

Université Paris Descartes  
Laboratoire Mémoire et Cognition

**Thèse de doctorat de l'Université Paris Descartes**  
**Gaëlle BORCHARDT**  
**SOUS LA DIRECTION DE SEBASTIEN PACTON**

# **L'influence des connaissances graphotactiques sur l'acquisition de l'orthographe lexicale : étude chez l'enfant d'école élémentaire et chez l'adulte.**

Membres du jury :

**Michel FAYOL** : Université Blaise Pascal, Clermont-Ferrand. (Examineur)

**Jean-Noël FOULIN** : Université Bordeaux Segalen, Bordeaux. (Rapporteur)

**Sébastien PACTON** : Université Paris Descartes, Paris. (Directeur)

**Pierre PERRUCHET** : Université de Bourgogne, Dijon. (Rapporteur)

- 29 novembre 2012 -

## Remerciements

A Sébastien Pacton, qui m'a encadrée durant ces trois années, pour son aide et ses précieux conseils apportés lors des différentes étapes de cette thèse.

A Michel Fayol, Jean-Noël Foulin et Pierre Perruchet pour avoir accepté de participer au jury de cette thèse.

Aux enseignants et aux directeurs des écoles, pour leur collaboration et le temps qu'ils m'ont accordé afin de réaliser les différentes passations individuelles et collectives.

Aux élèves et aux adultes, qui ont participé aux expériences.

A mes collègues et amis, Jennifer, Servane, Yannick, Marion, ainsi qu'à ma famille, pour leur soutien et leur présence au quotidien.

Aux membres du Laboratoire Mémoire et Cognition.

# SOMMAIRE

INTRODUCTION.....	6
CHAPITRE THEORIQUE .....	9
I) Les systèmes d'écritures alphabétiques.....	9
II) Modèles du langage écrit adulte (scripteurs experts).....	10
III) Critiques des modèles à doubles voies.....	12
IV) L'apprentissage de l'orthographe .....	14
1) Modèles en stades de l'acquisition de l'orthographe.....	14
2) L'apprentissage de l'orthographe de mots spécifiques.....	18
2.1) Effet de fréquence.....	18
2.2) Les écritures par analogie.....	19
2.3) L'hypothèse d'auto-apprentissage.....	21
2.4) Variables influençant l'apprentissage orthographique dans des situations de <i>self-teaching</i> .....	24
2.5) Le rôle des connaissances orthographiques dans l'apprentissage orthographique dans des situations de <i>self-teaching</i> .....	26
3) L'apprentissage de régularités graphotactiques.....	29
3.1) Connaissances graphotactiques examinées seules.....	29
3.2) De l'apprentissage implicite en laboratoire à l'apprentissage implicite de régularités graphotactiques.....	35
3.3) L'influence des régularités graphotactiques même lorsque le recours à des règles morphologiques est possible.....	40
3.4) Sensibilité aux régularités graphotactiques et apprentissage de l'orthographe lexicale.....	42
CHAPITRE EXPERIMENTAL I .....	46
Méthode.....	47
Participants.....	47
Matériel .....	48
Procédure.....	50
Résultats .....	51
1. Tests standards de lecture et orthographe.....	51
2. Epreuve de décision orthographique sur des mots.....	51
3. Epreuve de jugement de l'orthographe de pseudo-mots évaluant les connaissances graphotactiques.....	52
3.1. Identité des consonnes pouvant être doublées.....	52
3.2. Illégalité des doublets en position initiale.....	54
3.3. Illégalité des doublets après une consonne simple.....	55
3.4. Comparaison des scores pour Illégalité des doublets en position initiale et Illégalité des doublets après une consonne simple.....	56
3.5 Eau en position finale en fonction de l'environnement gauche .....	57
Discussion .....	58

CHAPITRE EXPERIMENTAL II .....	59
Méthode.....	61
Participants.....	61
Matériel .....	61
Procédure.....	65
Résultats .....	66
1. Sensibilité aux régularités graphotactiques.....	66
2. Analyses de corrélation et régressions .....	68
Discussion .....	72
CHAPITRE EXPERIMENTAL III.....	76
Expérience 1 .....	80
Méthode.....	80
Participants.....	80
Matériel .....	80
Procédure.....	82
Résultats .....	82
Groupe expérimental.....	82
Groupe contrôle.....	87
Comparaisons entre les groupes expérimentaux et contrôle .....	87
Discussion .....	88
Expérience 2.....	91
Méthode.....	91
Participants.....	91
Matériel .....	91
Procédure.....	92
Résultats .....	92
Condition rappel.....	92
Condition reconnaissance.....	95
Discussion .....	97
Expérience 3.....	98
Méthode.....	98
Participants.....	98
Matériel .....	99
Matériel pour tester l'apprentissage de l'orthographe de non-mots.....	99
Matériel pour évaluer la connaissance de l'orthographe de mots spécifiques .....	99
Procédure.....	99
Résultats .....	100
Discussion .....	104

Expériences 4 et 5 .....	105
Expérience 4.....	106
Méthode.....	106
Participants.....	106
Matériel .....	106
Procédure.....	106
Résultats .....	107
Groupe expérimental, condition rappel.....	112
Groupe expérimental, condition reconnaissance.....	114
Groupe contrôle.....	117
Discussion .....	118
Expérience 5.....	121
Méthode.....	121
Participants.....	121
Procédure.....	121
Résultats .....	122
Condition rappel.....	123
Condition reconnaissance.....	126
Discussion .....	130
CHAPITRE EXPERIMENTAL IV .....	131
Méthode.....	132
Participants.....	132
Matériel .....	132
Procédure.....	134
Résultats .....	135
Connaissances en orthographe lexicale et connaissances graphotactiques.....	135
Apprentissage de l'orthographe des non-mots — Condition rappel.....	137
Apprentissage de l'orthographe des non-mots — Condition reconnaissance.....	143
Discussion .....	147
DISCUSSION GENERALE .....	149
1) Etude des connaissances graphotactiques.....	149
2) Rôle des connaissance graphotactique dans l'apprentissage de l'orthographe lexicale	
.....	151
Conclusion.....	160
BIBLIOGRAPHIE .....	162
ANNEXES .....	172

## INTRODUCTION

De nombreux mots français incluent des phonèmes pouvant s'orthographier de différentes façons (e.g., /o/ dans *moto*, *bateau*, *landau*, *mot*) ou des lettres muettes, notamment en fin de mots (e.g., *buvard*, *bavard*). Une difficulté spécifique du système orthographique français concerne le doublement de certaines consonnes qui, le plus souvent, n'a pas de contrepartie phonologique (Jaffré et Fayol, 1997). Ainsi, les suites de lettres «ule» / «ulle» et «bale»/ « balle », se prononcent de façon strictement identique dans les mots *formule* / *bulle* et *timbale* / *balle*. Ces caractéristiques expliquent que l'application de correspondances phonèmes - graphèmes ne permet pas d'orthographier correctement plus de la moitié des mots français (Véronis, 1988 ; Ziegler, Jacobs, et Stone, 1996). La nécessité d'apprendre l'orthographe de nombreux mots a souvent été soulignée (Baron et Treiman, 1980) mais il a également été suggéré que la mémorisation de l'orthographe de mots spécifiques pourrait (en partie au moins) être évitée en se fondant sur des régularités relatives à la fréquence de combinaisons des graphèmes appelées régularités graphotactiques (e.g., Treiman et Kessler, 2006).

Les recherches sur l'acquisition de connaissances relatives à l'orthographe de mots spécifiques (orthographe lexicale) ont étudié si les élèves étaient influencés par la fréquence des mots et les régularités des correspondances phonème – graphème et s'ils utilisaient des mots connus pour orthographier par analogie d'autres mots ou des non-mots (pour des revues, voir Bosse & Pacton, 2006 ; Foulon, 2005). Selon Share (1995, 1999, 2004), l'acquisition de représentations orthographiques spécifiques résulte essentiellement du décodage phonologique. Toutefois, les habiletés de décodage phonologique ne semblent pas pouvoir expliquer à elles seules l'apprentissage de l'orthographe lexicale (Castle & Nation, 2006 ; Cunningham, Perry, Stanovich & Share, 2002 ; Cunningham, 2006 ; Nation, Angell & Castles, 2007). La part de variance inexplicée a souvent été attribuée à des habiletés "générales" de traitement orthographique (Berninger, 1995) mais cela a été soit suggéré mais non testé (e.g., Nation et al., 2007) soit testé mais de façon inappropriée (e.g., Cunningham et al., 2006).

Les recherches sur l'apprentissage de régularités graphotactiques ont étudié la sensibilité des enfants à diverses régularités graphotactiques, certaines de type probabiliste (e.g., que *eau* est plus fréquent après des consonnes comme *t* qu'après des consonnes comme *f*, Pacton, Fayol & Perruchet, 2002, 2005), d'autres descriptibles sous forme de règle (e.g., les consonnes ne peuvent pas être doublées en début de mots en Anglais et en Français, Cassar & Treiman, 1997 ; Pacton et al., 2001). L'apprentissage de ces régularités a parfois été étudié en adaptant certains paradigmes d'étude de l'apprentissage implicite (e.g., apprentissage de grammaires artificielles) pour étudier si les sujets apprennent implicitement des règles après une pratique beaucoup plus prolongée (au moins 5 ans chez des élèves de CM2) que dans les situations de laboratoire (souvent 15 – 30 minutes, beaucoup plus rarement quelques heures) (voir Pacton & Perruchet, 2006 ; Perruchet & Pacton, 2004).

Les diverses sources d'information utilisées dans l'acquisition de l'orthographe ont typiquement été étudiées indépendamment les unes des autres (Lennox & Siegel, 1994 ; Deacon, Conrad & Pacton, 2008). Toutefois, les études ayant exploré l'intégration de différentes sources d'information suggèrent que les individus se fondent sur leur sensibilité aux régularités graphotactiques même lorsque le recours à des règles sémantiquement fondées est possible. Par exemple, des enfants et des adultes anglophones semblent se fonder sur leur sensibilité à des régularités graphotactiques même lorsqu'ils peuvent se fonder sur des règles morphologiques s'appliquant à un large ensemble de mots comme celle spécifiant qu'en Anglais le pluriel des noms s'orthographie avec *s* qu'il se prononce /s/ ou /z/ (e.g., Deacon & Pacton, 2007 ; Kemp & Bryant, 2003 ; Pacton & Deacon, 2008). Concernant l'intégration des connaissances orthographiques lexicales et des connaissances graphotactiques, seules trois études ont suggéré un impact des connaissances graphotactiques sur l'acquisition de l'orthographe de nouveaux mots (Campbell & Coltheart, 1984, Fayol et al. 2010, Wright & Ehri, 2007). Ces études doivent néanmoins être poursuivies car elles présentent certaines limites (détaillées dans la partie théorique) comme l'utilisation d'un seul item dans l'étude de Campbell et Coltheart (1984).

L'objectif de cette thèse est de mieux comprendre l'intégration des connaissances orthographiques spécifiques et générales dans l'acquisition de l'orthographe lexicale. Le chapitre théorique commence par la présentation des modèles adultes de la production orthographique, d'études sur l'acquisition de l'orthographe lexicale puis d'études ayant mis en évidence une connaissance, souvent précoce, de diverses régularités graphotactiques. Le chapitre théorique se poursuit par la présentation d'études montrant que les connaissances

graphotactiques sont utilisées alors même que des règles phono-graphologiques ou morphologiques peuvent être utilisées et que les connaissances graphotactiques ont un impact sur l'acquisition de l'orthographe de nouveaux mots. Les différentes expériences réalisées dans cette thèse sont ensuite regroupées en quatre chapitres expérimentaux. Un dernier chapitre est consacré à la discussion générale des résultats rapportés dans les chapitres expérimentaux et à la conclusion.

Le premier objectif du chapitre expérimental I était de préciser les connaissances graphotactiques d'élèves de l'école élémentaire, notamment en confirmant que certaines régularités graphotactiques sont plus difficilement apprises que d'autres. Le second objectif était d'étudier le degré d'abstraction des connaissances graphotactiques pour des propriétés de l'emploi des doubles lettres descriptibles sous la forme de règles chez des adultes dont la pratique de l'écrit est encore plus importante que chez les sujets les plus âgés des études antérieures (i.e., CM2 pour Danjon & Pacton, 2009 ; Pacton et al., 2001).

L'objectif de l'expérience rapportée dans le chapitre expérimental II était de préciser le rôle des connaissances orthographiques lexicales et celui des connaissances graphotactiques lors de l'apprentissage implicite de l'orthographe de non-mots insérés dans des textes lus par des élèves de CM1 et CM2 (situation de *self-teaching*, Share, 1995, 1999). Des analyses de régression ont été menées afin de déterminer si ces deux types de connaissances orthographiques expliquaient des parts de variance indépendantes du niveau d'apprentissage de l'orthographe des non-mots.

Les six expériences rapportées dans les chapitres expérimentaux III et IV ont examiné si et comment les connaissances graphotactiques influencent l'apprentissage de l'orthographe lexicale dans diverses situations d'apprentissage de non-mots en manipulant le degré de convergence entre les orthographe des non-mots et les régularités graphotactiques de la langue française. Deux propriétés graphotactiques ont été manipulées, l'une de type probabiliste (certaines consonnes sont plus fréquemment doublées que d'autres), l'autre descriptible sous forme de règle (les consonnes doubles peuvent survenir avant mais pas après les consonnes simples) dans diverses situations d'apprentissage (implicite vs. explicite ; non-mots insérés dans des textes vs. présentés isolément ; lecture silencieuse vs. à haute voix) et de test (rappel vs. reconnaissance).



## CHAPITRE THEORIQUE

### D) Les systèmes d'écritures alphabétiques

Le système orthographique français est fondé sur le principe alphabétique. Selon ce principe, les phonèmes, les plus petites unités de la parole, sont représentés à l'écrit par des graphèmes, les plus petites unités orthographiques. Dans un système alphabétique idéal, les correspondances entre phonèmes et graphèmes seraient biunivoques, un phonème correspondrait à un seul graphème et réciproquement (Gak, 1976). En fait, parmi les langues à base alphabétique, les correspondances entre phonèmes et graphèmes sont plus ou moins régulières. Dans les langues dites transparentes, comme l'Italien ou l'Espagnol, les correspondances phonèmes-graphèmes (CPG), relatives à l'écriture, et les correspondances graphèmes-phonèmes (CGP), relatives à la lecture, sont assez consistantes. En Italien et en Espagnol, le nombre de phonèmes et le nombre de graphèmes diffèrent peu. En revanche, dans des langues dites opaques, comme l'Anglais ou le Français, il y a beaucoup plus de graphèmes que de phonèmes (Sprenger-Charolles, 2003). En Anglais, les CPG et CGP sont inconsistantes. En Français les CGP sont beaucoup plus consistantes que les CPG. Ainsi, de nombreux mots français, qui ne comportent aucune ambiguïté concernant la façon de les lire, peuvent être orthographiés de différentes façons (manteau – menteau – mantau – mentot). En fait, l'application des seules règles de CPG permet d'écrire seulement la moitié des mots français, (Véronis, 1988; Ziegler, Jacobs, & Stone, 1996). Cette spécificité du français s'ajoute à la différence de tâche, commune elle à toutes les langues, qui est que la lecture est une tâche de reconnaissance alors que l'orthographe est une tâche de rappel. De ce fait, s'il est souvent possible de lire un mot correctement en prenant des indices partiels, il en va tout autrement dans le cas de l'orthographe puisque l'intégralité des lettres formant un mot doit être restituée. Ceci explique pourquoi il est fréquent de trouver des enfants normo-lecteurs qui rencontrent des difficultés en orthographe alors que le patron de performances opposé est beaucoup plus rare, voire inexistant (Frith, 1980) et pourquoi les difficultés des dyslexiques persistent souvent davantage en orthographe qu'en lecture (Frith, 1985 ; Nicolson & Fawcet, 1994 ; Snowling, 2000).

Dans une étude comparative entre enfants italiens et anglais, Thorstad (1991), a mis en évidence des compétences en lecture et orthographe nettement supérieures chez les enfants italiens dès 6 ans. Même lorsque l'analyse des productions orthographiques portait sur huit mots ayant une écriture similaire dans les deux langues (e.g. special en Anglais et speciale en Italien), 90% des productions étaient correctes en italien contre seulement 19% en Anglais. Ces difficultés de traitement et d'apprentissage en fonction de la transparence des relations graphèmes – phonèmes et phonèmes – graphèmes ont été confirmées dans des études comparant la vitesse d'apprentissage de la lecture. Ainsi, Seymour, Aro et Erskine (2003) ont montré que l'apprentissage de la lecture en Anglais prend deux fois plus de temps que dans des langues transparentes.

## **II) Modèles du langage écrit adulte (scripteurs experts)**

Les modèles à deux voies (e.g., Ellis & Young, 1988; Roeltgen & Heilman, 1984) sont les plus utilisés pour rendre compte des performances des scripteurs experts. De nombreux tests sont construits afin de tester l'efficacité des deux voies postulées par ces modèles. La voie non lexicale (ou indirecte) ou procédure d'assemblage utilise des règles de traduction phonème-graphème. La voie lexicale (ou directe) ou procédure d'adressage récupère l'orthographe des mots connus dans un lexique orthographique.

Selon le modèle d'écriture à deux voies d'Ellis et Young (1988) représenté à la Figure 1, la production de mots par la voie lexicale se caractérise par l'activation de connaissances sur la forme orthographique de mots spécifiques. Lors de la dictée d'un mot, la forme phonologique du mot serait activée dans le lexique phonologique d'entrée, et de là, il y aurait accès à la forme orthographique de ce mot qui est stockée dans le lexique orthographique de sortie. Cet accès pourrait se faire via le système sémantique (accès au sens) ou directement du lexique phonologique d'entrée au lexique orthographique de sortie, ce qui peut conduire à des erreurs dans le cas de la production de mots homophones (e.g., vers/ verre / vert). Le mot doit évidemment faire partie du lexique orthographique du scripteur pour pouvoir être produit par la procédure lexicale. Dans le cas inverse, la voie non lexicale serait mobilisée. Les mots entendus seraient d'abord segmentés en phonèmes, puis ces phonèmes seraient traduits en graphèmes (en utilisant le graphème le plus fréquent de la

langue) et enfin, lors de l'étape d'assemblage, les différents graphèmes seraient assemblés pour former une séquence orthographique qui serait maintenue en mémoire de travail (dans un buffer graphémique) le temps nécessaire à la transcription écrite. Les mots *irréguliers* (e.g. *oignon*, *femme*) ne pourraient être orthographiés que par la procédure d'adressage car l'application de CPG conduirait à des erreurs phonologiquement plausibles comme OGNON ou FAME. Inversement, les mots jamais rencontrés à l'écrit ne pourraient être retranscrits qu'en utilisant la procédure d'assemblage. L'efficacité de la procédure d'assemblage est évaluée avec des dictées de mots inventés qui respectent les contraintes phonotactiques de la langue appelés non-mots ou pseudo-mots dans les expériences de psychologie expérimentale et dans les tests (e.g., BELEC ; Mousty, Leybaert, Alegria, Content & Morais, 1994).

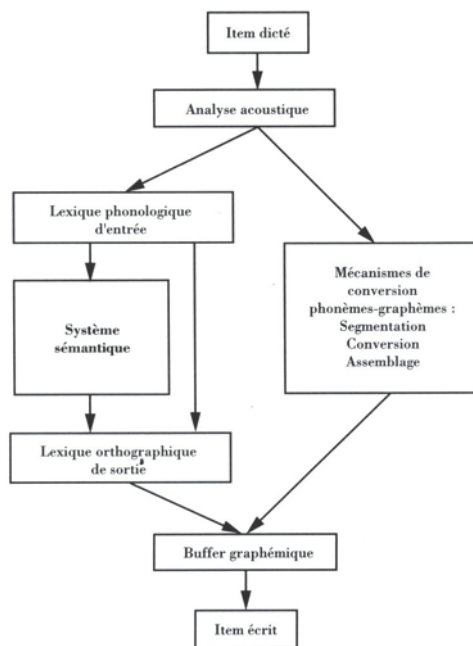


Figure 1 : Modèle à double voie (Ellis et Young, 1988)

Un argument en faveur des modèles à double voie est que, suite à une atteinte du système nerveux central, des adultes préalablement bons en lecture et orthographe ont des capacités d'écriture diminuées qui suggèrent que l'une des deux procédures peut être sélectivement altérée (e.g. Barry, 1994 ; Bonin, 2007 ; Tainturier & Rapp, 2001). Ainsi, les dysgraphiques phonologiques, chez qui c'est la procédure d'assemblage qui est

sélectivement atteinte, parviennent à produire des orthographe correctes pour des mots connus, réguliers et irréguliers, mais ne peuvent plus produire des orthographe plausibles de non-mots sous dictée (e.g. Shallice, 1981). Les difficultés concernent à la fois la segmentation et la conversion chez certains patients, comme pour le patient rapporté par Shallice, et seulement des difficultés de segmentation chez d'autres patients (Bolla-Wilson et al., 1985 ; Roeltgen et al., 1983). Inversement, les dysgraphiques de surface, chez qui c'est la procédure d'adressage qui est sélectivement atteinte, ont une capacité à orthographier correctement des mots réguliers et des non-mots relativement préservée (Beauvois & Dérouesné, 1981 ; Goodman & Caramazza, 1986 ; Hatfield & Patterson, 1983) mais ils ne parviennent plus à orthographier correctement les mots irréguliers. Pour ces items, ils produisent des erreurs phonologiquement plausibles, comme le mot *sœur* orthographié SEUR par le patient RG rapporté par Beauvois et Dérouesné (1981). Si ces dysgraphiques de surface utilisent très majoritairement les CPG les plus fréquentes de la langue, il leur arrive d'utiliser des correspondances plus rares, ce qui suggère que la procédure d'assemblage dispose de plusieurs correspondances pour les phonèmes à transcriptions multiples.

### **III) Critiques des modèles à doubles voies.**

Pour certains, l'existence de doubles dissociations, avec des patients dysgraphiques phonologiques qui ont une atteinte sélective de la procédure d'assemblage et des patients dysgraphiques de surface qui ont une atteinte sélective de la procédure d'adressage, montrerait que ces deux voies sont fonctionnellement indépendantes. Cette indépendance a néanmoins été remise en cause par des études conduites chez des sujets tout venants. Ainsi, Campbell (1983) a suggéré que l'écriture de non-mots ne s'effectuait pas indépendamment des connaissances lexicales avec une technique d'amorçage lexical. Des sujets adultes, entendaient des suites de mots et des non-mots. Ils devaient ignorer les mots et écrire les non-mots. Les listes d'items étaient construites de sorte que des non-mots comme /prein/ étaient précédés de mots comme *crane* ou *brain*. Campbell a montré que l'orthographe de non-mots différait en fonction des mots préalablement entendus. Par exemple /prein/ était plus souvent orthographié PRAIN après avoir entendu *brain* mais était plus souvent orthographié PRANE après avoir entendu *crane*. Il est à noter que cet effet d'amorçage est observé même lorsque le mot amorce ne précède pas immédiatement le non-mot mais qu'un autre mot distracteur est intercalé entre le mot amorce et le non-mot. De plus, Campbell a rapporté que cet effet d'amorçage était absent chez un patient dysgraphique de surface (E.E)

incapable d'écrire des mots connus.

Barry et Seymour (1988) ont repris le paradigme de Campbell (1983) et ont ajouté une condition dans laquelle des sujets adultes devaient écrire les mêmes non-mots non précédés de mots amorces. Comme dans l'étude de Campbell, un amorçage lexical était observé avec l'écriture des non-mots qui était biaisée par les mots précédemment entendus. Néanmoins, dans les deux situations expérimentales, la fréquence d'utilisation des différentes graphies était également influencée par la fréquence des associations phonèmes-graphèmes, sans interaction entre la fréquence des associations phonèmes-graphèmes et l'amorçage lexical. Dans le modèle interactif de Barry et Seymour, il existerait un pool de correspondances phonèmes-graphèmes hiérarchisées par leur fréquence et leur sélection serait influencée par le lexique orthographique.

Ainsi, les modèles à double voie supposent que tout individu dispose de deux procédures, l'une d'adressage, l'autre d'assemblage afin d'écrire tout type de mots. L'existence de patients chez qui chacune de ces procédures est sélectivement altérée est souvent avancée comme un argument en faveur d'une indépendance fonctionnelle. Toutefois, des données expérimentales ne remettent pas nécessairement en cause l'existence de ces deux voies mais suggèrent qu'elles interagissent. De plus, même des études en neuropsychologie chez des patients dysgraphiques corroborent l'idée d'une activation en parallèle des voies lexicales et non lexicales (Miceli, Capasso & Caramazza, 1994 ; Rapp, Epstein & Tainturier, 2002 ; Laiacina, Capitani, Zonca & Scola, 2009). Par exemple Rapp et al (2002) rapportent le cas du patient LAT qui produisait des erreurs phonologiquement plausibles mais contenant des graphèmes peu fréquents (e.g. « Bouquet » écrit « Bouket »), ainsi il y aurait activation de la voie non lexicale (production d'erreurs phonologiquement plausibles) et de la voie lexicale (production de graphèmes rares).

## IV) L'apprentissage de l'orthographe

### 1) Modèles en stades de l'acquisition de l'orthographe.

Il existe un certain nombre de modèles en stades de l'apprentissage de l'orthographe postulant tous une succession de stades caractérisés par la mise en place de différents processus cognitifs (Ehri, 1986, 1992, 1997 ; Frith, 1985 ; Gentry, 1982 ; Henderson, 1985). Ces modèles accordent une place primordiale aux traitements phonologiques au début de l'acquisition de l'orthographe et n'envisagent pas d'influence des connaissances orthographiques (relatives à l'orthographe de mots spécifiques ou relatives à des régularités orthographiques) au tout début de l'acquisition du langage écrit (Deacon, Conrad & Pacton, 2008 ; Pacton & Deacon, 2008 ; Pollo, Treiman et Kessler, 2008). Le caractère successif des stades a fait l'objet de critiques (e.g., Treiman & Cassar, 1997) mais ces modèles ont engendré un nombre important de recherches visant justement à tester les prédictions de ces modèles. Nous détaillons ci-dessous le modèle de Frith (1985) qui a la particularité de rendre compte de l'apprentissage de la lecture et de l'orthographe (voir Tableau 1).

Tableau 1 : Modèle de U. Frith (1985) concernant le développement de la lecture et de l'écriture ainsi que l'apparition de troubles dans ces habiletés

Stade	Lecture	Ecriture	Troubles
1a	logographique	(symbolique)	
1b	logographique	logographique	dyslexie "dysphonétique"
2a	logographique	alphabétique	
2b	alphabétique	alphabétique	dyslexie "dyséidétique" dysorthographe type A
3a	orthographique	alphabétique	
3b	orthographique	orthographique	dysorthographe type B

Le premier stade postulé par Frith s'appelle le stade *logographique*. A ce stade, les enfants reconnaîtraient instantanément quelques mots familiers, en se fondant sur des indices prégnants comme les jambages, la première lettre des mots, voire des indices extralinguistiques (e.g., le rouge et le blanc de *Coca-Cola*). L'enfant n'a pas encore compris le principe alphabétique, il n'y a donc pas de médiation phonologique. Morton (1980) parle d'un système picturo-sémantique car le mot écrit est traité comme une image qui est associée au mot oral, ce qui fait que *Coca-Cola* écrit en noir ne sera pas reconnu mais *Coda-Coba* écrit en rouge et blanc le sera. Jusqu'à une centaine de mots pourraient être lus avec cette stratégie (Seymour & Elder, 1986 ; Smith, 1971). En orthographe, les enfants pourraient écrire quelques mots familiers (e.g., papa, maman, leur prénom). Les erreurs ne seraient pas phonologiquement plausibles, l'ordre et l'orientation des lettres ne seraient pas respectés (e.g., juile pour Julie ; d à la place de b ou l'inverse) et les productions pourraient inclure des formes ressemblant à des lettres ou à des nombres (Henderson, 1985).

Au stade suivant, dit *alphabétique*, les enfants prendraient conscience des relations entre des unités orales et des unités écrites et appliqueraient des CPG en orthographe et des CGP en lecture. La stratégie alphabétique serait d'abord utilisée en orthographe, la lecture demeurant logographique dans un premier temps, puis la stratégie alphabétique serait utilisée à la fois en lecture et en orthographe. L'idée selon laquelle la procédure alphabétique serait d'abord utilisée en écriture a été étayée par un certain nombre d'études (Ellis, 1994). Ainsi, dans certaines études longitudinales, la conscience phonémique, cruciale pour mobiliser une stratégie alphabétique, était davantage corrélée avec le niveau en orthographe qu'avec le niveau en lecture au début de l'acquisition du langage écrit (e.g. Cataldo & Ellis, 1988) et l'effet bénéfique d'entraînement visant à développer la conscience phonologique était plus précoce et plus marqué sur l'écriture que sur la lecture (e.g. Bradley & Bryant, 1983). Il a également été rapporté que certains enfants sont capables d'écrire des mots qu'ils sont incapables de relire (e.g. Bryant & Bradley, 1980), mais l'inverse a également été rapporté (Ehri, 1997).

L'application correcte de règles de CPG permet l'écriture de mots jamais rencontrés et conduit à des erreurs phonologiquement plausibles (e.g., *gâteau* écrit *gato*). Toutefois, certains auteurs ont souligné qu'au début, même si les enfants ont compris le principe

alphabétique, ils rencontreraient des difficultés pour représenter complètement la structure phonologique des mots, notamment les groupes consonantiques (e.g., *tr* et *pl* dans *tri* et *plein*) car ils ne perçoivent pas, ou mal, les deux consonnes comme des sons distincts (e.g., Bruck & Treiman, 1990 ; Ehri, 1986 ; Gentry, 1982 ; Treiman, 1993 ; Treiman, Zukowski, & Richmond-Welty, 1995). A ce niveau, les enfants pourraient aussi utiliser une stratégie dite *nom de lettre* (Henderson, 1985) qui consiste à utiliser une lettre pour représenter le son du nom de cette lettre plutôt que le son de cette lettre. Ainsi, des enfants américains écriraient le mot *car* CR avec la lettre *r* pour représenter /ar/ (Ellefson, Treiman, & Kessler, 2009, voir Foulin, 2005 pour une revue).

Dans le modèle de Frith, et plus généralement dans l'ensemble des modèles en stades (e.g., Henderson, 1985), à ce niveau, les enfants n'utiliseraient pas d'autres informations que celles relatives aux CPG ; ils ne prendraient par exemple pas en compte le fait que certaines séquences de lettres sont plus fréquentes que d'autres.

Le stade suivant est le stade *orthographique* dans le modèle de Frith (1985). Il correspond à la mise en place de la voie d'adressage dans les modèles à double voie et serait la conséquence de l'automatisation de la lecture alphabétique (Share, 1995). Il est important de ne pas confondre ce stade avec le stade logographique : la reconnaissance des mots s'effectuerait via un système pictural au stade logographique mais via une analyse linguistique au stade orthographique. Selon Frith, la stratégie *orthographique* est d'abord utilisée seulement pour la lecture, puis pour la lecture et l'orthographe. Lecture orthographique et écriture alphabétique coexistent donc dans un premier temps. Une fois atteint le stade orthographique, les enfants écriraient en utilisant des connaissances sur l'orthographe de mots spécifiques, sur des patterns orthographiques fréquents et sur les morphèmes. C'est également à ce stade que les enfants utiliseraient des mots dont ils connaissent l'orthographe pour orthographier d'autres mots ou des non-mots par analogie. Une fois automatisée, ce stade orthographique est celui de l'expertise.

Un intérêt du modèle de Frith est de rendre compte de l'existence de différents troubles dans ces habiletés : la dyslexie *dysphonétique* chez des élèves ne parvenant pas à atteindre le stade alphabétique ; la dyslexie *dyséidétique* et la dysorthographe de type A chez des élèves ne parvenant pas à atteindre le stade orthographique en lecture et en



orthographe ; la dysorthographie de type B chez des élèves parvenant à atteindre le stade orthographique en lecture mais pas en orthographe. Selon Frith (1985), les représentations lexicales doivent être suffisamment précises pour que leur utilisation en lecture se transfère à l'orthographe. C'est pourquoi le stade orthographique est atteint d'abord en lecture, puis en orthographe et c'est aussi pourquoi le transfert de la stratégie orthographique de la lecture à l'écriture n'aurait pas lieu chez des enfants bons en lecture mais faibles en orthographe qui prendraient des indices partiels (Frith, 1980). On retrouve des idées similaires dans le modèle à double fondation de Seymour (1997) dans lequel le développement de connaissances orthographiques requiert également de bonnes connaissances alphabétiques (phonographémiques) puisque, pour devenir orthographiques, les représentations logographiques doivent être redéfinies via les processus alphabétiques.

L'importance des traitements alphabétiques dans les premières phases de l'acquisition de l'orthographe est bien établie. Les orthographes inventées des jeunes enfants se fondent massivement sur la phonologie (Read, 1986). La conscience phonologique et les qualités des productions orthographiques sont fortement corrélées (Bruck & Treiman, 1990). La plupart des erreurs d'orthographe sont phonétiquement correctes (e.g., femme orthographié FAME, Bosman & Van Orden, 1997 ; Sprenger-Charolles, Siegel, Béchenec & Serniclaes, 2003). Dans l'étude de Sprenger-Charolles *et al.* (2003), conduite auprès d'élèves français, les mots réguliers étaient mieux orthographiés que les mots irréguliers (effet de régularité) mais il n'y a pas de différence entre les mots fréquents et les mots rares (pas d'effet de fréquence), ni entre les mots réguliers et les non-mots (pas d'effet de lexicalité), chez les élèves les plus jeunes. Enfin, certaines études ont suggéré, en accord avec les modèles en stades, que l'utilisation d'analogie pour orthographier d'autres mots serait assez tardive (e.g., Campbell, 1985). La qualité des traitements phonologiques serait importante même pour apprendre et produire l'orthographe de mots irréguliers pour lesquels, très souvent, une seule partie est irrégulière, par exemple *oi* dans *oignon*. Ainsi, un élève qui a de bonnes habiletés de traitement phonologique pourra simplement apprendre comment se transcrit le /o/ de *oignon*, plutôt que de mémoriser toutes ses lettres constitutives.

Des études suggèrent toutefois que les enfants prennent en compte d'autres sources d'information que les seules CPG beaucoup plus tôt que ne le prédisent les modèles en stades (e.g. Treiman & Cassar, 1997). Il en va ainsi des connaissances relatives à l'orthographe de mots spécifiques, des connaissances sur les régularités orthographiques et des connaissances morphologiques (e.g., Deacon, Conrad & Pacton, 2008). Dans la suite de

cette partie théorique, nous présentons les travaux relatifs aux deux premiers types de connaissances (voir Pacton & Deacon, 2008 pour les connaissances morphologiques).

## **2) L'apprentissage de l'orthographe de mots spécifiques.**

Un certain nombre d'études ont mis en évidence une utilisation des informations relatives à l'orthographe de mots spécifiques dès le début de l'acquisition de l'orthographe (e.g., Lennox & Siegel, 1994 ; Rittle-Johnson & Siegler, 1999 ; Share, 1995, 1999 ; Snowling, 1994, Treiman & Cassar, 1997 ; Varnhagen, McCallum, Burstow, 1997). Nous présentons d'abord les études ayant examiné quand les performances orthographiques variaient en fonction de la fréquence des mots puis celles ayant exploré quand les enfants utilisaient des analogies.

### 2.1) Effet de fréquence.

Le moment auquel des effets de fréquence sur les performances orthographiques ont été rapportés diverge fortement selon les études. Ainsi, dans une étude impliquant des enfants francophones (belges), ce n'est qu'à partir de la troisième primaire (CE2) qu'Alegria et Mousty (1996) ont observé un effet de fréquence alors que Sprenger-Charolles, Siegel, Bonnet (1998) ont observé un effet de fréquence après dix mois, mais non après quatre mois, de première primaire (CP). Dans l'étude de Sprenger-Charolles et al. (1998), l'effet de fréquence interagissait avec la régularité des mots : les mots irréguliers fréquents étaient mieux orthographiés que les mots irréguliers rares mais il n'y avait pas de différences entre les mots réguliers fréquents et les mots réguliers rares car les mots réguliers rares, contrairement aux mots irréguliers rares, peuvent être orthographiés correctement en utilisant une procédure alphabétique (la procédure orthographique ne peut pas être utilisée si le mot rare ne fait pas encore partie du lexique de l'enfant). Ces différences inter-études pourraient s'expliquer par l'utilisation de mots plus ou moins représentatifs de la fréquence des mots rencontrés par les enfants au début de l'école élémentaire selon Martinet, Valdois et Fayol (2004). En effet, dans une tâche de dictée, lorsque la fréquence des mots est très soigneusement contrôlée, par exemple en analysant les livres de lecture des enfants dans l'étude de Martinet, Bosse, Valdois et Tainturier, (1999), les enfants orthographiaient mieux

les mots irréguliers fréquents que les mots irréguliers moins fréquents dès trois mois (janvier) d'apprentissage.

## 2.2) Les écritures par analogie.

Différentes études ont cherché à préciser si, quand et comment des élèves utilisaient leurs connaissances de l'orthographe des mots spécifiques pour écrire de nouveaux mots par analogie. Selon Nation et Hulme (1996), l'écriture par analogie consiste en l'utilisation d'une combinaison de connaissances phonologiques (la ressemblance phonologique entre le mot familier et le mot non familier) et lexicales (le patron orthographique du mot familier). Toujours selon Nation et Hulme, les mots connus sur lesquels se fonde l'analogie seraient activés automatiquement, non consciemment, sans la mise en œuvre d'une stratégie délibérée de la part de celui qui écrit. Les premières études corroboraient le postulat des modèles en stades du développement de l'orthographe selon lesquels la capacité à écrire en utilisant des analogies apparaîtrait assez tardivement (e.g., stade orthographique dans le modèle de Frith, 1985). Ainsi Marsh, Friedman, Welch et Desberg (1980) ont rapporté que des non-mots tels que '*zoldier*' ou '*wenght*' n'étaient pas orthographiés par analogie avec les mots '*soldier*' ou '*lenght*' à sept ans et l'étaient dans seulement 33% des cas à dix ans. De façon similaire, Campbell (1985) a repris le paradigme d'amorçage qu'elle avait utilisé avec des adultes (Campbell, 1983, voir section III de cette partie théorique) et a rapporté une influence des mots préalablement entendus sur les orthographes de non-mots seulement chez des enfants dont l'âge de lecture était supérieur à 11 ans. Cependant, une limite importante de ces études est qu'elles ne contrôlaient pas si les enfants connaissaient l'orthographe des mots amorces (i.e., des mots pouvant être utilisés pour en orthographier d'autres par analogie). Dès lors, ces études ne permettent pas de déterminer si l'absence d'effet d'analogie chez les jeunes élèves reflète une incapacité à écrire par analogie (spontanément ou même avec une incitation) ou tout simplement l'absence de connaissance de l'orthographe des mots amorces.

Afin de départager ces alternatives, Goswami (1988) a étudié si des enfants de six ans écrivaient par analogie lorsque les mots amorces étaient présentés visuellement. Des mots amorces (e.g., *beak*, *peak*, *lake*) étaient prononcés et restaient visibles pendant que les enfants écrivaient un mot cible (également monosyllabique) et ils étaient incités à utiliser le

mot amorce pour écrire l'item cible. Des écritures par analogie étaient observées lorsque les mots amorces et les items cibles partageaient une rime (*peak*) ou, plus rarement, une séquence Consonne - Voyelle (*bean*), mais pas lorsqu'ils partageaient seulement des lettres communes (*lake*). Selon Goswami, ceci montre que les analogies ne portaient que sur des unités plus larges que le graphème. Cependant, cette conclusion a été critiquée car si cette expérience montre bien que les analogies n'étaient pas basées sur la simple ressemblance graphémique entre les mots amorces et cibles (la présentation de *lake* ne facilitant pas l'écriture de *beak*), elle ne montre pas que les enfants ne pouvaient pas effectuer des analogies basées sur des CGP. Pour tester cela, il aurait fallu utiliser des mots comme *leap* ou *team* dans lesquels *ea* s'écrit et se prononce comme dans *beak*.

Nation et Hulme (1996) ont conduit une telle expérience et ont montré que des enfants de six - sept ans étaient capables de fonder leurs écritures sur des analogies portant sur la correspondance phonème-graphème lorsque la similarité graphémique et la similarité phonologique étaient maintenues mais non lorsque seule la similarité graphémique était maintenue. Dans une autre expérience, des enfants de six-sept ans entendaient un mot-amorce et devaient produire une orthographe plausible s'il s'agissait d'un non-mot, comme dans les études de Campbell (1983, 1985). Les mots amorces leur avaient préalablement été dictés afin de contrôler que ces orthographe étaient connues des enfants. Des écritures par analogie étaient observées chez des élèves de 6-7 ans dans ces conditions. Les mots amorces n'étaient pas visibles mais on peut se demander si ces effets d'analogie auraient été observés si les enfants n'avaient pas entendu les mots amorces et ne les avaient pas écrits au préalable. En effet, certaines études suggèrent que la proportion de réponses analogiques est très sensible à des modifications du paradigme (e.g., Deavers & Brown, 1997).

La démonstration la plus convaincante de l'utilisation d'une procédure d'analogie lexicale a été apportée par Bosse, Valdois et Tainturier (2003). Leur expérience présentait les caractéristiques suivantes : (1) il n'y avait pas de mots amorces, ni d'incitation à utiliser une stratégie analogique, ni de procédure favorisant l'activation du lexique orthographique (moins de 20% des items étaient voisins de mots) ; (2) les mots voisins utilisés avaient des CPG irrégulières afin de pouvoir distinguer l'influence de l'orthographe lexicale de l'influence des connaissances des CPG ; (3) les mots utilisés étaient connus des enfants. Dans leur première expérience, des élèves de huit ans écrivaient les non-mots voisins de mots par analogie, mais pas les élèves plus jeunes. Bosse et al. ont montré que cette absence d'effet d'analogie chez des élèves de CP et CE1 était en fait due à une méconnaissance des

mots références. En effet, dans une autre expérience, Bosse et al. ont observé des effets d'analogie chez des élèves de CP et CE1 qui connaissaient l'orthographe des mots références parce que leurs enseignants avaient introduit ces mots dans leur enseignement (voir Martinet *et al.*, 2004, pour des résultats similaires après seulement trois mois d'apprentissage du langage écrit).

Ainsi, lorsque les mots utilisés correspondent aux mots rencontrés par les enfants en tout début d'acquisition (Bosse *et al.*, 2003 ; Martinet *et al.*, 2004), des effets de fréquence et d'analogie sont mis en évidence beaucoup plus précocement que ne le postulaient les modèles traditionnels de l'acquisition de l'orthographe (e.g., Frith, 1985) et que ne le montraient certaines études antérieures (e.g., Alegria & Mousty, 1996, pour les effets de fréquence ; Campbell, 1985 pour l'utilisation d'analogies). Souvent, les effets de fréquence et d'analogie ne sont pas observés durant les deux premières années de l'école élémentaire non pas parce qu'à ces niveaux les enfants ne peuvent pas utiliser des connaissances lexicales mais simplement parce que les mots utilisés ne font pas partie du lexique orthographique des élèves les plus jeunes.

### 2.3) L'hypothèse d'auto-apprentissage.

Selon Share et collaborateurs (Jorm & Share, 1983 ; Share, 1995, 1999), le décodage phonologique (i.e., la mise en œuvre de relations entre graphèmes et phonèmes au cours de la lecture), ce qui correspond à la procédure alphabétique dans le modèle de Frith (1985) ou à la procédure d'assemblage dans les modèles à doubles voies, joue un rôle fondamental dans l'acquisition de l'orthographe lexicale. Ce serait ce décodage phonologique, et non la simple exposition visuelle, qui permettrait la mémorisation de la forme orthographique des mots par auto-apprentissage.

Afin de valider son hypothèse, Share (1999) a réalisé une série d'expériences auprès d'élèves hébraïques scolarisés en deuxième primaire. Dans une première expérience, les élèves lisaient à voix haute des textes contenant des non-mots qui apparaissaient quatre à six fois. Les élèves ne recevaient pas de feedback, même s'ils commettaient des erreurs. Trois jours plus tard, l'apprentissage de l'orthographe était évalué avec trois épreuves : une épreuve de dénomination des non-mots et de leurs homophones, une dictée des non-mots et une épreuve de choix de l'orthographe correcte parmi quatre alternatives comprenant la

forme orthographique correcte, une forme orthographique homophone (même prononciation mais orthographe différente), une forme orthographique contenant une substitution de lettre (remplacement d'une lettre par une autre) et une forme orthographique contenant une transposition (inversion de deux lettres). Les résultats montraient que la dénomination était plus rapide pour les non-mots cibles que pour leurs homophones. Les performances en dictée et en choix forcé révélaient également un apprentissage de l'orthographe des non-mots. Par exemple, dans l'épreuve de choix forcé, le taux de sélections de l'orthographe correcte (74%) était significativement supérieur au hasard (25%, puisque quatre orthographe étaient proposées). Ces résultats étaient observés dès quatre expositions, sans différence entre quatre et six expositions, ce qui suggère qu'un nombre limité d'expositions aux mots est suffisant pour permettre l'apprentissage orthographique.

Afin de vérifier que l'apprentissage orthographique dépend bien du décodage phonologique, Share a réalisé une seconde expérience dans laquelle il a limité la possibilité de décodage phonologique. La phase d'étude ne consistait plus en la lecture de textes mais en une tâche de décision lexicale. Les items apparaissaient isolés sur un écran pendant 300 ms et les élèves devaient décider s'il s'agissait d'un mot ou non. Le décodage phonologique était perturbé par une double tâche : les élèves devaient prononcer en même temps le mot *dubba*. L'apprentissage de l'orthographe des non-mots était ensuite évalué avec les mêmes épreuves que dans la première expérience (dénomination, dictée et choix forcé). Une diminution de l'apprentissage orthographique était observée par rapport à la première expérience, ce qui suggère, selon Share, que la seule exposition visuelle aux mots ne suffirait pas à acquérir l'orthographe lexicale et que le recodage phonologique serait nécessaire.

Afin de tester si le plus faible niveau d'apprentissage dans la deuxième expérience ne résultait pas d'un temps d'exposition trop court et /ou de l'absence de contexte (puisque les mots étaient présentés isolément dans la seconde expérience alors qu'ils étaient insérés dans des textes dans la première expérience), Share (1999) a conduit une troisième expérience. Les élèves devaient lire à haute voix les items utilisés pour l'épreuve de décision lexicale de la deuxième expérience. Ces items étaient présentés pendant 300 ms, comme dans la deuxième expérience sans double tâche. Dans ces conditions, les résultats aux épreuves de dénomination, dictée et choix forcé montraient un meilleur niveau d'apprentissage orthographique que dans la seconde expérience, ce qui confirme l'importance du décodage

phonologique et indique que les plus faibles scores dans l'expérience 2 ne pouvaient pas s'expliquer par un temps d'exposition trop court ni par une absence de contexte.

Dans une quatrième expérience, Share (1999) a remplacé les non-mots par une suite de quatre symboles afin d'examiner si une représentation visuelle pouvait être mémorisée lorsque le décodage phonologique n'est pas possible. Chaque suite de symboles était présentée quatre fois durant l'apprentissage. Lors du test, les enfants devaient choisir parmi quatre suites celle vue durant l'apprentissage. Le taux de sélections correctes était seulement de 32%, le hasard étant à 25%. Selon Share, ce faible score montrerait que la part du traitement visuel dans l'acquisition des connaissances orthographiques est assez faible alors que le décodage phonologique joue lui un rôle crucial.

Depuis, l'hypothèse d'auto-apprentissage, selon laquelle le décodage phonologique est la condition sine qua non pour apprendre l'orthographe lexicale (Share, 1995, 1999), a été confirmée par d'autres expériences conduites auprès d'élèves anglophones (Bowey & Muller, 2005 ; Cunningham et al., 2002 ; Kyte & Johnson, 2006 ; Nation et al., 2007 ; Share & Shaley, 2004) ou néerlandais (De Jong & Share, 2007). Par exemple, Cunningham et al. (2002) ont demandé à des élèves anglophones scolarisés en seconde primaire de lire à haute voix des textes contenant des non-mots (six non-mots par texte). Trois jours plus tard, l'apprentissage de l'orthographe des non-mots était évalué avec une épreuve de dictée et une épreuve de choix forcé comme dans l'expérience de Share (1999). Les résultats aux deux épreuves montraient un apprentissage orthographique et le niveau d'apprentissage orthographique était positivement corrélé avec le nombre de non-mots correctement décodés durant la phase d'étude ( $r = .52$ ). Dans l'étude de Kyte et Johnson (2006), des élèves anglophones de grade 4 et 5 effectuaient une tâche de décision lexicale dans deux conditions expérimentales. Dans la condition *décodage*, les enfants lisaient à voix haute les mots et les non-mots avant de donner leur réponse sur la lexicalité de l'item. Dans la condition *articulation*, les enfants devaient répéter la syllabe /la/. Cette répétition était supposée empêcher la mise en œuvre de la procédure de décodage. Après un délai d'un jour, l'apprentissage orthographique était évalué. Les résultats montraient que les non-mots étaient mieux reconnus (en choix forcé) et mieux orthographiés (en dictée) dans la condition *décodage* que dans la condition *articulation*. De plus, De Jong, Bitter, Van Setten et Marinus (2009) ont montré que la baisse des performances dans la condition *articulation* ne pouvait pas être attribuée à la seule présence d'une double tâche en comparant trois conditions : décision lexicale sans double tâche, décision lexicale avec une double tâche

d'articulation concurrente et décision lexicale avec une double tâche de tapping. Seule la double tâche d'articulation concurrente engendrait une diminution des scores d'apprentissage orthographique évalué avec une épreuve à choix forcé.

#### 2.4) Variables influençant l'apprentissage orthographique dans des situations de *self-teaching*

L'impact d'un certain nombre de variables sur l'apprentissage orthographique a été étudié : le fait que la lecture soit silencieuse ou orale, le nombre d'expositions et le délai de rétention, ou le rôle du contexte. Nous présentons les principaux résultats de ces études.

Concernant la lecture silencieuse ou orale, un apprentissage orthographique a été rapporté chez des enfants d'âges différents avec les deux types de lecture, sans différence d'efficacité en fonction du type de lecture (Bowey & Muller, 2005 ; Bowey & Miller 2007 ; De Jong & Share, 2007 ; De Jong et al., 2009).

Concernant le nombre d'expositions, si un apprentissage a été mis en évidence après seulement une exposition dans des langues transparentes (e.g., Share, 2004 en hébreu) et opaques (e.g., Nation et al., 2007 en Anglais), l'impact du nombre d'expositions semble étroitement lié à la transparence des langues. Ainsi, dans l'étude de Share (2004), le niveau d'apprentissage orthographique était similaire après une, deux ou quatre expositions. En revanche, dans les études conduites en Anglais, Nation et al. ont rapporté un niveau d'apprentissage orthographique plus élevé après quatre expositions qu'après une exposition et Bowey et Muller ont rapporté un niveau d'apprentissage orthographique plus élevé après huit expositions qu'après quatre expositions. L'impact du délai semble également varier en fonction de la transparence des langues. Ainsi, en hébreu, Share (2004, Expérience 1) a rapporté un apprentissage orthographique stable après des délais de 1, 7 et 30 jours alors que le niveau d'apprentissage orthographique dans les études impliquant des élèves anglophones était plus élevé après un délai d'un jour qu'après des délais de sept jours dans l'étude de Nation et al. (2007) et de six jours dans l'étude de Bowey et Muller (2005).

L'objectif des études ayant exploré le rôle du contexte était de déterminer si les nouveaux mots sont mieux mémorisés lorsqu'ils sont présentés en contexte que lorsqu'ils sont présentés isolément. Share (1995, 1999) a suggéré que le contexte jouait un rôle, en



particulier lorsque le décodage n'est pas très efficient. Peu d'études permettent de tester cette prédiction car la grande majorité d'entre elles utilisent le paradigme d'auto-apprentissage seulement avec des non-mots insérés dans des textes. Nation et al. (2007) ont comparé l'apprentissage de l'orthographe de non-mots selon qu'ils étaient insérés dans des textes ou présentés de façon isolée et n'ont pas trouvé d'effet du contexte. Cunningham (2006) a conduit une étude d'auto-apprentissage avec de vrais mots dont des enfants de première primaire connaissaient la forme orale (i.e., ces mots faisaient partie de leur vocabulaire) mais pas leur orthographe. Les mots étaient présentés soit dans des histoires (qui avaient du sens) soit mélangés dans des passages (qui n'avaient pas de sens). Les mots étaient mieux décodés dans la première condition que dans la seconde mais il n'y avait pas de différence entre les deux conditions lorsque l'acquisition des représentations orthographiques était évaluée avec une dictée ou un choix forcé entre quatre orthographes. Plus récemment, Wang, Castles, Nickels et Nation (2011) ont étudié si l'effet du contexte variait en fonction de la régularité des non-mots auprès d'élèves anglophones de deuxième primaire. Cette expérience incluait une première phase qui consistait à faire apprendre aux enfants des formes orales des non-mots avant qu'ils ne lisent ces non-mots insérés dans des histoires ou présentés isolément. Chaque orthographe correspondait à une orthographe régulière dans la première expérience et à une orthographe irrégulière dans la seconde. Wang et al. n'ont pas observé de différence en fonction du contexte pour les non-mots réguliers (comme Nation et al., 2007) mais ont trouvé un effet facilitateur du contexte pour les non-mots irréguliers dans une épreuve de décision orthographique (indiquer si une orthographe est correcte ou non). Il est à noter toutefois que cet effet du contexte pour les non-mots irréguliers n'était pas observé lorsque l'apprentissage orthographique était évalué avec des épreuves de dictée et de choix forcé, les deux épreuves les plus utilisées dans ce type d'études. Enfin, Landi, Perfetti, Bolger, Dunlap et Foorman (2006) ont trouvé un effet négatif du contexte sur l'apprentissage orthographique. Les enfants de 5–8 ans de leur étude essayaient de lire des mots qu'ils ne pouvaient pas lire lors d'un pré-test dans une situation de *self-teaching*. Ces mots étaient insérés dans des histoires ou présentés isolément. Les mots étaient mieux lus dans la condition contexte que dans la condition isolée mais l'apprentissage orthographique était meilleur dans la condition isolée que dans la condition contexte. Landi et al. ont expliqué ces résultats en termes de traitement attentionnel : les enfants feraient moins attention aux caractéristiques des mots lorsqu'ils les lisent en contexte, ce qui engendre une moins bonne mémorisation de leur orthographe.

En résumé, l'hypothèse d'auto-apprentissage selon laquelle les connaissances orthographiques lexicales sont acquises lors de la lecture, a été validée par de nombreuses études dans différentes langues (Hébreu, Anglais, Néerlandais), en lecture silencieuse et orale, avec des non-mots et des mots, avec un nombre d'expositions plus ou moins important, avec des délais plus ou moins longs, avec des items présentés en contexte ou isolément. Ces études confirment chez des enfants de l'école élémentaire les résultats d'études, parfois anciennes, montrant un apprentissage implicite de l'orthographe de mots spécifiques chez des adultes. Par exemple, les études de Gilbert (1934a, 1934b, 1935) montraient que des lycéens et des étudiants orthographiaient mieux des mots après les avoir rencontrés dans des textes qu'ils devaient lire afin de pouvoir répondre à des questions sur leur contenu, donc sans instruction d'apprendre l'orthographe des mots. L'étude d'Ormrod (1986a, 1986b) montrait que des étudiants de l'université apprenaient l'orthographe de nouveaux mots lorsqu'ils lisaient des textes même lorsque rien ne les y incitait mais qu'ils les apprenaient mieux lorsqu'on leur demandait explicitement de les apprendre. Ces résultats peuvent aussi être rapprochés de ceux montrant l'effet positif de l'exposition à des orthographes correctes et l'effet négatif de l'exposition à des orthographes incorrectes (Brown, 1988 ; Dixon & Kaminska, 1997).

### 2.5) Le rôle des connaissances orthographiques dans l'apprentissage orthographique dans des situations de *self-teaching*.

L'existence de corrélations entre la lecture-décodage et l'apprentissage de l'orthographe des non-mots est évidemment en accord avec l'idée que le décodage phonologique joue un rôle crucial dans cet apprentissage. Toutefois, ces corrélations sont souvent modérées, notamment dans des langues avec des correspondances graphèmes-phonèmes peu consistantes comme l'Anglais, ce qui suggère que les habiletés de décodage phonologique ne semblent pas pouvoir expliquer à elles seules l'apprentissage de l'orthographe lexicale (Castles & Nation, 2006 ; Cunningham et al., 2002, 2006 ; Nation et al., 2007). La part de variance inexpliquée par le décodage phonologique a souvent été attribuée à des habiletés 'générales' de traitement orthographique (Berninger, 1995), souvent assez mal définies (e.g., Castles & Nation, 2006). Castles et Nation (2006) ont souligné

l'importance de distinguer la connaissance de l'orthographe de mots spécifiques (orthographe lexicale) de la connaissance de régularités sur le langage écrit (régularités graphotactiques) (voir aussi Conners, Loveall, Moore, Hume & Maddox, 2011; Nation et al., 2007).

Les chercheurs ayant exploré le lien entre le décodage phonologique et l'apprentissage orthographique ont le plus souvent calculé des corrélations entre le score moyen en décodage phonologique et le score moyen d'apprentissage orthographique. Nation et al. (2007) ont étudié si la relation entre décodage phonologique et mémorisation de l'orthographe se maintenait dans des analyses item par item afin de déterminer si tous les items lus correctement étaient des items dont l'orthographe était correctement mémorisée et si tous les items lus incorrectement étaient des items dont l'orthographe était incorrectement mémorisée. Nation et al. ont trouvé que la relation générale entre le décodage phonologique et l'apprentissage orthographique (analyse sur les valeurs moyennes) ne se maintenait pas lorsque des analyses item par item étaient conduites. En effet, des items bien lus n'étaient pas correctement reconnus dans l'épreuve de choix forcé évaluant l'apprentissage orthographique mais aussi, une proportion non négligeable d'items mal lus étaient correctement reconnus. Nation et al. ont suggéré que les connaissances graphotactiques pourraient jouer un rôle dans l'apprentissage de l'orthographe lexicale mais leur étude ne fournit pas d'éléments empiriques étayant cette interprétation.

Cunningham (2006) a étudié le rôle des connaissances orthographiques dans une situation d'auto-apprentissage de mots dont les enfants connaissaient la forme orale mais pas l'orthographe (voir section précédente). Elle a rapporté, qu'après avoir contrôlé la qualité du recodage phonologique de non-mots, un score composite de connaissances orthographiques reflétant des connaissances en orthographe lexicale (choisir laquelle de deux orthographe comme rain - rane est correcte) et des connaissances graphotactiques (choisir entre une orthographe de non-mot respectant les régularités graphotactiques de l'Anglais et une orthographe de non-mot violant une régularité graphotactique, celle qui ressemble le plus à un mot, e.g., beff - ffeb) expliquait une part de variance significative de l'apprentissage de l'orthographe des non-mots. Malheureusement, l'utilisation d'un score ne distinguant pas les deux types de connaissances orthographiques ne permet pas de déterminer le poids respectif de ces deux types de connaissances orthographiques lors de l'apprentissage de l'orthographe de nouveaux mots. Il est à noter que, si le rôle des seules connaissances graphotactiques sur l'apprentissage de l'orthographe de nouveaux mots n'a jamais été étudié, le rôle des seules

connaissances orthographiques lexicales l'a été (Cunningham, Perry, Stanovich & Share, 2002). Dans cette étude, Cunningham et al. avaient montré qu'après contrôle du nombre de non-mots correctement décodés, les connaissances orthographiques lexicales préalables, évaluées avec une épreuve de choix entre deux orthographes (e.g., take – taik), ajoutaient 11% en première primaire et 20% en deuxième primaire de la variance en auto-apprentissage, alors que ni les scores de dénomination rapide (RAN) ni les d'habiletés cognitives générales n'ajoutaient une part de variance significative.

Les deux types de connaissances orthographiques (orthographe lexicale et connaissances graphotactiques) ont en revanche été distingués dans une étude récente portant sur la lecture de « vrais » mots chez des élèves de 7-9 ans (Connors, Loveall, Moore, Hume & Maddox, 2011). Dans cette étude, après contrôle des habiletés de recodage phonologique, les connaissances en orthographe lexicale et les connaissances graphotactiques expliquaient des parts de variance indépendantes. Dans l'étude rapportée dans le deuxième chapitre expérimental, nous examinerons le rôle de ces deux types de connaissances orthographiques dans l'apprentissage de l'orthographe lexicale dans une situation d'auto-apprentissage.

### **3) L'apprentissage de régularités graphotactiques**

#### **3.1) Connaissances graphotactiques examinées seules.**

Les connaissances graphotactiques ont été explorées de trois façons. La première consiste à étudier par exemple dans les cahiers d'enfants si les productions orthographiques (de vrais mots) reflètent les régularités graphotactiques de la langue. Dans cette perspective, qualifiée de naturalistique, les informations les plus intéressantes proviennent de l'analyse des erreurs. En effet, si un enfant écrit correctement le mot *tableau*, il n'est pas possible de déterminer si cela reflète la connaissance de l'orthographe de ce mot spécifique ou une sensibilité au fait que /o/ est fréquemment transcrit *eau* en fin de mot en Français. En revanche, si lorsque les enfants transcrivent /o/ erronément *eau* au lieu de *o* ou *au*, ils le font plus souvent en fin de mots, où *eau* est fréquent en Français, qu'en position initiale ou médiane, où *eau* est rare (position médiane) ou illégal (position initiale), cela suggère une influence des régularités graphotactiques. Le deuxième type d'études consiste à dicter des non-mots afin d'étudier non seulement si les enfants sont capables d'appliquer correctement des règles de CPG mais aussi s'ils varient leurs transcriptions en fonction de certaines régularités graphotactiques, par exemple la probabilité que /o/ soit transcrit *eau* en fonction de la position dans le mot. Le troisième type d'études implique des tâches de jugement de paires de non-mots construites de telle sorte qu'une seule des deux orthographe soit légale du point de vue des régularités graphotactiques (e.g., une orthographe avec un doublet en position initiale, l'autre avec un doublet en position médiane, comme dans la paire *tunnor* - *ttunor*) ou de sorte que les deux orthographe soient légales mais que l'une soit plus probable que l'autre (e.g., une orthographe avec un doublet fréquent en position médiane, l'autre avec un doublet légal mais rare en position médiane comme dans la paire *tunnor* - *tuddor*). L'avantage de cette dernière tâche est qu'elle permet de proposer un choix entre deux orthographe que les sujets ne produiraient pas systématiquement.

### 3.1.1) *Les études naturalistiques.*

Treiman (1993) a examiné les orthographes d'enfants américains scolarisés en première primaire. Ces enfants étaient inscrits dans des classes dans lesquelles ils ne recevaient pas d'enseignement explicite sur l'orthographe. En particulier, leurs productions orthographiques erronées n'étaient pas corrigées. Treiman a montré que les orthographes produites par ces enfants étaient le plus souvent compatibles avec les régularités du système écrit anglais. En particulier, ils doublaient plus souvent des lettres qui sont fréquemment doublées en Anglais (e.g., *e, l*) que des lettres qui le sont rarement ou jamais (e.g., *u, h*) et ils produisaient ces doublets presque toujours en position légale, c'est-à-dire en milieu ou fin de mots, mais pas en début de mots. De façon similaire, il leur arrivait de transcrire le son /k/ *ck*, pour des mots dont c'est l'orthographe correcte (e.g., *black*) mais aussi pour des mots dans lesquels /k/ n'est pas transcrit *ck*. Toutefois, lorsque *ck* était utilisé de façon erronée à la place de *c* ou de *ck*, ce n'était jamais en début de mots ou *ck* est illégal en Anglais.

### 3.1.2) *Les études impliquant des dictées de non-mots.*

En Anglais, Treiman, Berch et Weatherston (1993) ont investigué si les écritures de non-mots produites par des enfants de maternelle et de grade 1 reflétaient une sensibilité précoce à la position légale des doubles consonnes. Seize non-mots de type CVC (e.g., /des/) étaient dictés. Aux deux niveaux scolaires, les élèves utilisaient rarement des doublets mais, alors que les élèves de maternelle les utilisaient aussi bien en position légale qu'en position illégale, ceux de grade 1 utilisaient des doublets seulement en position légale : sur les 384 productions de chaque niveau scolaire (24 élèves x 16 non-mots), il y avait 11 doublements en position initiale et 13 doublements en position finale en maternelle alors qu'il y avait 16 doublements en position finale mais aucun en position initiale en grade 1. Toujours en Anglais, Nation (1997) a trouvé une influence de la fréquence des correspondances son-orthographe au niveau de la rime sur l'écriture de non-mots monosyllabiques. Il convient de noter que les non-mots utilisés dans cette étude différaient très peu de vrais mots (e.g., /zisk/ - 'disk'). De ce fait, il est difficile de déterminer si les effets observés reflétaient la

sensibilité à la fréquence des correspondances son-orthographe au niveau de la rime ou l'écriture des non-mots par analogie à de vrais mots (Bosse et al., 2003).

Des non-mots qui différaient davantage de vrais mots ont été utilisés dans une étude conduite en Français par Pacton, Fayol et Perruchet (2002). Ces auteurs ont exploré si des élèves de l'école élémentaire, scolarisés en CE1, CE2 et CM1, étaient sensibles au fait que les transcriptions du phonème /o/ varient en fonction de la position et du contexte consonantique. Ils leur ont dicté des non-mots tri-syllabiques tels que /obidar/, /ribore/ ou /bylevo/. Ces non-mots étaient construits afin de pouvoir examiner si les élèves transcrivaient /o/ *eau* plus souvent en position finale, où il est fréquent en Français, qu'en positions médiane et initiale, où il est rare ou inexistant et s'ils utilisaient *eau* plus souvent après des consonnes après lesquelles *eau* est fréquent (e.g., *veau*) qu'après des consonnes après lesquelles *eau* est rare (e.g., *feau*). Les résultats montraient que même les enfants les plus jeunes utilisaient plusieurs graphèmes pour transcrire /o/ et qu'à tous les niveaux scolaires *eau* était davantage utilisé en position finale qu'en positions initiale et médiane et, en position finale, *eau* était plus souvent utilisé après une consonne après laquelle *eau* est fréquent qu'après une consonne après laquelle *eau* est rare.

Des résultats convergents ont été ultérieurement rapportés en Anglais (e.g., Hayes, Treiman, & Kessler, 2006 ; Treiman & Kessler, 2006). Par exemple, Treiman et Kessler (2006) ont étudié comment des élèves âgés de 7 à 14 ans transcrivaient certaines voyelles en fonction des consonnes gauches ou droites. Ils ont trouvé que la transcription des voyelles variait en fonction de la consonne la précédant immédiatement (e.g., /*ɑ*/ plus souvent transcrit *a* après *w* mais plus souvent transcrit *o* après *p*) à partir du grade 4. Ils ont également trouvé que la transcription des voyelles variait fonction de la consonne la suivant immédiatement (e.g., /*I*/ transcrit plus souvent *ee* avant un *p* mais plus souvent transcrit *ea* avant un *m*) mais encore plus tardivement, à partir du grade 7.

### 3.1.3) Les études impliquant des jugements de plausibilité de pseudo-mots.

Afin d'étudier la connaissance de l'identité des consonnes pouvant être doublées, Cassar et Treiman (1997) ont présenté à des élèves de maternelle et de l'école élémentaire

des paires de non-mots comme *affe - ahhe* ou *yill - yihh* et leur ont demandé de désigner l'orthographe inventée de chaque paire qui ressemblait le plus à un mot. Un non-mot de chaque paire incluait une consonne qui est fréquemment doublée en Anglais (e.g., *f, l*) et l'autre incluait une consonne qui n'est jamais doublée (e.g., *h*). Dès le grade 1, mais pas en maternelle, les élèves sélectionnaient plus souvent les non-mots incluant des consonnes fréquemment doublées (e.g., *yill*) que ceux incluant des consonnes jamais doublées en Anglais (e.g., *yihh*). Cassar et Treiman en ont conclu que les enfants connaissaient l'identité des lettres pouvant être doublées. Selon Pacton et al. (2001), cette préférence pourrait ne refléter qu'une sensibilité à la fréquence des lettres simples car une consonne comme *l* est à la fois plus fréquente et plus souvent doublée qu'une consonne comme *h*. Autrement dit, les enfants pourraient juger que *yill* ressemble davantage à un mot que *yihh* simplement parce que *l* est une consonne plus fréquente que *h*, indépendamment du fait que *l* est plus souvent doublé que *h*.

Afin de tester si des élèves de l'école élémentaire sont vraiment sensibles à la fréquence de doublement des consonnes, Pacton et al. (2001) et Danjon et Pacton (2009) ont exploité le fait qu'en Français certaines consonnes sont fréquentes en formats simple et double (e.g., *m*) alors que d'autres ne sont fréquentes qu'en format simple (e.g., *d*). Pacton et al. ont proposé des paires comme *imose - idose* et *ummise - uddise* à des élèves français de chacun des cinq niveaux de l'école élémentaire. Dans chaque paire, un non-mot incluait une consonne cible qui est fréquemment doublée en Français (e.g., *m* comme dans *imose* et *ummise*) alors que l'autre non-mot incluait une consonne cible qui est rarement doublée en Français (e.g., *d* comme dans *idose* et *uddise*). Ces deux types de consonnes étaient en format simple dans la moitié des paires (e.g., *imose - idose*) et en format double dans l'autre moitié (e.g., *ummise - uddise*). L'hypothèse de Pacton et al. était que si les élèves sont sensibles à la fréquence de doublement des consonnes, les non-mots incluant des consonnes fréquentes en format simple uniquement devraient être plus souvent sélectionnés lorsque ces consonnes apparaissent en format simple (e.g., *idose* pour la paire *imose - idose*) que lorsqu'elles apparaissent en format double (e.g., *uddise* pour la paire *ummise - uddise*). Les résultats ont confirmé cette hypothèse puisque dès le CP. De plus, l'amplitude de cet effet augmentait chez des élèves plus âgés. Danjon et Pacton (2009) ont répliqué ce résultat en utilisant un matériel qui contrastait également des consonnes fréquentes en formats simple et double et des consonnes fréquentes seulement en format simple mais avec les deux non-mots de chaque paire qui se prononçaient pareillement. Ainsi, dès le CE1 (l'étude ne



comportait pas de CP), les élèves sélectionnaient plus souvent des non-mots comme *gattaver* pour la paire *gataver – gattaver* (avec le *t* qui est fréquemment doublé en Français) que des non-mots comme *abbader* pour la paire *abader – abbader* (avec le *d* qui est rarement doublé en Français).

La connaissance d'autres régularités, également de type probabiliste, a été étudiée avec des tâches de jugement de non-mots. Par exemple, Danjon et Pacton (2009) ont étudié si les enfants étaient sensibles au fait que certaines consonnes qui sont fréquemment doublées en Français, comme le *n*, le sont beaucoup plus après certaines voyelles qu'après d'autres voyelles. Par exemple, la consonne *n* est beaucoup plus fréquente après le graphème *o* qu'après le graphème *oi*. Danjon et Pacton (2009) ont rapporté une sensibilité à cette propriété à partir du CE2 mais pas en CE1. Ainsi, les élèves de CE2, CM1 et CM2 de cette étude choisissaient plus souvent un non-mot comme *ligédonne* pour la paire *ligédone – ligédonne* qu'un non-mot comme *ligédoinne* pour la paire *ligédoine – ligédoinne*.

La fréquence de doublement des lettres et la probabilité de doublement en fonction de l'environnement gauche sont des régularités de type probabiliste. Aucune règle générale ne permet de les décrire. En revanche, d'autres propriétés de l'emploi des doubles lettres peuvent être décrites sous la forme de règles. Il en va ainsi du fait que les doublets ne surviennent jamais en début de mots en Français et en Anglais ou du fait qu'une consonne double peut survenir avant mais non après une consonne simple.

Dans l'étude de Cassar et Treiman (1997), mentionnée précédemment concernant la connaissance du fait que seules certaines lettres peuvent être doublées, des paires de non-mots comme *baff – bbaf* étaient utilisées pour tester si les élèves savaient que les consonnes ne peuvent pas être doublées en début de mots en Anglais. Dès la maternelle, et encore plus en grade 1, les élèves américains préféraient des non-mots incluant une double consonne en position légale (médiane ou finale, e.g. *baff*) à des non-mots incluant une double consonne en position illégale (i.e., initiale, e.g. *bbaf*). Selon Pacton et al. (2001), cette préférence pourrait toutefois simplement refléter l'apprentissage des propriétés idiosyncrasiques de chaque consonne, par exemple, le fait que *bb* ne survient jamais au début des mots anglais. Leur objectif était de déterminer si les enfants acquièrent implicitement une règle générale relative à la position légale des doubles lettres en Français en utilisant deux types de paires de non-mots. Des paires *sans transfert* incluaient des doublets formés avec des consonnes fréquemment doublées en Français, en position légale dans un non-mot et en position

illégal dans l'autre (e.g., *fommir* - *ffomir*). Des paires avec transfert incluait des doublets formés de consonnes jamais doublées en Français, en position légale dans un non-mot et en position illégale dans l'autre (e.g., *xihhel* - *xxihel*). Leur hypothèse était que si les sujets recouraient à une règle spécifiant que les doublets ne peuvent pas survenir en début de mots, alors leurs performances devraient être aussi bonnes pour les deux types de paires de non-mots.

Pacton et al. (2001) ont rapporté une sensibilité au fait que les consonnes doubles ne peuvent pas survenir en début de mots chez les élèves des cinq niveaux de l'école élémentaire. Dès le CP, cette sensibilité s'étendait à des consonnes jamais doublées en Français (e.g., les enfants estimaient que *xihhel* ressemblait davantage à un mot que *xxihel*). Les performances étaient toutefois inférieures à celles observées avec des non-mots incluant des doublets formés de consonnes fréquemment doublées (e.g., *tunnor* – *ttunor*) du CP au CM2 et l'amplitude de la différence selon le type de matériel utilisé pour évaluer les connaissances des élèves ne variait ni en fonction du niveau scolaire ni en fonction du niveau orthographique (évalué avec un test d'orthographe lexicale standard).

Danjon et Pacton (2009) ont répliqué ce résultat et ont rapporté des résultats similaires avec une autre propriété de l'emploi des doubles lettres. Ils ont étudié si des élèves français des cinq niveaux de l'école élémentaire possédaient une connaissance relative au fait qu'une double consonne peut survenir avant mais non après une consonne simple. Ils ont proposé des paires de non-mots incluant un groupe consonantique comme *apprulir* – *aprrulir*. Dans chaque paire, un non-mot incluait une consonne double avant une consonne simple (ce qui est légal) alors que l'autre incluait une consonne simple avant une consonne double (ce qui est illégal). De plus, ils ont manipulé la voyelle précédant immédiatement le groupe consonantique. Dans certaines paires, il s'agissait soit d'une voyelle après laquelle ce groupe consonantique est fréquent en Français (e.g., *ppr* fréquent après un *a* dans la paire *apprulir* – *aprrulir*) ; dans d'autres, d'une voyelle après laquelle ce groupe consonantique est rare (e.g., *ppr* rare après un *i* dans la paire *ippraler* – *iprraler*). L'hypothèse était que, si les élèves recouraient à une règle spécifiant que les consonnes doubles peuvent survenir avant mais non après les consonnes simples, ils devraient davantage sélectionner les non-mots incluant un doublet avant une consonne simple et cette préférence ne devrait pas varier en fonction du type de voyelle précédant le groupe consonantique. En CP, les performances étaient au niveau du hasard pour les deux types de paires de non-mots. A partir du CE1, les non-mots incluant un doublet avant une consonne simple (e.g., *apprulir*) étaient plus souvent

sélectionnés que les non-mots incluant un doublet après une consonne simple (e.g., *aprrulir*) pour les paires incluant un matériel familier (i.e., avec un groupe consonantique placé après une voyelle après laquelle il est fréquent en Français). Toutefois, en CE1 et CE2, les performances ne différaient pas du hasard pour les paires dans lesquelles le groupe consonantique était placé après une voyelle après laquelle il est rare en Français. Les performances n'étaient supérieures au hasard qu'en CM1 et CM2 pour ces items. Toutefois, du CE1 au CM2, les performances étaient meilleures lorsque les connaissances graphotactiques étaient évaluées avec un matériel fréquent. Par exemple, *aprrulir* était davantage sélectionné dans la paire *aprrulir* – *aprrulir* que ne l'était *ippraler* dans la paire *ippraler* – *iprraler*. De plus, l'amplitude de cet effet ne variait ni en fonction du niveau scolaire, ni en fonction du niveau orthographique.

Deux résultats importants ressortaient de l'analyse des performances pour les deux types de propriétés descriptibles sous forme de règle. Premièrement, certaines régularités semblent plus difficilement repérées que d'autres. Deuxièmement, même après au moins cinq ans de pratique de l'écrit, les connaissances des élèves ne semblaient pas correspondre à des règles générales, abstraites, spécifiant par exemple que les consonnes ne sont jamais doublées en début de mots ou après une consonne simple. En effet, si tel était le cas, les performances n'auraient pas dû varier en fonction de la familiarité du matériel utilisé pour tester les connaissances des élèves. Ces résultats ont des implications importantes concernant la nature des connaissances acquises implicitement. La question de la nature des connaissances acquises implicitement a typiquement été abordée dans des études de laboratoire impliquant l'apprentissage de matériels artificiels, comme des séquences de consonnes générées par des grammaires artificielles (Pacton & Perruchet, 2006). Nous présentons brièvement ces études en montrant l'intérêt d'une mise en relation de ces deux domaines que sont l'apprentissage implicite d'une part et l'acquisition du langage écrit d'autre part.

### 3.2) De l'apprentissage implicite en laboratoire à l'apprentissage implicite de régularités graphotactiques.

Selon Pacton et Perruchet (2006), étudier l'apprentissage implicite d'aspects du langage écrit présente au moins trois intérêts. Premièrement, cela permet d'étudier la validité

écologique des études conduites en laboratoire en étudiant si les conclusions issues des études menées en laboratoire se généralisent en dehors du laboratoire. Deuxièmement, les situations d'apprentissage en dehors du laboratoire présentent des caractéristiques qui diffèrent des situations expérimentales de laboratoire, notamment la très longue période sur laquelle se déroulent ces apprentissages, qui permettent d'aborder des questions théoriques comme celle de savoir si et comment évolue la nature des connaissances acquises implicitement après une pratique très prolongée. Troisièmement, des méthodes d'investigation de l'apprentissage implicite, habituellement utilisées uniquement en laboratoire, peuvent être transposées pour tenter de mieux comprendre la nature des connaissances orthographiques acquises implicitement.

Les études sur l'apprentissage implicite conduites en laboratoire miniaturisent les conditions d'apprentissage en milieu naturel, dans la dimension temporelle et dans la complexité du matériel. Ainsi, dans une situation d'apprentissage de grammaires artificielles (Reber, 1967), les sujets sont confrontés à des séquences de lettres dont l'ordre est régi par une grammaire à états finis telle que celle représentée à la Figure 2. Les lettres peuvent être remplacées par des syllabes, des carrés de différentes couleurs, des notes de musique afin d'engendrer des séquences qui respectent la grammaire. Typiquement, l'expérimentateur demande aux sujets d'essayer de mémoriser par cœur ces séquences, souvent pendant une dizaine de minutes. A l'issue de cette phase d'étude, les sujets sont informés de l'existence d'une grammaire. De nouvelles séquences de lettres leur sont alors présentées. Certaines sont grammaticales en ce sens qu'elles peuvent être générées par la grammaire. D'autres sont agrammaticales en ce sens qu'elles violent les règles de la grammaire (voir Figure 2 pour des exemples de séquences grammaticales et agrammaticales). La tâche des sujets consiste à décider, pour chacune de ces nouvelles séquences de lettres, s'il s'agit d'une séquence grammaticale ou non.

Les résultats montrent que 75-80% des nouvelles séquences présentées lors de la phase test sont catégorisées correctement, ce qui est significativement supérieur au hasard (50%), alors que les sujets ont l'impression de répondre au hasard (Reber, 1967). Selon Reber (1967, 1993), ces performances au-dessus du niveau du hasard ne seraient pas accompagnées de la connaissance explicite des règles. Les sujets auraient appris implicitement (inconsciemment) les règles de la grammaire ayant servi à générer les items

durant la phase d'étude et utiliseraient cette connaissance inconsciente pour juger de la grammaticalité des items présentés en phase test.

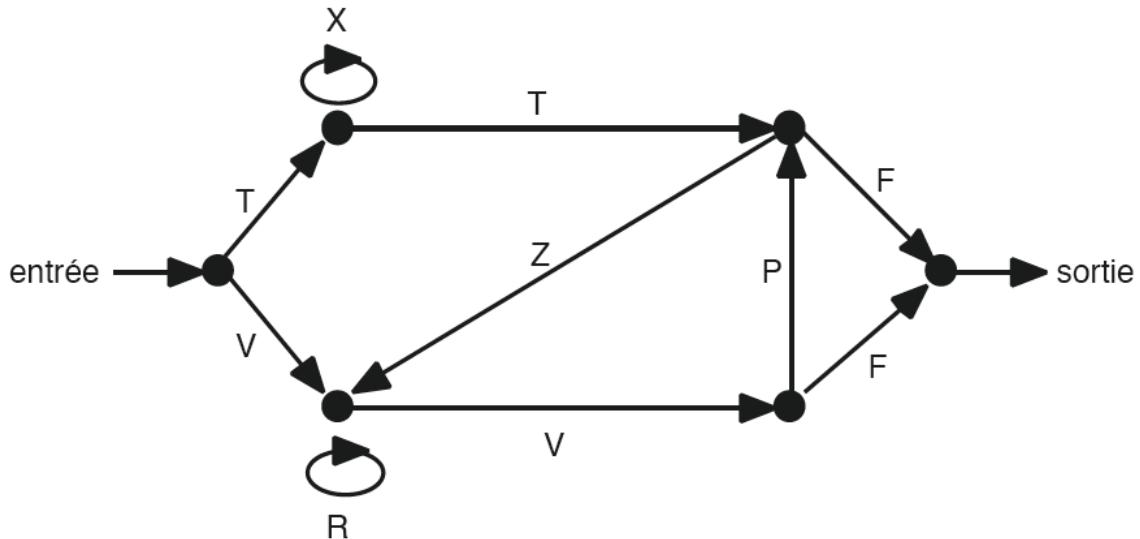


Figure 2. Exemple de grammaire artificielle à état fini. Les chaînes de consonnes grammaticales sont engendrées en suivant les flèches en passant d'un nœud de la grammaire à un autre, de l'entrée à la sortie (e.g., TXXTF, TXTZR VF). Les séquences agrammaticales ne respectent pas les transitions de la grammaire (e.g., TXXZVF ; VRPVF). Dans la situation de *transfert*, un nouvel item grammatical comme TXXTF sera remplacé en établissant des correspondances systématiques entre les différents types d'éléments, par exemple : KNKM (en remplaçant par d'autres consonnes), rouge-bleu-rouge-vert (en remplaçant par des carrés de couleurs), ou do-mi-do-fa (en remplaçant par des notes de musique).

Les résultats de Reber ont été répliqués mais des interprétations alternatives, radicalement différentes, ont été proposées. Ainsi, les sujets pourraient mémoriser les séquences entières ou les fragments de celles-ci (paires ou triplets de lettres) lors de la phase d'étude puis juger les nouvelles séquences présentées lors de la phase de test comme grammaticales si elles sont similaires aux séquences d'apprentissage (Brooks, 1978) ou si elles contiennent les mêmes fragments que les séquences d'apprentissage (Perruchet &

Pacteau, 1990). Cette dernière interprétation, selon laquelle les sujets se fonderaient en grande partie sur leur sensibilité à la fréquence de certains fragments, est étayée par des expériences montrant des niveaux de catégorisation supérieurs au hasard même chez des sujets ayant vu des fragments de lettres (e.g., des triplets de lettres) plutôt que des séquences entières lors de la phase d'étude (Perruchet & Pacteau, 1990). En effet, cette situation empêche à la fois l'abstraction de règle de haut niveau et la mémorisation de séquences entières.

Afin d'essayer de privilégier l'une de ces interprétations, certaines expériences d'apprentissage de grammaires artificielles ont utilisé une procédure dite de *transfert*. Cette procédure consiste à modifier la seule forme de surface des séquences entre les phases d'étude et de test de sorte que les séquences d'étude et de test possèdent la même structure sous-jacente mais des structures de surface différentes. Ainsi, si les séquences d'étude sont générées avec un ensemble de consonnes, celles de test le sont avec de nouvelles consonnes, avec des carrés colorés, ou des notes de musique, en établissant des correspondances systématiques entre les différents types d'éléments. Dans ces conditions, les sujets demeurent capables de discriminer les séquences tests qui respectent les règles de celles qui les violent. Leurs performances demeurent supérieures au hasard et supérieures à celles observées chez des sujets contrôles qui n'ont pas effectué la phase d'étude (Redington & Chater, 1996). Toutefois, les performances demeurent inférieures dans la *condition transfert* à ce qui est observé dans la condition *sans transfert*. Ce phénomène est appelé le *transfer-decrement*.

Le transfert a été initialement avancé comme un argument en faveur d'une interprétation abstractionniste (Manza & Reber, 1997). Cependant, il est assez bien établi que ce transfert reflète la sensibilité des sujets à des propriétés abstraites partagées par les séquences d'étude et de test comme les répétitions et les alternances (e.g., Gomez, Gerken, & Schvaneveldt, 2000 ; Shanks, Johnstone, & Staggs, 1997). Par exemple, Gomez et al. (2000) ont trouvé une absence de transfert dans une expérience avec une grammaire artificielle qui ne permettait pas d'engendrer des répétitions. Dans une autre expérience, Gomez et al. (2000) ont modifié la grammaire de sorte que certains items comportaient des répétitions et ont trouvé un apprentissage de dépendances séquentielles mais pas de transfert pour les items sans répétition mais un apprentissage de dépendances séquentielles et un transfert (ainsi qu'un *transfer-decrement*) pour les items incluant des répétitions.

Le *transfer-decrement*, c'est-à-dire la diminution des performances associée au changement des formes de surface entre les phases d'étude et de test, est problématique pour les conceptions supposant que les participants apprennent des règles abstraites, indépendantes des caractéristiques de surface (e.g., Manza & Reber, 1997). Il ne l'est pas en revanche pour les interprétations fondées sur l'exploitation des régularités statistiques pour lesquelles l'amplitude du transfert dépend du degré de similarité entre les situations familières et nouvelles. Selon Manza et Reber (1997), le *transfer-decrement* typiquement observé dans les études d'apprentissage de grammaires artificielles en laboratoire serait dû à la pratique limitée (une dizaine de minutes) qui les caractérise. Il correspondrait à un niveau de pratique où le processus d'abstraction ne serait pas complet. Si Manza et Reber ont suggéré qu'une pratique plus prolongée permettrait l'acquisition d'une connaissance abstraite des règles de la grammaire, complètement indépendante des caractéristiques de surface, ils n'ont pas testé cette prédiction.

Comme nous l'avons vu dans la section précédente, Pacton et al. (2001) et Danjon et Pacton (2009) ont adapté le paradigme de transfert pour étudier si après une pratique beaucoup plus prolongée qu'en laboratoire (des années d'exposition à l'écrit) les enfants avaient appris implicitement des règles générales, abstraites, spécifiant par exemple que les consonnes ne sont jamais doublées en début de mots. Il n'y avait donc pas à proprement parlé de phase d'étude dans leur expérience. L'exposition aux mots, en particulier à travers les lectures, constituait la phase d'étude. De la même façon que la phase test d'une situation de grammaire artificielle consiste en la présentation d'items tous nouveaux qui respectent ou non les règles de la grammaire, dans l'étude de Pacton et al. (2001), les items étaient tous nouveaux puisqu'il s'agissait de non-mots et certains respectaient la propriété graphotactique testée alors que d'autres la violaient. Dans la condition *sans transfert*, les non-mots incluaient des doublets fréquents en Français (*tunnor – ttunor*) alors que dans la condition *transfert* les non-mots incluaient des doublets formés de consonnes jamais doublées en Français (e.g., *xihhel – xxihel*).

Les résultats de Pacton et al. (2001) et Danjon et Pacton (2009) ont confirmé le transfert à un nouveau matériel et la différence de performances entre items nouveaux et familiers rapportés dans les études d'apprentissage de grammaires artificielles. Mais, le résultat majeur concernant le degré d'abstraction de la connaissance acquise implicitement était qu'aucune règle ne semblait acquise, même chez des enfants exposés à un très grand nombre de mots au fil de cinq années d'école. Dans le premier chapitre expérimental de cette thèse,

ce type d'expérience a été repris avec deux groupes d'enfants de l'école élémentaire (CE2 et CM2) et un groupe d'adultes afin de préciser la nature des connaissances graphotactiques descriptibles sous forme de règles après une pratique encore beaucoup plus importante que chez les élèves les plus âgés (scolarisés en CM2) des études de Pacton et al. (2001) et Danjon et Pacton (2009). Un résultat particulièrement intéressant est que des élèves de l'école élémentaire et même des adultes semblent se fonder sur leur sensibilité aux régularités graphotactiques même lorsqu'il est possible de se fonder sur des règles morphologiques (Deacon & Pacton, 2007 ; Kemp & Bryant, 2003 ; Pacton, Fayol & Perruchet, 2005).

### 3.3) L'influence des régularités graphotactiques même lorsque le recours à des règles morphologiques est possible.

La plupart des études ont examiné soit l'impact des seules régularités graphotactiques, soit l'impact des seules informations morphologiques. Trois études ont abordé la question de l'intégration des connaissances morphologiques et graphotactiques. En Anglais, ces études ont exploité le fait que la transcription du son /z/ est contrainte à la fois par des régularités morphologiques et par des régularités graphotactiques. Concernant la morphologie, le pluriel des noms réguliers s'écrit *-s* qu'il se prononce /s/ (e.g., *cats*) ou /z/ (e.g., *dogs*) et concernant les régularités graphotactiques, le son /z/ est toujours transcrit *-s* après une consonne alors qu'il peut se transcrire *s*, *z*, *zz*, *ze* ou *se* après une voyelle longue (e.g., *fleas*, *please*, *breeze*). L'hypothèse de Kemp et Bryant (2003) était que si les sujets appliquent une règle spécifiant que le pluriel des noms réguliers anglais s'écrit toujours *-s*, alors /z/ devrait être orthographié correctement aussi souvent après une consonne qu'après une voyelle longue. En revanche, s'ils se fondent uniquement, ou au moins en partie, sur leurs connaissances graphotactiques, le marquage du pluriel nominal (/z/ orthographié *-s*) devrait être plus souvent correct après une consonne qu'après une voyelle longue. Kemp et Bryant ont trouvé un impact des régularités graphotactiques en production orthographique de mots chez des élèves de 7-9 ans et en production orthographique de non-mots chez des adultes. Par exemple, les élèves de 7-9 ans transcrivaient le pluriel nominal /z/ plus souvent correctement avec *-s* après une consonne (e.g., *dogs* – i.e. *chiens*) qu'après une voyelle



longue (e.g., *fleas* – i.e. *puces*), en dépit de la possibilité d'appliquer la règle spécifiant que les noms pluriels réguliers se terminent par *-s*. Dans la tâche de production orthographique de non-mots, les enfants et les adultes utilisaient les informations morphologiques fournies par la phrase puisqu'ils transcrivaient plus souvent le son /z/ final avec un *-s* lorsque les non-mots étaient insérés dans des phrases comme des noms pluriels que lorsqu'ils apparaissaient comme des noms singuliers ou des infinitifs verbaux. Toutefois, les non-mots insérés dans des phrases comme des noms pluriels étaient plus souvent orthographiés correctement avec *-s* après une consonne qu'après une voyelle longue.

Dans l'étude de Kemp et Bryant (2003), il y avait toujours des contraintes morphologiques puisque les non-mots étaient toujours insérés dans des phrases dans lesquelles ils correspondaient à des noms pluriels, des noms singuliers ou des infinitifs verbaux. Cette situation permettait donc d'examiner l'influence des régularités graphotactiques lorsqu'il est possible de se fonder sur une règle mais elle ne permettait pas d'étudier si l'impact des contraintes graphotactiques variait selon que des contraintes morphologiques sont présentes ou absentes. Cette question a été abordée par Deacon et Pacton (2007) en demandant à des adultes d'écrire des non-mots se terminant par consonne + /z/ et des non-mots se terminant par voyelle longue + /z/ dans deux conditions expérimentales. Dans la première condition, les non-mots étaient dictés de façon isolée. Dans la deuxième, les non-mots dictés étaient insérés dans des phrases comme dans l'étude de Kemp et Bryant (2003). L'effet graphotactique (la différence d'utilisation de *-s* selon que les non-mots se terminent par /consonne + z/ ou /voyelle + z/) était observé dans les deux conditions. Il était plus large dans la condition *isolée* (70%) que dans la condition *non-mots insérés dans des phrases* (52%), ce qui indique que, même si les contraintes morphologiques réduisaient un peu l'impact des contraintes graphotactiques, l'impact des contraintes graphotactiques demeurait très important.

En Français, Pacton, Fayol et Perruchet (2005) ont également comparé l'impact des régularités graphotactiques selon qu'il y avait ou non des contraintes morphologiques afin d'investiguer si des élèves de l'école élémentaire et des adultes recouraient à des règles relatives à la transcription des suffixes diminutifs /o/ et /ɛt/. La transcription de ces deux sons peut être contrainte par la morphologie (e.g., /ɛt/ est toujours transcrit *ette* quand il correspond à un suffixe diminutif) et par des régularités graphotactiques (e.g., *ette* est plus fréquent après *r* qu'après *f*). Dès le CE1, les élèves transcrivaient /ɛt/ plus souvent *ette*

lorsque des non-mots tels que /vitarɛt/ et /vitafɛt/ étaient insérés dans des phrases comme “une petite /vitar/ est une /vitarɛt/” qui révèlent leur structure morphologique (*i.e.*, un radical suivi d’un suffixe diminutif) que lorsqu’ils étaient dictés isolément ou dans des phrases qui n’indiquaient pas qu’il s’agissait d’un suffixe diminutif. Les orthographe des élèves et des étudiants étaient également influencées par les régularités graphotactiques puisqu’ils transcrivaient /ɛt/ plus souvent *ette* pour des non-mots comme /vitarɛt/ que pour des non-mots comme /vitafɛt/ même lorsque ces non-mots étaient insérés dans des phrases indiquant qu’il s’agissait de diminutifs. Cet effet n’aurait pas dû être observé si les sujets avaient recouru à une règle spécifiant que /ɛt/ se transcrit *ette* quand il correspond à un suffixe diminutif. De plus, l’amplitude de cet effet ne variait pas en fonction de l’âge.

Ainsi, les rares études ayant abordé la question de l’intégration des contraintes graphotactiques et morphologiques, en Anglais avec de la morphologie flexionnelle (marquage du pluriel des noms réguliers) et Français avec de la morphologie dérivationnelle (marquage des suffixes diminutifs), suggèrent que des élèves de l’école élémentaire et des adultes utilisent les informations morphologiques mais se fondent également sur leur connaissances graphotactiques, même lorsque le recours à des règles morphologiques est possible.

#### 3.4). Sensibilité aux régularités graphotactiques et apprentissage de l’orthographe lexicale.

L’influence de la sensibilité à des régularités graphotactiques sur l’acquisition de connaissances lexicales a été suggérée par Campbell et Coltheart (1984) lors de la sortie du film *Gandhi*. Même si les 57 étudiants anglais en licence de cette étude avaient probablement souvent vu ce nom correctement écrit, 49 d’entre eux commettaient l’erreur GHANDI. L’absence d’erreurs telles que GANDY ou GANDI indique que les sujets savaient que l’orthographe *Gandhi* comportait un *h* mais ne savaient pas où le placer. L’erreur GHANDI serait particulièrement fréquente parce que cette orthographe est beaucoup plus congruente avec les régularités graphotactiques de l’Anglais que des orthographe incorrectes comme GANHDI, GANDIH et que l’orthographe correcte *Gandhi* (*gh* est beaucoup plus fréquent que *dh* ou *ih*). Les conclusions pouvant être tirées de cette étude sont néanmoins limitées puisqu’elle portait sur la production d’un seul item et ne contrôlait pas le nombre d’expositions à l’orthographe correcte *Gandhi*.

Un tel contrôle a été effectué dans des études investiguant l'apprentissage de l'orthographe de nouveaux mots. Dans l'étude de Wright et Ehri (2007), des élèves américains de 6 ans apprenaient explicitement comment se prononçaient de nouvelles orthographe. Certaines commençaient par une consonne double (e.g., *rrug*), ce qui est illégal en Anglais, alors que d'autres se terminaient par des doublets, ce qui est légal en Anglais (e.g., *rugg*). Dans un test ultérieur, les élèves rappelaient mieux les doublets en position finale qu'en position initiale. Pour les items incluant un doublet en position initiale, les erreurs d'omission (i.e., orthographier un item incluant un doublet sans doublet, e.g., *rrug* orthographié RUG) étaient les plus fréquentes. Les élèves commettaient également des erreurs de transposition, doublant la consonne finale à la place de la consonne initiale (e.g., *rrug* orthographié RUGG). En revanche, ils ne doublaient presque jamais la consonne initiale à la place de la consonne finale et ne doublaient presque jamais la consonne finale pour les items appris sans doublet. Ce résultat suggère que, de la même façon que les étudiants anglais se rappelaient que le mot *Gandhi* inclut un *h*, les enfants se rappelaient qu'une lettre était doublée et donc doublaient une consonne, mais pas nécessairement la correcte. Leur tendance à doubler la consonne en position légale, plutôt qu'illégale, suggère que les élèves se fondaient à la fois sur leur connaissance relative à l'orthographe d'un mot spécifique (que ce mot inclut un doublet) et sur leur connaissance orthographique générale (que les doublets ne peuvent pas survenir en début de mots).

Fayol, Treiman, Lété et Pacton (2010) ont étudié l'impact de la sensibilité d'adultes à la fréquence de doublement des consonnes sur leur apprentissage de l'orthographe de nouveaux mots contenant des doubles lettres. Alors que l'étude de Campbell et Coltheart (1984) portait sur le seul mot *Gandhi*, avec la combinaison de lettre rare *dh*, Fayol et al. ont utilisé des orthographe plus ou moins consistantes avec les régularités graphotactiques du Français afin d'explorer si certaines erreurs sont restreintes à des orthographe contenant des patterns rares. Alors que les enfants de l'étude de Wright et Ehri (2007) apprenaient explicitement la prononciation de mots présentés de façon isolée, Fayol et al. ont utilisé une situation d'apprentissage incident. Les sujets étaient exposés à de nouvelles orthographe placées dans des textes qu'ils devaient lire afin de les comprendre. Une autre différence entre ces deux études tenait à ce que Wright et Ehri comparaient l'apprentissage d'orthographe respectant les régularités graphotactiques à l'apprentissage d'orthographe ouvertement illégales (e.g., *rrug* avec un doublet en position illégale) alors que Fayol et al. (2010) comparaient l'apprentissage d'orthographe incluant des patterns fréquents à l'apprentissage d'orthographe incluant des patterns rares mais légaux. Fayol et al. (2010) ont exploité le fait

que les consonnes *n*, *r* et *t* sont fréquemment doublées en Français alors que *b*, *d* et *g* sont plus rarement doublées (35 fois plus de mots avec *nn*, *rr* et *tt* que de mots avec *bb*, *dd* et *gg* selon la base de données lexicale Lexique, New, Pallier, Ferrand & Matos, 2001).

Dans la phase d'apprentissage de l'étude de Fayol et al. (2010), des étudiants de l'université étaient exposés à trois types d'orthographe de non-mots se prononçant de façon identique : des items incluant seulement des consonnes simples (e.g., *tidunar*), des items incluant un doublet fréquent (e.g., *tidunnar*) et des items incluant un doublet rare (e.g., *tiddunar*). Après la lecture de chaque texte, les participants devaient répondre à des questions sur le contenu de l'histoire. L'une nécessitait l'écriture du nouvel item. Les participants n'avaient pas d'instruction explicite d'écrire l'item comme il était écrit dans l'histoire. Les résultats ont montré que les participants orthographiaient correctement (i.e., comme dans l'histoire) plus souvent les items incluant seulement des consonnes simples et les items incluant un doublet fréquent (sans différence entre ces deux types d'items) que les items incluant un doublet rare. De plus, l'analyse des erreurs montrait que les transpositions du doublement à une autre consonne étaient plus nombreuses pour les items incluant un doublet rare que pour les items incluant un doublet fréquent. Par exemple, *tiddunar* était plus souvent orthographié incorrectement *tidunnar*, avec la transposition du doublet du *d*, qui est rarement doublé en Français, au *n*, qui est fréquemment doublé en Français que *tidunnar* était orthographié *tiddunar*.

Ainsi, les trois études rapportées dans cette section suggèrent que la concordance entre la structure orthographique des mots et les régularités graphotactiques de la langue déterminerait en partie les difficultés d'acquisition des connaissances lexicales. Ce résultat semble robuste puisqu'il a été rapporté avec différents types de régularités (présence du *h* dans l'étude de Cappel & Coltheart, 1984 ; présence de doublets dans les études de Wright & Ehri, 2007 et de Fayol et al., 2010), dans différentes situations expérimentales (apprentissage de listes de mots présentés isolément dans l'étude de Wright & Ehri, 2007 vs. apprentissage incident de mots insérés dans des textes dans l'étude de Fayol et al., 2010 ; en contrastant des patterns fréquents à des patterns illégaux dans l'étude de Wright & Ehri, 2007 vs. en contrastant des patterns fréquents et légaux à des patterns rares mais également légaux dans l'étude de Fayol et al., 2010), avec des populations différentes (en tout début d'apprentissage de l'écrit dans l'étude de Wright & Ehri, 2007 vs. avec des adultes dans les études de Fayol et al., 2010 et Campbell & Coltheart, 1984) et dans différentes langues (e.g., Français et Anglais).

Les expériences rapportées dans les chapitres expérimentaux III et IV visent à prolonger l'étude de Fayol et al. (2010) afin de mieux comprendre le rôle de différentes connaissances graphotactiques dans l'acquisition de l'orthographe lexicale chez des élèves de l'école élémentaire et chez des adultes. Un de ces prolongements consistera à déterminer si l'influence de connaissances sur la fréquence d'occurrences de certains patterns orthographiques rapportée par Fayol et al. demeure lorsqu'il est explicitement demandé aux sujets d'orthographier les nouveaux mots tels qu'ils apparaissaient dans les histoires.

## CHAPITRE EXPERIMENTAL I

### Sensibilité à trois propriétés graphotactiques de l'emploi des doubles lettres chez des élèves de l'école élémentaire (CE2 et CM2) et chez des adultes

Des études antérieures ont mis en évidence une sensibilité à des régularités graphotactiques, notamment avec des épreuves de jugement de pseudo-mots. Par exemple, dès la première primaire, des élèves français estiment que le non-mot *tunnor* constitue une meilleure orthographe que le non-mot *ttunor*, ce qui montre une certaine connaissance du fait que les consonnes peuvent être doublées en position médiane mais pas en début de mots en français. D'autres propriétés de l'emploi des doubles lettres seraient plus difficilement repérées. Il en va ainsi du fait qu'une consonne double peut survenir avant mais non après une consonne simple (e.g., apprendre mais pas aPRRendre ; danse mais pas daNSSe). Ainsi, lorsque Danjon et Pacton (2009) ont proposé des paires comme *apprulir* – *aprrulir* à des élèves de primaire, les performances des élèves de première primaire ne différaient pas du hasard. Celles d'élèves plus âgés étaient certes supérieures au hasard mais elles demeuraient plus faibles que celles pour des paires testant la sensibilité au fait que les consonnes ne peuvent pas être doublées en début de mots.

De plus, dans les études de Danjon et Pacton (2009) et Pacton et al. (2001), tout au long de l'école élémentaire, les performances des élèves variaient en fonction de la familiarité du matériel utilisé pour tester la connaissance de propriétés de l'emploi des doubles lettres qui peuvent être décrites sous forme de règle. Par exemple, les performances étaient meilleures lorsque les non-mots utilisés pour tester la connaissance de l'illégalité des doublets en position initiale incluaient des doublets fréquents (e.g., *tunnor* – *ttunor*) plutôt que des doublets formés avec des consonnes jamais doublées en français (e.g., *kixxor* – *kkixor*). De façon très similaire, les performances étaient meilleures lorsque les non-mots utilisés pour tester la connaissance de l'illégalité des doublets après une consonne simple incluaient un groupe consonantique placé après une voyelle après laquelle il est fréquent (e.g., *apprulir* – *aprrulir*) plutôt qu'après une voyelle après laquelle il est rare (e.g., *ippraler* – *iprraler*). Cette différence, selon la familiarité du matériel utilisé pour évaluer les connaissances graphotactiques, n'aurait pas dû être observée si les élèves avaient eu recours à des règles spécifiant que les consonnes ne peuvent pas être doublées en début de mots ou après une

consonne simple. Selon Pacton et al. (2001), ce résultat est similaire au phénomène de *transfer-decrement* observé dans les études d'apprentissage de grammaires artificielles dans lesquelles le paradigme de transfert est utilisé. De plus, sa persistance tout au long de l'école élémentaire suggère qu'il ne peut pas être attribué à la pratique nécessairement limitée, tant en nombre d'items qu'en durée d'apprentissage, qui caractérise les situations d'apprentissage en laboratoire.

Le premier objectif de cette étude est de confirmer que certaines régularités graphotactiques (l'illégalité des doublets après une consonne simple) sont plus difficilement apprises que d'autres (l'illégalité des doublets en début de mots) chez des élèves de primaire en utilisant deux tâches : un jugement de pseudo-mots (Cassar & Treiman, 1997 ; Pacton et al., 2001) et une décision orthographique sur des mots. Cette nouvelle tâche consiste en un choix entre une orthographe correcte et une orthographe incorrecte, phonologiquement plausible, mais qui viole une règle graphotactique par la présence d'un doublet en début de mot (e.g., *menotte* – *mnotenote*) ou après une consonne simple (e.g., *affront* – *afront*). Nous faisons l'hypothèse que dans les deux tâches, les performances d'élèves de primaire seront meilleures pour les items testant la connaissance de l'illégalité des doublets en début de mots que pour les items testant l'illégalité des doublets après une consonne simple. Le second objectif est d'étudier si des adultes, dont la pratique de l'écrit est beaucoup plus prolongée que celle des élèves les plus âgés des études antérieures (CM2 dans les études de Danjon & Pacton, 2009 et Pacton et al., 2001), recourent à des règles graphotactiques lorsque cela est possible. Notre hypothèse est que leurs performances devraient être très élevées aussi bien lorsque les connaissances graphotactiques sont évaluées avec un matériel familier (e.g., doublet fréquent ; environnement gauche fréquent) que lorsqu'elles sont évaluées avec un matériel non familier (doublet formé de consonnes jamais doublées ; environnement gauche rare) mais qu'elles devraient néanmoins être significativement plus faibles avec un matériel non familier.

## **Méthode**

### **Participants**

Cinquante étudiants (15 hommes, âge moyen = 26 ans et 2 mois, EC = 6 ans et 8 mois), 50 élèves de CE2 (28 garçons, âge moyen = 8ans et 3 mois, EC = 6 mois) et 50 élèves de CM2 (24 garçons, âge moyen = 10 and 4 mois, EC = 6 mois), tous de langue maternelle française ont participé à cette expérience.

## Matériel

### Paires de non-mots évaluant les connaissances graphotactiques.

Cinquante-six paires de non-mots ont été construites afin d'évaluer la sensibilité des sujets à quatre régularités graphotactiques du français. Trois concernent le doublement des consonnes. Une concerne l'utilisation de *eau* pour transcrire le phonème /o/. Pour chaque paire expérimentale, les deux non-mots se prononçaient de la même manière (cf., *Annexe 1*).

#### **Items identité des consonnes pouvant être doublées (n = 12)**

Ces paires évaluaient la sensibilité au fait que seuls certains doublets sont permis en français. Chaque paire incluait un non-mot avec un doublet fréquent (e.g., *ll*) et un non-mot avec un doublet inexistant en français (e.g., *jj*), en position médiane dans les deux cas. Un exemple de ces paires est *ullate* – *ujjate*. Si les participants sont sensibles à cette propriété, ils devraient plus souvent sélectionner les non-mots incluant un doublet fréquent (e.g., *ullate*) que ceux incluant un doublet inexistant (e.g., *ujjate*).

#### **Items illégalité des doublets en position initiale (n = 16)**

Ces paires évaluaient la sensibilité au fait que les doublets sont illégaux en position initiale en français. Pour chaque paire, un non-mot incluait un doublet cible en position médiane (légale) et l'autre en position initiale (illégal). Ce doublet cible était formé avec des consonnes fréquemment doublées en français pour huit paires notées *position F+* (e.g., *nummar* – *nnummar*) versus jamais doublées pour huit autres paires notées *position F0* (e.g., *jjaxir* – *jaxxir*). Si les participants sont sensibles à cette propriété, ils devraient plus souvent sélectionner les non-mots incluant des doublets en position médiane (e.g., *nummar*, *jaxxir*) qu'en position initiale (e.g., *nnummar*, *jjaxir*).

#### **Items illégalité des doublets après une consonne simple (n = 16)**

Ces paires évaluaient la sensibilité au fait qu'un doublet peut précéder, mais non suivre, une consonne simple en français. Pour chaque paire, le doublet était placé avant la consonne simple dans un non-mot et après la consonne simple dans l'autre (e.g., *ppr* / *prp*). Dans huit paires notées *groupes consonantiques F+*, le groupe consonantique suivait une voyelle après laquelle il apparaît souvent en français, (e.g., *ppr* fréquent après *a*, comme dans la paire *apprulir* – *aprrulir*). Dans huit paires notées *groupes consonantiques F-*, le groupe consonantique suivait une voyelle après laquelle il apparaît rarement en français, (e.g., *ppr* rare après *i* comme dans la paire *ippraler* – *iprraler*). Si les participants sont sensibles à cette propriété, ils devraient plus souvent sélectionner les non-mots incluant des doublets avant une



consonne simple (e.g., *apprulir*, *ippraler*) que ceux incluant des doublets après une consonne simple (e.g., *aprrulir*, *iprraler*).

### **Probabilité de *eau* en fonction de l'environnement gauche ( $n=12$ )**

Ces paires évaluaient la sensibilité au fait que la probabilité que /o/ soit transcrit *eau* varie en fonction de la consonne gauche. Pour chaque paire, /o/ était transcrit *eau* dans un non-mot et *ot* dans l'autre. Dans six paires, *eau* suivait une consonne après laquelle il est fréquent en français (e.g., *eau* après /v/ dans la paire *lavireau* – *lavivot*). Dans six paires, *eau* suivait une consonne après laquelle il est rare en français (e.g., *eau* après /f/ dans la paire *durifeau* – *durifot*). Si les participants sont sensibles à cette propriété, ils devraient plus souvent sélectionner les non-mots se terminant par *eau* pour le premier type de paires que pour le second (i.e., *eau* devrait être davantage sélectionné pour des paires comme *lavireau* – *lavivot* que pour des paires comme *durifeau* – *durifot*).

Pour les quatre types d'items décrits ci-dessus, un type de non-mot (e.g., les non-mots avec un doublet en position médiane plutôt qu'en position initiale pour les paires testant la sensibilité au fait que les consonnes ne sont jamais doublées en début de mots) était placé à droite pour la moitié des paires et à gauche pour l'autre moitié. Ces 56 items étaient placés juste après cinq items essais, dans un ordre pseudo-aléatoire, avec la contrainte qu'il y ait au moins deux items de catégories différentes entre deux items d'une même catégorie. Différentes aléatorisations ont été utilisées pour chaque participant. L'ensemble des items était imprimé sur des feuilles de format A4 en police *New York font*, taille 14, avec environ 10 items par page.

### Décision orthographique sur des mots

Trente paires ont été créées pour la tâche de décision orthographique sur des mots (*cf.*, *Annexe 2*). Chaque paire incluait l'orthographe correcte et une orthographe incorrecte mais phonologiquement plausible. Dans neuf paires notées *doublet illégal en début de mots*, l'orthographe incorrecte incluait une consonne médiane simple (alors qu'elle doit être doublée) et la consonne en position initiale doublée (alors qu'elle doit être simple) (e.g., *menotte* – *mmenote* ; *raffinage* – *rffinage*). Dans neuf paires notées *doublet illégal après une consonne simple*, l'orthographe incorrecte incluait la première consonne du groupe consonantique simple (alors qu'elle doit être doublée) et la seconde consonne du groupe consonantique doublée (alors qu'elle doit être simple) (e.g., *coffrage* – *coffrage* ; *affront* – *afront*). Dans 12 paires *tamppons*, une orthographe se terminait par *eau* et l'autre par *o*, *eau*

étant le graphème correct pour certains mots et le graphème incorrect pour d'autres (e.g., écriteau – écrito ; recto – recteau). L'orthographe correcte était placée à droite pour la moitié des items et à gauche pour l'autre moitié. Ces 30 paires étaient placées dans un ordre pseudo-aléatoire, avec la contrainte qu'il y ait au moins un item d'une autre catégorie entre deux items d'une même catégorie. Différentes aléatorisations ont été utilisées pour chaque participant. L'ensemble des items était imprimé sur des feuilles de format A4 en police *New York font*, taille 14, avec 10 items par page.

### Tests standards de lecture et orthographe

**Test de lecture.** Le niveau de lecture des enfants a été évalué avec le sous-test L3 de la batterie ORLEC (Lobrot, 1973), qui fournit une mesure globale du niveau de compréhension en lecture. Ce test de complétion en choix forcé est constitué de 36 phrases avec pour chacune d'entre elles, 5 alternatives, comme par exemple : *Il est arrivé une drôle d'aventure à un pêcheur, il a attrapé une (carpe, tanche, godasse, truite, perche)*. Au fur et à mesure du test, la complexité de la tâche augmente, les mots deviennent moins fréquents et les caractéristiques pragmatiques et syntaxiques des phrases deviennent plus complexes. Les enfants doivent répondre à autant de phrases qu'ils le peuvent en 5 minutes, et leur score correspond au nombre de réponses correctes fournies dans ce laps de temps. (cf., *Annexe 3*)

**Test d'orthographe.** Le sous-test de dictée "Le Corbeau" de la batterie L2MA (batterie du Langage oral, du Langage écrit, de la Mémoire et de l'Attention ; Chevrier-Muller et al., 1997) fournit quatre scores. Un *score phonologique* correspond à l'habilité à produire des orthographe phonologiquement plausibles même si elles ne correspondent pas forcément aux formes orthographiques conventionnelles. Un *score lexical*, ou d'orthographe d'usage, correspond à l'habilité à produire des orthographe conventionnelles. Un *score grammatical* correspond à l'habilité à marquer les aspects morphosyntaxiques du langage écrit. Le score global correspond à la somme des trois scores précédents (cf., *Annexe 4*).

### Procédure

L'expérience s'est déroulée de façon collective pour les enfants (une douzaine d'enfants par groupe) et de façon individuelle pour les adultes. L'expérience débutait par l'épreuve de

jugement de pseudo-mots. Les participants étaient avertis que les mots imprimés étaient tous des mots inventés qu'ils n'avaient jamais vus ou entendus auparavant. Il leur était ensuite demandé de bien regarder chaque paire de non-mots essais et d'entourer le non-mot qui constituerait la meilleure orthographe pour un nouveau mot français. Les participants répondaient d'abord aux cinq items essais pendant que l'expérimentateur s'assurait de la bonne compréhension de la consigne puis ils traitaient seuls les 56 paires expérimentales. Ensuite, les participants effectuaient la tâche de décision orthographique sur les mots. L'expérimentateur leur expliquait qu'une série de mots orthographiés de deux façons était écrite sur les feuilles qu'il venait de leur fournir et leur demandait d'entourer l'orthographe correcte pour chaque paire. Ensuite, les enfants passaient les épreuves de lecture du Lobrot (Orlec, 1963) et la dictée du corbeau du L2MA (Chevrier-Muller, 1997).

## **Résultats**

### **1. Tests standards de lecture et orthographe.**

Les élèves avaient des scores dans la norme au test de lecture du Lobrot (Orlec, 1983) et au test d'orthographe « dictée du corbeau (Chevrier-Muller et al., 1997). Les scores étaient significativement meilleurs en CM2 qu'en CE2 au test de lecture ( $M=19.6$  (sur 36),  $EC=6.9$  en CE2 vs.  $M=25.9$ ,  $EC=5.6$  en CM2,  $t(98) = 5.01$ ,  $p < .001$ ) et à la dictée du corbeau ( $M=37.9$  (sur 50),  $EC=6.1$  en CE2 vs.  $M=43.4$ ,  $EC=4.3$  en CM2,  $t(98) = 5.20$ ,  $p < .001$ ).

### **2. Epreuve de décision orthographique sur des mots.**

La Figure 1 présente les résultats à l'épreuve de décision orthographique sur des mots. Les performances, proches de 100% (98.9%) pour les deux types d'items chez les adultes, étaient plus faibles chez les élèves de primaire. Elles étaient meilleures en CM2 (89.8%) qu'en CE2 (78.0%) et elles étaient meilleures aux deux niveaux scolaires pour les items *doublet illégal en début de mots* que pour les items *doublet illégal après une consonne simple*. Une ANOVA sur le nombre de sélections correctes avec les variables groupe (ce2, cm2 et adultes) et type d'item (doublet illégal en début de mots et doublet illégal après une consonne simple), utilisant les sujets ( $F_1$ ) et les items ( $F_2$ ) comme variables aléatoires, révèle un effet significatif du groupe ( $F_1(2, 147) = 22.23$ ,  $MSE = 3.99$ ,  $p < .001$ ,  $\eta_p^2 = .23$  ;  $F_2(2, 32) = 59.91$ ,

MSE = 8.23,  $p < .001$ ,  $\eta_p^2 = .79$ ), un effet du type d'item ( $F_1(1, 147) = 75.58$ , MSE = 1.21,  $p < .001$ ,  $\eta_p^2 = .34$  ;  $F_2(1, 16) = 24.37$ , MSE = 20.94,  $p < .001$ ,  $\eta_p^2 = .60$ ) et une interaction ( $F_1(2, 147) = 18.31$ , MSE = 1.21,  $p < .001$ ,  $\eta_p^2 = .20$  ;  $F_2(2, 32) = 15.01$ , MSE = 8.23,  $p < .001$ ,  $\eta_p^2 = .48$ ). Les scores moyens des CM2 étaient meilleurs que ceux des CE2 ( $F_1(1, 147) = 14.06$ , MSE = 3.99,  $p < .001$ ,  $\eta_p^2 = .09$  ;  $F_2(1, 16) = 80.23$ , MSE = 12.23,  $p < .001$ ,  $\eta_p^2 = .83$ ) et plus faibles que ceux des adultes ( $F_1(1, 147) = 8.41$ , MSE = 3.99,  $p = .004$ ,  $\eta_p^2 = .05$  ;  $F_2(1, 16) = 55.21$ , MSE = 5.65,  $p < .001$ ,  $\eta_p^2 = .78$ ). L'interaction reflète le fait que les scores étaient meilleurs pour les items *doublet illégal en début de mot* que pour les items *doublet illégal après une consonne simple* en CE2 (différence de 22.2%,  $F_s > 24.39$ ,  $p_s < .001$ ) et CM2 (différence de 13.3%,  $F_s > 14.30$ ,  $p_s < .002$ ) alors qu'ils ne différaient pas en fonction du type d'items chez les adultes dont les deux scores approchaient 100%.

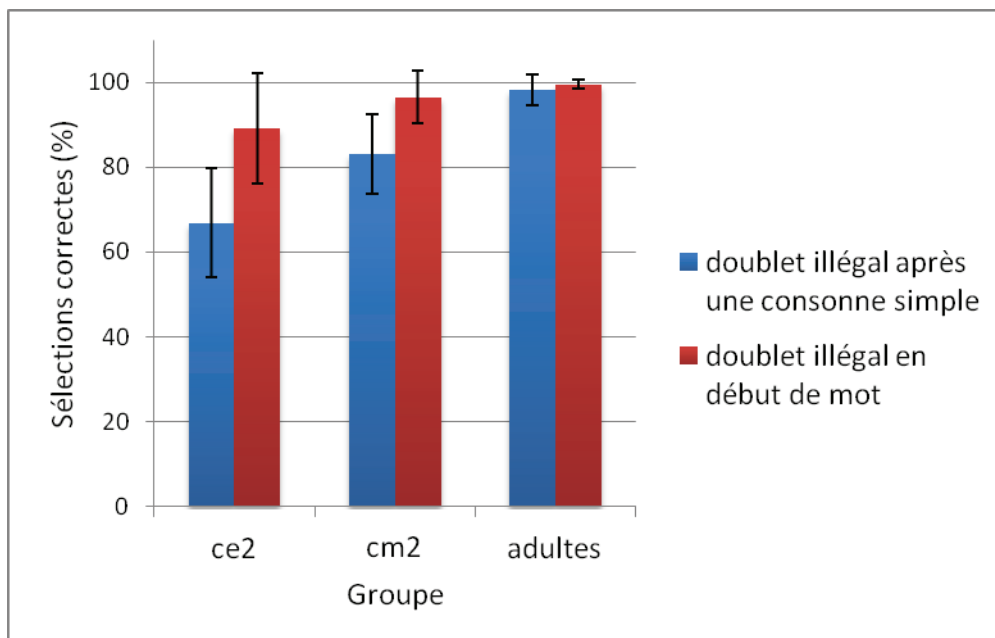


Figure 1. Pourcentages de sélection de l'orthographe correcte en fonction du type d'item et du groupe (Ecart-types en barre d'erreur).

### **3. Epreuve de jugement de l'orthographe de pseudo-mots évaluant les connaissances graphotactiques.**

#### **3.1. Identité des consonnes pouvant être doublées**

Les taux de sélections de l'orthographe incluant un doublet légal plutôt qu'illégal (items identité) représentés à la Figure 2 étaient significativement supérieurs au hasard (50%)

dans les trois groupes ( $ts > 9.24$ ,  $ps < .001$ ). Une ANOVA sur le nombre de sélections correctes avec la variable groupe (ce2, cm2 et adultes) montre que cet effet est significatif ( $F_1(2, 147) = 15.38$ ,  $MSE = 3.60$ ,  $p < .001$ ,  $\eta_p^2 = .17$  ;  $F_2(2, 22) = 30.30$ ,  $MSE = 7.62$ ,  $p < .001$ ,  $\eta_p^2 = .73$ ). Les scores des CM2 étaient meilleurs que ceux des CE2 ( $F_1(1, 147) = 9.66$ ,  $MSE = 3.60$ ,  $p = .002$ ,  $\eta_p^2 = .06$  ;  $F_2(1, 11) = 21.14$ ,  $MSE = 6.86$ ,  $p < .001$ ,  $\eta_p^2 = .66$ ) mais plus faibles que ceux des adultes ( $F_1(1, 147) = 5.87$ ,  $MSE = 3.60$ ,  $p = .016$ ,  $\eta_p^2 = .04$  ;  $F_2(1, 11) = 24.34$ ,  $MSE = 3.62$ ,  $p < .001$ ,  $\eta_p^2 = .69$ ).

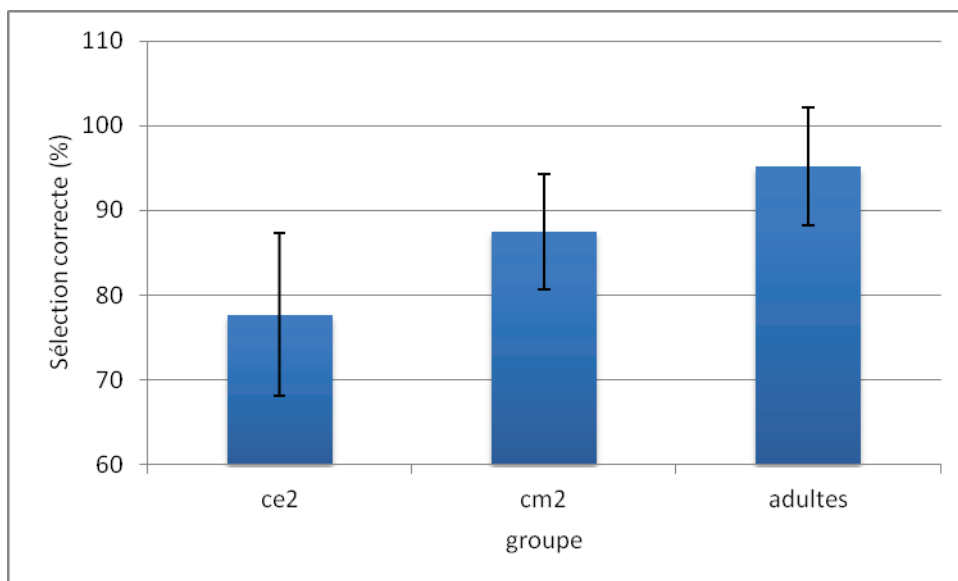


Figure 2. Pourcentages de sélection de l'orthographe incluant un doublet légal en fonction du groupe (Ecart-types en barre d'erreur).

### 3.2. Illégalité des doublets en position initiale

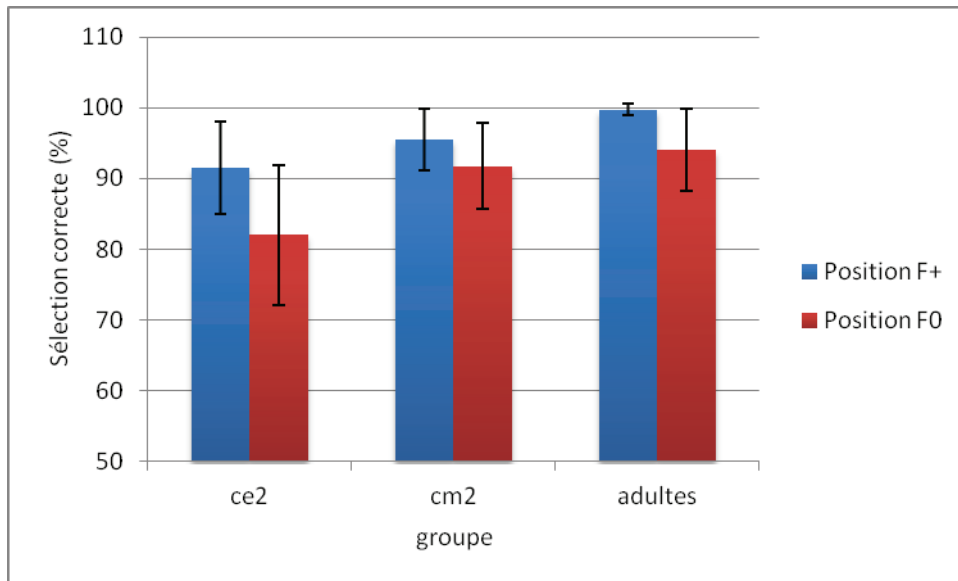


Figure 3. Pourcentages de sélection de l'orthographe avec un doublet en position légale (médiane) en fonction du groupe et du type d'items (Ecart-types en barre d'erreur)

Les taux de sélections de l'orthographe incluant un doublet en position médiane plutôt qu'en position initiale représentés à la Figure 3 étaient significativement supérieurs au hasard (50%) pour les items *Position F+* et *Position F0* dans les trois groupes ( $t_s > 11.46$ ,  $p_s < .001$ ). Une ANOVA sur le nombre de sélections correctes avec les variables groupe (ce2, cm2 et adultes) et type d'items (*Position F+* et *Position F0*) montre que les scores variaient significativement en fonction du groupe ( $F_1(2, 147) = 12.00$ ,  $MSE = 1.42$ ,  $p < .001$ ,  $\eta_p^2 = .14$  ;  $F_2(2, 28) = 36.28$ ,  $MSE = 2.94$ ,  $p < .001$ ,  $\eta_p^2 = .72$ ) et du type d'item ( $F_1(1, 147) = 34.51$ ,  $MSE = .56$ ,  $p < .001$ ,  $\eta_p^2 = .19$  ;  $F_2(1, 14) = 17.87$ ,  $MSE = 6.73$ ,  $p = .001$ ,  $\eta_p^2 = .56$ ). L'interaction Groupe x Type n'atteint pas le seuil de .05 ( $F_1(2, 147) = 2.44$ ,  $MSE = .56$ ,  $p = .09$ ,  $\eta_p^2 = .03$  ;  $F_2(2, 28) = 2.89$ ,  $MSE = 2.89$ ,  $p = .07$ ,  $\eta_p^2 = .17$ ). Des comparaisons planifiées indiquent que les scores des CM2 (93.6%) étaient significativement meilleurs que ceux des CE2 (86.8%) ( $F_1(1, 147) = 10.62$ ,  $MSE = 1.42$ ,  $p = .001$ ,  $\eta_p^2 = .07$  ;  $F_2(1, 14) = 33.56$ ,  $MSE = 2.82$ ,  $p < .001$ ,  $\eta_p^2 = .71$ ) et significativement plus faibles que ceux des adultes (96.9%) ( $F_1(1, 147) = 2.37$ ,  $MSE = 1.42$ ,  $p = .13$ ,  $\eta_p^2 = .02$  ;  $F_2(1, 14) = 7.23$ ,  $MSE = 2.92$ ,  $p = .018$ ,  $\eta_p^2 = .34$ ). En moyenne, les scores étaient meilleurs pour les items *Position F+* que pour les items *Position F0* (95.6% vs. 89.3%). Si la différence entre les scores pour les items *Position F+* et *Position F0* tendait à varier en fonction du groupe (9.5% en CE2, 3.8% en

CM2 et 5.8% chez les adultes), elle était significative dans les trois groupes ( $F_{1s} > 5.19$ ,  $ps < .027$  ;  $F_{2s} > 3.16$ ,  $ps < .017$ ).

Les scores sur les items *Position F+* et *Position F0* étaient corrélés ( $r = .54$ ,  $p < .001$ ), même après avoir contrôlé le facteur groupe ( $r = .48$ ,  $p < .001$ ). Les scores sur ces deux types d'items étaient également corrélés aux scores en décision orthographique pour les items *doublet illégal en début de mots* ( $rs = .52$ ,  $ps < .001$ ), même après avoir contrôlé le facteur groupe ( $rs > .48$ ,  $ps < .001$ ).

### 3.3. Illégalité des doublets après une consonne simple

Les taux de sélections de l'orthographe incluant un doublet avant plutôt qu'après une consonne simple représentés à la Figure 4 étaient significativement supérieurs au hasard (50%) pour les items *groupe consonantique F+* dans les trois groupes ( $ts > 6.92$ ,  $ps < .001$ ). Pour les items *groupe consonantique F-*, ils étaient significativement supérieurs au hasard en CM2 ( $t(49) = 7.42$ ,  $p < .001$ ) et chez les adultes ( $t(49) = 19.59$ ,  $p < .001$ ) mais pas en CE2 ( $t(49) = 1.33$ ,  $p = .19$ ).

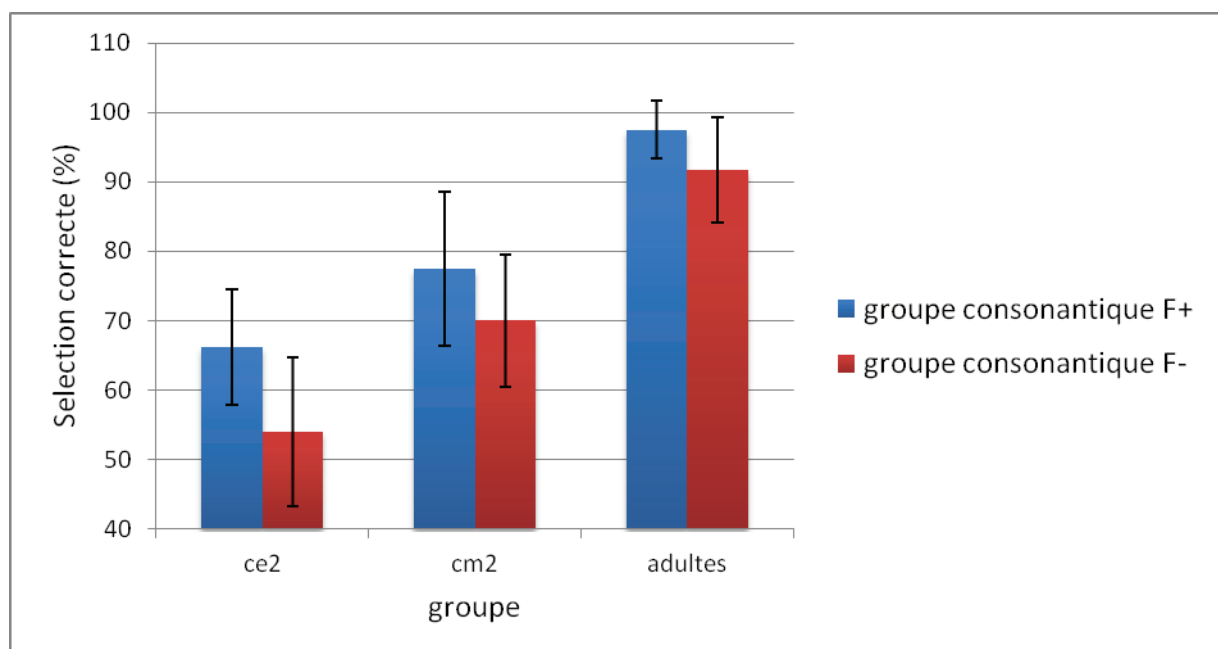


Figure 4. Pourcentages de sélection de l'orthographe avec un doublet en position légale (avant plutôt qu'après la consonne simple) en fonction du groupe et du type d'items (Ecart-types en barre d'erreur)

Une ANOVA sur le nombre de sélections correctes avec les variables groupe (ce2, cm2 et adultes) et type d'items (*groupe consonantique F+* et *groupe consonantique F-*) révèle un effet du groupe ( $F_1(2, 147) = 66.70$ ,  $MSE = 2.90$ ,  $p < .001$ ,  $\eta_p^2 = .48$  ;  $F_2(2, 28) = 133.58$ ,  $MSE = 9.04$ ,  $p < .001$ ,  $\eta_p^2 = .91$ ), un effet du type d'item ( $F_1(1, 147) = 30.58$ ,  $MSE = 1.13$ ,  $p < .001$ ,  $\eta_p^2 = .17$  ;  $F_2(1, 14) = 7.45$ ,  $MSE = 29.11$ ,  $p = .016$ ,  $\eta_p^2 = .35$ ) mais pas d'interaction ( $ps > .21$ ). Des comparaisons planifiées montrent que les scores des CM2 (73.8%) étaient significativement meilleurs que ceux des CE2 (60.1%) ( $F_1(1, 147) = 20.50$ ,  $MSE = 2.90$ ,  $p < .001$ ,  $\eta_p^2 = .12$  ;  $F_2(1, 14) = 53.62$ ,  $MSE = 6.92$ ,  $p < .001$ ,  $\eta_p^2 = .79$ ) mais plus significativement faibles que ceux des adultes (94.6%) ( $F_1(1, 147) = 48.13$ ,  $MSE = 2.90$ ,  $p < .001$ ,  $\eta_p^2 = .25$  ;  $F_2(1, 14) = 115.72$ ,  $MSE = 7.53$ ,  $p < .001$ ,  $\eta_p^2 = .89$ ). En moyenne, les scores étaient meilleurs pour les items *groupe consonantique F+* que pour les items *groupe consonantique F-* (80.4% vs. 71.9%). Des analyses conduites pour chaque groupe montrent que l'effet du type d'items était significatif dans les trois groupes ( $F_{1s} > 4.74$ ,  $ps < .034$  ;  $F_{2s} > 3.03$ ,  $ps < .013$ ).

Les scores sur les items *groupe consonantique F+* et *groupe consonantique F-* étaient corrélés ( $r = .66$ ,  $p < .001$ ), même après avoir contrôlé le facteur groupe ( $r = .44$ ,  $p < .001$ ). Les scores sur ces deux types d'items étaient corrélés aux scores de décision orthographique pour les items *doublet illégal après une consonne simple* ( $rs = .58$ ), même après avoir contrôlé le facteur groupe ( $rs > .35$ ,  $ps < .001$ ).

#### 3.4. Comparaison des scores pour Illégalité des doublets en position initiale et Illégalité des doublets après une consonne simple

Afin de confirmer que les élèves de l'école élémentaire étaient plus sensibles à l'illégalité des doublets en début de mots qu'à l'illégalité des doublets après des consonnes simples, le nombre de sélections de l'orthographe correcte sur ces items (sans distinguer *Position F+* et *Position F0* pour la première propriété et sans distinguer *groupe consonantique F+* et *groupe consonantique F-* pour la seconde propriété) a été soumis à une ANOVA avec les variables groupe (CE2, CM2 et adultes) et propriété graphotactique (*doublet illégal en début de mot*, *doublet illégal après une consonne simple*). Les réponses correctes variaient en fonction du groupe (73.4% en CE2, 83.7% en CM2 et 95.8% chez les adultes,  $F_1(2, 147) = 54.00$ ,  $MSE = 5.91$ ,  $p < .001$ ,  $\eta_p^2 = .42$  ;  $F_2(2, 60) = 159.52$ ,  $MSE = 6.26$ ,  $p < .001$ ,  $\eta_p^2 = .84$ ) ; elles étaient en moyenne plus nombreuses pour les items *doublet illégal*



en début de mot que pour les items *doublet illégal après une consonne simple* (92.4% vs. 76.2%,  $F_1(1, 147) = 185.7$ ,  $MSE = 2.73$ ,  $p < .001$ ,  $\eta_p^2 = .56$  ;  $F_2(2, 60) = 56.66$ ,  $MSE = 27.96$ ,  $p < .001$ ,  $\eta_p^2 = .65$ ) et ces deux variables interagissaient ( $F_1(2, 147) = 37.14$ ,  $MSE = 2.73$ ,  $p < .001$ ,  $\eta_p^2 = .34$  ;  $F_2(2, 60) = 50.64$ ,  $MSE = 6.26$ ,  $p < .001$ ,  $\eta_p^2 = .63$ ). Cette interaction est due au fait que l'effet du type d'item était significatif seulement chez les élèves de l'école élémentaire (différence de 26.6% en CE2 et 19.9% en CM2,  $F_s > 60.43$ ,  $p_s < .001$  ; différence de 2.3% chez les adultes,  $F_1(1, 49) = 3.32$ ,  $MSE = .97$ ,  $p = .074$ ,  $\eta_p^2 = .06$  ;  $F_2(1, 30) = 2.06$ ,  $MSE = 4.91$ ,  $p = .16$ ,  $\eta_p^2 = .06$ ).

### 3.5 Eau en position finale en fonction de l'environnement gauche

Les taux de sélections de l'orthographe incluant le graphème *eau* sont représentés à la Figure 5. Une ANOVA avec les variables groupe (ce2, cm2 et adultes) et type d'items (*eau fréquent* et *eau rare*) révèle un effet du type d'items, avec *eau* plus souvent choisi pour les items *eau fréquent* que pour les items *eau rare* (50.3% vs. 33.4%,  $F_1(1, 147) = 84.57$ ,  $MSE = .91$ ,  $p < .001$ ,  $\eta_p^2 = .37$  ;  $F_2(1, 10) = 14.05$ ,  $MSE = 45.68$ ,  $p < .004$ ,  $\eta_p^2 = .58$ ), sans effet du groupe ( $p_s > .81$ ) ni interaction ( $p_s > .54$ ).

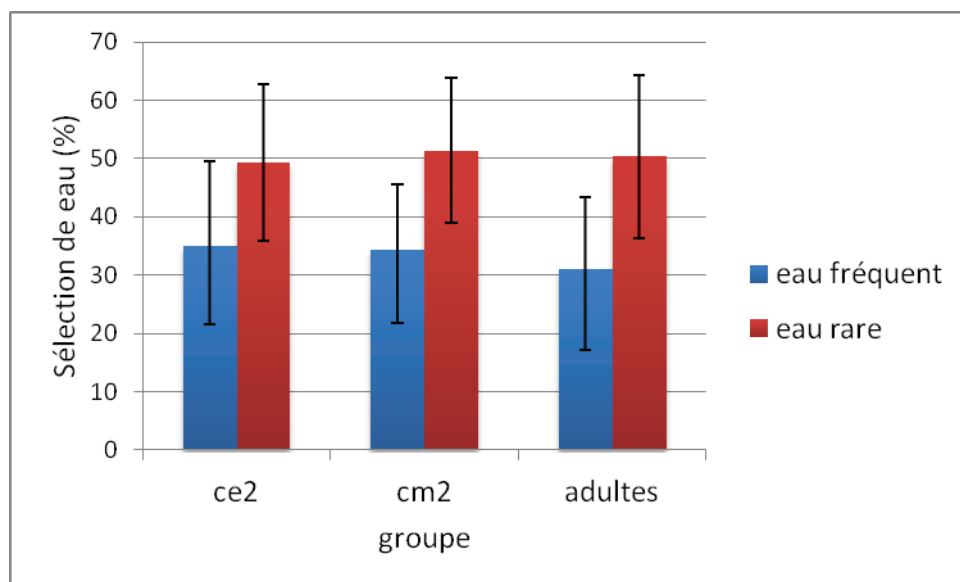


Figure 5. Pourcentages de sélection de l'orthographe avec le graphème eau en fonction du groupe et du type d'item (Ecart-types en barre d'erreur)

## Discussion

Cette expérience confirme que certaines régularités graphotactiques sont plus difficilement apprises que d'autres chez des élèves de primaire avec deux épreuves, l'une utilisée dans des études antérieures et consistant en un jugement de paires de non-mots, l'autre nouvelle et consistant en un choix entre une orthographe correcte et une orthographe incorrecte, phonologiquement plausible, mais qui viole une règle graphotactique par la présence d'un doublet en début de mot ou après une consonne simple. Comme dans l'étude de Danjon et Pacton (2009), les scores des CE2 et CM2 étaient plus faibles pour les items testant l'illégalité des doublets après une consonne simple que pour les items testant l'illégalité des doublets en position initiale en CE2 et CM2. Les scores des adultes ne différaient pas en fonction du type de propriété étudié (scores plafonds).

Pour les deux propriétés graphotactiques descriptibles sous la forme de règle, les scores moyens variaient en fonction de la familiarité du matériel utilisé pour tester les connaissances graphotactiques chez les élèves de l'école élémentaire mais aussi chez les adultes et les scores sur les items familiers et non familiers étaient corrélés même après avoir contrôlé l'âge des sujets. Ce résultat confirme et prolonge ceux d'études antérieures (e.g., Danjon & Pacton, 2009 ; Pacton et al., 2001) en montrant que l'effet de familiarité du matériel utilisé persiste même chez des adultes. Ce résultat, après des années d'exposition à l'écrit, va à l'encontre de la proposition de Manza et Reber (1997) selon laquelle le phénomène de *transfer-drecrement* dans les études d'apprentissage de grammaires artificielles serait dû à une pratique insuffisante pour permettre l'abstraction de règles abstraites, indépendantes des caractéristiques de surface du matériel.

Si cet effet de familiarité du matériel utilisé pour tester les connaissances d'adultes était significatif, les scores moyens étaient très proches de 100%. De ce fait, un nombre important de sujets obtenaient le score maximum pour les items familiers et non familiers. Mais, parmi les sujets qui obtenaient des scores différents sur les items familiers et non familiers, ceux ayant des scores plus élevés pour les items familiers étaient beaucoup plus nombreux que ceux ayant des scores plus élevés pour les items non familiers. Ces proportions pour les items testant l'illégalité des doublets en début de mots étaient de 50% vs. 8% en CE2, 26% vs. 12% en CM2 et 26% vs. 0% chez les adultes. Ces proportions pour les items testant l'illégalité des doublets après une consonne simple étaient de 66% vs. 16% en CE2, 62% vs. 24% en CM2 et 30% vs. 6% chez les adultes.

## CHAPITRE EXPERIMENTAL II

### **Rôle des connaissances en orthographe lexicale et des connaissances graphotactiques dans l'apprentissage de l'orthographe de nouveaux mots dans une situation de self-teaching chez des élèves français de CM1-CM2**

De nombreuses études ont mis en évidence le rôle clef de la conscience phonologique et de la qualité du décodage phonologique dans l'acquisition de connaissances relatives à l'orthographe de mots spécifiques (e.g., Share, 2008 pour une revue). Dans les expériences examinant l'apprentissage de l'orthographe de non-mots, le plus souvent insérés dans des textes que les sujets doivent lire pour les comprendre (e.g., Share, 1999), les corrélations entre la qualité du décodage phonologique des non-mots et l'apprentissage de leur orthographe sont le plus souvent significatives mais peu élevées, notamment dans des langues avec des correspondances graphèmes-phonèmes peu consistantes comme l'anglais (e.g., Nation et al., 2007). Ceci suggère que d'autres variables jouent un rôle dans cet apprentissage. Parmi ces variables, les connaissances orthographiques sont souvent mentionnées mais demeurent mal définies. Selon Castles et Nation (2006), il convient de distinguer la connaissance de l'orthographe de mots spécifiques (orthographe lexicale) de la connaissance de régularités sur le langage écrit (régularités graphotactiques) (voir aussi Connors et al., 2011 ; Nation et al., 2007).

Dans une étude récente portant sur la lecture de « vrais » mots chez des élèves de 7-9 ans, Connors et al. (2011) ont montré qu'après contrôle des habiletés de décodage phonologique, les connaissances en orthographe lexicale et les connaissances graphotactiques expliquaient des parts de variance indépendantes. Les rares études ayant exploré le rôle de ces deux types de connaissances dans des situations d'apprentissage de l'orthographe de non-mots ne permettent pas d'évaluer leur impact respectif. Ainsi, dans une étude conduite auprès d'élèves de deuxième primaire, Cunningham et al. (2006) ont rapporté, qu'après avoir contrôlé la qualité du décodage phonologique de non-mots, un score composite de connaissances orthographiques reflétant des connaissances en orthographe lexicale (choisir laquelle de deux orthographe comme rain - rane est correcte) et des connaissances

graphotactiques (choisir entre une orthographe de non-mot respectant les régularités graphotactiques de l'anglais et une orthographe de non-mot violant une régularité graphotactique, celle qui ressemble le plus à un mot) expliquait une part de variance significative de l'apprentissage de l'orthographe des non-mots. Malheureusement, l'utilisation d'un score ne distinguant pas les deux types de connaissances orthographiques ne permet pas de déterminer le poids respectif de ces deux types de connaissances orthographiques lors de l'apprentissage de l'orthographe de nouveaux mots.

L'objectif de l'expérience rapportée dans ce chapitre était d'étudier le rôle de ces deux types de connaissances orthographiques sur l'apprentissage de l'orthographe de non-mots dans une situation de *self-teaching* auprès d'élèves français de CM1/CM2, des niveaux scolaires où la mise en place du lexique orthographique a déjà bien débuté. Les élèves devaient lire des textes dans lesquels des non-mots étaient insérés. L'apprentissage de l'orthographe des non-mots était ensuite évalué avec une tâche de rappel (dictée). Les élèves passaient également diverses tâches évaluant leur niveau de conscience phonémique, leurs habiletés à produire des orthographes de non-mots phonologiquement plausibles, leurs connaissances en orthographe lexicale et leurs connaissances graphotactiques. Les connaissances en orthographe lexicale étaient évaluées avec trois épreuves, une dictée de mots isolés, la dictée du corbeau en utilisant le score d'orthographe lexicale (L2MA, Chevrier-Muller et al., 1997) et une épreuve de jugement consistant à indiquer l'orthographe correcte parmi trois orthographes phonologiquement plausibles (e.g., pissenlit – pissenlid – pissenlie). Cette dernière épreuve permettait de comparer la part de variance expliquée par les connaissances en orthographe lexicale et celle expliquée par les connaissances graphotactiques (e.g., *tunnor* – *ttunor* pour évaluer si les élèves savent que les consonnes doubles ne surviennent jamais en début de mots) avec deux tâches de jugement. Ainsi, si ces deux types de connaissances orthographiques, toutes deux évaluées avec une épreuve de jugement, expliquent des parts de variance indépendantes de l'apprentissage de l'orthographe de non-mots, cela suggèrera l'influence de différents types de connaissances, plutôt que l'influence de différentes tâches (ce qui aurait été le cas si les connaissances graphotactiques avaient été évaluées avec une tâche de jugement et les connaissances en orthographe lexicale avec une tâche de dictée).

Notre hypothèse était que les connaissances en orthographe lexicale (évaluées avec des épreuves de production ou de jugement) et les connaissances graphotactiques devraient être corrélées et que ces deux types de connaissances devraient expliquer des parts de variance indépendantes de l'apprentissage de l'orthographe des non-mots, même après divers

contrôles (niveau scolaire, niveau en conscience phonémique, qualité du décodage phonémique, capacité à produire des orthographe de non-mots phonologiquement plausibles).

## **Méthode**

### **Participants.**

Trente-huit élèves de CM1 (20 filles) et vingt-huit élèves de CM2 (17 filles), tous de langue maternelle française, ont participé à cette expérience. L'âge moyen des CM1 était de 9 ans et 11 mois (EC = 4 mois) ; celui des CM2 de 10 ans 8 mois (EC = 4 mois).

### **Matériel**

#### 1. Non-mots et textes utilisés pour l'étude de l'apprentissage de l'orthographe de non-mots.

Douze non-mots bi-syllabiques ont été créés. La rime de la deuxième syllabe était phonologiquement identique mais orthographiquement différente dans deux non-mots : *on / ond* (e.g., loivon ; carmond) ; *and / ant* (e.g., borvand ; loufant) ; *au / eau* (e.g., menlau ; pirdeau) ; *ard / art* (e.g., polgart ; tonvard) ; *ait / ais* (e.g., farnait ; moudais) ; *in / ain* (e.g., dirtain ; vondin). Chaque non-mot était orthographié avec une graphie pour la moitié des sujets (e.g., farnait et moudais) et avec l'autre graphie pour l'autre moitié (e.g., farnais et moudait). Ces 12 non-mots étaient insérés dans six textes (deux par texte) d'une longueur moyenne de 157 mots. Les non-mots correspondaient à des noms communs, par exemple un outil ou un instrument de musique. Les deux non-mots se terminant par la même rime (e.g., farnait et modais) n'apparaissaient jamais au sein d'un même texte. A la fin de chaque texte, quatre questions de compréhension, qui ne portaient jamais sur les non-mots, étaient posées. La première question consistait à choisir, parmi trois titres, le plus approprié. Les suivantes consistaient en trois affirmations pour lesquelles les élèves devaient indiquer si elles étaient vraies ou fausses. Un exemple est fourni dans l'encadré ci-après. L'ordre des histoires et des non-mots insérés dans ces histoires étaient randomisés sur l'ensemble des sujets.

## **Encadré.**

### **Exemple d’histoire et de questions sur ces histoires**

Autrefois, les habitants de la campagne se retrouvaient pour le polgart, la fête du village. C’était l’occasion de chanter, de danser et de déguster les spécialités de la région. Au polgart, la coutume était de s’amuser toute la nuit et d’attendre que le jour se lève pour partager un énorme gâteau au pirdeau. Le pirdeau est un fruit délicieux que l’on trouvait dans les forêts. Lors d’un polgart, une drôle d’histoire est arrivée à Loura, une jeune villageoise curieuse et intrépide qui aimait faire des blagues. La fillette a voulu goûter le gâteau avant la fin du polgart parce qu’elle adorait le goût du pirdeau. Malheureusement, le gâteau se renversa. La pauvre fillette passa le reste de la nuit à chercher la plante qui donne le pirdeau, pour que sa mère prépare un autre gâteau. Finalement, personne ne se rendit compte de rien et le polgart se termina comme prévu. Le pirdeau fit encore beaucoup d’heureux.

Questions :

#### **1. Entoure le titre qui convient le plus**

Chantons et dansons

Attention à la gourmandise

La fête au village

#### **2. Réponds par vrai ou faux**

La fête commence le matin                      Vrai / Faux

Tout le monde adore le gâteau                Vrai / Faux

La fête du village s’est mal terminée        Vrai / Faux

## 2. Connaissance de l’orthographe de mots spécifiques.

La connaissance de l’orthographe de mots spécifiques a été évaluée avec deux dictées et un choix entre plusieurs orthographe. La première dictée est la “dictée du corbeau” (Chevrier-Muller et al., 1997, *cf.*, *Annexe 4*). Cette étude portant sur l’acquisition de l’orthographe lexicale, seul le score d’usage sera utilisé pour les analyses de régression.

La deuxième dictée est une dictée de mots isolés (Martinet & Valdois, 1999) comportant 66 mots monomorphémiques, de 4 à 9 lettres et de complexité phonographémique variable. Vingt-deux mots sont constitués uniquement des correspondances

phono-graphémiques les plus fréquentes (e.g., marmite), 22 autres contiennent un ou plusieurs graphèmes inconsistants non rares (e.g., faucon) et les 22 derniers mots contiennent au moins un graphème rare (e.g., femme). Les trois types de mots sont appariés en longueur (6 lettres en moyenne pour chaque type). La fréquence moyenne des 66 mots dans les manuels scolaires d'école primaire est de 59.4 (score U issu de Manulex, Lété, Sprenger-Charolles & Colé, 2004), chaque type de mots comprenant des items de diverses fréquences (fréquence moyenne par type : 42.6, 45 et 90.7 respectivement pour mots simples, à graphème inconsistant et à graphème rare) (cf., *Annexe 5*).

L'épreuve de jugement comprend 42 triplets de mots écrits. Chaque triplet est formé par un mot correctement orthographié et deux homophones qui ne se distinguent du mot correct que sur un phonème écrit de trois façons différentes (e.g., *pissenlit*, *pissenlid*, *pissenlie*). De plus, les trois graphies d'un triplet (e.g., *it*, *id*, et *ie*) sont utilisées dans deux autres triplets, chacune d'elles étant la graphie exacte dans un triplet. Ainsi *id* est la graphie correcte dans le triplet *nit*, *nid*, *nie* et *ie* la graphie correcte dans le triplet *fourberie*, *fourberit*, *fourberid*. La fréquence moyenne des items est de 4.69 pour 1 million, avec un minimum de 2.94 (quadrille) et un maximum de 62,52 (repas). Ces 42 triplets étaient imprimés dans un ordre aléatoire, différent pour chaque élève, sur une feuille A4, en deux colonnes, police Times new roman, taille 10, après un item essai (mèson, maison, maizon) (cf., *Annexe 6*).

### 3. Connaissances graphotactiques.

Les connaissances graphotactiques ont été évaluées avec les mêmes items que dans l'expérience rapportée au chapitre expérimental I. Douze paires testaient l'identité des consonnes pouvant être doublées (e.g., *buffor* – *bukkor*) ; 16 paires testaient la connaissance du fait que les doublets peuvent survenir en position médiane mais non en position initiale, avec des doublets fréquents pour la moitié des paires (e.g., *tunnor* – *ttunor*, noté *Doublet illégal en début de mots F+*) et des doublets formés à partir de consonnes jamais doublées en français pour l'autre moitié (e.g., *xikkor* – *xxikor*, noté *Doublet illégal en début de mots F0*). Seize paires testaient la connaissance du fait que les doublets peuvent survenir avant, mais non après, une consonne simple avec le groupe consonantique. Le groupe consonantique était placé après une voyelle après laquelle il est fréquent pour huit paires, (e.g., *apprulir* – *aprrulir*, noté *Doublet illégal après une consonne simple F+*) et après une voyelle après laquelle il est rare pour huit autres paires (e.g., *ippraler* – *iprraler*, noté *Doublet illégal après une consonne simple F-*). Le matériel comportait également 38 paires tampons qui ne feront pas l'objet

d'analyse<sup>1</sup>. A l'exception des paires testant la connaissance du fait que seules certaines consonnes peuvent être doublées, qui incluaient deux doublets se prononçant différemment, les deux pseudo-mots de toutes les autres paires se prononçaient de la même façon.

#### 4. Conscience phonologique et habiletés à effectuer des correspondances phonèmes – graphèmes.

Les trois épreuves de Bosse et Valdois (2009) ont été utilisées pour évaluer la conscience phonologique : décomposition en phonèmes (15 items) ; contrepèterie (10 items) omission du premier phonème (20 items) (*cf.*, Annexe 7).

**Segmentation en phonèmes.** Quinze mots étaient présentés oralement aux élèves qui devaient prononcer chaque phonème constituant le mot (e.g., cadeau → /k-a-d-o/). Les mots étaient constitués en moyenne de 4 phonèmes (variant entre 3 et 5), sept mots se terminaient par une syllabe CVC ou CVCC (e.g., *four*, *porte*) ; huit par une syllabe CV (e.g., *cadeau*, *crapaud*).

**Contrepèterie.** Les élèves devaient échanger le premier phonème de dix couples de mots présentés oralement (e.g., banane – ficelle → /fanan/ - /bisèl/). Les réponses étaient toujours deux non-mots. Les mots étaient constitués en moyenne de 4,8 phonèmes (entre 4 et 6), Tous commençaient par une consonne simple.

**Omission du phonème.** Les élèves devaient supprimer le premier son de 20 mots présentés oralement, créant ainsi un non-mot (e.g., *outil* → /ti/; *drapeau* → /rapo/). Sept mots commençaient avec un phonème vocalique correspondant à un graphème complexe dans lequel la suppression de la première lettre conduit à une réponse incorrecte (e.g., *outil*), neuf mots commençaient par un groupe consonantique (e.g., *drapeau*) et quatre par une consonne simple (e.g., *capitaine*).

**Habilité à effectuer des correspondances phonèmes – graphèmes.** Vingt non-mots, 10 bi-syllabiques et 10 tri-syllabiques issus du test BALE (Jacquier-Roux, Lequette, Pouget, Valdois, & Zorman, 2010) ont été dictés pour évaluer les habiletés à effectuer des correspondances phonèmes – graphèmes.

---

<sup>1</sup> Parmi les 38 paires tampons, 10 testaient le fait qu'un doublet ne peut pas survenir après une lettre accentuée (e.g., *abérrol* – *aberrol*), 12 le fait que la probabilité que le phonème /o/ soit transcrit eau varie en fonction de l'environnement gauche de /o/ (e.g., *driveau* – *drivot*) et 16 le fait que la probabilité d'occurrences de certaines consonnes muettes en fin de mots, comme *d* ou *t*, varie en fonction de l'environnement gauche (e.g., *cilunar* – *cilunard*).



## Procédure

Les passations individuelles se déroulaient dans l'ordre suivant.

Dans un premier temps, l'expérimentateur expliquait aux élèves qu'il leur donnerait un livret comprenant des histoires ainsi qu'une série de questions sur celles-ci. Les élèves devaient lire à haute voix une histoire, tourner la page pour répondre aux questions situées sur la page suivante sans revenir en arrière pour lire l'histoire et ainsi de suite jusqu'à la dernière histoire du livret. Leur lecture était enregistrée afin de permettre l'analyse du décodage phonologique des non-mots insérés dans les textes.

Immédiatement après avoir fini de répondre aux questions sur le dernier texte, les élèves effectuaient les trois épreuves évaluant la conscience phonologique dans l'ordre suivant : segmentation en phonèmes, contrepèterie, omission du premier phonème. L'expérimentateur leur proposait des essais pour chaque tâche. Il fournissait un feedback relatif à l'exactitude de la réponse et corrigeait l'élève si nécessaire pendant les items essais. En revanche, aucun feedback n'était fourni pour les items tests.

Une fois terminée la passation des trois épreuves évaluant la conscience phonologique, les 12 non-mots insérés dans les histoires étaient dictés aux élèves. L'expérimentateur leur demandait de les écrire comme ils étaient écrits dans les histoires qu'ils venaient de lire.

La dernière épreuve individuelle, proposée après la dictée des non-mots insérés dans les histoires, était celle évaluant les connaissances graphotactiques. L'expérimentateur expliquait aux élèves que des mots inventés seraient présentés sur l'écran d'ordinateur, orthographiés de deux façons et qu'ils devraient indiquer laquelle des deux orthographes constitue la meilleure orthographe en français. L'expérimentateur expliquait qu'ils devraient appuyer sur la touche Q, sur laquelle était collée une gommette de couleur, s'ils pensaient que c'était l'orthographe de gauche ou sur la touche M, également recouverte d'une gommette de couleur, s'ils pensaient que c'était l'orthographe de droite. Les élèves effectuaient les items essais avec l'expérimentateur puis poursuivaient la tâche seuls, sans recevoir de feedback relatif à l'exactitude de leurs réponses. Les paires de non-mots étaient présentées dans un ordre aléatoire, différent pour chaque élève, avec la contrainte que deux paires testant une même propriété graphotactique n'étaient jamais présentées d'affilée. Pour chaque propriété graphotactique, un type d'orthographe (e.g., incluant un doublet fréquent pour la propriété *identité des lettres pouvant être doublées*) était placé à gauche pour la moitié des items et à droite pour l'autre moitié. Chaque paire de non-mots était présentée au centre d'un écran d'ordinateur portable PC Toshiba 15 pouces à l'aide du logiciel E-Prime. Le logiciel

enregistrait les réponses des élèves.

Les passations collectives avaient lieu lorsque l'ensemble des élèves avaient passé les épreuves individuelles.

Les 66 mots isolés, répartis de façon aléatoire, étaient dictés aux élèves en effectuant trois pauses. Le lendemain, les élèves effectuaient la tâche de jugement comprenant 42 triplets (e.g., *pissenlit*, *pissenlid*, *pissenlie*). L'expérimentateur leur expliquait que des mots étaient orthographiés de trois façons et que leur tâche consistait à entourer l'orthographe correcte. Le même jour, les 20 non-mots issus du test BALE (Jacquier-Roux et al., 2010) leur étaient dictés. L'expérimentateur expliquait aux élèves qu'il allait leur dicter des mots inventés, qu'ils n'avaient jamais vus ni entendus, et qu'ils devaient les écrire comme s'il s'agissait de vrais mots rencontrés pour la première fois dans une dictée.

## **Résultats**

### **1. Sensibilité aux régularités graphotactiques.**

Tableau 1. Pourcentages de sélections correctes pour chacune des trois propriétés de l'emploi des doubles lettres ; les écart-types sont entre parenthèses.

	CM1	CM2
Identité des lettres pouvant être doublées.	87.5 <sup>**</sup> (14.7)	93.8 <sup>**</sup> (9.0)
Doublet illégal en début de mots F+	85.2 <sup>**</sup> (16.9)	87.9 <sup>**</sup> (12.0)
Doublet illégal en début de mots F0	76.0 <sup>**</sup> (20.2)	79.9 <sup>**</sup> (14.2)
Doublet illégal après une consonne simple F+	61.8 <sup>**</sup> (19.3)	66.5 <sup>**</sup> (25.0)
Doublet illégal après une consonne simple F-	51.6 (22.0)	58.0 <sup>*</sup> (17.7)

Note. Scores significativement supérieurs au hasard à  $p < .01$  <sup>\*\*</sup> ;  $p < .05$  <sup>\*</sup>

Concernant la connaissance des lettres qui peuvent être doublées, comme le montre la première ligne du Tableau 1, les taux de sélections des orthographes incluant un doublet légal (plutôt qu'illégal) étaient significativement supérieurs au hasard en CM1 et CM2 ( $t_s > 15.7$ ,  $ps < .001$ ). Des tests de Student avec les sujets et les items comme variables aléatoires (notés respectivement  $t_1$  et  $t_2$ ) révèlent des scores significativement meilleurs en CM2 qu'en CM1 ( $t_1(64) = 1.99$ ,  $p = .05$  ; et  $t_2(11) = 4.3$ ,  $p = .001$ ).

Concernant l'illégalité des doublets en début de mots, comme le montrent les deuxième et troisième lignes du Tableau 1, les taux de sélection de l'orthographe incluant un doublet en position légale (plutôt qu'illégal) étaient supérieurs au hasard en CM1 et CM2 pour les items incluant des doublets fréquents et pour ceux incluant des doublets formés avec des consonnes jamais doublées en français ( $t_s > 7.9$ ,  $ps < .001$ ). Une ANOVA avec les variables Niveau scolaire et Type d'items (*Doublet illégal en début de mots F+* vs. *Doublet illégal en début de mots F0*) sur le nombre de choix corrects révèle un effet du type d'item ( $F_1(1, 64) = 16.4$ ,  $MSE=0.9$   $p < .001$ ,  $\eta_p^2 = 0.20$  ;  $F_2(1, 14) = 7.6$ ,  $MSE=.01$   $p = .016$ ,  $\eta_p^2 = 0.35$ ), avec de meilleurs scores pour les items incluant des doublets fréquents, sans effet du niveau scolaire ( $ps > .17$ ), ni interaction ( $ps > .78$ ).

Concernant l'illégalité des doublets après une consonne simple, les taux de sélections correctes, indiqués aux deux dernières lignes du Tableau 1, sont supérieurs au hasard en CM1 en CM2 pour les items *Doublet illégal après une consonne simple F+* ( $t_s > 3.5$ ,  $ps < .002$ ) mais seulement en CM2 pour les items *Doublet illégal après une consonne simple F-* ( $t(27) = 2.4$ ,  $p = .024$  en CM2 ;  $t(37) = .5$ ,  $p = .65$  en CM1). Une ANOVA sur le nombre de sélections correctes avec les variables Niveau scolaire (CM1, CM2) et Type d'item (*Doublet illégal après une consonne simple F+* vs. *Doublet illégal après une consonne simple F-*) révèle un effet du type d'item ( $F_1(1, 64) = 10.8$ ,  $MSE=1.7$ ,  $p = .002$ ,  $\eta_p^2 = 0.14$  ;  $F_2(1, 14) = 2.4$ ,  $MSE=.01$ ,  $p = .14$ ,  $\eta_p^2 = 0.15$ ), avec de meilleurs scores pour les items *Doublet illégal après une consonne simple F+*, sans effet du niveau scolaire, ni interaction ( $ps > .18$ ).

En résumé, 1) les élèves des deux niveaux scolaires étaient sensibles aux trois propriétés de l'emploi des doubles lettres évaluées dans cette expérience ; 2) ils étaient néanmoins moins sensibles à l'illégalité des doublets après une consonne simple qu'à l'illégalité d'un doublet en début de mots et au fait que seules certaines consonnes peuvent être doublées ; 3) pour les deux propriétés descriptibles sous la forme de règles, leurs

performances variaient en fonction de la familiarité du matériel utilisé pour tester leurs connaissances de ces propriétés graphotactiques.

## 2. Analyses de corrélation et régressions

Plusieurs scores ont été créés pour conduire les analyses de corrélation / régression :

1) Un score d'orthographe lexicale en production obtenu en sommant le score *orthographe usage* de la dictée du corbeau (L2MA, Chevrier-Muller et al., 1997) et le score à la dictée de mots isolés (Martinet & Valdois, 1999) — deux scores qui étaient très corrélés ( $r = .79$ ), même après contrôle du niveau scolaire ( $r = .75$ ).

2) Un score graphotactique reflétant le nombre de sélections correctes pour chacune des trois propriétés de l'emploi des doubles lettres : identité des lettres pouvant être doublées, illégalité des doublets en début de mots et illégalité des doublets après une consonne simple.

3) Un score de conscience phonologique composite reflétant la réussite aux trois épreuves : contrepèterie, omission du phonème initial et décomposition phonémique.

4) Un score de décodage phonologique, correspondant au nombre de non-mots insérés dans les textes correctement lus.

Tableau 2. Scores aux différentes épreuves (sur 100) en CM1 et CM2 ; les écart-types sont entre parenthèses.

Epreuves	CM1	CM2
1) Dictée du corbeau – score phonologique	92.0 (10.6)	93.7 (7.1)
2) Dictée du corbeau – score usage	67.2 (14.4)	78.5 (12.6) **
3) Dictée du corbeau – score grammatical	55.5 (15.9)	72.4 (17.9) **
4) Dictée de mots isolés	72.2 (14.8)	84.3 (10.9) **
5) Orthographe lexicale en production	69.7 (13.6)	81.4 (11.2) **
6) Orthographe lexicale en jugement	62.6 (14.5)	75.9 (12.8) **
7) Connaissances graphotactiques	75.0 (11.5)	80.0 (9.6) <sup>s</sup>
8) Conscience phonologique	66.7 (19.8)	78.9 (16.7) **
9) Dictée de non-mots du test BALE	89.1 (13.4)	92.9 (6.6)
10) Décodage phonologique correct	87.1 (11.5)	90.0 (8.3)
11) Apprentissage de l'orthographe des non-mots	34.9 (11.8)	44.3 (16.0) **

Note. Scores significativement meilleurs en CM2 qu'en CM1 à  $p < .01^{**}$  ;  $p < .05^{*}$  ;  $p < .1^{§}$

Le Tableau 2 indique les scores en orthographe lexicale (lignes 1 à 6), connaissances graphotactiques (ligne 7), conscience phonologique (ligne 8), en dictée de non-mots du test BALE (ligne 9), décodage phonologique (i.e., le pourcentage de non-mots insérés dans les textes correctement décodés, à la ligne 10) et le niveau d'apprentissage de l'orthographe des non-mots (ligne 11). Les scores en orthographe lexicale, évaluée avec des épreuves de jugement ou de production montrent que les élèves des deux niveaux scolaires possédaient d'assez bonnes connaissances relatives à l'orthographe de mots spécifiques. En revanche, leur niveau d'apprentissage de l'orthographe des non-mots était assez faible (34.9% en CM1 et 44.3% en CM2 de non-mots avec la graphie finale correctement orthographiée). Il est à noter que cette difficulté d'apprentissage ne peut pas s'expliquer par des difficultés générales à produire des orthographe phonologiquement plausibles puisque la dictée de non-mots jamais vus ni entendus du test BALE était particulièrement bien réussie (environ 90% d'orthographe phonologiquement plausibles) et la grande majorité des erreurs à la dictée des non-mots insérés dans les histoires étaient phonologiquement plausibles (91% en CM1 et 88% en CM2). Parmi ces erreurs phonologiquement plausibles sur les items incluant une consonne muette finale, il y avait davantage d'orthographe n'incluant pas de consonnes muettes finales (e.g., *ar* au lieu de *ard* ou *art*) que d'orthographe incluant une autre consonne muette (e.g., *ard* au lieu de *art* ou *art* au lieu de *ard*) (61% vs. 39% en CM1 et 59% vs. 41% en CM2)

Des tests de *Student* révèlent des scores significativement meilleurs en CM2 qu'en CM1 pour les scores usage et grammatical de la dictée du corbeau, la dictée de mots isolés, l'orthographe lexicale en production (score reflétant le score aux deux dictées), l'orthographe lexicale en jugement, la conscience phonologique et l'apprentissage de l'orthographe des non-mots ( $ts > 2.63$ ,  $ps < .011$ ). Pour les connaissances graphotactiques, la supériorité des CM2 est marginalement significative ( $p = .064$ ). Les scores ne différaient pas en fonction du niveau scolaire pour le score phonologique de la dictée du corbeau ( $p = .45$ ), la dictée de non-mots de la BALE ( $p = .18$ ) et le décodage phonologique ( $p = .25$ ).

Le Tableau 3 fournit les corrélations entre les différents scores. L'apprentissage de l'orthographe des non-mots corrèle avec les différentes mesures de connaissances en orthographe lexicale ( $rs > .44$ ,  $ps < .01$ ), même après contrôle du niveau scolaire ( $rs > .32$ ,  $ps < .01$ ). Il corrèle également avec les connaissances graphotactiques ( $r = .40$ ,  $p < .01$ ), même après contrôle du niveau scolaire ( $r = .35$ ,  $p = .01$ ). En revanche, il ne corrèle pas

significativement avec la conscience phonologique ( $r = .19$ ), le décodage phonologique ( $r = .11$ ) et la dictée de non-mots de la BALE ( $r = .11$ ). Les différents scores en orthographe lexicale (score d'usage à la dictée du corbeau, dictée de mots isolés, orthographe lexicale en jugement) sont très corrélés ( $rs > .71$ ,  $ps < .01$ ), même une fois contrôlé le niveau scolaire ( $rs > .65$ ,  $ps < .01$ ). Les corrélations entre ces différents scores et les connaissances graphotactiques étaient plus faibles ( $rs$  comprises entre  $.28$  et  $.36$ ,  $ps < .05$  sans contrôle du niveau scolaire et  $rs$  comprises entre  $.21$  et  $.30$ ,  $ps < .1$ ).

Afin de tester si les connaissances en orthographe lexicale et les connaissances graphotactiques rendent compte de parts de variance indépendantes dans l'apprentissage de l'orthographe des non-mots insérés dans les textes, deux séries d'analyses de régression ont été conduites, la première en rentrant les connaissances en orthographe lexicale après les connaissances graphotactiques, la seconde en rentrant les connaissances graphotactiques après les connaissances en orthographe lexicale. Ces deux analyses ont été conduites en utilisant soit le score d'orthographe lexicale en production, soit le score d'orthographe lexicale en jugement pour contrôler que, si les connaissances en orthographe lexicale et les connaissances graphotactiques expliquent des parts de variance indépendantes, cela reflète l'impact de deux types de connaissances et non l'effet de deux tâches différentes (i.e., production pour la connaissance de l'orthographe de mots spécifiques et jugement pour les connaissances graphotactiques). Ces deux types de connaissances orthographiques ont toujours été rentrés après le niveau scolaire et les scores en conscience phonologique, décodage phonologique et dictée de non-mots du test BALE.

Comme l'indique le Tableau 4, après contrôle de ces quatre variables, qui expliquent 12.0% de la variance, tous les scores de connaissances orthographiques ajoutent un pourcentage de variance significatif : 12% pour l'orthographe lexicale en production, 9% pour l'orthographe lexicale en jugement et 10% pour les connaissances graphotactiques. Après contrôle des connaissances graphotactiques, la connaissance de l'orthographe de mots spécifiques ajoute 8% de variance quand elle est évaluée avec des épreuves de production et 6% de variance quand elle est évaluée avec l'épreuve de jugement. Les connaissances graphotactiques ajoutent 6% de variance après contrôle des connaissances en orthographe lexicale évaluées avec les épreuves de production et 7% de variance après contrôle des connaissances en orthographe lexicale évaluées avec l'épreuve de jugement.

Tableau 3.

Corrélations entre les différents scores ; les corrélations partialisées sur le niveau scolaire sont indiquées en italique en-dessous de la diagonale.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Connaissances graphotactiques	—	.36**	.25*	.28*	.32**	.34**	.18	.13	.40**
(2) CORBEAU – score d’usage	.30**	—	.19	.79**	.71**	.95**	.42**	.43**	.40**
(3) Dictée de non-mots (BALE)	.22 <sup>\$</sup>	.13	—	.18	.25*	.20	.33**	.02	.11
(4) Dictée de mots isolés	.21 <sup>\$</sup>	<i>0.75**</i>	.12	—	.82**	.94**	.47**	.41**	.44**
(5) Orthographe lexicale en jugement	.25*	.65**	.20 <sup>\$</sup>	.77**	—	.80**	.44**	.44**	.42**
(6) Orthographe lexicale en production	.27*	.94**	.14	.93**	.76**	—	.48**	.44**	.45**
(7) Conscience phonologique	.12	.35**	.30**	.40**	.36**	.40**	—	.16	.19
(8) Décodage phonologique correct	.10	.41**	.00	.39**	.42**	.43**	.12	—	.11
(9) Apprentissage orthographique	.35**	.32**	.06	.35**	.32**	.36**	.09	.06	—

Note. Corrélations significatives à  $p < .01$  \*\* ;  $p < .05$  \* ;  $p < .1$  <sup>\$</sup>

Tableau 4.

Analyses de régression cherchant à prédire le niveau d'apprentissage de l'orthographe des non-mots insérés dans les textes.

Etapes	Variables rentrées	R	R <sup>2</sup>	Changement du R <sup>2</sup>	significativité
1 à 4	Niveau scolaire, conscience phonologique, décodage phonologique et dictée de non-mots de la BALE	.34	.12	–	.001
5	Orthographe lexicale en jugement	.46	.21	.09**	.01
6	Connaissances graphotactiques	.53	.28	.07*	.02
5	Orthographe lexicale en production	.49	.24	.12**	.004
6	Connaissances graphotactiques	.55	.30	.06*	.03
5	Connaissances graphotactiques	.47	.22	.10**	.01
6	Orthographe lexicale en production	.55	.30	.08**	.01
6	Orthographe lexicale en jugement	.53	.28	.06*	.03

## **Discussion**

L'objectif de cette expérience était d'étudier si et comment les connaissances en orthographe lexicale (i.e., relatives à l'orthographe de mots spécifiques) et les connaissances graphotactiques expliquaient l'apprentissage de l'orthographe de nouveaux mots chez des élèves français de CM1 / CM2. Comme dans de nombreuses études, les élèves lisaient des histoires dans lesquelles des non-mots étaient insérés (e.g., Cunningham et al., 2006 ; Nation et al., 2007) puis leur apprentissage de l'orthographe de ces non-mots était évalué, ici avec une dictée.

L'apprentissage de l'orthographe des non-mots était difficile : seulement 34.9% en CM1 et 44.3% en CM2 des non-mots étaient correctement orthographiés. La constitution du



matériel explique certainement aussi les difficultés d'apprentissage observées : non seulement sept des douze orthographes incluait une lettre muette finale (*ond, and, ant, ard, art, ais, ait*), mais les six sons cibles étaient tous orthographiés de deux façons différentes dans deux non-mots. Cette situation crée des interférences mais elle constitue un bon test de l'apprentissage de l'orthographe de mots spécifiques, les élèves devant par exemple mémoriser qu'un non-mot incluait *and* et qu'un autre incluait *ant*.

Les élèves étaient sensibles aux diverses régularités graphotactiques testées et leurs résultats étaient très similaires à ceux d'études antérieures (e.g., Danjon & Pacton, 2009 ; chapitre expérimental I). En particulier, certaines régularités étaient moins bien repérées que d'autres et, pour les propriétés descriptibles sous la forme de règle, les performances variaient en fonction de la familiarité du matériel utilisé pour évaluer les connaissances graphotactiques, ce qui indique que les élèves ne recouraient pas (systématiquement) à des règles spécifiant dans quelle position un doublet peut survenir.

Les scores aux différentes épreuves évaluant les connaissances en orthographe lexicale étaient fortement corrélés. Ils l'étaient moins avec le score en connaissances graphotactiques, ce qui suggère qu'il s'agit bien là de deux facettes des connaissances orthographiques (Castles & Nation, 2006). L'idée de deux facettes des connaissances orthographiques, et non de différents modes d'évaluation des connaissances orthographiques, est étayée par le fait que les connaissances en orthographe lexicale évaluées avec une épreuve de jugement étaient davantage corrélées aux connaissances en orthographe lexicale évaluées avec des épreuves de production qu'avec les connaissances graphotactiques évaluées avec une épreuve de jugement.

Ces deux types de connaissances orthographiques, spécifiques et générales, étaient toutes deux corrélées au niveau d'apprentissage de l'orthographe des non-mots. De plus, les analyses de régression indiquent que ces deux facettes des connaissances orthographiques expliquaient des parts de variance indépendantes, ce qui est en accord avec l'étude de Connors et al. (2011) explorant le rôle de ces deux types de connaissances dans une tâche de lecture de 'vrais' mots. Ces résultats confirment ceux de Cunningham et al. (2006) montrant une influence des connaissances orthographiques sur l'apprentissage de l'orthographe de non-mots dans une situation d'apprentissage similaire à la nôtre une fois contrôlées les habiletés de décodage phonologique. De plus, ils les prolongent en montrant l'influence de deux types de connaissances orthographiques, spécifiques et générales, que Cunningham et al. n'avaient pas distinguées. Nos résultats sont également en accord avec Nation et al. (2007) qui évoquaient une possible influence des connaissances graphotactiques pour expliquer que, dans leur étude, la relation entre le niveau de décodage phonologique et le niveau

d'apprentissage de l'orthographe de nouveaux mots, observée avec des analyses sur les scores moyens, ne demeurait pas lorsque des analyses item par item étaient conduites.

Contrairement aux études mentionnées dans le paragraphe précédent, conduites en anglais, nous n'avons pas observé de corrélation significative entre le niveau de décodage phonologique et le niveau d'apprentissage orthographique. Les analyses rapportées ont été conduites avec le nombre total de décodages corrects (i.e., 12 non-mots présentés chacun cinq fois) mais le pattern de résultats était identique en considérant la seule première présentation de chaque non-mot. Aux deux niveaux scolaires, environ 90% des non-mots étaient décodés correctement. Un manque de variabilité des scores en décodage phonologique ne semble toutefois pas pouvoir expliquer l'absence de lien entre le décodage phonologique et l'apprentissage de l'orthographe des non-mots puisque les scores en décodage phonologique étaient corrélés avec les scores aux épreuves évaluant la connaissance de l'orthographe de 'vrais mots'. Des études devront être conduites pour explorer si une corrélation entre décodage phonologique et apprentissage de l'orthographe des non-mots est observée chez des élèves français plus jeunes dont les habiletés de décodage phonologique varient probablement davantage. Dans cette perspective, le nombre de non-mots à apprendre devrait probablement être restreint étant donné le niveau d'apprentissage de l'orthographe des non-mots déjà assez faible en CM1 / CM2.

Enfin, la dictée de non-mots (des items jamais vus ni entendus) du test BALE nous a permis de contrôler que les scores à la dictée des non-mots vus dans les histoires ne reflétaient pas seulement une capacité à produire des orthographe phonologiquement plausibles mais aussi et surtout l'apprentissage d'informations relatives à l'orthographe de mots spécifiques. A notre connaissance, ce contrôle n'a pas été effectué dans des études antérieures. Un autre argument en faveur d'un apprentissage de l'orthographe de mots spécifiques provient de la comparaison des corrélations entre les scores aux deux dictées de non-mots et les différentes mesures de connaissances orthographiques. En effet, ces différentes mesures étaient toutes plus corrélées avec le niveau d'apprentissage de l'orthographe des non-mots insérés dans les histoires qu'avec les scores à la dictée de non-mots du test BALE (.40 vs. .25 pour les connaissances graphotactiques ; .45 vs. .20 pour la connaissance de l'orthographe de mots spécifiques évaluée en production et .42 vs. .25 pour la connaissance de l'orthographe de mots spécifiques évaluée en jugement).

En conclusion, nous n'avons pas observé de corrélation entre la qualité du décodage phonologique, considérée comme cruciale dans la théorie du *self-teaching* de Share (1995), et l'apprentissage de l'orthographe de nouveaux mots. En revanche, nous avons montré que

deux types de connaissances orthographiques, générales et spécifiques, expliquaient des parts de variance indépendantes de l'apprentissage de l'orthographe de nouveaux mots. Dans les deux prochains chapitres expérimentaux, nous poursuivons notre investigation du rôle des connaissances en orthographe lexicale et des connaissances graphotactiques dans l'apprentissage de l'orthographe de non-mots en manipulant le degré de convergence entre et les régularités graphotactiques du français et les orthographe des non-mots.

## CHAPITRE EXPERIMENTAL III

### **Influence des connaissances graphotactiques de type probabiliste sur l'apprentissage de l'orthographe de nouveaux mots : Etudes chez l'adulte et l'enfant de l'école élémentaire.**

Fayol et al. (2010) ont étudié l'influence de la sensibilité d'étudiants français au fait que certaines consonnes sont plus fréquemment doublées que d'autres sur leur apprentissage de l'orthographe de nouveaux mots. Dans leur étude, les sujets devaient lire des histoires dans lesquelles étaient insérés à cinq reprises un non-mot qui correspondait au nom du héros de chaque histoire. Sur l'ensemble des six histoires, les étudiants étaient exposés à deux non-mots orthographiés sans double lettre (items *sans doublet*, e.g., *tidunar*), deux non-mots orthographiés avec un doublet fréquent (items *doublet fréquent*, e.g., *tidunnar*) et deux non-mots orthographiés avec un doublet rare (items *doublet rare*, e.g., *tiddunar*). Les étudiants lisaient les textes puis devaient répondre à des questions de compréhension dont l'une nécessitait l'écriture du non-mot. Les étudiants, qui n'avaient aucune instruction d'orthographier les non-mots tels qu'ils l'étaient dans les histoires, avaient plus de difficultés à rappeler des orthographe contenant des doublets rares que des orthographe contenant des doublets fréquents ou ne contenant pas de doublets, même s'ils avaient vu aussi souvent ces trois types d'items. Ils avaient acquis des informations spécifiques sur l'orthographe des non-mots, notamment concernant la présence d'un doublet puisque, par exemple, ils utilisaient davantage un doublet fréquent pour écrire des items présentés avec un doublet fréquent que pour écrire des items présentés sans doublet ou avec un doublet rare. Toutefois, l'analyse des erreurs indiquait qu'environ un tiers des erreurs sur les doublets rares étaient des transpositions (e.g., *tiddunar* orthographié TIDUNNAR). La prévalence des erreurs de transposition sur les items *doublet rare*, leur rareté sur les items *doublet fréquent*, et la rare utilisation des doublets fréquents pour les items *sans doublet* suggèrent que, parfois, les participants se rappelaient de la présence d'un doublet dans l'item mais ne se rappelaient pas de la lettre spécifique qui devait être doublée. Dans ce cas, ils reconstruisaient une orthographe en se fondant sur leur connaissance des lettres les plus susceptibles d'être doublées en français et de la position dans laquelle ces lettres sont susceptibles d'être

doublées. Cette reconstruction conduisait à une production orthographique correcte pour les items *doublet fréquent* mais à des erreurs de transpositions pour les items *doublet rare*.

L'étude de Fayol et al. (2010) présente néanmoins certaines limites. Une première est liée à la lecture silencieuse des textes. Même s'il s'agit d'une activité particulièrement écologique, plus fréquente que la lecture à voix haute, un inconvénient d'un point de vue expérimental est qu'elle ne fournit aucune preuve directe que les sujets prononçaient de la même façon les orthographes incluant des consonnes simples et celles incluant des consonnes doubles. C'est pourquoi nous utiliserons des situations de lecture à voix haute en plus des situations de lecture silencieuse dans ce chapitre expérimental.

Une seconde limite de l'étude de Fayol et al. (2010) est qu'à travers leur lecture des textes successifs, les sujets pourraient avoir remarqué que pour chaque histoire, une des questions requiert la production orthographique du non-mot. Par conséquent, ils pourraient avoir alloué plus d'attention au non-mot au cours de l'expérience. Cette stratégie aurait pu être renforcée par le fait que le non-mot correspondait toujours au nom du héros. Si les sujets utilisaient une telle stratégie, l'apprentissage de l'orthographe des non-mots ne pourrait plus être considéré comme implicite. Un argument allant à l'encontre de cette possibilité est que le pourcentage d'orthographe correctes des non-mots était relativement stable de la première à la sixième histoire (87.0%, 83.3%, 83.3%, 83.6%, 87.0%, et 88.7%, respectivement). Ainsi, si les participants avaient adopté une stratégie d'apprentissage explicite, intentionnel, celle-ci n'engendrait pas une amélioration de l'apprentissage orthographique. Dans les expériences rapportées dans ce chapitre, nous avons utilisé une procédure avec laquelle il est très improbable que les sujets portent de plus en plus leur attention sur les non-mots au fur et à mesure qu'ils lisent de nouvelles histoires. Les mêmes non-mots étaient insérés dans de nouvelles histoires dans lesquelles ils ne correspondaient plus au nom du héros. Une fois que les sujets avaient fini de lire une histoire, ils devaient répondre à quelques questions sur celle-ci, comme dans l'étude de Fayol et al. Toutefois, contrairement à ce qui était fait dans l'étude de Fayol et al., aucune des questions ne portait sur le non-mot. L'apprentissage de l'orthographe des non-mots était évalué à la fin de l'expérience, une fois l'ensemble des histoires lues.

Enfin, dans l'étude de Fayol et al. (2010), il n'était pas explicitement demandé aux sujets d'écrire les mots tels qu'ils étaient orthographiés dans les histoires. Cette

situation, où un individu rencontre un mot nouveau dans un texte sans intention d'apprendre comment il s'orthographe, et l'écrit ultérieurement sans chercher à se rappeler comment ce mot était orthographié dans les textes où il l'a rencontré est particulièrement fréquente. Nous pouvons néanmoins nous demander si les résultats de Fayol et al. (2010), notamment ceux relatifs aux erreurs de transposition, auraient également été observés si les participants avaient eu comme consigne d'orthographier les non-mots comme ils l'étaient dans les histoires qu'ils venaient de lire. Dans les expériences rapportées dans ce chapitre, il était explicitement demandé aux sujets d'écrire les non-mots tels qu'ils étaient écrits dans les histoires.

L'influence des connaissances graphotactiques sur l'apprentissage de l'orthographe de nouveaux mots a été étudiée a) avec des non-mots insérés dans des histoires que des sujets adultes lisaient à voix haute (Expériences 1 et 2) ; b) avec des non-mots présentés de façon isolée que des sujets adultes lisaient silencieusement ou à voix haute (Expérience 3) ; c) avec des non-mots insérés dans des histoires que des élèves de CE2 et CM2 lisaient silencieusement (Expérience 4) ou à voix haute (Expérience 5).

Nous faisons l'hypothèse que, comme dans l'étude de Fayol et al. (2010), les orthographe incluant des doublets rares seraient plus difficilement apprises que celles incluant des doublets fréquents ou n'incluant pas de doublets et que les erreurs de transposition du doublement d'une consonne rarement doublée vers une consonne fréquemment doublée seraient beaucoup plus nombreuses que les erreurs de transposition du doublement d'une consonne fréquemment doublée vers une consonne rarement doublée. Nous faisons cette hypothèse même si, contrairement à Fayol et al. (2010), nous demandions explicitement aux sujets d'orthographier les non-mots tels qu'ils l'étaient dans les histoires ou tels qu'ils apparaissaient à l'écran pour l'expérience 3 dans laquelle les non-mots étaient présentés de façon isolée. Toutefois, du fait du délai entre la lecture d'un non-mot et l'évaluation de l'apprentissage de son orthographe, plus long dans les expériences de ce chapitre expérimental que dans celle de Fayol et al. (2010), les erreurs d'omission devraient être plus nombreuses pour les items requérant un doublet fréquent ou un doublet rare, sans différence entre ces deux types d'items. De ce fait, alors que les items *sans doublet* et les items *doublet fréquent* étaient aussi souvent orthographiés correctement dans l'étude de Fayol et al. (2010), les items *sans doublet* devraient être mieux orthographiés que les items *doublet fréquent* dans les études

rapportées dans ce chapitre. Ce pattern de réussite et d'erreurs devrait être observé chez les adultes et chez les élèves de primaire, même si les erreurs d'omission devraient être encore plus fréquentes chez les élèves que chez les adultes (et encore plus en CE2 qu'en CM2), dans les différentes conditions d'étude (lecture silencieuse ou à haute voix ; mots insérés dans des textes ou présentés de façon isolée) et de test (rappel vs. reconnaissance).

## **Expérience 1**

### **Lecture à haute voix de textes incluant des non-mots chez l'adulte**

Dans cette expérience, les non-mots de l'étude de Fayol et al. (2010) ont été réutilisés mais les changements de procédure suivants ont été effectués : 1) les non-mots étaient insérés dans de nouvelles histoires dans lesquelles ils ne correspondaient plus au nom du héros ; 2) les sujets lisaient les histoires à haute voix ; 3) l'évaluation de l'apprentissage orthographique des six non-mots avait lieu une fois terminée la lecture de l'ensemble des histoires ; 4) les non-mots étaient dictés avec l'instruction de les orthographier comme ils l'étaient dans les histoires ; 5) un groupe contrôle de sujets devait écrire les mêmes non-mots que le groupe expérimental alors qu'ils ne les avaient ni entendus ni lus au préalable. L'inclusion de ce groupe permettait de déterminer si des étudiants français doublent spontanément certaines lettres, et si oui, lesquelles. Cela permettait également de mieux examiner l'apprentissage d'information relative à l'apprentissage de l'orthographe d'un mot spécifique, par exemple en explorant si les sujets qui ont lu un item incluant un doublet fréquent l'écrivent plus souvent avec un doublet fréquent mais moins souvent sans doublet que les sujets du groupe contrôle.

### **Méthode**

#### **Participants**

Quarante-huit étudiants (34 femmes) de l'Université Paris Descartes, tous de langue maternelle française, ont participé à cette étude. L'âge moyen était de 22 ans et 0 mois, tous compris entre 18 et 27 ans. La moitié des participants était assigné au groupe expérimental qui lisait les histoires incluant les non-mots qu'on leur dictait ultérieurement. L'autre moitié, le groupe contrôle, effectuait la même dictée de non-mots sans avoir lu ou entendu les non-mots au préalable.

#### **Matériel**

Les six non-mots trisyllabiques utilisés par Fayol et al. (2010) ont été réutilisés dans cette expérience (*cf.*, *Annexe 8*). Chaque non-mot incluait deux consonnes cibles, qui



correspondaient aux attaques de la seconde et de la troisième syllabe. Une des consonnes cibles de chaque mot était *n*, *t*, ou *r* (trois consonnes fréquemment doublées en français) et l'autre était *b*, *d*, ou *g* (trois consonnes rarement doublées en français). Chaque consonne cible était l'attaque de la deuxième syllabe dans un non-mot et l'attaque de la troisième syllabe dans un autre non-mot. Pour chaque non-mot, trois orthographes ont été créées : sans doublet, avec un doublet fréquent et avec un doublet rare, tous se prononçant de manière identique selon les règles de correspondances graphèmes-phonèmes du français, comme dans *bagotin* (item *sans doublet*), *bagottin* (item *doublet fréquent*), et *baggotin* (item *doublet rare*).

### **Encadré.**

#### **Exemple d'histoire et de questions sur ces histoires**

Autrefois, les habitants de la campagne se retrouvaient pour le tidunnar, la fête du village. C'était l'occasion de chanter, de danser et de déguster les spécialités de la région. Au tidunnar, la coutume était de s'amuser toute la nuit et d'attendre que le jour se lève pour partager un énorme gâteau au bagotin. Le bagotin est un fruit délicieux que l'on trouvait dans les forêts. Lors d'un tidunnar, une drôle d'histoire est arrivée à Loura, une jeune villageoise curieuse et intrépide qui aimait faire des blagues. La fillette a voulu goûter le gâteau avant la fin du tidunnar parce qu'elle adorait le goût du bagotin. Malheureusement, le gâteau se renversa. La pauvre fillette passa le reste de la nuit à chercher la plante qui donne le bagotin, pour que sa mère prépare un autre gâteau. Finalement, personne ne se rendit compte de rien et le tidunnar se termina comme prévu. Le bagotin fit encore beaucoup d'heureux.

Questions :

#### **1. Entoure le titre qui convient le plus**

Chantons et dansons

Attention à la gourmandise

La fête au village

#### **2. Réponds par vrai ou faux**

La fête commence le matin                      Vrai / Faux

Tout le monde adore le gâteau                Vrai / Faux

La fête du village s'est mal terminée        Vrai / Faux

Trois histoires, d'une longueur moyenne de 157 mots, ont été créées. Un exemple est fourni dans l'encadré ci-dessus. Deux non-mots étaient inclus dans chaque histoire, avec la

contrainte que deux non-mots d'un même type (i.e., item *sans doublet*, item *doublet fréquent* ou item *doublet rare*) n'étaient pas inclus dans la même histoire. Les non-mots étaient tous utilisés comme noms, par exemple, le nom d'un type de fruit. Chaque non-mot survenait cinq fois dans chaque histoire. Quatre questions étaient créées pour chacune des histoires. La première consistait à choisir le titre approprié parmi une liste de trois titres. Les trois suivantes étaient des questions en vrai/faux. Ces questions étaient très simples pour des étudiants à l'université. L'ordre des histoires et des non-mots insérés dans ces histoires étaient randomisés sur l'ensemble des sujets.

### **Procédure**

Les participants du groupe expérimental étaient testés individuellement. L'expérimentateur leur disait qu'il avait besoin d'aide pour calibrer des textes qui seraient utilisés pour évaluer les capacités de compréhension d'élèves de l'école élémentaire. L'expérimentateur leur disait qu'il leur donnerait un livret dans lequel se trouvent des histoires ainsi qu'une série de questions sur celles-ci. Les participants devaient lire à haute voix une histoire, tourner la page pour répondre aux questions situées sur la page suivante sans revenir en arrière pour lire l'histoire et ainsi de suite jusqu'à la dernière histoire du livret. La lecture des participants était enregistrée. Ensuite, les participants effectuaient une tâche de barrage pendant une dizaine de minutes. Enfin, l'expérimentateur prononçait chacun des six non-mots et demandait aux participants de les écrire comme ils étaient écrits dans les textes qu'ils venaient de lire. Les participants du groupe contrôle, qui étaient également testés individuellement, effectuaient seulement la dictée finale. Il leur était demandé d'orthographier les non-mots comme s'il s'agissait de mots réels qu'ils rencontreraient pour la première fois.

### **Résultats**

#### **Groupe expérimental**

Lorsqu'ils lisaient les histoires, les participants prononçaient toujours les consonnes cibles et les voyelles qui les entourent de la même façon qu'elles soient simples ou doubles. Ils réussissaient parfaitement aux questions évaluant leur compréhension des histoires.

Les différents types d'orthographes produites par les sujets du groupe expérimental pour chacun des trois types d'items et par les sujets de groupe contrôle sont indiqués au

Tableau 1. Le nombre d'orthographe correctes, définies comme celles dans lesquelles les deux consonnes cibles sont orthographiées comme dans les histoires, étaient plus nombreuses pour les items *sans doublet* que pour les items *doublet fréquent* et elles étaient plus nombreuses pour ces deux types d'items que pour les items *doublet rare*. Une ANOVA sur le nombre d'orthographe correctes avec la variable type d'item (*sans doublet*, *doublet fréquent*, ou *doublet rare*) confirme cet effet ( $F_1(2, 46) = 8.64$ ,  $MSE = .59$ ,  $p = .001$ ,  $\eta_p^2 = .27$ ;  $F_2(2, 10) = 10.61$ ,  $MSE = 20.39$ ,  $p = .003$ ,  $\eta_p^2 = .68$ ). Les comparaisons planifiées montrent des différences significatives entre les items *sans doublet* et les items requérant un doublet ( $F_1(1, 23) = 13.75$ ,  $MSE = .49$ ,  $p = .001$ ,  $\eta_p^2 = .37$ ;  $F_2(1, 5) = 17.10$ ,  $MSE = 1.56$ ,  $p = .009$ ,  $\eta_p^2 = .77$ ) et entre les items *doublet fréquent* et les items *doublet rare* ( $F_1(1, 23) = 5.07$ ,  $MSE = .69$ ,  $p = .03$ ,  $\eta_p^2 = .18$ ;  $F_2(1, 5) = 6.17$ ,  $MSE = 2.28$ ,  $p = .043$ ,  $\eta_p^2 = .55$ ).

Les erreurs de transposition sur les items *doublet rare* (déplacement du doublement d'une consonne rarement doublée en français vers une consonne fréquemment doublée en français, e.g., *tiddunar* orthographié TIDUNNAR) étaient beaucoup plus nombreuses que les erreurs de transposition sur les items *doublet fréquent* (déplacement du doublement d'une consonne fréquemment doublée vers une consonne rarement doublée, e.g., *tidunnar* orthographié TIDDUNAR). Des tests de *Student* avec les sujets ( $t_1$ ) et les items ( $t_2$ ) comme variables aléatoires indiquent que cette différence est significative ( $t_1(23) = 2.69$ ,  $p = .013$ ;  $t_2(5) = 3.05$ ,  $p = .028$ ). La prévalence des erreurs de transposition sur les items *doublet rare* apparaît également dans l'analyse des profils individuels, comme le montre le Tableau 2. Quinze participants sur 24 commettaient des erreurs de transposition. Ceux commettant des erreurs de transposition seulement sur les items *doublet rare* étaient beaucoup plus nombreux que ceux commettant des erreurs de transposition seulement sur les items *doublet fréquent* (12 vs. 3).

La prévalence des erreurs de transposition sur les items *doublet rare* ne reflétait pas une tendance générale à utiliser des doublets fréquents indépendamment du type d'item qui a été présenté, parce que les items *doublet rare* étaient orthographiés avec un doublet fréquent (29.2% d'erreurs de transposition) à la fois moins souvent que les items *doublet fréquent* (66.7% correct) et plus souvent que les items *sans doublet* (8.3% d'erreurs d'addition, définies comme l'utilisation d'un doublet pour un item *sans doublet*, e.g., *tidunar* orthographié TIDUNNAR ou TIDDUNAR). Une ANOVA sur l'utilisation d'un doublet fréquent avec la variable type d'item (*sans doublet*, *doublet fréquent*, et *doublet rare*) révèle un effet principal du type d'item ( $F_1(2, 46) = 26.51$ ,  $MSE = .32$ ,  $p < .001$ ,  $\eta_p^2 = .54$  ;

$F_2(2, 10) = 17.16$ ,  $MSE = 1.96$ ,  $p = .001$ ,  $\eta_p^2 = .77$ ), avec davantage de doublets fréquents utilisés pour les items *doublet fréquent* que pour les items *sans doublet* et *doublet rare* ( $F_1(1, 23) = 39.12$ ,  $MSE = .38$ ,  $p < .001$ ,  $\eta_p^2 = .63$  ;  $F_2(1, 5) = 49.91$ ,  $MSE = 1.18$ ,  $p = .001$ ,  $\eta_p^2 = .91$ ) et davantage de doublets fréquents utilisés pour les *doublet rare* que pour les items *sans doublet* lorsque les participants étaient utilisés comme variable aléatoire ( $F_1(1, 23) = 8.10$ ,  $MSE = 0.26$ ,  $p = .009$ ,  $\eta_p^2 = .26$ ;  $F_2(1, 5) = 3.05$ ,  $MSE = 2.73$ ,  $p = .14$ ,  $\eta_p^2 = .38$ ).

Tableau 1

Pourcentages des différents types d'orthographe produites dans l'Expérience 1. Les orthographe correctes apparaissent en gras, les transpositions en italique, et les écart-types sont entre parenthèses.

Orthographe produite	Groupe contrôle	Type d'item présenté au groupe expérimental		
		sans doublet	doublet fréquent	doublet rare
Orthographe sans doublet	86.1 (19.4)	<b>85.4 (27.5)</b>	27.1 (38.9)	29.2 (29.2)
Orthographe avec un doublet fréquent	12.5 (17.9)	8.3 (19.0)	<b>66.7 (45.8)</b>	29.2 (32.7)
Orthographe avec un doublet rare	1.4 (4.7)	6.3 (16.89)	6.3 (1.4)	<b>39.6 (32.9)</b>

Note. La somme des différentes orthographe produites diffère de 100% dans la condition *doublet rare* car une erreur dans cette condition, 2.1%, impliquait le doublement des deux consonnes.

Tableau 2

Nombre de participants commettant des erreurs de transposition dans chaque expérience

Expérience et condition	N total	N avec des erreurs de transposition		N avec des erreurs de transposition uniquement sur les items <i>doublet fréquent</i>	
		erreurs de transposition	sur les items <i>doublet rare</i>	transposition uniquement	les items <i>doublet fréquent</i>
Expérience 1	24	15	12	3	3
Expérience 2, condition rappel	36	20	16	3	3
Expérience 2, condition reconnaissance	36	22	18	3	3
Expérience 3, lecture à haute voix	24	14	9	3	3
Expérience 3, lecture silencieuse	24	19	15	2	2

Note. Les cas où le nombre de participants commettant uniquement des erreurs de transposition sur les items *doublet fréquent* plus le nombre de participants commettant uniquement des erreurs de transposition sur les items *doublet rare* diffère du nombre total sont des cas de participants commettant des erreurs de transposition sur les deux types d'items.

Comme le montre le Tableau 1, les participants de l'expérience 1 commettaient beaucoup plus d'erreurs d'omission (production d'une orthographe sans doublet pour des items vus avec un doublet fréquent ou rare) que ceux de l'expérience de Fayol et al. (2010). Ce résultat n'est guère surprenant compte tenu du délai entre la lecture d'un non-mot et sa production orthographique (plus de 10 minutes dans l'Expérience 1 comparé à moins d'une minute dans l'expérience de Fayol et al., 2010). Il n'y avait pas de différence significative entre le nombre d'erreurs d'omission pour les items *doublets fréquents* et pour les items *doublets rares* ( $t_1(23) = .27, p = .79$ ;  $t_2(5) = .67, p = .53$ ).

Aucune erreur ne correspondait au doublement d'une voyelle ou d'une consonne jamais doublée en français et aucune erreur ne correspondait au doublement d'une consonne en position initiale ou finale.

### Groupe contrôle

Comme le montre le Tableau 1, la plupart des orthographes produites par le groupe contrôle n'incluait pas de doublet. Quand des doublets étaient utilisés, il s'agissait beaucoup plus souvent d'un doublet fréquent que d'un doublet rare. Les participants du groupe contrôle ne doublaient jamais des lettres qui ne peuvent pas être doublées en français et n'utilisaient jamais des doublets légaux dans des positions illégales (i.e., début et fin de mots).

### Comparaisons entre les groupes expérimentaux et contrôle

Une série de tests de *Student* a été conduite pour comparer les orthographes produites par le groupe expérimental et par le groupe contrôle pour chaque type d'item. Les deux groupes ne différaient pas significativement quand le groupe expérimental voyait les items *sans doublet*, avec les variables dépendantes orthographe sans doublet, orthographe avec un doublet fréquent et orthographe avec un doublet rare ( $ts < 1.36, ps < .18$ ). De même, il n'y avait pas de différence entre les deux groupes concernant l'utilisation des doublets rares quand le groupe expérimental voyait des items *doublet fréquent* ( $t_1(36) = 1.35, p = .18$ ;  $t_2(5) = 1.66, p = .16$ ). En revanche, le groupe expérimental écrivait les items *doublet fréquent* moins souvent sans doublet mais plus souvent avec un doublet fréquent que le groupe contrôle (respectivement,  $t_1(46) = 6.64, p < .001$ ;  $t_2(5) = 6.45, p = .001$  et  $t_1(46) = 5.4, p < .001$ ;  $t_2(5) = 6.15, p = .002$ ). Une seconde différence était que lorsque le groupe

expérimental voyait des items *doublet rare*, ils écrivaient ces items moins souvent sans doublet mais plus souvent avec un doublet rare ainsi que plus souvent avec un doublet fréquent que le groupe contrôle (respectivement,  $t_1(36) = 7.95, p < .001$ ;  $t_1(5) = 5.52, p = .003$ ,  $t_1(36) = 5.63, p < .001$ ;  $t_1(5) = 4.20, p = .008$ , et  $t_1(36) = 2.19, p = .034$ ;  $t_1(5) = 6.15, p = .001$ ).

## **Discussion**

Un résultat clef de l'étude de Fayol et al. (2010) —le taux d'erreurs de transposition plus élevé pour les items *doublet rare* que pour les items *doublet fréquent*— a été répliqué dans cette étude. Cette réplification était observée alors même qu'il existe plusieurs différences procédurales entre ces expériences. Une différence est que les participants de l'Expérience 1 lisaient les histoires à voix haute (ce qui nous a permis de contrôler que les orthographes incluant seulement des consonnes simples et celles incluant des consonnes doubles étaient prononcées pareillement). Une seconde différence était que l'apprentissage orthographique était évalué dans l'Expérience 1 avec une tâche de dictée en demandant explicitement aux participants d'écrire les non-mots comme ils étaient orthographiés dans les histoires. Une troisième différence est que l'apprentissage orthographique était évalué une fois les six histoires lues (une procédure avec laquelle il est très improbable que les participants se focalisent de plus en plus sur l'orthographe des non-mots au fur et à mesure qu'ils progressent dans la lecture des histoires).

La plupart des erreurs de transposition que nous avons observées impliquent la transposition du doublet d'une consonne rarement doublée à une consonne fréquemment doublée en français. Ce résultat montre que les sujets possèdent une connaissance relative à la fréquence de doublement des consonnes et qu'ils utilisent cette connaissance en production orthographique. Les orthographes produites par le groupe contrôle montrent également une connaissance de certaines régularités de l'emploi des doubles lettres. Même si la plupart de leurs orthographes incluaient seulement des consonnes simples, ceux qui incluaient des doublets utilisaient toujours des doublets dans une position légale (médiane) en français. De plus, les doublets utilisés étaient beaucoup plus souvent des doublets fréquents en français que des doublets rares.

Les comparaisons entre les orthographes produites par les groupes expérimental et



contrôle nous renseignent sur les interactions entre des connaissances générales sur l'écrit et des connaissances relatives à l'orthographe de mots spécifiques. Les participants du groupe contrôle qui n'ont pas de connaissances relatives à l'orthographe de mots spécifiques, utilisent davantage les consonnes simples que les doublets fréquents dans leurs productions orthographiques et n'utilisent presque jamais des doublets rares. Les participants du groupe expérimental ont acquis des informations sur l'orthographe de mots spécifiques, notamment sur la présence d'un doublet dans un item spécifique, mais ils sont également influencés par leurs connaissances générales sur l'écrit.

Même si le pattern d'erreurs de transposition dans l'Expérience 1 était similaire à celui rapporté par Fayol et al. (2010), les erreurs d'omission étaient beaucoup plus fréquentes dans l'Expérience 1 que dans l'expérience de Fayol et al. (2010) (28.2% vs. 3.7%). Les omissions sur les items *doublet fréquent* et *doublet rare*, expliquent pourquoi les items *doublet fréquent* étaient orthographiés correctement moins souvent que les items *sans doublet* dans l'Expérience 1 alors que les taux d'orthographe correctes étaient très élevés (effet plafond) et ne différaient pas entre ces deux types d'items dans l'étude de Fayol et al. (2010). Les omissions étaient plus nombreuses dans l'Expérience 1 que dans celle de Fayol et al. (2010) car le délai entre la lecture d'un non-mot et l'évaluation de l'apprentissage de son orthographe était plus long dans l'Expérience 1. En effet, des études antérieures ont déjà rapporté de meilleures performances sur des mesures d'apprentissage orthographique après des délais courts qu'après des délais longs (e.g., Nation et al., 2007).

En résumé, les résultats de l'étude de Fayol et al. (2010) et ceux de l'Expérience 1 suggèrent que les individus se rappellent parfois de la présence d'un doublet mais ne se rappellent pas de la lettre spécifique qui est doublée. Dans de telles situations, ils se fondent sur leur connaissance du fait que certaines consonnes sont plus fréquemment doublées que d'autres. Ceci les conduit à orthographier correctement les items *doublet fréquent* mais à commettre des erreurs de transposition sur les items *doublet rare*. Ce résultat semble assez robuste étant donné que les deux expériences diffèrent sur plusieurs dimensions comme le délai entre l'exposition à un non-mot et la récupération de son orthographe, l'instruction d'orthographier le non-mot tel qu'il l'était dans le texte, les non-mots qui correspondent à des noms propres ou à des noms communs et enfin la lecture silencieuse ou à haute voix des histoires.

Afin d'explorer davantage si ces résultats sont généralisables, deux autres expériences (Expériences 2 et 3) ont été conduites. Toutes deux utilisent un nouvel échantillon de non-

mots, plus large que celui utilisé dans l'expérience de Fayol et al. (2010) et dans l'Expérience 1. Dans l'Expérience 2, comme dans certaines études explorant le rôle du *self-teaching* dans l'apprentissage de l'orthographe (e.g., Share, 1995), l'apprentissage orthographique était évalué avec une tâche de rappel (dictée comme dans l'Expérience 1) et une tâche de reconnaissance. Dans l'Expérience 3, les sujets devaient lire les non-mots qui étaient présentés de façon isolée sur l'écran d'un ordinateur.

## Expérience 2

### Lecture à haute voix de textes incluant des non-mots chez l'adulte

#### Méthode

#### Participants

Soixante-douze étudiants (40 femmes) de l'Université Paris Descartes, tous de langue maternelle française ont participé à l'expérience. L'âge moyen était de 21 ans et 9 mois, tous compris entre 18 et 28 ans. L'apprentissage orthographique était testé avec une tâche de rappel (dictée) pour la moitié des participants et avec une tâche de reconnaissance pour l'autre moitié.

#### Matériel

Dix-huit non-mots trisyllabiques incluant deux consonnes cibles ont été construits. Une des consonnes cibles était *n*, *t*, ou *r* (trois consonnes fréquemment doublées en français) et l'autre était *b*, *d*, ou *g* (trois consonnes rarement doublées). Chacune des neuf combinaisons de ces deux types de lettres a été utilisée dans deux non-mots. Dans un non-mot, l'attaque de la seconde syllabe était la consonne fréquemment doublée et l'attaque de la troisième syllabe était la consonne rarement doublée (e.g., *cinabé* pour la combinaison *b/n*). Dans l'autre non-mot, l'attaque de la seconde syllabe était la consonne rarement doublée et l'attaque de la troisième syllabe était la consonne fréquemment doublée (e.g., *cibané* pour la combinaison *b/n*). Trois orthographe, se prononçant toutes de la même façon selon les règles de correspondances graphème-phonème du français, ont été créées pour chaque non-mot : une orthographe sans doublet (e.g., *cinabé*, *cibané*), une orthographe avec un doublet fréquent (*cinnabé*, *cibanné*) et une orthographe avec un doublet rare (*cinabbé*, *cibbané*) (cf., *Annexe 8*).

Les non-mots ont été insérés dans les mêmes histoires que pour l'Expérience 1, avec la contrainte que les deux non-mots de chaque histoire n'appartenaient pas à la même catégorie d'items (i.e., *sans doublet*, *doublet fréquent* ou *doublet rare*).

Comme trois orthographe étaient créées pour chacun des 18 non-mots expérimentaux

et que six non-mots étaient utilisés pour chaque participant, chaque forme orthographique de chaque non-mot était utilisée pour quatre participants. Pour le test de reconnaissance, pour chaque non-mot, trois orthographes qui ne différaient que par le format des consonnes cibles étaient écrites sur une page d'un livret. Une orthographe contenait seulement des consonnes simples, une orthographe incluait un doublet fréquent et une orthographe incluait un doublet rare. (e.g., *cibané*, *cibanné*, *cibbané*).

### **Procédure**

La procédure dans la condition rappel était strictement identique à celle de l'Expérience 1. La procédure pour la condition reconnaissance différait seulement au niveau du test au cours duquel les participants devaient indiquer laquelle de trois orthographes était utilisée dans les textes qu'ils venaient de lire.

### **Résultats**

Quand ils lisaient les histoires à voix haute, les participants prononçaient de la même façon les consonnes cibles et leurs voyelles adjacentes, qu'elles soient doublées ou non doublées. Ils répondaient toujours correctement aux questions évaluant leur compréhension des textes.

#### **Condition rappel**

Le nombre de non-mots avec les deux consonnes cibles correctement orthographiées, indiqué au Tableau 3, a été soumis à une ANOVA utilisant les sujets et les items comme variables aléatoires. Dans l'analyse sur les items, nous avons utilisé les scores pour chaque combinaison de consonnes fréquemment et rarement doublées. C'est pourquoi, bien que six items soient présentés à chaque participant dans l'Expérience 1 et dans l'Expérience 2, l'analyse par item a été conduite sur six items dans l'Expérience 1 et sur neuf items dans l'Expérience 2. Il y avait un effet principal du type d'item ( $F_1(2, 70) = 26.89$ ,  $MSE = .44$ ,  $p < .001$ ,  $\eta_p^2 = .43$ ;  $F_2(2, 16) = 24.90$ ,  $MSE = 46.93$ ,  $p < .001$ ,  $\eta_p^2 = .76$ ). Les comparaisons planifiées indiquaient que les orthographes correctes étaient significativement plus nombreuses pour les items *sans doublet* que pour les items *doublet fréquent* et *doublet rare* ( $F_1(1, 35) = 50.91$ ,  $MSE = .37$ ,  $p < .001$ ,  $\eta_p^2 = .59$ ;  $F_2(1, 8) = 117.87$ ,  $MSE = 0.64$ ,  $p < .001$ ,

$\eta_p^2 = .94$ ) et, parmi les items requérant un doublet, pour les items *doublet fréquent* que pour les items *doublet rare* ( $F_1(1, 35) = 9.00$ ,  $MSE = .50$ ,  $p = .005$ ,  $\eta_p^2 = .20$ ;  $F_2(1, 8) = 5.76$ ,  $MSE = 3.13$ ,  $p = .043$ ,  $\eta_p^2 = .42$ ).

Tableau 3

Pourcentages des différents types d'orthographe produites dans l'Expérience 2. Les orthographe correctes apparaissent en gras, les transpositions en italique, et les écart-types sont entre parenthèses.

Orthographe produite (condition rappel) ou choisie (condition reconnaissance)	condition rappel		condition reconnaissance	
	item	item	item	item
	<i>sans</i>	<i>doublet</i>	<i>sans</i>	<i>doublet</i>
	<i>doublet</i>	<i>fréquent</i>	<i>doublet</i>	<i>rare</i>
Orthographe sans doublet	<b>87.5 (22.0)</b>	38.9 (39.8)	<b>75.0 (25.4)</b>	30.6 (36.4)
		40.3 (33.4)		27.8 (25.2)
Orthographe avec un doublet fréquent	4.2 (14.0)	<b>55.6 (42.7)</b>	13.9 (22.7)	<b>63.9 (38.9)</b>
		29.2 (34.6)		31.9 (34.1)
Orthographe avec un doublet rare	8.3 (18.9)	5.6 (15.9)	11.1 (21.1)	<b>40.3 (28.8)</b>
		<b>30.6 (27.5)</b>		5.6 (15.9)

Les erreurs de transposition étaient significativement plus nombreuses pour les items *doublet rare* que pour les items *doublet fréquent* ( $t_1(35) = 3.66, p = .001; t_2(8) = 4.15, p = .003$ ). Ce résultat est confirmé par l'analyse des profils individuels, comme le montre le Tableau 2. En effet, parmi les 20 participants qui commettaient des erreurs de transposition, ceux commettant des erreurs uniquement sur les items *doublet rare* étaient beaucoup plus nombreux que ceux commettant des erreurs uniquement sur les items *doublet fréquent* (16 vs. 3). De plus, la prévalence des erreurs de transposition pour les items *doublet rare* ne reflétait pas une tendance générale à utiliser des doublets fréquents indépendamment du type d'item présenté. En effet, les items *doublet rare* étaient orthographiés avec un doublet fréquent (29.2% d'erreurs de transposition) à la fois moins souvent que les items *doublet fréquent* (55.6% correct) et plus souvent que les items *sans doublet* (4.2% erreurs d'addition). Une ANOVA sur l'utilisation d'un doublet fréquent avec la variable type d'item (*sans doublet*, *doublet fréquent*, ou *doublet rare*) révèle un effet significatif ( $F_1(2, 70) = 29.82, \text{MSE} = .32, p < .001, \eta_p^2 = .46$  ;  $F_2(2, 16) = 17.94, \text{MSE} = 2.12, p < .001, \eta_p^2 = .69$ ), avec davantage de doublets fréquents utilisés pour les items *doublet fréquent* que pour les items *doublet rare* et *sans doublet* ( $F_1(1, 35) = 39.65, \text{MSE} = .36, p < .001, \eta_p^2 = .53$  ;  $F_2(1, 8) = 20.26, \text{MSE} = 2.86, p = .002, \eta_p^2 = .72$ ) et davantage de doublets fréquents utilisés pour les items *doublet rare* que pour les items *sans doublet* ( $F_1(1, 35) = 16.58, \text{MSE} = .27, p < .001, \eta_p^2 = .32$ ;  $F_2(1, 8) = 13.09, \text{MSE} = 1.37, p = .007, \eta_p^2 = .62$ ).

Les erreurs d'omission étaient assez fréquentes dans l'Expérience 2 et leur taux ne différait pas significativement entre les items *doublet fréquent* et *doublet rare* ( $t_1(35) = .25, p = .80; t_2(8) = .15, p = .89$ ). Aucune erreur n'impliquait le doublement d'une lettre jamais doublée en français et aucune erreur n'impliquait un doublement en position initiale ou finale.

### Condition reconnaissance

Les performances pour la tâche de reconnaissance sont indiquées au Tableau 3. Une ANOVA sur le nombre de sélections d'orthographe correctes révèle un effet du type d'item ( $F_1(2, 70) = 9.43, \text{MSE} = 4.53, p < .001, \eta_p^2 = .21$ ;  $F_2(2, 16) = 12.91, \text{MSE} = 18.11, p < .001, \eta_p^2 = .62$ ), avec les orthographe correctes plus souvent choisies pour les items *sans doublet* que pour les items *doublet fréquent* et *doublet rare* ( $F_1(1, 35) = 10.30, \text{MSE} = .49, p = .003, \eta_p^2 = .23$ ;  $F_2(1, 8) = 32.27, \text{MSE} = .63, p < .001, \eta_p^2 = .80$ ) et, parmi les items requérant un doublet, de meilleurs scores pour les items *doublet fréquent* que pour les items *doublet rare* ( $F_1(1, 35) = 8.52, \text{MSE} = .47, p = .006, \eta_p^2 = .20$ ;  $F_2(1, 8) = 7.36, \text{MSE} = 2.18, p = .027,$

$\eta_p^2 = .48$ ). En effet, le taux de sélection de l'orthographe correcte ne différait pas significativement du hasard (1/3) pour les items *doublet rare* ( $t(35) = 1.45, p = .16$ ), alors qu'ils étaient significativement au-dessus du hasard pour les items *sans doublet* ( $t(35) = 9.86, p < .001$ ) et *doublet fréquent* ( $t(35) = 4.71, p < .001$ ).

Les participants sélectionnaient significativement plus les orthographe qui impliquaient des transpositions pour les items *doublet rare* que pour les items *doublet fréquent* ( $t_1(35) = 3.66, p = .001$ ;  $t_2(8) = 4.15, p = .003$ ). Comme le montre le Tableau 2, ce pattern de résultat était confirmé par l'analyse des profils individuels, avec 18 des 22 participants commettant des erreurs de transposition qui les commettaient uniquement sur les items *doublet rare*.

La prévalence des erreurs de transposition sur les items *doublet rare* ne résultait pas d'une préférence générale pour les items incluant un doublet fréquent, indépendamment du type d'items présenté. En effet, les orthographe avec un doublet fréquent étaient sélectionnées pour les items *doublet rare* (31.9% d'erreurs de transposition) à la fois moins souvent que pour les items *doublet fréquent* (63.9% correct) et plus souvent que pour les items *sans doublet* (13.9% d'erreurs d'addition). Une ANOVA avec le nombre de sélections de l'orthographe incluant un doublet fréquent avec la variable type d'item (*sans doublet*, *doublet fréquent*, et *doublet rare*) révèle un effet ( $F_1(2, 70) = 30.0, MSE = .30, p < .001, \eta_p^2 = .46$ ;  $F_2(2, 16) = 43.8, MSE = .84, p < .001, \eta_p^2 = .85$ ), avec les orthographe incluant un doublet fréquent plus souvent choisies pour les items *doublet fréquent* que pour les items *doublet rare* et *sans doublet* ( $F_1(1, 35) = 42.1, MSE = .38, p < .001, \eta_p^2 = .55$ ;  $F_2(1, 8) = 49.7, MSE = 1.3, p < .001, \eta_p^2 = .86$ ) et les orthographe incluant un doublet fréquent plus souvent choisies pour les items *doublet rare* que pour les items *sans doublet* ( $F_1(1, 35) = 10.1, MSE = .23, p = .003, \eta_p^2 = .22$ ;  $F_2(1, 8) = 24.1, MSE = .39, p = .001, \eta_p^2 = .75$ ). Ces résultats confirment donc que les participants ont bien mémorisé les items qui incluait un doublet et ceux qui n'en incluait pas.

Alors que les erreurs de transposition étaient presque restreintes aux items *doublet rare*, le nombre d'erreurs d'omission ne différaient pas en fonction du type d'item ( $t_1(35) = .53, p = .60$ ;  $t_2(8) = .32, p = .75$ ).



## Discussion

L'utilisation d'un nouvel échantillon de non-mots, plus varié que les expériences précédentes (Fayol et al. (2010) et Expérience 1) et l'utilisation d'une tâche de reconnaissance en plus du seul rappel dans l'Expérience 2 suggère que les résultats rapportés dans ces trois expériences sont robustes. Un résultat important est que les participants ont davantage de difficulté à rappeler et reconnaître les orthographes contenant des doublets rares que les orthographes contenant un doublet fréquent ou ne contenant pas de doublet. La difficulté avec les orthographes contenant des doublets rares était flagrante dans la tâche de reconnaissance, avec des taux de sélection des orthographes correctes au-dessus du hasard pour les items *doublet fréquent* et *sans doublet* mais qui ne différaient pas significativement du hasard pour les items *doublet rare*. Un second résultat important est que les participants commettaient des erreurs de transposition, dont la plupart impliquait la transposition du doublement d'une consonne rarement doublée à une consonne fréquemment doublée en français. Ceci ne reflète pas une tendance générale à inclure des doublets fréquents dans les productions orthographiques (condition rappel) ou à choisir des orthographes contenant des doublets fréquents (condition reconnaissance), car les participants produisaient et choisissaient les orthographes contenant un doublet beaucoup moins souvent pour les items *sans doublet* que pour les items *doublet fréquent* et *doublet rare*.

## Expérience 3

### Lecture silencieuse et à haute voix de non-mots présentés isolément chez l'adulte

Les Expériences 1 et 2 de ce chapitre expérimental ainsi que celle conduite par Fayol et al. (2010) ont mis en évidence un impact des connaissances orthographiques générales, relatives à la fréquence de doublement des consonnes, sur l'apprentissage de l'orthographe de non-mots insérés dans des histoires que les sujets lisent silencieusement (Fayol et al., 2010) ou à voix haute (Expériences 1 et 2). Cette situation est la plus utilisée dans les études dites de *self-teaching* (e.g., Share, 2008) mais l'apprentissage de l'orthographe de non-mots présentés de façon isolée a parfois été étudié (Nation et al., 2007 ; Wang et al., 2011). L'objectif de l'Expérience 3 est de répliquer les résultats des Expériences 1 et 2 dans une situation d'apprentissage au cours de laquelle les non-mots sont présentés de façon isolée sur un écran d'ordinateur. Un groupe de participants les lisait silencieusement et un second groupe les lisait à haute voix. Une seconde différence par rapport aux Expériences 1 et 2 concerne le nombre d'items présentés durant la phase d'étude. Dans les expériences précédentes, quatre non-mots incluaient un doublet (deux non-mots incluaient un doublet fréquent et deux non-mots incluaient un doublet rare) et deux non-mots n'incluaient pas de doublet. Dans l'Expérience 3, les participants lisaient six non-mots incluant un doublet et six non-mots sans doublet. Notre hypothèse était que, en dépit de ces différences de procédures concernant le mode de présentation des non-mots, le nombre total de non-mots présentés et la proportion de non-mots incluant des doublets, les résultats devraient être identiques à ceux rapportés dans les Expériences 1 et 2.

### Méthode

#### Participants

Quarante-huit étudiants (32 femmes) de l'Université Paris Descartes, tous de langue maternelle française ont participé à l'expérience. L'âge moyen était de 21 ans et 3 mois, tous compris entre 18 et 25 ans. La moitié des participants était assignée à la condition lecture silencieuse et l'autre à la condition lecture à voix haute.

## **Matériel**

### **Matériel pour tester l'apprentissage de l'orthographe de non-mots**

Dix-huit non-mots expérimentaux et trois non-mots tampons ont été utilisés. Les 18 non-mots expérimentaux étaient ceux utilisés dans l'Expérience 2, chacun orthographié de trois façons différentes : une orthographe sans doublet (e.g., *lobanin*, *lonabin*), une orthographe avec un doublet fréquent (*lobannin*, *lonnabin*) et une orthographe avec un doublet rare (*lobbanin*, *lonabbin*). Dans les trois items tampons, utilisés pour que les sujets soient exposés à autant de non-mots incluant un doublet que de non-mots n'incluant pas de doublet, l'attaque de la deuxième syllabe et l'attaque de la troisième syllabe ne correspondaient pas à l'une des six consonnes *n*, *t*, *r*, *b*, *d*, et *g* (*bupila*, *tocali* and *moufopin*). Comme trois orthographes étaient créés pour chacun des 18 non-mots expérimentaux et que neuf non-mots étaient utilisés pour chaque participant, chaque forme orthographique de chaque non-mot expérimental était utilisée pour quatre participants de la condition *lecture silencieuse* et pour quatre participants de la condition *lecture à voix haute*.

### **Matériel pour évaluer la connaissance de l'orthographe de mots spécifiques**

Deux épreuves ont été utilisées pour évaluer le niveau d'orthographe lexicale des étudiants afin de vérifier que le groupe lisant les mots silencieusement et le groupe lisant les mots à voix haute étaient de niveau orthographique comparable. La première épreuve est une dictée de 32 mots (cf Annexes). La seconde est une épreuve de reconnaissance. Elle consiste en 30 orthographes de mots présentées dans un ordre aléatoire sur une page A4. Quinze orthographes sont correctes et 15 autres sont incorrectes mais phonologiquement plausibles (e.g., *pantin* orthographié *pantain*, cf Annexes).

## **Procédure**

Les passations étaient individuelles. Dans les deux conditions expérimentales, les non-mots étaient présentés les uns après les autres au centre d'un écran d'ordinateur portable (Macbook 13 pouces) pendant 1.5 seconde à l'aide du logiciel Psyscope. Chaque participant voyait chacun des 12 non-mots cinq fois, répartis selon un ordre aléatoire. Dans la condition *lecture à voix haute*, les participants devaient lire les non-mots à haute voix. Dans la condition *lecture silencieuse*, les participants devaient lire les non-mots silencieusement. Dans la

condition *lecture à voix haute*, les participants étaient enregistrés afin de pouvoir vérifier qu'ils prononçaient correctement les non-mots.

A l'issue de cette phase d'étude, les participants des deux groupes effectuaient une tâche de barrage pendant une dizaine de minutes puis ils passaient le Trail Making Test et une épreuve de mémoire à court terme (rappel de chiffres dans l'ordre direct et dans l'ordre indirect) puis les 12 non-mots leur étaient dictés, avec la consigne de les orthographier comme ils l'étaient dans la phase d'étude. Ensuite, leur niveau en orthographe lexicale était évalué, d'abord avec la dictée de 32 mots puis avec l'épreuve de reconnaissance. Pour la dictée, les mots étaient dictés dans un ordre différent pour chaque sujet. Pour l'épreuve de reconnaissance, l'expérimentateur donnait aux participants la feuille A4 sur laquelle se trouvaient les orthographes de 30 mots, 15 correctes et 15 incorrectes puis leur demandait de barrer les mots mal orthographiés.

## **Résultats**

Les performances des deux groupes aux épreuves évaluant les connaissances en orthographe lexicale, au test de mémoire à court terme et au Trail Making Test, rapportées au Tableau 4, ne différaient pas entre les deux groupes ( $ts < .72$ ,  $ps > .47$ ). Les participants de la condition lecture à haute voix prononçaient toujours correctement les non-mots.

Comme le montre le Tableau 5, dans les deux groupes, les productions orthographiques correctes étaient plus nombreuses pour les items *sans doublet* que pour les items *doublet fréquent* et elles étaient plus nombreuses pour les items *doublet fréquent* que pour les items *doublet rare*. Une ANOVA avec les variables groupe (lecture silencieuse et lecture à voix haute) et type d'items (*sans doublet*, *doublet fréquent* et *doublet rare*) sur le nombre de productions avec les deux consonnes cibles correctement orthographiées confirme l'effet du type d'item ( $F_1(2, 92) = 160.02$ ,  $MSE = .35$ ,  $p < .001$ ,  $\eta_p^2 = .78$  ;  $F_2(2, 16) = 75.66$ ,  $MSE = 1.96$ ,  $p < .001$ ,  $\eta_p^2 = .90$ ), sans effet du groupe ( $ps > .26$ ), ni interaction ( $ps > .68$ ). Des comparaisons planifiées révèlent des scores significativement meilleurs pour les items *sans doublet* que pour les items *doublet fréquent* et *doublet rare* (respectivement,  $F_1(1, 46) = 170.45$ ,  $MSE = .34$ ,  $p < .001$ ,  $\eta_p^2 = .79$  ;  $F_2(1, 8) = 100.00$ ,  $MSE = 1.56$ ,  $p < .001$ ,  $\eta_p^2 = .93$  et  $F_1(1, 46) = 273.90$ ,  $MSE = .37$ ,  $p < .001$ ,  $\eta_p^2 = .86$  ;  $F_2(1, 8) = 272.25$ ,  $MSE = 1.00$ ,  $p < .001$ ,  $\eta_p^2 = .97$ ) et, parmi les items requérant un doublet, de meilleurs scores pour les items *doublet fréquent* que pour les items *doublet rare* de façon significative pour l'analyse

sur les sujets ( $F_1(1, 46) = 18.45$ ,  $MSE = .33$   $p < .001$ ,  $\eta_p^2 = .29$ ) et de façon marginalement significative pour l'analyse sur les items ( $F_2(1, 8) = 4.83$ ,  $MSE = 3.31$ ,  $p = .059$ ,  $\eta_p^2 = .38$ ).

Tableau 4.

Performances des deux groupes aux épreuves évaluant les connaissances en orthographe lexicale, au test de mémoire à court terme et au Trail Making Test ; les écart-types sont entre parenthèses.

	lecture à voix haute	lecture silencieuse
Orthographe lexicale : dictée (/32)	20.2 (5.9)	19.5 (5.8)
Orthographe lexicale : reconnaissance (/30)	20.0 (2.7)	20.0 (3.5)
Mémoire à court terme : ordre direct (/16)	10.9 (2.1)	10.5 (2.3)
Mémoire à court terme : ordre indirect (/16)	8.8 (2.9)	8.9 (2.8)
TMT A (en msec)	24.8 (11.2)	22.6 (10.5)
TMT B (en msec)	53.7 (28.4)	50.8 (29.3)
Différence TMT B – TMT A (en msec)	28.9 (26.5)	28.2 (26.9)

Tableau 5

Pourcentages des différents types d'orthographe produites en fonction de la condition d'apprentissage et du type d'items présentés. Les orthographe correctes apparaissent en gras, les erreurs de transposition en italique et les écart-types sont entre parenthèses.

Orthographe produite	Lecture à haute voix			Lecture silencieuse		
	Item	Item	Item	Item	Item	Item
Orthographe produite	<i>sans</i>	<i>doublet</i>	<i>doublet</i>	<i>sans</i>	<i>doublet</i>	<i>doublet</i>
	<i>doublet</i>	<i>fréquent</i>	<i>rare</i>	<i>doublet</i>	<i>fréquent</i>	<i>rare</i>
Orthographe sans doublet	<b>90.3 (15.5)</b>	56.9 (30.3)	61.1 (35.0)	<b>91.7 (14.7)</b>	52.8 (21.8)	45.8 (21.6)
Orthographe avec un doublet fréquent	6.9 (13.8)	<b>36.1 (29.4)</b>	20.8 (27.5)	6.9 (13.8)	<b>41.7 (22.5)</b>	27.8 (21.2)
Orthographe avec un doublet rare	2.8 (9.4)	6.9 (13.8)	<b>18.1 (26.0)</b>	1.4 (6.8)	5.6 (12.7)	<b>26.4 (21.9)</b>

Les erreurs de transposition et d'omission ont été successivement soumises à une ANOVA avec les variables groupe (lecture silencieuse et lecture à voix haute) et type d'items (*doublet fréquent* et *doublet rare*). L'ANOVA sur les omissions ne révèle aucun effet ( $ps > .71$  pour le type d'item ;  $ps > .17$  pour le groupe et  $ps > .15$  pour l'interaction). Par contraste, les erreurs de transposition étaient significativement plus fréquentes pour les items *doublet rare* que pour les items *doublet fréquent* (24.3% vs. 6.3%,  $F_1(1, 46) = 19.53$ ,  $MSE = .36$ ,  $p < .001$ ,  $\eta_p^2 = .30$  ;  $F_2(1, 8) = 15.45$ ,  $MSE = 1.21$ ,  $p = .004$ ), sans différence entre les deux groupes ( $ps > .49$ ) et sans interaction groupe x type d'item ( $ps > .33$ ). Comme indiqué au Tableau 2, ce pattern de résultats était confirmé par l'analyse des profils individuels : dans les deux groupes, parmi les sujets commettant des erreurs de transposition, ceux commettant des erreurs de transposition uniquement sur les items *doublet rare* étaient beaucoup plus nombreux que ceux commettant des erreurs de transposition uniquement sur les items *doublet fréquent* (24 vs. 5 sur l'ensemble des deux groupes).

Comme dans les Expériences 1 et 2, la prévalence des erreurs de transposition sur les items *doublet rare* ne pouvait pas s'expliquer par une plus fréquente utilisation des doublets fréquents indépendamment du type d'items. En effet, les orthographe incluant un doublet fréquent étaient produites moins souvent pour les items *doublet rare* (24.3%) que pour les items *doublet fréquent* (43.8%) mais plus souvent pour les items *doublet rare* que pour les items *sans doublet* (6.9%). Une ANOVA avec les variables groupe (lecture silencieuse et lecture à voix haute) et type d'items (*sans doublet*, *doublet fréquent* et *doublet rare*) sur le nombre de productions orthographiques incluant un doublet fréquent confirme l'effet du type d'item ( $F_1(2, 92) = 32.51$ ,  $MSE = .45$ ,  $p < .001$ ,  $\eta_p^2 = .41$ ;  $F_2(2, 16) = 20.43$ ,  $MSE = 1.44$ ,  $p < .001$ ,  $\eta_p^2 = .72$ ), sans interaction avec le groupe ( $ps > .25$ ). Des analyses de contraste confirment que les orthographe incluant un doublet fréquent étaient produites moins souvent pour les items *doublet rare* que pour les items *doublet fréquent* ( $F_1(1, 46) = 16.01$ ,  $MSE = .51$ ,  $p < .001$ ,  $\eta_p^2 = .26$  ;  $F_2(1, 8) = 7.54$ ,  $MSE = 1.63$ ,  $p = .025$ ,  $\eta_p^2 = .49$ ) mais plus souvent pour les items *doublet rare* que pour les items *sans doublet* ( $F_1(1, 46) = 20.33$ ,  $MSE = .32$ ,  $p < .001$ ,  $\eta_p^2 = .31$  ;  $F_2(1, 8) = 18.80$ ,  $MSE = .92$ ,  $p = .002$ ,  $\eta_p^2 = .70$ ). Les erreurs de transposition tendaient à être plus nombreuses dans le groupe *lecture silencieuse* que dans le groupe *lecture à haute voix* mais l'effet de la condition était seulement marginalement significatif pour l'analyse sur les sujets ( $F_1(1, 46) = 3.52$ ,  $MSE = .50$ ,  $p = .067$ ,  $\eta_p^2 = .07$ ) et n'était pas significatif dans l'analyse sur les items ( $F_2(1, 8) = .59$ ,  $MSE = 2.54$ ,  $p = .46$ ,  $\eta_p^2 = .07$ ).

Enfin, les participants ne doublaient jamais des lettres qui ne peuvent pas être doublées en français, n'utilisaient jamais de doublet en début ou en fin de mots. Quelques productions orthographiques étaient phonologiquement incorrectes mais ces erreurs ne portaient jamais sur les consonnes cibles.

## **Discussion**

Les résultats des Expériences 1 et 2, dans lesquelles les non-mots étaient insérés dans des histoires, étaient répliqués dans l'Expérience 3 dans laquelle un nombre de non-mots plus important était présenté de façon isolée : trois items de chaque type dans l'Expérience 3 au lieu de deux dans les Expériences 1 et 2 et la présence de trois items tampons sans doublet de sorte que le matériel de l'Expérience 3 comportait six non-mots avec un doublet et six non-mots sans doublet alors que le matériel des Expériences 1 et 2 comportait quatre items avec un doublet et deux sans doublet. D'une part, les items *sans doublet* étaient mieux orthographiés que les items *doublet fréquent* et ces derniers étaient mieux orthographiés que les items *doublet rare*. D'autre, part, alors que les erreurs d'omission étaient aussi nombreuses pour les items *doublet fréquent* et *doublet rare*, les erreurs de transposition étaient beaucoup plus nombreuses pour les items *doublet rare* que pour les items *doublet fréquent*. De plus, aucune différence n'était observée entre les sujets qui lisaient les items silencieusement et ceux qui les lisaient à haute voix, ce qui est en accord avec des études dans lesquelles le niveau d'apprentissage de l'orthographe de non-mots insérés dans des histoires ne variait pas selon que les sujets les lisaient silencieusement ou à haute voix (De Jong & Share 2007).

Ainsi, les Expériences 1, 2 et 3 utilisaient des procédures différentes (e.g., non-mots insérés dans des textes vs. présentés de façon isolée ; lecture orale vs. silencieuse) mais toutes suggèrent que les sujets acquéraient des informations sur l'orthographe de mots spécifiques (notamment sur la présence d'un doublet dans un item spécifique) mais se fondaient également sur leurs connaissances générales sur l'écrit (notamment relatives à la fréquence de doublement des lettres). Ce patron de résultats semble donc particulièrement robuste.



## Expériences 4 et 5

L'objectif des Expériences 4 et 5 était d'étudier l'influence de la sensibilité à la fréquence de doublement des consonnes sur l'apprentissage de l'orthographe de nouveaux mots chez des élèves de l'école élémentaire. Ces expériences ont été conduites auprès d'élèves scolarisés en CE2, un niveau où la mise en place du lexique orthographique a déjà débuté (Bosse & Pacton, 2006) et où la connaissance du fait que certaines lettres sont plus fréquemment doublées que d'autres est bien établie (e.g., Pacton et al., 2001; Chapitre expérimental I), et en CM2 où le lexique orthographique est encore plus développé. Dans l'Expérience 4, le matériel utilisé est celui de l'Expérience 1 chez l'adulte, la lecture est silencieuse et l'apprentissage orthographique est évalué avec une épreuve de rappel (dictée) pour la moitié des sujets et avec une épreuve de reconnaissance pour l'autre moitié. Dans l'Expérience 5, le matériel utilisé est celui de l'Expérience 2 chez l'adulte, la lecture est à voix haute et l'apprentissage orthographique est également évalué avec une dictée pour la moitié des sujets et avec une épreuve de reconnaissance pour l'autre moitié.

Nous faisons l'hypothèse que, comme chez les adultes, l'orthographe des items *sans doublet* devrait être mieux apprise que celle des items *doublet fréquent* et que l'orthographe des items *doublet fréquent* devrait être mieux apprise que celle des items *doublet rare*. Les erreurs d'omissions devraient être relativement fréquentes, en particulier en CE2 lorsque l'apprentissage orthographique est évalué avec l'épreuve de rappel, et ces omissions devraient être aussi nombreuses pour les items *doublet fréquent* que pour les items *doublet rare*. Les élèves de CM2 devraient commettre davantage d'erreurs de transposition que les CE2 s'ils mémorisent mieux la présence d'un doublet dans certains items. Toutefois, aux deux niveaux scolaires, les transpositions devraient être plus nombreuses pour les items *doublet rare* que pour les items *doublet fréquent*. Ce pattern de résultats sur les réponses correctes et les erreurs (omissions aussi nombreuses sur les items *doublet fréquent* et *doublet rare*; transpositions plus nombreuses sur les items *doublet rare* que sur les items *doublet fréquent*) devrait être observé que la lecture soit silencieuse (Expérience 4) ou à voix haute (Expérience 5) et que l'apprentissage orthographique soit évalué avec une tâche de rappel ou de reconnaissance.

## Expérience 4

### Lecture silencieuse de textes incluant des non-mots chez des CE2 et CM2

#### Méthode

#### Participants

Les participants étaient 68 élèves de CE2 (36 filles; âge = 9 ans, 4 mois; EC = 4.6 mois) et 68 élèves de CM2 (31 filles ; âge moyen = 11 ans, 5 mois; EC = 6.2 mois), tous de langue maternelle française. Pour chaque niveau scolaire, 48 élèves étaient assignés aux groupes expérimentaux, 24 dans la condition rappel (dictée) et 24 dans la condition reconnaissance, et 20 constituaient le group contrôle.

#### Matériel

Les non-mots et les histoires utilisés dans l'Expérience 1 avec des adultes ont été réutilisés dans l'Expérience 4. Les non-mots incluaient donc tous deux consonnes cibles, l'une qui est fréquemment doublée en français (*n*, *t*, ou *r*), l'autre qui l'est rarement (*b*, *d*, ou *g*) et, pour chaque non-mot, trois orthographes étaient créées : sans doublet (e.g., *bagotin*), avec un doublet fréquent (*bagottin*), et avec un doublet rare (*baggotin*). Ces non-mots étaient insérés dans trois histoires (deux non-mots dans chaque histoire apparaissant chacun cinq fois). Chacune de ces histoires était suivie de quatre questions, la première consistant à sélectionner le titre approprié parmi une liste de trois titres, les trois autres étaient des questions vrai/faux. Un livret de six pages était créé pour le test de reconnaissance, avec sur chaque page un des non-mots qui était orthographié de trois façons : sans doublet, avec un doublet fréquent, et avec un doublet rare (e.g., *bagotin*, *bagottin*, *baggotin*).

#### Procédure

Les élèves des groupes expérimentaux étaient testés en groupes de huit à dix. Ils lisaient silencieusement une histoire, puis tournaient la page pour répondre aux questions sur cette histoire et ainsi de suite pour chacune des trois histoires. Après ceci, ils effectuaient une

tâche de barrage pendant 10 minutes. Enfin, dans la condition rappel, l'expérimentateur leur dictait les six non-mots en précisant qu'il fallait les orthographier comme ils l'étaient dans les histoires et, dans la condition reconnaissance, l'expérimentateur leur demandait d'indiquer laquelle de trois orthographes était utilisée dans les histoires. Les élèves du groupe contrôle, qui étaient testés par groupe de 10, ne lisaient pas les histoires et commençaient l'expérience par la dictée. Il leur était demandé d'écrire les six non-mots comme s'il s'agissait de nouveaux mots qu'ils rencontraient pour la première fois. Deux jours après, les élèves du groupe contrôle effectuaient la tâche de jugement dans laquelle ils recevaient le même livret que les élèves du groupe expérimental dont l'apprentissage de l'orthographe des non-mots était évalué avec une épreuve de reconnaissance entre trois orthographes (e.g., *bagotin*, *bagottin*, *baggotin*). Pour chaque non-mot, ils devaient indiquer laquelle des trois orthographes ressemblait le plus à un mot.

Une fois l'expérience passée, les niveaux de lecture et d'orthographe des élèves étaient évalués avec le test de lecture de Orlec (Lobrot, 1967) et la dictée du corbeau du L2MA (Chevrier-Muller et al., 1997) (pour des détails concernant ces deux tests, voir Chapitre expérimental I).

## **Résultats**

Le Tableau 6 présente les performances des enfants aux tests standards de lecture et orthographe et l'exactitude de leurs réponses aux questions sur les histoires. Les scores aux tests de lecture et d'orthographe ont été soumis à des ANOVAs avec les variables niveau scolaire (CE2, CM2) et condition (rappel, reconnaissance, contrôle). Les scores en lecture étaient meilleurs en CM2 qu'en CE2 ( $F(1, 130) = 160.24$ ,  $MSE = 24.28$ ,  $p < .001$ ,  $\eta_p^2 = .55$ ) et il n'y avait pas d'effet de la condition ni d'interaction ( $ps > .77$ ). Les scores en orthographe étaient meilleurs en CM2 qu'en CE2 ( $F(1, 130) = 91.38$ ,  $MSE = 34.85$ ,  $p < .001$ ,  $\eta_p^2 = .41$ ). L'effet de la condition était marginalement significatif ( $F(2, 130) = 2.57$ ,  $MSE = 34.85$ ,  $p = .08$ ,  $\eta_p^2 = .04$ ), avec les scores du groupe contrôle qui tendaient à être meilleurs que ceux des groupes rappel et reconnaissance, et il n'y avait pas d'interaction ( $p = .70$ ). Une ANOVA avec les variables niveau scolaire (CE2, CM2) et condition (rappel, reconnaissance) sur le nombre de réponses correctes aux questions sur les histoires révélait de meilleurs scores en CM2 qu'en CE2 ( $F(1, 92) = 25.92$ ,  $MSE = 2.44$ ,  $p < .001$ ,  $\eta_p^2 = .22$ ) mais pas d'effet de la condition ni d'interaction ( $ps > .38$ ). Ces résultats confirment que les élèves étaient assez bien

appariés entre les conditions et montrent, comme attendu, que les CM2 étaient meilleurs en lecture et orthographe que les CE2.

Tableau 6.

Scores aux tests standards de lecture et d'orthographe, pourcentages de réponses correctes aux questions sur les histoires et pourcentages de non-mots correctement décodés ; les écart-types sont indiqués entre parenthèses.

Expérience	Condition	Niveau scolaire	Lecture (Lobrot) (sur 36)	Orthographe (dictée du corbeau) (sur 50)	Questions sur les histoires (%)	Décodage (%)
Expérience 4	Rappel	CE2	21.3 (6.3)	33.8 (7.6)	73.3 (14.2)	—
Expérience 4	Reconnaissance	CE2	20.3 (6.7)	32.0 (8.1)	70.8 (16.7)	—
Expérience 4	Contrôle	CE2	20.3 (6.0)	35.9 (6.2)	—	—
Expérience 5	Rappel	CE2	20.7 (6.5)	35.4 (6.2)	78.3 (14.2)	78.7 (17.7)
Expérience 5	Reconnaissance	CE2	19.9 (6.9)	29.6 (8.4)	70.0 (14.2)	86.3 (13.0)
Expérience 4	Rappel	CM2	31.2 (2.5)	43.6 (4.0)	85.0 (10.8)	—
Expérience 4	Reconnaissance	CM2	31.4 (3.3)	42.7 (4.3)	85.0 (10.0)	—
Expérience 4	Contrôle	CM2	31.6 (2.5)	44.5 (2.8)	—	—
Expérience 5	Rappel	CM2	29.6 (5.3)	43.4 (5.3)	78.3 (15.8)	90.3 (13.0)
Expérience 5	Reconnaissance	CM2	32.0 (2.7)	43.4 (4.5)	87.5 (8.3)	97.0 (3.7)

Tableau 7

Pourcentage des différents types d'orthographe produites par le groupe expérimental dans la condition rappel de l'Expérience 4. Les orthographe correctes sont indiquées en gras et les erreurs de transposition en italique, les écart-types sont entre parenthèses.

	CE2			CM2		
	Items	Items	Items	Items	Items	Items
Orthographe sans doublet	<b>93.8</b> (16.9)	75.0 (33.0)	79.2 (32.7)	<b>91.7</b> (19.0)	52.1 (40.3)	43.8 (42.5)
Orthographe avec un doublet fréquent	0.0 (0.0)	<b>16.7</b> (28.2)	4.2 (14.1)	6.3 (16.9)	<b>39.6</b> (36.1)	27.1 (32.9)
Orthographe avec un doublet rare	2.1 (10.2)	2.1 (10.2)	<b>12.5</b> (22.1)	2.1 (10.2)	6.3 (27.1)	<b>27.1</b> (39.0)

Note. La somme des pourcentages des trois types d'orthographe diffère de 100% parce que 4.9% des erreurs en CE2 et 1.4% des erreurs en CM2 correspondaient à une représentation phonologiquement incorrecte d'au moins une consonne ciblée<sup>2</sup>.

<sup>2</sup> Une ANOVA avec les variables niveau scolaire (CE2, CM2) et type d'item (*sans doublet*, *doublet fréquent*, *doublet rare*) sur le nombre d'orthographe avec une représentation phonologique incorrecte d'au moins une consonne cible ne révélait aucun effet significatif ( $ps > .26$ ).

Tableau 8

Nombre d'élèves commettant des erreurs de transposition pour chaque expérience

Expérience	Condition	Niveau scolaire	N total	N commettant des erreurs de transposition		N commettant des erreurs de transposition seulement pour les items <i>doublet fréquent</i>	
				N commettant des erreurs de transposition	N commettant des erreurs de transposition seulement pour les items <i>doublet rare</i>	N commettant des erreurs de transposition	N commettant des erreurs de transposition seulement pour les items <i>doublet fréquent</i>
Expérience 4	Rappel	CE2	24	3	2	1	1
Expérience 4	Rappel	CM2	24	13	10	2	2
Expérience 4	Reconnaissance	CE2	24	17	9	0	0
Expérience 4	Reconnaissance	CM2	24	17	12	3	3
Expérience 5	Rappel	CE2	27	5	3	2	2
Expérience 5	Rappel	CM2	27	16	12	3	3
Expérience 5	Reconnaissance	CE2	27	19	10	1	1
Expérience 5	Reconnaissance	CM2	27	17	13	2	2

Note. Les cas où le nombre d'erreurs de transposition seulement sur les items *doublet rare* plus le nombre d'erreurs de transposition seulement sur les items *doublet fréquent* diffère du nombre total d'erreurs de transposition sont des cas d'élèves commettant des erreurs de transposition à la fois sur les items *doublet rare* et sur les items *doublet fréquent* (e.g., un élève de CE2 dans la condition rappel de l'Expérience 4 et 8 élèves de CE2 dans la condition reconnaissance de l'Expérience 4).

### Groupe expérimental, condition rappel

Le Tableau 7 présente le pourcentage d'orthographe correctes, définies comme celles dans lesquelles les deux consonnes cibles sont orthographiées comme dans les histoires. Les orthographe correctes étaient plus nombreuses pour les items *sans doublet* que pour les items *doublet fréquent* et *doublet rare*. Des ANOVAs avec la variable niveau scolaire (CE2, CM2) et type d'item (*sans doublet*, *doublet fréquent*, *doublet rare*) révèle un effet principal du type d'item ( $F_1(2, 92) = 91.92$ ,  $MSE = .33$ ,  $p < .001$ ,  $\eta_p^2 = .67$  ;  $F_2(2, 10) = 131.07$ ,  $MSE = .93$ ,  $p < .001$ ,  $\eta_p^2 = .96$ ). Les comparaisons planifiées indiquent que les orthographe correctes étaient significativement plus nombreuses pour les items *sans doublet* que pour les items *doublet fréquent* et *doublet rare* (respectivement,  $F_1(1, 46) = 127.39$ ,  $MSE = .31$ ,  $p < .001$ ,  $\eta_p^2 = .73$  ;  $F_2(1, 5) = 165.69$ ,  $MSE = .97$ ,  $p < .001$ ,  $\eta_p^2 = .97$  et  $F_1(1, 46) = 176.65$ ,  $MSE = .29$ ,  $p < .001$ ,  $\eta_p^2 = .79$  ;  $F_2(1, 5) = 306.25$ ,  $MSE = .67$ ,  $p < .001$ ,  $\eta_p^2 = .98$ ), sans différence significative entre les items *doublet fréquent* et *doublet rare* ( $ps > .20$ ). Il y avait également un effet principal du niveau scolaire, avec en moyenne plus d'orthographe correctes en CM2 qu'en CE2 (52.8% vs. 41.0%,  $F_1(1, 46) = 7.05$ ,  $MSE = .28$ ,  $p = .011$ ,  $\eta_p^2 = .13$  ;  $F_2(1, 5) = 7.81$ ,  $MSE = 1.03$ ,  $p = .038$ ,  $\eta_p^2 = .61$ ). L'interaction grade x type d'item n'était pas significative ( $F_1(2, 92) = 2.34$ ,  $MSE = .33$ ,  $p = .10$ ,  $\eta_p^2 = .05$  ;  $F_2(2, 10) = 3.41$ ,  $MSE = .91$ ,  $p = .074$ ,  $\eta_p^2 = .41$ ).

Les erreurs de transposition, qui impliquent le déplacement du doublement vers la mauvaise consonne cible (e.g., *baggotin* orthographié *bagottin* pour les items *doublet rare* ; *bagottin* orthographié *baggotin* pour les items *doublet fréquent*) sont indiquées au Tableau 7. Les erreurs de transposition étaient rares en CE2 pour les deux types d'items. Elles étaient plus fréquentes en CM2 pour les items *doublet rare* mais demeuraient rares pour les items *doublet fréquent*. Ces résultats étaient confirmés par une ANOVA avec les variables niveau scolaire (CE2, CM2) et type d'item (*doublet fréquent*, *doublet rare*). Il y avait un effet principal du type d'item, avec davantage d'erreurs impliquant le déplacement du doublement de la consonne rarement doublée en français vers la consonne fréquemment doublée (15.6% pour les items *doublet rare*) que d'erreurs impliquant le déplacement du doublement de la consonne fréquemment doublée en français vers la consonne rarement doublée (4.2% pour les items *doublet fréquent*) ( $F_1(1, 46) = 6.91$ ,  $MSE = .18$ ,  $p = .012$ ,  $\eta_p^2 = .13$ ;  $F_2(1, 5) = 20.86$ ,  $MSE = .24$ ,  $p = .006$ ,  $\eta_p^2 = .81$ ). L'effet principal du niveau scolaire, avec davantage d'erreurs de transposition en CM2 qu'en CE2, était significatif dans l'analyse sur les sujets ( $F_1(1, 46) = 11.60$ ,  $MSE = .15$ ,  $p = .001$ ,  $\eta_p^2 = .20$ ) et marginalement significatif dans



l'analyse sur les items ( $F_2(1, 5) = 5.67$ ,  $MSE = 1.24$ ,  $p = .063$ ,  $\eta_p^2 = .53$ ). L'interaction significative dans l'analyse sur les sujets ( $F_1(1, 46) = 4.62$ ,  $MSE = .18$ ,  $p = .037$ ,  $\eta_p^2 = .09$ ) et marginalement significative dans l'analyse sur les items ( $F_2(1, 5) = 5.87$ ,  $MSE = .58$ ,  $p = .06$ ,  $\eta_p^2 = .54$ ) reflète le fait qu'il y avait une différence entre les items *doublet fréquent* et les items *doublet rare* seulement en CM2 ( $F_1(1, 23) = 6.93$ ,  $MSE = .30$ ,  $p = .015$ ,  $\eta_p^2 = .23$  ;  $F_2(1, 5) = 11.36$ ,  $MSE = .73$ ,  $p = .020$ ,  $\eta_p^2 = .69$  en CM2;  $ps > .37$  en CE2). L'augmentation des erreurs de transposition du CE2 au CM2 et leur prévalence sur les items *doublet rare* apparaissaient également dans l'analyse des profils individuels, comme le montre le Tableau 8. La plupart des élèves qui commettaient des erreurs de transposition les commettaient seulement sur les items *doublet rare* ; les élèves commettant des erreurs de transposition étaient plus nombreux en CM2 qu'en CE2.

La fréquence des erreurs d'omission (lorsqu'un item *doublet fréquent* ou *doublet rare* est écrit sans doublet) peut être dérivée des valeurs indiquées dans le Tableau 7. Une ANOVA sur le nombre d'erreurs d'omission avec les variables niveau scolaire (CE2, CM2) et type d'item (*doublet fréquent* et *doublet rare*) montre que les élèves de CE2 commettaient davantage d'erreurs d'omission que ceux de CM2 (77.1% vs. 47.9%,  $F_1(1, 46) = 10.82$ ,  $MSE = .81$ ,  $p = .002$ ,  $\eta_p^2 = .19$  ;  $F_2(1, 5) = 45.00$ ,  $MSE = 1.20$ ,  $p = .001$ ,  $\eta_p^2 = .90$ ). Il n'y avait pas d'effet du type d'item ni d'interaction ( $ps > .18$ ). Le Tableau 7 montre que les erreurs d'addition (production d'une orthographe incluant un doublet pour un item *sans doublet*) étaient assez rares. Aucune erreur n'impliquait le doublement d'une lettre jamais doublée en français et aucune erreur n'impliquait le doublement d'une consonne en position initiale ou finale.

Un examen du pattern de réponses correctes et d'erreurs pour les trois types d'items montre que la prévalence des erreurs de transposition pour les items *doublet rare* ne survenait pas du fait d'une tendance générale à inclure des doublets fréquents indépendamment du type d'item présenté. En effet, les items *doublet rare* étaient orthographiés avec un doublet fréquent (15.6% d'erreurs de transposition) à la fois moins souvent que les items *doublet fréquent* (28.1% d'orthographe correctes) et plus souvent que les items *sans doublet* (3.1% d'erreurs d'addition). Une ANOVA sur l'utilisation de doublets fréquents avec les variables niveau scolaire (CE2, CM2) et type d'item (*sans doublet*, *doublet fréquent* et *doublet rare*) révèle un effet principal du type d'item ( $F_1(2, 92) = 16.78$ ,  $MSE = .18$ ,  $p < .001$ ,  $\eta_p^2 = .27$ ;  $F_2(2, 10) = 4.24$ ,  $MSE = 2.83$ ,  $p = .047$ ,  $\eta_p^2 = .46$ ) et un effet principal du niveau scolaire, avec plus d'items orthographiés avec des doublets fréquents en CM2 qu'en CE2 (24.3% vs. 6.9%,  $F_1(1, 46) = 11.56$ ,  $MSE = .37$ ,  $p = .001$ ,  $\eta_p^2 = .20$ ;  $F_2(1, 5) = 15.86$ ,  $MSE = 1.09$ ,

$p = .011$ ,  $\eta_p^2 = .76$ ). L'interaction n'était pas significative ( $F_1(2, 92) = 2.49$ ,  $MSE = .18$ ,  $p = .089$ ,  $\eta_p^2 = .05$ ;  $F_2(2, 10) = 1.10$ ,  $MSE = 1.61$ ,  $p = .37$ ,  $\eta_p^2 = .18$ ). Les comparaisons planifiées indiquent que les items *sans doublet* étaient moins souvent orthographiés avec un doublet fréquent que les items *doublet fréquent* et *doublet rare* (respectivement,  $F_1(1, 46) = 29.57$ ,  $MSE = .20$ ,  $p < .001$ ,  $\eta_p^2 = .39$ ;  $F_2(1, 5) = 6.15$ ,  $MSE = 3.90$ ,  $p = .05$ ,  $\eta_p^2 = .55$  et  $F_1(1, 46) = 10.10$ ,  $MSE = .15$ ,  $p = .003$ ,  $\eta_p^2 = .18$ ;  $F_2(1, 5) = 15.00$ ,  $MSE = .40$ ,  $p = .012$ ,  $\eta_p^2 = .75$ ), ce qui montre que les élèves avaient mémorisé, dans une certaine mesure, les items qui incluaient un doublet et ceux qui n'en incluaient pas. La différence entre les items *doublet fréquent* et *doublet rare* était significative seulement pour l'analyse sur les sujets ( $F_1(1, 46) = 8.11$ ,  $MSE = .18$ ,  $p = .007$ ,  $\eta_p^2 = .15$ ;  $F_2(1, 5) = 1.43$ ,  $MSE = 4.20$ ,  $p = .29$ ,  $\eta_p^2 = .22$ ).

#### Groupe expérimental, condition reconnaissance

Aux deux niveaux scolaires, les orthographes correctes étaient plus souvent choisies pour les items *sans doublet* que pour les items *doublet fréquent* et les orthographes correctes étaient beaucoup plus souvent choisies pour ces deux types d'items que pour les items *doublet rare*, comme le montre le Tableau 9. Le nombre de sélections de l'orthographe correcte a été soumis à une ANOVA avec les variables niveau scolaire (CE2, CM2) et type d'item (*sans doublet*, *doublet fréquent* et *doublet rare*). Il y avait un effet principal du type d'item ( $F_1(2, 92) = 22.39$ ,  $MSE = .45$ ,  $p < .001$ ,  $\eta_p^2 = .33$ ;  $F_2(2, 10) = 39.43$ ,  $MSE = 1.03$ ,  $p < .001$ ,  $\eta_p^2 = .89$ ). L'effet du niveau scolaire, avec davantage de sélections de l'orthographe correcte en CM2 qu'en CE2 était significatif dans l'analyse sur les sujets ( $F_1(1, 46) = 6.37$ ,  $MSE = .31$ ,  $p = .015$ ,  $\eta_p^2 = .12$ ) et marginalement significatif dans l'analyse sur les items ( $F_2(1, 5) = 4.73$ ,  $MSE = 1.69$ ,  $p = .081$ ,  $\eta_p^2 = .49$ ) et il n'y avait pas d'interaction niveau scolaire x type d'item ( $ps > .82$ ). Les comparaisons planifiées montraient que les orthographes correctes étaient significativement moins souvent choisies pour les items *doublet rare* que pour les items *sans doublet* et *doublet fréquent* ( $F_1(1, 46) = 31.63$ ,  $MSE = .58$ ,  $p < .001$ ,  $\eta_p^2 = .41$ ;  $F_2(1, 5) = 144.03$ ,  $MSE = .51$ ,  $p < .001$ ,  $\eta_p^2 = .97$ ). Le plus fort taux de sélection des orthographes correctes pour les items *sans doublet* que pour les items *doublet fréquent* était significatif dans l'analyse sur les sujets ( $F_1(1, 46) = 5.50$ ,  $MSE = .32$ ,  $p = .023$ ,  $\eta_p^2 = .11$ ) et marginalement significatif dans l'analyse par item ( $F_2(1, 5) = 4.57$ ,  $MSE = 1.54$ ,  $p = .085$ ,  $\eta_p^2 = .48$ ). En fait, pour les deux niveaux scolaires, le taux de sélection de l'orthographe correcte ne différait pas du hasard (1/3) pour les items *doublet rare* ( $t(23) = 1.04$ ,  $p = .31$  en

CE2 et  $t(23) = .27, p = .79$  en CM2), alors qu'ils étaient significativement au-dessus du hasard pour les items *sans doublet* ( $t(23) = 6.03, p < .001$  en CE2 et  $t(23) = 10.17, p < .001$  en CM2) et pour les items *doublet fréquent* ( $t(23) = 3.03, p = .006$  en CE2 et  $t(23) = 5.36, p < .001$  en CM2).

Le Tableau 9 montre qu'aux deux niveaux scolaires, les enfants effectuaient plus de choix qui impliquaient des transpositions pour les items *doublet rare* que pour les items *doublet fréquent*. Une ANOVA sur le nombre d'erreurs de transposition avec les variables niveau scolaire (CE2, CM2) et type d'item (*doublet fréquent* et *doublet rare*) confirme l'effet du type d'item ( $F_1(1, 46) = 18.13, \text{MSE} = .30, p < .001, \eta_p^2 = .41$  ;  $F_2(1, 5) = 13.42, \text{MSE} = 1.64, p = .01, \eta_p^2 = .73$ ) et ne révèle pas d'effet du niveau scolaire ( $ps > .20$ ) ni d'interaction ( $ps > .85$ ). Comme le montre le Tableau 8, les élèves commettant des erreurs de transposition seulement sur les items *doublet rare* étaient beaucoup plus nombreux que ceux commettant des erreurs de transposition seulement sur les items *doublet fréquent* aux deux niveaux scolaires.

La prévalence des erreurs de transposition sur les items *doublet rare* ne résultait pas d'une préférence générale pour les items incluant un doublet fréquent, indépendamment du type d'item présenté. En effet, les orthographe avec un doublet fréquent étaient sélectionnées pour les items *doublet rare* (37.5% d'erreurs de transposition) moins souvent que pour les items *doublet fréquent* (62.5% de choix corrects) et plus souvent que pour les items *sans doublet* (14.6% d'erreurs d'addition). Une ANOVA sur le nombre de sélections des orthographe incluant un doublet fréquent avec les variables niveaux scolaires (CE2, CM2) et type d'item (*sans doublet*, *doublet fréquent* et *doublet rare*) révèle un effet principal du type d'item ( $F_1(2, 92) = 29.94, \text{MSE} = .37, p < .001, \eta_p^2 = .39$  ;  $F_2(2, 10) = 23.08, \text{MSE} = 1.91, p < .001, \eta_p^2 = .82$ ) mais pas d'effet du niveau scolaire ( $ps > .29$ ) et pas d'interaction ( $ps > .16$ ). Des comparaisons planifiées indiquent que les choix d'orthographe incluant un doublet fréquent étaient moins nombreux pour les items *sans doublet* que pour les items *doublet fréquent* et *doublet rare* (respectivement,  $F_1(1, 46) = 56.59, \text{MSE} = .39, p < .001, \eta_p^2 = .55$ ;  $F_2(1, 5) = 240.45, \text{MSE} = .36, p < .001, \eta_p^2 = .98$  et  $F_1(1, 46) = 16.61, \text{MSE} = .30, p < .001, \eta_p^2 = .27$ ;  $F_2(1, 5) = 6.80, \text{MSE} = 2.97, p = .048, \eta_p^2 = .58$ ), ce qui confirme que les élèves avaient une certaine mémoire des items qui contenaient un doublet et de ceux qui n'en contenaient pas. De plus, les comparaisons planifiées montraient que les orthographe avec un doublet fréquent étaient plus souvent choisies pour les items *doublet fréquent* que pour les items *doublet rare* ( $F_1(1, 46) = 14.56, \text{MSE} = .41, p < .001, \eta_p^2 = .24$ ;  $F_2(1, 5) = 10.00, \text{MSE} = 2.40, p = .025, \eta_p^2 = .67$ ).

Tableau 9

Pourcentages des différents types d'orthographe sélectionnées par le groupe expérimental dans la condition reconnaissance de l'Expérience 4. Les orthographe correctes sont indiquées en gras et les erreurs de transposition en italique ; les écart-types sont entre parenthèses.

	CE2				CM2			
	Items	Items	Items	Items	Items	Items	Items	Items
Orthographe sans doublet	<i>sans</i>	<i>doublet</i>	<i>doublet</i>	<i>sans</i>	<i>doublet</i>	<i>doublet</i>	<i>doublet</i>	<i>doublet</i>
	<b>68.8</b>	27.1	31.3	<b>83.3</b>	20.8	31.3	20.8	31.3
Orthographe avec un doublet fréquent	<i>doublet</i>	<i>fréquent</i>	<i>rare</i>	<i>doublet</i>	<i>fréquent</i>	<i>rare</i>	<i>fréquent</i>	<i>rare</i>
	<b>(28.8)</b>	(32.9)	(41.2)	<b>(24.1)</b>	(29.2)	(41.2)	(29.2)	(41.2)
Orthographe avec un doublet rare	18.8	<b>56.3</b>	41.7	10.4	<b>68.8</b>	33.3	<b>32.3</b>	33.3
	(24.7)	<b>(37.0)</b>	(31.9)	(20.7)	<b>(32.3)</b>	(31.9)	(31.9)	(31.9)
Orthographe avec un doublet rare	12.5	16.7	<b>27.1</b>	6.3	10.4	<b>35.4</b>	10.4	<b>35.4</b>
	(22.1)	(24.1)	<b>(29.4)</b>	(16.9)	(20.7)	<b>(37.5)</b>	(20.7)	<b>(37.5)</b>

Alors que les erreurs de transposition étaient presque restreintes aux items *doublet fréquent*, le nombre d'omissions ne différait pas en fonction du type d'item. Une ANOVA sur le nombre d'omissions avec les variables niveaux scolaires (CE2, CM2) et type d'item (*doublet fréquent* et *doublet rare*) ne révélait pas d'effets significatifs ( $ps > .31$ ).

### Groupe contrôle

Comme le montre le Tableau 10, les élèves du groupe contrôle n'incluaient presque jamais des doublets dans leurs orthographes dans la tâche de dictée. En fait, seulement trois enfants de CE2 et six enfants de CM2 produisaient des orthographes qui incluaient des doubles consonnes. Lorsqu'ils utilisaient des doublets, ils doublaient davantage des consonnes qui sont fréquemment doublées en français que des consonnes qui le sont rarement : deux élèves de CE2 et quatre élèves de CM2 utilisaient seulement des doublets fréquents alors qu'un élève de chaque niveau scolaire utilisait seulement des doublets rares. Les élèves du groupe contrôle ne doublaient jamais une lettre qui ne peut pas être doublée en français et n'utilisaient jamais un doublet légal dans une position illégale.

Dans la tâche de jugement, les élèves du groupe contrôle choisissaient les items qui n'incluaient pas de doubles consonnes beaucoup plus souvent que les items qui incluaient un doublet, même si cette différence était moindre que dans la tâche de dictée. Alors que la plupart des élèves n'utilisaient jamais les consonnes doubles dans la tâche de dictée, un seul élève de chaque niveau scolaire ne sélectionnait jamais une orthographe contenant un doublet dans la tâche de jugement. Quand les élèves sélectionnaient une orthographe incluant un doublet, c'était plus souvent une consonne qui est fréquemment doublée en français qu'une consonne qui l'est rarement. Aux deux niveaux scolaires, le taux de sélection des orthographes n'incluant pas de doublet était significativement au-dessus du hasard ( $t(9) = 6.47, p < .001$  en CE2;  $t(19) = 7.96, p < .001$  en CM2) alors que le taux de sélection des orthographes incluant un doublet rare était significativement en dessous du hasard ( $t(9) = -8.30, p < .001$  en CE2 ;  $t(19) = -16.17, p < .001$  en CM2). Le taux de sélection des orthographes incluant un doublet fréquent ne différait pas du hasard ( $t(9) = -1.14, p = .27$  en CE2 ;  $t(19) = -1.67, p = .11$  en CM2).

Tableau 10

Pourcentage des différents types d'orthographe produites dans la tâche de dictée et sélectionnées dans la tâche de jugement par le groupe contrôle dans l'Expérience 4 ; les écart-types sont entre parenthèses.

	Dictée		Jugement	
	CE2	CM2	CE2	CM2
Orthographe sans doublet	85.0 (14.2)	84.2 (13.8)	61.7 (19.6)	66.7 (18.7)
Orthographe avec un doublet fréquent	1.7 (5.1)	5.8 (11.2)	28.3 (19.6)	28.3 (13.4)
Orthographe avec un doublet rare	0.8 (3.7)	1.7 (5.1)	10.0 (12.6)	5.0 (7.8)

Note. Dans la tâche de dictée, la somme des pourcentages des trois types d'orthographe diffère de 100% car la représentation phonologique d'au moins une consonne cible était incorrecte pour 12.5% des orthographe en CE2 et 8.3% en CM2.

## Discussion

Des élèves français de CE2 et CM2 avaient plus de difficultés à rappeler des orthographe contenant des doublets fréquents (e.g., *tidunnar*) ou des doublets rares (e.g., *tiddunar*) que des orthographe ne contenant pas de doublet (e.g., *tidunar*), même s'ils ont vu autant de fois ces trois types d'orthographe. Les difficultés d'apprentissage des orthographe contenant des doublets, en particulier des doublets rares, étaient manifestes dans la condition reconnaissance où les élèves des deux niveaux scolaires choisissaient plus souvent l'orthographe correcte pour les items *doublet fréquent* et *sans doublet* (les deux scores au-dessus du hasard) que pour les items *doublet rare* (scores ne différant pas du hasard). Les erreurs d'omission étaient aussi fréquentes pour les items *doublet fréquent* que pour les items *doublet rare* aux deux niveaux scolaires et dans les deux conditions de test (rappel et reconnaissance).

Contrairement aux erreurs d'omissions, les erreurs de transposition étaient beaucoup plus nombreuses pour les items *doublet rare* dans les deux conditions de test. Cet effet ne reflétait pas une tendance générale à utiliser des doublets fréquents puisque les élèves

utilisaient (condition rappel) et choisissaient (condition reconnaissance) les doublets fréquents significativement plus souvent pour les items qu'ils avaient vus avec un doublet rare (erreurs de transposition sur les items *doublet rare*) que pour les items qu'ils avaient vus sans doublet (erreurs d'addition sur les items *sans doublet*).

La prévalence des erreurs de transposition sur les items *doublet rare*, leur rareté sur les items *doublet fréquent*, et la rare utilisation de doublets fréquents pour les items *sans doublet* suggèrent que les élèves se rappelaient parfois de la présence d'un doublement mais pas de la lettre spécifique qui devait être doublée. Dans ce cas, comme les adultes des trois premières expériences rapportées dans cette partie expérimentale, ils se fondaient parfois sur leur connaissance du fait que certaines lettres sont plus fréquemment doublées que d'autres et surviennent seulement dans certaines positions pour reconstruire une orthographe. Pour les items *doublet fréquent*, cela conduisait à une orthographe correcte mais pour les items *doublet rare*, cela conduisait à des erreurs de transposition.

La connaissance de certaines régularités de l'emploi des doubles lettres apparaissait également dans le groupe contrôle. Dans la dictée, la plupart des orthographes produites par ce groupe incluait seulement des consonnes simples. Toutefois, ceux qui incluaient des doublets utilisaient toujours des doublets qui sont possibles en français (et des doublets fréquents plutôt que des doublets rares) dans une position légale. Dans la tâche de jugement, le groupe contrôle avait également une préférence pour les orthographes ne contenant pas de doublets (environ deux tiers de leurs choix) mais lorsqu'ils choisissaient une orthographe contenant un doublet, c'était beaucoup plus souvent un doublet fréquent, en CE2 et CM2. Certains résultats étaient néanmoins différents selon que l'apprentissage orthographique était évalué avec une tâche de rappel ou de reconnaissance. Ces différences provenaient du fait que beaucoup de CE2, qui typiquement utilisaient les correspondances graphème-phonème les plus fréquentes et les plus simples telles que *n* pour /n/ dans la tâche de dictée, n'hésitaient pas à choisir des orthographes incluant des doublets. Du fait que les CE2 incluaient très rarement des doublets dans leurs productions orthographiques, ils écrivaient très souvent correctement les items *sans doublet* mais commettaient souvent des erreurs d'omission sur les items *doublet fréquent* et *doublet rare*. Pour cette même raison, ils commettaient très peu d'erreurs de transposition sur les items *doublet fréquent* et *doublet rare*. Ceci explique pourquoi les erreurs d'omission étaient aussi nombreuses dans la condition reconnaissance en CE2 qu'en CM2 mais étaient plus nombreuses dans la condition rappel en CE2 (environ deux tiers des erreurs sur les items requérant un doublet) qu'en CM2 (environ la moitié des erreurs sur les items requérant un doublet). Ceci explique aussi pourquoi, dans la condition reconnaissance, la prévalence des erreurs de

transposition sur les items *doublet rare* était observée aux deux niveaux scolaires, avec 71% des erreurs de transposition en CE2 et 84% des erreurs de transposition en CM2 commises sur les items *doublet rare*, alors que, dans la condition rappel, la prévalence des erreurs de transposition sur les items *doublet rare* (81% des erreurs de transposition) était trouvée uniquement en CM2.

L'impact de la tâche utilisée pour évaluer l'apprentissage orthographique, plutôt que des différences de groupe, est étayée par le fait que les élèves de la condition rappel et ceux de la condition reconnaissance avaient des niveaux similaires en lecture et orthographe, ainsi que par les performances du groupe contrôle. En effet, les mêmes élèves effectuaient la tâche de dictée et la tâche de jugement dans le groupe contrôle et ils choisissaient les orthographe incluant des doublets (principalement des doublets fréquents) beaucoup plus souvent qu'ils ne les produisaient.

Dans l'expérience 4, les enfants lisaient les histoires silencieusement pour les comprendre et répondre à des questions sur celles-ci. Comme nous l'avons signalé à propos de l'expérience de Fayol et al. (2010), un inconvénient de cette tâche est qu'elle ne fournit aucune preuve que les consonnes sont prononcées de la même façon en format simple ou double. C'est pourquoi, comme les adultes des Expériences 1 et 2, les élèves de l'Expérience 5 devaient lire à voix haute les histoires dans lesquelles étaient insérés les mêmes non-mots que dans l'Expérience 2 chez l'adulte (échantillon de non-mots plus varié que dans les Expériences 1 et 4). Notre hypothèse était que les élèves, comme les adultes, prononceraient de la même façon les consonnes en format simple et double et que les résultats, notamment le pattern d'erreurs, devrait être similaire à celui rapporté dans l'Expérience 4.



## Expérience 5

### Lecture à haute voix de textes incluant des non-mots chez des CE2 et CM2

#### Méthode

#### Participants

Les participants étaient 54 élèves de CE2 (26 filles; âge = 8 ans, 6 mois; EC = 4.5 mois) et 54 élèves de CM2 (30 filles ; âge moyen = 10 ans, 5 mois; EC = 5.7 mois), tous de langue maternelle française. Pour chaque niveau scolaire, 27 élèves étaient assignés à la condition rappel (dictée) et 27 à la condition reconnaissance.

Le matériel était constitué de 18 non-mots trisyllabiques incluant deux consonnes cibles, l'une *n*, *t*, ou *r* (fréquemment doublée en français), l'autre *b*, *d*, ou *g* (rarement doublée). Chacune des neuf combinaisons de ces deux types de lettres était utilisée dans deux non-mots (e.g., *cinabé* et *cibané* pour la combinaison *b/n*) et trois orthographes, se prononçant toutes de la même façon selon les règles de correspondances graphème- phonème du français, étaient créées pour chaque non-mot : une orthographe sans doublet (e.g., *cinabé*, *cibané*), une orthographe avec un doublet fréquent (*cinnabé*, *cibanné*) et une orthographe avec un doublet rare (*cinabbé*, *cibbané*). Ces non-mots étaient insérés dans des histoires. Pour le test de reconnaissance, pour chaque non-mot, trois orthographes qui ne différaient que par le format des consonnes cibles étaient écrites sur une page d'un livret (e.g., *cibané*, *cibanné*, *cibbané*).

#### Procédure

Les élèves, testés individuellement, lisaient à haute voix une histoire, tournaient la page pour répondre aux questions situées sur la page suivante sans revenir en arrière pour lire l'histoire et ainsi de suite jusqu'à la dernière histoire du livret. Leur lecture était enregistrée. Ensuite, les participants effectuaient une tâche de barrage pendant une dizaine de minutes puis, dans la condition rappel, l'expérimentateur leur dictait chacun des six non-mots et leur demandait de les écrire comme ils étaient écrits dans les histoires qu'ils venaient de lire. La procédure pour la condition reconnaissance différait seulement au niveau du test au cours duquel les participants devaient indiquer laquelle de trois orthographes était utilisée dans les textes qu'ils venaient de lire. Ensuite, les élèves passaient le test de lecture du Lobrot (1967) et

la dictée « le corbeau » du L2MA (Chevrier-Muller, 1997).

## **Résultats**

Les scores aux tests standards de lecture et orthographe et le nombre de réponses correctes aux questions sur les histoires, qui apparaissent au Tableau 6, ont été soumis à des ANOVAs avec les variables niveau scolaire (CE2, CM2) et groupe (rappel, reconnaissance). Les scores en lecture étaient meilleurs en CM2 qu'en CE2 ( $F(1, 104) = 93.91$ ,  $MSE = 31.58$ ,  $p < .001$ ,  $\eta_p^2 = .47$ ) et il n'y avait pas d'effet de la condition ( $p = .43$ ) ni d'interaction ( $p = .13$ ). Les scores en orthographe étaient meilleurs en CM2 qu'en CE2 ( $F(1, 104) = 62.72$ ,  $MSE = 62.72$ ,  $p < .001$ ,  $\eta_p^2 = .38$ ). L'effet de la condition, qui était marginalement significatif ( $F(1, 104) = 3.36$ ,  $MSE = 12.76$ ,  $p = .07$ ,  $\eta_p^2 = .03$ ), interagissait avec le niveau scolaire ( $F(1, 104) = 6.41$ ,  $MSE = 12.76$ ,  $p = .012$ ,  $\eta_p^2 = .06$ ) parce que les scores étaient meilleurs en CM2 qu'en CE2 dans les deux conditions mais la différence était moindre dans la condition reconnaissance ( $F(1, 52) = 55.73$ ,  $MSE = 12.51$ ,  $p < .001$ ,  $\eta_p^2 = .52$ ) que dans la condition rappel ( $F(1, 52) = 14.23$ ,  $MSE = 13.01$ ,  $p < .001$ ,  $\eta_p^2 = .21$ ). Les scores aux questions sur les histoires étaient meilleurs en CM2 qu'en CE2 ( $F(1, 104) = 12.08$ ,  $MSE = 2.58$ ,  $p < .001$ ,  $\eta_p^2 = .10$ ). Il n'y avait pas d'effet de la condition ( $p = .90$ ) mais une interaction ( $F(1, 104) = 12.08$ ,  $MSE = 2.58$ ,  $p < .001$ ,  $\eta_p^2 = .10$ ) car les scores, qui étaient exactement les mêmes en CE2 et CM2 dans la condition rappel, étaient meilleurs en CM2 qu'en CE2 dans la condition reconnaissance ( $F(1, 52) = 30.9$ ,  $MSE = 2.01$ ,  $p < .001$ ,  $\eta_p^2 = .37$ ).

Comme le montre le Tableau 6, même si les élèves ne lisaient pas les non-mots parfaitement (82.7% en CE2 et 93.7% en CM2), il n'y avait aucune indication que les élèves prononçaient différemment les orthographes incluant des consonnes simples et celles incluant des consonnes doubles. Une ANOVA sur le nombre de non-mots correctement décodés avec les variables niveau scolaire (CE2, CM2), type de test (rappel, reconnaissance), et type d'item type (*sans doublet*, *doublet fréquent* et *doublet rare*) révèle que les décodages corrects étaient plus nombreux en CM2 qu'en CE2 (94.1% vs. 82.5%,  $F_1(1, 104) = 21.75$ ,  $MSE = 4.96$ ,  $p < .001$ ,  $\eta_p^2 = .17$  ;  $F_2(1, 8) = 57.37$ ,  $MSE = .14$ ,  $p < .001$ ,  $\eta_p^2 = .88$ ) et dans la condition reconnaissance que dans la condition rappel (92.0% vs. 84.6%,  $F_1(1, 104) = 8.81$ ,  $MSE = 4.96$ ,  $p = .004$ ,  $\eta_p^2 = .72$  ;  $F_2(1, 8) = 20.63$ ,  $MSE = .21$ ,  $p = .002$ ,  $\eta_p^2 = .72$ ). Toutefois, il n'y avait pas d'effet du type d'item ( $ps > .43$ ) ni d'interaction ( $ps > .38$ ).

Ainsi, les scores aux tests de lecture et orthographe, les scores aux questions sur les histoires et le niveau de décodage étaient meilleurs en CM2 qu'en CE2 dans les deux conditions de test à l'exception des questions sur les histoires pour lesquelles les scores étaient similaires en CE2 et CM2 dans la condition rappel. Même si les non-mots étaient plus souvent décodés correctement dans la condition reconnaissance que dans la condition rappel, le résultat important était l'absence de différence en fonction du type d'item qui montre que les élèves ne pouvaient pas se souvenir de la présence d'un doublet dans un non-mot parce qu'ils prononçaient différemment les non-mots selon qu'ils incluaient ou non un doublet.

### Condition rappel

Les différentes orthographes produites pour les trois types d'items sont indiquées au Tableau 11. Les orthographes avec les deux consonnes cibles orthographiées comme dans l'histoire étaient plus nombreuses pour les items *sans doublet* que pour les items requérant un doublet et, parmi ces derniers, elles étaient plus nombreuses pour les items *doublet fréquent* que pour les items *doublet rare*. Une ANOVA sur le nombre d'orthographes correctes avec les variables niveau scolaire (CE2, CM2) et type d'item (*sans doublet*, *doublet fréquent*, *doublet rare*) révèle un effet principal du type d'item ( $F_1(2, 104) = 75.07$ ,  $MSE = .37$ ,  $p < .001$ ,  $\eta_p^2 = .59$  ;  $F_2(2, 16) = 116.60$ ,  $MSE = .02$ ,  $p < .001$ ,  $\eta_p^2 = .94$ ) et un effet principal du niveau scolaire, avec de meilleurs scores en CM2 qu'en CE2 (46.9% vs. 36.4%,  $F_1(1, 52) = 5.13$ ,  $MSE = .35$ ,  $p = .028$ ,  $\eta_p^2 = .09$  ;  $F_2(1, 8) = 14.80$ ,  $MSE = .01$ ,  $p = .005$ ,  $\eta_p^2 = .65$ ) et une absence d'interaction ( $ps > .16$ ). Les comparaisons planifiées montrent que les orthographes correctes étaient significativement plus nombreuses pour les items *sans doublet* que pour les items *doublet fréquent* et *doublet rare* (respectivement  $F_1(1, 52) = 81.25$ ,  $MSE = .41$ ,  $p < .001$ ,  $\eta_p^2 = .61$  ;  $F_2(1, 8) = 135.64$ ,  $MSE = .02$ ,  $p < .001$ ,  $\eta_p^2 = .94$  et  $F_1(1, 52) = 143.39$ ,  $MSE = .33$ ,  $p < .001$ ,  $\eta_p^2 = .73$  ;  $F_2(1, 8) = 283.16$ ,  $MSE = .01$ ,  $p < .001$ ,  $\eta_p^2 = .97$ ). La différence entre les items *doublet fréquent* et *doublet rare* était marginalement significative pour l'analyse sur les sujets ( $F_1(1, 52) = 3.74$ ,  $MSE = .36$ ,  $p = .058$ ,  $\eta_p^2 = .07$ ) et significative pour l'analyse sur les items ( $F_2(1, 8) = 5.96$ ,  $MSE = .03$ ,  $p = .04$ ,  $\eta_p^2 = .43$ ).

Tableau 11

Pourcentages des différents types d'orthographe produites dans la condition rappel de l'Expérience 5. Les orthographe correctes sont indiquées en gras et les erreurs de transposition en italique ; les écart-types sont entre parenthèses.

	CE2			CM2		
	Items	Items	Items	Items	Items	Items
Orthographe sans doublet	<b>83.3 (24.0)</b>	74.1 (32.1)	70.4 (37.4)	<b>81.5 (28.2)</b>	55.6 (42.4)	44.4 (42.4)
Orthographe avec un doublet fréquent	<i>sans doublet</i>	<i>doublet</i>	<i>doublet</i>	<i>sans doublet</i>	<i>doublet</i>	<i>doublet</i>
	0.0 (0.0)	<b>16.7 (31.0)</b>	5.6 (16.0)	9.3 (19.8)	<b>37.0 (35.6)</b>	29.6 (34.7)
Orthographe avec un doublet rare	3.7 (13.3)	3.7 (13.3)	<b>9.3 (19.8)</b>	1.9 (9.6)	7.4 (18.1)	<b>22.2 (37.6)</b>
		<i>fréquent</i>	<i>rare</i>		<i>fréquent</i>	<i>rare</i>

Note. La somme des pourcentages des trois types d'orthographe diffère de 100% parce que 10.5% des erreurs en CE2 et 3.7% des erreurs en CM2 correspondent à une représentation phonologique incorrecte de la consonne cible et parce que 1.9% des items avec un doublet rare en CE2 étaient orthographiés avec les deux consonnes cibles doublées<sup>3</sup>.

<sup>3</sup> Une ANOVA avec les variables niveau scolaire (CE2 et CM2) et type d'item (*sans doublet*, *doublet fréquent* et *doublet rare*) sur le nombre d'orthographe avec une représentation phonologiquement incorrecte d'au moins une des consonnes cibles montre que ces erreurs sont plus nombreuses en CE2 qu'en CM2 ( $F_1(2, 92) = 16.78$ ,  $MSE = .18$ ,  $p < .001$ ,  $\eta_p^2 = .27$ ;  $F_2(2, 10) = 4.24$ ,  $MSE = 2.83$ ,  $p = .047$ ,  $\eta_p^2 = .46$ ) mais ne révèle pas d'effet du type d'item ( $ps > .10$ ), ni d'interaction ( $ps > .83$ ).

Le Tableau 11 montre que les élèves de CM2 commettaient davantage d'erreurs de transposition que les élèves de CE2 (18.5% vs. 4.6%). Alors que les erreurs de transposition étaient très rares à la fois pour les items *doublet fréquent* et *doublet rare* en CE2, les erreurs de transposition étaient beaucoup plus nombreuses sur les items *doublet rare* que sur les items *doublet fréquent* en CM2. Une ANOVA sur le nombre d'erreurs de transposition avec la variable niveau scolaire (CE2, CM2) et type d'item (*doublet fréquent*, *doublet rare*) confirme l'effet principal du type d'item ( $F_1(1, 52) = 6.89$ ,  $MSE = .23$ ,  $p = .011$ ,  $\eta_p^2 = .12$ ;  $F_2(1, 8) = 7.61$ ,  $MSE = .02$ ,  $p = .025$ ,  $\eta_p^2 = .89$ ), l'effet principal du niveau scolaire ( $F_1(1, 52) = 12.55$ ,  $MSE = .17$ ,  $p < .001$ ,  $\eta_p^2 = .19$ ;  $F_2(1, 8) = 48.53$ ,  $MSE = .004$ ,  $p < .001$ ,  $\eta_p^2 = .89$ ) et l'interaction ( $F_1(1, 52) = 4.93$ ,  $MSE = .23$ ,  $p = .03$ ,  $\eta_p^2 = .09$ ;  $F_2(1, 8) = 8.01$ ,  $MSE = .01$ ,  $p = .022$ ,  $\eta_p^2 = .89$ ). Cette interaction est due au fait que les erreurs de transposition étaient plus fréquentes pour les items *doublet rare* que pour les items *doublet fréquent* en CM2 ( $F_1(1, 26) = 7.43$ ,  $MSE = .36$ ,  $p = .011$ ,  $\eta_p^2 = .22$ ;  $F_2(1, 8) = 11.11$ ,  $MSE = .02$ ,  $p = .01$ ,  $\eta_p^2 = .58$ ) alors qu'il n'y avait pas de différence significative en CE2 ( $ps > .49$ ). Comme le montre le Tableau 8, ce pattern de résultats était confirmé par l'analyse de profils individuels : Davantage d'élèves commettaient des erreurs de transpositions en CM2 qu'en CE2 (16 vs. 5 sur 27). La plupart des élèves commettant des erreurs de transpositions les commettaient uniquement sur les items *doublet rare*, en particulier en CM2 (12 en CM2 ; 3 en CE2).

Comme dans les expériences précédentes, la prédominance des erreurs de transposition pour les items *doublet rare* n'était pas due à une tendance à utiliser des doublets fréquents indépendamment du type d'item présenté. En effet, les items *doublet rare* étaient orthographiés avec un doublet fréquent (17.6% d'erreurs de transposition) à la fois moins souvent que les items *doublet fréquent* (26.9% de correct) et plus souvent que les items *sans doublet* (4.6% d'erreurs d'addition qui impliquent le doublement d'une des consonnes cibles pour les items *sans doublet*). Une ANOVA sur l'utilisation de doublets fréquents avec les variables niveau scolaire (CE2, CM2) et type d'item (*sans doublet*, *doublet fréquent*, *doublet rare*) révèle un effet principal du type d'item ( $F_1(2, 104) = 13.13$ ,  $MSE = .20$ ,  $p < .001$ ,  $\eta_p^2 = .20$ ;  $F_2(2, 16) = 14.49$ ,  $MSE = .23$ ,  $p < .001$ ,  $\eta_p^2 = .99$ ), un effet principal du niveau scolaire, avec les doublets fréquents plus souvent utilisés en CM2 qu'en CE2 (25.3% vs. 7.4%,  $F_1(1, 52) = 12.77$ ,  $MSE = .41$ ,  $p < .001$ ,  $\eta_p^2 = .20$ ;  $F_2(1, 8) = 29.11$ ,  $MSE = .01$ ,  $p < .001$ ,  $\eta_p^2 = .85$ ) mais pas d'interaction ( $ps > .21$ ). Les comparaisons planifiées indiquent que les items *sans doublet* étaient moins souvent orthographiés avec un doublet fréquent que les items *doublet fréquent* et *doublet rare* (respectivement,  $F_1(1, 52) = 22.49$ ,  $MSE = .24$ ,

$p < .001$ ,  $\eta_p^2 = .30$  ;  $F_2(1, 8) = 21.14$ ,  $MSE = .02$ ,  $p < .001$ ,  $\eta_p^2 = .95$  et  $F_1(1, 52) = 10.98$ ,  $MSE = .16$ ,  $p = .002$ ,  $\eta_p^2 = .17$  ;  $F_2(1, 8) = 27.90$ ,  $MSE = .006$ ,  $p < .001$ ,  $\eta_p^2 = .97$ ), ce qui confirme que les enfants ont mémorisé, dans une certaine mesure, quels items incluait un doublet et quels items n'en incluait pas. Concernant la différence entre les items *doublet fréquent* et *doublet rare*, la plus forte utilisation des doublets fréquents pour les premiers était significative pour l'analyse sur les sujets ( $F_1(1, 52) = 4.33$ ,  $MSE = .21$ ,  $p = .04$ ,  $\eta_p^2 = .08$ ) et marginalement significative pour l'analyse sur les items ( $F_2(1, 8) = 3.94$ ,  $MSE = .02$ ,  $p = .08$ ,  $\eta_p^2 = .80$ ).

Les erreurs d'omission (un item *doublet fréquent* ou *doublet rare* orthographié sans doublet) étaient de loin les plus nombreuses en CE2 pour les items *doublet fréquent* et *doublet rare*. Les erreurs d'omission étaient moins nombreuses en CM2, même si elles demeuraient très nombreuses, avec environ la moitié des items vus avec un doublet qui étaient orthographiés sans doublet. Une ANOVA sur le nombre d'erreurs d'omission avec les variables niveau scolaire (CE2, CM2) et type d'item type (*doublet fréquent* et *doublet rare*) confirme que les CE2 commettaient plus d'erreurs d'omission que les CM2 (72.2% vs. 50.0%,  $F_1(1, 52) = 5.74$ ,  $MSE = .93$ ,  $p = .02$ ,  $\eta_p^2 = .10$  ;  $F_2(1, 8) = 41.27$ ,  $MSE = .01$ ,  $p < .001$ ,  $\eta_p^2 = .89$ ) et ne révèle pas d'effet du type d'item, ni d'interaction ( $ps > .15$ ). Enfin, les élèves ne doublaient jamais des lettres qui ne sont jamais doublées en français et ne doublaient jamais les lettres du début ou de la fin des mots.

### Condition reconnaissance

Les orthographes correctes étaient plus souvent choisies en CM2 qu'en CE2, en particulier pour les items *sans doublet* et *doublet fréquent*. Aux deux niveaux scolaires, les orthographes correctes étaient plus souvent choisies pour les items *sans doublet* que pour les items *doublet fréquent* et pour les items *doublet fréquent* que pour les items *doublet rare*, comme le montre le Tableau 12. Une ANOVA sur le nombre de sélections de l'orthographe correcte avec les variables niveau scolaire (CE2 et CM2) et type d'item (*sans doublet*, *doublet fréquent* et *doublet rare*) montre un effet principal du type d'item ( $F_1(2, 104) = 20.89$ ,  $MSE = .47$ ,  $p < .001$ ,  $\eta_p^2 = .29$  ;  $F_2(2, 16) = 14.83$ ,  $MSE = .06$ ,  $p < .001$ ,  $\eta_p^2 = .65$ ), un effet principal du niveau scolaire, avec 47.5% de choix de l'orthographe correcte en CE2 et 61.7% en CM2 ( $F_1(1, 52) = 8.08$ ,  $MSE = .40$ ,  $p = .006$ ,  $\eta_p^2 = .13$ ;  $F_2(1, 8) = 15.76$ ,  $MSE = .05$ ,  $p = .004$ ,  $\eta_p^2 = .66$ ) et pas d'interaction ( $ps > .36$ ). Les comparaisons planifiées montrent que les orthographes correctes sont significativement moins souvent choisies pour les items

*doublet rare* que pour les items *sans doublet* et *doublet fréquent* ( $F_1(1, 52) = 35.35$ ,  $MSE = .41$ ,  $p < .001$ ,  $\eta_p^2 = .40$  ;  $F_2(1, 8) = 31.58$ ,  $MSE = .04$ ,  $p < .001$ ,  $\eta_p^2 = .80$ ) et que les orthographes correctes étaient significativement plus souvent choisies pour les items *sans doublet* que pour les items *doublet fréquent* ( $F_1(1, 52) = 15.42$ ,  $MSE = .32$ ,  $p = .003$ ,  $\eta_p^2 = .23$  ;  $F_2(1, 8) = 5.65$ ,  $MSE = .09$ ,  $p = .04$ ,  $\eta_p^2 = .41$ ). En fait, le taux de sélection de l'orthographe correcte ne différait pas du hasard (1/3) pour les items *doublet rare* aux deux niveaux scolaires ( $t(26) = .28$ ,  $p = .78$  en CE2 et  $t(26) = .25$ ,  $p = .81$  en CM2). Par contraste, le taux de sélection de l'orthographe correcte était au-dessus du hasard pour les items *sans doublet* ( $t(26) = 5.38$ ,  $p < .001$  en CE2 et  $t(26) = 12.50$ ,  $p < .001$  en CM2) et pour les items *doublet fréquent* en CE2 où cet effet était marginalement significatif ( $t(26) = 1.85$ ,  $p = .076$ ) et en CM2 où cet effet était significatif ( $t(26) = 4.32$ ,  $p < .001$ ).

Table 12

Pourcentages des différents types d'orthographe sélectionnées dans la condition reconnaissance de l'Expérience 5. Les orthographe correctes sont indiquées en gras et les erreurs de transposition en italique ; les écart-types sont entre parenthèses.

	CE2			CM2		
	Items	Items	Items	Items	Items	Items
Orthographe sans doublet	<b>64.8 (30.4)</b>	37.0 (35.6)	31.5 (39.6)	<b>87.0 (22.3)</b>	29.6 (34.7)	29.6 (39.9)
Orthographe avec un doublet fréquent	20.4 (25.0)	<i>doublet</i>	<i>doublet</i>	<i>sans doublet</i>	<i>doublet</i>	<i>doublet</i>
		<i>fréquent</i>	<i>rare</i>		<i>fréquent</i>	<i>rare</i>
Orthographe avec un doublet rare	14.8 (23.3)	<i>16.7 (24.0)</i>	<b>31.5 (34.4)</b>	7.4 (18.1)	7.4 (18.1)	<b>35.2 (38.8)</b>



Le Tableau 12 montre qu'aux deux niveaux scolaires, les élèves effectuaient plus de choix impliquant des transpositions pour les items *doublet rare* que pour les items *doublet fréquent*. Une ANOVA sur le nombre d'erreurs de transposition avec les variables niveaux scolaires (CE2 et CM2) et type d'item (*doublet fréquent* et *doublet rare*) révèle un effet principal du type d'item ( $F_1(1, 52) = 20.87$ ,  $MSE = .30$ ,  $p < .001$ ,  $\eta_p^2 = .29$  ;  $F_2(1, 8) = 20.59$ ,  $MSE = .02$ ,  $p = .002$ ,  $\eta_p^2 = .72$ ) mais pas d'effet du niveau scolaire ( $ps > .31$ ) ni d'interaction ( $ps > .52$ ). Le Tableau 8 montre, qu'aux deux niveaux scolaires, les élèves commettant des erreurs de transposition seulement sur les items *doublet rare* étaient beaucoup plus nombreux que ceux commettant des erreurs de transposition seulement sur les items *doublet fréquents* (23 vs. 3 pour les deux niveaux rassemblés). Ce résultat ne reflétait pas une tendance à sélectionner les orthographes incluant un doublet fréquent, indépendamment du type d'items présentés, puisque les erreurs de transpositions pour les items *doublet rare* (36.1%) étaient à la fois moins nombreuses que les réponses correctes sur les items *doublet fréquent* (54.6%) et plus nombreuses que les erreurs d'addition pour les items *sans doublet* (13.0%). Une ANOVA sur le nombre de sélection de l'orthographe incluant un doublet fréquent avec les variables niveau scolaire (CE2, CM2) et type d'item (*sans doublet*, *doublet fréquent* et *doublet rare*) révèle un effet principal du type d'item ( $F_1(2, 104) = 24.19$ ,  $MSE = .39$ ,  $p < .001$ ,  $\eta_p^2 = .32$  ;  $F_2(2, 16) = 17.68$ ,  $MSE = .04$ ,  $p < .001$ ,  $\eta_p^2 = .69$ ) qui interagit avec le niveau scolaire ( $F_1(2, 104) = 3.47$ ,  $MSE = .39$ ,  $p = .035$ ,  $\eta_p^2 = .06$  ;  $F_2(2, 16) = 11.80$ ,  $MSE = .02$ ,  $p < .001$ ,  $\eta_p^2 = .60$ ). En CE2, l'effet du type d'item était significatif sur l'analyse sur les sujets ( $F_1(2, 52) = 4.78$ ,  $MSE = .39$ ,  $p = .012$ ,  $\eta_p^2 = .16$ ) et marginalement significatif sur l'analyse sur les items ( $F_2(2, 16) = 3.05$ ,  $MSE = .037$ ,  $p = .08$ ,  $\eta_p^2 = .28$ ). En CM2, l'effet du type d'item était significatif dans l'analyse par sujets ( $F_1(2, 52) = 22.92$ ,  $MSE = .39$ ,  $p < .001$ ,  $\eta_p^2 = .47$ ) ainsi que dans l'analyse par item ( $F_2(2, 16) = 34.84$ ,  $MSE = .02$ ,  $p < .001$ ,  $\eta_p^2 = .81$ ). Les comparaisons planifiées montrent que les choix de l'orthographe avec un doublet fréquent sont moins nombreuses pour les items *sans doublet* que pour les items *doublet fréquent* et *doublet rare* en CE2 ( $ps < .036$ ) et en CM2 ( $ps < .033$ ), ce qui confirme que les élèves avaient dans une certaine mesure mémorisé quels items incluait un doublet et quels items n'en incluait pas. La différence entre les deux niveaux scolaires concernait le choix de l'orthographe incluant un doublet fréquent pour les deux types d'items présentés avec un doublet: Ils ne différaient pas entre les items *doublet fréquent* et *doublet rare* en CE2 ( $ps > .31$ ) mais ils étaient plus nombreux pour les items *doublet fréquent* en CM2 ( $ps < .008$ ).

Alors que les erreurs de transposition étaient presque restreintes aux items *doublet fréquent*, le nombre d'omissions ne différait pas en fonction du type d'item. En effet, une

ANOVA sur le nombre d'omissions avec les variables niveau scolaire (CE2, CM2) et type d'item (*doublet fréquent* et *doublet rare*) ne révèle aucun effet ( $ps > .56$  pour l'analyse sur les sujets et  $ps > .14$  pour l'analyse sur les items).

## **Discussion**

Les résultats de l'Expérience 4 étaient répliqués dans l'Expérience 5. Premièrement, les erreurs d'omission, très fréquentes en CE2, étaient moins nombreuses en CM2. Elles étaient aussi nombreuses pour les items *doublet fréquent* que pour les items *doublet rare* aux deux niveaux scolaires. Deuxièmement, les erreurs de transposition étaient plus nombreuses sur les items *doublet rare* que sur les items *doublet fréquent* en CM2 dans la condition rappel et aux deux niveaux scolaires dans la condition reconnaissance. Troisièmement, la prévalence des erreurs de transposition sur les items *doublet rare* ne résultait pas d'une préférence générale pour les orthographes qui incluent un doublet fréquent, indépendamment du type d'items présentés, parce que les orthographes incluant des doublets fréquents étaient moins souvent produites (condition rappel) et choisies (condition reconnaissance) pour les items vus sans doublet que pour les items vus avec un doublet. Ce pattern de résultats semble robuste puisque l'Expérience 5 répliquait les résultats de l'Expérience 4 en dépit de différences de procédures entre les deux expériences: Les élèves de l'Expérience 5 lisaient les histoires à voix haute alors que ceux de l'Expérience 4 les lisaient silencieusement et l'Expérience 5 utilisait un échantillon de non-mots plus variés que l'Expérience 4. De plus, l'utilisation d'une situation de lecture à haute voix a permis de montrer que les élèves prononçaient les orthographes incluant seulement des lettres simples et celles incluant des doublets de façon similaire, et donc qu'ils ne pouvaient pas mémoriser si un item spécifique incluait ou non un doublet en se fondant sur des prononciations différentes.

## CHAPITRE EXPERIMENTAL IV

### **Influence des connaissances graphotactiques descriptibles sous la forme de règle sur l'apprentissage de l'orthographe de nouveaux mots : Etudes chez l'adulte.**

Les expériences rapportées dans le chapitre expérimental III étudiaient l'impact de la sensibilité d'adultes et d'élèves de l'école élémentaire à la fréquence de doublement des consonnes sur leur apprentissage de l'orthographe de non-mots, avec un intérêt particulier pour les erreurs de transposition qui étaient plus nombreuses d'un doublet rare vers un doublet fréquent (*tiddunar* orthographié TIDUNNAR) que d'un doublet fréquent vers un doublet rare (*tidunnar* orthographié TIDDUNAR). La propriété graphotactique manipulée était de nature probabiliste : des doublets légaux fréquents versus des doublets légaux mais rares en français. L'objectif de cette étude était d'examiner l'impact d'une autre régularité graphotactique, descriptible sous la forme d'une règle, sur l'apprentissage de l'orthographe de non-mots chez des adultes. Cette règle spécifie qu'une consonne double peut survenir avant mais non après une consonne simple en français. L'apprentissage orthographique a été évalué pour trois types d'items ne différant que par l'orthographe du groupe consonantique médian : des orthographe sans doublet (items *AB*, e.g., *guprane*), des orthographe avec la première consonne doublée, ce qui est légal en français (items *AAB*, e.g., *gupprane*) et des orthographe avec la seconde consonne doublée, ce qui est illégal en français (items *ABB*, *guprrane*). Trois conditions d'apprentissage ont été utilisées : apprentissage implicite avec les mots insérés dans des textes ; apprentissage implicite avec les non-mots présentés de façon isolée et apprentissage explicite avec les non-mots présentés de façon isolée.

Si, les sujets apprennent qu'une des consonnes du groupe consonantique d'un item est doublée, mais se fondent également sur leur connaissance du fait que les consonnes peuvent être doublées avant mais non après une consonne simple en français, alors ils devraient plus souvent doubler une des consonnes du groupe consonantique pour les items *AAB* et *ABB* que pour les items *AB* et devraient moins souvent orthographier correctement les items *ABB* que les items *AAB* et *AB*. De plus, parmi les items *AAB* et *ABB*, les erreurs de transposition sur les items *ABB*, avec transposition du doublement de la seconde vers la première consonne (e.g., *guprrane* orthographié GUPPRANE)

devraient être plus fréquentes que les erreurs de transposition sur les items *AAB*, avec transposition du doublement de la première vers la seconde consonne (e.g., *gupprane* orthographié GUPRRANE). Des erreurs d'omission devraient également être observées, avec des items *AAB* et *ABB* orthographiés sans doublet mais, contrairement aux erreurs de transposition, les omissions devraient être aussi nombreuses pour les items *AAB* et *ABB*. Sur la base d'études montrant un apprentissage de l'orthographe de nouveaux mots dans des situations d'apprentissage implicite et explicite, mais plus efficient lorsque l'apprentissage est explicite (Ormrod, 1986), nous faisons l'hypothèse que l'orthographe des différents types de non-mots sera meilleur dans la condition d'apprentissage explicite que dans les deux conditions d'apprentissage implicite. Toutefois, le même pattern d'erreurs est attendu dans les trois conditions d'apprentissage avec, en particulier beaucoup plus d'erreurs de transposition pour les items *ABB* que pour les items *AAB* et des erreurs d'omission commises dans des proportions similaires pour ces deux types d'items. Enfin, ce pattern de résultats devrait être observé que l'apprentissage orthographique soit évalué avec une épreuve de rappel (dictée) ou de reconnaissance.

## **Méthode**

### **Participants**

Soixante-douze étudiants de l'Université Paris Descartes, tous de langue maternelle française, ont participé à l'expérience. L'âge moyen était de 22 ans et 8 mois, tous compris entre 21 et 33 ans. Ils ont été répartis de façon aléatoire dans trois groupes de 24 participants correspondant aux trois conditions d'apprentissage de l'expérience : *implicite contexte*, *implicite isolé* et *explicite isolé*.

### **Matériel**

#### **Matériel pour tester l'apprentissage de l'orthographe de non-mots**

Six non-mots bi-syllabiques ont été construits. Tous incluent un groupe consonantique au début de la seconde syllabe : *dufline*, *guprane*, *mifron*, *nocrile*, *toplire*, *viclare*. Ces six groupes consonantiques sont légaux en français, aussi bien avec les deux consonnes non doublées qu'avec la première consonne doublée. Pour chaque non-mot, trois orthographes ne différant que par l'orthographe du groupe consonantique, et se

prononçant de la même façon selon les règles de correspondances graphème-phonème en français, ont été construites. Le groupe consonantique pouvait s'orthographier soit sans doublet, soit avec la seule première consonne doublée, ce qui est légal en français, soit avec la seule seconde consonne doublée, ce qui est illégal en français. Un exemple est *guprane* (item *AB*), *gupprane* (item *AAB*) et *guprrane* (item *ABB*). Chaque groupe comportant 24 participants, chaque participant lisant six non-mots et le matériel étant constitué de six non-mots, chacun orthographié de trois façons différentes, une des trois orthographes de chaque non-mot était donc présentée à huit participants.

Pour la condition *implicite contexte*, les non-mots étaient insérés dans les trois histoires utilisées pour les Expériences 1, 2, 4 et 5 du chapitre expérimental III testant l'influence de la sensibilité à la fréquence de doublement des consonnes sur l'apprentissage de l'orthographe de non-mots. Chaque non-mot apparaissait cinq fois dans chaque histoire. Les deux orthographes d'une même catégorie (i.e., *AB*, *AAB*, ou *ABB*) n'apparaissaient jamais dans une même histoire. L'ordre des histoires et les non-mots insérés dans celles-ci étaient randomisés sur l'ensemble des sujets.

Pour le test de reconnaissance, trois orthographes ne différant que par l'orthographe du groupe consonantique étaient écrites sur chaque page d'un carnet. Le groupe consonantique était écrit sans doublet dans une orthographe, avec la première consonne doublée dans une deuxième orthographe et avec la seconde consonne doublée dans une troisième orthographe (e.g., *guprane*, *gupprane*, *guprrane*).

#### Matériel pour évaluer les connaissances graphotactiques

Trente-six des 56 paires de non-mots utilisées dans le chapitre expérimental I ont été utilisées pour tester la sensibilité des sujets aux régularités graphotactiques. Dans les 12 paires *consonnes pouvant être doublées*, un non-mot incluait un doublet légal fréquent, l'autre un doublet rare (e.g., *onnave* – *ojjave*). Dans les 12 paires *doublet illégal en début de mot*, un non-mot incluait un doublet en position initiale, l'autre un doublet en position médiane. Le doublet était un doublet fréquent pour six paires (e.g., *tammir* – *ttamir*) et un doublet formé de consonnes jamais doublées pour six autres paires (e.g., *jikkol* – *jjikol*). Dans les 12 paires *doublet illégal après une consonne simple*, un non-mot incluait une consonne double avant une consonne simple, l'autre une consonne double après une consonne simple. Le groupe consonantique suivait une voyelle après laquelle il est fréquent en français pour six paires (e.g., *apprulir* – *aprrulir*) et une voyelle après laquelle il est rare en français pour six autres paires (e.g., *ipprulir* – *iprrulir*). Ces 36

paires étaient placées selon un ordre aléatoire et réparties sur trois feuilles A4, avec la réponse correcte à droite pour la moitié des items de chaque catégorie et à gauche pour l'autre moitié, juste après trois items essais (e.g., *vxyat – lytale*).

#### Matériel pour évaluer la connaissance de l'orthographe de mots spécifiques

Trente orthographes de mots étaient présentées dans un ordre aléatoire sur une page A4. Quinze orthographes étaient correctes et 15 autres étaient incorrectes mais phonologiquement plausibles (e.g., *pantin* orthographié *pantain*, voir Annexe 9).

#### **Procédure**

Les passations étaient individuelles pour les trois conditions d'apprentissage. Dans la condition *implicite contexte*, les participants lisaient à voix haute les textes puis répondaient aux questions exactement comme pour les Expériences 1, 2, 4 et 5 du chapitre expérimental III.

Dans les conditions *implicite isolé* et *explicite isolé*, les non-mots étaient présentés les uns après les autres au centre d'un écran d'ordinateur portable (Macbook 13 pouces) pendant 1.5 seconde à l'aide du logiciel Psyscope. Les participants voyaient chaque non-mot cinq fois, répartis selon un ordre aléatoire. Dans la condition *implicite isolé*, les participants devaient lire les non-mots à haute voix, sans instruction particulière. Dans la condition *explicite isolé*, ils devaient en plus essayer de mémoriser leur orthographe. Dans les trois conditions, les participants étaient enregistrés afin de pouvoir vérifier qu'ils prononçaient correctement les non-mots.

A l'issue de la phase d'apprentissage, les participants des trois groupes effectuaient une tâche de barrage pendant une dizaine de minutes puis les six non-mots leur étaient dictés, avec consigne de les orthographier comme ils l'étaient dans les histoires pour la condition *implicite contexte* et comme ils étaient présentés à l'écran pour les conditions *implicite isolé* et *explicite isolé*. Ensuite, les participants passaient le Trail Making Test, effectuaient une nouvelle épreuve de barrage pendant environ 10 minutes puis l'épreuve de choix forcé évaluant leur apprentissage de l'orthographe des non-mots. Pour cette dernière, un carnet comportant sur chaque page trois orthographes d'un des non-mots était fourni aux sujets qui devaient indiquer quelle orthographe était présentée dans la phase d'étude.

Enfin, les participants effectuaient l'épreuve évaluant l'orthographe lexicale. L'expérimentateur leur donnait la feuille A4 sur laquelle se trouvaient les orthographes

de 30 mots, 15 correctes et 15 incorrectes puis leur demandait de barrer les mots mal orthographiés.

## **Résultats**

### Connaissances en orthographe lexicale et connaissances graphotactiques

Les scores à l'épreuve évaluant le niveau en orthographe lexicale étaient équivalents pour chacune des trois conditions d'apprentissage (67.5% de réponses correctes dans la condition *explicite isolé*, 68.6% dans la condition *implicite isolé* et 71.5% dans la condition *implicite contexte*,  $p = .26$ ).

Les pourcentages de sélections des items respectant les trois régularités graphotactiques évaluées, indiqués au Tableau 1, étaient très élevés et largement au-dessus du hasard (50%) pour les trois types d'items dans chacune des trois conditions d'apprentissage ( $ts(23) > 14.13$ ,  $ps < .001$ ). Une ANOVA avec les variables condition d'apprentissage (*implicite contexte*, *implicite isolé* et *explicite isolé*) et type d'items (*doublet illégal en début de mot*, *doublet illégal après une consonne simple* et *consonnes pouvant être doublées*) sur le nombre de réponses correctes révèle un effet du type d'item ( $F_1(2, 138) = 8.02$ ,  $MSE = 1.27$ ,  $p < .001$ ,  $\eta_p^2 = .10$  ;  $F_2(2, 33) = 3.81$ ,  $MSE = 5.35$ ,  $p = .03$ ,  $\eta_p^2 = .19$ ), sans effet de la condition d'apprentissage ( $ps > .26$ ) ni interaction ( $ps > .15$ ). Les scores étaient significativement meilleurs pour les items *doublet illégal en début de mot* que pour les items *doublet illégal après une consonne simple* et *consonnes pouvant être doublées* ( $F_1(1, 69) = 23.14$ ,  $MSE = .72$ ,  $p < .001$ ,  $\eta_p^2 = .25$  ;  $F_2(1, 33) = 6.47$ ,  $MSE = .43$ ,  $p = .016$ ,  $\eta_p^2 = .16$ ), sans différence entre ces deux derniers ( $ps > .16$ )<sup>4</sup>.

Ainsi, les participants ont une très bonne connaissance des trois régularités graphotactiques testées, notamment du fait que les consonnes doubles peuvent survenir avant mais pas après les consonnes simples. De plus, le niveau des connaissances

---

<sup>4</sup> Des ANOVAs prenant en compte la familiarité du matériel pour les items *doublet illégal en début de mot* et *doublet illégal après une consonne simple* révèlent de meilleures performances lorsqu'un matériel familier est utilisé ( $F_s > 4.41$ ,  $ps < .04$ ), sans effet de la condition d'apprentissage ni interaction ( $ps > .29$ ).

graphotactiques et le niveau en orthographe lexicale ne différaient pas entre les trois conditions d'apprentissage.



Tableau 1. Pourcentages de sélections des items respectant chacune des trois régularités graphotactiques évaluées (écart-types entre parenthèses).

Régularités graphotactiques	Implicite contexte	Implicite isolé	Explicite isolé
Doublet illégal après une consonne simple *	92.0 (12.1)	93.8 (6.6)	89.6 (12.1)
Doublet illégal en début de mots **	99.3 (3.7)	96.9 (7.7)	97.9 (3.7)
Consonnes pouvant être doublées ***	92.7 (14.2)	96.5 (6.5)	94.1 (14.2)

Note. Choix de la consonne double placée avant plutôt qu'après la consonne simple (\*), choix de la consonne double en position médiane plutôt qu'en position initiale (\*\*) et choix du doublet légal plutôt que du doublet illégal (\*\*\*)

#### Apprentissage de l'orthographe des non-mots — Condition rappel

Dans les différentes conditions d'apprentissage, les participants prononçaient correctement les non-mots. Ainsi, l'item *AB gruprane* était prononcé exactement comme l'item *AAB gupprane*, ce qui exclut que les sujets aient pu mémoriser ces deux types orthographiques en se fondant sur des prononciations différentes (e.g., *gruprane* prononcé /gypran/ et *gupprane* prononcé /gyp/-/pran/). Dans la condition d'apprentissage *implicite contexte*, les participants répondaient tous correctement aux questions sur les textes.

Dans un premier temps, nous avons considéré les non-mots dont le groupe consonantique était correctement orthographié. Le Tableau 2 montre qu'ils sont plus nombreux, d'une part dans la condition *explicite isolé* (75.7%) que dans les conditions *implicite isolé* (57.6%) et *implicite contexte* (43.8%); d'autre part pour les items *AB* (81.9%) que pour les items *AAB* (62.4%) et pour les items *ABB* (32.8%). Une ANOVA avec les variables condition d'apprentissage (*implicite contexte*, *implicite isolé* et *explicite isolé*) et types d'items (*AB*, *AAB* et *ABB*) sur le nombre de groupes consonantiques correctement orthographiés confirme les effets de la condition d'apprentissage ( $F_1(2, 69) = 16.44$ ,  $MSE = .43$ ,  $p < .001$ ,  $\eta_p^2 = .32$ ;  $F_2(2, 10) = 12.79$ ,  $MSE = .03$ ,  $p = .002$ ,  $\eta_p^2 = .72$ ) et du type d'item ( $F_1(2, 138) = 48.99$ ,  $MSE = .36$ ,

$p < .001$ ,  $\eta_p^2 = .42$  ;  $F_2(2, 10) = 44.98$ ,  $MSE = .02$ ,  $p < .001$ ,  $\eta_p^2 = .90$ ) et révèle une interaction significative dans l'analyse sur les items ( $F_1(4, 138) = 1.88$ ,  $MSE = .36$ ,  $p = .12$ ,  $\eta_p^2 = .05$  ;  $F_2(4, 20) = 4.02$ ,  $MSE = .02$ ,  $p = .014$ ,  $\eta_p^2 = .45$ ). Concernant l'effet de la condition d'apprentissage, les orthographes correctes étaient plus nombreuses dans la condition *explicite isolé* que dans les conditions *implicite contexte* et *implicite isolé* (respectivement,  $F_1(1, 69) = 32.61$ ,  $MSE = .43$ ,  $p < .001$ ,  $\eta_p^2 = .32$  ;  $F_2(1, 5) = 19.00$ ,  $MSE = .04$ ,  $p = .007$ ,  $\eta_p^2 = .79$  et  $F_1(1, 69) = 10.89$ ,  $MSE = .43$ ,  $p = .001$ ,  $\eta_p^2 = .14$  ;  $F_2(1, 5) = 17.67$ ,  $MSE = .014$ ,  $p = .008$ ,  $\eta_p^2 = .78$ ). La différence entre les conditions *implicite contexte* et *implicite isolé* est significative seulement dans l'analyse sur les sujets ( $F_1(1, 69) = 5.81$ ,  $MSE = .43$ ,  $p = .018$ ,  $\eta_p^2 = .08$  ;  $F_2(1, 5) = 3.96$ ,  $MSE = .04$ ,  $p = .10$ ,  $\eta_p^2 = .44$ ). Concernant l'effet du type d'item, les items *ABB* étaient significativement moins souvent orthographiés correctement que les items *AB* et *AAB* (respectivement,  $F_1(1, 69) = 14.00$ ,  $MSE = .36$ ,  $p < .001$ ,  $\eta_p^2 = .17$  ;  $F_2(1, 5) = 3.46$ ,  $MSE = .03$ ,  $p = .12$ ,  $\eta_p^2 = .41$  et  $F_1(1, 69) = 32.69$ ,  $MSE = .41$ ,  $p < .001$ ,  $\eta_p^2 = .32$  ;  $F_2(1, 5) = 35.04$ ,  $MSE = .03$ ,  $p = .002$ ,  $\eta_p^2 = .88$ ) et les items *AB* étaient plus souvent orthographiés correctement que les items *AAB* ( $F_1(1, 69) = 109.48$ ,  $MSE = .32$ ,  $p < .001$ ,  $\eta_p^2 = .61$  ;  $F_2(1, 5) = 266.79$ ,  $MSE = .01$ ,  $p < .001$ ,  $\eta_p^2 = .98$ ).

Dans un deuxième temps, nous avons considéré deux types d'erreurs commises sur les items contenant un doublet. Les erreurs d'omissions consistent à orthographier *AB* un item *AAB* ou un item *ABB* ; les erreurs de transposition consistent à orthographier *AAB* un item *ABB* (donc à transformer une orthographe illégale en français en une orthographe légale) ou, inversement, à orthographier *ABB* un item *AAB* (donc à transformer une orthographe légale en français en une orthographe illégale). Le nombre de chacune de ces erreurs a été successivement soumis à une ANOVA avec les variables condition d'apprentissage (*implicite contexte*, *implicite isolé* et *explicite isolé*) et types d'items (*AAB* et *ABB*). Les omissions variaient en fonction de la condition d'apprentissage ( $F_1(2, 69) = 24.27$ ,  $MSE = .34$ ,  $p < .001$ ,  $\eta_p^2 = .41$  ;  $F_2(2, 10) = 9.93$ ,  $MSE = .05$ ,  $p < .001$ ,  $\eta_p^2 = .67$ ) mais il n'y avait pas d'effet du type d'item ni d'interaction ( $ps > .43$ ). Les omissions étaient significativement moins nombreuses dans la condition *explicite isolé* (5.2%) que dans les conditions condition *implicite isolé* (24.0%) et *implicite contexte* (46.9%) (respectivement,  $F_1(1, 69) = 14.63$ ,  $MSE = .34$ ,  $p < .001$ ,  $\eta_p^2 = .17$  ;  $F_2(1, 5) = 7.52$ ,  $MSE = .05$ ,  $p = .04$ ,  $\eta_p^2 = .60$  et  $F_1(1, 69) = 48.37$ ,  $MSE = .34$ ,  $p < .001$ ,  $\eta_p^2 = .41$  ;  $F_2(1, 5) = 16.20$ ,  $MSE = .06$ ,  $p = .01$ ,  $\eta_p^2 = .76$ ). La différence entre les deux conditions implicites était significative seulement pour l'analyse

sur les sujets ( $F_1(1, 69) = 9.80$ ,  $MSE = .34$ ,  $p = .002$ ,  $\eta_p^2 = .12$  ;  $F_2(1, 5) = 3.28$ ,  $MSE = .036$ ,  $p = .13$ ,  $\eta_p^2 = .40$ ).

Les erreurs de transposition étaient plus nombreuses pour les items *ABB* (37.5%), avec la transposition du doublement de la seconde à la première consonne du groupe consonantique, que pour les items *AAB* (6.3%), avec la transposition du doublement de la première à la seconde consonne ( $F_1(1, 69) = 47.86$ ,  $MSE = .29$ ,  $p < .001$ ,  $\eta_p^2 = .41$  ;  $F_2(1, 5) = 20.43$ ,  $MSE = .04$ ,  $p = .006$ ,  $\eta_p^2 = .80$ ). Il n'y avait pas d'effet de la condition d'apprentissage, ni d'interaction ( $ps > .63$ ). Comme le montre le Tableau 3, ce pattern de résultats est confirmé par l'analyse des profils individuels. En effet, parmi les participants qui commettaient des erreurs de transposition (entre 15 et 17 sur 24 selon les conditions d'apprentissage), ceux commettant des erreurs de transposition seulement sur les items *ABB* étaient beaucoup plus nombreux que ceux commettant des erreurs uniquement sur les items *AAB* (40 vs. 2 toutes conditions d'apprentissage confondues).

Tableau 2.

Pourcentages d'orthographe AB, AAB et ABB produites en fonction de la condition d'apprentissage et du type d'items présentés. Les orthographes correctes apparaissent en gras, les erreurs de transposition en italique et les écart-types sont entre parenthèses.

Orthographe produites	Implicite contexte			Implicite isolé			Explicite isolé		
	items	items	items	items	items	items	items	items	items
Orthographe AB	<i>AB</i>	<i>AAB</i>	<i>ABB</i>	<i>AB</i>	<i>AAB</i>	<i>ABB</i>	<i>AB</i>	<i>AAB</i>	<i>ABB</i>
	<b>77.1</b>	47.8	50.0	<b>79.2</b>	25.0	22.9	<b>89.6</b>	4.2	6.3
Orthographe AAB	18.8	<b>43.5</b>	32.6	10.4	<b>60.4</b>	41.7	6.3	<b>83.3</b>	35.4
	(24.7)	<b>(37.9)</b>	(28.6)	(25.4)	<b>(25.4)</b>	(31.9)	(16.9)	<b>(35.1)</b>	(37.5)
Orthographe ABB	0.0	4.3	<b>10.9</b>	2.1	6.3	<b>33.3</b>	2.1	8.3	<b>54.2</b>
	(0.0)	(14.4)	<b>(21.1)</b>	(10.2)	(16.9)	<b>(38.1)</b>	(10.2)	(16.9)	<b>(35.9)</b>
Autres doublets	4.2	4.3	6.5	8.3	8.3	2.1	2.1	4.2	4.2
	(14.1)	(14.4)	(22.9)	(19.0)	(19.0)	(10.2)	(10.2)	(14.1)	(20.4)

Note. Orthographe AB = deux consonnes du groupe consonantique non doublées ; Orthographe AAB = seule la première consonne du groupe consonantique doublée ; Orthographe ABB = seule la seconde consonne du groupe consonantique doublée ; Autres doublets = doublement d'une consonne n'appartenant pas au doublet.

Un résultat important est que la prévalence des erreurs de transposition sur les items *ABB* ne pouvait pas s'expliquer par une utilisation plus fréquente des orthographes *AAB* indépendamment du type d'item. En effet, dans les trois conditions d'apprentissage, les orthographes *AAB* étaient utilisées plus souvent pour les items *AAB* que pour les items *ABB* et plus souvent pour les items *ABB* que pour les items *AB*. Une ANOVA avec les variables condition d'apprentissage (*implicite contexte*, *implicite isolé* et *explicite isolé*) et types d'items (*AB*, *AAB* et *ABB*) sur le nombre de production d'orthographes *AAB* révèle un effet du type d'item ( $F_1(2, 138) = 54.95$ ,  $MSE = .35$ ,  $p < .001$ ,  $\eta_p^2 = .44$  ;  $F_2(2, 10) = 107.96$ ,  $MSE = .01$ ,  $p < .001$ ,  $\eta_p^2 = .96$ ) et une interaction ( $F_1(4, 138) = 4.97$ ,  $MSE = .35$ ,  $p < .001$ ,  $\eta_p^2 = .13$  ;  $F_2(4, 20) = 2.92$ ,  $MSE = .03$ ,  $p = .05$ ,  $\eta_p^2 = .37$ ). Des comparaisons planifiées montrent que les orthographes *AAB* étaient produites pour les items *ABB* moins souvent que pour les items *AAB* ( $F_1(1, 69) = 23.20$ ,  $MSE = .41$ ,  $p < .001$ ,  $\eta_p^2 = .25$  ;  $F_2(1, 5) = 72.46$ ,  $MSE = .01$ ,  $p < .001$ ,  $\eta_p^2 = .94$ ) mais plus souvent que pour les items *AB* ( $F_1(1, 69) = 33.82$ ,  $MSE = .28$ ,  $p < .001$ ,  $\eta_p^2 = .33$  ;  $F_2(1, 5) = 40.79$ ,  $MSE = .01$ ,  $p = .001$ ,  $\eta_p^2 = .89$ ). L'interaction s'explique par le fait que l'utilisation des orthographes *AAB* variait en fonction de la condition d'apprentissage seulement pour les items *AAB* ( $F_1(2, 69) = 7.61$ ,  $MSE = .45$ ,  $p = .001$ ,  $\eta_p^2 = .18$  ;  $F_2(2, 10) = 6.28$ ,  $MSE = .04$ ,  $p = .017$ ,  $\eta_p^2 = .56$  ;  $ps > .12$  pour les items *ABB* et *AB*). Les productions d'orthographes *AAB* pour les items *AAB* étaient plus nombreuses dans la condition *explicite isolé* que dans les deux conditions d'apprentissage implicite ( $F_1(1, 69) = 9.62$ ,  $MSE = .45$ ,  $p = .003$ ,  $\eta_p^2 = .12$  ;  $F_2(1, 5) = 8.23$ ,  $MSE = .04$ ,  $p = .03$ ,  $\eta_p^2 = .62$ ) ; la plus forte utilisation des orthographes *AAB* dans la condition *implicite isolé* que dans la condition *implicite contexte* était significative lorsque les analyses étaient conduites sur les sujets ( $F_1(1, 69) = 5.59$ ,  $MSE = .45$ ,  $p = .02$ ,  $\eta_p^2 = .07$  ;  $F_2(1, 5) = 3.92$ ,  $MSE = .04$ ,  $p = .10$ ,  $\eta_p^2 = .44$ ).

Tableau 3.

Nombre de participants faisant des erreurs de transposition en fonction du type de test et de la condition d'apprentissage.

Epreuve	condition	N avec des erreurs de transposition	N avec des erreurs de transposition seulement sur les items <i>ABB</i>	N avec des erreurs de transposition seulement sur les items <i>AAB</i>
rappel	explicite	15	11	2
rappel	implicite isolé	17	14	0
rappel	implicite contexte	17	15	0
reconnaissance	explicite	15	10	1
reconnaissance	implicite isolé	19	13	2
reconnaissance	implicite contexte	16	10	3

Note. Les cas où le nombre de participants commettant uniquement des erreurs de transposition sur les items *ABB* plus le nombre de participants commettant uniquement des erreurs de transposition sur les items *AAB* diffère du nombre total sont des cas de participants commettant des erreurs de transposition à la fois sur les items *ABB* et *AAB*.

Comme indiqué à la dernière ligne du Tableau 1, les participants doubleraient parfois une consonne n'appartenant pas au groupe consonantique, par exemple lorsqu'ils produisaient l'orthographe GUPRANNE après avoir vu des orthographe comme *gugrane* (item *AB*), *gupprane* (item *AAB*) ou *gupprane* (item *ABB*). En revanche, les participants ne doubleraient jamais des lettres qui ne peuvent pas être doublées en français, n'utiliseraient jamais un doublet en position initiale ou finale et ne doubleraient jamais les deux consonnes d'un groupe consonantique.

#### Apprentissage de l'orthographe des non-mots — Condition reconnaissance

Le Tableau 4 montre que les sélections de l'orthographe correcte sont plus nombreuses d'une part dans la condition *explicite isolé* (74.3%) que dans les conditions *implicite contexte* (50.0%) et *implicite isolé* (56.9%) et d'autre part pour les items *AB* (71.5%) et *AAB* (70.8%) que pour les items *ABB* (38.9%). Une ANOVA avec les variables condition d'apprentissage (*implicite contexte*, *implicite isolé* et *explicite isolé*) et type d'items (*AB*, *AAB* et *ABB*) sur le nombre de sélections correctes confirme les effets de la condition d'apprentissage ( $F_1(2, 69) = 8.51$ ,  $MSE = .53$ ,  $p < .001$ ,  $\eta_p^2 = .20$  ;  $F_2(2, 10) = 24.25$ ,  $MSE = .74$ ,  $p < .001$ ,  $\eta_p^2 = .83$ ) et du type d'item ( $F_1(2, 138) = 21.17$ ,  $MSE = .47$ ,  $p < .001$ ,  $\eta_p^2 = .23$  ;  $F_2(2, 10) = 18.68$ ,  $MSE = 2.14$ ,  $p < .001$ ,  $\eta_p^2 = .79$ ) et ne révèle pas d'interaction ( $ps > .66$ ). Concernant l'effet de la condition d'apprentissage, les orthographe correctes étaient plus souvent sélectionnées dans la condition *explicite isolé* que dans les conditions *implicite contexte* et *implicite isolé* (respectivement,  $F_1(1, 69) = 16.04$ ,  $MSE = .53$ ,  $p < .001$ ,  $\eta_p^2 = .19$  ;  $F_2(1, 5) = 41.11$ ,  $MSE = .82$ ,  $p = .001$ ,  $\eta_p^2 = .89$  ;  $F_1(1, 69) = 8.18$ ,  $MSE = .53$ ,  $p = .006$ ,  $\eta_p^2 = .11$  ;  $F_2(1, 5) = 25.00$ ,  $MSE = .70$ ,  $p = .004$ ,  $\eta_p^2 = .83$ ). La différence entre les conditions *implicite contexte* et *implicite isolé* n'était pas significative ( $F_1(1, 69) = 1.31$ ,  $MSE = .53$ ,  $p = .26$ ,  $\eta_p^2 = .02$  ;  $F_2(1, 5) = 3.90$ ,  $MSE = .71$ ,  $p = .11$ ,  $\eta_p^2 = .44$ ). Concernant l'effet du type d'item, l'orthographe correcte était significativement moins souvent sélectionnée pour les items *ABB* que pour les items *AB* ( $F_1(1, 69) = 23.71$ ,  $MSE = .65$ ,  $p < .001$ ,  $\eta_p^2 = .26$  ;  $F_2(1, 5) = 59.70$ ,  $MSE = 1.02$ ,  $p < .001$ ,  $\eta_p^2 = .92$ ) et *AAB* ( $F_1(1, 69) = 39.31$ ,  $MSE = .37$ ,  $p < .001$ ,  $\eta_p^2 = .36$  ;  $F_2(1, 5) = 26.19$ ,  $MSE = 2.24$ ,  $p = .004$ ,  $\eta_p^2 = .84$ ), sans différence entre les items *AB* et *AAB* ( $ps > .89$ ).

Tableau 4.

Pourcentages de sélection des orthographe AB, AAB et ABB en fonction de la condition d'apprentissage et du type d'items présentés. Les orthographe correctes apparaissent en gras, les erreurs de transposition en italique et les écart-types sont entre parenthèses.

Orthographe sélectionnées	Implicite contexte		Implicite isolé		Explicite isolé				
	items	items	items	items	items	items			
Orthographe AB	AB	AAB ABB	AB	AAB ABB	AB	AAB ABB			
	<b>60.4</b>	29.2 (35.9)	35.4 (37.5)	<b>68.8</b>	20.8 (29.2)	20.8 (29.2)	<b>87.5</b>	2.1 (10.2)	2.1 (10.2)
Orthographe AAB	37.5	<b>58.3</b>	31.3 (32.3)	22.9 (32.9)	<b>66.7</b>	43.8 (34.0)	10.4 (20.7)	<b>87.5</b>	39.6 (39.0)
	2.1 (10.2)	12.5 (22.1)	<b>33.3</b>	8.3 (19.0)	12.5 (22.1)	<b>35.4</b>	2.1 (10.2)	<b>47.9</b>	<b>40.3</b>

Note. AB = deux consonnes du groupe consonantique non doublées ; AAB = seule la première consonne du groupe consonantique doublée ; ABB = seule la seconde consonne du groupe consonantique doublée.



La comparaison des taux de sélection de l'orthographe correcte pour chaque type d'items dans chaque condition d'apprentissage par rapport au hasard (33.3%) montre clairement les difficultés d'apprentissage de l'orthographe des items *ABB*. En effet, dans les trois conditions d'apprentissage, les taux de sélection de l'orthographe correcte étaient supérieurs au hasard pour les items *AB* ( $t(23) > 2.98$ ,  $ps < .007$ ) et pour les items *AAB* ( $t(23) > 3.47$ ,  $ps < .002$ ) mais ne différaient pas significativement du hasard pour les items *ABB* ( $t(23) = 1.75$ ,  $p = .093$  pour la condition *explicite isolé*,  $t(23) = .30$ ,  $p = .77$  pour la condition *implicite isolé* et  $t(23) = -.02$ ,  $p = .98$  pour la condition *implicite contexte*).

Le nombre d'erreurs d'omissions (choix de l'orthographe *AB* pour un item *AAB* ou *ABB*) et de transpositions (choix de l'orthographe illégale *ABB* pour un item *AAB* ou choix de l'orthographe légale *AAB* pour un item *ABB*) ont été successivement soumis à des ANOVAs avec les variables condition d'apprentissage (*implicite contexte*, *implicite isolé* et *explicite isolé*) et types d'items (*AAB* et *ABB*). Les omissions variaient en fonction de la condition d'apprentissage ( $F_1(2, 69) = 6.77$ ,  $MSE = .44$ ,  $p = .002$ ,  $\eta_p^2 = .16$  ;  $F_2(2, 10) = 7.39$ ,  $MSE = 1.62$ ,  $p = .01$ ,  $\eta_p^2 = .60$ ) mais il n'y avait pas d'effet du type d'item ni d'interaction ( $ps > .18$ ). Elles étaient significativement moins nombreuses dans la condition *explicite isolé* (7.3%) que dans les conditions condition *implicite isolé* (20.8%,  $F_1(1, 69) = 3.97$ ,  $MSE = .44$ ,  $p = .05$ ,  $\eta_p^2 = .05$  ;  $F_2(1, 5) = 3.62$ ,  $MSE = 1.94$ ,  $p = .11$ ,  $\eta_p^2 = .42$ ) et *implicite contexte* (32.3%,  $F_1(1, 69) = 13.52$ ,  $MSE = .44$ ,  $p < .001$ ,  $\eta_p^2 = .16$  ;  $F_2(1, 5) = 16.00$ ,  $MSE = 1.5$ ,  $p = .01$ ,  $\eta_p^2 = .76$ ), sans différence significative entre les deux conditions implicites ( $F_1(1, 69) = 2.84$ ,  $MSE = .44$ ,  $p = .10$ ,  $\eta_p^2 = .04$  ;  $F_2(1, 5) = 3.50$ ,  $MSE = 1.44$ ,  $p = .12$ ,  $\eta_p^2 = .41$ ).

Les erreurs de transposition étaient significativement plus nombreuses pour les items *ABB* que pour les items *AAB* (38.2% vs. 11.8%,  $F_1(1, 69) = 30.70$ ,  $MSE = .33$ ,  $p < .001$ ,  $\eta_p^2 = .31$  ;  $F_2(1, 5) = 12.62$ ,  $MSE = 3.18$ ,  $p = .02$ ,  $\eta_p^2 = .72$ ). Il n'y avait pas d'effet de la condition d'apprentissage ni d'interaction ( $ps > .43$ ). Ce pattern de résultats est confirmé par l'analyse des profils individuels (cf Tableau 3). En effet, parmi les participants commettant des erreurs de transposition (entre 15 et 19 sur 24 selon les conditions d'apprentissage), ceux en commettant seulement sur les items *ABB* étaient beaucoup plus nombreux que ceux en commettant uniquement sur les items *AAB* (33 vs. 6 toutes conditions d'apprentissage confondues).

Comme dans la condition rappel, cette prévalence des erreurs de transposition sur les items *ABB* ne peut pas s'expliquer par une préférence générale pour les orthographes *AAB*. Une ANOVA avec les variables condition d'apprentissage (*implicite contexte*, *implicite isolé*

et *explicite isolé*) et types d'items (*AB*, *AAB* et *ABB*) sur le nombre de sélections des orthographes *AAB* révèle une absence de différence en fonction de la condition d'apprentissage ( $ps > .77$ ) mais un effet du type d'item ( $F_1(2, 138) = 36.38$ ,  $MSE = .43$ ,  $p < .001$ ,  $\eta_p^2 = .35$  ;  $F_2(2, 10) = 27.34$ ,  $MSE = 2.46$ ,  $p < .001$ ,  $\eta_p^2 = .85$ ) et une interaction ( $F_1(4, 138) = 4.57$ ,  $MSE = .43$ ,  $p = .001$ ,  $\eta_p^2 = .12$  ;  $F_2(4, 20) = 6.16$ ,  $MSE = 1.37$ ,  $p = .002$ ,  $\eta_p^2 = .55$ ). Des comparaisons planifiées montrent que les orthographes *AAB* étaient choisies pour les items *ABB* moins souvent que pour les items *AAB* ( $F_1(1, 69) = 33.67$ ,  $MSE = .46$ ,  $p < .001$ ,  $\eta_p^2 = .33$  ;  $F_2(1, 5) = 16.61$ ,  $MSE = 3.69$ ,  $p = .009$ ,  $\eta_p^2 = .77$ ) mais plus souvent que pour les items *AB* ( $F_1(1, 69) = 6.19$ ,  $MSE = .49$ ,  $p = .015$ ,  $\eta_p^2 = .08$  ;  $F_2(1, 5) = 7.73$ ,  $MSE = 1.58$ ,  $p = .038$ ,  $\eta_p^2 = .61$ ). L'interaction s'explique par le fait que le choix des orthographes *AAB* variait en fonction de la condition d'apprentissage pour les items *AB* et *AAB* ( $F_s > 4.28$ ,  $ps < .017$ ) mais pas pour les items *ABB* ( $ps > .45$ ). Les orthographes *AAB* étaient choisies moins souvent pour les items *AB* et plus souvent pour les items *AAB* dans la condition *explicite* que dans les deux conditions *implicites* ( $F_s > 6.09$ ,  $ps < .016$ ). Les orthographes *AAB* étaient choisies moins souvent pour les items *AB* dans la condition *implicite isolé* que dans la condition *implicite contexte* mais cet effet n'était significatif que dans l'analyse sur les items ( $F_1(1, 69) = 2.48$ ,  $MSE = .41$ ,  $p = .12$ ,  $\eta_p^2 = .03$  ;  $F_2(1, 5) = 14.41$ ,  $MSE = .28$ ,  $p = .013$ ,  $\eta_p^2 = .74$ ) et les orthographes *AAB* étaient choisies plus souvent pour les items *AAB* dans la condition *implicite isolé* que dans la condition *implicite contexte* mais cet effet n'était significatif que dans l'analyse sur les sujets ( $F_1(1, 69) = 5.59$ ,  $MSE = .45$ ,  $p = .02$ ,  $\eta_p^2 = .07$  ;  $F_2(1, 5) = 1.82$ ,  $MSE = .73$ ,  $p = .23$ ,  $\eta_p^2 = .27$ ).

## Discussion

L'objectif de cette expérience était de généraliser le résultat rapporté dans les expériences du chapitre expérimental III montrant que des individus se fondent à la fois sur leur connaissance relative à des mots spécifiques (la présence d'un doublet) et leurs connaissances générales sur l'écrit (régularités graphotactiques) lors du rappel ou de la reconnaissance de l'orthographe de nouveaux mots présentés au préalable. Alors que la propriété graphotactique exploitée dans les expériences du chapitre précédent était de nature probabiliste, celle exploitée dans cette expérience était descriptible sous forme de règle (les consonnes doubles peuvent survenir avant mais non après les consonnes simples). Par ailleurs, alors que le mode d'apprentissage était toujours implicite dans les expériences du chapitre précédent, des conditions d'apprentissage implicite et explicite étaient utilisées dans cette expérience.

Les orthographe correctes étaient moins bien rappelées et reconnues pour les items *ABB*, qui violent la régularité graphotactique, que pour les deux types d'items *AB* et *AAB*, tous deux légaux du point de vue des régularités graphotactiques. Les difficultés d'apprentissage de l'orthographe des items *ABB* était manifeste dans la tâche de reconnaissance dans laquelle les taux de reconnaissance de l'orthographe correcte étaient significativement supérieurs au hasard pour les items *AB* et *AAB* mais n'en différaient pas pour les items *ABB*. Les orthographe correctes étaient mieux rappelées pour les items ne contenant pas de doublet (items *AB*) que pour les items incluant un doublet en position légale (items *AAB*) mais les taux de reconnaissances correctes ne différaient pas entre ces deux types d'items. Les orthographe correctes étaient mieux rappelées et reconnues dans la condition d'apprentissage explicite que dans les deux conditions d'apprentissage implicite dans lesquelles les erreurs d'omission étaient plus nombreuses.

Les erreurs d'omission ne différaient pas en fonction du type d'items, de façon stable entre les différentes conditions d'apprentissage comme l'atteste l'absence d'interaction Type d'item x Condition d'apprentissage. Par contraste, les erreurs de transpositions, dont l'occurrence ne variait pas en fonction de la condition d'apprentissage, étaient plus nombreuses sur les items *ABB* (transposition du doublement de la seconde consonne du groupe consonantique, ce qui est illégal en français, vers la première consonne du groupe consonantique, ce qui est légal) que pour les items *AAB* (transposition du doublement de la

première vers la seconde consonne du groupe consonantique), de façon stable entre les différentes conditions d'apprentissage, comme l'atteste l'absence d'interaction Type d'item x Condition d'apprentissage. Enfin, la prévalence des erreurs de transposition n'était pas due à une préférence générale, indépendante du type d'items, pour les orthographe avec un doublement de la première consonne du groupe consonantique (orthographe *AAB*) car celles-ci étaient produites (rappel) et sélectionnées (reconnaissance) pour les items *ABB* moins souvent que pour les items *AAB* mais plus souvent que pour les items *AB*.

Ainsi, cette expérience montre que des individus se fondent à la fois sur leur connaissance relative à des mots spécifiques (la présence d'un doublet dans un groupe consonantique) et leurs connaissances relatives à une propriété graphotactique descriptible sous la forme d'une règle générale spécifiant que les consonnes doubles peuvent survenir avant mais pas après une consonne simple. Ceci conduit à des réponses correctes pour les items *AAB* mais à des erreurs de transposition pour les items *ABB*. Cette expérience confirme les résultats rapportés dans les cinq expériences du chapitre précédent avec des régularités graphotactiques de nature probabiliste. De plus, cette expérience suggère que les régularités graphotactiques ont un impact sur l'apprentissage de l'orthographe de mots spécifiques chez l'adulte même dans une situation d'apprentissage explicite.

## DISCUSSION GENERALE

L'objectif principal de cette thèse était de préciser le rôle des connaissances graphotactiques (i.e., relatives à la fréquence de combinaisons des graphèmes) dans l'acquisition de l'orthographe lexicale (i.e., orthographe de mots spécifiques). Cette question a été abordée dans des situations d'apprentissage de l'orthographe de non-mots de deux façons. Dans le chapitre expérimental II, nous avons étudié à l'aide d'analyses de corrélation et de régression si une épreuve évaluant les connaissances graphotactiques prédisait le niveau d'apprentissage de l'orthographe de non-mots contenant des phonèmes à transcriptions multiples après avoir contrôlé diverses variables, dont les connaissances en orthographe lexicale, chez des élèves de CM1 et CM2. Dans les chapitres expérimentaux III et IV, nous avons manipulé le degré de convergence entre les orthographe des non-mots et les régularités graphotactiques du Français afin d'examiner si et comment le niveau d'apprentissage orthographique et la nature des erreurs variaient en fonction du degré de convergence entre les orthographe des non-mots et les régularités graphotactiques chez des élèves de CE2 et CM2 et chez des adultes. Avant de discuter les résultats rapportés dans ces trois chapitres, nous discutons brièvement ceux du premier chapitre expérimental dont l'objectif était de préciser les connaissances graphotactiques d'élèves de l'école élémentaire et d'adultes.

### **1) Etude des connaissances graphotactiques**

Dans l'expérience rapportée au chapitre expérimental I, les connaissances graphotactiques étaient examinées avec deux épreuves de jugement. La première était une tâche de jugement de paires de non-mots, similaire à celles utilisées dans des études antérieures (e.g., Cassar & Treiman, 1997 ; Danjon & Pacton, 2009 ; Pacton et al., 2001), qui explorait trois propriétés des doubles lettres : 1) le fait que seules certaines consonnes peuvent être doublées ; 2) l'illégalité des doublets en début de mot et 3) l'illégalité des doublets après une consonne simple. Ces deux dernières propriétés descriptibles sous la forme de règle générale ont également été étudiées avec une nouvelle tâche de jugement consistant en un choix entre l'orthographe correcte et une orthographe incorrecte, phonologiquement plausible, mais qui viole une règle graphotactique par la présence d'un doublet en début de mot ou après une consonne simple. De plus, dans la tâche de jugement de paires de non-mots, la familiarité

du matériel utilisé pour tester les connaissances graphotactiques était manipulée, l'idée étant que cette manipulation ne devait pas affecter les performances des sujets si ceux-ci appliquaient des règles spécifiant que les consonnes doubles ne peuvent pas survenir en début de mot ni après une consonne simple.

Les résultats indiquaient que la sensibilité aux propriétés graphotactiques ne peut pas être considérée comme une composante homogène. En effet, pour les deux épreuves, les scores étaient plus faibles pour les items testant l'illégalité des doublets après une consonne simple que pour les items testant l'illégalité des doublets en début de mots en CE2 et CM2. Ce résultat confirme ceux rapportés par Danjon et Pacton (2009) dans une étude qui n'utilisait qu'une tâche de jugement de paires de non-mots. Des scores plafonds étaient observés chez les adultes dont les performances ne différaient pas en fonction du type de propriété.

Concernant la nature des connaissances graphotactiques pour les deux propriétés descriptibles sous forme de règle, les performances des élèves variaient en fonction de la familiarité du matériel utilisé chez des élèves de l'école élémentaire, ce qui confirme les résultats de Pacton et al. (2001) et Danjon et Pacton (2009). Le résultat nouveau est que cet effet était observé même chez des adultes dont la pratique de l'écrit était encore beaucoup plus importante que celle d'enfants scolarisés en fin de l'école élémentaire. Les performances sur les items familiers et non familiers étaient corrélées de façon significative même après avoir contrôlé l'âge des sujets ( $r_s > .48$ ), ce qui montre une relation étroite entre les performances sur du matériel non familier et celles sur du matériel familier.

La persistance d'effets liés à la familiarité du matériel test utilisé même chez les adultes est difficile à réconcilier avec l'idée selon laquelle les sujets acquièrent une connaissance de plus en plus abstraite, fondée sur des règles, des régularités présentes dans le matériel auquel ils sont confrontés. Elle suggère que, contrairement à ce qu'avaient suggéré certains auteurs (e.g., Manza & Reber, 1997), la pratique limitée qui caractérise de nombreuses études de l'apprentissage implicite conduites en laboratoire, comme celles portant sur l'apprentissage de grammaires artificielles, ne peut pas expliquer (à elle seule) des phénomènes comme le *transfer-decrement*. La persistance d'effets liés à la familiarité du matériel test utilisé est en revanche en accord avec les conceptions distributionnelles de l'apprentissage implicite (Shanks et al., 1997) selon lesquelles l'abstraction est un processus graduel qui dépend de la similarité (similarité qui ne change pas au cours du temps) entre les items familiers et les items nouveaux.

Il convient de noter que cet effet de familiarité du matériel utilisé pour tester les connaissances d'adultes était significatif alors même que leurs scores moyens étaient très

proches de 100%. De tels scores impliquent que de nombreux sujets obtenaient le score maximum pour les items familiers et non familiers mais que ceux qui obtenaient des scores différents pour les deux types d'items obtenaient (quasi) systématiquement des scores plus faibles pour les items non familiers. Il serait intéressant, par exemple en interviewant les sujets après la passation de l'expérience, d'essayer de préciser en quoi ces sujets se distinguent de ceux dont les scores ne diffèrent pas en fonction de la familiarité du matériel utilisé.

## **2) Rôle des connaissances graphotactiques dans l'apprentissage de l'orthographe lexicale**

L'étude rapportée au chapitre expérimental II, conduite auprès d'élèves de CM1 et CM2 nous a permis de préciser le rôle des connaissances en orthographe lexicale et celui des connaissances graphotactiques dans l'apprentissage de l'orthographe de nouveaux mots. Pour ce faire, nous avons utilisé le paradigme classique d'auto-apprentissage, qui consiste à faire lire des histoires incluant des non-mots puis à tester l'apprentissage de leur orthographe (avec une épreuve de dictée). Les connaissances en orthographe lexicale des élèves étaient évaluées avec une épreuve de dictée et une épreuve de reconnaissance ; les connaissances graphotactiques avec une épreuve de jugement de paires de non-mots (e.g., Cassar & Treiman, 1997 ; Pacton et al., 2001). Des corrélations significatives mais modérées étaient observées entre le score en connaissances graphotactiques et les scores en orthographe lexicale qu'il s'agisse d'épreuve de jugement ou de production. Les corrélations étaient plus fortes entre deux épreuves de nature différente (reconnaissance et production) évaluant toutes deux les connaissances en orthographe lexicale qu'entre la tâche de reconnaissance évaluant les connaissances en orthographe lexicale et la tâche de jugement évaluant les connaissances graphotactiques, ce qui suggère qu'il s'agit de deux facettes des connaissances orthographiques (Castles & Nation, 2006).

Ces deux types de connaissances orthographiques, spécifique et générale, étaient toutes deux corrélées au niveau d'apprentissage de l'orthographe des non-mots. De plus, elles expliquaient des parts de variance indépendantes. Ce résultat confirme dans le cas de l'apprentissage de l'orthographe de nouveaux mots celui de Connors et al. (2011) dans le cas de la lecture de mots. Notre étude permet de préciser le rôle des connaissances orthographiques dans l'apprentissage de l'orthographe de nouveaux mots par rapport aux

rare études antérieures qui soit avaient examiné le rôle des connaissances lexicales mais pas celui des connaissances graphotactiques (Cunningham et al., 2002), soit avaient examiné le rôle de ces deux types de connaissances sans les distinguer (Cunningham, 2006). Cependant, contrairement à d'autres études antérieures (Cunningham, 2006 ; Nation et al., 2006), nous n'avons pas observé de corrélation entre la qualité du recodage phonologique, considérée comme cruciale dans la théorie du *self-teaching* de Share (1995), et l'apprentissage de l'orthographe de nouveaux mots. Cette différence pourrait s'expliquer par des différences d'âges : les élèves plus âgés dans notre étude ayant de meilleures habiletés de recodage phonologique (e.g., 90% dans notre étude vs. 78% dans celle de Nation et al., 2006), les scores en recodage phonologique variaient moins que dans d'autres études. Il serait intéressant de conduire une expérience de ce type avec des élèves plus jeunes et/ou dont les habiletés en recodage phonologique diffèrent davantage.

Sur la base des études antérieures et de notre étude, un modèle des relations entre le recodage phonologique et les différentes connaissances orthographiques peut être proposé (cf. Figure 1.). L'exposition à des mots permettrait l'acquisition de deux types de connaissances orthographiques, générales et spécifiques. La relation entre l'exposition visuelle et chacune de ces connaissances orthographiques serait soit directe, soit médiatisée par le recodage phonologique. En effet, des études ont montré que l'apprentissage orthographique de nouveaux mots n'est pas dû à une simple exposition visuelle et que le recodage phonologique joue un rôle important (Share, 1999 ; Kyte & Johnson, 2006 ; Dejong et al, 2009). D'une part, l'utilisation de doubles tâches, minimisant les possibilités de recodage phonologique (Kyte & Johnson, 2006 ; Dejong et al, 2009), diminue le niveau d'apprentissage orthographique. D'autre part, le niveau d'apprentissage est plus faible avec des symboles non alphabétiques qu'avec des symboles alphabétiques (Share, 1999). Toutefois, un apprentissage est observé dans ces situations qui limitent les possibilités de recodage phonologique.

Ce modèle représente aussi les relations entre les deux types de connaissances orthographiques. La première, des connaissances spécifiques vers les connaissances générales, correspond à l'extraction de régularités graphotactiques à partir de mots stockés dans le lexique orthographique, en plus de l'extraction des régularités graphotactiques à partir de l'exposition visuelle (via le recodage phonologique ou non) pour des items ne faisant pas (encore) partie du lexique orthographique. La seconde, des connaissances générales vers les connaissances spécifiques, représente l'influence des connaissances graphotactiques sur la mémorisation de l'orthographe de mots spécifiques, mise en évidence dans les chapitres expérimentaux III et IV. Enfin, ce modèle prévoit une influence des connaissances générales



vers le recodage phonologique car les sujets décoderaient des mots qui ne font pas (encore) partie de leur lexique orthographique en utilisant leurs connaissances relatives à des unités plus larges que le graphème mais plus petites que le mot qui surviennent dans de nombreux mots (e.g., *able*, *tion*).

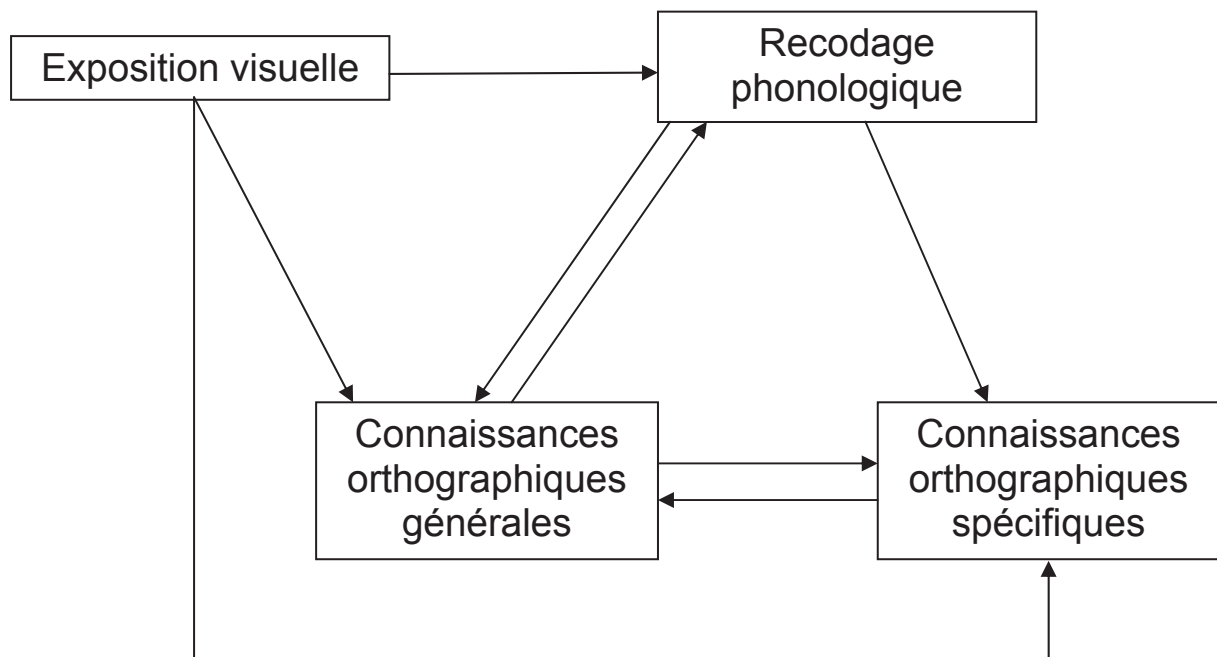


Figure 1 : Modèle des relations entre recodage phonologique et connaissances orthographiques (générales et spécifiques) dans l'apprentissage de l'orthographe de nouveaux mots.

Dans les expériences rapportées dans les chapitres expérimentaux III et IV, nous avons étudié l'influence des connaissances graphotactiques sur l'apprentissage de l'orthographe lexicale en examinant si et comment le niveau d'apprentissage de non-mots et la nature des erreurs variaient en fonction du degré de convergence entre les orthographe des non-mots et les régularités graphotactiques de la langue française. Deux propriétés graphotactiques ont été manipulées, l'une probabiliste (certaines consonnes sont plus fréquemment doublées que d'autres) chez des adultes et des élèves de CE2 et CM2, l'autre descriptible sous forme de règle (les consonnes doubles peuvent survenir avant mais pas après les consonnes simples) seulement chez des adultes. Nous avons utilisé différentes situations d'apprentissage (lecture silencieuse ou à haute voix chez les élèves et les adultes ; non-mots insérés dans des textes ou

présentés isolément seulement chez l'adulte ; apprentissage implicite ou explicite chez l'adulte pour la propriété graphotactique descriptible sous la forme de règle) et deux types de test pour évaluer l'apprentissage orthographique (rappel et reconnaissance).

Afin d'explorer l'impact de la sensibilité à la fréquence de doublement des consonnes sur l'apprentissage de l'orthographe de non-mots, trois types d'items ont été utilisés : des items sans doublet, des items incluant un doublet fréquent et des items incluant un doublet rare. Dans différentes conditions d'apprentissage incident (non-mots insérés dans des textes lus à voix haute ; non-mots présentés isolément lus silencieusement ou à voix haute) et de test (rappel et reconnaissance), les performances des adultes étaient meilleures pour les items sans doublet que pour les items incluant un doublet fréquent et elles étaient meilleures pour les items incluant un doublet fréquent que pour les items incluant un doublet rare. Les erreurs d'omissions étaient aussi nombreuses pour les items incluant un doublet fréquent que pour les items incluant un doublet rare (e.g., tidunnar ou tiddunar orthographié TIDUNAR). En revanche, les erreurs de transposition étaient toujours beaucoup plus fréquentes sur les items incluant un doublet rare (e.g., tiddunar orthographié TIDUNNAR) que sur les items incluant un doublet fréquent (e.g., tidunnar orthographié TIDDUNAR). En fait, les erreurs de transposition du doublement d'une consonne rarement doublée vers une consonne fréquemment doublée en Français représentaient entre 75% et 85% des erreurs de transposition selon les conditions expérimentales chez les adultes.

Un résultat important est que cette prévalence des erreurs de transposition sur les items incluant un doublet rare ne pouvait pas s'expliquer par une tendance générale à utiliser des doublets fréquents indépendamment de toute connaissance sur la présence d'un doublet dans un item spécifique puisque les orthographes incluant un doublet fréquent étaient moins souvent produites (test de dictée) et choisies (test de reconnaissance) pour les items présentés sans doublet dans la phase d'étude que pour les items présentés avec un doublet fréquent ou un doublet rare. Ceci était vrai dans les expériences dont deux tiers des items incluaient des doublets (deux items *sans doublet*, deux items *doublet fréquent* et deux items *doublet rare* dans les expériences avec les non-mots insérés dans des textes) et dans celles dont la moitié des items incluaient des doublets (quatre items *sans doublet*, deux items *doublet fréquent* et deux items *doublet rare* dans les expériences avec les non-mots présentés isolément). Un résultat intéressant est que le patron de performances concernant les erreurs de transposition dans les trois premières expériences du chapitre expérimental III, dans lesquelles il était explicitement demandé aux sujets d'orthographier les items comme ils l'étaient lors de la

phase d'étude, était similaire à celui de l'étude de Fayol et al. (2010) dans laquelle les sujets ne recevaient aucune instruction de ce type. Ceci indique donc que les résultats de Fayol et al. (2010) concernant les erreurs de transposition ne pouvaient pas être attribués (uniquement) à l'absence d'instruction d'orthographier les nouveaux mots comme ils l'étaient dans les textes.

Les résultats des élèves de CM2 (Expériences 4 et 5 du chapitre expérimental III) étaient très similaires à ceux des adultes aussi bien lorsque leur apprentissage orthographique était évalué avec une épreuve de rappel que lorsqu'il était évalué avec une épreuve de reconnaissance. Les productions et sélections correctes étaient plus nombreuses pour les items sans doublet que pour les items incluant un doublet et parmi ces derniers, elles étaient plus nombreuses pour les items incluant un doublet fréquent que pour les items incluant un doublet rare. Concernant les erreurs, les omissions étaient aussi nombreuses pour les items incluant un doublet rare que pour les items incluant un doublet fréquent alors que les transpositions étaient beaucoup plus fréquentes pour les items incluant un doublet rare (entre 76% et 82% des erreurs de transposition) et cela ne correspondait pas à une préférence générale pour les orthographe qui incluent un doublet fréquent, indépendamment du type d'items présentés.

Les résultats des élèves de CE2 montraient que les enfants de ce niveau scolaire peuvent acquérir implicitement des connaissances sur l'orthographe de mots spécifiques lorsqu'ils lisent des textes mais que la mise en évidence de cet apprentissage dépend étroitement de la façon dont il est évalué. En effet, les élèves de CE2 qui ont lu les textes, comme ceux du groupe contrôle qui n'avaient pas lu de textes, incluaient rarement des doubles lettres dans leurs productions orthographiques, pour les trois types d'items. De ce fait, ils orthographiaient très bien les items présentés sans doublet mais commettaient énormément d'erreurs d'omission sur les items présentés avec un doublet, fréquent ou rare. Comme ils incluaient très rarement des doublets dans leurs orthographe, les productions correctes pour les items présentés avec un doublet, fréquent ou rare, ainsi que les erreurs de transposition étaient rares à l'épreuve de dictée. En revanche, le pattern de résultats des CE2 ressemblait davantage à celui des CM2 et des adultes, lorsque leur apprentissage orthographique était évalué avec une épreuve de reconnaissance, avec en particulier une prévalence des erreurs de transposition sur les items incluant des doublets rares qui ne pouvait pas s'expliquer par une tendance générale à sélectionner les items incluant un doublet fréquent. Cette différence en fonction de la tâche utilisée a déjà été soulignée dans des études antérieures utilisant des situations d'auto-apprentissage. Par exemple, dans l'étude de Wang et al. (2011), des effets

observés avec un test de reconnaissance en oui/non ne l'étaient pas avec un test de reconnaissance en choix forcé ou avec une épreuve de dictée (voir aussi Pacton & Deacon, 2008 pour des différences en fonction du type d'épreuves proposées dans des études explorant l'utilisation d'informations morphologiques).

Une influence des connaissances graphotactiques sur l'apprentissage de l'orthographe de nouveaux mots a également été observée dans l'expérience rapportée au chapitre expérimental IV. Dans cette expérience, conduite uniquement avec des sujets adultes, trois types d'items étaient utilisés pour étudier si la connaissance du fait que les consonnes doubles peuvent survenir avant mais non après les consonnes simples — une propriété graphotactique descriptible sous forme de règle — influençait l'apprentissage de l'orthographe de nouveaux mots. Ces items incluaient un groupe consonantique orthographié soit avec les deux consonnes simples, soit avec la première consonne doublée, soit avec la deuxième consonne doublée. Par ailleurs, en plus des deux conditions d'apprentissage implicite utilisées dans le chapitre expérimental III (i.e., non-mots insérés dans des textes ou présentés isolément), nous avons utilisé une condition d'apprentissage explicite dans laquelle les sujets devaient essayer de mémoriser l'orthographe des non-mots qui étaient présentés isolément.

Les orthographes correctes étaient moins bien rappelées et reconnues pour les items incluant un doublet après une consonne simple, qui violent la régularité graphotactique, que pour les deux autres types d'items, tous deux légaux du point de vue des régularités graphotactiques dans les trois conditions d'apprentissage. L'apprentissage orthographique était meilleur dans la condition d'apprentissage explicite que dans les deux conditions d'apprentissage implicite, ce qui est en accord avec des études antérieures dans lesquelles si des conditions d'apprentissage implicite conduisaient à un apprentissage orthographique, des conditions d'apprentissage explicite étaient néanmoins plus efficaces (Ormrod, 1986a, 1986b).

Par delà cette différence d'efficacité entre les modes d'apprentissage implicite et explicite, le résultat principal était qu'alors que les erreurs d'omission étaient aussi nombreuses pour les deux types d'items incluant un doublet, les transpositions du doublement de la seconde consonne du groupe consonantique (illégal) vers la première consonne du groupe consonantique (légal) étaient beaucoup plus nombreuses que les transpositions du doublement de la première vers la seconde consonne du groupe consonantique dans les trois conditions d'apprentissage. De plus, ce résultat n'était pas dû à une préférence générale, indépendante du type d'items, pour les orthographes avec la première consonne du groupe consonantique doublée car ces orthographes étaient produites (rappel) et sélectionnées

(reconnaissance) beaucoup moins souvent pour les items sans doublet que pour les deux types d'items incluant un doublet.

La prévalence des erreurs de transposition sur les items incluant des doublets rares (plutôt que des doublets fréquents pour la propriété *fréquence de doublement des lettres*) et sur les items incluant un doublet après une consonne simple (plutôt qu'avant une consonne simple pour la propriété *illégalité d'un doublet après une consonne simple*), leur rareté sur les items incluant un doublet fréquent (pour la propriété *fréquence de doublement des lettres*) et sur les items incluant un doublet avant une consonne simple (pour la propriété *illégalité d'un doublet après une consonne simple*) et le fait que les productions et sélections d'orthographe incluant un doublet (fréquent ou rare pour la propriété *fréquence de doublement des lettres* / avant ou après la consonne simple pour la propriété *illégalité d'un doublet après une consonne simple*) suggèrent que les sujets se rappellent parfois de la présence d'un doublet mais plus de la lettre spécifique qui doit être doublée. Dans ce cas, ils reconstruiraient une orthographe en se fondant sur leurs connaissances du fait que certaines lettres sont plus susceptibles d'être doublées que d'autres et que ces doublements sont eux mêmes plus probables dans certaines positions que d'autres. Ceci expliquerait que les enfants et les adultes ne doublaient jamais des lettres qui ne sont jamais doublées en français et ne doublaient jamais des lettres en début et fin de mots. Ceci expliquerait également qu'ils doublaient plus souvent des lettres qui sont souvent doublées en français que des lettres qui le sont rarement (d'où la prévalence des erreurs de transposition sur les items incluant des doublets rares) et qu'ils doublaient plus souvent la première que la deuxième consonne du groupe consonantique (d'où la prévalence des erreurs de transposition sur les items incluant des doublets après les consonnes simples).

Les expériences rapportées dans les chapitres expérimentaux III et IV et les études antérieures (Campbell & Coltheart, 1984 ; Fayol et al. 2010 ; Wright & Ehri, 2007) montrent que les connaissances graphotactiques d'élèves de l'école élémentaire et d'adultes influencent la mémorisation de l'orthographe des mots. Ce résultat semble particulièrement robuste puisqu'il apparaît avec différents types de mots (e.g., noms propres vs. noms communs), dans différentes situations d'apprentissage (apprentissage incident vs. intentionnel ; lecture à haute voix vs. silencieuse ; non-mots insérés dans des textes vs. présentés isolément), différentes situations de test (rappel et reconnaissance ; instruction ou non d'orthographier les items tels qu'ils étaient présentés lors de la phase d'étude), avec des populations différentes (élèves en

tout début d'apprentissage de l'écrit ; élèves en milieu et fin de l'école élémentaire ; étudiants de l'université) et dans des langues différentes (anglais et français).

Ces résultats suggèrent que les processus reconstructifs de la mémoire, étudiés dans d'autres domaines comme la mémorisation d'histoires (e.g., Bartlett, 1932 ; Bergman & Roediger, 1999) ou de chansons (Hyman & Rubin, 1990), sont également à l'œuvre dans le cas de l'orthographe. Selon une conception reconstructive de la mémoire, la mémoire est influencée à la fois par des items ou événements spécifiques et par des connaissances générales sur un domaine (Suprenant & Neath, 2009). Ces recherches montrent que les processus reconstructifs facilitent le plus souvent la récupération d'informations parce qu'ils sont fondés sur les régularités du monde environnant. Toutefois, ces processus engendrent des erreurs quand la réponse la plus probable n'est pas la réponse correcte. Ces erreurs sont plus ou moins systématiques. Elles présentent un biais en faveur de l'événement / item le plus probable. Ce biais, qui conduit à la prévalence des erreurs de transposition sur les items incluant un doublet rare ou incluant un doublet situé après une consonne simple dans nos expériences, a été rapporté dans d'autres domaines (Botvinick & Bylsma, 2005; Dell, Reed, Adams & Meyer, 2000; Loftus & Palmer, 1974). Par exemple, les étudiants de l'Université de l'étude de Bartlett (1932), qui devaient reproduire une histoire plutôt incohérente compte tenu de leur culture, effectuaient des « régularisations » en la rendant plus similaire à leur culture (voir aussi Bergman & Roediger, 1999). Ces phénomènes de reconstruction sont également centraux dans les études sur les fausses mémoires (e.g., Hyman & Loftus, 1998 ; Loftus & Palmer, 1974), la mémoire sémantique (Hyman & Rubin, 1990) et la mémoire d'ordre sériel (e.g., Botvinick & Bylsma, 2005).

Ainsi, si les performances orthographiques des élèves et des adultes dépendaient principalement de processus de mémorisation « par cœur », l'apprentissage de l'orthographe d'un nouveau mot serait avant tout déterminé par le nombre de fois que cette orthographe a été vue. Dans les différentes expériences rapportées aux chapitres expérimentaux III et IV, alors que le nombre d'expositions était identique pour les différents types d'items, le degré de convergence entre l'orthographe d'un mot et les régularités graphotactiques de la langue influençait l'apprentissage orthographique.

Dans ce qui précède, nous nous sommes focalisés sur le rôle des connaissances générales (connaissances graphotactiques) sur l'apprentissage de l'orthographe de mots spécifiques mais nos résultats apportent également des informations intéressantes concernant la nature des connaissances orthographiques lexicales, c'est-à-dire spécifiques à un mot. Ils

suggèrent que la mémoire d'un mot spécifique inclut une information sur la présence d'un doublement qui peut être détachée de la lettre à laquelle cette caractéristique (abstraite) s'applique. L'idée selon laquelle les représentations orthographiques ne contiennent pas seulement les lettres individuelles formant ce mot mais aussi la présence de caractéristiques comme le doublement a été suggérée par des auteurs ayant étudié les productions orthographiques de patients cérébro-lésés. Certains patients semblent en effet se rappeler qu'un mot particulier contient un doublet mais pas de la lettre qui est doublée. Par exemple, les patients LB (Caramazza & Miceli, 1990) et HE (McCloskey et al., 1994) commettaient des erreurs de substitutions qui impliquaient systématiquement les deux lettres formant le doublet (e.g., *sorella* orthographié *soretta*) et des erreurs de transposition dans lesquelles ils doubleraient la mauvaise lettre (e.g., *sorella* orthographié *sorrela*). Ces patients ne produisaient presque jamais de doublet pour les mots n'en contenant pas. De plus, le patient FM (Tainturier & Caramazza, 1996) produisait plus souvent une seule de deux lettres adjacentes lorsque ces deux lettres ne correspondaient pas à un doublet (*basket* orthographié *baket*) que lorsqu'elles formaient un doublet (*rabbit* orthographié *rabit*). FM utilisait des doublets incorrects plus souvent pour les mots contenant un doublet (e.g., *rabbit* orthographié *rappit*) que pour les mots sans doublet (e.g., *basket* orthographié *bappet*). Enfin, FM ne présentait pas ce pattern de performances pour des paires de lettres répétées mais non adjacentes (e.g., *cactus*). Sur la base de ces résultats, les auteurs ont suggéré que l'identité des graphèmes serait spécifiée indépendamment des informations sur leur quantité.

On retrouve cette idée dans les travaux de Rumelhart et Norman (1982) cherchant à modéliser les erreurs de frappe notamment pour des mots incluant un doublet. Dans leur modèle, chaque lettre était représentée par un nœud simple qui doit être activé afin qu'une lettre soit frappée. Un schéma de doublement doit être activé pour qu'une lettre soit doublée parce que le rappel sériel utilisé dans leur modèle nécessite que les réponses sélectionnées soient ensuite inhibées pour éviter des persévérations. Lorsque ce schéma de doublement est activé au mauvais moment, par exemple parce que du bruit a été ajouté, des erreurs de transposition surviennent (scrren à la place de screen). Le modèle « Competitive Cueing model of serial order representation in spelling » (Glasspool & Houghton, 2005; Houghton, Glasspool & Shallice, 1994) comporte un mécanisme similaire, sans lequel le modèle peut produire deux fois la même lettre si celle-ci est séparée par d'autres lettres (e.g., le *c* dans le mot *CaCtus*) mais pas deux lettres identiques immédiatement l'une après l'autre. Glasspool et Houghton (2005) ont appliqué leur modèle à des données neuropsychologiques mais ils ont souligné qu'ils manquaient d'informations sur les performances orthographiques des patients

pour pouvoir tester de façon plus approfondie leur modèle. Parmi les informations manquantes, ils mentionnaient la prise en compte ou non de régularités / conventions du langage écrit, comme le fait que certaines lettres sont souvent doublées alors que d'autres le sont rarement. Nos résultats fournissent donc de telles données dans le cas de sujets (enfants et adultes) sains.

## **Conclusion**

L'ensemble des expériences rapportées dans cette thèse contribue à une meilleure compréhension du rôle des connaissances graphotactiques dans l'acquisition et la production de l'orthographe de mots spécifiques, et plus généralement, de l'intégration de différentes sources de connaissances dans l'acquisition du langage écrit. Ces études mériteraient d'être prolongées d'au moins trois façons. Premièrement, l'expérience rapportée au deuxième chapitre expérimental devrait être conduite avec des élèves plus jeunes afin de préciser chez ces élèves le rôle du recodage phonologique (qui n'était pas corrélé avec l'apprentissage orthographique chez les CM1/CM2 de notre étude) et des connaissances orthographiques spécifiques et générales (qui étaient toutes deux corrélées avec l'apprentissage orthographique chez les CM1/CM2 de notre étude). Deuxièmement, concernant les expériences comparant le niveau d'apprentissage orthographique selon que l'orthographe des non-mots correspond ou non aux régularités graphotactiques de la langue, il serait intéressant d'étudier si et comment l'impact des connaissances graphotactiques sur l'apprentissage de l'orthographe lexicale varie lorsque les situations d'apprentissage sont plus longues, réparties sur plusieurs jours (apprentissage distribué), et/ou lorsque le délai entre la phase d'apprentissage et l'évaluation de l'apprentissage orthographique est plus long que dans les études rapportées dans cette thèse. Troisièmement, dans l'expérience rapportée au chapitre expérimental IV, l'apprentissage explicite permettait un meilleur apprentissage orthographique que l'apprentissage implicite chez des adultes mais le pattern d'erreurs de transposition (qui révèle l'impact des connaissances graphotactiques) ne variait pas en fonction du mode d'apprentissage. De nouvelles études devront examiner si ce résultat se généralise avec



d'autres propriétés graphotactiques et si et comment l'impact du mode d'apprentissage varie en fonction de l'âge des sujets.

## BIBLIOGRAPHIE

- Alegria, J., & Mousty, P. (1996). The development of spelling procedures in French-Speaking, normal and reading-disabled children: Effects of frequency and lexicality. *Journal of Experimental Child Psychology*, 63, 312-338.
- Army Individual Test Battery. (1944). *Manual of Directions and Scoring*. Washington, DC: War Department, Adjutant General's Office.
- Baron, J. & Treiman, R. (1980). Some problems in the study of differences in cognitive processes. *Memory and Cognition*, 8, 313-321.
- Barry, C. (1994). Spelling routes (or roots or ruts). In G.D.A. Brown & N.C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 27- 49). New York: John Wiley & Sons.
- Barry, C., & Seymour, P. H. K. (1988). Lexical priming and sound-to-spelling contingency effects in nonword spelling. *The Quarterly Journal of Experimental Psychology*, 40A, 5-40.
- Bartlett, F.C. (1932). *Remembering : A study in experimental and social psychology*. Cambridge : Cambridge University Press.
- Beauvois, M.F. & Desrouesné, J. (1981) Lexical or ortographic agraphia. *Brain*, 104, 21-49.
- Bergman, E., & Roediger, H. L. (1999). Can Bartlett's repeated reproduction experiments be replicated? *Memory & Cognition*, 27, 937-947.
- Berninger, V. W. (1995). The varieties of orthographic knowledge, *Vol. 2: Relations to phonology, reading, and writing*. Dordrecht, Netherlands: Kluwer Academic.
- Bolla-Wilson, K., Speedie, L. J., & Robinson, R. G. (1985). Phonological agraphia in a left handed patient after a right-hemisphere lesion. *Neurology*, 35, 1778-1781.
- Bosman, A. M. T., & Van Orden, G. C. (1997). Why spelling is more difficult than reading. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: Research, theory, and practice across languages* (pp. 173-194). Mahwah, NJ: Lawrence Erlbaum Associates.
- Bosse M-L. & Pacton, S. (2006). Comment l'enfant produit-il l'orthographe des mots ? In P. Dessus & E. Gentaz (Eds), *Apprendre et enseigner à l'école*, Paris, Dunod.
- Bosse, M. L., Valdois, S., & Tainturier, M. J (2003). Analogy without priming in early spelling development. *Reading and Writing*, 16, 693-716.
- Bosse, M.L. & Valdois, S. (2009). Influence of the visual attention span on child reading performance: a cross-sectional study. *Journal of Research in Reading*, 32, 2, 230-253.

- Bosse, M.L., Valdois, S., Tainturier, M. J. (2003). Analogy without priming in early spelling development, *Reading and Writing: An Interdisciplinary Journal*, 16, 693-716.
- Botvinick, M. & Bylisma, L. M. (2005). Regularization in short-term memory for serial order. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 31, 351-358.
- Bowey, J. A., & Miller, R. (2007). Correlates of orthographic learning in third-grade children's silent reading. *Journal of Research in Reading*, 30, 115–128.
- 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.
- Bradley, L. & Bryant, P.E. (1983). Categorising sounds and learning to read - a causal connection. *Nature*, 301, 419-421.
- Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and Categorization* (pp. 169-215). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brown A.S. (1988). Encountering misspellings and spelling performance: Why wrong isn't right. *Journal of Educational Psychology*, 4, 488–494.
- Brown, G. D. A., & Ellis, N. C. (1994). *Handbook of spelling: Theory, process and application*. Chichester: John Wiley & Sons.
- Bruck, M. & Treiman, R. (1990). Phonological awareness and spelling in normal children and dyslexics: the case of the initial consonant clusters. *Journal of Experimental Child Psychology*, 50, 156-178.
- Bryant, P. & Bradley, L. (1980). Why children sometimes write words which they do not read. In U. Frith (Ed.). *Cognitive Processes in Spelling*. (355-372). London: Academic Press.
- Campbell, R. & Butterworth, B. (1985). Phonological dyslexia and dysgraphia in a highly literate subject : A developmental case with associated deficits of phonemic processing awareness. *Quarterly Journal of Experimental Psychology*, 37A, 435-475.
- Campbell, R. (1983). Writing nonwords to dictation. *Brain and Language*, 19, 153-178.
- Campbell, R. (1985). When children write nonwords to dictation. *Journal of Experimental Child Psychology*, 57, 133–151.
- Campbell, R., & Coltheart, M. (1984). Gandhi: the nonviolent route to spelling reform? *Cognition*, 17, 185-192.
- Caramazza, A., & Miceli, G. (1990). The structure of graphemic representations. *Cognition*, 37, 243-297.
- Cassar, M., & Treiman, R. (1997). The Beginnings of Orthographic Knowledge: Children's Knowledge of Double Letters in Words, *Journal of Educational Psychology*, 89, 631-644.

- Castles, A., & Nation, K. (2006). How does orthographic learning happen? In S. Andrews (Ed.), *From inkmarks to ideas: Challenges and controversies about word recognition and reading* (pp. 151-179). Hove, East Sussex: Psychology Press.
- Cataldo, S. & Ellis, N. (1988). Interactions in the development of spelling, reading and phonological skills. *Journal of Research in Reading, 11*, 86-109.
- Coltheart, M. (1978), Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies in Information Processing* (pp. 151-216). London: Academic Press.
- Connors, F. A., Loveall, S. J., Moore, M. S., Hume, L. E., & Maddox, C. D. (2011). An individual differences analysis of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 108*, 402–410.
- Cunningham, A. E. (2006). Accounting for children’s orthographic learning while reading text: Do children self-teach? *Journal of Experimental Child Psychology, 95*, 56–77.
- 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.
- Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: a study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1355–1367.
- De Jong, P. F., & Share, D. L. (2007). Orthographic learning during oral and silent reading. *Scientific Studies of Reading, 11*, 55–71.
- 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*, 267–282.
- Deacon, S. H., & Pacton, S. (2007). Using spelling as an empirical test of rules versus statistics. *Society for the Scientific Study of Reading, Prague, CZ, July 12-15*.
- Deacon, S.H., Conrad, N., & Pacton, S. (2008). A statistical learning perspective on children’s learning about graphotactic and morphological regularities in spelling. *Canadian Psychology, 49*, 118-124.
- Deavers, R. P., & Brown, G. D. A. (1997). Rules vs. Analogies in Children's Spelling: Evidence for Task Dependence. *Reading and Writing, 9*, 339-361.
- Dixon, M., & Kaminska, Z. (1997). Is it misspelled or is it misspelled? The influence of fresh orthographic information on spelling. *Reading and Writing: An Interdisciplinary Journal, 9*, 483–498.
- Ehri, L. C. (1986). Sources of difficulty in learning to spell and read. *Advances in Developmental and Behavioural Pediatrics, 7*, 121-195.

Ehri, L. C. (1989). Apprendre à lire et à écrire les mots, in L. Rieben et C.A. Perfetti (Edit.), *L'apprenti lecteur : recherches empiriques et implications pédagogiques*, Neuchatel, Delachaux & Niestlé, 102-127.

Ehri, L. C. (1991). The development of reading and spelling in children : An overview, in M. Snowling et M. Thomson (Edit.), *Dyslexia : Integrating theory and practice*, Londres, Whurr Publishers, 63-79.

Ehri, L. C. (1992). Review and commentary: stages of spelling development. In S. Templeton, & D. R. Bear (Eds.), *Development of orthographic knowledge and the foundations of literacy: a memorial festschrift for Edmund H. Henderson* (pp. 307-332). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ehri, L. C. (1997). Learning to read and learning to spell are one and the same, almost. In C. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: research, theory and practice across languages* (pp. 237-269). Mahwah, NJ: Lawrence Erlbaum Associates.

Ellefson, M., Treiman, R., & Kessler, B. (2009). Learning to label letters by sounds or names: A comparison of England and the United States. *Journal of Experimental Child Psychology*, 102, 323-341.

Ellis, A.W. & Young, A.W. (1988), *Human Cognitive Neuropsychology*. London : Lawrence Erlbaum Ass.

Ellis, N.C. (1994). Longitudinal studies of spelling development. In Brown, G.D.A., & Ellis, N.C. (Eds.). *Handbook of Spelling: Theory, Process and Intervention*. (pp. 155-178). Chichester: John Wiley & Sons Ltd.

Foulin, J.N., (2005). Why is letter-name knowledge such a good predictor of learning to read? *Reading and Writing*, 18, 129-155.

Frith U. (1985). Beneath the surface of developmental dyslexia, in K. Patterson, J. Marshall et M. Coltheart (Edit.), *Surface dyslexia : Neuropsychological and cognitive studies of phonological reading*, Hillsdale (Nj), Lawrence Erlbaum, 301-330.

Frith, U. (1980). *Cognitive Processes in Spelling*. London: Academic Press.

Frith, U. (1986). A Developmental Framework for Developmental Dyslexia. *Research Reflections. Neuropsychology*, p. 69-81.

Gak, V. G. (1976). *L'orthographe française*. Paris: SELAF.

Gentry, J. R. (1982). An analysis of developmental spelling in GNYS AT WRK. *The Reading Teacher*, 36, 192-200.

Gilbert, L. C. (1934a). Effect of reading on spelling in the ninth grade. *School Review*, 42, 197-204.

Gilbert, L. C. (1934b). Effect of reading on spelling in the secondary schools. *California Quarterly of Secondary Education*, 9, 269-275.

- Gilbert, L. C. (1935). A study of the effect of reading on spelling. *Journal of Educational Research*, 28, 570–576.
- Glasspool, D. W. and Houghton, G. (2005) Serial order and consonant-vowel structure in a graphemic output buffer model. *Brain and Language*, 94(3), 304-330.
- Gomez, R. L., Gerken, L., & Schvaneveldt, R. W. (2000). The basis of transfer in artificial grammar learning. *Memory and Cognition*, 28, 253-263.
- Goodman, R. A. & Caramazza, A. (1986) Aspects of the spelling process: Evidence from a case of acquired dysgraphia. *Language and Cognitive processes*, 1 (4), 263-296.
- Goswami, U. (1988). Children's use of analogy in learning to spell. *British Journal of Developmental Psychology*, 6, 21-33.
- Hatfield, F. M. & Patterson K. E. (1983) Phonological spelling. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 35, 451-468.
- Hayes, H., Treiman, R., & Kessler, B. (2006). Children use vowels to help them spell consonants. *Journal of Experimental Child Psychology*, 94, 27-42.
- Henderson, L. (1985). *Teaching spelling*. Boston: Houghton Mifflin.
- Houghton, G., Glasspool, D.W., and Shallice, T., (1994). Spelling and serial recall: Insights from a competitive queueing model. In G.D.A. Brown and N.C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention*. Wiley: Chichester.
- Hyman, I.E. & Loftus, E.F. (1998). Errors in autobiographical memory. *Clinical Psychology Review*, 18, 933–947.
- Jaffré, J.P. & Fayol, M. (1997). *Orthographes. Des systèmes aux usages*. Paris : Flammarion.
- Jorm, A. F., & Share, D. L. (1983). Phonological recoding and reading acquisition. *Applied Psycholinguistics*, 4, 103–147.
- Kemp, N., & Bryant, P. (2003). Do bees buzz? Rule-based and frequency-based knowledge in learning to spell plural –s. *Child Development*, 74, 63-74.
- Kyte, C. S., & Johnson, C. J. (2006). The role of phonological recoding in orthographic learning. *Journal of Experimental Child Psychology*, 93, 166–185.
- Laiacona, M., Capitani, E., Zonca, G., & Scola, I. (2009). Integration of lexical and sublexical processing in the spelling of regular words: A multiple single–case study in Italian dysgraphic patients, *Cortex*, 45, 804-815.
- Landi, N., Perfetti, C., Bolger, D., Dunlap, S., & Foorman, B. (2006). The role of discourse context in developing word form representations: A paradoxical relation between reading and learning. *Journal of Experimental Child Psychology*, 94, 114–133.

- Lehtonen, A., & Bryant, P. (2005). Doublet challenge : form comes before function in children's understanding of their orthography, *Developmental Science*, 8, 211-217.
- Lennox, C. & Siegel, L.S. (1994). The role of phonological and orthographic processes in learning to spell. In G.D.A Brown & N.Ellis (Eds.). *Handbook of spelling: Theory, process and intervention* (pp. 93-109). Toronto: Wiley.
- Loftus, E.F., & Palmer, J.C. (1974). Reconstruction of Automobile Destruction : An Example of the Interaction Between Language and Memory. *Journal of verbal learning and verbal behavior*, 13, 585-589.
- Manza, L., & Reber, A. S. (1997). Representing artificial grammar: Transfer across stimulus forms and modalities. In D. Berry (Ed.). *How implicit is implicit learning* (pp. 431-454). Oxford: Oxford University Press.
- Marsh, G., Friedman, M., Welch, V., & Desberg, P. (1980). The development of strategies in spelling. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 339-354). London: Academic Press.
- Martinet, C., Bosse M. L., Valdois, S. & Tainturier, M. J. osèphe (1999). Existe-t-il des stades successifs dans l'acquisition de l'orthographe d'usage ? *Langue française*, 124, 58-73.
- Martinet, C., Bosse, M. L., Valdois, S., & Tainturier, M. J. (1999). Existe-t-il des stades successifs dans l'acquisition de l'orthographe d'usage ? *Langue française*, 124, 58-73.
- Martinet, C., Valdois, S., & Fayol, M. (2004) Lexical orthographic knowledge develops from the beginning of literacy acquisition, *Cognition*, 91, 2004, pp. B11-B22.
- Matinet, C., Valdois, S., Fayol, M. (2004). Lexical orthographic knowledge develops from the beginning of literacy acquisition, *Cognition*, 91, B11-B22.
- McCloskey, M., Badecker, W., Goodman-Schulman, R. A., & Aliminosa, D. (1994). The structure of graphemic representations in spelling: Evidence from a case of acquired dysgraphia. *Cognitive Neuropsychology*, 11, 341-392.
- Miceli, G., Capasso, R., & Caramazza, A. (1994). The interaction of lexical and sublexical processes in reading, writing and repetition. *Neuropsychologia*, 32, 317-333
- Morton, J. (1980). The logogen model and orthographic structure. In U. Frith (Ed.), *Cognitive processes in spelling*. London: Academic Press.
- Mousty, P. Leybaert, J. Alegria, J., Content, A., & Moraïs, J. (1994) Bélec. In, Grégoire, J., Piérart, B. (Eds.), *Evaluer les troubles de la lecture. Les nouveaux modèles théoriques et leurs implications diagnostiques*. 127-145. Bruxelles : De Boeck.
- Nation, K. (1997). Children's sensitivity to rime unit frequency when spelling words and nonwords. *Reading and Writing*, 9, 321-338.

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, 71–84.

Nation, K., Hulme, C. (1996). The automatic activation of sound-letter knowledge: An alternative interpretation of analogy and priming effects in early spelling development. *Journal of Experimental Child Psychology* 63, 416-435

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicale du français contemporain sur internet : Lexique. *L'Année Psychologique*, 101, 447-462.

Nicolson, R.I. and Fawcett, A.J. (1994). Reaction times and dyslexia. *Quarterly Journal of Experimental Psychology*, 47A, 29-48.

Ormrod, J.E. (1986). Learning to spell while reading : a follow-up study. *Perceptual and Motor Skills*, 63, 652-654.

Pacton, S. & Perruchet, P. (2006). L'apprentissage implicite : du labo à l'école. In P. Dessus & E. Gentaz (Eds), *Comprendre les apprentissages, Vol 2*, Paris, Dunod, pp. 59-73.

Pacton, S., & Deacon, S.H. (2008). The timing and mechanisms of children's use of morphological information in spelling: A review of evidence from French and English. *Cognitive Development*, 23, 339-359.

Pacton, S., Fayol, M., & Perruchet, P. (2002). The acquisition of untaught orthographic regularities in French. In *Precursors of Functional Literacy*, L. Verhoeven, C. Erlbro, d P. Reitsma (eds), 121-136. Dordrecht : Kluwer.

Pacton, S., Fayol, M., & Perruchet, P. (2002). The acquisition of untaught orthographic regularities in French . In L. Verhoeven, C. Elbro, et P. Reitsma (Eds.), *Precursors of Functional Literacy* (pp 121-136). Dordrecht: Kluwer.

Pacton, S., Fayol, M., & Perruchet, P. (2005). Children's implicit learning of graphotactic and morphological regularities. *Child Development*, 76, 324-339.

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, 401-426.

Perruchet, P. & Pacton, S. (2004). Qu'apportent à la pédagogie les travaux de laboratoire sur l'apprentissage implicite ? *L'Année psychologique*, 14, 121-146.

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, 119, 264-275.

Pollo, T. C., Treiman, R., & Kessler, B. (2008). Three perspectives on spelling development. In E. J. Grigorenko & A. Naples (Eds.), *Single-word reading: Cognitive, behavioral, and biological perspectives* (pp. 175-189). Mahwah, NJ: Lawrence Erlbaum Associates.



- Rapp, B., Epstein, C., & Tainturier, M.J. (2002). The integration of information across lexical and sublexical processes in spelling, *Cognitive Neuropsychology*, *19* (1), 1-29.
- Read, C. (1986). *Children's creative spelling*. London: Routledge and Kegan Paul.
- Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*, 855-863.
- Reber, A.S. (1993). *Implicit learning and tacit knowledge: an essay on the cognitive unconscious*. New York: Oxford University Press.
- Redington, M., & Chater, N. (1996). Transfer in artificial grammar learning: A re-evaluation. *Journal of Experimental Psychology: General*, *125*, 123-138.
- Rittle-Johnson, B. R., & Siegler, R. S. (1999). Learning to spell: Variability, choice, and change in children's strategy use. *Child Development*, *70*, 332-349.
- Roeltgen, D. P., & Heilman, K. M. (1984). Lexical agraphia: Further support for the two system hypothesis of linguistic agraphia. *Brain*, *101*, 811-827.
- Roeltgen, D. P., Sevush, S., & Heilman, K. M. (1983). Phonological agraphia: Writing by the lexical-semantic route. *Neurology*, *33*, 755-765.
- Rumelhart, D.E. and Norman, D.A. (1982) Simulating a skilled typist: a study of skilled cognitive-motor performance, *Cognitive Science*, *6*, 1-36.
- Seymour, P. H. K. (1997). Foundations of orthographic development. In C. Perfetti, L. Rieben & M. Fayol (Eds.), *Learning to spell: Research, theory, and practice across languages* (pp. 319-337). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Seymour, P. H. K., & Elder, L. (1986). Beginning reading without phonology. *Cognitive Neuropsychology*, *3*, 1-36.
- Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143-174.
- Shallice, T. (1981) Phonological agraphia and the lexical route in writing. *Brain*, *104*, 413-429.
- Shanks, D. R., Johnstone, T., & Staggs, L. (1997). Abstraction processes in artificial grammar learning. *The Quarterly Journal of Experimental Psychology*, *50A*, 216-252.
- Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, *55*, 151-218.
- Share, D. L. (1999). Phonological recoding and orthographic learning: a direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *72*, 95-129.
- Share, D. L. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology*, *87*, 267-298.

- Share, D.L. & Shalev, C. (2004). Self-teaching in normal and disabled readers. *Reading and Writing*, 17, 1-31.
- Smith, F. (1971). *Understanding reading: a psycholinguistic analysis of reading and learning to read*. New York: Holt, Rinehart & Winston.
- Snowling, M. J. (1994). Towards a model of spelling acquisition : the development of some component skills. In G.D.A. Brown, & N.C. Ellis (Eds), *Handbook of Spelling : Theory, Process and Intervention*. Chichester. John Wiley & Sons Ltd. 111-128.
- Sprenger-Charolles, L. & Béchennec, D. (2004), Variability and invariance in learning alphabetic orthographies. *Written Language*, 9-33.
- Sprenger-Charolles, L. & Casalis, S. (1995). Reading and spelling acquisition in French first graders : Longitudinal evidence. *Reading and Writing: An Interdisciplinary Journal*, 7,39-63.
- Sprenger-Charolles, L. (2003). Linguistic processes in reading and spelling. The case of alphabetic writing systems: English, French, German and Spanish. In T. Nunes, & P. Bryant (Eds) *Handbook of children's literacy*. (pp. 43-65). Dordrecht: Kluwer Academic Publishers.
- Sprenger-Charolles, L., Siegel, L. S., Béchennec, D., & Serniclaes, W (2003). Development of phonological and orthographic processing in reading aloud, in silent reading and in spelling: a four-year longitudinal study. *Journal of Experimental Child Psychology*,84, 194-217.
- Sprenger-Charolles, L., Siegel, L.S. & Bonnet, P. (1998) Reading and spelling acquisition in French: the role of phonological mediation and orthographic factors. *Journal of Experimental Child Psychology*, 68(2), 134–165.
- Surprenant, A. M., & Neath, I. (2009). The 9 lives of short-term memory. In A. Thorn & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 16-43). Hove, UK: Psychology Press.
- Tainturier, M.J., & Caramazza, A. (1996). The status of double letters in graphemic representations. *Journal of Memory and Language*, 35, 53-73.
- Tainturier, M.J. & Rapp, B. (2001),The spelling process. In B. Rapp (Ed.), *What deficits reveal about the human mind: A handbook of cognitive neuropsychology*. Philadelphia, PA: Psychology Press.
- Thorstad, G. (1991), The effect of orthography on the acquisition of literacy skills. *British Journal of Psychology*, 82, 527-537.
- Treiman, R. (1993) *Beginning to spell: A study of first-grade children*. New York: Oxford University Press.
- Treiman, R. (1993). *Beginning to spell: A study of first-grade children*. New-York: Oxford University Press.

Treiman, R., Berch, D. & Weatherston, S. (1993). Children's use of phoneme-grapheme correspondences in spelling: Roles of position and stress. *Journal of Educational Psychology*, 85, 1-12.

Treiman, R., & Cassar, M. (1997). Spelling acquisition in English. In C.A. Perfetti, L. Rieben, M. Fayol (Eds.), *Learning to spell: Research theory, and practice across languages* (pp. 61-80). Mahwah, NJ: Erlbaum.

Treiman, R., & Kessler, B. (2006). Spelling as statistical learning: Using consonantal context to spell vowels. *Journal of Educational Psychology*, 98, 642-652.

Treiman, R., Zukowski, A. & Richmond-Welty, E.D. (1995). What happened to the "n" of sink ? Children's spellings of final consonant clusters. *Cognition*, 55, 1-38.

Valdois, S., Bosse, M.L., Ans B., Zorman, M., Carbonnel, S., David, D. & Pellat, J. (2003). Phonological and visual processing deficits are dissociated in developmental dyslexia: Evidence from two case studies. *Reading and Writing*, 16, 543-572.

Varnhagen, C.K, McCallum, M., & Burstow, M. (1997). Is children's spelling naturally stage-like? *Reading and Writing: An Interdisciplinary Journal*, 9, 451-481.

Véronis, J. (1988), From sound to spelling in french: simulation on a computer. *Cahiers de Psychologie Cognitive*, 8, 315-334.

Wang, H.-C., Castles, A., Nickels, L., & Nation, K. (2011). Context effects on orthographic learning of regular and irregular words. *Journal of Experimental Child Psychology*, 109, 39-57.

Wimmer, H., & Mayringer, H. (2002). Dysfluent reading in the absence of spelling difficulties : A specific disability in regular orthographies. *Journal of Educational Psychology*, 94, 272-277.

Wright, D.M., & Ehri, E. (2007) Beginners remember orthography when they learn to read words: the case of doubled letters. *Applied Psycholinguistics*, 28, 115-133.

Ziegler, J., Jacobs, A., & Stone, G. (1996). Statistical analysis of the bi-directional inconsistency of spelling and sound in French. *Behavior, Research Methods, Instruments, & Computers*, 28, 504-515.

## ANNEXES

### ANNEXE 1

#### Paires de non-mots évaluant les connaissances graphotactiques :

Items identité des consonnes pouvant être doublées (n = 12) :

ojjave	onnavé
bihhot	billot
oxxile	ommile
boxxit	bottit
akkoge	attoge
ujjate	ullate
tummet	tukket
exxiva	effiva
irramé	ijjamé
merral	mehhal
befful	bekkul
tinnas	tihhas

Items illégalité des doublets en position initiale (n = 16) :

*position F+ :*

nnumar	nummar
raffout	rrafout
niffor	nnifor
ffanous	fannous
tinnot	ttinot
rramin	rammin
ttamir	tammir
ffurois	furrois

*position F0 :*

jjikol	jikkol
xxakir	xakkir
zikkel	zzikel
zaxxil	zzaxil
kkaxort	kaxxort
jjaxir	jaxxir
xijjer	xxijer
kkujire	kujjire

Items illégalité des doublets après une consonne simple (n = 16) :

*groupes consonantiques F+ :*

aplover	aplover
aprrulir	apprulir
afrrunir	affrunir
acrriver	acrriver
acclomir	acclomir
gabiffler	gabiffler
moufrrive	mouffrive
ripllave	ripllave

*groupes consonantiques F- :*

maiffrude	maiffrude
rupllave	rupplave
yffroner	yffroner
iprraler	iprraler
icrraver	iccraver
ucclimer	ucllimer
opllavir	opplavir
giboffler	giboffler

Probabilité de eau en fonction de l'environnement gauche (n= 12)

*Eau f+ :*

suléveau	sulévot
driveau	drivot
frasseau	frassot
falisseau	falissot
glirot	glireau
lavireau	lavirot

*Eau f-*

tilagueau	tilagot
birgot	birqueau
gircot	girqueau
vuraqueau	vuracot
lirfeau	lirfot
durifot	durifeau

## ANNEXE 2

### Décision orthographique sur des mots

#### Doublet illégal en début de mots (n=9)

risotto	rrisoto
mmapemonde	mappemonde
bbaricade	barricade
raffut	rrafut
nnaperon	napperon
nnaratrice	narratrice
raffinage	rraifinage
mmenote	menotte
commutateur	ccomutateur

#### Doublet illégal après une consonne simple (n=9)

supplications	supllications
acroc	accroc
affront	afrfront
appréhension	aprréhension
raprochement	rapprochement
suffrages	suffrages
coffrage	cofrrage
appréciation	aprréciation
aggravation	agrravation

#### Paires tampons n = 12

alto	alteau
sombrero	sombrereau
sureau	suro
recteau	recto
apéreau	apéro
boqueteau	boqueto
coto	coteau
bigarreau	bigarro
écriteau	écrito
perdro	perdreau
porto	porteau
brasero	brasereau

## ANNEXE 3

### Test de lecture : Sous-test L3 de la batterie ORLEC (Lobrot, 1973)

Epreuve préliminaire :

- 01) Je ferai la vaisselle demain matin, car je suis fatigué et je préfère aller au (bout, loup, jour, lit, loin).
- 02) Si on fait marcher trop fort sa radio, on risque de déranger les (poissons, mains, coins, voisins, trains).
- 03) Mon frère a fait un voyage en Afrique et a rapporté une très originale (ville, statue, chaleur, estrade, saison).
- 04) Un homme qui conduit un véhicule s'appelle un (mécanicien, compagnon, conducteur, chanteur).

Epreuve :

- 1) Prends le panier et va m'acheter des (armoires, oranges, ordures, ombres, ordres).
- 2) Si vous mangez ce gâteau, dit ma mère, vous verrez comme il est (long, rond, bon, doux, chou).
- 3) Tous les chiens ont quatre (bouches, pattes, pinces, prunes, oreilles).
- 4) J'aimerais aller sur la plage pour me baigner dans la (guerre, mer, mère, marche, marque).
- 5) La petite fille a mis sa (roche, cloche, roue, rue, robe).
- 6) La gare se trouve au milieu de la (fille, ville, bille, boule, poule).
- 7) Il a ouvert la radio et a écouté les (nouvelles, chandelles, voiles, vitres, navires).
- 8) Il a déchiré son tablier et il s'est fait (rouler, grandir, sonder, craindre, gronder).
- 9) Un endroit où on range des livres s'appelle une (pêche, cuisine, galerie, bibliothèque, porte).
- 10) Il y a eu un grand accident : la locomotive est sortie des ( tiroirs, rails, rayons, routes, rangs).
- 11) Ils travaillent toute la journée et le soir ils se (noient, brisent, sèchent, répondent, reposent).
- 12) Vous pourriez enlever la poussière avec un (palais, balai, boeuf, lard, four).
- 13) Il est parti à la chasse, c'est pourquoi il a pris son (outil, feu, fusil, gentil, foin).
- 14) Mon oncle, après de longues études, est devenu (nouille, médecin, moisi, monsieur, moyen).
- 15) Il s'est penché sur le puits et il est tombé au (fond, front, frein, fard, four).
- 16) Il fait chaud sur la terrasse, pourquoi ne mettez-vous pas le (paravent, radiateur, parasol, passage, patin) ?

- 17) Quand on est dans la rue, il faut faire très attention aux autos afin de ne pas se faire (laver, transporter, casser, pousser, écraser).
- 18) Quand vous dormirez, j'espère que vous ferez de jolis (rêves, yeux, trous, rires, cous).
- 19) Parmi tous les jeux préférez-vous le ping-pong, le billard, les dominos, ou les (douches, astres, bras, cartes, cadres) ?
- 20) Il s'est pris la main dans la porte et il s'est mis à pleurer en poussant des (bruits, lits, nuits, cris, cas).
- 21) Un camarade l'a poussé et il est tombé sur les (roues, mains, nains, vins, ponts).
- 22) Tout le monde est parti en voiture jusqu'à la forêt et là, nous nous sommes assis sur l'herbe, où nous avons mangé notre (rat, rang, repas, quart, pas).
- 23) Ils comptent aller aux courses dimanche prochain car ils aiment voir les chevaux courir sur la (piste, liste, voûte, route, mine).
- 24) Il est arrivé une drôle d'aventure à un pêcheur, il a attrapé une (carpe, tanche, godasse, truite, perche).
- 25) Du cratère du volcan s'échappent peu à peu des flots de (vague, lave, bave, cave, rage).
- 26) Pourquoi ne vous servez-vous pas d'un couteau pour manger votre (vin, voiture, viande, voisin, ville) ?
- 27) Tous les gens sont sortis de leur maison et ont regardé les dégâts produits par l' (explosion, exposition, ascension, expédition, exagération).
- 28) Nos voisins ont acheté un gros chien méchant qui doit rester devant la porte pour monter la (corde, fuite, chaîne, grade, garde).
- 29) C'est l'hiver, et cette nuit sont tombés de gros (flacons, cocons, flocons, sapins, sabots).
- 30) Nous sommes allés nous promener dans la forêt et nous avons rapporté des (chalets, champignons, châtaignes, châteaux, chapeaux).
- 31) C'est le printemps, les bois sont fleuris de (quilles, jongleurs, jonques, jonquilles, feuilles).
- 32) La fatigue, le surmenage, ont rendu cette personne (alerte, petite, aimable, maligne, souffrante).
- 33) Le prestidigitateur en plantant un couteau dans la paume de sa main nous a (payés, effacés, fouillés, effrayés, ensanglantés).
- 34) Les hommes aiment ce qui est nouveau parce que cela satisfait leur (bonté, amitié, curiosité, vanité, justice).
- 35) Le mari d'une fille est pour la mère de cette fille un (géant, agent, gendre, geôlier, gendarme).
- 36) Les réfrigérateurs empêchent la nourriture de (mourir, rouiller, se souiller, geler, pourrir).



## **ANNEXE 4**

### **Test orthographique : Sous-test "Le Corbeau" de la batterie L2MA (Chevrier-Muller, Simon & Fournier, 1997)**

Chaque "/" indique les pauses faites par l'expérimentateur. La dictée est lue une première fois entièrement, puis chaque segment est lu deux fois. Enfin l'expérimentateur relit lentement la dictée dans sa totalité. La deuxième partie n'est dictée qu'à partir du CM2.

Le corbeau :

Un corbeau / perché sur l'antenne d'un bâtiment / tient dans son bec / une souris blessée /.  
Rendus furieux / par cet oiseau cruel, / des enfants lancent des cailloux / pour l'obliger à  
s'envoler /.

Le corbeau les a observés / puis a déployé ses ailes / et s'est élancé / en lâchant la souris / que  
les enfants / vont recueillir et soigner /.

## ANNEXE 5

### **Test orthographique : dictée de mots isolés (Martinet et Valdois, 1999)**

amour, carpe, confiture, finale, louche, moto, ourson, rage, tribune, verbe, bouture, charge, frite, marmite, soucoupe, individu, miroir, mérite, soda, globule, culbute, inutilité, bain, dentiste, relation, terre, figue, océan, vigne, boucher, cuvette, cuisson, pharmacie, sauvetage, cuisinier, témoin, flan, freiner, mimosa, ruban, gitan, gourdin, faucon, angine, femme, hiver, œuf, pied, piscine, monsieur, tabac, examen, faon, haricot, rayure, agenda, gentil, habit, aquarium, clown, gruyère, technique, poêle, cageot, beignet, baptême.

## ANNEXE 6

### Test orthographique : épreuve de jugement

tablette	tablète	tablète
abracadabrat	abracadabras	abracadabra
ascidulé	acidulé	assidulé
quête	quette	quête
encard	encart	encar
pissenlid	pissenlie	pissenlit
bassinnet	bascinet	bacinet
indigos	indigo	indigot
contemporein	contemporin	contemporain
cabanon	quabanon	kabanon
proue	prout	proup
bananier	bananié	bananiet
nénuphar	nénuphard	nénuphart
artichaut	artichau	articheau
atoup	atoue	atout
joyeau	joyaut	joyau
fourberie	fourberit	fourberid
loyautet	loyauter	loyauté
sélérat	scélérat	célérat
angora	hangora	engora
bénévola	bénévolat	bénévolas
épinart	épinar	épinard
acension	assension	ascension
quangourou	kangourou	cangourou
chao	chaot	chaos
frain	frein	frin
repas	repa	repat
ailier	ailied	ailiet
saynête	saynète	saynette
tourtau	tourteau	tourtaut
baudé	baudet	bauder
nit	nid	nie
endicap	andicap	handicap
coussined	coussinet	coussiner
ciboulette	siboulette	sciboulette
ballot	ballos	ballo
gradin	gradain	gradein
hantraver	entraver	antraver
cout	coup	coue
trépiet	trépier	trépied
sciamois	ciamois	siamois
kadrille	cadrille	quadrille

## ANNEXE 7

### Conscience phonologique

#### Contrepéties

Exemples : pipe-robe (ripe-pobe)    sucre-tigre (tucre-sigre)    vase-lion (lase-vion)

	Réponse attendue	Réponse juste	erreur
Banane - ficelle	Fanane Bicelle	+	
Fourmi – journal	Jourmi Fournal	+	
Mouton – tulipe	Touton Mulipe	+	
Manège – volcan	Vanège Molcan	+	
Jardin – bagage	Bardin Jagage	+	
Tomate – camion	Comate Tamion	+	
Verrou – falaise	Ferrou Valaise	+	
Panier – bureau	Banier Pureau	+	
Baleine – miroir	Maleine Biroir	+	
Bougie – poupée	Pougie Boupée	+	
<u>SCORE TOTAL</u>			

## Omission du phonème

Exemples : envoi

orange

caverne

placard

	Réponse attendue	Réponse juste	erreur
Outil	Ti	+	
Fontaine	Ontaine	+	
Drapeau	Rapeau	+	
Courage	Ourage	+	
Ombrage	Brage	+	
Planète	Lanète	+	
Droit	Roit	+	
Frite	Rite	+	
Capitaine	Apitaine	+	
Ampoule	Poule	+	
Dragon	Ragon	+	
Virgule	Irgule	+	
Statue	Tatue	+	
Gravier	Ravier	+	
Hiver	Ver	+	
Impossible	Possible	+	
Envoyer	Voyer	+	
Planche	Lanche	+	
Grappe	Rappe	+	
Oubli	bli	+	
<b>SCORE TOTAL</b>			

### Décomposition en phonèmes

Exemples : pour                      chasse                      menton                      planche

	Réponse attendue	Réponse juste	erreur
Four	f-u-r	+	
Epaule	e-p-o-l	+	
Chauffage	S-o-f-a-j	+	
Cri	k-r-i	+	
Porte	p-o-r-t	+	
Allumer	a-l-y-m-e	+	
Outil	u-t-i	+	
Plage	p-l-a-j	+	
Crapaud	k-r-a-p-o	+	
Encre	ã-k-r	+	
Cadeau	k-a-d-o	+	
Régler	r-e-g-l-e	+	
jouet	j-u-e	+	
abri	a-b-r-i	+	
Armure	a-r-m-y-r	+	
<b>SCORE TOTAL</b>			

## ANNEXE 8

### Non-mots de l' Expérience 1 et 4 du Chapitre Expérimental III

Sans Doublet	Doublet fréquent	Doublet rare
bagotin	bagottin	baggotin
finodus	finnodus	finoddus
liroban	lirroban	lirobban
maboron	maborron	mabboron
pitegure	pittegure	piteggure
tidunar	tidunnar	tiddunar

### Non-mots de l' Expérience 2, 3 et 5 du Chapitre Expérimental III

Sans Doublet	Doublet fréquent	Doublet rare		Items Tampons
bagotou	bagottou	baggotou		bupila
batogou	battogou	batoggou		moufopin
coubatan	coubattan	coubbatan		tocali
coutaban	couttaban	coutabban		
fagona	fagonna	faggona		
fanoga	fannoga	fanogga		
lobanin	lobannin	lobbanin		
lonabin	lonnabin	lonabbin		
maboron	maborron	mabboron		
marobon	marrobon	marobbon		
miduré	midurré	midduré		
mirudé	mirrudé	miruddé		
sidaté	sidatté	siddaté		
sidadé	sittadé	sitaddé		
tidunar	tidunnar	tiddunar		
tinadur	tinnadur	tinaddur		
vagoran	vagorran	vaggoran		
varogan	varrogan	varoggan		

## ANNEXE 9

### Test orthographique : épreuve de décision orthographique

hirondèle
sollicitude
facination
balistique
susurrer
abri
hameau
impotente
carrie
planette
paquebeau
collis
dégrafer
fourrager
esquimeau
épous
étayage
barraque
hublot
antenne
remou
layette
pantain
poulin
vacciner
allouette
haleine
occurrence
parcimonie
baffouer



## Résumé :

L'objectif de cette thèse est d'étudier le rôle des connaissances orthographiques générales, appelées connaissances graphotactiques, dans l'acquisition de l'orthographe lexicale (i.e., l'orthographe de mots spécifiques) chez des élèves de l'école élémentaire (CE2 et CM2) et des adultes.

Une première expérience (chapitre expérimental I) confirme des études antérieures montrant que des élèves de l'école élémentaire sont sensibles à diverses régularités graphotactiques, certaines plus difficilement repérées que d'autres (e.g., Danjon & Pacton, 2009 ; Pacton et al., 2001), et les prolonge en montrant que même des adultes ne semblent pas (systématiquement) recourir à des règles spécifiant par exemple que les consonnes doubles peuvent survenir avant mais pas après les consonnes simples.

Une seconde expérience (chapitre expérimental II) précise le rôle des connaissances en orthographe lexicale et celui des connaissances graphotactiques lors de l'apprentissage implicite de l'orthographe de non-mots insérés dans des textes que des élèves de CM1 et CM2 lisent pour les comprendre (situation de *self-teaching*, Share, 1995, 1999). Des analyses de régression montrent en effet que les deux types de connaissances orthographiques expliquent des parts de variance indépendantes du niveau d'apprentissage de l'orthographe des non-mots, même après avoir contrôlé diverses habiletés (conscience phonémique, recodage phonologique, capacité à produire des orthographe phonologiquement plausibles).

Les six expériences suivantes (chapitres expérimentaux III et IV) explorent si et comment les connaissances graphotactiques influencent l'apprentissage de l'orthographe lexicale dans diverses situations d'apprentissage de non-mots en manipulant le degré de convergence entre les orthographe des non-mots et les régularités graphotactiques de la langue française. Deux propriétés graphotactiques sont manipulées, l'une de type probabiliste (certaines consonnes sont plus fréquemment doublées que d'autres), l'autre descriptible sous forme de règle (les consonnes doubles peuvent survenir avant mais pas après les consonnes simples). Dans diverses situations d'apprentissage (implicite vs. explicite ; non-mots insérés dans des textes vs. présentés isolément ; lecture silencieuse vs. à haute voix) et de test (rappel vs. reconnaissance), les résultats montrent que des élèves de CE2 et CM2 et des adultes acquièrent des connaissances relatives à l'orthographe de mots spécifiques (e.g., la présence d'un doublet) mais se fondent également sur leurs connaissances relatives aux régularités graphotactiques pour reconstruire l'orthographe des mots. Ce processus de reconstruction

explique que certaines orthographes sont plus difficiles à mémoriser que d'autres ainsi que le caractère systématique de certaines erreurs. Par exemple, les élèves et les adultes, qui se rappellent de la présence d'un doublet dans un item spécifique mais ne se rappellent pas de la lettre spécifique qui est doublée, semblent se fonder sur leur sensibilité au fait que certaines consonnes sont plus souvent doublées que d'autres, ce qui explique que les erreurs de transposition d'une consonne rarement doublée vers une consonne fréquemment doublée (*tiddunar* orthographié TIDUNNAR) sont beaucoup plus nombreuses que les erreurs de transposition d'une consonne fréquemment doublée vers une consonne rarement doublée (*tidunnar* orthographié TIDDUNAR).

L'ensemble de ces expériences contribue donc à une meilleure compréhension du rôle des connaissances graphotactiques dans l'acquisition et la production de l'orthographe de mots spécifiques, et plus généralement, de l'intégration de différentes sources de connaissances dans l'acquisition du langage écrit.