
S42	21	2	1	18	88%	76%	2.5	.04
S54	21	2	6	13	76%	52%	5.0	.11
EI								
S12	21	8	7	6	45%	-9.5%*	11.5	.64
S37	21	1	3	17	88%	76.2%	2.5	.04
S39	21	2	6	13	76%	52.4%	5.0	.11
S41	21	0	0	21	100%	100%	0.0	.01

* La valeur négative du NAPr indique un effet adverse de l'intervention.

EE = Groupe entraîné explicitement. EI = Groupe entraîné implicitement. NAP = *Non-overlap of All Pairs*. NAPr = *Non-overlap of All Pairs* ré-échelonné.

Les NAP obtenus par les sujets 6^e NL sont présentés dans le Tableau 19. Aucun participant du groupe EE ni du groupe EI ne présente de NAPr significatif ; toutefois, un NAPr médium et tendanciel (62% ; $p = .06$) est observé pour le sujet 57, qui faisait partie du groupe EE.

Tableau 19. Indices de non-recouvrement entre les phases A et B chez les sujets de 6^{ème} NL

	Paires	Overlaps	Ties	Non-overlaps	NAP	NAPr	U_L	p
EE								
S6	21	4	3	14	74%	48%	5.5	.15
S20	21	4	2	15	76%	52%	5.0	.12
S57	21	0	8	13	81%	62%	4.0	.06
EI								
S21	21	3	10	8	62%	24%	8.0	.30
S59	21	2	12	7	62%	24%	8.0	.26

EE = Groupe entraîné explicitement. EI = Groupe entraîné implicitement. NAP = *Non-overlap of All Pairs*. NAPr = *Non-overlap of All Pairs* ré-échelonné.

Enfin, les NAP obtenus par les sujets 6^e TL sont présentés dans le Tableau 20. À nouveau, aucun sujet du groupe EE comme du groupe EI ne semblent présenter de NAPr significatif. Le sujet 11, qui a suivi l'EI, présente toutefois un NAPr tendanciel et médium (52% ; $p = .10$).

Tableau 20. Indices de non-recouvrement entre les phases A et B chez les sujets de 6^{ème} TL

	Paires	Overlaps	Ties	Non-overlaps	NAP	NAPr	U_L	p
EE								
S34	9	7	2	0	11%	-78%*	8.0	.97
EI								
S4	21	3	4	14	76%	52%	5.0	.12
S11	21	4	1	16	79%	57%	4.5	.10
S32	15	3	3	9	70%	40%	4.5	.22

* La valeur négative du NAPr indique un effet adverse de l'intervention.

EE = Groupe entraîné explicitement. EI = Groupe entraîné implicitement. NAP = *Non-overlap of All Pairs*. NAPr = *Non-overlap of All Pairs* ré-échelonné.

Résumé des indices de recouvrement et conclusion

Le calcul des indices NAP rééchelonnés confirme en partie les conclusions tirées des analyses visuelles. En effet, le profil de participants semblant le plus positivement impacté par l'EE semble être les apprenants 2^e NL, puisque 6 sur les 8 sujets de cette classe d'âge à avoir suivi cet entraînement ont obtenu un NAPr significatif médium ou fort. Toutefois, contrairement à ce que les analyses visuelles semblaient indiquer, les apprenants 4^e TL ne semblent bénéficier que modestement de ce type d'entraînement, puisqu'un seul parmi les 4 obtient une taille d'effet significative et modérée. Concernant l'EI, et ainsi que les analyses visuelles semblaient le suggérer, le profil d'apprenant le plus à même d'en bénéficier semble bien être les 6 sujets de 4^e NL : la moitié d'entre eux ont obtenu un NAPr significatif, médium ou fort. En revanche, aucun des 5 sujets de 4^e NL ne semblait bénéficier d'un EE. Il est à noter que la moitié également des 4 sujets de 4^e TL à avoir suivi l'EI ont obtenu un NAPr significatif.

En ce qui concerne les profils de 2^e TL, de 6^e NL et de 6^e TL, il est délicat de conclure à partir des résultats obtenus, en raison du faible nombre de participants dans les deux groupes d'entraînement. L'efficacité de l'entraînement, quel que soit son mode de délivrance, paraît limitée, puisqu'aucun sujet de 6^e NL ou TL n'obtient de NAPr significatif, que ce soit dans le groupe EE ou dans le groupe EI. Signalons toutefois que, parmi les 3 sujets présentant un profil 2^e TL, le seul à avoir obtenu un NAPr significatif avait reçu un EE.

7.4 Discussion

L'objectif de cette étude était d'explorer l'efficacité d'un entraînement orthographique sublexical (EOS) sur la production orthographique et l'identification des mots d'apprenants de 2^{ème}, 4^{ème} et 6^{ème} année d'instruction, avec ou sans trouble de la lecture, dispensé selon un mode explicite (EE) ou implicite (EI). Nous avons pour cela constitué deux groupes d'entraînement, respectivement de 25 et 26 participants. Nous avons comparé les performances pré- et post-test de nos groupes EE et EI en précision et fluence de lecture au test de l'Alouette-R, ainsi que sur la production orthographique de la terminaison /o/ (DIPHO). En outre, nous avons réalisé chez 45 d'entre eux des mesures répétées via une dictée (MEREP) ciblant également la transcription de la terminaison, et ce sur toute la durée de leur entraînement, pendant les trois premières séances consacrées au travail des prérequis (phase A), puis à l'introduction du travail sur les régularités (phase B).

Les ANOVA à mesures répétées réalisées sur nos groupes EE et EI ont mis en évidence un effet simple du temps, suggérant que nos participants se sont améliorés entre le pré- et le post-test sur l'ensemble des variables d'intérêt (production orthographique de mots entraînés et non-entraînés ; indices de précision et de fluence de lecture). Concernant la DIPHO, un effet du type de mot (entraîné versus non-entraîné) est également retrouvé, suggérant une amélioration plus importante sur les mots entraînés (ME) que sur les mots non-entraînés (NE), indépendamment du groupe d'entraînement. En revanche, nous n'avons pas retrouvé d'interaction entre le point de mesure et le groupe d'entraînement, ce qui ne nous permet pas d'exclure un potentiel effet développemental ou un effet placebo / Hawthorne sur la base de ces analyses. Il convient cependant de signaler deux arguments en faveur d'un effet de l'intervention, qu'elle soit explicite ou implicite. En premier lieu, les tailles d'effet retrouvées sur nos variables d'intérêt en identification des mots et en production orthographique sont systématiquement élevées voire très élevées, y compris lorsque le gain à

la tâche de choix orthographique Lex OP 2, notre variable contrôle, était intégré à l'analyse en tant que covariable. De plus, l'effet du point de mesure sur la tâche Lex OP 2 était modéré, ce qui est consistant avec l'interprétation que l'entraînement a pu impacter les variables d'identification des mots et de production orthographique au-delà d'un simple effet développemental. Néanmoins, ainsi que nous l'avons rappelé précédemment, l'absence de groupe contrôle borne cette interprétation à l'état d'hypothèse.

Nous avons réalisé des corrélations entre nos variables langagières évaluées au pré-test et le gain en production orthographique et en identification des mots suite à l'entraînement, afin de déterminer la potentielle existence de lien entre cette amélioration et les compétences langagières évaluées. Concernant l'identification des mots, sur l'ensemble des liens testés, seule la corrélation entre la tâche de choix orthographique Lex OP 1 BF et l'indice de précision de l'Alouette-R donne lieu à un coefficient significatif, et uniquement pour les apprenants de 2^{ème} année d'instruction. Un coefficient tendanciel est également obtenu entre cette tâche et l'indice de fluence de l'Alouette-R, pour les apprenants de 4^{ème} année d'instruction. Ainsi que nous l'avons rappelé dans le chapitre 4 du présent manuscrit, le lien entre l'identification des mots et la compétence orthographique a été exploré dans l'analyse factorielle confirmatoire (*confirmatory factor analysis*, CFA) menée par Deacon et al. (2019). Cette dernière avait conduit à l'élaboration d'un modèle dans lequel les compétences orthographiques, qu'elles soient lexicales ou sublexicales, étaient sous-tendues par un facteur sous-jacent commun à celui de l'identification des mots écrits. Au vu des liens forts qui semblent exister entre identification des mots et compétence orthographique, une interprétation plausible de la corrélation retrouvée dans notre étude serait que la performance initiale en compétence orthographique constitue un facteur prédictif des gains en identification des mots suite à l'EOS. Cependant, il convient de garder à l'esprit qu'une seule tâche de choix orthographique (Lex OP 1 BF ; Commissaire & Besse, 2019) paraissait

corrélés avec les indices de précision et, dans une moindre mesure, de fluence de lecture. L'analyse ayant été réalisée sans corriger un potentiel biais de multiples comparaisons, et sur un nombre important de variables, il nous paraît préférable de considérer ces résultats avec prudence.

Les corrélations réalisées entre les mêmes variables pré-test et le gain à la DIPHO n'ont donné lieu à des coefficients significatifs que chez les élèves de 4^{ème} année. Ces coefficients étaient par ailleurs tous négatifs, et concernaient l'empan visuo-attentionnel partiel (significativité uniquement tendancielle), la performance à la dictée de mots réguliers et irréguliers, les tâches de choix orthographiques Sublex OP 2 (corrélation avec le gain aux mots entraînés uniquement), Lex OP 1 HF et Lex OP 2. Ces corrélations négatives constituent un résultat surprenant, en ce qu'elles semblent indiquer que plus le niveau initial des participants aux tâches évoquées était faible, plus l'entraînement a donné lieu à un gain important sur la production orthographique de mots entraînés et/ou non-entraînés. Si cette interprétation semble cohérente pour ce qui concerne la dictée de mots réguliers et irréguliers – les moins bons orthographes étant les plus susceptibles de bénéficier d'un entraînement ciblant la production orthographique –, elle est davantage difficile à comprendre en ce qui concerne l'empan visuo-attentionnel (EVA) partiel et les tâches de choix orthographiques. Peu d'études se sont encore intéressées à l'impact de l'EVA sur le développement de la production orthographique, mais l'intégralité des travaux dont nous avons connaissance concluent soit à une corrélation positive entre les capacités de production orthographique et l'EVA (Bosse et al., 2015; van den Boer et al., 2015), soit à l'absence d'une telle relation (Banfi et al., 2018), mais jamais à la présence d'une corrélation négative entre ces deux compétences, comme ce qui semble être observé dans notre étude. Il nous faut cependant nuancer cette observation surprenante. Cette corrélation négative pourrait en effet simplement être le reflet indirect d'un progrès plus important de la part des apprenants les plus en

difficulté de lecture de notre échantillon. De fait, un faible niveau d'EVA en pré-test traduirait des compétences en lecture déficitaires, et l'amélioration de la performance à la DIPHO mettrait en évidence un plus grand potentiel de progression chez les mauvais lecteurs par rapport aux bons lecteurs, dont les capacités d'EVA seraient également préservées. Une telle interprétation pourrait également être envisagée pour expliquer les corrélations négatives observées entre les performances aux tâches de choix orthographique en pré-test (Sublex OP 2, Lex OP 1 HF, Lex OP 2) et le gain à la DIPHO, du fait du lien fort entre identification des mots et compétence orthographique (Conrad et al., 2013; Cunningham et al., 2001; Deacon et al., 2019). Dans cette optique, nous avons réalisé plusieurs modèles de médiation sur les données des apprenants de 4^{ème} année d'instruction, testant alternativement l'effet de médiation de l'indice CM et de l'indice CTL, à la fois sur le gain sur les mots entraînés et non-entraînés ; toutefois, ni l'un ni l'autre de ces indices ne semblait constituer un médiateur significatif de l'effet des tâches susnommées sur le gain aux scores de la DIPHO.

En somme, les analyses réalisées semblent mettre en évidence des corrélations significatives ou tendanciellement significatives entre les variables pré-tests et le gain à l'entraînement, à la fois sur l'identification des mots et la transcription correcte de la terminaison. Néanmoins, le nombre important d'analyses statistiques réalisées, ainsi que les résultats contradictoires obtenus au regard de la littérature limitent les possibilités d'interprétation des coefficients obtenus.

L'introduction de l'EOS a produit des patterns de progression intéressants sur les scores à la dictée MEREP, différents pour la plupart des profils d'apprenants. Concernant les apprenants normo-lecteurs de 2^{ème} année d'instruction (2^e NL), ils avaient tous tendance à stabiliser leurs performances autour de leur niveau en phase B, qu'ils aient reçu l'EI ou l'EE ; toutefois, ce plafond était plus élevé et plus rapidement atteint dans le cas de ceux ayant suivi l'EE. Les apprenants avec un trouble de la lecture de 2^{ème} année d'instruction (2^e TL)

présentaient un pattern similaire, mais traduisant des résultats plus modestes. Les apprenants normo-lecteurs de 4^{ème} année d’instruction (4^e NL), quant à eux, faisaient état de patterns hétérogènes, tant dans le groupe EE que le groupe EI, témoignant dans le premier cas d’un plafonnement rapide aux scores maximaux ou d’une importante variabilité des performances ; et dans le second cas d’une amélioration progressive des performances, s’illustrant par des courbes d’apprentissage diverses, allant d’une stabilisation autour du niveau, à une oscillation, en passant par une progression par paliers réguliers. Concernant les apprenants avec un trouble de la lecture de 4^{ème} année d’instruction (4^e TL), ceux ayant suivi l’EI ont présenté une amélioration lente de leur pente de progression en phase B, et seulement la moitié d’entre eux présentait une augmentation de niveau ; à l’inverse, ceux qui ont suivi l’EE ont augmenté rapidement de niveau, mais seulement la moitié présentait une amélioration de leur pente de progression. Enfin, en ce qui concerne les apprenants de 6^e, nous n’avons pu analyser que les résultats de deux sujets de l’EE pour les normo-lecteurs (NL), et les résultats de trois participants de l’EI pour les apprenants avec un trouble de la lecture (TL). Nous remarquons, dans le cas des premiers, une augmentation de niveau en phase B qui n’a lieu qu’à partir de la seconde moitié de l’entraînement ; dans le cas des seconds, une telle augmentation de niveau est également retrouvée, mais la pente de la courbe de progression de chaque participant, de même que la rapidité d’occurrence des progrès constatés, étaient propres à chaque sujet.

En complément de ces analyses visuelles, un indice de non-recouvrement rééchelonné (NAPr) a également été calculé afin d’estimer la valeur et la significativité de la taille d’effet de l’entraînement pour chaque sujet. Cette dernière semble donner lieu à des résultats différents en fonction de la modalité d’entraînement (explicite versus implicite) et du profil d’apprenant : les 2^e NL semblent avoir en effet presque tous amélioré leurs performances après introduction de l’EE, contrairement aux autres profils d’apprenants, chez qui ce bénéfice était moins systématiquement retrouvé. Par contraste, le profil de sujets ayant le plus

bénéficié de l'EI semble être les 4^e TL. Dans l'ensemble, les sujets de 6^{ème} TL et NL, qu'ils aient suivi l'EE ou l'EI, paraissent avoir peu bénéficié de l'EOS. L'intérêt d'un EE pour des enfants jeunes, et d'un EI pour des enfants avec un trouble de l'apprentissage, font écho à l'étude de Kemper et al. (2012). En effet, les résultats obtenus par ces auteurs indiquaient de manière claire que les enfants de 1^{ère} année d'instruction, au début de leur apprentissage et sans trouble de l'orthographe, s'emparaient d'autant mieux d'une règle orthographique sur des mots entraînés et non-entraînés lorsque cette dernière leur était explicitement présentée, par rapport à une présentation plus implicite. Cet avantage était moins nettement retrouvé pour le groupe d'enfants plus âgés d'un ou deux ans, et présentant un trouble de l'orthographe. Un pattern similaire est reflété dans la présente étude, bien que dans l'ensemble, l'EE semble avoir profité au plus grand nombre.

Nous remarquons, à la marge de ces constatations générales, que lorsqu'un effet était présent chez nos sujets, il était le plus souvent noté dès la première moitié de la période d'intervention. Or, étant donné que les différentes graphies du phonème final /o/ étaient enseignées successivement, à raison d'une régularité différente par séance, des progrès à la MEREP très incrémentiels, et non immédiats, auraient davantage été attendus. Il est toutefois possible que l'acquisition d'une régularité ait provoqué une certaine désambiguation, permettant aux sujets de choisir par procédé d'élimination la bonne graphie pour un mot. Par exemple, si un enfant hésitait initialement entre la graphie O et AUX pour transcrire le mot FRIGO, il se pourrait qu'il décide d'exclure cette dernière à l'issue de la séance 4 abordant la règle de la graphie AUX, bien que la règle d'orthographe relative aux apocopes (O) ne lui ait pas encore été enseignée. L'amélioration de son score ne serait donc pas due qu'à la bonne transcription des graphies selon les règles abordées, mais également à la « non-application » de ces dernières pour les mots restants.

Notre étude est, à notre connaissance, la première à évaluer l'intérêt d'un EOS sur le développement de la production orthographique et de l'identification des mots pour le français. Une autre étude de ce type, menée par Ise et Schulte-Körne (2010) auprès d'enfant de 5^{ème} et de 6^{ème} année d'instruction, a été réalisée en langue allemande. Ces auteurs ont retrouvé une amélioration entre le pré- et le post-test, attribuée à l'intervention, mais qui ne pouvait être distinguée, pour la lecture, d'un effet développemental, et pour la production orthographique, d'un effet placebo / Hawthorne. L'effet spécifique de l'intervention proposée était ainsi difficile à inférer. Dans le cas de notre étude, et bien que nous concluons également à un effet bénéfique de l'EOS, une telle généralisation n'est pas non plus possible, du fait d'une part de l'absence de supériorité d'un des deux types d'intervention par rapport à l'autre, et d'autre part du faible nombre de participants correspondant aux différents profils analysés. Les tailles d'effet retrouvées pour nos variables d'intérêt sont plus importantes que pour notre variable contrôle. De plus, nos analyses sur nos mesures d'intérêt donnent lieu à des indices de non-recouvrement élevés et significatifs. En dépit de ces éléments encourageants, l'inclusion d'un groupe contrôle demeure nécessaire avant que de pouvoir conclure à l'efficacité de l'entraînement. Au-delà d'un potentiel effet développemental, il paraît important d'évoquer l'éventualité d'un effet test / re-test. L'usage des mêmes mesures d'intérêt entre le pré-test et le post-test nous semble susceptible d'avoir participé à l'amélioration des scores lors du post-test, en particulier en ce qui concerne l'Alouette-R. En effet, ce test reposant sur la lecture d'un texte asémantique, il est probable que cette dernière soit plus aisée à la deuxième tentative ; l'émotion de surprise ressentie lors de la découverte du texte aux phrases inattendues a même pu en favoriser la mémorisation (Kim et al., 2021). L'intervalle entre les passations pré-test et post-test étant de 5 semaines, nous estimons que les progrès tendanciellement constatés sur la tâche Lex OP 2 sont probablement davantage à mettre en lien avec un effet test / re-test qu'avec un effet développemental. En ce qui concerne nos

variables d'intérêt, tout ou partie de l'effet constaté sur nos deux groupes d'entraînement pourrait également être expliqué par cette familiarisation avec les mesures d'évaluation.

Additionnellement à ces premières limites relevées, notre étude présente plusieurs défauts méthodologiques qu'il convient de souligner. Tout d'abord, concernant la MEREP, il existait, chez les apprenants de 6^{ème} année d'instruction, une différence tendancielle entre les formes B et C de cette dictée à l'issue de la première séance de ligne de base, suggérant des performances plus élevées pour la forme C. Cette absence d'équivalence entre ces deux formes de la MEREP a pu impacter le pattern de résultat de ces apprenants, en induisant de la variabilité dans les scores obtenus en phase A et en interférant avec un potentiel effet de l'intervention en phase B. Par ailleurs, il paraît important de rappeler qu'il existait déjà des différences entre l'EE et l'EI en phase A, certaines consignes ou certains éléments de matériel étant spécifiques à l'EE. Ceci a pu se refléter notamment dans les coefficients directeurs négatifs en phase A observés pour 3 apprenants 2^e NL de l'EI, contre seulement 1 dans le cas de l'EE. Concernant l'évaluation de l'effet de l'intervention, nous identifions également plusieurs limites importantes. En premier lieu, nous n'avons pas testé la connaissance déclarative des régularités enseignées chez nos sujets au-delà de la transcription orthographique. Une telle vérification nous aurait permis, par exemple via un questionnaire, de tester l'effet spécifique de l'intervention sur nos variables d'intérêt. Ceci apparaît d'autant plus essentiel qu'Ise et Schulte-Körne (2010) ne trouvaient pas, dans leur propre étude, de corrélation entre la connaissance de l'algorithme et les scores aux tests standardisés de production orthographique. En second lieu, le choix de nos mesures d'intérêt comporte des inconvénients : en ne tenant compte que de l'orthographe de la terminaison ciblée par l'EOS, le bénéfice de l'intervention peut être surestimé ; un entraînement qui ne permettrait d'améliorer l'orthographe que d'une partie du mot, et non de son entièreté, possède une application écologique limitée. Toutefois, cette limite en particulier nous semble concerner

tous les entraînements orthographiques sublexicaux, et non seulement notre entraînement. Un autre facteur ayant potentiellement surestimé le bénéfice de l'intervention dans le cadre de la dictée MEREP concerne le recours à l'indice du NAP. En effet, Petersen-Brown et al. (2012) notent, dans leur méta-analyse, que les études évaluant l'efficacité d'un dispositif d'apprentissage avec un *design* SCD en lignes de base multiples (MBD) étaient intrinsèquement exposées au risque de surestimer la valeur cet indice. Ces auteurs suggèrent ainsi de nuancer les valeurs d'interprétation du NAP proposées par Parker et Vannest (2009) dans le cadre des MBD, ne préconisant de considérer comme « fortes » que les tailles d'effet supérieures ou égales à 96% (notons que dans cette étude, le seuil préconisé pour un NAP « faible » était de 93%, compromettant la capacité discriminante même de cet indice). Enfin, d'un point de vue méthodologique, notre *Multiple Baseline Design* (MBD) ne comportait pas d'échelonnage de notre phase A : tous les sujets ont commencé l'EOS à proprement parler à la 4^{ème} séance. Dès lors, il n'est donc pas possible de déterminer avec certitude si la tendance à l'amélioration observée dans la phase B est due à l'introduction de l'intervention, ou à un effet non-spécifique de la prise en charge en elle-même.

En dépit des limites évoquées, l'entraînement orthographique sublexical mis au point, bien que ne permettant pas de réduire entièrement l'inconsistance du phonème final /o/ via l'algorithme enseigné, paraît pouvoir guider, même chez des enfants relativement jeunes, le choix d'une graphie appropriée, en proposant un ensemble de stratégies phonologiques, morphologiques et graphotactiques. L'intérêt de ces stratégies est d'être proposées non pas en opposition, e.g., stratégie phonologique versus morphologique (Casalis & Colé, 2009; Lyster, 2002), ou morphologique versus orthographique (Berninger et al., 2008; Wolter & Dilworth, 2014), mais bien de manière conjointe et complémentaire.

Au final, cette étude confirme, à notre sens, l'intérêt d'investiguer plus précisément l'efficacité d'un entraînement orthographique sublexical métalinguistique pour le

développement de l'orthographe. Les recherches futures devraient inclure un groupe contrôle afin de distinguer l'effet d'un tel entraînement d'un effet développemental, et devraient se concentrer sur un seul profil de participant.

Discussion générale

Cette thèse visait à explorer l'impact d'entraînements ciblant les unités morphologiques et orthographiques pour favoriser le développement de l'identification des mots et de la production orthographique, i.e., le développement de représentations orthographiques mobilisables aussi bien en lecture qu'en transcription. L'entraînement morphologique (EM) a été évalué de deux manières : premièrement, d'une façon purement bibliographique, via la réalisation d'une méta-analyse synthétisant les résultats de 40 études d'intervention morphologique indépendantes ; et, dans un second temps d'une façon expérimentale, par la mise en œuvre d'une étude d'entraînement au *design* pré/post-test, conduite auprès de 51 élèves scolarisés de la 3^{ème} à la 5^{ème} année d'instruction. Un entraînement orthographique sublexical (EOS) métalinguistique, présentant des éléments à la fois phonologiques, morphologiques et graphotactiques, a également été évalué via une étude d'entraînement, selon un double *design* pré/post-test et *Single-Case Design* (SCD), auprès de 51 enfants scolarisés en 2^{ème}, 4^{ème} ou 6^{ème} année d'instruction.

Par le biais de ces méthodologies diversifiées, nous avons souhaité explorer comment la prise de conscience via un entraînement de ces informations sur la structure du langage écrit pouvaient contribuer au développement de l'identification des mots et de la production orthographique. Notre questionnement s'inscrivait tout d'abord dans une perspective fondamentale, en ce qu'il permettait d'étayer le validité de la théorie de l'Intégration de Multiples Patterns (IMP ; Treiman & Kessler, 2014) développée dans le point 1.4 du présent manuscrit, selon laquelle la prise en compte de ces dimensions permet un meilleur apprentissage orthographique, à la fois en lecture et en transcription. Notre démarche était également appliquée, dans la mesure où la validation d'entraînements dispensés dans des conditions écologiques (deux fois 30 minutes par semaine, selon des formats compatibles avec celui d'une remédiation) renseigne sur la pertinence d'approches morphologique ou orthographique pour compléter l'instruction phonographique, dont la validité a déjà été

éprouvée (voir point 2.2). Plusieurs éléments de réflexion se dégagent des discussions respectives de nos différentes études. Nous avons choisi de rapprocher ces éléments sous deux grands axes, que nous détaillons dans les paragraphes suivants : la question de l'impact de l'entraînement morphologique sur l'identification des mots, ainsi que celle de l'impact d'une stratégie morphologique « pure » et d'une stratégie métalinguistique sublexicale sur la production orthographique de mots non-entraînés.

La conscience morphologique (CM), bien que moins étudiée que la conscience phonologique (CP), constitue une compétence métalinguistique désormais bien connue et consensuellement reconnue comme essentielle au développement de la littératie (Passaretti et al., 2023). Toutefois, son impact réel sur le développement de la lecture, et plus précisément sur l'identification des mots, semble partagé entre deux types de résultats dans la littérature, aux interprétations quelque peu contrastées.

D'une part, plusieurs études longitudinales ont permis de mettre en évidence le caractère prédictif de la CM sur le niveau ultérieur de lecture (Casalis & Louis-Alexandre, 2000; Desrochers et al., 2018; Kirby et al., 2012). Ceci est constaté non pas sur des mesures construites pour figurer un nombre important de mots complexes, c'est-à-dire possédant au moins deux morphèmes distincts, mais sur des tests de lecture standardisés, cherchant à évaluer l'efficacité du processus d'identification sur la base d'un corpus de mots représentatif de la langue cible. Ce constat est appuyé par de très nombreuses études corrélationnelles (e.g., (Haase & Steinbrink, 2022; Oliveira et al., 2020; Robertson & Deacon, 2019) retrouvant systématiquement un lien entre CM et identification des mots sur des tâches de lecture « classiques ». Ce résultat, s'il ne permet pas de statuer sur une relation causale entre CM et identification des mots, suggère toutefois que ces deux compétences sont très liées au sein de leur développement, et qu'elles ont tendance à s'influencer dans leur développement respectif. Toutefois, dans l'hypothèse où la CM permettrait effectivement de bénéficier à

l'identification des mots dans son ensemble et non seulement aux mots complexes, le mécanisme sous-jacent n'est pas aisé à expliquer. L'étude 2 de notre chapitre 6 traitant de la mise en place d'un entraînement morphologique a permis de dégager quelques hypothèses à ce sujet. Toutefois, à ce stade, elles ne sauraient constituer une piste interprétative fiable, du fait du caractère inédit des résultats obtenus et des biais méthodologiques relevés dans les points 6.2.3, 6.3.3 et 6.4 (que nous rappelons également plus bas). Les participants de notre groupe entraîné à la conscience morphologique et à la production orthographique de morphèmes (EM), ainsi que de notre groupe entraîné au calcul mental et à cette même production orthographique (EC-PO), semblent avoir progressé significativement entre le pré-test et le post-test sur l'identification de mots irréguliers fréquents (composés seulement à 15% de mots complexes, dont aucun travaillés lors de l'intervention), contrairement au groupe entraîné uniquement au calcul mental (EC-M). En nous appuyant sur le modèle ST-DRC de Pritchard et collaborateurs (2018 ; voir §1.3.3), nous avons suggéré, dans le point 6.4, que cette exposition à des morphèmes (même si elle n'était qu'implicite dans le groupe EC-PO) pourrait avoir encouragé les participants des groupes concernés à solliciter davantage leur système sémantique durant la lecture, leur permettant de compenser l'insuffisance de leur recodage phonologique, due aux inconsistances des CGP des mots irréguliers. Ces mots étant par ailleurs « fréquents », ils avaient en effet de bonnes chances de posséder une entrée dans le lexique phonologique des participants. Cependant, cette hypothèse repose sur deux prémices : tout d'abord, pour que l'appui du système sémantique constitue le facteur déterminant entre les participants ayant reçu ou non une instruction morphologique, cela implique que les mots irréguliers fréquents ne possédaient pas d'entrée dans le lexique orthographique des participants, en ce qu'ils doivent avoir été lus prioritairement via la voie sublexicale. Par ailleurs, cette hypothèse suppose que le seul fait d'avoir pris conscience des particularités des morphèmes suffise à inciter les participants des groupes EM et EC-PO à

chercher le sens de ce qu'ils décodent. Une interprétation alternative au gain en lecture constaté pour les deux groupes entraînés à la morphologie (au-delà de celle d'un artefact statistique due à nos faibles échantillons), pourrait être que ces derniers ont, du fait des entraînements, davantage exercé leur processus de production orthographique général que le groupe EC-M. En effet, les groupes EM et EC-PO ont réalisé une dictée de mots suffixés avec feedback correctif immédiat à la fin de chaque session, contrairement au groupe EC-M. La production nécessitant de traiter la séquence à transcrire dans son entièreté (Shahar-Yames & Share, 2008), et l'activité de transcription sans erreur ayant montré un impact positif sur l'identification des mots (Wanzek et al., 2006, 2018), ce ne serait pas tant la complexité morphologique des mots qui aurait permis aux participants d'améliorer leur lecture, mais plutôt le fait de s'exercer à produire sans s'exposer à des formes orthographiques erronées. Il n'est pas possible, dans le cadre méthodologique de notre étude d'entraînement morphologique, de statuer sur ce point avec certitude. Un groupe contrôle supplémentaire préférentiellement issu des mêmes écoles que les participants des groupes EM et EC-PO, et entraîné à la production de mots simples, serait nécessaire.

L'hypothèse d'une activation renforcée des liens avec le système sémantique du fait d'une exposition à des morphèmes écrits, permettant de mieux décoder les mots irréguliers, bien qu'il puisse permettre d'expliquer le caractère prédictif de la CM sur l'identification de mots issus de tests standardisés rapporté par les études longitudinales, constitue une hypothèse tout à fait inédite. Celle-ci n'est, à notre connaissance, étayée d'aucune autre étude. Notre propre méta-analyse (voir chapitre 5) ayant exploré l'impact des entraînements morphologiques sur plusieurs mesures de lecture tend à ne pas retrouver de bénéfice de cet entraînement sur la lecture, hors identification de mots complexes. À notre connaissance, seule Rassel, dans sa thèse soutenue en 2020, a obtenu un résultat similaire, en ce que ses participants de 2^{ème} année de primaire entraînés à la morphologie ont progressé plus

significativement dans la fluence de lecture d'un texte que son groupe de participants entraînés au vocabulaire. Toutefois, cet avantage n'était pas retrouvé lorsque le groupe entraîné à la morphologie était comparé à un groupe sans entraînement.

Aucune étude d'entraînement s'intéressant à l'impact de la CM sur l'identification des mots n'a, à notre connaissance, investigué cet impact selon les propriétés psycholinguistiques des mots, au-delà de la distinction mots simples / mots complexes. Notre étude d'entraînement est la première à s'intéresser aux progrès différenciés entre mots réguliers et irréguliers, qui avait par ailleurs la particularité de ne présenter qu'une minorité de mots complexes. Nous n'avons par ailleurs pas identifié d'étude corrélationnelle ayant évalué spécifiquement les liens entre CM et mots réguliers versus irréguliers. Cette différenciation pourrait constituer un facteur modérateur de l'impact de la CM sur l'identification des mots, qui gagnerait selon nous à être davantage exploré, afin de mieux comprendre les liens entre CM et lecture.

Les bases théoriques et empiriques de l'impact de la CM sur l'identification des mots complexes sont, en revanche, bien plus étayées. De nombreux travaux Ainsi que nous l'avons développé dans le point 3.2.1 du présent manuscrit, de nombreux travaux ont été répertoriés dans la littérature en faveur d'un décodage morphologique (Levesque et al., 2021), c'est-à-dire d'une capacité pour un individu de lire plus vite et plus précisément un mot, dès lors que celui-ci présente une structure complexe, sensément via la reconnaissance de l'unité orthographique du morphème, avec ou sans accès immédiat à sa signification. Selon le *Morphological Pathways Framework* (MPF ; Levesque et al., 2021), évoqué dans le point 3.2.1, la CM permet de renforcer les représentations morphémiques dans le système orthographique, utilisées dans le décodage morphologique. En retour, l'identification des unités morphémiques en lecture renforce les compétences morphologiques. Cette hypothèse formalisée dans le MPF la relation entre CM et lecture est très probablement bidirectionnelle,

les études d'entraînement permettent de renseigner sur la capacité pour la CM de contribuer au développement de l'identification des mots, au cas où le renforcement de la première via des activités de manipulation de morphèmes engendrerait des progrès dans la seconde.

Afin de mieux comprendre l'impact de la CM sur l'acquisition du langage écrit, nous avons distingué, dans notre méta-analyse (cf. chapitre 5 du présent manuscrit), quatre variables dépendantes relatives à l'identification des mots : la précision de lecture, la fluence de lecture, la lecture morphologique (i.e., la lecture de mots ou non-mots complexes) et le décodage phonologique (i.e., la lecture de non-mots simples). Respectivement 8, 13, 9 et 11 études d'entraînement ont été incluses dans le calcul de la taille d'effet relative à chaque variable, totalisant 489, 922, 1 368 et 833 participants. Parmi ces quatre variables d'identification des mots, seule la lecture morphologique a donné lieu à une taille d'effet significative. Cette dernière était positive et faible, et sa valeur restait la même avec ou sans prise en compte de la variabilité inter-études. Ainsi, l'effet causal de la CM sur la lecture de mots complexes apparaît modeste mais relativement robuste à travers les études l'ayant investigué. Notre méta-analyse semble donc confirmer que la CM contribue bien à l'identification des mots, mais via la reconnaissance rapide de morphèmes (i.e., décodage morphologique), expliquant que cet effet ne soit observé que sur les mots complexes.

Une étude récente (Levesque & Deacon, 2022) a cherché à clarifier de quelle manière la CM contribuait à l'identification des mots au-delà des mots complexes. Cette étude est la première à notre connaissance à avoir inclus dans un modèle structurel longitudinal à la fois des listes de mots complexes (évaluant le décodage morphologique) et des tests de lecture standardisés (TOWRE ; Torgesen et al., 1999) présentant des mots simples et complexes, réguliers et irréguliers (évaluant les capacités plus générales d'identification des mots). Il apparaît que lorsque le décodage morphologique est pris en compte dans le modèle, la CM en 3^{ème} année d'instruction prédit significativement pour l'année suivante les capacités de

décodage morphologique, mais pas les capacités plus générales d'identification des mots. Il semblerait donc que l'effet de la CM sur la lecture, retrouvé par les études longitudinales précédentes, soit expliqué par le fait que ces études n'évaluaient pas le décodage morphologique (i.e., des listes de mots complexes), permettant aux capacités générales d'identification des mots (évaluées par des tests standardisés) de récupérer la variance associée.

Pris ensemble, le résultat de cette étude ainsi que de notre méta-analyse paraissent apporter des arguments solides en faveur d'une contribution de la CM pour l'identification des mots, médiatisée principalement par la lecture des mots complexes. L'étude longitudinale de Levesque et Deacon (2022) suggère que l'effet prédictif de la CM sur la l'identification des mots évaluée par des tests standardisés serait rattachable à la présence de mots complexes dans ces tests. Ainsi, selon cette piste interprétative, l'effet retrouvé dans notre propre étude d'entraînement morphologique sur la lecture de mots irréguliers serait également attribuable à la présence de mots complexes dans cette liste. Deux arguments nous incitaient initialement à écarter cette hypothèse : d'une part, la faible proportion de mots complexes sur l'ensemble des mots de la liste (15%), et d'autre part, le fait que cette proportion soit plus importante sur d'autres listes n'ayant pourtant pas montré d'effet significatif de l'entraînement (e.g., les mots réguliers peu fréquents présentaient une proportion de 25% mots complexes). Toutefois, ni l'effet de consistance, ni l'effet de fréquence, n'ayant – à notre connaissance – fait l'objet d'une investigation spécifique tant dans les études longitudinales que dans les études d'entraînement, il est possible que ces effets soient de nature à moduler l'impact de la CM sur l'identification des mots.

En conclusion, si notre méta-analyse semble apporter un argument supplémentaire en faveur d'un effet de la CM sur l'identification des mots médiatisé par la présence de mots complexes dans les corpus concernés (i.e., via un décodage morphologie ; Levesque et al.,

2021), notre étude d'entraînement morphologique suggère que d'autres paramètres psycholinguistiques (fréquence, consistance) devraient être explorés afin d'affiner encore cette interprétation.

Le second grand axe exploré dans l'ensemble de nos études concernait la production orthographique. Notre étude d'entraînement morphologique, tout comme notre méta-analyse, semblent retrouver un impact positif de la CM sur la transcription de morphèmes et des mots entiers. Cette interprétation est plus délicate à tirer dans le cadre de notre entraînement orthographique sublexical (EOS), en raison de limites méthodologiques importantes (absence de groupe contrôle pour le *design* pré/post-test ; absence d'échelonnage de la ligne de base pour le *design* en mesures répétées). Toutefois, les tailles d'effet obtenues et les patterns d'évolution paraissent compatibles avec un effet sur la production orthographique des terminaisons en /o/ qui ne serait pas attribuable uniquement à un effet développemental ou placebo / Hawthorne.

Outre cet effet bénéfique « global » des interventions mises en œuvre, il est intéressant, voire surprenant, de constater que dans le cas de l'entraînement morphologique tout comme dans celui de l'EOS, la version plus « implicite » de l'entraînement semblait donner lieu à des effets d'entraînement comparables à ceux de la version plus explicite. Ainsi, dans l'étude d'entraînement morphologique, le groupe EM (i.e., entraîné à la CM et à la production orthographique de morphèmes) n'obtenait pas de résultats significativement supérieurs au groupe EC-PO (i.e., entraîné aux mathématiques et à la production orthographique de morphèmes, sans CM) en production orthographique, et ce quel que soit le score (mots entiers, bases ou suffixes ; mots entraînés ou non-entraînés). De même, dans l'étude d'EOS, le groupe d'entraînement explicite (EE ; groupe exposé à l'algorithme et auquel les régularités étudiées étaient présentées comme des « règles » formalisées), n'a pas obtenu au post-test de performances significativement supérieures au groupe d'entraînement

implicite (EI ; groupe auquel l’algorithme n’était pas présenté, et auquel les régularités étaient introduites comme de simples patterns récurrents, sans autre commentaire de la part de l’examinatrice) dans la transcription des terminaisons en /o/, que cela concerne les mots entraînés ou non-entraînés (voir § 7.3.1.3). La méthodologie en *Single-Case Design* (SCD) suggère des effets qualitativement différents de ces deux interventions, en fonction du profil d’apprenant. D’un point de vue plus quantitatif, l’EE semble être celui qui bénéficie au plus grand nombre lorsqu’on considère les changements de niveau (i.e., comparaison de la médiane en phase A versus phase B). Toutefois, lorsqu’on considère l’évolution de la pente (i.e., comparaison de la ligne de tendance en phase A versus phase B), c’est l’EI qui semble générer les résultats les plus encourageants (voir § 7.3.2.1).

Ce point constitue, au regard des données de la littérature, un résultat surprenant. Sauf cas relativement isolés (e.g., Kemper et al., 2012), il ressort en effet que l’approche explicite obtient systématiquement des gains plus importants que tout autre approche non-structurée en ce qui concerne les apprentissages langagiers, en particulier pour les apprenants en difficulté. C’est notamment le constat de la méga-analyse de Bissonnette et collaborateurs, menée sur 11 méta-analyses et totalisant ainsi 362 études indépendantes et plus de 30 000 participants (Bissonnette et al., 2010). Malheureusement, cette méga-analyse n’incluait pas de méta-analyse portant sur la production orthographique. Toutefois, les travaux de Wanzek et collaborateurs (Wanzek et al., 2006, 2018; Williams et al., 2017) soulignent que l’instruction explicite demeure l’approche la plus efficace pour l’enseignement de la production orthographique, en particulier chez les enfants présentant des difficultés en lecture/écriture.

Plusieurs études ont pourtant démontré que les enfants d’âge scolaire étaient capables d’apprendre l’orthographe des mots de manière implicite (e.g., Pacton et al., 2014; Share, 1999). Toutefois, la majorité des études sur le sujet ont exploré les effets de cet apprentissage incident sur les capacités de transcription chez des enfants tout-venants, sans trouble de la

lecture. Il est probable que dans la population présentant des troubles de la lecture en particulier, l'apprentissage implicite soit particulièrement difficile, probablement du fait d'une procédure de recodage phonologique moins efficace, gênant l'apprentissage orthographique (Frith, 1985; Share, 1995), voire même d'un déficit spécifique dans l'apprentissage implicite des séquences, qui s'estompe dès lors que la structure de la séquence est explicitée (Jiménez-Fernández et al., 2011).

En ce qui concerne les liens entre CM et production orthographique de manière plus spécifique, les études d'entraînement morphologique ayant comparé les bénéfices d'un entraînement explicite versus implicite témoignent pour la majorité d'une supériorité de l'approche explicite (e.g., Burton et al., 2021; Lee & Choi, 2020) par rapport à l'approche implicite. Burton et collaborateurs se sont notamment intéressés à la transcription de morphèmes après une session unique d'entraînement morphologique de 35 minutes ciblant les morphogrammes, soit de manière explicite (i.e., l'examinatrice prodiguait aux élèves des conseils et des règles d'orthographe morphologique pour déduire l'orthographe d'un dérivé, telle que l'invariance morphémique de la base et des suffixes), soit de manière implicite (i.e., les élèves devaient essayer de déduire l'orthographe du dérivé sans aucune aide de la part de l'examinatrice). Après avoir suggéré des dérivés possibles, les deux groupes étaient exposés à l'orthographe correcte du dérivé étudié. Les performances de ces deux groupes étaient comparées entre elles avant et après cette session d'entraînement, ainsi qu'à celle d'un groupe sans entraînement. Le groupe entraîné implicitement ne se distinguait ainsi pas du groupe sans entraînement, contrairement au groupe entraîné explicitement, qui obtenait des performances significativement supérieures aux deux autres groupes, à la fois au post-test immédiat et différé, et sur des vrais mots complexes (non-entraînés) comme sur des non-mots complexes.

Il semblerait donc que pour l'entraînement morphologique, l'instruction explicite génère des résultats supérieurs à la simple exposition à des morphèmes écrits, et que cette

dernière ne soit pas systématiquement plus efficace qu'une absence d'entraînement. Les deux études d'entraînement implicite de Bar-Kochva et collaborateurs (Bar-Kochva et al., 2020; Bar-Kochva & Hasselhorn, 2017), incluses dans notre méta-analyse, n'ont d'ailleurs pas obtenu de taille d'effet significative, que ce soit pour l'identification des mots comme pour la production orthographique.

En ce qui concerne l'entraînement orthographique sublexical, seule l'étude de Cordewener et collaborateurs (2016) a, à notre connaissance, comparé l'efficacité d'une telle approche à une intervention implicite, i.e., à la simple exposition à la forme correcte du mot faisait l'objet d'un entraînement. De même que ce que nous avons rapporté au sujet de l'entraînement morphologique, les auteurs retrouvaient un impact bénéfique d'une modalité d'administration explicite par rapport à une administration implicite.

Ainsi, l'absence de différence significative entre ces deux modalités d'administration dans nos propres études morphologique et orthographique s'inscrivent clairement en porte-à-faux par rapport aux données de la littérature. Nous évoquerons dans les paragraphes suivants deux hypothèses explicatives qui nous semblent des pistes intéressantes, non simplement pour rendre compte de cette situation divergente, mais également dans une perspective de recherches futures cherchant à explorer dans quelle mesure l'administration explicite est réellement préférable à une administration implicite.

Tout d'abord, il est possible que la variante « implicite » proposée dans nos deux études d'entraînement ait en réalité contenu des éléments plus structurés et dirigés que la plupart des approches implicites auxquelles nous nous référons. En effet, dans le groupe EC-PO (entraînement à la production de morphèmes sans entraînement à la conscience morphologique), les participants ont dû adopter une posture active prodiguée par la situation de transcription, qui a certainement généré un apprentissage plus explicite que ne l'aurait été la simple exposition à des mots morphologiquement construits, sans entraînement à les écrire.

Bien que l'examinatrice n'ait pas fourni aux participants d'instruction d'ordre morphologique, elle a tout de même montré l'orthographe correcte et demandé aux participants de la recopier lorsqu'elle n'était pas obtenue dès le premier essai. Même si l'objectif d'apprentissage n'était pas apparent, les participants bénéficiaient ainsi de plusieurs occasions d'apprentissage des unités sublexicales ciblées. Ce mode d'administration, s'il n'était pas entièrement explicite, ne peut certainement pas être considéré comme étant aussi implicite que la simple présentation de mots suffixés. De même, en ce qui concerne le groupe EI de notre étude d'EOS, plusieurs éléments de l'entraînement pourraient être considérés comme au moins en partie explicite. En effet, en plus de devoir régulièrement écrire les mots présentés lors des séances, toujours avec un feedback correctif immédiat en cas des erreurs, l'examinatrice orientait l'attention des participants sur les régularités étudiées, bien qu'aucune « règle » ne soit formulée. Là encore, un tel entraînement, s'il n'est pas aussi structuré que sa contrepartie explicite, ne saurait être considéré comme entièrement implicite.

Par ailleurs, au-delà d'une approche qui n'était ainsi sans doute pas aussi implicite que ce qui est habituellement constaté dans la littérature, il est possible que les différences observées sur nos propres études d'entraînement, par rapport aux études précédentes, puissent être attribués à d'autres particularités dans l'administration de l'entraînement. La durée d'intervention, par exemple, était plus longue dans nos études que ce qui est typiquement rapporté dans les *designs* cherchant à comparer l'efficacité d'un entraînement explicite versus implicite (entre 35 minutes et 4 heures dans les études évoquées). En augmentant la durée de l'entraînement, il nous paraît envisageable que les différences entre approche explicite et approche implicite deviennent moins évidentes, potentiellement parce que les participants recevant un entraînement implicite commencent alors à cumuler suffisamment d'expériences d'apprentissage pour déduire l'objectif de l'instruction qui leur est présentée. Ceci est d'autant plus cohérent dans la mesure où la majorité des enfants que nous avons recrutés ne

rencontraient pas de difficultés de lecture, et avaient ainsi probablement des capacités d'apprentissage implicite préservées.

En somme, de même que ce que nous avons constaté sur les liens entre CM et identification des mots, le fait que nos études d'entraînement n'aient pas mis en évidence le résultat unanimement retrouvé dans la littérature pourrait s'expliquer par la présence de variables modératrices (la présence de quelques éléments d'enseignement structuré dans l'instruction implicite ; durée d'intervention supérieure à 4 heures) qui n'ont pas, à notre connaissance, été prises en compte lorsqu'il s'agissait de comparer les bénéfices de l'entraînement explicite par rapport à l'entraînement implicite.

Ces différences dans la mise en œuvre de nos études d'entraînement s'accompagnent également de limitations méthodologiques, qui ont été évoquées dans leurs discussions respectives. Nos deux études d'entraînement souffrent ainsi de tailles d'échantillon faibles (inférieures à 30 participants, et même à 20 dans la seconde partie de notre étude morphologique) ainsi que d'un groupe contrôle absent dans l'étude d'EOS, et rajouté « après-coup » dans l'étude morphologique, résultant en des facteurs confondants nuisant à l'interprétation des résultats. Cette limite en particulier est rattachable à la situation de pandémie mondiale du COVID-19. En effet, nous avons rencontré une mortalité expérimentale dans l'étude morphologique, car n'ayons pu commencer notre étude d'entraînement qu'au courant du mois de mars, l'entraînement s'est achevé à la fin de l'année scolaire, et certains enfants étaient déjà partis en vacances. Concernant l'étude d'entraînement orthographique, nous avons préféré intervenir à domicile et non au sein des écoles, ce qui nous assurait de pouvoir continuer à intervenir même si ces dernières fermaient, mais ce qui limitait mécaniquement notre nombre d'interventions possible. Outre nos petits échantillons, nous n'avons, ni dans l'étude d'entraînement morphologique, ni dans l'étude d'EOS, évalué la fidélité du traitement dispensé par les examinatrices. Bien qu'un scénario ait été rédigé en

amont des entraînements pour s'assurer de la reproductibilité de l'intervention et pour uniformiser l'administration à travers les différentes instructrices dans l'étude morphologique, nous ne nous sommes pas assurées que ces dernières dispensaient effectivement les entraînements de manière conforme au scénario. Nous aurions pu, notamment, intégrer une check-list à ces scénarios, afin d'obtenir une indication de la fidélité de l'administration (e.g., Crosson et al. 2018). Bien que nous ayons encouragé les examinatrices à nous contacter en cas de problème rencontré pendant les entraînements, et que nous leur ayons régulièrement demandé si tout se passait bien, nous n'avons pas réalisé de réunion formelle au cours desquelles nous aurions pu nous assurer de leur compréhension et de leur aisance dans l'administration des entraînements, comme cela a été fait dans l'étude d'Ise et Schulte-Körne (2010). Additionnellement, nous aurions pu demander aux examinatrices de remplir des rapports après chaque séance sur la façon dont la session s'était déroulée, comme dans l'étude de Gellert et al. (2021).

Nous avons évoqué le fait, dans la discussion générale de notre étude morphologique, qu'une différence dans les capacités de production orthographique entre le groupe EM et EC-PO aurait potentiellement pu être observée si l'évaluation pré/post-test comportait des épreuves plus écologiques qu'une dictée de mots, telles qu'une épreuve de production écrite libre ou guidée. Une même limite est constatée dans l'étude d'EOS, mais également dans notre méta-analyse, puisqu'une seule étude proposant ce type d'épreuve a pu être incluse (i.e., Allen & Lembke, 2020), ne permettant pas de calculer une taille d'effet moyenne. Dans le même ordre d'idée, aucune de nos études d'entraînement ne comportait une épreuve de compréhension écrite. Même si l'identification des mots constitue un prérequis à cette dernière (Gough & Tunmer, 1986; Tunmer & Chapman, 2012; Wang et al., 2019), elle constitue la finalité de toute activité d'identification des mots et il paraît essentiel que d'éventuels gains sur cette compétence soient évalués dès lorsqu'un entraînement possédant

une ambition applicative est proposé à des enfants d'âge scolaire. Pour cette même raison, nous considérons que l'absence de vérification d'un potentiel maintien à long terme suite aux entraînements, qu'il concerne la méta-analyse aussi bien que l'entraînement morphologique et l'EOS, constitue un manquement majeur pour s'assurer de la pertinence applicative des interventions proposées.

En dépit des limitations relevées, nous considérons également que notre travail de thèse comporte des intérêts inédits pour le champ des entraînements métalinguistiques sublexicaux. Tout d'abord, notre méta-analyse proposait une étude de l'impact de l'entraînement morphologique sur 11 variables dépendantes en lien avec le langage écrit, dont certaines n'avaient jamais été explorées jusqu'à lors (lecture morphologique, précision et fluence de lecture, production orthographique de morphèmes), de même que de nombreuses variables modératrices (système de transcription, opacité du système de transcription, type d'entraînement morphologique, etc.). De plus, elle était la première depuis la revue systématique de Reed (2008) à ne s'intéresser qu'aux interventions dans lesquelles la conscience morphologique (CM) constituait au minimum 2/3 de l'entraînement, permettant d'explorer le véritable impact de cette compétence, en-dehors d'autres éléments dont l'efficacité a déjà été reconnue (e.g., entraînement phonologique ou graphophonologique). Notre entraînement morphologique a permis d'apporter des éclairages inédits sur le rôle de la CM dans l'identification des mots dans le cadre d'un test standardisé de lecture (possible valeur modératrice des effets de consistance et de fréquence des mots simples ou complexes), ainsi que sur la production orthographique (remise en question de la pertinence d'ajouter des activités de manipulation morphologique par rapport au seul entraînement à la production orthographique de morphèmes). Enfin, notre étude orthographique sublexicale, en plus de proposer un protocole d'entraînement alliant les différentes unités ciblées par la Théorie de l'Intégration des Patterns Multiples (Treiman, 2017a; Treiman & Kessler, 2014), ce qui

n'avait pas encore été proposé pour la langue française, et a permis, grâce à l'analyse en *Single-Case Design*, de dégager de potentiels profils d'apprenants qui seraient les plus à même de bénéficier de cet entraînement (i.e., les normo-lecteurs de deuxième année d'instruction pour l'entraînement explicite, et les apprenants de quatrième année d'instruction présentant un trouble de la lecture pour l'entraînement implicite).

Conclusion

Cette thèse de doctorat avait pour but de contribuer à la compréhension de l'intérêt d'approches complémentaires aux entraînements graphophonologiques pour favoriser l'acquisition de l'identification des mots et de la production orthographique. Nous nous sommes concentrées sur les entraînements morphologique et orthographique sublexical, en nous proposant d'explorer des aspects encore peu compris dans la littérature. Bien que nos études d'entraînement comme notre méta-analyse comportent des biais limitant la portée de nos résultats, elles permettent d'enrichir les connaissances sur ces entraînements, et de dégager des pistes de recherches ultérieures qui n'avaient pas encore été évoquées.

Nommément, notre méta-analyse a mis en évidence des manques dans la littérature concernant les études morphologiques : très peu d'entre elles s'intéressaient à des apprenants scolarisés dans les dernières années d'instruction (10^{ème} à 12^{ème} année), à la morphologie flexionnelle seule, ou à l'impact des entraînements morphologiques pour les langues non-alphabétiques. Notre étude d'entraînement morphologique a retrouvé un impact des deux groupes entraînés à la production orthographique de morphèmes (avec ou sans activité de conscience morphologique) sur l'identification de mots irréguliers fréquents, suggérant que les effets de consistance et de fréquence pourraient moduler la façon dont la conscience morphologique contribue à l'identification des mots complexe et à la lecture en général. Elle ne retrouvait pas de supériorité d'activités de manipulation de morphème, par rapport à l'activité de transcription de ces morphèmes, ce qui n'a, à notre connaissance pas été investigué de la sorte jusqu'à présent. Ces deux résultats inédits demandent à être reproduits en adressant les limites méthodologiques évoquées au sujet du groupe contrôle et de la taille d'échantillon. Enfin, notre entraînement orthographique sublexical, à visée exploratoire, semble confirmer l'intérêt de proposer une intervention alliant les patterns phonologiques, morphologiques et graphotactiques, en particulier pour les apprenants les plus jeunes ou présentant des troubles de la lecture.

Annexes

Annexe 1

ANOVA non-paramétrique entre les items entraînés et non-entraînés de la COMOS

	χ^2	p-valeur
U mot dérivé	0.00	1
U base	0.83	0.36
NLET dérivé	0.29	0.59
NLET base	1.15	0.28

U = Occurrence pour un million de mots. NLET = nombre de lettres.

NB. Les valeurs de U sont issues de la base de données Manulex (classes CP-CM2).

Annexe 2

ANOVA non-paramétrique entre les items entraînés et non-entraînés de la DIMO

	χ^2	p-valeur
U	0.66	0.42
NLET	0.19	0.68
NSYL	0.06	0.81
NPHON	0.96	0.33
NGRAPH	0.18	0.67
FRGPTTO	0.86	0.36
COPGTTO	1.08	0.30

U = Occurrence pour un million de mots. NLET = nombre de lettres. NSYL = nombre de syllabes. NPHON = nombre de phonèmes. NGRAPH = nombre de graphèmes. FRGPTTO = fréquence des correspondances graphème-phonème (*token count*). COPGTTO = consistance des correspondances phonème-graphème (*token count*).

NB. Les valeurs de U sont issues de la base de données Manulex (classes CP-CM2).

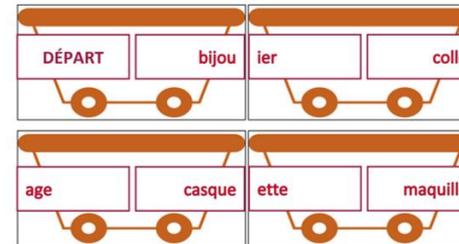
Annexe 3

Exemples des jeux utilisés dans l'entraînement morphologique

Jeu 1 : Séances 1 et 2



Jeu 2 : Séances 3 et 4



Jeu 3 : Séances 5 et 6

Bijouterie	Coiffeur
Crémier	Jardinier
Epicier	Poissonnerie
Fromagerie	Maquillage

Jeu 4 : Séances 7 et 8



Jeu 5 : Séances 9 et 10



Annexe 4

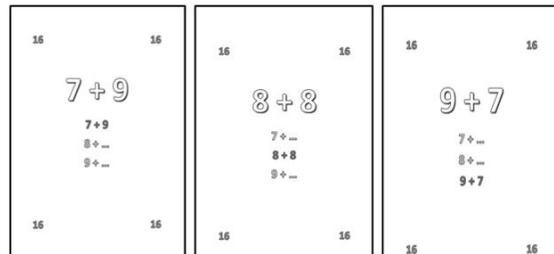
Mots suffixés dictés à la fin de chaque séance des groupes EM et EC-PO

Mot suffixé	Numéro du jeu dans lequel le mot apparaît	Numéro de séance où le mot est dicté
Bijoutier	1, 2, 4, 5	1, 3, 7, 9
Bijouterie	1, 3, 4, 5	2, 6, 8, 10
Poissonnier	1, 2, 3, 5	2, 3, 5, 9
Poissonnerie	1, 2, 3, 4, 5	1, 4, 5, 6, 8, 10
Crémier	1, 3, 4, 5	2, 7, 10
Crèmerie	1, 2, 3, 5	1, 4, 6, 9
Epicier	1, 3, 4, 5	1, 5, 7, 9
Epicerie	1, 2, 4, 5	2, 3, 7, 10
Fromager	1, 2, 3, 4	2, 4, 6, 8
Fromagerie	1, 3, 4, 5	1, 5, 7, 9
Collage	1, 2, 3, 5	1, 3, 5, 9
Jardinier	1, 2, 4, 5	2, 4, 7, 10
Jardinierie	1, 2, 3, 5	1, 3, 5, 9
Jardinage	1, 2, 3, 4	2, 4, 6, 8
Maquilleur	1, 2, 4, 5	1, 3, 7, 9
Maquillage	1, 2, 3, 4	2, 4, 6, 8
Coiffeur	1, 2, 3, 4	1, 3, 5, 7, 9
Coiffure	1, 2, 3, 4, 5	2, 4, 6, 8, 10
Nageur	2, 3, 4, 5	3, 5, 7
Fillette	2, 3, 4, 5	4, 6, 8, 10
Cigarette	2, 3, 4, 5	3, 5, 8, 10
Casquette	2, 3, 4, 5	4, 6, 8, 10

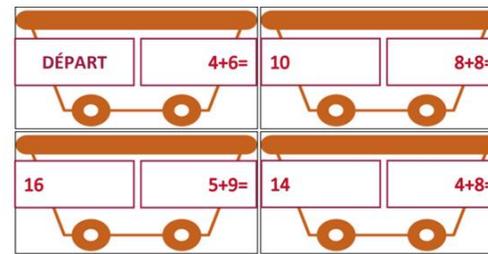
Annexe 5

Exemples des jeux utilisés dans l'entraînement aux mathématiques

Jeu 1 : Séances 1 et 2



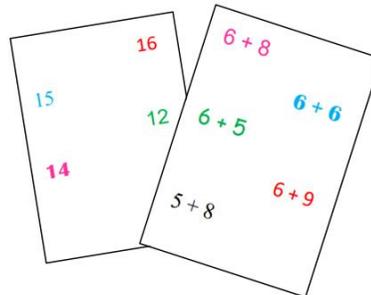
Jeu 2 : Séances 3 et 4



Jeu 3 : Séances 5 et 6



Jeu 4 : Séances 7 et 8



Jeu 5 : Séances 9 et 10



Annexe 6

Exemple de scénario d'une session d'entraînement morphologique

SESSION 4 – SCENARIO		
MORPHO		
Matériel nécessaire :		
- Jeux de wagonnets morpho		
- Feuilles à grands carreaux		
- Un crayon et une gomme personnels		
Durée du jeu : environ 30 minutes		
Durée de l'entraînement orthographique : environ 15 minutes		
Plan de la session		
PHASE DE JEU		
Etape	Description	Temps
1	Saluer les élèves. Leur dire que vous allez utiliser le même jeu que la fois précédente. Leur dire qu'à la fin du jeu, vous ferez des petits exercices pour s'entraîner à orthographier certains mots.	~1m
2	Sortir le jeu de wagonnets morpho. Expliquer que vous allez rejouer avec le même jeu, mais avec des règles différentes, comme à un jeu de domino. Enlever les cartes DEPART et ARRIVEE. Distribuer deux cartes à tous les élèves ; les cartes restantes forment la pioche. Prendre une des cartes restantes, la retourner et la placer au milieu de la table. Jouer la première pour montrer l'exemple aux élèves : on regarde dans son jeu si on peut placer une carte avant ou après la carte posée sur la table pour former un mot existant. Si on peut jouer, on place la carte ; sinon, on pioche et c'est au joueur à notre gauche de jouer. Comme pour la fois précédente, il faudra peut-être rajouter une lettre ou superposer la carte pour former un mot qui existe. Lorsqu'un joueur pose une carte, il doit lire le mot ainsi formé et essayer de donner la signification du mot. Les autres élèves peuvent l'aider s'il n'est pas sûr de lui, mais seulement après qu'il ait formulé une première tentative, ou s'il affirme ne pas connaître la réponse. Si la signification donnée n'est pas la bonne, valoriser la tentative de l'élève, mais lui indiquer que ce n'est pas tout à fait ça. Demander si un autre élève a une autre idée, sinon, définir le mot soi-même.	~5mn

3	Jouer aux dominos avec les élèves jusqu'à ce que toutes les cartes aient été posées, que plus personne ne puisse poser de cartes, ou jusqu'à ce que les 30 minutes soient écoulées. Si on finit le jeu avec au moins 15 minutes d'avance, rejouer. Sinon, passer directement à la phase d'entraînement orthographique.	~15m
PHASE D'ENTRAÎNEMENT ORTHOGRAPHIQUE		
Etape	Description	Temps
1	Leur expliquer qu'on va maintenant s'entraîner à écrire certains des mots qu'ils viennent de voir en faisant une dictée de 9 mots. Leur dire qu'il faut écrire si possible sans faire de faute, en essayant de se rappeler des cartes.	~3 mn
2	<p>Dictier les mots suivants. Après que chaque mot a été dicté, vérifier que tous les enfants l'aient bien orthographié. Si ce n'est pas le cas, leur demander de gommer et leur donner la bonne orthographe (ne pas laisser une fausse orthographe, la corriger immédiatement)</p> <ol style="list-style-type: none"> 1. Poissonnerie 2. Crèmerie 3. Fromager 4. Jardinier 5. Jardinage 6. Maquillage 7. Coiffure 8. Fillette 9. Casquette 	~12 mn
3	Finir la session en remerciant les enfants pour leurs efforts et leur aide. Leur rappeler la date à laquelle on les reverra, et ce qu'on fera la prochaine fois : un autre jeu avec des mots, une petite dictée comme celle qu'ils viennent de faire.	~1 mn

Annexe 7

Items de la DIPHO

Item	Entraîné / Non-entraîné	Régularité
Hôpitaux	Entraîné	AUX pluriel
Vitraux	Entraîné	AUX pluriel
Royaux	Entraîné	AUX pluriel
Promo	Entraîné	O apocope
Radio	Entraîné	O apocope
Magnéto	Entraîné	O apocope
Loto	Entraîné	O étranger
Judo	Entraîné	O étranger
Domino	Entraîné	O étranger
Bureau	Entraîné	EAU règle générale
Manteau	Entraîné	EAU règle générale
Corbeau	Entraîné	EAU règle générale
Normaux	Non-entraîné	AUX pluriel
Coraux	Non-entraîné	AUX pluriel
Locaux	Non-entraîné	AUX pluriel
Info	Non-entraîné	O apocope
Zoo	Non-entraîné	O apocope
Studio	Non-entraîné	O étranger
Mikado	Non-entraîné	O étranger
Couteau	Non-entraîné	EAU règle générale
Morceau	Non-entraîné	EAU règle générale
Museau	Non-entraîné	EAU règle générale

Annexe 8

Items de la MEREP

Dictée A

Item	Régularité
Végétaux	AUX pluriel
Amicaux	AUX pluriel
Signaux	AUX pluriel
Frigo	O apocope
Métro	O apocope
Brio	O étranger
Tango	O étranger
Casino	O étranger
Poireau	EAU règle générale
Chapiteau	EAU règle générale
Anneau	EAU règle générale

Dictée B

Item	Régularité
Régionaux	AUX pluriel
Cristaux	AUX pluriel
Bocaux	AUX pluriel
Kilo	O apocope
Météo	O apocope
Poncho	O étranger
Rodéo	O étranger
Duo	O étranger
Berceau	EAU règle générale
Panneau	EAU règle générale
Carreau	EAU règle générale

Dictée C

Item	Régularité
Nationaux	AUX pluriel
Spéciaux	AUX pluriel
Métaux	AUX pluriel
Stylo	O apocope
Chrono	O apocope
Solo	O étranger

Diabolo	O étranger
Flamenco	O étranger
Poteau	EAU règle générale
Râteau	EAU règle générale
Radeau	EAU règle générale

Annexe 9

Exemple de scénario d'une session d'entraînement orthographique sublexical explicite

SESSION 4 : Travail sur les pluriels

SCENARIO

Matériel nécessaire :

- Fichier « 4. Travail sur les pluriels »
- Jeu Tangram
- Pièces d'or
- Feuille mesures répétées
- Feuille et crayon ou stylo effaçable
- Gel hydroalcoolique, masque

Durée de la phase d'entraînement : 20 à 25 minutes

Durée des mesures répétées : 5 à 10 minutes

Plan de la session

PHASE D'ENTRAINEMENT		
Etape	Description	Temps
1	Saluer l'enfant. Si besoin, lui réexpliquer qu'on va faire plusieurs exercices d'orthographe qui visent à aider les enfants qui ont des difficultés. Lui redire qu'à la fin du jeu, on fera une petite dictée (11 mots) avec lui pour voir comment il fait pour orthographier.	<1mn
2	Lui demander s'il se souvient de ce qu'on a fait la dernière fois, l'aider à se rappeler qu'on a vu des mots qui n'avaient qu'une seule syllabe. Dire : « à partir de maintenant, on va uniquement travailler sur des mots avec au moins 2 syllabes. Aujourd'hui, on va voir des mots qui ont comme point commun d'être des pluriels. Est-ce que tu sais ce qu'est, un pluriel ? Un pluriel , c'est tout simplement un mot qui désigne plusieurs choses. Souvent, en français, le pluriel ne se prononce pas différemment du singulier. Par exemple, si je te dis « un chat », et si je te dis « des chats », le mot « chat », il va se prononcer pareil, qu'il soit au singulier ou au pluriel. A l'oral, on ne va pas entendre la différence, mais par contre, à l'écrit, on va écrire un -s à la fin de « chat » pour expliquer que le mot est au pluriel , qu'on ne parle pas d'un seul chat mais de plusieurs chats. Il y a quand même des pluriels qui peuvent s'entendre en français, et c'est justement ceux qu'on va voir aujourd'hui. Tu as peut-être déjà entendu qu'on dit « un cheval, des chevaux » ? Ce sont ces pluriels-là qu'on va voir aujourd'hui.	4mn

3	Désigner la colonne de gauche. Dire : « Ici, tu as des mots qui sont au pluriel. Lis le premier mot à haute voix (« chevaux ») puis essaye de deviner ce que donne ce mot au singulier. Essaie d'abord de trouver à l'oral, puis trace un trait entre le mot au pluriel et le même mot au singulier. » Effectuer la même chose pour tous les mots de la colonne de gauche.	3mn
4	Dire : « Tu as vu, tous les mots qui finissent par le son /o/ ici, ils s'écrivent de la même manière. Ils s'écrivent tous A-U-X, et les singuliers, ils s'écrivent tous soit A-L soit A-I-L. C'est ça, la première règle qui va nous aider pour écrire le son /o/ à la fin des mots. Si tu dois écrire un mot qui finit par le son /o/, ce que tu peux faire en tout premier, c'est voir si tu peux remplacer le son /o/ par « al » ou par « ail ». Par exemple, pour journaux, ça donne soit « journal » soit « journalil ». Comme « journal », c'est un mot qui existe, tu es sûr que journaux, il va s'écrire A-U-X à la fin. C'est obligatoire, tu ne peux pas te tromper. » Dire : « Maintenant, je vais te demander d'écrire ces mots. Avant de les écrire, dis-moi ce qu'ils donnent au singulier. Quand tu les écris, rappelle-toi, comme ce sont des pluriels de mots en « al » ou « ail », ils vont TOUJOURS s'écrire avec « aux » à la fin. A chaque fois que tu réussiras à en écrire un correctement, tu gagneras une pièce. Si tu parviens à gagner toutes les pièces, tu pourras jouer avec le Tangram jusqu'à la fin de la séance. »	~5mn
5	Le laisser jouer au Tangram de façon à ce qu'il reste du temps pour réaliser les mesures répétées (environ 10 minutes pour les CE1 et entre 5 et 10 minutes pour les CM1/6e en fonction de leur niveau.)	Entre 5 et 10mn
PHASE DE MESURES REPETEES		
Etape	Description	Temps
1	Lui expliquer qu'on va maintenant faire une petite dictée pour voir comment il fait pour orthographier.	~1 mn
2	Avant de dicter les mots, dire : « certains des mots de la dictée vont s'écrire comme les mots qu'on a vus aujourd'hui, mais pas tous. C'est à toi de voir si on peut les écrire comme ça ou non. » Faire passer la dictée de mesure répétée correspondant à l'ordre indiqué sur la feuille.	Entre 3 et 7 mn
3	Finir la session en remerciant l'enfant pour ses efforts et son aide. Lui rappeler la date à laquelle on le reverra, et ce qu'on fera la prochaine fois : un autre jeu avec des mots, et quelques petits exercices comme ceux qu'il vient de faire.	~1 mn

Annexe 10

Inventaire des graphies existantes pour transcrire le phonème /o/ en fin de mot

EAU – O – AUX – AUD – AU – AUT

OS – OT – OC – OP – ÔT

Annexe 11

Support imagé pour la présentation des mots étrangers finissant par le phonème /o/



0



Annexe 12

Devinettes utilisées pour travailler la terminaison -EAU

Bateau	Tableau	Chapeau	Nouveau
<p>C'est un mot avec deux syllabes.</p> <p>Il finit par les lettres « E – A – U ».</p> <p>C'est un moyen de transport.</p> <p>Il peut avoir un moteur ou des rames.</p> <p>Il sert à voyager sur l'eau.</p>	<p>C'est un mot avec deux syllabes.</p> <p>Il finit par les lettres « E – A – U ».</p> <p>C'est une œuvre d'art.</p> <p>Il peut être blanc ou noir.</p> <p>Le maître ou la maîtresse de classe écrit dessus.</p>	<p>C'est un mot avec deux syllabes.</p> <p>Il finit par les lettres « E – A – U ».</p> <p>C'est un vêtement.</p> <p>Il peut être en paille ou en tissu.</p> <p>On le porte sur la tête pour se protéger du soleil.</p>	<p>C'est un mot avec deux syllabes.</p> <p>Il finit par les lettres « E – A – U ».</p> <p>C'est un synonyme de « neuf ».</p> <p>C'est le contraire d'« ancien » ou de « vieux ».</p> <p>Quand quelqu'un vient d'arriver dans un endroit, on dit que c'est le petit n...</p>

Annexe 13

Exemple de « famille » utilisée pour présenter les exceptions à la règle générale -EAU



Annexe 14

Valeurs de niveau et de pente pour chaque participant entre la phase A et B

	Immédiateté					
	Médiane (niveau)		SMTL (tendance)		Médiane (niveau)	SMTL (tendance)
	A	B	A (pente)	B (pente)	B (3 premiers points)	B (pente ; 3 premiers points)
2^eNL						
EE						
S25	4	4	1	0	4	1
S38	3	5	1	0.25	5	0
S43	5	7	0.5	0.75	6	0.5
S52	2	7	1	0.25	6	0.5
S53	3	7	0	0	7	2
S58	1	6	1.5	0.25	5	2.5
S64	2	5	0	-0.5	5	-0.5
S65	3	6	-0.5	0.5	4	1.5
EI						
S40	2	4	0	0.5	3	0.5
S46	4	3	-1	0.5	3	0.5
S48	3	3	-0.5	0.25	3	0
S50	3	3	0.5	-0.25	4	0.5
S55	2	3	0.5	0.25	3	1.5
S61	3	5	-0.5	0.75	4	1.5
2^eTL						
EE						
S44	2	3	0.5	0	3	0.5
S47	1	3	0	0.25	3	-1.5
EI						
S60	3	3	-1	-0.5	4	1.5
4^eNL						
EE						
S9	8	9	1	0.25	9	0
S16	9	9	-0.5	0	9	-0.5
S28	7	8	0.5	0.5	8	0
S36	6	8	0	0.75	6	-1
S45	5	6	-2	-0.25	6	-0.5

Annexe 14 (suite)

					Immédiateté	
	Médiane (niveau)		SMTL (tendance)		Médiane (niveau)	SMTL (tendance)
	A	B	A (pente)	B (pente)	B (3 premiers points)	B (pente ; 3 premiers points)
4^eNL						
EI						
S5	4	5	1	0.25	5	-0.5
S8	4	6	-0.5	0.25	5	-1
S27	7	7	0.5	0.5	7	0.5
S35	5	9	0	0.5	8	0.5
S51	4	7	-0.5	0.5	6	-0.5
S62	7	9	-1	0.5	8	-0.5
4^eTL						
EE						
S1	6	7	0.5	0.25	6	-1
S29	3	5	-1	0.5	4	0
S42	6	9	0	0.5	7	0
S54	5	6	1	0	6	0.5
EI						
S12	4	4	-0.5	0.25	4	-1
S37	5	8	0	0.75	6	0.5
S39	4	4	-1.5	0.25	4	1
S41	3	5	0.5	0.75	4	0.5
6^eNL						
EE						
S6	7	8	0.5	0.25	7	1
S20	6	8	0	1	5	0
S57	9	9	0	-0.25	10	0.5
EI						
S21	9	10	0.5	-0.25	10	-0.5
S59	10	10	-1	0	10	0
6^eTL						
EE						
S34	8	7	0	-0.5	7	-0.5
EI						
S4	8	9	-0.5	0.25	8	1
S11	2	4	1.5	0.75	4	1
S32	6	7	1	-0.5	8	-0.5

EE = Groupe entraîné explicitement. EI = Groupe entraîné implicitement. NL = Normo-lecteurs.
 SMTL = *split-middle trend line*. TL = Trouble de la lecture.

Références bibliographiques

- Abbott, S. P., & Berninger, V. W. (1999). It's never too late to remediate : Teaching word recognition to students with reading disabilities in grades 4–7. *Annals of Dyslexia*, 49(1), 221-250. <https://doi.org/10.1007/s11881-999-0025-x>
- Académie française. (1932). *Dictionnaire de l'Académie française* (Huitième édition, 1-2). Librairie Hachette. <http://gallica.bnf.fr/ark:/12148/bpt6k1280382r>
- Allen, A. A., & Lembke, E. S. (2020). The Effect of a Morphological Awareness Intervention on Early Writing Outcomes. *Learning Disability Quarterly*. <https://doi.org/10.1177/0731948720912414>
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition). American Psychiatric Association. <https://doi.org/10.1176/appi.books.9780890425596>
- Amirjalili, F., Jabbari, A. A., & Rezai, M. J. (2018). The effect of explicit instruction on derivational morphological awareness amongst Iranian EFL learners. *Linguistic Research*, 35, 47-82. <https://doi.org/10.17250/KHISLI.35..201809.003>
- Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval : Resolving neighborhood conflicts. *Psychonomic Bulletin & Review*, 4(4), 439-461. <https://doi.org/10.3758/BF03214334>
- Andrews, S., & Scarratt, D. R. (1998). Rule and analogy mechanisms in reading nonwords : Hough dou peapel rede gnew wirds? *Journal of Experimental Psychology: Human Perception and Performance*, 24(4), 1052-1086. <https://doi.org/10.1037/0096-1523.24.4.1052>
- Anthony, J. L., Lonigan, C. J., Driscoll, K., Phillips, B. M., & Burgess, S. R. (2003). Phonological sensitivity : A quasi-parallel progression of word structure units and cognitive operations. *Reading Research Quarterly*, 38(4), 470-487. <https://doi.org/10.1598/RRQ.38.4.3>
- Apel, K. (2011). What is orthographic knowledge? *Language, Speech, and Hearing Services in Schools*, 42(4), 592-603. [https://doi.org/10.1044/0161-1461\(2011/10-0085\)](https://doi.org/10.1044/0161-1461(2011/10-0085))
- Apel, K., & Diehm, E. (2014). Morphological Awareness Intervention With Kindergarteners and First and Second Grade Students From Low SES Homes : A Small Efficacy Study. *Journal of Learning Disabilities*, 47(1), 65-75. <https://doi.org/10.1177/0022219413509964>
- Apel, K., Diehm, E., & Apel, L. (2013). Using Multiple Measures of Morphological Awareness to Assess its Relation to Reading. *Topics in Language Disorders*, 33(1), 42-56. <https://doi.org/10.1097/TLD.0b013e318280f57b>
- Apel, K., & Henbest, V. S. (2016). Affix Meaning Knowledge in First Through Third Grade Students. *Language, Speech, and Hearing Services in Schools*, 47(2), 148-156. https://doi.org/10.1044/2016_LSHSS-15-0050
- Ardanouy, E., Zesiger, P., & Delage, H. (2023). Intensive and explicit derivational morphology training in school-aged children : An effective way to improve morphological awareness, spelling and reading? *Reading and Writing*. <https://doi.org/10.1007/s11145-023-10454-y>
- Arfé, B., Cona, E., & Merella, A. (2018). Training Implicit Learning of Spelling in Italian Children With Developmental Dyslexia. *Topics in Language Disorders*, 38(4), 299-315. <https://doi.org/10.1097/TLD.000000000000163>
- Arnbak, E., & Elbro, C. (2000). The Effects of Morphological Awareness Training on the Reading and Spelling Skills of Young Dyslexics. *Scandinavian Journal of Educational Research*, 44(3), 229-251. <https://doi.org/10.1080/00313830050154485>

- Bahr, R. H., Sillian, E. R., Berninger, V. W., & Dow, M. (2012). Linguistic pattern analysis of misspellings of typically developing writers in grades 1-9. *Journal of Speech, Language, and Hearing Research: JSLHR*, 55(6), 1587-1599. [https://doi.org/10.1044/1092-4388\(2012/10-0335\)](https://doi.org/10.1044/1092-4388(2012/10-0335))
- Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R : A practical tutorial. *Evidence Based Mental Health*, 22(4), 153-160. <https://doi.org/10.1136/ebmental-2019-300117>
- Banfi, C., Kemény, F., Gangl, M., Schulte-Körne, G., Moll, K., & Landerl, K. (2018). Visual attention span performance in German-speaking children with differential reading and spelling profiles : No evidence of group differences. *PLOS ONE*, 13(6), e0198903. <https://doi.org/10.1371/journal.pone.0198903>
- Bar-Kochva, I., & Hasselhorn, M. (2017). The Training of Morphological Decomposition in Word Processing and Its Effects on Literacy Skills. *Frontiers in Psychology*, 8, 1583. <https://doi.org/10.3389/fpsyg.2017.01583>
- Bar-Kochva, I., Korinth, S. P., & Hasselhorn, M. (2020). Effects of a morpheme-based training procedure on the literacy skills of readers with a reading disability. *Applied Psycholinguistics*, 41(5), 1061-1082. <https://doi.org/10.1017/S0142716420000120>
- Barron, R. W., & Baron, J. (1977). How Children Get Meaning from Printed Words. *Child Development*, 48(2), 587-594. <https://doi.org/10.2307/1128657>
- Beauvois, M. F., & Dérouesné, J. (1979). Phonological alexia : Three dissociations. *Journal of Neurology, Neurosurgery & Psychiatry*, 42(12), 1115-1124. <https://doi.org/10.1136/jnnp.42.12.1115>
- Berko, J. (1958). The Child's Learning of English Morphology. *WORD*, 14(2-3), 150-177. <https://doi.org/10.1080/00437956.1958.11659661>
- Berninger, V. W. (1987). Global, component, and serial processing of printed words in beginning reading. *Journal of Experimental Child Psychology*, 43(3), 387-418. [https://doi.org/10.1016/0022-0965\(87\)90015-4](https://doi.org/10.1016/0022-0965(87)90015-4)
- Berninger, V. W., Abbott, R. D., Nagy, W., & Carlisle, J. (2010). Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *Journal of Psycholinguistic Research*, 39(2), 141-163. <https://doi.org/10.1007/s10936-009-9130-6>
- Berninger, V. W., Nagy, W. E., Carlisle, J., Thomson, J., Hoffer, D., & Abbott, S. (2003). Effective treatment for children with dyslexia in grades 4-6 : Behavioral and brain evidence. In *Preventing and remediating reading difficulties : Bringing science to scale* (B. R. Foorman, p. 381-417). Baltimore: York Press.
- Berninger, V. W., Winn, W. D., Stock, P., Abbott, R. D., Eschen, K., Lin, S.-J. (Cindy), Garcia, N., Anderson-Youngstrom, M., Murphy, H., Lovitt, D., Trivedi, P., Jones, J., Amtmann, D., & Nagy, W. (2008). Tier 3 specialized writing instruction for students with dyslexia. *Reading and Writing*, 21(1), 95-129. <https://doi.org/10.1007/s11145-007-9066-x>
- Berthiaume, R., Besse, A.-S., & Daigle, D. (2010). L'évaluation de la conscience morphologique : Proposition d'une typologie des tâches. *Language Awareness*, 19, 153-170. <https://doi.org/10.1080/09658416.2010.482992>
- Besner, D., & Smith, M. C. (1992). Chapter 3 Basic Processes in Reading : Is the Orthographic Depth Hypothesis Sinking? In R. Frost & L. Katz (Éds.), *Advances in Psychology* (Vol. 94, p. 45-66). North-Holland. [https://doi.org/10.1016/S0166-4115\(08\)62788-0](https://doi.org/10.1016/S0166-4115(08)62788-0)

- Beyersmann, E., Ziegler, J. C., Castles, A., Coltheart, M., Kezilas, Y., & Grainger, J. (2016). Morpho-orthographic segmentation without semantics. *Psychonomic Bulletin & Review*, 23(2), 533-539. <https://doi.org/10.3758/s13423-015-0927-z>
- Biederman, I., & Cooper, E. E. (2009). Translational and reflectional priming invariance : A retrospective. *Perception*, 38(6), 809-817. <https://doi.org/10.1068/pmkbie>
- Bissonnette, S., Richard, M., Gauthier, C., & Bouchard, C. (2010). Quelles sont les stratégies d'enseignement efficaces favorisant les apprentissages fondamentaux auprès des élèves en difficulté de niveau élémentaire? Résultats d'une mégaanalyse. *Revue de recherche appliquée sur l'apprentissage*, 3, 1-35.
- Blachman, B. A., Schatschneider, C., Fletcher, J. M., Murray, M. S., Munger, K. A., & Vaughn, M. G. (2014). Intensive Reading Remediation in Grade 2 or 3 : Are There Effects a Decade Later? *Journal of Educational Psychology*, 106(1), 46-57. <https://doi.org/10.1037/a0033663>
- Bonin, P., Collay, S., & Fayol, M. (2008). La consistance orthographique en production verbale écrite : Une brève synthèse. *L'Année psychologique*, 108(3), 517-546.
- Booij, G. (2015). Morphology : The structure of words. In *The Routledge Handbook of Linguistics* (Keith Allan). Routledge.
- Bosse, Chaves, N., Largy, P., & Valdois, S. (2015). Orthographic learning during reading : The role of whole-word visual processing: orthographic learning and visual processing. *Journal of Research in Reading*, 38(2), 141-158. <https://doi.org/10.1111/j.1467-9817.2012.01551.x>
- Bosse, M.-L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia : The visual attention span deficit hypothesis. *Cognition*, 104(2), 198-230. <https://doi.org/10.1016/j.cognition.2006.05.009>
- Bourassa, D., & Treiman, R. (2003). Spelling in Children With Dyslexia : Analyses From the Treiman-Bourassa Early Spelling Test. *Scientific Studies of Reading*, 7(4), 309-333. https://doi.org/10.1207/S1532799XSSR0704_1
- Bowers. (2006). *Adding transparency to morphologically opaque words through instruction* [Master's dissertation in Education]. Queen's University.
- Bowers, J. S. (2000). In defense of abstractionist theories of repetition priming and word identification. *Psychonomic Bulletin & Review*, 7(1), 83-99. <https://doi.org/10.3758/BF03210726>
- Bowers, J. S. (2020). Reconsidering the Evidence That Systematic Phonics Is More Effective Than Alternative Methods of Reading Instruction. *Educational Psychology Review*, 32(3), 681-705. <https://doi.org/10.1007/s10648-019-09515-y>
- Bowers, J. S., & Bowers, P. N. (2017). Beyond Phonics : The Case for Teaching Children the Logic of the English Spelling System. *Educational Psychologist*, 52(2), 124-141. <https://doi.org/10.1080/00461520.2017.1288571>
- Bowers, Kirby, J. R., & Deacon, S. H. (2010). The Effects of Morphological Instruction on Literacy Skills : A Systematic Review of the Literature. *Review of Educational Research*, 80(2), 144-179. <https://doi.org/10.3102/0034654309359353>
- Bowey, J. A., & Muller, D. (2005). Phonological recoding and rapid orthographic learning in third-graders' silent reading : A critical test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, 92, 203-219. <https://doi.org/10.1016/j.jecp.2005.06.005>

- Bradley, L., & Bryant, P. E. (1979). Independence of Reading and Spelling in Backward and Normal Readers. *Developmental Medicine & Child Neurology*, 21(4), 504-514. <https://doi.org/10.1111/j.1469-8749.1979.tb01655.x>
- Brimo, D. (2016). Evaluating the Effectiveness of a Morphological Awareness Intervention : A Pilot Study. *Communication Disorders Quarterly*, 38(1), 35-45. <https://doi.org/10.1177/1525740115604592>
- Bruck, M., & Treiman, R. (1990). Phonological awareness and spelling in normal children and dyslexics : The case of initial consonant clusters. *Journal of Experimental Child Psychology*, 50(1), 156-178. [https://doi.org/10.1016/0022-0965\(90\)90037-9](https://doi.org/10.1016/0022-0965(90)90037-9)
- Bryant, P. E., & Bradley, L. (1980). Why children sometimes write words which they do not read. In *Cognitive Processes in Spelling* (U. Frith). Academic Press.
- Bryson, S. E., & Werker, J. F. (1989). Toward understanding the problem in severely disabled readers Part 1 : Vowel errors. *Applied Psycholinguistics*, 10(1), 1-12. <https://doi.org/10.1017/S0142716400008389>
- Budoff, M., & Quinlan, D. (1964). Auditory and visual learning in primary grade children. *Child Development*, 35, 583-586. <https://doi.org/10.2307/1126722>
- Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language and Cognitive Processes*, 2(3-4), 217-227. <https://doi.org/10.1080/01690968708406932>
- Burani, C., Marcolini, S., De Luca, M., & Zoccolotti, P. (2008). Morpheme-based reading aloud : Evidence from dyslexic and skilled Italian readers. *Cognition*, 108(1), 243-262. <https://doi.org/10.1016/j.cognition.2007.12.010>
- Burt, J. S., & Tate, H. (2002). Does a Reading Lexicon Provide Orthographic Representations for Spelling? *Journal of Memory and Language*, 46(3), 518-543. <https://doi.org/10.1006/jmla.2001.2818>
- Burton, L., Nunes, T., & Evangelou, M. (2021). Do children use logic to spell logician? Implicit versus explicit teaching of morphological spelling rules. *The British Journal of Educational Psychology*, e12414. <https://doi.org/10.1111/bjep.12414>
- Bus, A. G., & van IJzendoorn, M. H. (1999). Phonological awareness and early reading : A meta-analysis of experimental training studies. *Journal of Educational Psychology*, 91(3), 403-414. <https://doi.org/10.1037/0022-0663.91.3.403>
- Butterworth, B. (1983). *Lexical representation*. na.
- Byrne, B., & Fielding-Barnsley, R. (1989). Phonemic awareness and letter knowledge in the child's acquisition of the alphabetic principle. *Journal of Educational Psychology*, 81(3), 313-321. <https://doi.org/10.1037/0022-0663.81.3.313>
- Calhoun, M. B., Sandow, A., & Hunter, C. V. (2010). Reorganizing the instructional reading components : Could there be a better way to design remedial reading programs to maximize middle school students with reading disabilities' response to treatment? *Annals of Dyslexia*, 60(1), 57-85. <https://doi.org/10.1007/s11881-009-0033-x>
- Campbell, R. (1983). Writing nonwords to dictation. *Brain and Language*, 19(1), 153-178. [https://doi.org/10.1016/0093-934x\(83\)90061-5](https://doi.org/10.1016/0093-934x(83)90061-5)
- Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, 28(3), 297-332.

- Carioti, D., Masia, M. F., Travellini, S., & Berlinger, M. (2021). Orthographic depth and developmental dyslexia : A meta-analytic study. *Annals of Dyslexia*, 71(3), 399-438. <https://doi.org/10.1007/s11881-021-00226-0>
- Carlisle, J. F. (1988). Knowledge of derivational morphology and spelling ability in fourth, sixth, and eighth graders. *Applied Psycholinguistics*, 9(3), 247-266. <https://doi.org/10.1017/S0142716400007839>
- Carlisle, J. F. (1995). Morphological awareness and early reading achievement. In *Morphological aspects of language processing* (p. 189-209). Lawrence Erlbaum Associates, Inc.
- Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words : Impact on reading. *Reading and Writing*, 12(3), 169-190. <https://doi.org/10.1023/A:1008131926604>
- Carlisle, J. F., McBride-Chang, C., Nagy, W., & Nunes, T. (2010). Effects of Instruction in Morphological Awareness on Literacy Achievement : An Integrative Review. *Reading Research Quarterly*, 45(4), 464-487. <https://doi.org/10.1598/RRQ.45.4.5>
- Carlisle, J. F., & Stone, C. A. (2005). Exploring the role of morphemes in word reading. *Reading Research Quarterly*, 40(4), 428-449. <https://doi.org/10.1598/RRQ.40.4.3>
- Carrillo, M. S., Alegría, J., & Marín, J. (2013). On the acquisition of some basic word spelling mechanisms in a deep (French) and a shallow (Spanish) system. *Reading and Writing*, 26(6), 799-819. <https://doi.org/10.1007/s11145-012-9391-6>
- Casalis, S., & Colé, P. (2009). On the relationship between morphological and phonological awareness : Effects of training in kindergarten and in first-grade reading. *First Language*, 29(1), 113-142. <https://doi.org/10.1177/0142723708097484>
- Casalis, S., Deacon, S. H., & 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>
- Casalis, S., & Louis-Alexandre, M.-F. (2000). Morphological analysis, phonological analysis and learning to read French : A longitudinal study. *Reading and Writing*, 12(3), 303-335. <https://doi.org/10.1023/A:1008177205648>
- Casalis, S., Pacton, S., Lefevre, F., & Fayol, M. (2017). Morphological training in spelling : Immediate and long-term effects of an interventional study in French third graders. *Learning and Instruction*, 53, 89-98. <https://doi.org/10.1016/j.learninstruc.2017.07.009>
- Casalis, S., Quémart, P., & Duncan, L. G. (2015). How language affects children's use of derivational morphology in visual word and pseudoword processing : Evidence from a cross-language study. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00452>
- Cassar, M., & Treiman, R. (1997). The beginnings of orthographic knowledge : Children's knowledge of double letters in words. *Journal of Educational Psychology*, 89(4), 631-644. <https://doi.org/10.1037/0022-0663.89.4.631>
- Castles, A., & Nation, K. (2006). How does orthographic learning happen? In *From inkmarks to ideas : Current issues in lexical processing* (p. 151-179). Psychology Press.
- Castles, A., Rastle, K., & Nation, K. (2018). Ending the Reading Wars : Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest*, 19(1), 5-51. <https://doi.org/10.1177/1529100618772271>
- Chaney, C. (1992). Language development, metalinguistic skills, and print awareness in 3-year-old children. *Applied Psycholinguistics*, 13(4), 485-514. <https://doi.org/10.1017/S0142716400005774>

- Chow, B. W.-Y., McBride-Chang, C., Cheung, H., & Chow, C. S.-L. (2008). Dialogic reading and morphology training in Chinese children : Effects on language and literacy. *Developmental Psychology*, 44(1), 233-244. <https://doi.org/10.1037/0012-1649.44.1.233>
- Christensen, C. A., & Bowey, J. A. (2005). The Efficacy of Orthographic Rime, Grapheme–Phoneme Correspondence, and Implicit Phonics Approaches to Teaching Decoding Skills. *Scientific Studies of Reading*, 9(4), 327-349. https://doi.org/10.1207/s1532799xssr0904_1
- Clark, E. V. (1998). Lexical creativity in French-speaking children. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 17(2), 513-530.
- Cohen, Feinstein, A., Masuda, A., & Vowles, K. E. (2014). Single-Case Research Design in Pediatric Psychology : Considerations Regarding Data Analysis. *Journal of Pediatric Psychology*, 39(2), 124-137. <https://doi.org/10.1093/jpepsy/jst065>
- Colchester, E. J. R. (2010). *Can we improve children's spelling ability by teaching morphemes through text reading ? An intervention study exploring the relationship between morphological awareness and literacy* [Doctoral dissertation in Professional Educational, Child and Adolescent Psychology]. Institute of Education, University of London.
- Colé, P., Bouton, S., Leuwers, C., Casalis, S., & Sprenger-Charolles, L. (2012). Stem and derivational-suffix processing during reading by French second and third graders. *Applied Psycholinguistics*, 33(1), 97-120. <https://doi.org/10.1017/S0142716411000282>
- Colenbrander, D., Parsons, L., Bowers, J. S., & Davis, C. J. (2021). Assessing the Effectiveness of Structured Word Inquiry for Students in Grades 3 and 5 With Reading and Spelling Difficulties : A Randomized Controlled Trial. *Reading Research Quarterly*, rrq.399. <https://doi.org/10.1002/rrq.399>
- Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language*, 15(3), 245-286.
- Coltheart, M., Davelaar, E., Jonasson, J. F., & Besner, D. (1977). Access to the internal lexicon. In *Attention and performance VI* (S. Dornic, p. 535-555). Erlbaum.
- Coltheart, M., Langdon, R., & Haller, M. (1996). Computational cognitive neuropsychology. In *Evaluating theories of language : Evidence from disordered communication* (B. Dodd, R. Campbell, L. Worrall, p. 9-36). Whurr Publishers.
- Coltheart, M., Masterson, J., Byng, S., Prior, M., & Riddoch, J. (1983). Surface dyslexia. *The Quarterly Journal of Experimental Psychology Section A*, 35(3), 469-495. <https://doi.org/10.1080/14640748308402483>
- Coltheart, M., & Rastle, K. (1999). Serial and Strategic Effects in Reading Aloud. *Journal of Experimental Psychology – Human Perception and Performance*, 25. <https://doi.org/10.1037//0096-1523.25.2.482>
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC : A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204-256. <https://doi.org/10.1037/0033-295x.108.1.204>
- Commissaire, E., & Besse, A.-S. (2019). Investigating lexical and sub-lexical orthographic processing skills in French 3rd and 5th graders. *Journal of Research in Reading*, 42(2), 268-287. <https://doi.org/10.1111/1467-9817.12268>
- Commissaire, E., Duncan, L. G., & Casalis, S. (2011). Cross-language transfer of orthographic processing skills : A study of French children who learn English at school. *Journal of Research in Reading*, 34(1), 59-76. <https://doi.org/10.1111/j.1467-9817.2010.01473.x>

- Condry, S. M., McMahon-Rideout, M., & Levy, A. A. (1979). A developmental investigation of selective attention to graphic, phonetic, and semantic information in words. *Perception & Psychophysics*, 25, 88-94. <https://doi.org/10.3758/BF03198791>
- Conners, Loveall, Moore, Hume, & Maddox. (2011). An individual differences analysis of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, 108(2). <https://doi.org/10.1016/j.jecp.2010.09.009>
- Conrad, N. J., Harris, N., & Williams, J. (2013). Individual differences in children's literacy development : The contribution of orthographic knowledge. *Reading and Writing: An Interdisciplinary Journal*, 26(8), 1223-1239. <https://doi.org/10.1007/s11145-012-9415-2>
- Content, A. (1991). The effect of spelling-to-sound regularity on naming in French. *Psychological Research*, 53(1), 3-12. <https://doi.org/10.1007/BF00867327>
- Content, A., & Zesiger, P. (2005). *L'acquisition du langage écrit* (p. 1-46).
- Cooke, A. (2001). Critical response to dyslexia, literacy and psychological assessment. (Report by a working party of the division of educational and child psychology of the British psychological society). A view from the chalk face. *Dyslexia (Chichester, England)*, 7(1), 47-52. <https://doi.org/10.1002/dys.181>
- Cordewener, K. A. H., Verhoeven, L., & Bosman, A. M. T. (2016). Improving Spelling Performance and Spelling Consciousness. *The Journal of Experimental Education*, 84(1), 48-74. <https://doi.org/10.1080/00220973.2014.963213>
- Crosson, A. C., McKeown, M. G., Lei, P., Zhao, H., Li, X., Patrick, K., Brown, K., & Shen, Y. (2021). Morphological analysis skill and academic vocabulary knowledge are malleable through intervention and may contribute to reading comprehension for multilingual adolescents. *Journal of Research in Reading*, 44(1), 154-174. <https://doi.org/10.1111/1467-9817.12323>
- Cunningham. (2006). Accounting for children's orthographic learning while reading text : Do children self-teach? *Journal of experimental child psychology*, 95, 56-77. <https://doi.org/10.1016/j.jecp.2006.03.008>
- Cunningham, A. E., Perry, K. E., & Stanovich, K. E. (2001). Converging evidence for the concept of orthographic processing. *Reading and Writing*, 14(5), 549-568. <https://doi.org/10.1023/A:1011100226798>
- Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Share, D. L. (2002). Orthographic learning during reading : Examining the role of self-teaching. *Journal of Experimental Child Psychology*, 82, 185-199. [https://doi.org/10.1016/S0022-0965\(02\)00008-5](https://doi.org/10.1016/S0022-0965(02)00008-5)
- Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children : A quick measure of reading experience. *Journal of Educational Psychology*, 82(4), 733-740. <https://doi.org/10.1037/0022-0663.82.4.733>
- Daffern, T., Mackenzie, N. M., & Hemmings, B. (2015). The development of a spelling assessment tool informed by Triple Word Form Theory. *Australian Journal of Language and Literacy*, The, 38(2), 72.
- Dallasheh-Khatib, R., Ibrahim, R., & Karni, A. (2014). Longitudinal Data on the Relations of Morphological and Phonological Training to Reading Acquisition in First Grade : The Case of Arabic Language. *Psychology*, 05(08), 918-940. <https://doi.org/10.4236/psych.2014.58103>
- Danjon, J., & Pacton, S. (2009). Children's learning about properties of double letters : The case of French. *Presented at the 16th European Society for Cognitive Psychology Conference*.

- Darch, C., Eaves, R. C., Crowe, D. A., Simmons, K., & Conniff, A. (2006). Teaching Spelling to Students with Learning Disabilities : A Comparison of Rule-Based Strategies versus Traditional Instruction. *Journal of Direct Instruction*, 6(1), 1-16.
- Davidson, S. J., & O'Connor, R. E. (2019). An intervention using morphology to derive word meanings for English language learners. *Journal of Applied Behavior Analysis*, 52(2), 394-407. <https://doi.org/10.1002/jaba.539>
- Davis, R. D. (1997). *The gift of dyslexia*. Souvenir.
- Deacon, H., Comissaire, E., Chen, X., & Pasquarella, A. (2012). Learning about print : The development of orthographic processing and its relationship to word reading in first grade children in French immersion. *Reading and Writing*. <https://doi.org/10.1007/s11145-012-9407-2>
- Deacon, S. H., Benere, J., & Pasquarella, A. (2013). Reciprocal relationship : Children's morphological awareness and their reading accuracy across grades 2 to 3. *Developmental Psychology*, 49(6), 1113-1126. <https://doi.org/10.1037/a0029474>
- Deacon, S. H., & Kirby, J. R. (2004). Morphological awareness : Just "more phonological"? The roles of morphological and phonological awareness in reading development. *Applied Psycholinguistics*, 25(2), 223-238. <https://doi.org/10.1017/S0142716404001110>
- Deacon, S. H., Kirby, J. R., & Casselman-Bell, M. (2009). How Robust is the Contribution of Morphological Awareness to General Spelling Outcomes? *Reading Psychology*, 30(4), 301-318. <https://doi.org/10.1080/02702710802412057>
- Deacon, S. H., Pasquarella, A., Marinus, E., Tims, T., & Castles, A. (2019). Orthographic processing and children's word reading. *Applied Psycholinguistics*, 40(2), 509-534. <https://doi.org/10.1017/S0142716418000681>
- Deacon, S. H., Tong, X., & Francis, K. (2017). The relationship of morphological analysis and morphological decoding to reading comprehension : Morphological Analysis and Morphological Decoding. *Journal of Research in Reading*, 40(1), 1-16. <https://doi.org/10.1111/1467-9817.12056>
- de Jong, P. F., Bitter, D. J. L., van Setten, M., & Marinus, E. (2009). Does phonological recoding occur during silent reading, and is it necessary for orthographic learning? *Journal of Experimental Child Psychology*, 104(3), 267-282. <https://doi.org/10.1016/j.jecp.2009.06.002>
- de Jong, P. F., & Olson, R. K. (2004). Early predictors of letter knowledge. *Journal of Experimental Child Psychology*, 88(3), 254-273. <https://doi.org/10.1016/j.jecp.2004.03.007>
- de Jong, P. F., & Share, D. L. (2007). Orthographic learning during oral and silent reading. *Scientific Studies of Reading*, 11(1), 55-71. https://doi.org/10.1207/s1532799xssr1101_4
- DEPP. (2016). *Les performances en orthographe des élèves en fin d'école primaire (1987-2007-2015)* (Note d'information 28). Direction de l'évaluation, de la prospective et de la performance. <https://www.education.gouv.fr/les-performances-en-orthographe-des-eleves-en-fin-d-ecole-primaire-1987-2007-2015-1991>
- Desrochers, A., Manolitsis, G., Gaudreau, P., & Georgiou, G. (2018). Early contribution of morphological awareness to literacy skills across languages varying in orthographic consistency. *Reading and Writing*, 31(8), 1695-1719. <https://doi.org/10.1007/s11145-017-9772-y>
- Devonshire, V., & Fluck, M. (2010). Spelling development : Fine-tuning strategy-use and capitalising on the connections between words. *Learning and Instruction*, 20(5), 361-371. <https://doi.org/10.1016/j.learninstruc.2009.02.025>

- Di Folco, C., Guez, A., Peyre, H., & Ramus, F. (2022). Epidemiology of reading disability : A comparison of DSM-5 and ICD-11 criteria. *Scientific Studies of Reading, 26*(4), 337-355. <https://doi.org/10.1080/10888438.2021.1998067>
- Diamanti, V., Mouzaki, A., Ralli, A., Antoniou, F., Papaioannou, S., & Protopapas, A. (2017). Preschool Phonological and Morphological Awareness As Longitudinal Predictors of Early Reading and Spelling Development in Greek. *Frontiers in Psychology, 8*, 2039. <https://doi.org/10.3389/fpsyg.2017.02039>
- Diependaele, K., Sandra, D., & Grainger, J. (2005). Masked cross-modal morphological priming : Unravelling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes, 20*(1-2), 75-114. <https://doi.org/10.1080/01690960444000197>
- Diependaele, K., Sandra, D., & Grainger, J. (2009). Semantic transparency and masked morphological priming : The case of prefixed words. *Memory & Cognition, 37*(6), 895-908. <https://doi.org/10.3758/MC.37.6.895>
- Diependaele, K., Ziegler, J. C., & Grainger, J. (2010). Fast phonology and the Bimodal Interactive Activation Model. *European Journal of Cognitive Psychology, 22*(5), 764-778. <https://doi.org/10.1080/09541440902834782>
- Doignon-Camus, N., & Zagar, D. (2014). The syllabic bridge : The first step in learning spelling-to-sound correspondences*. *Journal of Child Language, 41*(5), 1147-1165. <https://doi.org/10.1017/S0305000913000305>
- Duke, N. K., & Pearson, P. D. (2009). Effective Practices for Developing Reading Comprehension. *Journal of Education, 189*(1-2), 107-122. <https://doi.org/10.1177/0022057409189001-208>
- Dunn, L. M., Dunn, L. M., & Thériault-Whalen, C. M. (1993). *Echelle de vocabulaire en images Peabody : EVIP*. PSYCAN.
- Egger, M., Smith, G. D., & Phillips, A. N. (1997). Meta-analysis : Principles and procedures. *BMJ (Clinical Research Ed.), 315*(7121), 1533-1537. <https://doi.org/10.1136/bmj.315.7121.1533>
- Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic Awareness Instruction Helps Children Learn to Read : Evidence From the National Reading Panel's Meta-Analysis. *Reading Research Quarterly, 36*(3), 250-287. <https://doi.org/10.1598/RRQ.36.3.2>
- Ehri, L. C., & Wilce, L. S. (1985). Movement into Reading : Is the First Stage of Printed Word Learning Visual or Phonetic? *Reading Research Quarterly, 20*(2), 163-179. <https://doi.org/10.2307/747753>
- Ehri, L. C., & Wilce, L. S. (1987). Does Learning to Spell Help Beginners Learn to Read Words? *Reading Research Quarterly, 22*(1), 47-65. JSTOR. <https://doi.org/10.2307/747720>
- Elbro, C., & Arnbak, E. (1996). The role of morpheme recognition and morphological awareness in dyslexia. *Annals of Dyslexia, 46*(1), 209-240. <https://doi.org/10.1007/BF02648177>
- Elliott, J. G. (2020). It's Time to Be Scientific About Dyslexia. *Reading Research Quarterly, 55*(S1). <https://doi.org/10.1002/rrq.333>
- Elliott, J., & Gibbs, S. (2008). Does Dyslexia Exist? *Journal of Philosophy of Education, 42*, 475-491. <https://doi.org/10.1111/j.1467-9752.2008.00653.x>
- Ellis, A. W. (1982). Spelling and Writing (and Reading and Speaking). In *Normality and Pathology in Cognitive Functions* (p. 113-146). Academic Press.

- Ellis, A. W. (1988). Normal writing processes and peripheral acquired dysgraphias. *Language and Cognitive Processes*, 3(2), 99-127. <https://doi.org/10.1080/01690968808402084>
- Engen, L., & Høien, T. (2002). Phonological skills and reading comprehension. *Reading and Writing*, 15(7), 613-631. <https://doi.org/10.1023/A:1020958105218>
- Erbeli, F., Peng, P., & Rice, M. (2021). No Evidence of Creative Benefit Accompanying Dyslexia : A Meta-Analysis. *Journal of Learning Disabilities*, 55, 242-253. <https://doi.org/10.1177/00222194211010350>
- Faber, G. (2005). Systematische Rechtschreibförderung mit algorithmischen und selbstinstruktiven Lernhilfen : Die Leistungen von konzeptgemäß trainierten und untrainierten Schülern im Vergleich. *Heilpädagogische Forschung*, 31, 129-144.
- Fejzo, A. (2016). The contribution of morphological awareness to the spelling of morphemes and morphologically complex words in French. *Reading and Writing*, 29(2), 207-228. <https://doi.org/10.1007/s11145-015-9586-8>
- Fejzo, A., Desrochers, A., & Deacon, H. (2018). The Acquisition of Derivational Morphology in Children. In *Morphological Processing and Literacy Development*. Routledge.
- Fejzo, A., Saidane, R., Whissell-Turner, K., & Chapleau, N. (2019, juillet 19). *Effects of morphological awareness training on the development of morphological strategy use in polymorphemic word spelling among 4th grade French-speaking students* [Poster]. 26th annual meeting of Society of Scientific Studies of Reading (SSSR), Toronto, Canada.
- Feldman, L. B., Kostić, A., Gvozdenović, V., O'Connor, P. A., & Moscoso del Prado Martín, F. (2012). Semantic similarity influences early morphological priming in Serbian : A challenge to form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review*, 19(4), 668-676. <https://doi.org/10.3758/s13423-012-0250-x>
- Filippini, A. L. (2007). *Effects of a vocabulary-added instructional intervention for at-risk English learners : Is efficient reading instruction more effective?* [Unpublished doctoral dissertation]. University of California, Santa Barbara.
- Firth, I. (1972). *Components of reading disability* [Unpublished doctoral dissertation]. University of New South Wales.
- Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2019). *Learning disabilities : From identification to intervention, 2nd ed* (p. xiv, 418). The Guilford Press.
- Folk, J. R., Rapp, B., & Goldrick, M. (2002). The interaction of lexical and sublexical information in spelling : What's the point? *Cognitive Neuropsychology*, 19(7), 653-671. <https://doi.org/10.1080/02643290244000184>
- Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, 12(6), 627-635. [https://doi.org/10.1016/S0022-5371\(73\)80042-8](https://doi.org/10.1016/S0022-5371(73)80042-8)
- Fradin, B. (2003). *Nouvelles approches en morphologie* (1. éd). Presses universitaires de France.
- Francis, H. (1982). *Learning to read : Literate behaviour and orthographic knowledge*. G. Allen & Unwin.
- Frank, A. R. (2008). *The effect of instruction in orthographic conventions and morphological features on the reading fluency and comprehension skills of high-school freshmen* [Doctoral dissertation in Education, University of San Francisco]. <https://repository.usfca.edu/diss/263>

- Frauenfelder, U. H., & Schreuder, R. (1992). Constraining psycholinguistic models of morphological processing and representation : The role of productivity. In G. Booij & J. van Marle (Éds.), *Yearbook of Morphology 1991* (p. 165-183). Springer Netherlands. https://doi.org/10.1007/978-94-011-2516-1_10
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In *Surface dyslexia : Neurological and cognitive studies of phonological reading* (K. Patterson, J. Marshall, M. Coltheart, p. 301-330). Lawrence Erlbaum.
- Frith, U. (1999). Paradoxes in the definition of dyslexia. *Frith, U. (1999) Paradoxes in the definition of dyslexia. Dyslexia, 5 (4). pp.192 - 214 . ISSN 10769242, 5.* [https://doi.org/10.1002/\(SICI\)1099-0909\(199912\)5:43.0.CO;2-N](https://doi.org/10.1002/(SICI)1099-0909(199912)5:43.0.CO;2-N)
- Gallet, C., Viriot-Goedel, C., & Leclercq, V. (2020). Effects of an early reading intervention based on grapho-syllabic decoding and fluency training in French elementary schools. *European Review of Applied Psychology, 70*(1), 100471. <https://doi.org/10.1016/j.erap.2019.100471>
- Galuschka, K., Goergen, R., Kalmar, J., Haberstroh, S., Schmalz, X., & Schulte-Koerne, G. (2020). Effectiveness of spelling interventions for learners with dyslexia : A meta-analysis and systematic review. *Educational Psychologist, 55*(1), 1-20. <https://doi.org/10.1080/00461520.2019.1659794>
- Galuschka, K., & Schulte-Körne, G. (2016). The Diagnosis and Treatment of Reading and/or Spelling Disorders in Children and Adolescents. *Deutsches Arzteblatt International, 113*(16), 279-286. <https://doi.org/10.3238/arztebl.2016.0279>
- Galuschka, K., Von Wirth, E., Krick, K., & Schulte-Körne, G. (2014). Effectiveness of Treatment Approaches for Children and Adolescents with Reading Disabilities : A Meta-Analysis of Randomized Controlled Trials. *PloS one, 9*, e89900. <https://doi.org/10.1371/journal.pone.0089900>
- Ganopole, S. J. (1987). The Development of Word Consciousness Prior to First Grade. *Journal of Reading Behavior, 19*(4), 415-436. <https://doi.org/10.1080/10862968709547614>
- Gebauer, D., Fink, A., Kargl, R., Reishofer, G., Koschutnig, K., Purgstaller, C., Fazekas, F., & Enzinger, C. (2012). Differences in brain function and changes with intervention in children with poor spelling and reading abilities. *PloS One, 7*(5), e38201. <https://doi.org/10.1371/journal.pone.0038201>
- Georgiou, G., Savage, R., Dunn, K., Bowers, P., & Parrila, R. (2021). Examining the Effects of Structured Word Inquiry on the Reading and Spelling Skills of Persistently Poor Grade 3 Readers. *Journal of Research in Reading, 44*(1). <https://doi.org/10.1111/1467-9817.12325>
- Given, B. K., Wasserman, J. D., Chari, S. A., Beattie, K., & Eden, G. F. (2008). A randomized, controlled study of computer-based intervention in middle school struggling readers. *Brain and Language, 106*(2), 83-97. <https://doi.org/10.1016/j.bandl.2007.12.001>
- Gleitman, L. R., & Rozin, P. (1977). The structure and acquisition of reading I: Relations between orthographies and the structure of language. In *Toward a psychology of reading* (p. 1-54).
- Goodwin, A. P. (2016). Effectiveness of word solving : Integrating morphological problem-solving within comprehension instruction for middle school students. *Reading and Writing, 29*(1), 91-116. <https://doi.org/10.1007/s11145-015-9581-0>
- Goodwin, A., Petscher, Y., & Tock, J. (2021). Multidimensional morphological assessment for middle school students. *Journal of Research in Reading, 44*(1), 70-89. <https://doi.org/10.1111/1467-9817.12335>

- Goodwin, & Ahn. (2010). A meta-analysis of morphological interventions : Effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia*, 60(2), 183-208. <https://doi.org/10.1007/s11881-010-0041-x>
- Goodwin, & Ahn. (2013). A Meta-Analysis of Morphological Interventions in English : Effects on Literacy Outcomes for School-Age Children. *Scientific Studies of Reading*, 17(4), 257-285. <https://doi.org/10.1080/10888438.2012.689791>
- Goswami, U. (1990). Phonological priming and orthographic analogies in reading. *Journal of Experimental Child Psychology*, 49, 323-340. [https://doi.org/10.1016/0022-0965\(90\)90062-D](https://doi.org/10.1016/0022-0965(90)90062-D)
- Goswami, U., & Bryant, P. (2016). *Phonological Skills and Learning to Read* (0 éd.). Routledge. <https://doi.org/10.4324/9781315695068>
- Goswami, U., Gombert, J. E., & Barrera, L. F. de. (1998). Children's orthographic representations and linguistic transparency : Nonsense word reading in English, French, and Spanish. *Applied Psycholinguistics*, 19(1), 19-52. <https://doi.org/10.1017/S0142716400010560>
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, Reading, and Reading Disability. *Remedial and Special Education*, 7(1), 6-10. <https://doi.org/10.1177/074193258600700104>
- Gough, P. B., & Walsh, M. A. (1991). Chinese, Phoenicians, and the orthographic cipher of English. In *Phonological processes in literacy : A tribute to Isabelle Y. Liberman* (p. 199-209). Lawrence Erlbaum Associates, Inc.
- Graham. (1999). Handwriting and Spelling Instruction for Students with Learning Disabilities : A Review. *Learning Disability Quarterly*, 22, 78-98. <https://doi.org/10.2307/1511268>
- Graham, N. L., Patterson, K., & Hodges, J. R. (1997). Progressive Dysgraphia : Co-occurrence of Central and Peripheral Impairments. *Cognitive Neuropsychology*, 14(7), 975-1005.
- Grigorakis, I., & Manolitsis, G. (2021). Longitudinal effects of different aspects of morphological awareness skills on early spelling development. *Reading and Writing*, 34(4), 945-979. <https://doi.org/10.1007/s11145-020-10098-2>
- Grund, M. (2003). *RST 4-7. Rechtschreibtest für 4-7. Klassen [RST 4-7. Spelling test for grades 4 to 7]*.
- Haase, A., & Steinbrink, C. (2022). Associations between morphological awareness and literacy skills in German primary school children : The roles of grade level, phonological processing and vocabulary. *Reading and Writing*. <https://doi.org/10.1007/s11145-021-10247-1>
- Hagiliassis, N., Pratt, C., & Johnston, M. (2006). Orthographic and Phonological Processes in Reading. *Reading and Writing*, 19(3), 235-263. <https://doi.org/10.1007/s11145-005-4123-9>
- Halls, C. (2021). *The struggles, successes and disappointments of carrying out Action Research : A reflective analysis of a 25 week syllable intervention across two schools focusing on improving reading and spelling in young children aged between five and nine years old*. [Thesis, University of Cambridge]. <https://doi.org/10.17863/CAM.72292>
- Han, B., Koh, P. W., Zhang, S., Joshi, R. M., & Li, H. (2022). The relative contributions of facets of morphological awareness to vocabulary development in Chinese : A longitudinal study in grades one to three. *Contemporary Educational Psychology*, 69, 102063. <https://doi.org/10.1016/j.cedpsych.2022.102063>
- Hanna, P. R., Hanna, J. S., Hodges, R. E., & Rudorf, E. H. (1966). *Phoneme-grapheme correspondences as cues to spelling improvement*. U. S. Government Printing Office.

- Harris, M. L., Schumaker, J. B., & Deshler, D. D. (2011). The Effects of Strategic Morphological Analysis Instruction on the Vocabulary Performance of Secondary Students with and without Disabilities. *Learning Disability Quarterly*, 34(1), 17-33. <https://doi.org/10.1177/073194871103400102>
- Hasenäcker, J., Beyersmann, E., & Schroeder, S. (2016). Masked Morphological Priming in German-Speaking Adults and Children : Evidence from Response Time Distributions. *Frontiers in Psychology*, 7, 929. <https://doi.org/10.3389/fpsyg.2016.00929>
- Hazard, M.-C., De Cara, B., Chanquoy, L., & Negro, I. (2020). Influence des caractéristiques de consistance orthographique et fréquence lexicale sur la nature des « fautes » d'orthographe en français : Profils développementaux du CE1 à la troisième. *Psychologie Française*, 65(3), 225-241. <https://doi.org/10.1016/j.psfr.2019.06.001>
- Heathcote, L., Nation, K., Castles, A., & Beyersmann, E. (2018). Do « blacheap » and « subcheap » both prime « cheap »? An investigation of morphemic status and position in early visual word processing. *Quarterly Journal of Experimental Psychology (2006)*, 71(8), 1645-1654. <https://doi.org/10.1080/17470218.2017.1362704>
- Heim, S., Tschierse, J., Amunts, K., Wilms, M., Vossel, S., Willmes, K., Grabowska, A., & Huber, W. (2008). Cognitive subtypes of dyslexia. *Acta Neurobiologiae Experimentalis*, 68(1), 73-82.
- Høien, T., Lundberg, I., Stanovich, K. E., & Bjaalid, I. K. (1995). Components of phonological awareness. *Reading and writing*, 7(2), 171-188.
- Houghton, G., & Zorzi, M. (1998). A model of the sound spelling mapping in English and its role in word and nonword spelling. In *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*. Lawrence Erlbaum Associates Inc.
- Houghton, G., & Zorzi, M. (2003). Normal and impaired spelling in a connectionist dual-route architecture. *Cognitive Neuropsychology*, 20(2), 115-162. <https://doi.org/10.1080/02643290242000871>
- Hughes, C. A., Morris, J. R., Therrien, W. J., & Benson, S. K. (2017). Explicit Instruction : Historical and Contemporary Contexts: LEARNING DISABILITIES RESEARCH. *Learning Disabilities Research & Practice*, 32(3), 140-148. <https://doi.org/10.1111/ldrp.12142>
- Huot, H. (2010). *La morphologie : Forme et sens des mots du français* (2. éd. rev. et actual). Colin.
- Ise, E., & Schulte-Körne, G. (2010). Spelling deficits in dyslexia : Evaluation of an orthographic spelling training. *Annals of Dyslexia*, 60(1), 18-39. JSTOR.
- Jacquier-Roux, M., Pouget, G., Valdois, S., & Zorman, M. (2010). *BALE: batterie analytique du langage écrit*. Groupe Cogni-Sciences, Laboratoire de Psychologie et NeuroCognition.
- Jacquier-Roux, M., Valdois, S., Zorman, M., Lequette, C., & Pouget, G. (2005). *ODÉDYS*. Laboratoire Cogni-Sciences.
- JASP Team. (2023). *JASP* (0.17.3) [Logiciel].
- Jiménez-Fernández, G., Vaquero, J. M. M., Jiménez, L., & Defior, S. (2011). Dyslexic children show deficits in implicit sequence learning, but not in explicit sequence learning or contextual cueing. *Annals of Dyslexia*, 61(1), 85-110. <https://doi.org/10.1007/s11881-010-0048-3>
- Jin, H. (2021). *Hello again, ANOVA : Rethinking ANOVA in the context of confirmatory data analysis*. <https://doi.org/10.31234/osf.io/yhmzg>

- Jöbstl, V., Kargl, R., Prattes, A. E., Beyersmann, E., & Landerl, K. (2021). Effects of a Morpheme-based Spelling Intervention Challenging Previous Results. *International Electronic Journal of Elementary Education*, 13(5), 651-671. <https://doi.org/10.26822/iejee.2021.219>
- Jonas, N. (2012). *Pour les générations les plus récentes, les difficultés des adultes diminuent à l'écrit, mais augmentent en calcul* (1426; Insee Première). INSEE. <https://www.insee.fr/fr/statistiques/1281410>
- Jorm, (1977). Effect of word imagery on reading performance as a function of reader ability. *Journal of Educational Psychology*, 69, 46-54. <https://doi.org/10.1037/0022-0663.69.1.46>
- Jorm, (1981). Children with reading and spelling retardation : Functioning of whole-word and correspondence-rule mechanisms. *Journal of Child Psychology and Psychiatry*, 22(2), 171-178. <https://doi.org/10.1111/j.1469-7610.1981.tb00540.x>
- Jorm, & Share, D. (1983). Phonological Recoding and Reading Acquisition. *Applied Psycholinguistics*, 4, 103-147. <https://doi.org/10.1017/S0142716400004380>
- Joseph, L. M., Konrad, M., Cates, G., Vajcner, T., Eveleigh, E., & Fishley, K. M. (2012). A meta-analytic review of the cover-copy-compare and variations of this self-management procedure. *Psychology in the Schools*, 49(2), 122-136. <https://doi.org/10.1002/pits.20622>
- Kazanina, N. (2011). Decomposition of prefixed words in Russian. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(6), 1371-1390. <https://doi.org/10.1037/a0024335>
- Kazanina, N., Dukova-Zheleva, G., Geber, D., Kharlamov, V., & Tonciulescu, K. (2008). Decomposition into multiple morphemes during lexical access : A masked priming study of Russian nouns. *Language and Cognitive Processes*, 23(6), 800-823. <https://doi.org/10.1080/01690960701799635>
- Kazdin, A. E. (2011). *Single-case research designs : Methods for clinical and applied settings* (2nd ed). Oxford University Press.
- Ke, S. (Echo), & Zhang, D. (2021). Morphological instruction and reading development in young L2 readers : A scoping review of causal relationships. *Studies in Second Language Learning and Teaching*, 11(3), 331-350. <https://doi.org/10.14746/ssllt.2021.11.3.2>
- Kearns, D. M., Steacy, L. M., Compton, D. L., Gilbert, J. K., Goodwin, A. P., Cho, E., Lindstrom, E. R., & Collins, A. A. (2016). Modeling Polymorphemic Word Recognition : Exploring Differences Among Children With Early-Emerging and Late-Emerging Word Reading Difficulty. *Journal of Learning Disabilities*, 49(4), 368-394. <https://doi.org/10.1177/0022219414554229>
- Kemp, N., & Bryant, P. (2003). Do Beez Buzz? Rule-Based and Frequency-Based Knowledge in Learning to Spell Plural -s. *Child Development*, 74(1), 63-74. <https://doi.org/10.1111/1467-8624.00521>
- Kemper, M. J., Verhoeven, L., & Bosman, A. M. T. (2012). Implicit and explicit instruction of spelling rules. *Learning and Individual Differences*, 22(6), 639-649. <https://doi.org/10.1016/j.lindif.2012.06.008>
- Kenner, B. B., Terry, N. P., Friehling, A. H., & Namy, L. L. (2017). Phonemic awareness development in 2.5- and 3.5-year-old children : An examination of emergent, receptive, knowledge and skills. *Reading and Writing*, 30(7), 1575-1594. <https://doi.org/10.1007/s11145-017-9738-0>
- Kieffer, M. J., & Lesaux, N. K. (2012). Effects of Academic Language Instruction on Relational and Syntactic Aspects of Morphological Awareness for Sixth Graders from Linguistically Diverse

- Backgrounds. *The Elementary School Journal*, 112(3), 519-545. JSTOR.
<https://doi.org/10.1086/663299>
- Kim, Y., Sidtis, D. V. L., & Sidtis, J. J. (2021). Emotional Nuance Enhances Verbatim Retention of Written Materials. *Frontiers in Psychology*, 12.
<https://www.frontiersin.org/articles/10.3389/fpsyg.2021.519729>
- Kimura, Y., & Bryant, P. (1983). Reading and writing in English and Japanese : A cross-cultural study of young children*. *British Journal of Developmental Psychology*, 1(2), 143-153.
<https://doi.org/10.1111/j.2044-835X.1983.tb00552.x>
- Kirby, Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing*, 25(2), 389-410.
<https://doi.org/10.1007/s11145-010-9276-5>
- Kirby, P. (2019). Gift from the gods? Dyslexia, popular culture and the ethics of representation. *Disability & Society*, 34(10), 1-22. <https://doi.org/10.1080/09687599.2019.1584091>
- Kirby, P. (2020). Dyslexia debated, then and now : A historical perspective on the dyslexia debate. *Oxford Review of Education*, 46(4), 472-486. <https://doi.org/10.1080/03054985.2020.1747418>
- Kirk, C., & Gillon, G. T. (2009). Integrated morphological awareness intervention as a tool for improving literacy. *Language, Speech, and Hearing Services in Schools*, 40(3), 341-351.
[https://doi.org/10.1044/0161-1461\(2008/08-0009\)](https://doi.org/10.1044/0161-1461(2008/08-0009))
- Kolne, K. (2019). *Disambiguating the benefits of morphology and vocabulary instruction for improving spelling among English-speaking children* [Doctoral dissertation in Communication Sciences and Disorders]. School of Communication Sciences and Disorders, McGill University.
- Krashen, S. (2004). *The Power of Reading : Insights from the Research* (2d edition). Heinemann.
- Krasny-Pacini, A., & Evans, J. (2018). Single-case experimental designs to assess intervention effectiveness in rehabilitation : A practical guide. *Annals of Physical and Rehabilitation Medicine*, 61(3), 164-179. <https://doi.org/10.1016/j.rehab.2017.12.002>
- Kratochwill, T., Horner, R., Levin, J., Machalicek, W., Ferron, J., & Johnson, A. (2021). Single-case design standards : An update and proposed upgrades. *Journal of School Psychology*, 89, 91-105.
<https://doi.org/10.1016/j.jsp.2021.10.006>
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). What Works Clearinghouse : Single-Case Design Technical Documentation. *Institute of Educational Sciences*. <https://ies.ed.gov/ncee/wwc/Document/229>
- Kruk, R. S., & Bergman, K. (2013). The reciprocal relations between morphological processes and reading. *Journal of Experimental Child Psychology*, 114(1), 10-34.
<https://doi.org/10.1016/j.jecp.2012.09.014>
- Kwong, T. E., & Varnhagen, C. K. (2005). Strategy Development and Learning to Spell New Words : Generalization of a Process. *Developmental Psychology*, 41(1), 148-159. <https://doi.org/10.1037/0012-1649.41.1.148>
- Kyte, C. S., & Johnson, C. J. (2006). The role of phonological recoding in orthographic learning. *Journal of Experimental Child Psychology*, 93, 166-185. <https://doi.org/10.1016/j.jecp.2005.09.003>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science : A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00863>

- Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P. H. T., Lohvansuu, K., O'Donovan, M., Williams, J., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., Tóth, D., Honbolygó, F., Csépe, V., Bogliotti, C., Iannuzzi, S., Chaix, Y., Démonet, J.-F., ... Schulte-Körne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity : Cross-linguistic predictors of dyslexia. *Journal of Child Psychology and Psychiatry*, *54*(6), 686-694. <https://doi.org/10.1111/jcpp.12029>
- Landerl, K., & Reitsma, P. (2005). Phonological and morphological consistency in the acquisition of vowel duration spelling in Dutch and German. *Journal of Experimental Child Psychology*, *92*(4), 322-344. <https://doi.org/10.1016/j.jecp.2005.04.005>
- Lane, J. D., & Gast, D. L. (2014). Visual analysis in single case experimental design studies : Brief review and guidelines. *Neuropsychological Rehabilitation*, *24*(3-4), 445-463. <https://doi.org/10.1080/09602011.2013.815636>
- Langsrud, Ø. (2003). ANOVA for unbalanced data : Use type II instead of type III sums of squares. *Statistics and Computing*, *13*, 163-167. <https://doi.org/10.1023/A:1023260610025>
- Lavine, L. O. (1977). Differentiation of letterlike forms in prereading children. *Developmental Psychology*, *13*(2), 89-94. <https://doi.org/10.1037/0012-1649.13.2.89>
- Law, J. M., Wouters, J., & Ghesquière, P. (2015). Morphological Awareness and Its Role in Compensation in Adults with Dyslexia. *Dyslexia (Chichester, England)*, *21*(3), 254-272. <https://doi.org/10.1002/dys.1495>
- Laxon, V., Masterson, J., Pool, M., & Keating, C. (1992). Nonword naming : Further exploration of the pseudohomophone effect in terms of orthographic neighborhood size, graphemic changes, spelling-sound consistency, and reader accuracy. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *18*(4), 730-748. <https://doi.org/10.1037//0278-7393.18.4.730>
- Lee, H.-W., Rayner, K., & Pollatsek, A. (2001). The Relative Contribution of Consonants and Vowels to Word Identification during Reading. *Journal of Memory and Language*, *44*(2), 189-205. <https://doi.org/10.1006/jmla.2000.2725>
- Lee, M., & Choi, S. (2020). Effects of Implicit and Explicit Instruction on Morphological Awareness and Knowledge. *Studies in English Education*, *25*(1), 1-20. <https://doi.org/10.22275/SEE.25.1.01>
- Lefavrais, P. (2005). *Alouette-R : Test d'analyse de la vitesse en lecture à partir d'un texte*. les Éditions du centre de psychologie appliquée.
- Lenhard, W., & Schneider, W. (2006). *ELFE 1-6 : Ein Leseverständnistest für Erst- bis Sechstklässler [A Reading Comprehension Test for Grade 1 to 6]*. Hogrefe.
- Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The Effectiveness and Ease of Implementation of an Academic Vocabulary Intervention for Linguistically Diverse Students in Urban Middle Schools. *Reading Research Quarterly*, *45*(2), 196-228. <https://doi.org/10.1598/RRQ.45.2.3>
- Leslie, L., & Shannon, A. J. (1981). Recognition of Orthographic Structure during Beginning Reading. *Journal of Reading Behavior*, *13*(4), 313-324. <https://doi.org/10.1080/10862968109547420>
- Lété, B. (2008). La consistance phonographique : Une mesure statistique de la complexité orthographique. Limoge : Lambert-Lucas. In *Nouvelles recherches en orthographe* (C. Brissaud, J.-P. Jaffré & J.-C. Pellat, p. 85-99). Lambert-Lucas Limoges.
- Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX : A grade-level lexical database from French elementary school readers. *Behavior Research Methods, Instruments, & Computers*, *36*(1), 156-166. <https://doi.org/10.3758/BF03195560>

- Levesque, K. C., Breadmore, H. L., & Deacon, S. H. (2021). How morphology impacts reading and spelling : Advancing the role of morphology in models of literacy development. *Journal of Research in Reading, 44*(1), 10-26. <https://doi.org/10.1111/1467-9817.12313>
- Levesque, K. C., & Deacon, S. H. (2022). Clarifying links to literacy : How does morphological awareness support children's word reading development? *Applied Psycholinguistics, 43*(4), 921-943. <https://doi.org/10.1017/S0142716422000194>
- Levesque, K. C., Kieffer, M. J., & Deacon, S. H. (2017). Morphological awareness and reading comprehension : Examining mediating factors. *Journal of Experimental Child Psychology, 160*, 1-20. <https://doi.org/10.1016/j.jecp.2017.02.015>
- Levesque, K. C., Kieffer, M. J., & Deacon, S. H. (2019). Inferring Meaning From Meaningful Parts : The Contributions of Morphological Skills to the Development of Children's Reading Comprehension. *Reading Research Quarterly, 54*(1), 63-80. <https://doi.org/10.1002/rrq.219>
- Levy, B. A., Gong, Z., Hessels, S., Evans, M. A., & Jared, D. (2006). Understanding print : Early reading development and the contributions of home literacy experiences. *Journal of Experimental Child Psychology, 93*(1), 63-93. <https://doi.org/10.1016/j.jecp.2005.07.003>
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*(6), 431-461. <https://doi.org/10.1037/h0020279>
- Lieberman, Shankweiler, D., Liberman, A. M., Fowler, C., & Fischer, F. W. (1977). Phonetic segmentation and recoding in the beginning reader. In *Towards a Psychology of Reading* (A. S. Reber & D. Scarborough). Lawrence Erlbaum Associates Inc.
- Lobo, M. A., Moeyaert, M., Cunha, A. B., & Babik, I. (2017). Single-Case Design, Analysis, and Quality Assessment for Intervention Research. *Journal of neurologic physical therapy : JNPT, 41*(3), 187-197. <https://doi.org/10.1097/NPT.0000000000000187>
- Longtin, C.-M., Segui, J., & Halle, P. A. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes, 18*(3), 313-334. <https://doi.org/10.1080/01690960244000036>
- Lonigan, C. J., Burgess, S. R., Anthony, J. L., & Barker, T. A. (1998). Development of Phonological Sensitivity in 2- to 5-Year-Old Children. *Journal of Educational Psychology, 90*(2), 294-311.
- Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia, 53*(1), 1-14. <https://doi.org/10.1007/s11881-003-0001-9>
- Lyster, S.-A. H. (2002). The effects of morphological versus phonological awareness training in kindergarten on reading development. *Reading and Writing, 15*(3), 261-294. <https://doi.org/10.1023/A:1015272516220>
- Lyster, S.-A. H., Lervåg, A. O., & Hulme, C. (2016). Preschool morphological training produces long-term improvements in reading comprehension. *Reading and Writing, 29*, 1269-1288. <https://doi.org/10.1007/s11145-016-9636-x>
- Maclean, M., Bryant, P., & Bradley, L. (1987). Rhymes, nursery rhymes, and reading in early childhood. *Merrill-Palmer Quarterly, 33*(3), 255-281.
- Maggin, D. M., Cook, B. G., & Cook, L. (2018). Using Single-Case Research Designs to Examine the Effects of Interventions in Special Education. *Learning Disabilities Research & Practice, 33*(4), 182-191. <https://doi.org/10.1111/ldrp.12184>

- Mailhot, H., Wilson, M. A., Macoir, J., Deacon, S. H., & Sánchez-Gutiérrez, C. (2020). Morpholex-FR : A derivational morphological database for 38,840 French words. *Behavior Research Methods*, 52(3), 1008-1025. <https://doi.org/10.3758/s13428-019-01297-z>
- Maïonchi-Pino, N., De Cara, B., Ecalle, J., & Magnan, A. (2012). Do Consonant Sonority and Status Influence Syllable-Based Segmentation Strategies in a Visual Letter Detection Task? Developmental Evidence in French Children. *Scientific Studies of Reading - SCI STUD READ*, 16, 1-13. <https://doi.org/10.1080/10888438.2011.620672>
- Maïonchi-Pino, N., Magnan, A., & Écalle, J. (2010). Syllable frequency effects in visual word recognition : Developmental approach in French children. *Journal of Applied Developmental Psychology*, 31, 70-82. <https://doi.org/10.1016/j.appdev.2009.08.003>
- Manesse, D., Cogis, D., & Chervel, A. (2007). *Orthographe : À qui la faute ?* ESF Editeur.
- Mann, V. A., & Foy, J. G. (2003). Phonological awareness, speech development, and letter knowledge in preschool children. *Annals of Dyslexia*, 53(1), 149-173. <https://doi.org/10.1007/s11881-003-0008-2>
- Manolitsis, G., Grigorakis, I., & Georgiou, G. K. (2017). The longitudinal contribution of early morphological awareness skills to reading fluency and comprehension in Greek. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01793>
- Manza, L., & Reber, A. S. (1997). Representing artificial grammars : Transfer across stimulus forms and modalities. In *How implicit is implicit learning?* (p. 73-106). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198523512.003.0004>
- Marelli, M., Traficante, D., & Burani, C. (2020). Reading morphologically complex words : Experimental evidence and learning models. In V. Pirrelli, I. Plag, & W. U. Dressler (Éds.), *Word Knowledge and Word Usage* (p. 553-592). De Gruyter. <https://doi.org/10.1515/9783110440577-014>
- Marinelli, C. V., Romani, C., Burani, C., & Zoccolotti, P. (2015). Spelling Acquisition in English and Italian : A Cross-Linguistic Study. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01843>
- Marsh, G., Desberg, P., & Cooper, J. (1977). Developmental Changes in Reading Strategies. *Journal of Reading Behavior*, 9(4), 391-394. <https://doi.org/10.1080/10862967709547244>
- Marslen-Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, 101(1), 3-33. <https://doi.org/10.1037/0033-295X.101.1.3>
- Marulis, L. M., & Neuman, S. B. (2013). How Vocabulary Interventions Affect Young Children at Risk : A Meta-Analytic Review. *Journal of Research on Educational Effectiveness*, 6(3), 223-262. <https://doi.org/10.1080/19345747.2012.755591>
- Masonheimer, P. E., Drum, P. A., & Ehri, L. C. (1984). Does environmental print identification lead children into word reading? *Journal of Reading Behavior*, 16, 257-271.
- McArthur, G., Sheehan, Y., Badcock, N. A., Francis, D. A., Wang, H.-C., Kohonen, S., Banales, E., Anandakumar, T., Marinus, E., & Castles, A. (2018). Phonics training for English-speaking poor readers. *Cochrane Database of Systematic Reviews*, 11. <https://doi.org/10.1002/14651858.CD009115.pub3>
- McCann, R., & Besner, D. (1987). Reading pseudohomophones : Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance*, 13(1), 14-24. <https://doi.org/10.1037/0096-1523.13.1.14>

- McCaughey, M. W., Juola, J. F., Schadler, M., & Ward, N. J. (1980). Whole-word units are used before orthographic knowledge in perceptual development. *Journal of Experimental Child Psychology*, 30(3), 411-421. [https://doi.org/10.1016/0022-0965\(80\)90047-8](https://doi.org/10.1016/0022-0965(80)90047-8)
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception : I. An account of basic findings. *Psychological Review*, 88, 375-407. <https://doi.org/10.1037/0033-295X.88.5.375>
- McCloskey, M., Macaruso, P., & Rapp, B. (2006). Grapheme-to-lexeme feedback in the spelling system : Evidence from a dysgraphic patient. *Cognitive neuropsychology*, 23, 278-307. <https://doi.org/10.1080/02643290442000518>
- McQuillan, J. (2020). *The Effects of Morphological Training on Vocabulary Knowledge : A Reanalysis of Goodwin & Ahn*. 19-38.
- Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read : A meta-analytic review. *Psychological Bulletin*, 138(2), 322-352. <https://doi.org/10.1037/a0026744>
- Miciak, J., & Fletcher, J. M. (2020). The Critical Role of Instructional Response for Identifying Dyslexia and Other Learning Disabilities. *Journal of Learning Disabilities*, 53(5), 343-353. <https://doi.org/10.1177/0022219420906801>
- Bulletin officiel spécial n°11 du 26 novembre 2015 : Programmes d'enseignement de l'école élémentaire et du collège, Pub. L. No. 11 (2015). <https://www.education.gouv.fr/au-bo-special-du-26-novembre-2015-programmes-d-enseignement-de-l-ecole-elementaire-et-du-college-3737>
- Morris, S. B. (2008). Estimating Effect Sizes From Pretest-Posttest-Control Group Designs. *Organizational Research Methods*, 11(2), 364-386. <https://doi.org/10.1177/1094428106291059>
- Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76(2), 165-178. <https://doi.org/10.1037/h0027366>
- Mukaka, M. M. (2012). A guide to appropriate use of Correlation coefficient in medical research. *Malawi Medical Journal : The Journal of Medical Association of Malawi*, 24(3), 69.
- Nagy, & Anderson, R. C. (1984). How Many Words Are There in Printed School English? *Reading Research Quarterly*, 19(3), 304. <https://doi.org/10.2307/747823>
- Nagy, Berninger, V. W., & Abbott, R. D. (2006). Contributions of Morphology Beyond Phonology to Literacy Outcomes of Upper Elementary and Middle-School Students. *Journal of Educational Psychology*, 98(1), 134-147. <https://doi.org/10.1037/0022-0663.98.1.134>
- Nation, K., Angell, P., & Castles, A. (2007). Orthographic learning via self-teaching in children learning to read English : Effects of exposure, durability, and context. *Journal of Experimental Child Psychology*, 96(1), 71-84. <https://doi.org/10.1016/j.jecp.2006.06.004>
- Neumann, M. M., Hood, M., Ford, R. M., & Neumann, D. L. (2012). The role of environmental print in emergent literacy. *Journal of Early Childhood Literacy*, 12(3), 231-258. <https://doi.org/10.1177/1468798411417080>
- Nicholson, T., & Hill, D. (1985). Good readers don't guess : Taking another look at the issue of whether children read words better in context or in isolation. *Reading Psychology*, 6, 181-198.
- Noel Foulin, J. (2005). Why is letter-name knowledge such a good predictor of learning to read? *Reading and Writing*, 18(2), 129-155. <https://doi.org/10.1007/s11145-004-5892-2>

- Nokes-Malach, T. J., Richey, J. E., & Gadgil, S. (2015). When Is It Better to Learn Together? Insights from Research on Collaborative Learning. *Educational Psychology Review*, 27(4), 645-656. <https://doi.org/10.1007/s10648-015-9312-8>
- Nunes, T., Bryant, P., & Olsson, J. (2003). Learning Morphological and Phonological Spelling Rules : An Intervention Study. *Scientific Studies of Reading - SCI STUD READ*, 7, 289-307. https://doi.org/10.1207/S1532799XSSR0703_6
- Oliveira, M., Levesque, K. C., Deacon, S. H., & Peruzzi Elia da Mota, M. M. (2020). Evaluating models of how morphological awareness connects to reading comprehension : A study in Portuguese. *Journal of Research in Reading*, 43(2), 161-179. <https://doi.org/10.1111/1467-9817.12296>
- Olson, R., Forsberg, H., Wise, B., & Rack, J. (1994). Measurement of word recognition, orthographic, and phonological skills. In *Frames of reference for the assessment of learning disabilities : New views on measurement issues* (p. 243-277). Paul H Brookes Publishing Co.
- Olson, R., Kliegl, R., Davidson, B., & Foltz, G. (1985). Individual and developmental differences in reading disability. *Reading research : advances in theory and practice, Vol. 4. - San Diego : Academic, 1985, S. 1-64.*
- Orton, S. T. (1937). *Reading, writing and speech problems in children* (p. 215). W W Norton & Co.
- Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound : Still a good horse race. *Psychological Research*, 53(1), 13-24. <https://doi.org/10.1007/BF00867328>
- Packard, J. L., Chen, X., Li, W., Wu, X., Gaffney, J. S., Li, H., & Anderson, R. C. (2006). Explicit instruction in orthographic structure and word morphology helps Chinese children learn to write characters. *Reading and Writing*, 19(5), 457-487. <https://doi.org/10.1007/s11145-006-9003-4>
- Pacton, S., & Afonso Jaco, A. (2015). How do children learn spelling words? *Revue Française de Linguistique Appliquée*, 20, 51-61.
- Pacton, S., Afonso Jaco, A., Nys, M., Foulon, J. N., Treiman, R., & Peereman, R. (2018). Children benefit from morphological relatedness independently of orthographic relatedness when they learn to spell new words. *Journal of Experimental Child Psychology*, 171, 71-83. <https://doi.org/10.1016/j.jecp.2018.02.003>
- Pacton, S., Borchardt, G., Treiman, R., Lété, B., & Fayol, M. (2014). Learning to spell from reading : General knowledge about spelling patterns influences memory for specific words. *Quarterly Journal of Experimental Psychology (2006)*, 67(5), 1019-1036. <https://doi.org/10.1080/17470218.2013.846392>
- Pacton, S., & Deacon, S. H. (2008). The timing and mechanisms of children's use of morphological information in spelling : A review of evidence from English and French. *Cognitive Development*, 23(3), 339-359. <https://doi.org/10.1016/j.cogdev.2007.09.004>
- Pacton, S., Fayol, M., & Perruchet, P. (2005). Children's Implicit Learning of Graphotactic and Morphological Regularities. *Child Development*, 76(2), 324-339. JSTOR.
- Pacton, S., Foulon, J., Casalis, S., & Treiman, R. (2013). Children benefit from morphological relatedness when they learn to spell new words. *Frontiers in psychology*, 4, 696. <https://doi.org/10.3389/fpsyg.2013.00696>
- Pacton, S., Perruchet, P., Fayol, M., & 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>

- Pagliuca, G., Arduino, L. S., Barca, L., & Burani, C. (2008). Fully transparent orthography, yet lexical reading aloud : The lexicality effect in Italian. *Language and Cognitive Processes*, 23(3), 422-433. <https://doi.org/10.1080/01690960701626036>
- Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research : Nonoverlap of all pairs. *Behavior Therapy*, 40(4), 357-367. <https://doi.org/10.1016/j.beth.2008.10.006>
- Passaretti, B., Turkstra, L., Gallagher, T., Jiang, A., Cahill, P., & Campbell, W. (2023). Reporting of Classroom-Based Morphological Awareness Instruction and Intervention for Kindergarten to Grade 3 Students in the Literature : A Scoping Review. *Language, Speech, and Hearing Services in Schools*, 1-22. https://doi.org/10.1044/2022_LSHSS-22-00090
- Peereman, R., L  t  , B., & Sprenger-Charolles, L. (2007). Manulex-infra : Distributional characteristics of grapheme-phoneme mappings, and infralexical and lexical units in child-directed written material. *Behavior Research Methods*, 39(3), 593-603. <https://doi.org/10.3758/BF03193029>
- Peereman, R., Sprenger-Charolles, L., & Messaoud-Galusi, S. (2013). The contribution of morphology to the consistency of spelling-to-sound relations : A quantitative analysis based on French elementary school readers. *L'Ann  e Psychologique*, Vol. 113(1), 3-33.
- Pennington, B. F., McGrath, L. M., & Peterson, R. L. (2019). *Diagnosing learning disorders : From science to practice*, 3rd ed (p. xvi, 399). The Guilford Press.
- Perfetti, C. A. (2007). Reading Ability : Lexical Quality to Comprehension. *Scientific Studies of Reading*, 11(4), 357-383.
- Perry, C., & Ziegler, J. C. (2004). Beyond the two-strategy model of skilled spelling : Effects of consistency, grain size, and orthographic redundancy. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 57A(2), 325-356. <https://doi.org/10.1080/02724980343000323>
- Perry, C., Ziegler, J. C., & Coltheart, M. (2002). How predictable is spelling? Developing and testing metrics of phoneme-grapheme contingency. *The Quarterly Journal of Experimental Psychology Section A*, 55(3), 897-915. <https://doi.org/10.1080/02724980143000640>
- Perry, C., Ziegler, J., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories : The CDP+ model of reading aloud. *Psychological review*, 114, 273-315. <https://doi.org/10.1037/0033-295X.114.2.273>
- Petersen-Brown, S., Karich, A. C., & Symons, F. J. (2012). Examining Estimates of Effect Using Non-Overlap of All Pairs in Multiple Baseline Studies of Academic Intervention. *Journal of Behavioral Education*, 21(3), 203-216. <https://doi.org/10.1007/s10864-012-9154-0>
- Peterson, R. L., & Pennington, B. F. (2015). Developmental dyslexia. *Annual Review of Clinical Psychology*, 11, 283-307. <https://doi.org/10.1146/annurev-clinpsy-032814-112842>
- Pfost, M., Hattie, J., D  rfler, T., & Artelt, C. (2014). Individual Differences in Reading Development : A Review of 25 Years of Empirical Research on Matthew Effects in Reading. *Review of Educational Research*, 84(2), 203-244. <https://doi.org/10.3102/0034654313509492>
- Pick, A. D., Unze, M. G., Brownell, C. A., Drozdal, J. G., & Hopmann, M. R. (1978). Young Children's Knowledge of Word Structure. *Child Development*, 49(3), 669-680. <https://doi.org/10.2307/1128234>
- Pring, L., & Snowling, M. (1986). Developmental changes in word recognition : An information-processing account. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 38, 395-418. <https://doi.org/10.1080/14640748608401605>

- Pritchard. (2013). *Incorporating Learning Mechanisms into the Dual-Route Cascaded (DRC) Model of Reading Aloud and Word Recognition*.
- Pritchard, Coltheart, M., Marinus, E., & Castles, A. (2018). A Computational Model of the Self-Teaching Hypothesis Based on the Dual-Route Cascaded Model of Reading. *Cognitive Science*, 42(3), 722-770. <https://doi.org/10.1111/cogs.12571>
- Pritchard, S. C., Coltheart, M., Marinus, E., & Castles, A. (2016). Modelling the implicit learning of phonological decoding from training on whole-word spellings and pronunciations. *Scientific Studies of Reading*, 20(1), 49-63. <https://doi.org/10.1080/10888438.2015.1085384>
- Pritchard, S. C., Coltheart, M., Palethorpe, S., & Castles, A. (2012). Nonword reading : Comparing dual-route cascaded and connectionist dual-process models with human data. *Journal of Experimental Psychology: Human Perception and Performance*, 38(5), 1268-1288. <https://doi.org/10.1037/a0026703>
- Pyykkönen, P., & Järvikivi, J. (2012). Children and situation models of multiple events. *Developmental Psychology*, 48(2), 521-529. <https://doi.org/10.1037/a0025526>
- Quémart, P., Casalis, S., & Duncan, L. G. (2012). Exploring the role of bases and suffixes when reading familiar and unfamiliar words : Evidence from French young readers. *Scientific Studies of Reading*, 16(5), 424-442. <https://doi.org/10.1080/10888438.2011.584333>
- R Core Team. (2022). *R: A language and environment for statistical computing* [Logiciel]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ramus, F., Altarelli, I., Jednoróg, K., Zhao, J., & Scotto di Covella, L. (2018). Neuroanatomy of developmental dyslexia : Pitfalls and promise. *Neuroscience and Biobehavioral Reviews*, 84, 434-452. <https://doi.org/10.1016/j.neubiorev.2017.08.001>
- Rapp, B., Epstein, C., & Tainturier, M.-J. (2002). The integration of information across lexical and sublexical processes in spelling. *Cognitive neuropsychology*, 19, 1-29. <https://doi.org/10.1080/0264329014300060>
- Rassel, A. (2020). *Contribution de la conscience morphologique à l'apprentissage de la lecture : Études en milieu défavorisé* [Doctoral dissertation in Psychology, University of Lille 3]. <http://www.theses.fr/2020LIL3H023>
- Rastle, K. (2019). The place of morphology in learning to read in English. *Cortex*, 116, 45-54. <https://doi.org/10.1016/j.cortex.2018.02.008>
- Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition : A time-course study. *Language and Cognitive Processes*, 15(4-5), 507-537. <https://doi.org/10.1080/01690960050119689>
- Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel : Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, 11(6), 1090-1098. <https://doi.org/10.3758/BF03196742>
- Raven, J. C., Court, J. H., & Raven, J. E. (1990). Section 2 : Coloured Progressive Matrices. In *Manual for the Raven's Progressive Matrices and Vocabulary Scales* (1990 Edition, with US Norms). Oxford Psychologist Press.
- Ravid, D. (1996). Accessing the mental lexicon : Evidence from incompatibility between representation of spoken and written morphology. *Linguistics*, 34(6). <https://doi.org/10.1515/ling.1996.34.6.1219>

- Ravid, D., & Geiger, V. (2009). Promoting morphological awareness in Hebrew-speaking grade-schoolers : An intervention study using linguistic humor. *First Language*, 29(1), 81-112. <https://doi.org/10.1177/0142723708097483>
- Read, C. (1971). Pre-School Children's Knowledge of English Phonology. *Harvard Educational Review*, 41(1), 1-34. <https://doi.org/10.17763/haer.41.1.91367v0h80051573>
- Reed, D. K. (2008). A Synthesis of Morphology Interventions and Effects on Reading Outcomes for Students in Grades K–12. *Learning Disabilities Research & Practice*, 23(1), 36-49. <https://doi.org/10.1111/j.1540-5826.2007.00261.x>
- Reis, A., Araújo, S., Morais, I. S., & Faísca, L. (2020). Reading and reading-related skills in adults with dyslexia from different orthographic systems : A review and meta-analysis. *Annals of Dyslexia*, 70(3), 339-368. <https://doi.org/10.1007/s11881-020-00205-x>
- Reitsma, P. (1983). Printed word learning in beginning readers. *Journal of Experimental Child Psychology*, 36(2), 321-339. [https://doi.org/10.1016/0022-0965\(83\)90036-X](https://doi.org/10.1016/0022-0965(83)90036-X)
- Rey-Debove, J. (1984). Le domaine de la morphologie lexicale. In *Cahiers de lexicologie* (p. 3-19).
- Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, 6(2), 135-147. <https://doi.org/10.1016/j.edurev.2010.12.001>
- Robertson, E. K., & Deacon, S. H. (2019). Morphological awareness and word-level reading in early and middle elementary school years. *Applied Psycholinguistics*, 40(4), 1051-1071. <https://doi.org/10.1017/S0142716419000134>
- Rollenhagen, J. E., & Olson, C. R. (2000). Mirror-image confusion in single neurons of the macaque inferotemporal cortex. *Science (New York, N.Y.)*, 287(5457), 1506-1508. <https://doi.org/10.1126/science.287.5457.1506>
- Rowland, C. A. (2014). The effect of testing versus restudy on retention : A meta-analytic review of the testing effect. *Psychological Bulletin*, 140, 1432-1463. <https://doi.org/10.1037/a0037559>
- Rozin, P., Poritsky, S., & Sotsky, R. (1971). American children with reading problems can easily learn to read English represented by Chinese characters. *Science (New York, N.Y.)*, 171(3977), 1264-1267. <https://doi.org/10.1126/science.171.3977.1264>
- Ruan, Y., Georgiou, G. K., Song, S., Li, Y., & Shu, H. (2018). Does writing system influence the associations between phonological awareness, morphological awareness, and reading? A meta-analysis. *Journal of Educational Psychology*, 110(2), 180-202. <https://doi.org/10.1037/edu0000216>
- Rueda-Sanchez, M. I., & Lopez-Bastida, P. (2016). Effects of morphological awareness training on reading, writing and comprehension : Meta-analysis. *Anales De Psicología*, 32(1), 60-71. <https://doi.org/10.6018/analesps.32.1.196261>
- Rumelhart, D. E., & Siple, P. (1974). Process of recognizing tachistoscopically presented words. *Psychological Review*, 81(2), 99-118. <https://doi.org/10.1037/h0036117>
- Saksida, A., Iannuzzi, S., Bogliotti, C., Chaix, Y., Démonet, J.-F., Bricout, L., Billard, C., Nguyen-Morel, M.-A., Le Heuzey, M.-F., Soares-Boucaud, I., George, F., Ziegler, J. C., & Ramus, F. (2016). Phonological skills, visual attention span, and visual stress in developmental dyslexia. *Developmental Psychology*, 52(10), 1503-1516. <https://doi.org/10.1037/dev0000184>
- Schiff, R. (2003). The effects of morphology and word length on the reading of Hebrew nominals. *Reading and Writing*, 16(4), 263-287. <https://doi.org/10.1023/A:1023666419302>

- Schiff, R., & Saiegh-Haddad, E. (2018). Development and Relationships Between Phonological Awareness, Morphological Awareness and Word Reading in Spoken and Standard Arabic. *Frontiers in Psychology*, 9. <https://www.frontiersin.org/article/10.3389/fpsyg.2018.00356>
- Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In *Morphological aspects of language processing* (p. 131-154). Lawrence Erlbaum Associates, Inc.
- Schulte-Körne, G., Deimel, W., Hülsmann, J., Seidler, T., & Remschmidt, H. (2001). Das Marburger Rechtschreib-Training—Ergebnisse einer Kurzzeit-Intervention. *Zeitschrift Für Kinder- und Jugendpsychiatrie Und Psychotherapie - Z KINDER JUGENDPSYCHIAT PSYCH*, 29, 7-15. <https://doi.org/10.1024//1422-4917.29.1.7>
- Schulte-Körne, G., Deimel, W., & Remschmidt, H. (2003). Rechtschreibtraining in schulischen Fördergruppen—Ergebnisse einer Evaluationsstudie in der Primarstufe. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 31(2), 85-98. <https://doi.org/10.1024/1422-4917.31.2.85>
- Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). *Meta-analysis with R*. Springer.
- Scott, J. A., & Ehri, L. C. (1990). Sight Word Reading in Prereaders : Use of Logographic vs. Alphabetic Access Routes. *Journal of Reading Behavior*, 22(2), 149-166. <https://doi.org/10.1080/10862969009547701>
- Seidenberg, M. S., & Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences*, 4(9), 353-361. [https://doi.org/10.1016/s1364-6613\(00\)01515-1](https://doi.org/10.1016/s1364-6613(00)01515-1)
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4), 523-568. <https://doi.org/10.1037/0033-295X.96.4.523>
- Selby, S. (1972). The development of morphological rules in children. *British Journal of Educational Psychology*, 42(3), 293-299. <https://doi.org/10.1111/j.2044-8279.1972.tb00722.x>
- Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology (London, England: 1953)*, 94(Pt 2), 143-174. <https://doi.org/10.1348/000712603321661859>
- Shahar-Yames, D., & Share, D. L. (2008). Spelling as a self-teaching mechanism in orthographic learning. *Journal of Research in Reading*, 31(1), 22-39. <https://doi.org/10.1111/j.1467-9817.2007.00359.x>
- Share. (1999). Phonological recoding and orthographic learning : A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, 72(2), 95-129. <https://doi.org/10.1006/jecp.1998.2481>
- Share, D. L. (1995). Phonological recoding and self-teaching : Sine qua non of reading acquisition. *Cognition*, 55(2), 151-218. [https://doi.org/10.1016/0010-0277\(94\)00645-2](https://doi.org/10.1016/0010-0277(94)00645-2)
- Share, D. L. (2008). On the Anglocentricities of current reading research and practice : The perils of overreliance on an « outlier » orthography. *Psychological Bulletin*, 134(4), 584-615. <https://doi.org/10.1037/0033-2909.134.4.584>
- Share, D. L. (2021a). Common Misconceptions about the Phonological Deficit Theory of Dyslexia. *Brain Sciences*, 11(11), 1510. <https://doi.org/10.3390/brainsci11111510>
- Share, D. L. (2021b). Is the Science of Reading Just the Science of Reading English? *Reading Research Quarterly*, 56(S1), S391-S402. <https://doi.org/10.1002/rrq.401>

- Singh, D., Wonnacott, E., & Samara, A. (2021). Statistical and explicit learning of graphotactic patterns with no phonological counterpart : Evidence from an artificial lexicon study with 6–7-year-olds and adults. *Journal of Memory and Language*, *121*, 104265. <https://doi.org/10.1016/j.jml.2021.104265>
- Snowling, M., & Frith, U. (1981). The role of sound, shape and orthographic cues in early reading. *British Journal of Psychology*, *72*, 83-87. <https://doi.org/10.1111/j.2044-8295.1981.tb02164.x>
- Sparks, E., & Deacon, H. (2015). Morphological awareness and vocabulary acquisition : A longitudinal examination of their relationship in English-speaking children. *Applied Psycholinguistics*, *36*, 299-321. <https://doi.org/10.1017/S0142716413000246>
- Sprengr-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. *Approche neuropsychologique des apprentissages chez l'enfant*, *29*(3), 247-256.
- Sprengr-Charolles, L., Colé, P., Béchenec, D., & Kipffer-Piquard, A. (2005). French normative data on reading and related skills from EVALEC, a new computerized battery of tests (end Grade 1, Grade 2, Grade 3, and Grade 4). *European Review of Applied Psychology*, *55*(3), 157-186. <https://doi.org/10.1016/j.erap.2004.11.002>
- Stanké, B., Moreau, A., Royle, P., Rezzonico, S., & Dumais, C. (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 3 année du primaire* (Programme Actions Concertées 2015-LC-187661). Université de Montréal. <https://policycommons.net/artifacts/2244020/nouvelle-approche-basee-sur-un-enseignement-orthographique-favorisant-lapprentissage-de-la-production-ecrite-et-de-lorthographe-lexicale-des-eleves-faibles-orthographes-de-3-annee-du-primaire/3002163/>
- Stanovich, K. E. (1986). Matthew Effects in Reading : Some Consequences of Individual Differences in the Acquisition of Literacy. *Reading Research Quarterly*, *21*(4), 360-407. <https://doi.org/10.1598/RRQ.21.4.1>
- Stanovich, K. E., & West, R. F. (1989). Exposure to Print and Orthographic Processing. *Reading Research Quarterly*, *24*(4), 402-433. JSTOR. <https://doi.org/10.2307/747605>
- Stuart, M. (1990). Factors influencing word recognition in pre-reading children. *British Journal of Psychology*, *81*, 135-146. <https://doi.org/10.1111/j.2044-8295.1990.tb02351.x>
- Stuart, M., & Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition*, *30*, 139-181. [https://doi.org/10.1016/0010-0277\(88\)90038-8](https://doi.org/10.1016/0010-0277(88)90038-8)
- Suggate, S. P. (2010). Why what we teach depends on when : Grade and reading intervention modality moderate effect size. *Developmental Psychology*, *46*(6), 1556-1579. <https://doi.org/10.1037/a0020612>
- Suggate, S. P. (2016). A Meta-Analysis of the Long-Term Effects of Phonemic Awareness, Phonics, Fluency, and Reading Comprehension Interventions. *Journal of Learning Disabilities*, *49*(1), 77-96. <https://doi.org/10.1177/0022219414528540>
- Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning & Verbal Behavior*, *14*(6), 638-647. [https://doi.org/10.1016/S0022-5371\(75\)80051-X](https://doi.org/10.1016/S0022-5371(75)80051-X)
- Taha, H. (2016). Deep and shallow in Arabic orthography : New evidence from reading performance of elementary school native Arab readers. *Writing Systems Research*, *8*. <https://doi.org/10.1080/17586801.2015.1114910>

- Taha, H., & Saiegh-Haddad, E. (2016). The Role of Phonological versus Morphological Skills in the Development of Arabic Spelling : An Intervention Study. *Journal of Psycholinguistic Research*, 45(3), 507-535. <https://doi.org/10.1007/s10936-015-9362-6>
- Tainturier, M.-J., & Rapp, B. (2001). *The spelling process* (p. 263-289).
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Taylor, H., & Vestergaard, M. D. (2022). Developmental Dyslexia : Disorder or Specialization in Exploration? *Frontiers in Psychology*, 13. <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.889245>
- Temple, C. M., & Marshall, J. C. (1983). A case study of developmental phonological dyslexia. *British Journal of Psychology*, 74(4), 517-533. <https://doi.org/10.1111/j.2044-8295.1983.tb01883.x>
- Templeton, S. (2020). Stages, Phases, Repertoires, and Waves : Learning to Spell and Read Words. *The Reading Teacher*, 74(3), 315-323. <https://doi.org/10.1002/trtr.1951>
- The jamovi project. (2021). *Jamovi* (1.6) [Logiciel]. <https://www.jamovi.org>
- Torgerson, C., Brooks, G., Gascoine, L., & Higgins, S. (2019). Phonics : Reading policy and the evidence of effectiveness from a systematic 'tertiary' review. *Research Papers in Education*, 34(2), 208-238. <https://doi.org/10.1080/02671522.2017.1420816>
- Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of Word Reading Efficiency (TOWRE)*. Pro-Ed.
- Torkildsen, J. von K., Bratlie, S. S., Kristensen, J. K., Gustafsson, J.-E., Lyster, S.-A. H., Snow, C., Hulme, C., Mononen, R.-M., Næss, K.-A. B., López-Pedersen, A., Wie, O. B., & Hagtvet, B. (2021). App-based morphological training produces lasting effects on word knowledge in primary school children : A randomized controlled trial. *Journal of Educational Psychology*. <https://doi.org/10.1037/edu0000688>
- Torres, A. R., Mota, N. B., Adamy, N., Naschold, A., Lima, T. Z., Copelli, M., Weissheimer, J., Pegado, F., & Ribeiro, S. (2021). Selective Inhibition of Mirror Invariance for Letters Consolidated by Sleep Doubles Reading Fluency. *Current Biology*, 31(4), 742-752.e8. <https://doi.org/10.1016/j.cub.2020.11.031>
- Torrey, J. W. (1979). Reading that comes naturally : The early reader. In *Reading research : Advances in theory and practice, Vol. 1* (T. G. Waller & G. E. MacKinnon, p. 115-144). Academic Press.
- Traficante, D., & Burani, C. (2014). List context effects in languages with opaque and transparent orthographies : A challenge for models of reading. *Frontiers in psychology*, 5, 1023. <https://doi.org/10.3389/fpsyg.2014.01023>
- Traficante, D., Marcolini, S., Luci, A., Zoccolotti, P., & Burani, C. (2011). How do roots and suffixes influence reading of pseudowords : A study of young Italian readers with and without dyslexia. *Language and Cognitive Processes*, 26(4-6), 777-793. <https://doi.org/10.1080/01690965.2010.496553>
- Treiman, R. (2017a). Learning to spell : Phonology and beyond. *Cognitive Neuropsychology*, 34(3-4), 83-93. <https://doi.org/10.1080/02643294.2017.1337630>
- Treiman, R. (2017b). Learning to Spell Words : Findings, Theories, and Issues. *Scientific Studies of Reading*, 21(4), 265-276. <https://doi.org/10.1080/10888438.2017.1296449>

- Treiman, R. (2018). Teaching and Learning Spelling. *Child Development Perspectives*, 12(4), 235-239. <https://doi.org/10.1111/cdep.12292>
- Treiman, R., & Kessler, B. (2014). *How children learn to write words* (p. xvii, 395). Oxford University Press.
- Truer, A., & Martray, S. (2018). *Validation d'un entraînement morphologique visant à améliorer l'orthographe lexicale des adolescents dyslexiques-dysorthographiques à l'aide d'une étude expérimentale en cas uniques* [Mémoire d'orthophonie, University of Strasbourg]. <https://www.idref.fr/24289822X>
- Tsesmeli, S. N. (2017). Spelling and Meaning of Compounds in the Early School Years through Classroom Games : An Intervention Study. *Frontiers in Psychology*, 8, 2071. <https://doi.org/10.3389/fpsyg.2017.02071>
- Tunmer, W. E., & Chapman, J. W. (2012). The Simple View of Reading Redux : Vocabulary Knowledge and the Independent Components Hypothesis. *Journal of Learning Disabilities*, 45(5), 453-466. <https://doi.org/10.1177/0022219411432685>
- Uhry, J. K., & Shepherd, M. J. (1993). Segmentation/Spelling Instruction as Part of a First-Grade Reading Program : Effects on Several Measures of Reading. *Reading Research Quarterly*, 28(3), 219-233. JSTOR. <https://doi.org/10.2307/747995>
- Ukrainetz, T. A. (2008). Phonemic awareness instruction for preschoolers : The evidence for pre-phonemic versus phonemic tasks. In *EBP Briefs* (Vol. 2, p. 47-58).
- Ukrainetz, T. A., Nuspl, J. J., Wilkerson, K., & Beddes, S. R. (2011). The effects of syllable instruction on phonemic awareness in preschoolers. *Early Childhood Research Quarterly*, 26(1), 50-60. <https://doi.org/10.1016/j.ecresq.2010.04.006>
- Vágvölgyi, R., Coldea, A., Dresler, T., Schrader, J., & Nuerk, H.-C. (2016). A Review about Functional Illiteracy : Definition, Cognitive, Linguistic, and Numerical Aspects. *Frontiers in Psychology*, 7. <https://www.frontiersin.org/articles/10.3389/fpsyg.2016.01617>
- Vaknin-Nusbaum, V., & Raveh, M. (2019). Cultivating morphological awareness improves reading skills in fifth-grade Hebrew readers. *Journal of Educational Research*, 112(3), 357-366. <https://doi.org/10.1080/00220671.2018.1528541>
- Vaknin-Nusbaum, V., & Sarid, M. (2021). The role of morphological awareness in the development of reading comprehension in Hebrew-speaking second-graders. *Reading and Writing*, 34(10), 2645-2671. <https://doi.org/10.1007/s11145-021-10162-5>
- Valdois, S., Reilhac, C., Ginestet, E., & Line Bosse, M. (2021). Varieties of Cognitive Profiles in Poor Readers : Evidence for a VAS-Impaired Subtype. *Journal of Learning Disabilities*, 54(3), 221-233. <https://doi.org/10.1177/0022219420961332>
- van den Boer, M., van Bergen, E., & de Jong, P. F. (2015). The specific relation of visual attention span with reading and spelling in Dutch. *Learning and Individual Differences*, 39, 141-149. <https://doi.org/10.1016/j.lindif.2015.03.017>
- 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.
- Wang, Castles, A., Nickels, L., & Nation, K. (2011). Context effects on orthographic learning of regular and irregular words. *Journal of Experimental Child Psychology*, 109(1), 39-57. <https://doi.org/10.1016/j.jecp.2010.11.005>

- Wang, Y., & McBride, C. (2016). Beyond Copying : A Comparison of Multi-Component Interventions on Chinese Early Literacy Skills. *International Journal of Behavioral Development*, 41(3), 380-389. <https://doi.org/10.1177/0165025416637212>
- Wang, Z., Sabatini, J., O'Reilly, T., & Weeks, J. (2019). Decoding and reading comprehension : A test of the decoding threshold hypothesis. *Journal of Educational Psychology*, 111(3), 387-401. <https://doi.org/10.1037/edu0000302>
- Wanzek, J., Stevens, E. A., Williams, K. J., Scammacca, N., Vaughn, S., & Sargent, K. (2018). Current Evidence on the Effects of Intensive Early Reading Interventions. *Journal of Learning Disabilities*, 51(6), 612-624. <https://doi.org/10.1177/0022219418775110>
- Wanzek, J., Vaughn, S., Wexler, J., Swanson, E. A., Edmonds, M., & 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.
- Wei, T.-Q., Bi, H.-Y., Chen, B.-G., Liu, Y., Weng, X.-C., & Wydell, T. N. (2014). Developmental changes in the role of different metalinguistic awareness skills in Chinese reading acquisition from preschool to third grade. *PloS One*, 9(5), e96240. <https://doi.org/10.1371/journal.pone.0096240>
- Weiser, B., & Mathes, P. (2011). Using Encoding Instruction to Improve the Reading and Spelling Performances of Elementary Students At Risk for Literacy Difficulties : A Best-Evidence Synthesis. *Review of Educational Research*, 81(2), 170-200. <https://doi.org/10.3102/0034654310396719>
- White, O. R., & Haring, N. G. (1980). *Exceptional teaching* (2^e éd.). Charles E. Merrill.
- Williams, K. J., Walker, M. A., Vaughn, S., & Wanzek, J. (2017). A Synthesis of Reading and Spelling Interventions and Their Effects on Spelling Outcomes for Students With Learning Disabilities. *Journal of Learning Disabilities*, 50(3), 286-297. <https://doi.org/10.1177/0022219415619753>
- Wolff, U., & Gustafsson, J.-E. (2022). Early phonological training preceding kindergarten training : Effects on reading and spelling. *Reading and Writing*, 35(8), 1865-1887. <https://doi.org/10.1007/s11145-022-10261-x>
- Wolter, J. A., & Dilworth, V. (2014). The Effects of a Multilinguistic Morphological Awareness Approach for Improving Language and Literacy. *Journal of Learning Disabilities*, 47(1), 76-85. <https://doi.org/10.1177/0022219413509972>
- Wu, X., Anderson, R., li, W., Wu, X., Li, H., Zhang, J., Zheng, Q., Zhu, J., Shu, H., Jiang, W., Chen, X., Wang, Q., Yin, L., He, Y., Packard, J., & Gaffney, J. (2009). Morphological Awareness and Chinese Children's Literacy Development : An Intervention Study. *SCIENTIFIC STUDIES OF READING*, 13, 26-52. <https://doi.org/10.1080/10888430802631734>
- Zesiger, P. (2004). Neuropsychologie développementale et dyslexie. *Enfance*, 56(3), 237-243. <https://doi.org/10.3917/enf.563.0237>
- Zhang, D., & Li, L. (2016). Morphological Teaching and Singaporean Children's English Word Learning. In R. Elaine Silver & W. D. Bokhorst-Heng (Éds.), *Quadrilingual Education in Singapore* (p. 199-218). Springer Singapore. https://doi.org/10.1007/978-981-287-967-7_12
- Ziegler, J. C. (2018). Différences inter-linguistiques dans l'apprentissage de la lecture. *Langue française*, 199(3), 35-49. <https://doi.org/10.3917/lf.199.0035>
- Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments & Computers*, 28(4), 504-515. <https://doi.org/10.3758/BF03200539>

- Ziegler, J. C., Perry, C., & Coltheart, M. (2003). Speed of lexical and nonlexical processing in French : The case of the regularity effect. *Psychonomic Bulletin & Review*, *10*(4), 947-953. <https://doi.org/10.3758/BF03196556>
- Ziegler, J. C., Stone, G. O., & Jacobs, A. M. (1997). What is the pronunciation for -ough and the spelling for /u/? A database for computing feedforward and feedback consistency in English. *Behavior Research Methods, Instruments, & Computers*, *29*(4), 600-618. <https://doi.org/10.3758/BF03210615>
- Zinna, D. R., Liberman, I. Y., & Shankweiler, D. (1986). Children's Sensitivity to Factors Influencing Vowel Reading. *Reading Research Quarterly*, *21*(4), 465. <https://doi.org/10.2307/747617>
- Zoski, J. L., & Erickson, K. A. (2017). Multicomponent Linguistic Awareness Intervention for At-Risk Kindergarteners. *Communication Disorders Quarterly*, *38*(3), 161-171. <https://doi.org/10.1177/1525740116660817>
- Zygomatic - Board Game Studio. (2019). *Dobble*. Dobble. <https://www.dobblegame.com/en/homepage/>

Impacts d'entraînements ciblant les unités morphologiques et orthographiques sur le développement de l'identification des mots et de la production orthographique des enfants avec et sans trouble de la lecture

Résumé

La maîtrise de la lecture et de l'écriture constitue un enjeu particulièrement critique pour les individus présentant un trouble de la lecture. Outre les correspondances graphème-phonème, indispensables à l'instauration d'une procédure d'auto-apprentissage, il existe des unités morphologiques et graphotactiques qui concourent au développement de l'identification des mots et à la production orthographique. La présente thèse visait à évaluer l'impact d'entraînements ciblant ces unités sur le développement de l'identification et de la production orthographique d'enfants avec ou sans trouble de la lecture.

Pour cela, nous avons réalisé une méta-analyse sur 40 études indépendantes portant sur les entraînements morphologiques (EM), qui retrouve un impact positif de l'EM pour la conscience morphologique, la production orthographique de morphèmes, la production orthographique de mots et la lecture morphologique (i.e., la lecture de mots présentant au moins deux morphèmes). Nous avons par ailleurs conduit une étude d'entraînement morphologique impliquant 51 enfants scolarisés de la 3^{ème} à la 5^{ème} année d'instruction. Nos résultats indiquent que l'entraînement à la production orthographique de morphèmes, qu'il soit ou non complété d'exercices à l'analyse morphologique, améliore l'identification de mots irréguliers et la production orthographique de mots complexes, comparativement à un entraînement aux mathématiques. Enfin, nous avons conduit une étude d'entraînement orthographique sublexical ciblant les unités phonologiques, morphologiques et graphotactiques, dispensé chez 51 participants, selon un mode soit explicite soit implicite. Des mesures répétées ont permis de dégager différents « profils d'apprenants », et suggèrent une efficacité différenciée en fonction du mode d'administration et du type d'apprenant (avec ou sans trouble de la lecture ; niveau de scolarisation).

Nos résultats confirment l'intérêt de considérer les unités morphologiques et orthographiques pour soutenir l'apprentissage de l'identification des mots et de la production orthographique. Toutefois, en raison de biais méthodologiques, des études ultérieures sont nécessaires pour compléter et confirmer ces résultats.

Mots-clés : entraînement morphologique, entraînement orthographique, identification des mots, production orthographique

Résumé en anglais

Achieving reading and writing mastery is a critical issue for individuals with a reading disorder. Besides grapheme-to-phoneme correspondences (CGP) which are essential to the acquisition of a self-teaching procedure, there are morphological and graphotactic units that concur to word identification and spelling development. The present thesis aimed to evaluate the impact of training those units on the development of word identification and spelling in children with or without a reading disorder.

To that end, we conducted a meta-analysis of 40 independent morphological training (MT) studies and found a positive impact of MT on morphological awareness, morpheme spelling, word spelling and morphological reading. We also conducted a MT study involving 51 3rd-to-5th grade schoolchildren. Our results show morpheme spelling training, whether or not it comprises morphological analysis training, seems to benefit irregular word identification and complex word spelling, when compared to a mathematical training. Lastly, we conducted a sublexical orthographic training targeting phonological, morphological and graphotactic units in 51 participants, which we delivered either explicitly or implicitly. Repeated measures analysis seemed to outline different "learning profiles", suggesting a differential efficiency of the program, depending on delivery mode and on the type of learner (with or without a reading disorder; class level).

All in all, our results appear to confirm that morphological and orthographic units are relevant in supporting the development of word identification and spelling. However, due to methodological bias, further studies are required to complete and confirm such results.

Keywords: morphological training, orthographic training, word identification, spelling.