

**Université de Haute-Bretagne – Rennes 2
UFR de Sciences Humaines
CRPCC – EA 1285**

**REGULARITES GRAPHOPHONOLOGIQUES, ORTHOGRAPHIQUES
ET MORPHOLOGIQUES : APPRENTISSAGE IMPLICITE ET IMPACT
PRECOCE SUR LA LECTURE**

**Thèse pour obtenir le grade de DOCTEUR DE L'UNIVERSITE DE HAUTE-
BRETAGNE RENNES 2**

Discipline : Psychologie

Présentée et soutenue le 13 septembre 2005 par Anne-Sophie Rocher

Sous la direction de Lucile Chanquoy et de Jean-Emile Gombert

Jury

**Pr. Lucile Chanquoy, Université de Nice Sophia-Antipolis, Nice
Pr. Michel Fayol, Université Blaise Pascal, Clermond-Ferrand
Pr. Jean-Emile Gombert, Université de Haute-Bretagne, Rennes 2
Pr. Annie Magnan, Université Lyon 2 Lumière, Lyon
Pr. Béatrice Pothier, Université Catholique de l'Ouest, Angers**

Remerciements

A Lucile Chanquoy,

Mes remerciements très chaleureux pour m'avoir suivie depuis le DEA jusqu'à la fin du doctorat avec beaucoup de rigueur, une très grande disponibilité sans oublier votre convivialité et votre écoute compréhensive qui se sont avérées indispensables dans les moments de doute. Sans vous, ce travail n'aurait jamais trouvé sa fin.

A Jean-Emile Gombert,

Vos conseils, votre présence, vos encouragements m'ont éclairée dans ce travail. Vous m'avez permis de faire des rencontres riches sur les plan scientifique et personnel que je n'oublierai pas. Très sincèrement merci.

Aux membres du jury,

Je vous remercie d'avoir accepté de participer au jury de cette thèse et de l'intérêt manifesté pour ce travail.

A Nathalie Marec-Breton pour tes précieux conseils et ta grande disponibilité à mon égard : nos échanges m'ont beaucoup aidée au niveau théorique, technique et personnel. A Gaspard Breton, un « informaticien » hors pair !

Aux enfants et aux enseignantes des écoles privées de La Chaume, de Nantes et de Sainte Luce-sur-Loire qui ont participé aux expériences de cette thèse et à qui j'ai demandé beaucoup. Leur enthousiasme et leur curiosité ont été salvateurs.

Aux étudiantes de Master 1 de Psychologie de l'Université de Nantes qui ont collecté une partie des données expérimentales : sans elles, ce travail n'aurait pas été possible.

A l'équipe de l'Institut Médico-Educatif Chanzy de Nantes pour m'avoir accueillie en stage durant l'année 2004-2005 : mes remerciements chaleureux à Bruno Alexandre, Claire Raulin, Fabienne Leguère, Cécile Lecoq, Gilles Sergent, Cécile Barotin et l'équipe éducative qui ont contribué indirectement à ce travail.

A Frédéric P., Magalie C., Anne-Muriel C-C., Solène M., Anne-Laure G., David I., Christophe J., Nadège V., Bruno C., Yvon G., Pierre-Henri G., Emmanuel A. et Laurent G. de l'Université de Nantes pour votre aide de tous les instants...et pour m'avoir supportée pendant trois années !

A Hélène L., Marine R., Laëtitia P., Mélissa A., Karen B., Grégory M., Stéphane B., Briac A., Arnaud B., Roxanne R., Morgane B., Damien B., Cathy A., Sandrine B., François G., Véronique C., Cédric V., Maïfe V., Joël G., Christophe G., Jérôme G., Liza L., Romain C., Anaïck C., Eve M., Thomas M., Anne L., Laëtitia C. et Pierre-Yves M. pour votre soutien et votre légèreté.

A mes parents, ma sœur Sandrine et Noëlla : depuis le début, vous avez toujours été présents pour m'encourager dans tous mes projets. Merci.

Enfin, à François pour son accompagnement sans égal dans tous mes choix... .

« ...ouvrez vos livres Paul et Valérie, page 3. Paul et Valérie, leur chien, leur maison, leurs parents... . Ve et a, ça fait va. Le et é, ça fait lé. Re et i, ça fait ri. Ce qui paraissait simple devient compliqué, c'est comme ça... . Des Valérie on en connaît des tonnes, mais la Valérie du livre, c'est différent, elle répond à d'autres règles. Tout devient abstraction. On dit place mais placard. Ce et e, ça fait s. Ce et a, ça fait ka. Ah bon mais se et e, ça fait ce ou ze ? Ça dépend, un secret ou une ruse, c'est selon. Les lettres dansent et tout ça ne veut rien dire. Vous parliez et bien, lisez maintenant ! ».

(Tania de Montaigne (2004). *Geneviève et la théorie du cinq*. Paris : Flammarion)

SOMMAIRE

AVANT-PROPOS	8
CHAPITRE 1 – LA LECTURE : MODELISATION ET TRAITEMENTS MIS EN OEUVRE	10
1- Les modèles développementaux	10
2- Les modèles connexionnistes et interactifs	13
2.1- Le modèle de Seidenberg et McClelland (1989)	13
2.2- Le modèle à double fondation de Seymour (1997)	15
2.3- Le modèle de lecture par analogies	17
2.4- Le modèle d'activation des niveaux orthographique et phonologique de Colé, Magnan et Grainger (1999)	19
3- Rôle de l'apprentissage implicite en lecture	21
3.1- L'apprentissage implicite	21
3.2- Implicite <i>versus</i> explicite	23
CHAPITRE 2 – LECTURE ET GRAPHOPHONOLOGIE	28
1- Les connaissances graphophonologiques	28
1.1- Qu'est-ce que la graphophonologie ?	28
1.2- Capacités métaphonologiques et lecture	30
2- Conscience phonologique et lecture : une relation établie et largement démontrée	33
3- Implicite versus explicite en graphophonologie : l'effet de consistance graphème-phonème	35
3.1- Notion de régularité et de consistance	35
3.2- Effet de consistance et de régularité sur la reconnaissance de mots écrits et sur la lecture	39
CHAPITRE 3 – LECTURE ET ORTHOGRAPHE	47
1- Acquisition de la lecture et de l'orthographe	48
1.1- Modèles en stades de l'acquisition de l'orthographe	48
1.2- Critiques de ces modèles	52
2- Rôle de la fréquence et des analogies en orthographe	54
3- Implicite versus explicite en orthographe	59
3.1- Notion de régularités orthographiques	59
3.2- Effet de fréquence de certains patrons orthographiques : l'exemple des bigrammes	63
CHAPITRE 4 – LECTURE ET MORPHOLOGIE	67
1- Les connaissances morphologiques	68
1.1- Définition de la morphologie	68
1.2- Les différents types de morphèmes	68
2- Le rôle de la morphologie en lecture	70
2.1- Lecture experte : les modèles de décomposition morphologique	70
2.2- Effets de la morphologie : la pseudo-affixation	76
3- Implicite versus explicite en morphologie	77
3.1- La morphologie dans l'apprentissage de la lecture	77
3.2- Les traitements morphologiques chez l'apprenti lecteur	80
CHAPITRE 5 – PROBLEMATIQUE	91
1- Prédications sur les aspects graphophonologiques : la consistance graphophonologique	92
2- Prédications sur les aspects orthographiques : la fréquence des bigrammes	92
3- Prédications sur les aspects morphologiques : l'effet de pseudo-affixation	94
4- Prédications sur les relations entre les connaissances graphophonologiques, orthographiques, morphologiques et la lecture	95
5- Présentation des études expérimentales	98

CHAPITRE 6 – GRAPHOPHONOLOGIE, ORTHOGRAPHE ET MORPHOLOGIE : ASPECTS DEVELOPPEMENTAUX	101
1- Expérience 1 : graphophonologie	101
1.1- Méthode	101
1.1.1- Population	101
1.1.2- Matériel	104
<i>Epreuves graphophonologiques implicite, explicite et lecture à haute voix</i>	104
Tâche graphophonologique implicite de plausibilité lexicale écrite	105
Tâche graphophonologique explicite de jugement portant sur des mots irréguliers et réguliers fréquents de la langue (à partir du niveau CP)	106
Tâche de lecture à haute voix (uniquement à partir du niveau CP)	106
1.1.3- Procédure	107
<i>Epreuves graphophonologiques implicite, explicite et lecture à haute voix</i>	107
Tâche graphophonologique implicite de plausibilité lexicale écrite	107
Tâche graphophonologique explicite de jugement portant sur des mots irréguliers et réguliers fréquents de la langue (à partir du niveau CP)	108
Tâche de lecture à haute voix (à partir du niveau CP)	109
1.2- Hypothèses opérationnelles	109
1.3- Résultats	110
1.3.1- Tâches graphophonologiques implicite et explicite	110
<i>Tâche graphophonologique implicite</i>	111
<i>Tâche graphophonologique explicite</i>	111
1.3.2- Tâche de lecture à haute voix	112
1.4- Discussion	116
1.4.1- Développement des connaissances graphophonologiques implicites et explicites	116
1.4.2- Traitements graphophonologiques et niveau de lecture	118
2- Expérience 2 : orthographe	120
2.1- Méthode	120
2.1.1- Population	120
2.1.2- Matériel	120
<i>Epreuves orthographiques implicite, explicite et lecture à haute voix</i>	120
Tâche orthographique implicite de plausibilité lexicale écrite	121
Tâche orthographique explicite de jugement de patrons orthographiques (bigrammes isolés) (à partir du niveau CP)	121
Tâche de lecture à haute voix (uniquement à partir du niveau CP)	122
2.1.3- Procédure	122
<i>Epreuves orthographiques implicite, explicite et lecture à haute voix</i>	122
Tâche orthographique implicite de plausibilité lexicale écrite	122
Tâche orthographique explicite de jugement de patrons orthographiques	123
Tâche de lecture à haute voix	124
2.2- Hypothèses opérationnelles	124
2.3- Résultats	124
2.3.1- Tâches orthographiques implicite et explicite	124
<i>Tâche orthographique implicite</i>	125
<i>Tâche orthographique explicite</i>	127
2.3.2- Tâche de lecture à haute voix	129
2.3.2.1- Position initiale	129
2.3.2.2- Position finale	133
2.4- Discussion	136
2.4.1- Développement des connaissances orthographiques implicites et explicites	137
2.4.2- Traitements orthographiques et niveau de lecture	140
3- Expérience 3 : morphologie	143
3.1- Méthode	143
3.1.1- Population	143
3.1.2- Matériel	143
<i>Epreuves morphologiques implicite, explicite et lecture à haute voix</i>	143

Tâche morphologique implicite de plausibilité lexicale écrite	144
Tâche morphologique explicite de jugement d'intrus (uniquement à partir du niveau CP)	144
Tâche de lecture à haute voix (uniquement à partir du niveau CP)	145
3.1.3- Procédure	146
<i>Epreuves morphologiques implicite, explicite et lecture à haute voix</i>	146
Tâche morphologique implicite de plausibilité lexicale écrite	146
Tâche morphologique explicite de jugement d'intrus	146
Tâche de lecture à haute voix	147
3.2- Hypothèses opérationnelles	147
3.3- Résultats	147
3.3.1- Tâches morphologiques implicite et explicite	147
<i>Tâche morphologique implicite</i>	148
<i>Tâche morphologique explicite</i>	150
3.3.2- Tâche morphologique de lecture à haute voix	151
3.3.2.1- Préfixés	151
3.3.2.2- Suffixés	154
3.4- Discussion	160
3.4.1- Développement des connaissances morphologiques implicites et explicites	160
3.4.2- Traitements morphologiques et niveau de lecture	162

CHAPITRE 7 – EVOLUTION DES CONNAISSANCES GRAPHOPHONOLOGIQUES, ORTHOGRAPHIQUES ET MORPHOLOGIQUES 165

1- Expérience 4 : adultes et enfants de niveau CM2	165
1.1- Méthode	165
1.1.1- Population	165
1.1.2- Matériel et procédure	165
1.2- Résultats du groupe adulte	167
1.2.1- Tâches graphophonologiques	167
<i>Tâches implicite et explicite</i>	167
<i>Tâche graphophonologique de lecture à haute voix</i>	167
1.2.2- Tâches orthographiques	168
<i>Tâches implicite et explicite</i>	168
<i>Tâche orthographique de lecture à haute voix</i>	169
Position initiale	169
Position finale	170
1.2.3- Tâches morphologiques	170
<i>Tâches implicites et explicites</i>	170
<i>Tâche morphologique de lecture à haute voix</i>	171
Préfixés	172
Suffixés	172
1.3- Discussion des résultats du groupe adulte	173
1.3.1- Graphophonologie	174
1.3.1.1- Développement des connaissances graphophonologiques implicites et explicites	174
1.3.1.2- Traitements graphophonologiques et expertise en lecture	174
1.3.2- Orthographe	175
1.3.2.1- Développement des connaissances orthographiques implicites et explicites	175
1.3.2.2- Traitements orthographiques et expertise de lecture	176
1.3.3- Morphologie	177
1.3.3.1- Développement des connaissances morphologiques implicites et explicites	177
1.3.3.2- Traitements morphologiques et expertise en lecture	178
1.4- Analyses comparatives adultes-enfants de niveau CM2	178
1.4.1- Graphophonologie	179
1.4.2- Orthographe	179
Position initiale	180
Position finale	181
1.4.3- Morphologie	182

Préfixés _____	182
Suffixés _____	183
1.5- Discussion des résultats issus des comparaisons adultes-enfants de niveau CM2 _____	184
1.5.1- Evolution des traitements graphophonologiques entre le niveau CM2 et l'expertise en lecture _____	184
1.5.2- Evolution des traitements orthographiques entre le niveau CM2 et l'expertise en lecture _____	185
1.5.3- Evolution des traitements morphologiques entre le niveau CM2 et les lecteurs experts _____	185
CHAPITRE 8 – RELATIONS ENTRE LES CONNAISSANCES GRAPHOPHONOLOGIQUES, ORTHOGRAPHIQUES, MORPHOLOGIQUES, LES NIVEAUX DE LECTURE ET DE VOCABULAIRE _____	187
1- Corrélations entre les notes globales pour les différents groupes _____	187
1.1- Corrélations en GSM _____	187
1.2- Corrélations au niveau CP _____	187
1.3- Corrélations au niveau CE1 _____	188
1.4- Corrélations au niveau CE2 _____	189
1.5- Corrélations au niveau CM1 _____	189
1.6- Corrélations au niveau CM2 _____	189
1.7- Corrélations chez le groupe adulte _____	190
2- Régressions multiples _____	190
3- Discussion _____	191
CHAPITRE 9 – DISCUSSION GENERALE _____	195
1- Connaissances graphophonologiques, orthographiques, morphologiques et apprentissage de la lecture _____	197
1.1- Les connaissances graphophonologiques, orthographiques et morphologiques _____	197
1.2- Liens avec la lecture et le vocabulaire _____	204
1.3- Liens entre les connaissances graphophonologiques, orthographiques et morphologiques _____	206
2- Influence des traitements graphophonologiques, orthographiques et morphologiques sur la reconnaissance de mots écrits _____	209
2.1- Traitements graphophonologiques, orthographiques et morphologiques en lecture _____	209
2.2- Evolution des traitements graphophonologiques, orthographiques et morphologiques _____	213
CONCLUSION _____	216
REFERENCES BIBLIOGRAPHIQUES _____	221

AVANT-PROPOS

Dans la littérature portant sur l'apprentissage des procédures de reconnaissance des mots écrits, il est admis que la lecture passe par la réalisation d'un ensemble de traitements conduisant à la construction des représentations nécessaires à la compréhension, finalité de l'activité de lecture. Ces traitements sont multiples, de niveaux et de natures différents : graphémiques, phonologiques, sémantiques, syntaxiques et textuels. Cependant, malgré cette multiplicité des traitements, la majorité des recherches s'est centrée surtout, voire de manière exclusive, sur les relations entre conscience phonologique et acquisition de la lecture. Ces travaux démontrent généralement que les habiletés phonologiques et les capacités d'identification de mots entretiennent des relations fortes et spécifiques (par exemple, Wagner & Torgensen, 1987). La conscience phonologique y apparaît comme bi-directionnelle, à la fois cause et conséquence de l'apprentissage de la langue écrite et constituerait un bon prédicteur de la réussite en lecture (*cf.*, entre autres, Mann, 1986 ; Wimmer, Landerl, Linorter, & Hummer, 1991). Ces recherches s'appuient généralement sur les modèles d'apprentissage à étapes qui, malgré les limites et les critiques soulevées, restent encore dominants. Les principales critiques sont évoquées en partie dans le chapitre 1 de cette thèse.

D'autres conceptions de l'apprentissage se sont développées, en particulier connexionnistes, qui reconnaissent l'interaction nécessaire des différents types de traitements. Dans ces modèles connexionnistes, une place importante est notamment accordée aux relations entre l'orthographe et la phonologie. Cependant, cette interaction entre représentations orthographiques et représentations phonologiques est évoquée essentiellement chez l'adulte et rarement chez le jeune apprenant. De plus, ces modèles délaissent largement les connaissances morphologiques dans l'analyse de la lecture et de son apprentissage. Pourtant, des données expérimentales récentes sur le rôle des connaissances morphologiques montrent que la conscience morphologique, même si elle n'occupe pas la même position que la conscience phonologique dans l'émergence des habiletés de lecture, intervient précocement dans les procédures de reconnaissance de mots écrits (*cf.* Carlisle, 1995 ; Marec-Breton, 2003).

Outre la prise en compte généralement trop tardive des connaissances morphologiques et orthographiques, la plupart des modélisations de la lecture se focalise sur le versant explicite de l'apprentissage : l'installation des compétences, notamment phonologiques, s'y réalise généralement sous l'effet de l'enseignement explicite. Or, si l'enseignement scolaire donne une large place à l'acquisition de règles explicites telles que les règles de grammaire, les

apprentissages naturels ne permettent généralement pas à l'apprenant de prendre conscience des règles qui structurent la part de l'environnement avec laquelle il interagit, règles qui semblent pourtant guider son comportement. Face à ce constat, Pacton et Perruchet (2004) considèrent notamment que deux types d'apprentissage coexisteraient. D'un côté, l'expérience scolaire permet à l'apprentissage d'opérer intentionnellement, par le biais d'un effort attentionnel orienté. De l'autre côté, un grand nombre d'aptitudes seraient acquises de façon apparemment « naturelle » ou spontanée, sans qu'un quelconque superviseur ait planifié une méthode d'instruction et sans effort ni intention particulière de la part de l'apprenant. Il est ainsi possible de citer la capacité à catégoriser les objets, à s'adapter aux lois de la physique ou encore à se conformer à des règles sociales mais c'est l'acquisition de la langue maternelle qui fait souvent figure d'exemple privilégié. Les auteurs des travaux conduits dans le domaine du langage considèrent que l'enfant pourrait extraire de son environnement linguistique des aspects perceptivement saillants et développer alors une sensibilité implicite à certaines régularités langagières. Dans le cadre du langage écrit, il est possible de considérer que l'enfant puisse développer très tôt non seulement des connaissances graphophonologiques mais aussi des connaissances orthographiques et morphologiques dans la mesure où il est confronté précocement aux régularités du système orthographique et morphologique de sa langue.

Dans cette perspective, l'objectif principal de ce travail est de montrer qu'avant même l'apprentissage de la lecture et le développement des capacités phonologiques explicites, l'enfant possède des connaissances implicites en graphophonologie, en orthographe et en morphologie et que ces connaissances sont susceptibles de faciliter la lecture. Le chapitre 2 propose un certain nombre de recherches montrant que les jeunes enfants sont sensibles à la fréquence des associations graphèmes-phonèmes et que, dans une langue comme le français, l'apprentissage de la lecture est lié en partie à la régularité de telles associations. Le chapitre 3 présente plusieurs études développementales qui suggèrent une sensibilité à certaines régularités orthographiques n'ayant jamais fait l'objet d'un enseignement explicite. Le chapitre 4 fournit un ensemble de données démontrant l'existence précoce de connaissances morphologiques et l'influence de telles connaissances sur les procédures de reconnaissance de mots écrits. Des expériences ont été réalisées auprès d'enfants et d'adultes afin d'éprouver et valider ces données théoriques. Les principaux résultats issus de ces expériences ainsi que les analyses subséquentes figurent dans les chapitres méthodologiques 6, 7 et 8. Enfin, le chapitre 9 propose une interprétation et une discussion de ces résultats ainsi qu'une conclusion sur les perspectives de recherche qui en découlent.

CHAPITRE 1 – LA LECTURE : MODELISATION ET TRAITEMENTS MIS EN OEUVRE

La conception qui a été longtemps dominante dans les modèles et les études sur l'apprentissage de la lecture est celle relative à des « étapes » dans la mise en place du processus d'identification des mots. Sur le versant pédagogique, la plupart des méthodes de lecture s'appuient ainsi encore implicitement sur les modèles développementaux en stades, en particulier sur le modèle de Frith (1985), qui repose sur l'idée selon laquelle les processus d'identification des mots se mettraient en place et se développeraient en suivant un certain nombre d'étapes, ce qui entraîne, pour l'enseignement, un ordre immuable et une progression à respecter.

1- Les modèles développementaux

Un des principaux modèles développementaux qui a fait figure de référence pendant de nombreuses années dans la littérature est celui de Frith (1985). Ce modèle se fonde lui-même sur les modèles d'accès au lexique à double voie (*cf.* Coltheart, 1978). Selon cette conception modulaire de la lecture, celle-ci mobilise deux voies d'accès au lexique interne :

- une voie directe (adressage) : la reconnaissance de mots bien connus du lecteur s'effectue par appariement direct de la configuration écrite du mot avec sa représentation visuelle en mémoire ;
- une voie indirecte (assemblage) : le traitement des mots inconnus est réalisé par la transformation de l'information orthographique en information phonologique par l'application des règles de correspondances graphèmes-phonèmes.

Dans ce type de modélisation, l'apprentissage de la lecture dépend de la mise en place de ces deux voies de façon strictement séquentielle. Cette conception se retrouve dans le modèle développemental à étapes décrit par Frith (1985). Selon elle, l'enfant passe par trois étapes pour mettre en place, développer et automatiser les processus d'identification de mots.

Le premier stade est qualifié de logographique. A ce niveau, l'association entre les objets visuels (les mots écrits) et la signification associée est directe. Ce stade, préalable à l'apprentissage de la lecture, se caractérise par le développement chez l'enfant de stratégies pour « deviner » les mots perçus. L'enfant s'appuie alors sur tous les indices visuels saillants (indices extralinguistiques) pris dans le mot lui-même et pouvant permettre son identification. Le mot est reconnu dans sa globalité sans que les lettres qui le composent puissent être utilisées pour lire d'autres mots. En fait, l'enfant reconnaît les mots comme il reconnaît des objets non

linguistiques. Le stade logographique correspond au moment où l'enfant n'a pas encore compris le principe alphabétique (c'est-à-dire le principe de la correspondance graphèmes-phonèmes) : l'enfant utilise alors un système sémantique pictural dans lequel l'information linguistique est traitée comme une image à laquelle un mot oral est secondairement associé (Morton, 1980). En écriture, ce stade correspond essentiellement au moment où l'enfant devient capable d'écrire quelques mots familiers, comme son prénom, sans avoir compris que les lettres des mots représentent les sons des mots correspondants.

Le stade alphabétique est déterminé par l'apprentissage explicite des règles de conversion graphèmes-phonèmes. Ce deuxième stade se caractérise ainsi par l'utilisation de la médiation phonologique. A ce niveau de l'apprentissage, l'effort attentionnel est consacré à la mise en correspondance de l'écrit avec l'oral. La procédure alphabétique serait d'abord utilisée en écriture avant de l'être en lecture (pour une revue des résultats obtenus dans ce domaine, voir Ellis, 1994). A ce stade, l'apprenti scripteur prend conscience de l'existence d'unités sublexicales constituant les mots parlés (phonèmes) et les mots écrits (graphèmes) et de la relation entre ces unités orales et écrites (correspondances graphèmes-phonèmes exigeant la connaissance de l'alphabet, une conscience de l'unité phonémique et une mise en relation de ces deux types de connaissances). Pour orthographier tous les types de mots, l'enfant utiliserait également cette procédure alphabétique, ce qui entraîne des erreurs dans le cas des mots irréguliers (*cf. femme écrit fame*). La procédure alphabétique en écriture stimulerait l'adoption d'une stratégie alphabétique en lecture. Grâce à cette procédure d'assemblage graphèmes-phonèmes, l'enfant va pouvoir reconnaître de plus en plus de mots et se constituer ainsi un lexique mental. L'automatisation de la lecture alphabétique et l'inefficacité de la stratégie alphabétique pour lire les mots irréguliers entraîneraient ensuite l'adoption d'une procédure orthographique.

Dans le troisième et dernier stade, orthographique, l'existence d'un lexique mental permet l'appariement direct des mots familiers et de leurs correspondants dans ce lexique. Dans la phase logographique initiale, les mots étaient reconnus par des procédures de reconnaissance visuelle non spécifiques au traitement d'un matériel verbal. Au stade orthographique, c'est par une analyse linguistique que le système de traitement de l'information active directement l'élément lexical pertinent. Ce stade est donc caractérisé par une analyse linguistique permettant au système de traitement de l'information d'accéder immédiatement au mot qui est récupéré directement dans le lexique orthographique. L'écriture orthographique correspond à la mise en place d'unités de taille supérieure au phonème, les morphèmes. Pour Frith (1985), ce serait à ce stade que des mots nouveaux pourraient être écrits par analogie avec des mots connus.

Selon cette conception, l'acquisition de la lecture passe par un traitement non linguistique qui a un impact direct sur la reconnaissance du mot écrit *via* la compréhension, puis l'application du principe alphabétique, enfin par son automatiser. Dans un tel modèle, un facteur joue un rôle central : l'apprentissage explicite des correspondances graphèmes-phonèmes car c'est sous l'effet de cet enseignement que peut se mettre en place le lexique mental puis un accès direct à ce lexique (Gombert, 2003). La phonologie semble donc être le facteur développemental crucial et un prérequis à l'apprentissage de la lecture (Lennox & Siegel, 1994) : elle serait notamment à la base de l'acquisition orthographique des mots. Selon Fayol et Gombert (1999), les modèles en cours confortent, à quelques variations près, la conception développée par Frith (1985). Le moteur de l'évolution de l'apprentissage serait essentiellement l'intervention pédagogique : l'enfant passe d'un stade au suivant sous l'effet d'une instruction qui vise à lui faire comprendre le fonctionnement des codes écrits. Ce modèle a cependant fait l'objet de diverses critiques, dont l'une concerne la remise en cause de la successivité des étapes. Ainsi, Sprenger-Charolles et Casalis (1996) remarquent que tous les enfants ne passent pas nécessairement par le stade logographique. En dépit de cette limite, ce modèle présente néanmoins l'intérêt de mettre en avant l'importance de la maîtrise des correspondances entre les graphèmes et les phonèmes dans les processus d'apprentissage (Fayol & Gombert, 1999). Toutefois, une autre des critiques apportées à ce type d'approche générale de l'apprentissage lecture est justement la réduction de cet apprentissage aux capacités de décodage et d'assemblage phonologique, suivies de l'automatisation de ces processus et de l'articulation avec le traitement visuo-orthographique. Le développement de la littéracie suit alors toujours le même cheminement : l'enfant prend d'abord en compte l'information phonologique, puis orthographique et enfin morphologique. Ces deux derniers types de connaissances n'apparaîtraient donc que tardivement chez l'enfant. En outre, ce type de modèle est un cadre descriptif général qui ne renseigne pas sur la dynamique de l'apprentissage : l'apprentissage de la lecture est y envisagé comme une suite d'étapes ou de stades dont l'ordre est strict et identique pour tous les enfants. Il ne prend pas en compte les différences interindividuelles et est plutôt descriptif (Ecalte & Magnan, 2002a). De plus, selon Gombert (2002), la composante morphologique ne joue qu'un rôle secondaire dans ce modèle. Une telle modélisation la lecture décrit essentiellement le versant explicite de l'apprentissage mais délaisse la part d'apprentissage implicite, c'est-à-dire les changements s'opérant à l'insu de l'individu sous le simple effet de la rencontre répétée d'un système écrit présentant de nombreuses régularités (orthographiques, morphologiques et graphophonologiques). Toujours selon Gombert (2000), une perspective connexionniste comme celle de Seidenberg et

McClelland (1989) est plus adaptée à la description de cette dimension implicite de l'apprentissage.

De façon générale, l'orientation actuelle suppose plutôt que les différentes étapes décrites dans les modèles développementaux correspondent à des procédures disponibles pour le lecteur de façon contemporaine. Les modèles de type connexionniste (e.g., Seidenberg & McClelland, 1989) et de type interactif (voir Seymour, 1997) remplacent ainsi progressivement les conceptions développementales en décrivant comment le système de traitement de l'information mobilise simultanément différents processeurs lors de la reconnaissance de mots écrits.

2- Les modèles connexionnistes et interactifs

Les modèles interactifs n'assimilent pas la lecture à la recherche d'un mot stocké dans le lexique mental comme le fait la plupart des modèles développementaux. En effet, pour les modèles interactifs comme celui de Seidenberg et McClelland (1989), lire découle de l'activation de différentes unités spécialisées dans le traitement orthographique, phonologique et sémantique. Toutes ces connaissances sont activées simultanément et automatiquement. Ce sont les configurations d'unités activées qui déterminent les mots reconnus.

2.1- Le modèle de Seidenberg et McClelland (1989)

L'objectif de Seidenberg et McClelland est d'utiliser une architecture minimale dans laquelle l'apprentissage joue un rôle majeur. Le modèle proposé comprend un processeur orthographique, un processeur phonologique, un processeur sémantique et un processeur contextuel qui interagissent, et des processus de transformation qui permettent la transmission d'activation et d'inhibition d'un processeur à l'autre. Concrètement, le système connexionniste décrit par Seidenberg et McClelland correspond à un ensemble de réseaux d'unités communiquant entre elles par des connexions. Chacune de ces unités peut être active ou inactive en fonction de l'influence activatrice ou inhibitrice communiquée par les autres unités *via* les connexions. Ce système est, par définition, un système qui apprend. De façon continue, le poids des liaisons entre les unités constitutives des réseaux du système varie. Chaque co-activation de deux unités va renforcer le lien entre ces unités de telle façon que l'information activatrice circulera de façon de plus en plus rapide et efficace entre ces unités. La Figure 1 ci-dessous résume ce modèle.

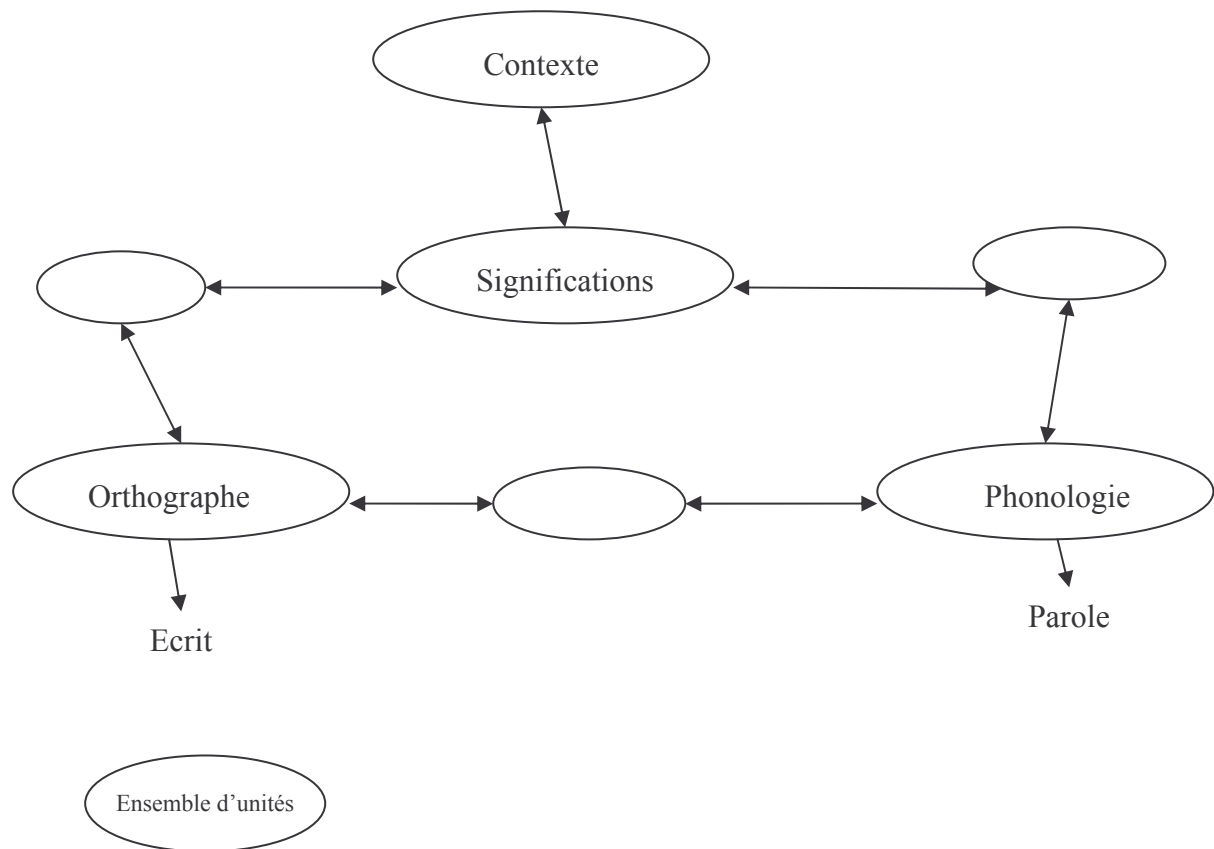


Figure 1. Modèle de Seidenberg et McClelland (1989)

Dans ce type de modèle, le lexique mental n'existe pas, à proprement parler, dans la mesure où il n'y a pas de mots stockés comme des unités en mémoire à long terme. La reconnaissance des mots est conçue comme le produit de la distribution d'une activation dans le système de traitement des lettres et des informations phonologiques et sémantiques associées à ces lettres. Ici, reconnaître un mot n'est pas retrouver ce mot quelque part en mémoire mais recouvrer un certain état d'activation de son système de traitement de l'information lexicale. Ainsi, pour les mots les plus familiers, l'ensemble des lettres activées l'aurait été si souvent que les liens entre ces lettres, et entre ces configurations orthographiques et les informations qui leur sont associées, deviendraient très forts, ce qui permettrait que les unes et les autres s'activent automatiquement autorisant la reconnaissance de ces mots comme des globalités. Cela serait de moins en moins vrai au fur et à mesure que les mots sont peu familiers, tandis que les mots familiers bénéficient de la rapidité d'activation consécutive aux lectures antérieures. Chaque configuration différente d'activation correspond alors à la reconnaissance d'un mot différent.

Le schéma développemental est le suivant : si des configurations visuelles (des lettres) perçues par l'individu sont de façon récurrente associées à des significations (dimensions lexicales et morphologiques) et à des prononciations (dimensions phonologiques), la répétition de la co-activation des unités activées par ces percepts (initialement des unités picturales) et des unités sémantiques et phonologiques aboutira, *via* le renforcement des liaisons, à une spécialisation orthographique de ces unités. Ainsi, la reconnaissance d'un mot est conçue comme le produit d'une distribution d'une activation dans le système de traitement des lettres et des informations phonologiques, orthographiques et sémantiques associées à ces lettres. Cette reconnaissance de mot s'effectue donc sur la base d'une interaction entre les différents types d'informations issues du mot écrit. Cette idée d'interaction est également développée dans le modèle à fondation duale de Seymour (1997) qui cherche aussi à pallier les critiques émises à l'égard des modèles développementaux.

2.2- Le modèle à double fondation de Seymour (1997)

Dans ce modèle, Seymour suppose que l'acquisition des connaissances orthographiques met en jeu la formation d'une structure qui encode des propriétés abstraites de la langue écrite et dont le développement dépendrait de deux fondations. Celles-ci doivent avoir atteint un certain niveau avant que le développement orthographique puisse être initié. Le modèle de Seymour comporte au total cinq composants distincts, chacun pouvant être conçu comme un processeur ou module : les processeurs logographique et alphabétique, le processeur de conscience linguistique et deux structures centrales orthographique et morphographique.

Le processeur logographique renvoie à un traitement holistique permettant la reconnaissance directe et le stockage des mots. La procédure logographique ne nécessite pas une organisation phonologique allant au-delà du niveau de segmentation du langage parlé en unités lexicales. Ainsi, aucune relation interactive avec la conscience linguistique n'existe ici.

Un traitement analytique s'opère ensuite avec la mise en place du processeur alphabétique. Une procédure séquentielle de conversion des graphèmes en phonèmes et des phonèmes en graphèmes est alors appliquée. La procédure alphabétique entretient donc une relation interactive avec le composant phonologique de la conscience linguistique : son acquisition nécessite, en effet, l'isolement des structures phonémiques de la parole. Une conséquence du développement du traitement alphabétique serait l'émergence de la conscience phonémique.

Dans l'apprentissage de l'orthographe, la procédure logographique est nécessaire dans la mesure où c'est par elle que les mots, à partir desquels l'information orthographique doit être

extraite, sont intériorisés. La procédure alphabétique contribue, quant à elle, à la formation de la structure orthographique.

La structure orthographique se développe progressivement : elle prend d'abord en compte des structures plutôt simples, telles que les graphèmes et les phonèmes, puis s'étend graduellement à des structures multi-lettres. Pendant toute la période au cours de laquelle cette structure s'élabore, le système logographique renferme des représentations de nombreux mots qui peuvent devenir des représentations orthographiques. Ces représentations orthographiques vont alors se construire petit à petit par une procédure interne, phonologiquement motivée, de redescriptions des données issues du système logographique. Le développement est ici conçu de façon interactive car la segmentation orthographique influence la segmentation phonologique, et *vice-versa*.

Enfin, la structure morphologique permet le traitement de mots plurisyllabiques. La construction de cette structure est directement dépendante de l'achèvement de la structure orthographique et est en étroite relation avec la conscience linguistique. L'influence morphologique établirait le statut des segments en tant que racines des mots et affixes.

L'intérêt du modèle de Seymour est de rendre compte d'une interactivité entre les processus orthographiques plutôt que d'une sérialité comme dans le modèle de Frith (1985). Il montre également que le développement de la conscience phonologique sous l'effet de l'enseignement s'opère des unités réduites vers des unités plus larges qui sont respectivement les phonèmes et les syllabes. L'unité morphologique est présente dans le modèle de Seymour au travers du processeur morphologique, contrairement au modèle de Frith (1985). Seymour met ainsi en évidence l'importance d'autres types de connaissances que celles phonologiques, notamment les connaissances orthographiques et morphologiques dans le développement orthographique. Cependant, la composante morphologique est retrouvée à un niveau de développement relativement avancé dans l'apprentissage de la lecture dans la mesure où elle s'établit à la suite du niveau orthographique qui lui-même apparaît tardivement et se développe sous la dépendance de la procédure alphabétique. Or, plusieurs théories mettent en avant l'idée selon laquelle les jeunes enfants possèderaient de façon précoce des connaissances sur les régularités orthographiques du langage écrit. Ces théories insistent également sur le rôle de la morphologie dans l'orthographe. Les enfants commenceraient ainsi à utiliser ces connaissances bien plus tôt que ne l'affirment les modèles développementaux décrits précédemment. Ces théories défendent également l'idée d'une interaction entre les différentes sources de connaissances. Quelques modèles mettent ainsi en avant l'utilisation des connaissances antérieures grâce à des processeurs analogiques (Gombert, Bryant, & Warrick, 1997) ou encore

via l'importance de l'exposition à l'écrit (Colé, Magnan, & Grainger, 1999). Ces modèles appelés modèles de lecture par analogie insistent notamment sur l'ancrage de nouvelles connaissances sur des connaissances antérieures.

2.3- Le modèle de lecture par analogies

Actuellement, la plupart des chercheurs accordent une place importante à l'utilisation des connaissances antérieures (phonologiques et visuo-orthographiques) qui résulterait de la mise en œuvre de processus analogiques. Les modèles de lecture proposés insistent alors fortement sur l'ancrage de nouvelles connaissances sur des connaissances antérieures, plus spécifiquement sur les liens entre les connaissances phonologiques et orthographiques. L'analogie est une procédure de transfert de connaissances qui rend possible la lecture d'un mot nouveau grâce à un mot connu. Par exemple, l'apprenti lecteur serait capable de lire le mot nouveau *main* grâce à sa capacité d'extraire (par analogie) la rime *-ain* qu'il sait déjà identifier dans un mot connu tel que *pain*.

Selon Goswami (1986, 1988a, 1990), la procédure de lecture de mots nouveaux sur la base d'une analogie avec des mots dont l'enfant connaît la forme orthographique se développerait de façon précoce. La lecture par analogie relève d'un transfert de connaissances basé sur la découverte de ressemblances entre les unités sonores que les enfants maîtrisent (unités plus grandes que le phonème) et les séquences de lettres connues dans leurs formes phonologiques et orthographiques. La modélisation initiale concernant la lecture par recours aux analogies a été réalisée par Goswami et Bryant (1990) puis enrichie par Gombert *et al.* (1997) et Goswami (1999). Selon cette modélisation, les enfants qui ont été entraînés à utiliser les analogies entre séquences phonologiques et séquences orthographiques apprennent plus facilement à lire (Goswami, 1998, 1999). Cet apprentissage précoce des analogies, qui s'effectue généralement sur les unités larges comme la rime, faciliterait l'apprentissage du code alphabétique (Goswami, 1995 ; Goswami & East, 2000).

Le modèle de lecture par analogies était au départ composé de quatre processeurs : orthographique, phonologique, sémantique et contextuel. Avant l'apprentissage de la lecture, le processeur orthographique serait inexistant et remplacé par un processeur pictural permettant une lecture globale ou logographique. Ainsi, l'information lexicale serait traitée conjointement par les processeurs pictural, phonologique et sémantique. Ce traitement conjoint aboutirait ensuite à la création d'un processeur lexical qui ne serait qu'une évolution du processeur pictural. Le processeur visuel (pictural, puis lexical) et le processeur phonologique interviendraient conjointement. Le processeur visuel établit une base de connaissances visuelles et est capable de

traiter globalement des mots et des configurations orthographiques fréquentes. Le processeur phonologique, quant à lui, participe à la construction d'une base de données phonologiques. Pour Gombert *et al.* (1997), cette correspondance entre analogies orthographiques et analogies phonologiques serait utilisée par le système cognitif dans un mécanisme d'auto-apprentissage pour permettre la lecture de certains mots nouveaux, voisins phonologiques et orthographiques de mots connus. Ainsi, la mise en œuvre de procédures analogiques implicites participerait au développement et à la maîtrise du code alphabétique appris de façon explicite *via* l'enseignement formel des règles de correspondance graphèmes-phonèmes. Dans un premier temps, le processeur pictural, composant du processeur visuel, permet l'identification globale des mots. Ce processeur pictural évolue ensuite en se spécialisant dans le repérage des régularités orthographiques de la langue. Cette spécialisation permet au processeur pictural de devenir un processeur lexical. A ce stade, le processeur lexical poursuit son développement par la maîtrise progressive des unités orthographiques en relation avec les unités phonémiques. Ce traitement très élaboré est réalisé par le processeur orthographique. Ce modèle met en évidence les fortes connexions établies entre les différents niveaux d'informations traitées par les processeurs visuels et phonologiques. Il tend également à confirmer la précocité de la mise en œuvre des procédures analogiques, avant même que l'enfant n'apprenne à lire.

De nombreuses recherches (*cf.* synthèses de Gineste, 1997 ou de Bastien, 1997) ont montré la précocité et l'efficacité des traitements analogiques dans les processus cognitifs. Ces traitements constitueraient un mode privilégié d'acquisition de connaissances nouvelles. Bastien (1997) considère ainsi que l'analogie est un processus général d'acquisition qui repose sur l'activation des connaissances antérieures avec lesquelles peuvent s'établir de nouveaux liens fonctionnels. Ainsi, Brun, Faure, Ecalle et Magnan (2000) ont observé un effet de l'analogie chez des enfants de CP de langue française possédant un niveau élevé de conscience phonologique lorsque les mots cibles et indices partagent une rime commune. Leur expérience consistait à proposer à ces enfants une tâche de reconnaissance de mots écrits : deux mots associés chacun à une image étaient présentés successivement. L'enfant devait indiquer si le second mot, après lecture du premier par l'expérimentateur, « allait bien » avec son image ou non. Les auteurs ont noté une réussite significativement supérieure pour les items partageant la rime dans la condition « analogie » alors qu'aucun effet significatif n'était obtenu dans la condition « sans analogie ».

Contrairement aux modèles en stades comme celui de Frith (1985), ce type de modèle par analogies est véritablement interactif, dans la mesure où il met l'accent sur les liens précoces entre unités phonologiques et unités orthographiques : les connaissances phonologiques

déterminent en partie les connaissances orthographiques qui en retour contribuent au développement et à la précision des connaissances phonologiques. En outre, cette modélisation met en avant les aspects implicite et explicite de l'apprentissage. En effet, ce modèle intègre à la fois les connaissances implicites décrites par la première connexion (i.e., connexion entre la connaissance phonologique et la connaissance des séquences orthographiques correspondantes : association « inconsciente » entre séquences de sons et séquences de lettres) et les connaissances explicites décrites par les suivantes. Les procédures analogiques se développeraient implicitement chez les lecteurs débutants parallèlement à l'instruction formelle (explicite) du code alphabétique. Ces procédures construisent alors un stock de séquences orthographiques de plus en plus complexes et spécifiques des mots. La mise en œuvre précoce des procédures analogiques, souvent implicites, participerait ainsi au développement et à la maîtrise du code alphabétique de nature explicite.

Cette idée d'interaction précoce entre les connaissances phonologiques et orthographiques ainsi que les notions d'implicite et d'explicite ont également été développées dans le modèle d'activation des niveaux orthographique et phonologique de Colé *et al.* (1999).

2.4- Le modèle d'activation des niveaux orthographique et phonologique de Colé, Magnan et Grainger (1999)

Ce modèle est une adaptation du modèle interactif d'activation du code phonologique en reconnaissance des mots écrits chez le lecteur expert de Jacobs, Rey, Ziegler et Grainger (1998). Dans ce modèle, l'identification d'un mot s'effectue sur la base de l'activation de l'information orthographique et phonologique qui serait irrépressible, générale, automatique, précoce et non optionnelle. Ce modèle postule une architecture spécifique pour rendre compte des relations entre les représentations orthographiques et phonologiques avec les unités orthographiques sublexicales (lettres ou groupes de lettres) et les unités phonologiques sublexicales (phonèmes, syllabes). Ces trois niveaux, celui des unités orthographiques infralexicales, celui des unités phonologiques infralexicales et celui des unités lexicales, sont interconnectés. Le traitement visuel du mot écrit active un ensemble d'unités orthographiques infralexicales qui active à la fois les unités lexicales et les unités phonologiques infralexicales correspondantes qui renvoient à leur tour l'activation vers les unités lexicales. En résumé, les représentations phonologiques lexicales sont activées simultanément par les représentations orthographiques lexicales et par les représentations phonologiques lexicales.

La principale modification apportée par Colé *et al.* (1999) au modèle de Jacobs *et al.* (1998) concerne l'existence d'un niveau intermédiaire entre les représentations des lettres et les

représentations phonologiques infralexicales, lequel implique les connexions graphèmes-phonèmes. L'idée de base est que la reconnaissance du mot parlé suppose le traitement des unités larges (syllabes et infrasyllabes telles que l'attaque et la rime) alors que la reconnaissance visuelle du mot écrit implique le traitement d'unités réduites (les lettres). L'hypothèse formulée est que lorsque ces associations sont explicitement apprises aux enfants, une interface entre les représentations des lettres et les représentations phonologiques sublexicales intervient au niveau des graphèmes et des phonèmes.

Avant tout apprentissage institutionnel (explicite) de la lecture, le jeune enfant constitue d'une part, des connaissances grâce à la pratique du langage parlé (versant implicite) et d'autre part, sous l'effet de l'exposition à l'écrit dans l'environnement quotidien (famille et école) et associe des représentations des lettres à des représentations phonologiques préexistantes construites pour accéder à la signification. En même temps, l'enfant apprend à associer des séquences de lettres à une signification *via* la construction de la représentation orthographique du mot (unités qui codent la présence d'une combinaison spécifique de lettres). La spécification des représentations orthographiques se réalise avec la multiplication des rencontres et des expériences avec l'écrit. La construction de telles représentations orthographiques permet d'expliquer l'effet de fréquence observé parfois chez les lecteurs. L'apprentissage explicite des correspondances graphèmes-phonèmes au cours de la première année d'apprentissage de la lecture conduit les enfants à développer des relations entre les représentations des lettres et les unités phonologiques larges telles que la syllabe ou les infrasyllabes.

Dans une telle conception, l'enfant parvient donc à développer des connaissances sur l'écrit avant même l'apprentissage explicite de la lecture. La multiplication et la fréquentation d'écrits favoriseraient l'émergence de connaissances dites implicites qui seraient acquises sans enseignement formel et qui renvoient également à la notion d'analogie développée par Gombert *et al.* (1997) ainsi qu'au versant implicite de l'apprentissage de la lecture.

L'interactivité entre les différents processus est ainsi omniprésente dans les modèles connexionnistes dans lesquels la reconnaissance d'un mot suppose que toutes les connaissances associées au mot - phonologique, orthographique, sémantique, morphologique - soient utilisées par le système, contrairement au modèle développemental de Frith (1985). Une des principales limites de ce dernier modèle est justement la prise en compte réductrice et/ou tardive des connaissances orthographiques et morphologiques. Or, il est vraisemblable que les enfants ne se réfèrent pas à une seule source d'informations : il n'y aurait pas une seule procédure utilisée à un moment donné, mais plutôt l'utilisation conjointe de différentes stratégies (Treiman, 1994). En

outre, les conceptions interactives accordent un rôle important à l'influence du langage parlé sur le langage écrit et visent à décrire comment un système devient progressivement capable de traiter le langage écrit sous l'effet de l'instruction (i.e., explicite) et de l'expérience avec l'écrit (i.e., implicite). Ainsi, les modèles connexionnistes décrivent les modifications observables induites par l'évolution continue de l'architecture du système qui lit. Ces modifications traduisent des apprentissages qui se font à l'insu du lecteur. Comme le souligne Gombert (2003), ces modèles ne traitent donc que du versant implicite de l'apprentissage de la lecture. Ces processus implicites ne sont cependant pas suffisants pour conduire à une lecture habile. En particulier, ils sont impuissants à installer chez le lecteur la connaissance consciente des règles à appliquer pour lire et orthographier correctement n'importe quel mot (ou pseudo-mot) dans n'importe quel contexte. C'est toutefois sur la base des apprentissages implicites que l'indispensable instruction explicite (l'enseignement) va permettre l'installation des traitements alphabétiques puis orthographiques. Ainsi, une bonne compréhension de l'ensemble des processus à l'œuvre dans l'apprentissage de la lecture exige de considérer de façon complémentaire les versants implicite et explicite de cet apprentissage.

3- Rôle de l'apprentissage implicite en lecture

3.1- L'apprentissage implicite

L'apprentissage implicite est un apprentissage généralement défini comme non intentionnel et inconscient. Les connaissances acquises *via* ce type d'apprentissage sont elles-mêmes difficilement accessibles à la conscience ou exprimables verbalement (Meulemans, 1998).

Dans les études portant sur l'apprentissage implicite, l'accent est essentiellement mis sur l'aspect incident ou inconscient de la phase d'apprentissage (Meulemans, 1998). Les définitions proposées diffèrent parfois selon les auteurs (*cf.* Seger, 1994) mais elles s'accordent cependant sur les caractéristiques suivantes : l'apprentissage se fait à l'insu du sujet et la connaissance acquise est difficilement accessible à la conscience ou exprimable verbalement (*cf.* Berry & Dienes, 1993 ; Seger, 1994). Perruchet, Vinter et Gallego (1997) ainsi que Perruchet et Nicolas (1998) définissent l'apprentissage implicite comme un mode d'adaptation dans lequel le comportement d'un individu apparaît sensible à la structure d'une situation, sans que cette adaptation ne soit imputable à l'exploitation intentionnelle de la connaissance explicite de cette structure. Il ne s'agit pas ici d'affirmer l'absence de toute connaissance explicite, mais de souligner que l'adaptation comportementale ne repose pas sur l'exploitation intentionnelle de cette connaissance. Cette définition renvoie au phénomène qui consiste à s'adapter à une situation complexe sans parvenir à comprendre les racines et les raisons de cette adaptation.

Selon Meulemans (1998), mieux comprendre le phénomène d'apprentissage implicite, c'est sans doute mieux comprendre le développement cognitif précoce. A cet égard, Reber (1993) défend l'idée de la primauté de l'apprentissage implicite. Il serait notamment à la base des premières étapes du développement cognitif. Dans ce cadre, l'apprentissage par le jeune enfant de sa langue maternelle est souvent cité comme exemple prototypique d'apprentissage implicite (*cf.* Lewicki, Hill, & Czyzewska, 1992). En effet, les enfants en âge préscolaire peuvent apprendre et utiliser des connaissances sémantiques complexes, produire des formes syntaxiques élaborées, tout en étant incapables de rapporter explicitement les règles du langage utilisées. L'enfant apprend ainsi à parler, c'est-à-dire à maîtriser l'outil complexe du langage, sans aborder cette tâche sur un mode analytique et surtout sans pouvoir verbaliser ce qu'il apprend. L'apprentissage implicite conduit alors, par l'action de processus inconscients de nature associative, à former des représentations en correspondance structurale avec le monde (Perruchet & Nicolas, 1998). Ceci signifie que les unités perceptives et représentationnelles tendent à reproduire, au fil des interactions avec un environnement donné, les unités structurellement pertinentes de cet environnement.

Ce processus d'apprentissage implicite de règles a été notamment étudié par Reber (1976) et Reber et ses collaborateurs (Reber & Allen, 1978 ; Reber & Lewis, 1977) en utilisant le paradigme d'apprentissage de grammaire artificielle. Dans les situations standards de grammaire artificielle, la tâche consiste pour les participants à mémoriser une série de séquences (chaînes) de consonnes générées à partir d'une grammaire à états finis qui détermine l'ordre possible des lettres. Les consignes ne mentionnent pas l'existence de cette grammaire et n'encouragent pas les participants à s'engager dans une analyse explicite et intentionnelle du matériel. Après cette première phase d'étude, les participants sont informés que les chaînes étudiées étaient engendrées par une grammaire. Il leur est alors demandé de classer de nouvelles séquences de lettres, selon qu'elles sont grammaticalement correctes ou non (c'est-à-dire selon qu'elles respectent ou non les règles de la grammaire). Les participants réussissent à distinguer les chaînes grammaticales de celles non grammaticales avec un taux de réussite supérieur au hasard, mais sans être capables de verbaliser les règles sur lesquelles leurs jugements semblent se fonder. Les participants ne parviennent donc pas à identifier les règles utilisées pour produire les séquences de lettres. Pour Reber (1976), cette dissociation entre performances et rapports verbaux atteste que l'apprentissage de grammaires artificielles est implicite. Son interprétation est que les participants abstraient inconsciemment la grammaire. Cette grammaire est stockée, et constitue une connaissance implicite, de nature abstraite. Cette acquisition serait donc incidente.

En situation naturelle, ce qui est vrai pour l'apprentissage de la langue maternelle peut l'être également pour d'autres formes d'acquisition de connaissances par l'enfant, par exemple la lecture. La psychologie cognitive du développement tend actuellement à être dominée par l'idée selon laquelle l'enfant acquiert des connaissances abstraites sur la structure et les régularités perçues de son environnement (Perruchet & Nicolas, 1998). Cependant, rares sont les études qui se sont intéressées à l'apprentissage implicite dans une perspective développementale. Les développementalistes, jusqu'à une époque récente, semblaient en effet manifester peu d'intérêt pour les phénomènes d'apprentissage implicite (Perruchet & Vinter, 1998). Le point suivant va mettre en évidence les modélisations récentes qui traitent de ces versants implicite et explicite et montrent que, de façon précoce, l'enfant développe des connaissances implicites sur sa langue qui pourraient jouer un rôle dans son apprentissage de la lecture.

3.2- Implicite *versus* explicite

Les connaissances implicites renvoient aux connaissances inconscientes que les locuteurs ont et qui gèrent, à leur insu, les traitements linguistiques qu'ils opèrent. Les connaissances explicites sont celles dont le locuteur a conscience et dont il peut piloter intentionnellement l'utilisation, par procéduralisation il peut toutefois y avoir dégageant de l'attention mais ces connaissances restent gérées attentionnellement. Les compétences explicites, contrairement aux compétences implicites, renvoient donc aux connaissances accessibles à la conscience (Meulemans, 1998).

Gombert (1990) a proposé un modèle du développement métalinguistique établissant des liens entre le développement du langage oral chez l'enfant et l'apprentissage de la lecture. Ce modèle propose une ébauche des versants implicite et explicite. En effet, pour établir les liens entre langage oral et langage écrit, Gombert (1990) distingue deux niveaux dans le contrôle linguistique qu'un individu peut exercer sur ses propres connaissances linguistiques : le niveau épilinguistique et le niveau métalinguistique. Au niveau épilinguistique, le contrôle sur les traitements linguistiques est exercé automatiquement, sans que l'individu en ait conscience, par la façon dont sont organisées les connaissances linguistiques en mémoire. Au niveau métalinguistique, le contrôle est exercé par l'individu à la suite d'une décision consciente et fait l'objet d'un pilotage intentionnel. Le niveau épilinguistique ou implicite de contrôle serait présent automatiquement dans tous les comportements impliquant des traitements linguistiques, tandis que le niveau métalinguistique ou explicite dépendrait de sollicitations externes ou d'une décision de l'individu. Dans un tel modèle, l'émergence du niveau explicite ou de la conscience métalinguistique (nécessaire à la lecture) dépend à la fois de l'existence préalable d'une connaissance implicite et de l'intervention d'une pression environnementale qui oblige l'individu

à faire l'effort cognitif nécessaire à la prise de conscience. L'enseignement de la lecture constituerait une telle pression.

Plus récemment, ces différents éléments ont été synthétisés par Gombert (2003) en un schéma global d'apprentissage de la lecture reposant sur les modèles de lecture par analogies. Selon lui, avant l'apprentissage formel de la lecture, l'enfant rencontre des mots écrits dans les livres ainsi que les significations associées à ces mots, notamment au cours de lectures effectuées avec ses parents. Ainsi, l'enfant peut accéder à la signification du mot à chaque fois qu'il le reconnaîtra. Ce processus est caractérisé par le fait qu'aucune analyse linguistique des mots écrits n'est effectuée. L'enfant réalise cette activité de manière implicite, inconsciente. Gombert (2003) suggère notamment que le lecteur débutant peut faire des analogies entre les mots qu'il sait reconnaître et ceux qu'il tente de découvrir. Par exemple, l'enfant qui connaît le mot *bien* utilisera l'analogie orthographique entre ce mot et le mot *rien* ainsi que sa capacité à reconnaître que ces mots partagent la même rime. Ces activités sont des processus implicites basés sur la prise en compte de régularités spécifiques de la langue lors des traitements analogiques. En effet, le système alphabétique de la langue française comporte de très nombreuses régularités et il n'est pas concevable qu'aucun apprentissage implicite ne se fasse à leur contact. Ces régularités concernent des configurations visuelles (régularités orthographiques), les mots oraux associés à ces configurations (régularités alphabético-phonologiques) et enfin les significations activées par ces configurations (régularités orthographico-morphologiques). La rencontre répétée de ces régularités, favorisée par l'exposition précoce à l'écrit, va permettre le développement de connaissances linguistiques et de capacités d'objets perçus visuellement qui seront présents avant la rencontre explicite de l'écrit. Ces connaissances et ces capacités évolueront ensuite au contact de l'écrit. Cette évolution constituera le versant implicite de l'apprentissage de la lecture. Cet apprentissage implicite peut être synthétisé de la façon suivante :

- (1) Dès la rencontre répétée de mots écrits stables (par exemple, les prénoms), les régularités concernant d'une part les configurations visuelles (régularités orthographiques), d'autre part les mots oraux et les significations qui leur sont associées (régularités alphabético-phonologiques et orthographico-morphologiques) engendrent des processus d'apprentissage implicite (voir Gombert & Peereman, 2001). Dans le cas de contacts précoces avec l'écrit, ces apprentissages préexistent donc à l'enseignement de la lecture ;
- (2) avec la multiplication de la fréquentation des mots écrits, consécutive à l'enseignement de la lecture, l'attention portée aux mots écrits, auparavant occasionnelle, devient systématique et fréquente. Les apprentissages implicites s'en trouvent ainsi décuplés ;

(3) l'apprentissage implicite se poursuit tant que l'individu lit et d'autant plus qu'il lit. Il ne s'arrête pas avec la fin de l'enseignement scolaire de la lecture.

Gombert (2003) considère que c'est sur la base des apprentissages implicites décrits précédemment que l'instruction explicite va pouvoir permettre l'installation de traitements alphabétiques puis orthographiques. Ces connaissances implicites sont donc nécessaires et fondamentales pour que les règles de l'apprentissage explicite puissent se mettre en place. L'apprentissage explicite résulte de l'effet direct de l'enseignement. Cet apprentissage concerne en particulier le décodage, c'est-à-dire l'apprentissage des règles de correspondances graphèmes-phonèmes qui permettent de lire tout item. Gombert (2003) suggère que la manipulation d'écrits a pour conséquence l'actualisation des connaissances métalinguistiques qui sont nécessaires au lecteur, c'est-à-dire des connaissances conscientes ou explicites. Ces connaissances explicites résultent de deux effets principaux :

- (1) Des effets de l'apprentissage explicite du code graphophonologique (le décodage), qui permet de lire tout item écrit, même s'il n'appartient pas (et n'est pas proche) au lexique écrit connu ;
- (2) des effets de l'apprentissage explicite de la morphologie écrite qui permet essentiellement de gérer les lettres muettes en lecture et de maîtriser l'orthographe en production.

L'ensemble des éléments développés précédemment peut être schématisé par la figure ci-après :

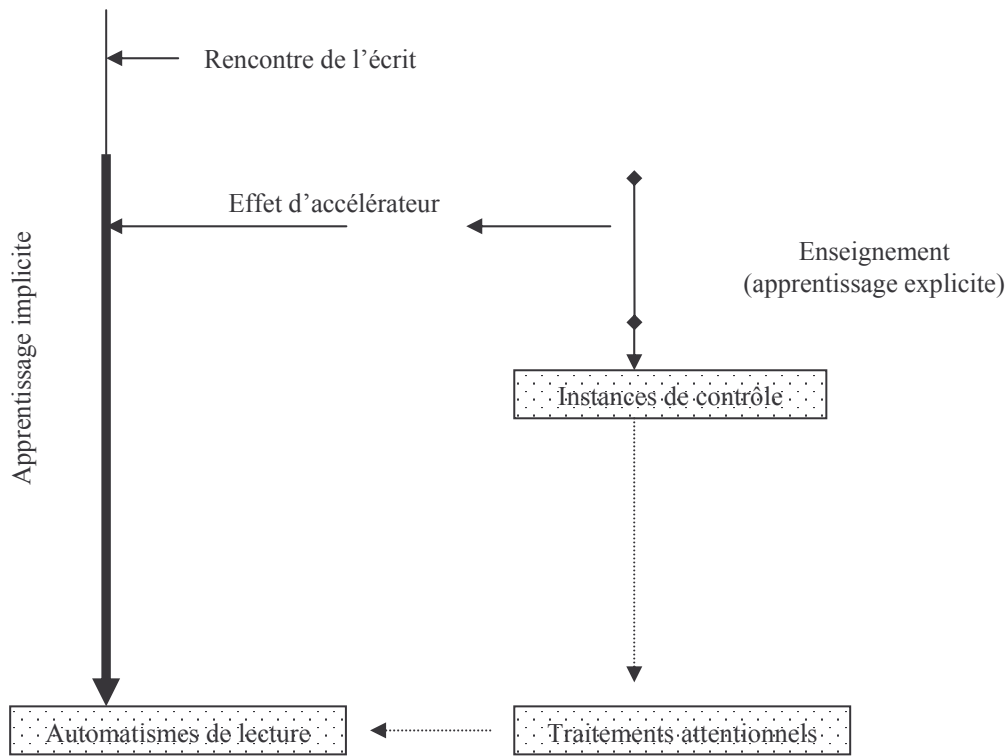


Figure 2. Apprentissages implicite et explicite de la lecture (Gombert, 2003)

Gombert (2003) postule ainsi que le caractère automatique de la lecture est toujours acquis par apprentissage implicite. Cette conception s'oppose à celle d'Anderson (1983) pour qui l'automatisation est le produit d'une transformation des traitements attentionnels par répétition de l'activité. Cette répétition permettrait le désengagement progressif de l'attention et l'accélération du traitement.

La maîtrise du langage écrit serait ainsi tributaire d'apprentissages implicites et explicites. La prise de conscience explicite des connaissances linguistiques, c'est-à-dire l'apparition des capacités métalinguistiques, n'est pas spontanée : elle nécessite un effort métacognitif qui n'est pas effectué spontanément par l'individu mais qui est généralement suscité par le contexte pédagogique (apprentissage scolaire, formel ou explicite). Le niveau implicite constitue une composante fondamentale : l'exposition à l'écrit et les activités répétées de lecture jouent notamment un rôle déterminant dans la mesure où elles permettent une association entre orthographe et morphologie et entre orthographe et phonologie.

Avant même d'apprendre de manière explicite la lecture et l'écriture, l'enfant est en mesure de développer une base de connaissances sur l'écrit (Ecalte, 2000). L'apprentissage des régularités, qui concerne d'une part les patterns visuels (régularités orthographiques), d'autre part les mots oraux et les significations qui leur sont associées (régularités graphophonologiques et orthographico-morphologiques) préexisterait donc à l'enseignement de la lecture et s'opérerait

sous une forme implicite dans la mesure où ces régularités ne répondent pas à des règles apprises explicitement (Gombert, 2003). Le système cognitif de l'enfant n'ignore aucunement ces régularités perçues et extraites de son environnement, dont l'importance a d'ailleurs déjà été montrée dans la reconnaissance de mots écrits chez le lecteur expert (Colé, 1999 ; Colé, Segui, & Taft, 1997).

Le modèle de Gombert (2003) rend compte à la fois de l'apprentissage implicite et explicite, alors que jusqu'à maintenant, les modèles développementaux en stades comme celui de Frith (1985) tendaient à ignorer l'apprentissage implicite en se centrant essentiellement sur les changements survenant sous l'effet de l'instruction et en décrivant avant tout les progrès scolaires de l'enfant (suite aux effets de l'enseignement explicite). Pourtant, d'un point de vue développemental, il est difficile de ne pas établir de liens entre les connaissances que l'enfant possède avant d'aborder explicitement l'apprentissage de la lecture et la façon dont il traite les mots écrits en début d'apprentissage. Les modèles par analogie développés précédemment (Gombert *et al.*, 1997 ; Goswami, 1999 ; Goswami & Bryant, 1990) appuient cette idée d'apprentissage implicite en prenant en compte le rôle des traitements implicites dans l'apprentissage.

En outre, la modélisation proposée par Gombert (2003) intègre la composante orthographique et morphologique et ce, à des niveaux très précoces de développement. Ces deux composantes feraient donc partie intégrante de l'apprentissage de la lecture au même titre que les aspects graphophonologiques. Au regard de ces constats, l'objectif des chapitres théoriques suivants consiste à mettre en évidence une sensibilité précoce chez l'enfant aux régularités morphologiques et orthographiques de la langue. Cette sensibilité permettrait l'émergence de connaissances implicites qui pourraient influencer l'acquisition même du système de règles de conversion graphème-phonème et faciliter la lecture (Martinet, 2001). Les aspects graphophonologiques ne sont pas en reste pour autant dans la mesure où cette sensibilité précoce concernerait également les régularités des correspondances graphèmes-phonèmes. Le chapitre 2 consiste justement en l'étude développementale de l'émergence de ces régularités graphophonologiques.

CHAPITRE 2 – LECTURE ET GRAPHOPHONOLOGIE

La mise en avant de l'importance du concept de conscience phonologique dans la littérature sur l'acquisition de la lecture a eu pour conséquence que bien souvent le rôle d'autres types de connaissances, notamment morphologiques et orthographiques, introduites à la fin du chapitre précédent, a été négligé. Il faut toutefois reconnaître l'importance fondamentale des connaissances phonologiques dans l'apprentissage de la lecture. Ainsi, depuis les années 70 et les nombreux travaux menés par Liberman (1973) et ses collaborateurs (Liberman, Shankweiler, Fisher, & Carter, 1974), la compétence phonologique est considérée comme la clé permettant de comprendre le processus complexe par lequel l'enfant apprend les relations entre langage oral et langage écrit. Les recherches établissant et discutant les liens entre l'apprentissage de la lecture et la capacité d'identifier les composants phonologiques des unités linguistiques et de les manipuler de façon intentionnelle (i.e., la maîtrise métaphonologique) ont été et sont toujours très abondantes (voir Gombert, 1992 pour une revue).

L'objectif de ce chapitre n'est pas de remettre en cause ces constats, une partie de ce chapitre sera d'ailleurs consacrée aux principales recherches démontrant la relation forte et établie entre conscience phonologique et lecture. Il s'agit néanmoins de mettre en évidence que la plupart de ces travaux renvoie surtout aux liens entre connaissances explicites des correspondances graphèmes-phonèmes et apprentissage de la lecture alors que certaines études récentes montrent qu'une forme de sensibilité précoce à la régularité et la fréquence de ces associations existe avant tout apprentissage explicite. Ces connaissances implicites pourraient elles-mêmes constituer une aide à la reconnaissance de mots écrits.

Avant de développer ces études, une définition précise de la conscience phonologique, des graphèmes et phonèmes s'avère nécessaire.

1- Les connaissances graphophonologiques

1.1- Qu'est-ce que la graphophonologie ?

Dans le terme générique de « graphophonologie », deux termes distincts entretenant de fortes relations apparaissent : le phonème et le graphème. Le phonème est une unité linguistique de l'oral : un « son » de la langue, une unité distinctive qui n'est pas porteuse de sens, mais permet de distinguer entre elles les unités signifiantes encore appelées morphèmes. En effet, les unités signifiantes minimales – les morphèmes – sont construites au moyen d'unités plus petites que

sont les phonèmes. Par exemple, *mer* contient trois phonèmes : /mɛʀ/. Ces phonèmes représentent les unités minimales de la deuxième articulation du langage. Le phonème est une unité segmentale et discrète se présentant dans chaque langue en nombre restreint et fini. Le graphème peut être défini comme le pendant graphique du phonème. Dans l'exemple précédent, le mot *mer* contient trois graphèmes ou lettres, à savoir [m], [e] et [r]. Le graphème n'est cependant pas toujours assimilable à la lettre. Ainsi, le mot *chat* fait apparaître deux phonèmes car il se prononce /ʃa/ et au phonème /ʃ/ correspond le graphème [ch] qui est lui-même composé de deux lettres (Ducard, Honvault, & Jaffré, 1995).

Outre les termes de phonème et de graphème, la syllabe constitue elle aussi une unité distinctive importante de la langue. La syllabe a une réalité articulatoire : c'est un groupe de sons qui se prononce en une seule émission de voix. Elle est facile à isoler sur le plan acoustique car elle est signalée par des variations d'intensité. La syllabe peut être décomposée en deux unités intermédiaires : l'attaque et la rime (Treiman, 1989). L'attaque (appelée aussi amorce) est la partie consonantique initiale de la syllabe. C'est une consonne isolée ou un groupe de consonnes. La rime est constituée par la voyelle (pic) et les éventuelles consonnes qui la suivent (coda). Par exemple, l'attaque du mot *bloc* est /bl/ et la rime est /ok/.

Ces unités sont fondamentales pour lire dans un système alphabétique comme le français. En effet, dans un tel système, le principe fondamental de la lecture consiste à coder les sons ou phonèmes des mots parlés à l'aide de symboles que sont les lettres ou graphèmes. De cette façon, le mot parlé /sak/ se représente à l'écrit avec la séquence de lettres *sac*. L'apprenti lecteur doit donc comprendre le principe alphabétique, ce qui lui permet d'utiliser une procédure de lecture phonologique qui consiste à traduire la séquence de lettres du mot lu en une séquence de sons correspondant à ces lettres. De cette façon, l'apprenti lecteur active le code phonologique du mot lu qui favorise l'activation du code sémantique correspondant, stocké dans son lexique oral. Au début de l'apprentissage, l'apprenti lecteur lirait donc les mots à partir des connexions entre les codes phonologiques et sémantiques des mots qui ont été développées au cours de l'acquisition du langage parlé. Selon Share (1995), cette procédure phonologique, utilisée systématiquement, fonctionnerait comme un mécanisme d'auto-apprentissage soutenant l'élaboration des codes orthographiques des mots.

Le terme de procédure phonologique est un terme générique pour désigner le recours au code phonologique pour lire les mots. Ce code phonologique peut être obtenu à partir de la mise en œuvre de deux procédures dénommées procédure d'adressage et procédure d'assemblage (voir chapitre précédent, Coltheart, 1978). Dans le cas de la procédure d'adressage (ou voie lexicale ou voie directe ou encore procédure orthographique), le code phonologique est récupéré

lexicalement sur la base d'une comparaison entre le mot écrit et son code orthographique stocké en mémoire. En recourant à la voie d'assemblage (ou voie sublexicale ou voie indirecte ou encore procédure par médiation phonologique), le lecteur « générerait » le code phonologique des mots à partir des codes orthographiques des lettres ou de groupes de lettres, *via* l'utilisation de correspondances entre les unités sublexicales orthographiques et phonologiques. Selon le point de vue adopté, cette procédure s'appuierait sur des règles (qui peuvent renvoyer au versant explicite de l'apprentissage) ou des régularités statistiques (i.e., versant implicite).

Pour comprendre le principe alphabétique, de nombreuses recherches ont montré que l'apprenti lecteur doit procéder à une analyse consciente de la structure du langage parlé nommé généralement « conscience phonologique » ou « capacité métaphonologique ».

1.2- Capacités métaphonologiques et lecture

La conscience phonologique est définie comme la compétence qui permet d'identifier les composants phonologiques des unités linguistiques et de les manipuler intentionnellement (Gombert, 1990 ; Gombert & Colé, 2000). Ces unités renvoient à des segments sans signification, tels que les syllabes et les phonèmes, mais également des unités infrasyllabiques entre la syllabe et le phonème, comme l'attaque et la rime. Cette conscience phonologique amène l'apprenti lecteur à concevoir que les mots parlés sont constitués d'une séquence de sons, ce qui lui permet – à terme – de comprendre que les sons les plus élémentaires (les phonèmes) correspondent à l'écrit, tout au moins dans une écriture alphabétique, à des lettres ([a], [t]) ou à des groupes de lettres ([ou], [ch]). Cette capacité à segmenter les mots parlés en ses unités les plus élémentaires s'avère indispensable pour que l'apprenti lecteur soit en mesure de découvrir les correspondances graphophonologiques.

L'enfant de langue française, lorsqu'il aborde la première année d'école primaire, dispose de représentations des sons de parole de sa langue bien spécifiées. Ces représentations phonologiques sont extrêmement importantes puisque c'est à partir de leur manipulation consciente que l'apprenti lecteur va développer les compétences métaphonologiques nécessaires pour apprendre à lire. Ces capacités métaphonologiques désignent, selon Gombert (1992), des connaissances phonologiques mentalisées et intentionnellement manipulées par l'enfant. L'apparition de ces capacités métalinguistiques doit être distinguée de la simple mise en place de la faculté de langage. Pour Gombert (2002), une distinction doit notamment être faite entre les capacités manifestées dans les comportements spontanés (par exemple, le trouble engendré chez le jeune enfant par l'audition d'une phrase agrammaticale comme *je vouloir le balle*) et les capacités qui sont fondées sur des connaissances mentalisées et intentionnellement appliquées

(par exemple : corriger la syntaxe d'un texte écrit). Pour Gombert (1990), la notion d'« épilinguistique » (voir paragraphe 3, chapitre 1) semble être parfaitement adaptée à la désignation des activités qui s'apparentent au comportement métalinguistique mais qui sont en réalité effectuées sans contrôle conscient. La métalinguistique est selon Gombert (1990) « l'attitude réflexive du sujet sur les objets langagiers, leur nature et leur fonction, et leur manipulation » (p. 11). Posant comme principe qu'un caractère réfléchi et intentionnel est inhérent à l'activité strictement métalinguistique, il propose d'utiliser le terme « épilinguistique » pour désigner les comportements qui, bien qu'isomorphes aux comportements métalinguistiques, ne sont pas le résultat d'un contrôle conscient par le sujet de ses propres traitements linguistiques. Toujours selon Gombert (1990), les compétences métalinguistiques se développent à partir des modèles linguistiques fournis par l'adulte, selon les quatre étapes suivantes :

- acquisition des premières habiletés linguistiques ;
- acquisition de la maîtrise épilinguistique ;
- acquisition de la maîtrise métalinguistique ;
- automatisation des processus.

Ces capacités métalinguistiques semblent être essentielles dans l'apprentissage de la lecture. En effet, la lecture étant une tâche métalinguistique formelle, son apprentissage nécessite de la part de l'enfant de développer une conscience explicite des structures linguistiques qui devront être intentionnellement manipulées. Dans ce cadre, les capacités métaphonologiques correspondent à « la capacité d'identifier les composants phonologiques des unités linguistiques et de les manipuler de façon délibérée » (Gombert, 1990, p. 29). Les capacités métaphonologiques se différencient des capacités épiphonologiques qui, elles, sont fondées sur l'expérience langagière et l'exposition répétée aux écrits. En effet, l'expérience langagière permettrait aux enfants d'acquérir une connaissance inconsciente (i.e., implicite) de certains aspects du système de sons de la langue. Ainsi, l'enfant pourrait organiser implicitement des segments phonologiques. L'habileté épiphonologique constituerait un prérequis à la mise en place d'une capacité métaphonologique : l'habileté métaphonologique apparaîtrait une fois le fonctionnement épiphonologique bien établi et ce, au travers d'activités de manipulations phonologiques dans l'apprentissage de la lecture.

Les résultats expérimentaux obtenus par Ecalle et Magnan (2002b) confortent l'idée selon laquelle les capacités épiphonologiques seraient à la base du développement des capacités métaphonologiques. En effet, leur étude longitudinale réalisée auprès de 36 enfants de GSM et de CP de langue française montre que les compétences épiphonologiques acquises en GSM

semblent prédictives des compétences métaphonologiques développées un an plus tard en CP. Dans cette recherche, deux types de tâches sont utilisées : l'une épiphonologique qui consiste à désigner deux mots partageant une unité phonologique parmi quatre mots (*zéro-ballon-zébu-toupie*) et l'autre métaphonologique dans laquelle il s'agit de détecter l'unité commune à deux mots (*chameau-chalet*). Les capacités métaphonologiques sont assimilables à la capacité d'analyser explicitement et automatiquement la parole grâce aux sons. Les résultats obtenus mettent en évidence que le niveau épiphonologique mesuré en GSM rend compte du développement métaphonologique ultérieur en CP.

Les capacités métaphonologiques renvoient aux connaissances « explicites » et se distinguent des connaissances phonologiques « implicites » (épiphonologiques) manipulées par l'enfant de manière automatique et sans focalisation attentionnelle. Ces dernières se développeraient « naturellement » au cours de l'acquisition du langage parlé et les premières seraient développées ou sollicitées par l'apprentissage de la lecture. Cet apprentissage explicite est bien évidemment fondamental à l'émergence des capacités métaphonologiques, émergence qui repose également sur la présence préalable de connaissances implicites. Une question encore débattue concerne la nature des unités impliquées dans ces traitements, qu'ils soient implicites ou non. Ces unités peuvent être syllabiques, phonémiques ou infrasyllabiques. Cependant, il semble que c'est le développement de la conscience phonémique, et non celui de la conscience d'unités plus larges – syllabiques ou rimiques – qui prédit le niveau de réussite en lecture (Hulme, Hatcher, Nation, Brown, Adams, & Stuart, 2002 ; Muter, Hulme, Snowling, & Taylor, 1998).

La majorité des chercheurs s'accorde ainsi sur l'importance du développement de la conscience phonologique dans l'apprentissage de la lecture : les recherches ont notamment démontré que la conscience phonologique jouait un rôle majeur dans cet apprentissage, en particulier lors du décodage (Bryant, 1993 ; Gombert, 1992 ; Goswami & Bryant, 1990 ; Rack, Hulme, & Snowling, 1993 ; Treiman, 1989). Plusieurs recherches ont aussi démontré qu'un entraînement spécifique aux habiletés phonologiques permettait d'améliorer considérablement le traitement phonologique (le décodage et la reconnaissance des mots). Le point suivant va mettre en perspective les différentes études établissant une relation forte entre conscience phonologique et lecture.

2- Conscience phonologique et lecture : une relation établie et largement démontrée

La question des relations entre les connaissances phonologiques et la réussite en lecture a été posée par de nombreux chercheurs au cours des 30 dernières années. Les travaux réalisés dans ce domaine ont établi que l'apprentissage de l'écrit se construit en partie sur la capacité à développer une connaissance phonologique des propriétés formelles de la langue parlée (Mattingly, 1972). Ainsi, le rôle de la médiation phonologique dans le déroulement du processus de traitement de l'information écrite a été mis en évidence par de très nombreux travaux, à la fois dès les premières étapes de l'acquisition du langage écrit et chez le lecteur expert (pour une revue, voir Content, 1990, 1993 ainsi que Sprenger-Charolles & Siegel, 1997). Les études longitudinales, notamment, démontrent que la conscience phonologique mesurée antérieurement à l'acquisition des habiletés de lecture prédit la performance en lecture. Un tel constat a été réalisé dans des langues très différentes de l'allemand au français en passant par le suédois et le chinois (pour une revue de ces études longitudinales, voir Ecalle & Magnan, 2002a). D'autres études ont par ailleurs montré que les enfants faibles lecteurs présentent des performances nettement inférieures aux bons lecteurs dans des tâches évaluant le niveau d'habiletés phonologiques (Bowey, Cain, & Ryan, 1992 ; Bruck & Treiman, 1992 ; Gillon & Dodd, 1994 ; Jimenez, 1997).

La relation entre conscience phonologique et habileté en lecture apparaît aujourd'hui indiscutable (pour une revue, voir également Adams, 1990 ; Brady & Shankweiler, 1991 ; Goswami & Bryant, 1990 ; Wagner & Torgensen, 1987). Toutefois, même si de nombreuses études ont démontré que les compétences phonologiques facilitent l'acquisition de la lecture, la nature des liens entre habiletés phonologiques et apprentissage de la lecture reste à préciser. En effet, pour certains (comme Content, 1985), le développement de la compétence phonologique serait fortement influencé par l'alphabétisation alors que pour d'autres (Bryant & Bradley, 1985), cette compétence existerait bien avant l'apprentissage de la lecture. Ainsi, certains chercheurs considèrent qu'il faut impérativement un enseignement explicite centré sur les phonèmes et les conversions graphèmes-phonèmes pour que les enfants puissent découvrir les unités abstraites de la langue : ce serait le processus de lecture lui-même qui inciterait le lecteur à se focaliser sur les segments phonologiques pertinents de la langue (Ehri, 1989 ; Morais, 1991 ; Morais, Alegria, & Content, 1987). Par exemple, Morais *et al.* (1987) montrent que des adultes portugais analphabètes réussissent moins bien aux tâches phonologiques que des ex-analphabètes après avoir appris à lire (Morais, Cary, Alegria, & Bertelson, 1979). Ce constat repose également sur une étude menée par Alegria et Morais (1979) dans laquelle les auteurs ont comparé deux

groupes d'enfants âgés de 6 ans soumis à une tâche de manipulation de phonèmes. Cette tâche consistait à effectuer des opérations explicites sur des segments telles que les compter, les supprimer, en ajouter ou encore en remplacer. Le premier groupe a été testé en novembre et le second en février. Les résultats montrent que les performances du second groupe sont significativement supérieures à celles du premier groupe. Les auteurs en concluent que la conscience phonémique est le résultat de l'enseignement explicite de la lecture. Une étude interlangues récente menée par Mann et Wimmer (2002) va dans le sens de cette conclusion. Cette étude avait pour objectif la comparaison des capacités de conscience phonémique d'enfants américains ayant appris les relations lettres-sons à l'école à celles d'enfants allemands qui ne les ont pas apprises. Les enfants américains réussissaient très bien les tâches phonologiques comparativement aux allemands : la conscience phonologique se développerait donc principalement en conséquence de l'exposition à la littéracie.

Ouzoulias (2001) nuance un tel point de vue en considérant qu'un enseignement explicite et direct des correspondances graphèmes-phonèmes n'est pas le seul moyen pour les enfants de découvrir ces règles de conversion. D'autres auteurs comme Bryant et Bradley (1985) considèrent également que les enfants peuvent découvrir les règles de conversion graphèmes-phonèmes sans enseignement explicite. Selon eux, les aptitudes phonologiques se développent bien avant l'apprentissage de la lecture. Dans le même sens, Bertelson et De Gelder (1990) constatent que les enfants acquièrent une certaine connaissance phonologique qui augmente tout au long de la période préscolaire et qui constitue une base pour établir la relation lettre/son au début de l'apprentissage de la lecture. Pour ces auteurs, la conscience phonologique jouerait un rôle causal dans l'acquisition de la lecture permettant ou assistant au moins au tout début l'acquisition des toutes premières capacités de lecture. La formulation d'une telle conception se focalise sur le niveau phonème : dans la mesure où les lettres sont représentées habituellement par des phonèmes individuels dans les systèmes alphabétiques, un enfant a besoin d'être conscient des segments phonémiques dans les mots oraux avant d'être capable d'apprendre leurs correspondances avec les lettres et les graphèmes (Gough & Hillinger, 1980 ; Mattingly, 1972).

Ces deux types de position renvoient en réalité à deux niveaux différents d'habiletés phonologiques, l'un explicite et l'autre implicite. En effet, si l'apprentissage explicite des règles de correspondances graphèmes-phonèmes apparaît bien comme une condition nécessaire à l'apprentissage de la lecture, une sensibilité de type implicite à certaines régularités dans ces correspondances pourrait également exister chez les enfants de façon précoce et constituer une aide à l'apprentissage de la lecture. L'objectif du paragraphe suivant est de montrer que dans le système alphabétique du français, il existe des correspondances graphèmes-phonèmes qui sont

plus régulières ou consistantes que d'autres : des connaissances implicites sur cette consistance pourraient faciliter la reconnaissance de certains mots et non mots écrits.

3- Implicite versus explicite en graphophonologie : l'effet de consistance graphème-phonème

Une écriture alphabétique comme le français est supposée transcrire principalement les phonèmes. Une des difficultés majeures des écritures alphabétiques vient de ce que, isolées, les lettres de l'alphabet ne permettent pas de transcrire tous les phonèmes des différentes langues. Par conséquent, l'unité de base de l'écrit n'est pas la lettre, mais le graphème qui renvoie au phonème, unité de base du système phonologique. Le graphème peut contenir soit une ou plusieurs lettres, soit des marques diacritiques (les accents ou les trémas sur les voyelles comme dans *château* ou *noël*), ce qui peut constituer une source de difficultés supplémentaire pour l'apprenti lecteur. Ainsi, en français, il existe 130 graphèmes permettant de transcrire 30-35 phonèmes (Catach, 1980). Une des difficultés de la langue française est notamment l'existence d'allographes, c'est-à-dire de graphèmes différents pour un même phonème. Un même graphème peut aussi correspondre à plusieurs phonèmes en fonction de sa place dans le mot : par exemple, en fonction de la voyelle qui suit, la lettre [c] peut ou non correspondre au phonème /k/ comme dans *cercle* (/s... k.../). De même, la lettre [g] peut ou non être prononcée /g/ comme dans *garage*. D'autres inconsistances ou irrégularités sont plus imprédictibles comme le fait que [u] suivant [g] est ou non prononcé suivant les mots (*guerre, aiguille, linguiste*). Ainsi, la connaissance explicite des correspondances entre graphèmes et phonèmes ne suffit pas à une maîtrise des mots écrits en raison de ces inconsistances.

3.1- Notion de régularité et de consistance

Les systèmes linguistiques écrits à base phonographique (i.e., utilisant les règles de correspondances graphèmes-phonèmes) varient, entre autres choses, selon le degré de régularité des relations graphèmes-phonèmes : dans certains cas, cette correspondance est quasi biunivoque, c'est-à-dire qu'à chaque phonème correspond un et un seul graphème et inversement. Dans d'autres cas, cette correspondance s'avère plus lâche, soit que plusieurs graphèmes renvoient à un même phonème (par exemple, en français, [o], [au], [eau] à /o/), soit qu'un même graphème est associé à plusieurs phonèmes (*cf. les notions / nous notions*).

Les systèmes d'écriture diffèrent donc au niveau de la consistance graphème-phonème. Le terme de consistance désigne de façon générale la consistance des régularités statistiques entre orthographe et phonologie. La présence d'irrégularités dans les correspondances

graphèmes-phonèmes permet de situer les systèmes alphabétiques sur un continuum allant des orthographe « transparentes » ou régulières aux orthographe « opaques » ou irrégulières. Cette notion de consistance graphophonologique est ainsi relative à la multiplicité des codes phonologiques qui peuvent être assignés à une unité orthographique particulière. Par exemple, la voyelle anglaise, [oa] a différentes prononciations selon qu'elle est incluse dans *road* ou dans *broad*. De même, l'unité [eaf] est prononcée différemment dans les mots anglais *deaf* et *leaf*. Par définition, la proportion de mots pour lesquels une unité orthographique donnée s'actualise par une prononciation donnée, eu égard au nombre total de mots qui comporte cette unité particulière, fournit une estimation du degré de consistance d'une correspondance graphophonie.

La consistance ou transparence de l'orthographe dépend du point de vue pris (lecture et écriture). Ainsi, en français les relations graphème-phonème (pour la lecture) sont plus simples à maîtriser que les relations phonème-graphème (pour l'écriture). Si le recours au traitement phonologique est bien au centre de l'apprentissage de la lecture, quelle que soit la transparence des relations entre l'écrit et l'oral, les enfants devraient utiliser précocement les correspondances graphophonémiques pour apprendre à lire. Toutefois, la facilité de cet apprentissage devrait également dépendre du degré de transparence des relations graphophonémiques. En effet, la connaissance du phonème ne suffit pas, à elle seule, à décider du graphème à utiliser. Cela se vérifie notamment en français : par exemple, le phonème /o/ est réalisé dans les trois-quarts des cas par la lettre [o], dans un cinquième par [au] et dans deux ou trois pour cent des cas par [eau] (Catach, Gruez, & Duprez, 1986). En français toujours, la correspondance /s/ - [c] est inconsistante : le phonème /s/ peut s'écrire [c], [s], [t]... . Dans le cas de relations inconsistantes, certaines correspondances sont cependant plus fréquentes que d'autres. Ainsi, le phonème /s/ se transcrit plus fréquemment par [s] que par [c] ou [t]. Lamarche (1986) a voulu tester si les enfants pouvaient être sensibles à telles consistances relatives. Pour cela, cet auteur a mené une étude partielle du /o/ et de façon complète du /s/ et du /e/ auprès d'une population francophone d'enfants âgés de 7 à 11 ans. Au début, pour /o/ comme pour /s/, un graphème domine, à savoir le [o] et le [s]. Or, [o] et [s] sont les réalisations graphémiques les plus fréquentes de /o/ et /s/ (cf. Catach *et al.*, 1986, p. 97 pour /o/ et pp. 161-162 pour /s/). Lamarche (1986) a ainsi constaté que les enfants faisaient appel très tôt à des séquences de lettres associées à des contextes d'occurrence sans que ces structures et ces correspondances n'aient été enseignées de manière explicite. De façon relativement précoce, les enfants témoigneraient d'une sensibilité à la diversité des graphèmes associés à un même phonème. Cette sensibilité pose la question de la nature de l'effet de consistance graphophonémique : cet effet est-il lié à la fréquence

d'occurrence par le truchement de connaissances implicites et/ou est-il la conséquence de l'instruction dispensée en classe ? Cette question a notamment été abordée par Fijalkow et Fijalkow (1994) qui constatent que 71 % des enseignants de CP utilisent un manuel, lequel comporte nécessairement des mots incluant des réalisations graphémiques de /o/ et /s/. Cette fréquentation d'écrits plus ou moins importante pourrait permettre l'extraction par l'enfant de régularités dans les correspondances graphèmes-phonèmes de façon implicite. Ce constat fait apparaître que les connaissances des enfants peuvent être attribuables au support : par exemple, les ouvrages utilisés par l'enseignant ou mis à disposition des enfants dans l'environnement scolaire et familial. Ces connaissances peuvent aussi être imputables à l'enseignement explicite réalisé en classe. Pour Lamarche (1986), l'accroissement de la diversité des graphèmes utilisés par les enfants peut être attribué à la fois au corpus d'items lexicaux rencontrés à l'écrit (implicite) et à l'instruction directe par les enseignants (explicite).

Cette question de la nature de l'effet de consistance (implicite *versus* explicite) a été également traitée en étudiant le sensibilité aux associations (ou non associations) entre certaines positions dans les items (début/milieu/fin) et certains graphèmes. Par exemple, Fayol, Lété et Gabriel (1996) ont étudié la distribution du graphème [o] et du graphème [s] en fonction des types de graphèmes et des positions en production écrite chez des enfants de CP et de CE1. En français, [eau] n'est jamais en début de mot tout comme [ss] ; [eau] est généralement en fin de mot et [ss] entre deux voyelles en milieu d'items tandis que [ç] se trouve devant une voyelle. L'objectif des auteurs était ainsi de vérifier si les enfants tenaient compte (i.e., étaient sensibles) de ces contraintes graphotactiques. La tâche des enfants consistait à transcrire les mêmes non mots à deux moments dans l'année, février et avril : neuf avec le phonème /o/ et neuf avec le phonème /s/. Les dix-huit non mots étaient dictés par blocs de trois. Les enfants devaient écrire les items comme ils pensaient qu'ils devraient s'écrire. Les non mots étaient formés de telle sorte que les phonèmes /o/ et /s/ apparaissaient dans trois positions : début, milieu et fin de mots. Les résultats indiquent que le graphème [o] est majoritaire aux deux niveaux scolaires : plus de 80 % des graphèmes en CP et plus de 50 % des graphèmes en CE1. De plus, [o], initialement distribué sur tous les emplacements (début, milieu, fin d'items), voit sa fréquence d'occurrence diminuer en début (de 70 % en CP à 40 % en CE1) et en fin d'items (80 % en CP à 47 % en CE1) mais non (ou très peu : de 85 à 70 % du CP au CE1) en milieu d'items. La proportion de graphèmes [au] augmente du CP au CE1, essentiellement en début d'items (de 30 % en CP à 60 % en CE1) mais également en milieu (de 3 à 27 %) et fin (de 11 à 25 %) d'items. La proportion de [eau] augmente mais seulement en fin d'items (de 8 % en CP à 20 % en CE1). Ainsi, au départ, il existe une dominance du [o] puis apparaît une diversification des graphèmes et la mise en place

d'une association régulière entre certains d'entre eux et les positions dans les non mots. Concernant la distribution du graphème [s], celui-ci est au départ dans 85 % des cas la transcription du phonème /s/, et ce quelles que soient les positions dans les items. Puis, l'évolution est lente au CP et en CE1 (février et entre février et avril) : le taux de [s] diminue (de 85 à 70 %) mais les graphèmes se distribuent inégalement en fonction des positions (début/milieu/fin) dans les items. Ainsi, [s] reste très fortement dominant (90 % toutes sessions confondues) en début d'item même si [c] et quelques rares [ss] y apparaissent aussi. En fin de mot, les enfants utilisent soit [s] s'ils terminent l'item par cet unique graphème (cas le plus fréquent) soit [ss] ou [c] s'ils ajoutent un [e] en fin d'item. Enfin, [s] est moins fréquent en milieu d'item, et cela dès le CP, là où [ss], [c] et [ç] sont le plus souvent employés. Au graphème [s] systématiquement employé, quel que soit son emplacement dans les items, succèdent donc différents graphèmes – [ss], [ç], [c] et [s] – qui tendent à s'associer de manière privilégiée avec certaines positions : [s] en début, [ss], [ç], [c] en milieu, [s] ou [ss] en fin. Ainsi, très tôt, l'enfant est capable d'établir des associations entre certaines positions dans les items (début/milieu/fin) et certains graphèmes. Ces associations probabilistes ne renvoient pas à des règles spécifiques d'utilisation : ce sont des régularités propres à un système donné (en l'occurrence, le système linguistique du français) qui ne peuvent être enseignées sous forme de règles. Les régularités seraient apprises par apprentissage implicite à partir de la fréquentation de mots écrits en lecture.

Plus récemment, Pacton, Fayol et Perruchet (1999) ont confirmé de telles conclusions avec le phonème /o/ en montrant à l'aide d'une tâche de production orthographique de non mots que de jeunes enfants de CP transcrivent le phonème /o/ en utilisant au moins deux graphèmes différents et utilisent [eau] en position finale davantage qu'en position initiale conformément aux régularités de la langue. Ces associations sont ici aussi de type probabiliste : leur « connaissance » renverrait donc au versant implicite.

Treiman (1993) avait déjà obtenu des données similaires en anglais. Elle avait notamment observé que des enfants de niveau CP n'utilisaient le graphème [ck] qu'en fin d'items alors qu'ils utilisaient les graphèmes [c] et [k] en début de mots pour transcrire le son /k/. Ces utilisations étant conformes aux régularités propres à la langue anglaise, Treiman (1993) en conclut qu'elles sont de type probabiliste et ne peuvent donc être enseignées sous forme de règles. Elles seraient acquises par apprentissage implicite à partir de la fréquentation – exposition répétée – en lecture et en production, d'un corpus d'exemples (Seger, 1994).

Ces exemples indiquent que très précocement, les enfants sont sensibles aux associations fréquentes entre certains graphèmes et phonèmes ainsi qu'à certains emplacements internes aux items lexicaux. Cette sensibilité précoce suggère que leur apprentissage s'effectue de manière

implicite au contact du corpus écrit et cela sans qu'il y ait eu d'instruction directe ou explicite. Les enfants découvrirait et acquerraient implicitement ces régularités au cours des lectures et des activités de copies de mots. Ensuite, ils pourraient généraliser les patrons ainsi induits à des mots nouveaux pour eux (lors de la première rencontre, les mots nouveaux ont le statut de non mots) qu'ils liraient et transcriraient par analogie avec certains mots connus (Englert, Hiebert, & Stewart, 1985).

En lecture notamment, cette notion de consistance peut s'avérer importante sur la façon dont les habiletés sont acquises et ce, en fonction des langues d'apprentissage. Par exemple, l'italien, l'espagnol, le grec et le finnois ont des orthographe régulières où les relations entre graphèmes et phonèmes sont plutôt transparentes, c'est-à-dire que les lettres ou groupes de lettres s'associent de façon consistante aux phonèmes. A la fin du niveau CP, les enfants de ces pays ont des scores très proches des plafonds en termes de précision de lecture (Goswami, Gombert, & Fraca de Barrera, 1998 ; Seymour, Aro, & Erskine, 2003). Dans d'autres orthographe, comme celle de l'anglais, ces relations sont plus opaques et de nombreux mots ne peuvent se lire correctement en utilisant les relations graphophonémiques. Ainsi, les enfants qui apprennent à lire en anglais sont face à un haut niveau d'inconsistance où les mêmes patrons orthographiques peuvent être prononcés de différentes manières et où les mêmes prononciations peuvent s'écrire de différentes façons (Perry, Ziegler, & Coltheart, 2002 ; Ziegler, Stone, & Jacobs, 1997). Dès lors, les enfants des pays anglophones mettent beaucoup plus de temps à obtenir un haut niveau de performances en lecture comparativement aux enfants apprenant des orthographe plus régulières (Goswami *et al.*, 1998 ; Seymour *et al.*, 2003). Les systèmes orthographiques de l'allemand et du français occupent une place intermédiaire, celui de l'allemand étant plus proche que celui de l'espagnol et du français. Le point suivant vise à montrer *via* des études comparatives interlangues et des études francophones en quoi le degré de transparence dans les relations graphèmes-phonèmes (i.e., la consistance) peut affecter la reconnaissance de mots et de non mots ainsi que le niveau de lecture.

3.2- Effet de consistance et de régularité sur la reconnaissance de mots écrits et sur la lecture

Les comparaisons interlangues renseignent sur les différents degrés de consistance graphophonologique en fonction des systèmes linguistiques. Il s'agit notamment de voir en quoi la plus ou moins grande transparence d'une orthographe dans un système donné comme le français peut avoir un impact sur la reconnaissance de mots et de non mots.

De nombreux chercheurs se sont intéressés à ce type de comparaison : Oney et Goldman (1984) pour les comparaisons turc-anglais, Thorstad (1991) pour l'italien et l'anglais ou encore Goswami, Porpodas et Weelwright (1997) qui ont comparé le grec à l'anglais. Tous ont mis en évidence une supériorité en lecture de non mots pour les enfants apprenant à lire des orthographe régulières telles que l'italien, le grec et le turc par rapport à l'orthographe opaque de l'anglais. Toutefois, Wimmer et Goswami (1994) ainsi que Jansen (1995) ont souligné les problèmes méthodologiques de telles études comparatives quand les stimuli de différentes orthographe sont comparés. Ces auteurs ont notamment montré qu'il n'existait pas de différence au niveau des habiletés à lire les mots les plus fréquents selon les différentes orthographe : les différences sont surtout importantes au niveau de l'habileté à lire des non mots. Ainsi, les principales difficultés des enfants anglais résident dans leurs capacités relativement pauvres en décodage phonologique. L'acquisition des habiletés de recodage phonologique en anglais est en effet lente et difficile. Le pourcentage moyen d'erreurs en lecture de non mots à la fin du niveau CP varie ainsi entre 40 et 80 % (Jorm, Share, MacLean, & Mathews, 1984 ; Juel, Griffith, & Gough, 1986 ; Seymour *et al.*, 2003 ; Treiman, Goswami, & Bruck, 1990). Par opposition, en grec, où l'orthographe est plus régulière, les enfants du même âge font seulement 10 % d'erreurs en lecture de mots et non mots (Porpodas, 1999). Dans une revue relativement récente, Landerl (2000) indique que les enfants dans des orthographe régulières comme l'italien, le néerlandais, l'allemand, le grec, le portugais ou le turc font moins de 25 % d'erreurs en lecture de mots et non mots à la fin du niveau CP. La plus ou moins forte consistance des correspondances graphèmes-phonèmes présente donc un impact réel sur la lecture de mots connus et nouveaux.

De nombreuses études comparatives se sont intéressées à l'analyse de l'anglais et de l'allemand en raison notamment de l'origine germanique des deux langues. Ces deux langues présentent des différences importantes au niveau de la régularité des associations lettres-sons mais possèdent une orthographe et une phonologie très similaires (Ziegler, Perry, & Coltheart, 2000). De nombreux mots ont ainsi une orthographe identique comme *zoo*, *bank*, *ball*, *land*, etc. Cette propriété rend possible l'étude du développement de la lecture et des habiletés en lecture des différentes orthographe en utilisant un matériel identique au niveau littéral. Plusieurs études ont montré de faibles capacités en lecture de non mots chez les enfants anglais comparés aux enfants allemands même lorsque le matériel utilisé est identique (Frith, Wimmer, & Landerl, 1998 ; Landerl, Wimmer, & Frith, 1997 ; Ziegler, Perry, Jacobs, & Braun, 2001). Par exemple, Frith *et al.* (1998) ont collecté des données sur l'exactitude en lecture auprès d'enfants autrichiens et anglais de 7, 8, 9 à 12 ans. Les résultats obtenus montrent une plus forte différence entre les enfants en lecture de non mots chez les plus jeunes. En allemand (orthographe

régulière), les enfants présentent dès 7 ans de meilleures performances (75 %). Après plusieurs années d'instruction, les enfants anglais rattrapent cependant les allemands même si, à l'âge de 12 ans, le décodage phonologique des enfants anglais reste toujours en deçà de celui des enfants allemands.

Le français a également fait l'objet de comparaison avec l'anglais et l'allemand. Ziegler, Jacobs et Stone (1996) ainsi que Ziegler, Stone et Jacobs (1997) ont examiné le degré de consistance (ou régularité) des relations orthographe-phonologie (OP) et phonologie-orthographe (PO) pour l'anglais, le français et l'allemand. Cette étude n'a pris en compte que des mots monosyllabiques, la consistance étant définie au niveau attaque-rime, et non à celui des correspondances graphophonémiques. Un mot est considéré comme inconsistant en lecture (OP) lorsque sa rime orthographique peut être lue de plusieurs façons (comme *-emme* de *femme* /fam/ et *flemme* /flɛm/) ou, pour l'écriture (PO), lorsque sa rime phonologique peut s'orthographier différemment (comme /o/ pour *sot*, *seau*, *saut*...). Pour la lecture, la proportion de mots inconsistants est plus élevée en anglais (12,9 %) qu'en allemand (5,8 %) et français (4,9 %). Une proportion nettement supérieure de mots inconsistants est relevée dans les trois langues pour l'écriture. Cette proportion est toutefois plus forte en français (50,3 %) qu'en anglais (28 %) et en allemand (25,9 %). Il existerait donc clairement un problème pour l'écriture en français. En revanche, pour la lecture, la différence entre l'anglais et les deux autres langues est moins forte que prévue. Ce résultat peut s'expliquer par le fait que, dans les études de Ziegler *et al.* (1996, 1997), les inconsistances ont été définies au niveau de la rime, et non à celui des correspondances graphophonémiques. Ce choix favorise l'anglais, les inconsistances pour les voyelles – tout au moins quand elles sont dans des mots monosyllabiques – étant alors réduites.

Pour le niveau graphophonémique, Peerean et Content (1999) ont réalisé une comparaison portant sur les mots monosyllabiques de l'anglais et du français. Les résultats sont alors plus proches des attentes. Pour la lecture, la proportion d'inconsistances pour les voyelles est très forte en anglais (52 %) mais pas en français (6 %). Pour l'écriture, les difficultés sont de même amplitude dans les deux langues (33 % en anglais et 32 % en français). Toutefois, dans cette étude, comme dans celle de Ziegler, Jacobs et Stone (1996) et de Ziegler, Stone et Jacobs (1997), les analyses ont uniquement pris en compte les mots monosyllabiques, ce qui conduit à sous-estimer les difficultés orthographiques de l'anglais et, peut-être, à surestimer celles du français. En effet, la prononciation des monosyllabiques en anglais est plus simple que celle des polysyllabiques. En français, les monosyllabiques ne représentent qu'une faible proportion des mots et ce sont généralement ceux qui peuvent être les plus inconsistants (6,70 % soit 2396 des 36000 mots du *Micro Robert*, 1986, cf. Content, Mousty, & Radeau, 1990).

Ces différentes données montrent que si les enfants utilisent les correspondances graphophonémiques pour apprendre à lire, cet apprentissage devrait être plus difficile pour les anglophones que pour des francophones ou des germanophones. De plus, l'acquisition de l'écriture devrait être plus difficile que celle de la lecture. Il n'en reste pas moins qu'en français, les régularités des correspondances graphèmes-phonèmes apparaissent jouer un rôle non négligeable en reconnaissance de mots écrits.

Les études interlangues montrent bien que le niveau en lecture dépend du degré de transparence de l'orthographe. L'anglais possède l'orthographe la plus opaque et l'italien ainsi que l'espagnol sont parmi les plus transparentes, l'allemand et le français occupent une position intermédiaire, l'orthographe de l'allemand étant plus transparente que celle du français. La hiérarchie des performances des enfants en lecture est conforme à ce qu'il est possible de prédire en fonction de l'opacité de l'orthographe. En effet, les scores les plus faibles se retrouvent chez les anglophones et les meilleurs chez les hispanophones (Goswami *et al.*, 1998). L'opacité de l'orthographe a donc un effet massif sur la reconnaissance de mots écrits et sur la lecture. Les points suivants vont s'attarder sur les études francophones proprement dites et sur l'impact de l'effet de consistance en lecture.

Comparativement à l'anglais, les relations graphème-phonème sont plus régulières en français (Peereman & Content, 1999), à condition toutefois de prendre comme unité de base du système d'écriture le graphème et non la lettre (*cf.* Catach, 1980). Ainsi, dans cette langue, dès le début de l'acquisition de la lecture devrait être observée une utilisation importante de la procédure sublexicale. Toutefois, les relations graphophonémiques sont moins transparentes que les relations phonographémiques.

L'étude présentée ci-après concerne l'évolution des performances en lecture et en écriture avec des enfants suivis longitudinalement depuis une étape très précoce (milieu du CP) jusqu'à la fin du CM1 (de 6 ans et demi à 10 ans). Les observations ont porté sur les effets de facteurs lexicaux (fréquence, lexicalité, analogie) et sublexicaux (régularité, longueur graphémique, complexité syllabique) en lecture à haute voix et en écriture (Sprenger-Charolles, Siegel, & Bonnet, 1998 ; Sprenger-Charolles, Siegel, Béchenec, & Serniclaes, 2003). Les réponses correctes et la nature des erreurs ont été analysées. Au départ de cette étude, en GSM, 60 non lecteurs de niveau cognitif moyen à supérieur et de milieux sociologiques représentatifs de la diversité de la population française ont été sélectionnés à partir d'un groupe d'une centaine d'enfants. Le niveau de « non lecture » a été évalué à l'aide des résultats obtenus à un test étalonné (*Bat-Elem*, Savigny, 1974) et le niveau cognitif par un test d'intelligence non verbale

(les *Matrices de Raven*, Raven, 1976). Les observations ont été recueillies en milieu et en fin de CP ainsi qu'à la fin du CE1, CE2 et CM1. En CP, les enfants étaient scolarisés dans 20 classes de 9 écoles de la banlieue parisienne. Cette diversité des environnements a permis de neutraliser, autant que possible, l'incidence du facteur enseignement.

Selon l'hypothèse de départ de cette étude, au début de l'acquisition, les enfants francophones devraient avoir recours essentiellement à la procédure sublexicale. Cette hypothèse a été testée à partir de l'examen des effets de fréquence, de lexicalité, d'analogie et de régularité en lecture à haute voix et en écriture. Pour vérifier cette hypothèse, un effet de régularité au cours des premiers mois de l'acquisition sans incidence de la fréquence, de la lexicalité ou de l'analogie devrait être observé. Ainsi, les mots réguliers devraient être mieux lus et écrits que les mots irréguliers qui conduiront à des erreurs de régularisation. En revanche, les mots fréquents ne devraient pas être mieux traités que les mots rares ou les pseudo-mots, tout comme les pseudo-mots analogues par rapport à ceux qui ne sont pas proches d'un mot de la langue. Des mots réguliers et irréguliers de différents niveaux de fréquence ont donc été utilisés. Certains mots réguliers n'étaient composés que de graphèmes simples (exemple : *sable*), d'autres incluaient un digraphe (exemples : *route*, *ruche*). Les mots réguliers ont été appariés en difficultés orthographiques à des pseudo-mots (exemple : *table-lople* ou *riche-soche*) qui étaient ou non analogues à des mots de la langue. Les pseudo-mots analogues ont été construits en modifiant la lettre initiale des mots fréquents (exemples : *porte/lorte* ; *poche/soche*).

Les résultats concernant les effets de régularité, de fréquence, de lexicalité et d'analogie (Sprenger-Charolles *et al.*, 1998) montrent qu'au milieu du CP, les performances en lecture et écriture sont très fortement affectées par la régularité mais pas par la fréquence, ni par la lexicalité. En effet, les mots réguliers sont mieux lus et écrits que les mots irréguliers. En revanche, la différence entre les mots fréquents et rares n'est pas significative. A la même époque, les performances pour les mots réguliers ne sont jamais supérieures à celles observées pour des pseudo-mots appariés en orthographe. En fait, il n'existe pas de différence entre ces deux catégories d'items en lecture alors que les scores pour les pseudo-mots sont supérieurs à ceux relevés pour les mots en écriture. Un résultat contradictoire (pour les auteurs) a toutefois été relevé en lecture dès le milieu du CP : l'effet de l'analogie, bien que faible (de l'ordre de 5 %), étant significatif. Selon les auteurs, cet effet s'explique probablement par une facilitation *via* le lexique oral – et non *via* le lexique orthographique –, les pseudo-mots analogues (exemples : *mable*, *lorte*, *soche*) étant construits à partir de mots fréquents (*table*, *porte*, *poche*) dont les codes articulatoires sont facilement récupérables, ce qui n'est pas le cas des pseudo-mots non analogues (exemples : *mirpe*, *lople*, *sulche*). Cette interprétation est renforcée par le fait que, lors

de la même session, pour les mêmes enfants et pour les mêmes items, cet effet n'était pas significatif en écriture.

Des changements quantitatifs mais également qualitatifs ont cependant été relevés en quelques mois. En effet, à la fin du CP, les principaux effets témoignant d'un recours à la procédure lexicale deviennent significatifs, en particulier, l'effet de la fréquence et celui de la lexicalité, tout au moins en lecture. Néanmoins, à la même époque, les signes de mise en œuvre de la procédure sublexicale sont toujours présents, en particulier l'effet de la régularité qui, de surcroît, est plus marqué que lors de la session précédente. Il en est de même pour la proportion des erreurs de régularisation. Ces erreurs ne représentaient que 5 % du nombre total d'erreurs en milieu de première année en lecture (CP) et 35 % en écriture. En fin de première année, 27 % des erreurs sont des régularisations en lecture et 64 % en écriture. Ces résultats indiquent que les enfants utilisent principalement la procédure sublexicale au début de l'apprentissage de la lecture et de l'écriture, le poids de cette procédure augmentant dans le temps alors qu'apparaissent progressivement quelques manifestations de traitements lexicaux.

Comme pour la lecture, les observations sur l'écriture ont été poursuivies jusqu'à la fin du CM1 (Sprenger-Charolles *et al.*, 2003) en tenant compte de l'évolution des performances pour les mots réguliers et irréguliers ainsi que pour les pseudo-mots non analogues. Les scores relevés pour les mots réguliers en lecture doublent entre les deux sessions de première année, passant de 43 à 89 % de réponses correctes. Il est possible de rendre compte de ce phénomène par le fait que ces items bénéficient à la fois de la régularité et de la fréquence d'exposition. La régularité seule ne suffit pas, comme l'indique la faible amélioration relevée pour les pseudo-mots (25 %), pourtant aussi bien lus que les mots réguliers en milieu de première année. La fréquence d'exposition n'est pas non plus suffisante, en témoigne l'évolution des mots irréguliers (33 %). Ces résultats permettent de comprendre pourquoi dans des langues qui possèdent une orthographe transparente, les progrès des enfants sont très rapides. Le tableau est totalement différent en écriture. En effet, les mots réguliers ne sont jamais mieux écrits que les pseudo-mots. Des résultats voisins ont été observés dans une autre étude francophone (Alegria & Morais, 1996).

La différence entre lecture et écriture peut s'expliquer par l'asymétrie des relations graphème-phonème (lecture) et phonème-graphème (écriture) en français, les premières étant plus prédictibles que les secondes (Peereman & Content, 1999). En effet, pour écrire un mot, même régulier, il ne suffit pas de connaître les relations phonème-graphème mais il faut également en maîtriser l'orthographe canonique. Par exemple, alors que les graphèmes [o], [au] et [eau] se lisent toujours /o/, il faut connaître la norme orthographique correspondante pour

écrire correctement *landau*, *bateau* ou *domino*. Cette contrainte ne s'applique pas pour l'écriture (ou la lecture) de pseudo-mots.

La réussite de l'apprentissage de la lecture dépendrait donc de la force des associations qui se créent entre unités graphiques et phonologiques, en fonction de la langue. Les enfants parviennent plus rapidement à apprendre à lire lorsque les relations entre unités graphiques et phonologiques sont transparentes (i.e., consistantes). Les nombreuses données théoriques et expérimentales développées précédemment ont permis de montrer que l'apprentissage dans une orthographe régulière était plus rapide que celui dans une orthographe moins régulière. Pour mieux comprendre l'apprentissage de la lecture, il faut donc tenir compte à la fois de la structure statistique de ce qui est appris, c'est-à-dire des régularités spécifiques d'une langue donnée et de l'environnement spécifique d'apprentissage. Les régularités renvoient aux unités orthophonologiques fréquentes que l'enfant peut extraire et catégoriser de façon implicite (Ecalte & Magnan, 2002a). Les connaissances implicites développées sur ces régularités faciliteraient la reconnaissance de mots et de non mots écrits.

La notion de transparence ou consistance permet aussi de rendre compte des différences dans les procédures de lecture relevées en fonction du degré d'opacité de l'orthographe. Ainsi, quand l'orthographe est relativement opaque, comme en anglais ou en français, les apprentis lecteurs pourraient s'appuyer davantage sur les informations lexicales de type orthographique et morphologique plutôt que sur des informations strictement phonologiques. En revanche, lorsque l'orthographe est relativement transparente, les informations phonologiques seraient activées quasi systématiquement. Ces phénomènes liés à la consistance ont mis en évidence les liens importants entretenus entre phonologie et orthographe. En effet, ce chapitre traitait de phonologie et graphophonologie et visait, en partie, à mettre en évidence l'effet de la consistance graphophonologique sur la reconnaissance de mots écrits. La notion d'orthographe a été constamment utilisée, ce qui montre bien que les connaissances phonologiques seules ne peuvent rendre compte des mécanismes d'acquisition des habiletés en lecture. Les codes orthographiques et phonologiques s'influencent mutuellement comme le montre l'importance des régularités ortho- ou graphophonologiques. Les systèmes d'écriture témoignent également de régularités orthographiques sans lien direct avec la phonologie ou non motivées phonologiquement. Ces régularités peuvent renvoyer à certaines spécificités d'un système linguistique donné comme les processus de gémination (i.e., le doublement de certaines consonnes) ou encore à la fréquence de certains patrons orthographiques tels que les bigrammes. Outre la sensibilité aux régularités graphophonologiques, les enfants pourraient ainsi être également sensibles aux régularités

orthographiques lexicales non apprises explicitement (Cassar & Treiman, 1997). L'objectif du chapitre suivant est de montrer l'influence de ces connaissances orthographiques, en particulier implicites sous forme de régularités, sur le traitement des mots et non mots écrits.

CHAPITRE 3 – LECTURE ET ORTHOGRAPHE

Dans le chapitre précédent, le rôle des régularités graphophonologiques dans la reconnaissance de mots et de non mots a été mis en évidence. Le chapitre 2, consacré en partie à la notion de consistance graphophonémique, a notamment montré que la lecture n'était pas uniquement influencée par des connaissances d'ordre phonologique mais que des aspects orthographiques pouvaient aussi entrer en ligne de compte.

Outre la question de la consistance des relations graphème-phonème et de la sensibilité précoce (et implicite) des enfants à ce type de régularités, les connaissances strictement orthographiques doivent donc être prises en compte dans l'apprentissage de la lecture. Wagner, Torgesen et Rashotte (1994) ou encore Elbro, Borstrom et Peterson (1998) ont souligné que certaines capacités préexistantes comme celles orthographiques pouvaient parfois contribuer à la fois aux performances phonologiques et aux performances en lecture. Ces capacités orthographiques renvoient à certaines spécificités de la langue écrite ou à des conventions orthographiques, telles que les connaissances relatives aux suites de lettres, auxquelles les enfants pourraient être sensibles précocement. Par exemple, une des spécificités du français écrit, partiellement partagée avec l'anglais, réside dans le fait que les consonnes peuvent être gémées (i.e., doublées). Plusieurs études qui seront développées dans ce chapitre ont montré que très tôt, les enfants témoignaient de connaissances implicites sur ce processus de gémation alors qu'aucune règle le concernant n'a été apprise. Certains patrons orthographiques (par exemple les bigrammes, i.e., des paires de lettres) sont aussi plus fréquents que d'autres dans la langue française. Une sensibilité à la fréquence de ces patrons pourrait faciliter la reconnaissance de mots et non mots écrits chez l'enfant dès le début de l'apprentissage de la lecture.

L'objectif de ce chapitre est de montrer que les enfants fortement exposés à l'écrit sont capables d'élaborer implicitement une sensibilité aux unités orthographiques composant les mots de la langue écrite et que cette sensibilité peut avoir un impact sur la reconnaissance de mots écrits. Au préalable, une présentation critique des principaux modèles d'acquisition du lexique orthographique en lecture sera effectuée afin de démontrer que ces modélisations ne prennent en compte les connaissances orthographiques que de façon tardive dans le décours de l'apprentissage.

1- Acquisition de la lecture et de l'orthographe

Plusieurs auteurs ont proposé des modèles dans lesquels l'acquisition de l'orthographe procède par stades successifs caractérisés par la mise en jeu de processus cognitifs qualitativement différents. Aujourd'hui, la notion de stade est difficile à maintenir (*cf.* Chapitre 1 et également Alegria, Leybaert, & Mousty, 1994 ; Martinet, Bosse, Valdois, & Tainturier, 1999 ; Snowling, 1994), mais il reste nécessaire de présenter ces modèles qui constituent des cadres de référence pour la plupart des recherches. Le modèle de Frith (1985) est le plus cité dans le domaine : il sera brièvement présenté dans la mesure où son fonctionnement a déjà fait l'objet d'un développement dans le chapitre 1. D'autres modèles d'acquisition en stades comme ceux de Gentry (1982), Henderson (1985) et de Ehri (1987) vont être plus largement décrits.

1.1- Modèles en stades de l'acquisition de l'orthographe

Gentry (1982) a développé un modèle dans lequel le développement de l'orthographe débute alors que l'enfant ne possède pas encore de connaissances sur les correspondances lettre-son. Au départ, les orthographe correspondent à des chaînes aléatoires de lettres pouvant comporter des nombres ou des marques ressemblant à des lettres. Ensuite, l'enfant commence à comprendre que les lettres sont utilisées pour représenter des sons. Toutefois, les orthographe débutent par des représentations phonémiques incomplètes des mots. Cette seconde étape est dénommée par Gentry « étape semi-phonétique ». Dès que les enfants peuvent symboliser dans leurs orthographe la structure sonore complète des mots, ils ont atteint l' « étape phonémique ». Tous les traits sonores de surface des mots sont désormais représentés dans les orthographe. La quatrième étape, « étape transitionnelle », commence lorsque les enfants deviennent plus conscients des orthographe conventionnelles. Les enfants utilisent alors des informations visuelles et morphologiques dans leurs orthographe plutôt que de s'attacher à leur seule sonorité. La dernière étape est appelée « étape correcte ». La connaissance du système orthographique et des règles de base est alors solide. Le scripteur continuera d'accroître sa maîtrise des formes peu communes et irrégulières. Des erreurs continueront à se produire mais la majorité des orthographe seront correctes.

Des étapes similaires dans l'acquisition de l'orthographe ont été esquissées par Henderson (1985). Selon lui, le développement de l'orthographe passe par cinq étapes successives permettant à l'enfant de passer du simple griffonnage à l'écriture par le nom des lettres. La première étape, « écriture prélettrée », débute pratiquement dès que l'enfant commence à faire des marques au crayon. L' « écriture » de l'enfant durant cette étape ressemble

plus à des griffonnages qu'à une véritable écriture : l'enfant ne comprend pas encore que l'écrit représente la parole. Quand les enfants commencent vraiment à comprendre que l'écriture est un moyen de communiquer la parole et que les lettres symbolisent les sons dans les mots, ils entrent alors dans la seconde étape du développement orthographique appelée « écriture par le nom des lettres ». Peu à peu, l'enfant a conscience des sons et des lettres utilisées pour les représenter. Il a également appris le nom des lettres de l'alphabet. Comme les enfants ne connaissent pas encore toutes les conventions du système écrit, leurs orthographes comportent des erreurs. Henderson affirme que les enfants utilisent une stratégie phonémique, ou fondée sur les sons, pour orthographier les mots qu'ils ne connaissent pas tout au long des deux premières étapes et jusqu'à la troisième, appelée « étape des patrons intra-mots ». Lorsque les enfants atteignent cette étape, ils ont commencé à mémoriser des mots au fil de leurs lectures. Selon Henderson, lorsque les mots sont en mémoire, ils sont généralement orthographiés correctement. Il faut alors attendre que beaucoup de mots soient mémorisés pour qu'ils influencent l'orthographe des mots nouveaux des enfants. Dès que l'écriture des enfants commence à se modifier en raison de cette influence, ils entrent dans l'étape des patrons intra-mots. Les connaissances que les enfants possèdent sur les mots en mémoire les aident à concevoir que l'écriture ne met pas toujours en correspondance des sons avec des lettres. Les enfants commencent à utiliser des patrons de lettres fréquents correspondant à des séquences de sons. L'utilisation de consonnes doubles pour transcrire une voyelle courte caractérise la quatrième étape dénommée « étape de jonction syllabique ». Les enfants qui progressent rapidement dans l'apprentissage de l'orthographe peuvent atteindre ce niveau vers le milieu des années de l'école élémentaire ; les autres enfants atteignent cette étape plus tardivement. Les enfants apprennent désormais que la présence *versus* l'absence d'une consonne double marque la différence entre les premières voyelles des mots, comme par exemple, dans *little* et *title*. Un autre apprentissage lors de cette étape amène la compréhension qu'il faut (ou non) doubler les consonnes lors de l'adjonction de suffixes. La dernière étape est dénommée « principes dérivationnels ». Cette étape peut débiter à partir des derniers niveaux de l'école élémentaire pour les enfants qui progressent rapidement ; elle débute plus tardivement pour les autres et se poursuit tout au long de la vie du scripteur. Lors de cette étape, les relations orthographiques entre les mots en termes de racines, origines et sens, sont explorées plus amplement.

Dans de tels modèles, l'acquisition de l'orthographe renvoie aux habiletés développées par le scripteur et donc au développement de l'écriture. Or, parler d'acquisition de l'orthographe-écriture est ambigu selon Ehri (1997) car le terme « orthographe » utilisé seul brouille la

distinction entre lecture et orthographe. En effet, parler d'orthographe revient bien souvent à l'action d'orthographier un mot en le produisant par écrit. Pourtant, le mot « orthographe » peut être également utilisé pour désigner le produit écrit, l'orthographe du mot, qui consiste en une séquence de lettres. L'orthographe des mots est ainsi la cible des comportements orthographiques et de lecture. De ce fait, lecture et orthographe ne sont pas en soit deux comportements différents. Les individus lisent l'orthographe des mots, produisent par écrit l'orthographe des mots et lisent enfin l'orthographe des mots produits. Lecture et orthographe entretiennent alors des relations d'interdépendance que plusieurs modèles dont celui de Frith (1985) et de Ehri (1987) ont mis en évidence.

Le modèle de Frith (1985) traite des relations étroites qui existent entre lecture et orthographe. Ces deux habiletés se construiraient selon trois stades dominés chacun par une stratégie différente. Lors du stade logographique, l'enfant identifie un mot grâce à des indices visuels saillants ; il ne tient alors aucunement compte de l'ordre des lettres et des facteurs phonologiques. Lors du stade alphabétique, l'enfant, ayant appris l'alphabet, est en mesure de comprendre que l'écriture transcrit la parole. Cette phase se caractérise par l'apprentissage de la correspondance graphème-phonème. Le système alphabétique étant arbitraire, le passage à cette étape résulte d'un apprentissage explicite (i.e., enseignement scolaire). A ce stade, l'enfant est particulièrement sensible aux aspects phonologiques des mots, et cette étape doit le conduire à lire et écrire tous les mots réguliers de la langue. Enfin, la dernière étape nommée « stade orthographique » correspond à un fonctionnement plus expert de la lecture. En écriture, ce serait à ce stade que des mots nouveaux pourraient être écrits par analogie avec des mots connus. L'enfant est alors capable d'analyser les mots auxquels il est confronté, de façon automatique, sans avoir recours à la médiation phonologique. Dans un tel modèle, la lecture et l'orthographe entretiennent des relations dynamiques : Frith (1985) émet notamment l'hypothèse que l'acquisition de ces deux habiletés se ferait en décalage, l'une étant le stimulateur de l'autre. En effet, la lecture logographique se développe avec l'écriture logographique, puis l'écriture alphabétique permet le développement de la lecture alphabétique puis orthographique qui elle-même conditionne une écriture orthographiquement correcte. Ce modèle dans son versant orthographique a fait l'objet de plusieurs critiques. Par exemple, selon Ellis (1997) le caractère logographique de l'écriture dans les premières phases reste peu probable. Le traitement logographique en tant que phase d'acquisition a été également remis en cause par Lecocq (1992) : selon lui, il s'agirait plutôt d'un mode de traitement plutôt que d'une véritable étape.

Ehri (1987), quant à elle, a proposé un modèle de développement des connaissances alphabétiques dans lequel elle distingue quatre niveaux. Ces niveaux font référence aux

connaissances possédées par les élèves sur le système alphabétique car elles constituent la capacité-clé qui distingue les niveaux et sous-tend le développement. Au niveau le plus précoce, désigné comme « préalphabétique » ou « logographique », les enfants connaissent peu de choses sur le système alphabétique notamment sur le fait qu'il représente la parole. Le mot est alors reconnu grâce à un indice visuel. En effet, lorsque les enfants se souviennent de la façon de lire des mots, ils stockent des traits visuels saillants. Cela rend compte de la façon dont ils lisent des labels familiers et des signes dans l'environnement. A ce niveau, la seule stratégie pour lire des mots non familiers est la devinette. En écriture, les enfants peuvent produire des gribouillages qui peuvent parfois approcher l'écriture cursive mais qui manquent de détails sur les lettres. Comme les capacités de lecture et d'écriture alphabétiques des enfants sont limitées, voire non existantes, leurs réalisations ont, à ce niveau, probablement peu d'impact sur les progrès du niveau suivant. Dès que les enfants apprennent les noms ou les sons des lettres de l'alphabet, ils sont capables de lire et d'écrire des mots au niveau de développement suivant, dénommé « alphabétique partiel » ou « semi-phonétique ». A ce niveau, les connaissances sur le système alphabétique des élèves sont rudimentaires et incomplètes. Elles incluent les sons trouvés dans les noms des lettres, parmi lesquelles les voyelles longues. Toutefois, les autres consonnes et voyelles sont bien moins connues. Pour se rappeler la façon de lire des mots déjà lus auparavant, ils utilisent leurs connaissances sur les noms de lettres ou de sons, afin d'établir des connexions entre des lettres saillantes, vues dans les mots et des sons détectés dans les prononciations. A ce niveau, les seuls moyens dont disposent les enfants pour lire des mots consistent à se les rappeler ou à les deviner à partir d'indices contextuels. Ils ne sont pas encore capables d'inventer les prononciations de mots non familiers en décodant les lettres car ils manquent de connaissances suffisantes sur le système graphophonologique, en particulier sur les voyelles. L'orthographe correcte des mots est très difficile à mémoriser pour les enfants de ce niveau car ils manquent de connaissances suffisantes sur le système alphabétique. Pour orthographier, les élèves détectent et orthographient certains sons des mots, principalement les consonnes et les voyelles les plus saillantes, dont les noms sont perçus dans les mots. Au niveau suivant, le niveau « alphabétique complet », les élèves doivent être capables de segmenter les mots en phonèmes. Ils ont également besoin de connaître les unités conventionnelles graphème-phonème, en particulier la façon dont les voyelles sont symbolisées par des lettres. Grâce à ces connaissances, ils peuvent saisir l'avantage du système alphabétique pour lire et orthographier des mots afin de se rappeler comment les lire lorsque les lettres sont compatibles avec leurs connaissances graphème-phonème. Les enfants opèrent alors principalement en appliquant leurs connaissances sur les relations graphème-phonème en lecture et les relations phonème-graphème en écriture. A ce

niveau, les enfants comprennent le système de base qui sous-tend et rend compte de la présence de plusieurs lettres en orthographe conventionnelle. Ils possèdent également des mots stockés en mémoire, avec suffisamment de détails sur les lettres qui rendent possible la lecture des mots nouveaux par analogie avec des mots familiers. Au dernier niveau, « alphabétique consolidé », lorsque les enfants pratiquent la lecture et l'écriture de mots conventionnels, ils apprennent des informations sur la structure d'unités plus larges dans les mots qui consistent en séquences de lettres revenant au sein de différents mots. Ces unités peuvent mettre en jeu l'orthographe de syllabes, ou de parties de syllabes, ou d'affixes qui apparaissent au début ou à la fin des mots. Au niveau précédent, de telles unités étaient analysées en unités graphophoniques plus petites alors qu'à ce niveau elles sont consolidées en unités plus larges. De plus, les enfants peuvent intérioriser d'autres types de régularités, par exemple les règles de doublement des lettres. Ce niveau alphabétique consolidé est encore appelé morphémique. Le fait d'opérer sur des groupes de lettres facilite le décodage et l'invention d'orthographe de mots plurisyllabiques plus longs. En outre, les enfants parviennent à former des connexions entre lettres et sons afin de retenir l'orthographe de mots spécifiques en mémoire. En résumé, dans un tel modèle, il apparaît que c'est à partir des expériences en lecture que l'apprenti lecteur stocke des associations graphème-phonème et développe des représentations orthographiques spécifiques des mots. Ehri (1997) appelle ce processus « apprentissage de mots visuels ».

En résumé, les modèles « en stades » supposent qu'à partir du moment où les enfants ont compris le principe alphabétique, ils utiliseraient leur connaissance des correspondances phonèmes-graphèmes et/ou leur connaissance des noms des lettres pour représenter les sons dans leurs écritures. Les informations orthographiques (i.e., la probabilité d'une transcription plutôt que d'une autre, en fonction du contexte) et morphologiques ne seraient utilisées que plus tard. Ce dernier point fait l'objet d'une des principales critiques adressées à ce type de modélisation. D'autres ont été formulées et sont synthétisées dans le point suivant.

1.2- Critiques de ces modèles

Les théories classiques en stades du développement de l'orthographe décrites précédemment postulent que les jeunes orthographes se reposent essentiellement sur le traitement phonologique au début de l'apprentissage de la lecture. Elles considèrent également que le développement orthographique (au cours du stade orthographique) repose sur l'activation de connaissances lexicales spécifiques (Ehri, 1986, 1992, 1997 ; Frith, 1985 ; Gentry, 1982 mais aussi Marsh, Friedman, Welch, & Desberg, 1980 ; Seymour, 1997 ; Seymour & Evans, 1993). Le stade orthographique serait alors établi seulement après l'acquisition de la connaissance des

correspondances graphèmes-phonèmes, le traitement phonologique permettant un début d'élaboration du lexique orthographique. Ainsi, Frith (1985) et Ehri (1986, 1997) considèrent qu'une pratique importante de la lecture au moyen de la procédure alphabétique permet au lecteur d'établir des représentations internes de la forme des mots. Frith (1985), notamment, soutient que les représentations lexicales doivent être suffisamment précises pour être transférées de la lecture à l'orthographe (donc à l'écrit); c'est pourquoi le stade orthographique se développerait plus tardivement en orthographe.

Même dans un modèle plus interactif comme celui à double fondation de Seymour (1997), une bonne connaissance alphabétique apparaît nécessaire pour le développement des représentations lexicales orthographiques. En effet, Seymour (1997) fait l'hypothèse que les représentations logographiques (visuelles) doivent être redéfinies, *via* le processus alphabétique, pour devenir orthographiques. Toutes ces théories décrivent ainsi le développement comme une succession de stades qualitativement différents et considèrent que le stade orthographique ne se développe en écriture qu'après une expérience importante en lecture. Elles se centrent donc sur le versant explicite de l'apprentissage de l'orthographe en considérant cet apprentissage comme tardif et conditionné par l'apprentissage préalable des correspondances graphophonémiques.

Pourtant, d'autres auteurs (Lennox & Siegel, 1994 ; Mousty & Alegria, 1996 ; Rittle-Johnson & Siegler, 1999) considèrent que les capacités de traitement orthographique et phonologique se développeraient en réalité de façon continue dès le début de l'apprentissage de la lecture. Par exemple, pour Rittle-Johnson et Siegler (1999), les orthographes débutants utilisent de multiples processus, incluant les connaissances lexicales, pour orthographier les mots. De même, Share (1995, 1999) dans son modèle d'auto-apprentissage précoce de la lecture développe l'idée selon laquelle même les très jeunes enfants sont susceptibles de posséder des représentations lexicales des mots dans la mesure où ils y ont été exposés de façon répétée. Ce modèle postule notamment que les représentations orthographiques se développent essentiellement en fonction du recodage phonologique et de la familiarité aux mots. Cette idée d'interaction entre les processus phonologiques et orthographiques ainsi que l'influence de l'exposition répétée et précoce sur le développement de connaissances lexicales ont déjà été mises en exergue dans les modèles de Colé *et al.* (1999) et de Gombert *et al.* (1997). D'autres modèles connexionnistes (Ans, Carbonnel, & Valdois, 1998 ; Harm & Seidenberg, 1999 ; Plaut, McClelland, Seidenberg, & Patterson, 1996 ; Seidenberg & McClelland, 1989) considèrent aussi, contrairement aux modèles développementaux en stades, que l'acquisition de la connaissance sur le mot peut être considérée en développement dès que l'enfant a été exposé aux formes écrites des mots simultanément à leurs formes orales.

Globalement, deux conceptions alternatives du développement orthographique semblent apparaître. D'une part, la modélisation en stades postule un développement des connaissances lexicales orthographiques tardif et conditionné par l'apprentissage explicite des correspondances graphèmes-phonèmes. D'autre part, les théories interactives prennent davantage en compte les relations entre les différentes sources d'informations (phonologique, lexicale, orthographique mais aussi morphologique) qui peuvent influencer tant l'orthographe que la lecture. Surtout, ces dernières théories considèrent que, sous l'effet de la fréquence d'exposition à l'écrit et *via* la capacité à utiliser des analogies, l'orthographe lexicale se développerait dès le début de l'apprentissage de la lecture. Le point suivant vise à mettre en évidence de façon expérimentale l'idée selon laquelle les enfants utilisent très tôt les informations orthographiques, liées à la fréquence et aux analogies, auxquelles ils sont exposés et se servent de ces connaissances (implicites) pour orthographier des mots.

2- Rôle de la fréquence et des analogies en orthographe

De nombreuses recherches expérimentales ont été menées afin de déterminer le moment où l'enfant commence à utiliser ses connaissances lexicales dans le développement de l'orthographe. Pour ce faire, l'étude des effets de régularité et de lexicalité a souvent été réalisée pour estimer la dépendance respective au traitement phonologique ou orthographique. Par exemple, Sprenger-Charolles *et al.* (2003) montrent que les enfants de CP de langue française orthographient de façon moins précise les mots irréguliers que les mots et pseudo-mots réguliers. De plus, les mots réguliers ne sont jamais orthographiés de façon plus précise que les pseudo-mots, même si ces derniers sont mieux orthographiés au milieu du CP. Ces données suggèrent que les orthographes novices ne se fondent pas sur les connaissances lexicales lorsqu'ils orthographient en dictée.

Il existe deux autres effets classiques d'utilisation de la procédure orthographique : l'effet de fréquence (les mots fréquents sont mieux orthographiés que les mots rares) et l'effet d'analogie (l'utilisation de l'orthographe d'un mot voisin). Cependant, de tels effets n'ont pas été rapportés chez les très jeunes orthographes. Ainsi, bien que des effets de fréquence soient constatés chez les enfants plus âgés, Mousty et Alegria (1996) ne trouvent pas ces effets aux étapes précoces de l'acquisition de l'orthographe suggérant que les jeunes enfants n'utilisent pas de procédure lexicale pour orthographier les mots. En revanche, Sprenger-Charolles *et al.* (1998) montrent un effet de fréquence chez des enfants de CP après dix mois d'apprentissage de la lecture mais pas après seulement quatre mois. Les divergences entre ces deux résultats peuvent provenir d'une absence de contrôle strict de la familiarité des mots dans l'expérience de Mousty

et Alegria (1996). En effet, dans la plupart des études, la fréquence est estimée en référence à des tables élaborées à partir de textes qui sont essentiellement à portée de lecteurs compétents. Ces tables ne sont donc pas représentatives de la fréquence des mots auxquels les enfants sont exposés au tout début de l'acquisition de la lecture. De plus, l'absence d'effet de fréquence ne peut pas être interprétée de façon unilatérale comme une preuve à l'encontre de l'utilisation d'une procédure s'appuyant sur l'orthographe lexicale lorsque les mots dictés sont consistants. En effet, les mots consistants peuvent tout à fait être orthographiés avec une égale précision à partir d'un traitement phonologique ou lexical, de telle sorte que les performances orthographiques devraient être similaires sur les mots consistants à la fois de basse et de haute fréquence, indépendamment de la connaissance lexicale de l'enfant. De fait, Sprenger-Charolles *et al.* (1998) rapportent que l'effet de fréquence observé à la fin du CP est beaucoup plus important pour les mots irréguliers que pour les réguliers.

Face à ces constats, Martinet, Valdois et Fayol (2004) ont mené deux expériences dont l'objectif principal était de mettre en évidence l'utilisation précoce de connaissances orthographiques chez l'apprenti lecteur et orthographeur. Dans la première expérience, il s'agissait de montrer un effet de fréquence en orthographe dès le début de l'apprentissage de la lecture. Avec cet objectif, la fréquence des mots a cette fois-ci, et contrairement à l'étude de Mousty et Alegria (1996), été strictement contrôlée à travers l'analyse des ouvrages de lecture des participants et une épreuve de dictée a été réalisée uniquement avec des mots irréguliers. Deux groupes d'enfants de CP ont été évalués à deux moments différents de l'année : en janvier, après seulement trois mois d'apprentissage de la lecture, et six mois après en juin. La présence d'un effet de fréquence seulement trois mois après le début de l'apprentissage de la lecture pourrait être interprétée comme attestant de l'utilisation précoce de procédures lexicales. Pour tester la présence d'un tel effet, les auteurs ont utilisé deux listes de mots irréguliers fréquents et peu fréquents extraits du livre de lecture des enfants de CP. La fréquence a été mesurée *via* le nombre d'occurrences des mots dans l'ouvrage. Concernant le facteur régularité des items, un mot était considéré comme irrégulier s'il contenait un graphème qui n'était pas associé de façon régulière à un phonème dans le système français. La fréquence des correspondances graphèmes-phonèmes a également été calculée à partir d'une analyse du livre de lecture des enfants. A chacune des deux sessions (janvier et juin), une liste de mots irréguliers dont la moitié étaient fréquents et l'autre moitié peu fréquents était dictée aux enfants. Ces mots contenaient un ou deux graphèmes peu familiers et les items fréquents et peu fréquents étaient appariés sur le nombre et le type de graphèmes peu familiers (*trois, bois, loup, sirop*). Ces graphèmes cibles non familiers n'étaient jamais des graphèmes associés fréquemment aux phonèmes

correspondants dans le livre de lecture. Aucun mot sélectionné n'avait, dans l'ouvrage de lecture, de voisin phonologique (Peereman & Content, 1997) incluant le graphème cible ; ainsi, une orthographe correcte ne devrait résulter que de l'activation des représentations lexicales des mots expérimentaux. Les résultats de cette première expérience indiquent que les mots fréquents étaient orthographiés avec moins d'erreurs que les mots peu fréquents aux deux sessions de janvier et de juin. L'effet de fréquence s'avérait donc significatif et ce, de façon précoce.

La seconde expérience de Martinet *et al.* (2004) visait à déterminer si des orthographes novices pouvaient utiliser des analogies pour orthographier les nouveaux mots. La plupart des études dans ce domaine utilisant une tâche d'amorçage lexical¹ avait trouvé un effet solide de l'analogie chez les enfants plus âgés mais pas chez les novices (Campbell, 1985 ; Gombert *et al.*, 1997 ; Marsh *et al.*, 1980 ; Nation & Hulme, 1996, 1998). En fait, l'utilisation très précoce de l'analogie chez des lecteurs débutants a été observée seulement dans le cas où est fournie aux enfants une base d'analogies qui constitue des mots indices (Deavers & Brown, 1997 ; Goswami, 1988b). Ces dernières données suggèrent que les jeunes enfants n'utilisent pas leurs connaissances lexicales pour orthographier des mots nouveaux sauf lorsque des mots indices et un enseignement explicite sont fournis afin de les aider. Cependant, les études utilisant l'amorçage n'indiquent pas si les mots amorces étaient familiers des participants. À l'inverse, en utilisant des mots amorces familiers, des effets précoces d'analogie apparaissent. Par exemple, Bosse, Valdois et Tainturier (2003) montrent des effets solides de l'analogie obtenus dès le CE1 dans une tâche de dictée utilisant des pseudo-mots : ces effets peuvent être obtenus dès le CP lorsque l'orthographe de mots référence – à partir desquels les pseudo-mots cibles étaient créés – est enseignée aux enfants dans les mois précédant la tâche expérimentale. De tels effets précoces de l'analogie sont justement confirmés dans la deuxième expérience de Martinet *et al.* (2004) dans laquelle une tâche de dictée ne comprenant que des pseudo-mots était proposée aux enfants de CP. Ces pseudo-mots étaient de trois types : des pseudo-mots expérimentaux avec des voisins phonologiques, des pseudo-mots contrôles et des pseudo-mots distracteurs. Les pseudo-mots voisins phonologiques étaient dérivés de mots sources de haute fréquence se terminant par une correspondance graphophonémique non familière (/diRo/ voisin de /siRo/ pour *siróp*). Les

¹ L'amorçage est une technique qui consiste en la présentation d'un mot ou d'un pseudo-mot amorce qui précède celle d'un mot cible sur lequel une tâche particulière à effectuer est demandée. Les « amorce et cible » partagent ou non des propriétés communes. Lorsqu'elles sont orthographiquement reliées (exemple : *chat/char*), le temps de reconnaissance de la cible est généralement réduit par rapport à celui obtenu dans une condition contrôle dans laquelle cible et amorce n'entretiennent aucune relation (exemple : *soleil/table*). Les tâches les plus fréquemment utilisées en technique d'amorçage sont la tâche de décision lexicale (déterminer si l'item constitue ou pas un mot de sa langue) ou de dénomination (lire à haute voix l'item présenté).

voisins ont été construits par substitution d'un seul phonème de la syllabe initiale du mot source, de telle sorte que le pseudo-mot voisin n'avait pas de voisin proche autre que le mot source. L'utilisation de graphèmes non familiers dans l'orthographe des pseudo-mots pourrait mettre en évidence un effet de l'analogie en orthographe. Par exemple, le pseudo-mot /teby/ devrait être orthographié *tébut* par analogie avec le mot source *début* au lieu de *tébu* (la correspondance la plus fréquente en français de /y/ est [u]). Dans la condition contrôle, le pseudo-mot n'avait pas de voisin lexical mais se terminait par le même phonème cible que le pseudo-mot voisin phonologique. Le pseudo-mot contrôle était utilisé pour vérifier que les phonèmes cibles en finale n'étaient pas spontanément orthographiés en utilisant le graphème peu familier. Par exemple, /liko/ était utilisé pour montrer que le graphème cible [op] n'était pas utilisé de façon spontanée pour orthographier les pseudo-mots se terminant par /o/ (l'association la plus fréquente est [eau] pour /o/ en français, pour cette position finale), indépendamment du voisinage lexical. Enfin, les distracteurs n'avaient pas de voisins phonologiques lexicaux et ne se terminaient par aucun phonème cible en final. Comme pour la première expérience, les pseudo-mots étaient dictés aux enfants aux deux sessions de janvier et de juin.

Selon les auteurs, la manifestation d'un effet d'analogie dans cette tâche où aucun mot indice ou amorce lexicale n'était utilisé pourrait être interprétée comme la démonstration que les enfants s'appuient bien sur un traitement orthographique au tout début de l'acquisition de l'orthographe. Les résultats obtenus montrent que les enfants sont plus enclins à utiliser les graphèmes cibles dans les pseudo-mots voisins que dans les pseudo-mots contrôles. Cet effet d'analogie était significatif en janvier comme en juin mais le poids de cet effet était plus important en juin qu'en janvier en raison d'une utilisation plus faible du graphème cible non familier pour orthographier les pseudo-mots contrôles en juin.

Ces deux expériences démontrent que des effets de fréquence pour orthographier les mots et d'analogie pour orthographier les pseudo-mots caractérisent les performances d'orthographiers novices après seulement trois mois d'apprentissage formel de la lecture. Les auteurs soulignent que c'est la première fois que de tels résultats sont obtenus chez des enfants aussi jeunes. Pour la première fois également, des expériences ont pu être réalisées en se basant sur un contrôle strict des mots présentés aux enfants et sur une estimation de la fréquence des mots à partir d'une analyse des livres de lecture des enfants. Martinet *et al.* (2004) en concluent que l'incapacité des études, déjà effectuées sur le même sujet, à démontrer une précocité des effets lexicaux en orthographe était probablement due au fait que les connaissances lexicales des enfants n'étaient pas strictement contrôlées. Les données obtenues par Martinet *et al.* (2004)

suggèrent que les enfants établissent des traces orthographiques des mots auxquels ils sont exposés de façon précoce et utilisent cette connaissance pour orthographier des mots nouveaux.

En outre, les résultats de la seconde expérience montrent que les enfants utilisent les graphèmes cibles non familiers moins souvent en juin qu'en janvier. Ces données indiquent que les associations graphèmes-phonèmes simultanées extraites des mots auxquels les enfants sont exposés ne diffèrent pas en fonction de la fréquence après seulement trois mois d'apprentissage de la lecture. Ainsi, les différentes associations simultanées peuvent être utilisées pour orthographier n'importe quel type de pseudo-mot et de graphèmes peu familiers apparaissant dans les pseudo-mots contrôles. Cependant, les différentes associations deviennent plus contrastées à mesure que l'exposition aux mots écrits augmente. Les correspondances graphèmes-phonèmes les plus fréquentes sont utilisées préférentiellement, ce qui explique que les graphèmes peu familiers apparaissent moins souvent dans l'orthographe des pseudo-mots contrôles en juin, après neuf mois d'apprentissage de la lecture. Ces résultats suggèrent que connaissances lexicales et connaissances générales sur les correspondances graphèmes-phonèmes seraient acquises simultanément.

Globalement, cette recherche montre qu'il est peu approprié de situer l'acquisition de l'orthographe dans les dernières étapes de l'apprentissage de la lecture, une fois que les capacités alphabétiques sont bien installées. En effet, les enfants manifestent en orthographe des connaissances lexicales précoces après seulement trois mois d'apprentissage de la lecture. Leurs connaissances non lexicales ne sont pas encore bien installées à ce moment mais continuent de se développer avec l'exposition aux mots écrits. De tels résultats sont consistants avec les modèles interactifs où connaissance lexicale orthographique et capacité alphabétique s'influencent mutuellement et se développent simultanément au tout début de l'apprentissage de la lecture. Ainsi, dans une perspective implicite, il est possible d'envisager que les connaissances orthographiques développées précocement par l'enfant contribuent à l'élaboration des connaissances graphophonologiques et du code alphabétique. En effet, comme en témoignent les effets d'analogie dans l'étude de Martinet *et al.* (2004) et comme l'avait déjà montré Goswami (1999), les procédures analogiques sont utilisées précocement et ces procédures pourraient relever au tout début de l'apprentissage d'un processus de stockage et de catégorisation des unités ortho-phonologiques fréquentes à partir desquelles se construit le code alphabétique (Ecalte & Magnan, 2000).

L'étude de Martinet *et al.* (2004) a montré de façon probante que les enfants possédaient des connaissances orthographiques précoces. Cependant, dans une telle étude, les aspects

orthographiques étudiés apparaissent fortement liés aux aspects phonologiques dans la mesure où le matériel utilisé s'appuie aussi sur la notion de fréquence des correspondances graphèmes-phonèmes. D'un point de vue strictement orthographique, plusieurs études ont montré que les enfants étaient également sensibles à certaines régularités orthographiques non motivées phonologiquement comme la gémiation ou encore à la fréquence de certains patrons orthographiques spécifiques d'une langue donnée. Ces régularités orthographiques pourraient être extraites par l'enfant de façon implicite *via* l'exposition fréquente à l'écrit. Ces connaissances orthographiques implicites sous-tendraient l'organisation des séquences de lettres. Très tôt, certaines séquences de lettres seraient ainsi écartées : la connaissance des régularités orthographiques, en restreignant les séquences admissibles de lettres pour former un mot, participerait ainsi à l'organisation du lexique orthographique.

L'objet du point suivant est de montrer, grâce aux travaux dans le domaine de l'apprentissage implicite appliqué à l'acquisition de l'orthographe, que les enfants fortement exposés à l'écrit élaborent implicitement une sensibilité aux unités orthographiques composant les mots de la langue écrite. Il s'agira en outre de poser l'hypothèse selon laquelle la connaissance implicite de certains patrons orthographiques fréquents peut faciliter la reconnaissance de mots et de non mots écrits et donc la lecture.

3- Implicite versus explicite en orthographe

Le rôle de l'apprentissage implicite dans le développement orthographique fait l'objet d'un intérêt récent de la recherche. Ainsi, les recherches en langue française qui montrent que, dès les premières années de scolarisation, l'enfant est sensible aux fréquences des suites de lettres de la langue écrite, se multiplient. Ces connaissances orthographiques s'élaboreraient par apprentissage implicite sans que l'apprentissage des relations graphèmes-phonèmes (explicite) soit impliqué (*cf.* Bonin, Pacton, & Fayol, 2001 ; Fayol, Pacton, & Totereau, 1998 ; Pacton & Fayol, 2000 ; Pacton, Fayol, & Perruchet, 1999, 2002 ; Pacton, Perruchet, Fayol, & Cleeremans, 2001). Les points suivants visent à démontrer, grâce à plusieurs études développementales en anglais et en français, cette sensibilité précoce aux régularités orthographiques et leur extraction de façon implicite par le jeune enfant.

3.1- Notion de régularités orthographiques

Une des premières recherches dont l'objectif était de montrer la sensibilité précoce des enfants à certaines régularités de la langue écrite a été réalisée par Treiman (1993) en anglais. Le but de cet auteur était de mettre en évidence la sensibilité des enfants aux conventions orthographiques

simples. Dans cette étude, de jeunes enfants de niveau CP étaient incités à écrire sans émettre de corrections sur leur orthographe. L'analyse de ces productions spontanées montre que les enfants tendent à produire des séquences de lettres compatibles avec les régularités orthographiques anglaises. Face à ces résultats, Treiman (1993) a voulu aller plus loin en étudiant expérimentalement les connaissances orthographiques implicites. Elle a proposé à des enfants de niveau CP et CE1 un test de contraintes orthographiques composé de 16 paires de pseudo-mots. Les enfants devaient choisir dans la paire l'item qui ressemblait le plus à un mot. Dans chacune de ces paires, seul un item était conforme aux patrons orthographiques de la langue anglaise (exemple : *nuck* vs. *ckun*). Les résultats indiquent, que dès le niveau CP, les enfants choisissent significativement plus l'item conforme aux régularités orthographiques. Ainsi, selon Treiman (1993), les connaissances orthographiques émergent bien plus précocement que ne le supposent les modèles développementaux de l'orthographe.

Dans la même perspective, Cassar et Treiman (1997) ont tenté de déterminer quelles étaient les connaissances que possédaient les enfants sur les consonnes et les voyelles doublées, c'est-à-dire sur les processus de gémination propres à la langue anglaise. Les résultats obtenus auprès d'enfants de niveau CP confirment ceux de Treiman (1993) sur l'émergence précoce de connaissances orthographiques. En effet, les enfants scolarisés en CP sont habituellement considérés comme étant au niveau de l'étape alphabétique dans laquelle ils se basent sur les relations phonétiques des mots pour les orthographier. Or, Cassar et Treiman (1997) démontrent que ces enfants ont également des connaissances sur le système orthographique anglais puisqu'ils choisissent significativement davantage les items corrects ou réguliers (ceux contenant des consonnes ou voyelles fréquemment doublées dans la langue) que les items contenant des irrégularités (i.e., avec des consonnes ou voyelles jamais doublées dans la langue). Les études de Treiman (1993) et de Cassar et Treiman (1997) tendent ainsi à montrer que les enfants ont appris à quoi ressemble un mot bien plus précocement que ne le suggéraient les modèles d'acquisition de l'orthographe comme ceux de Ehri (1987), Frith (1985) ou Henderson (1985).

En français, la sensibilité précoce des enfants aux régularités orthographiques de la langue a également fait l'objet de plusieurs recherches. En effet, une des spécificités du français écrit, partiellement partagée avec l'anglais, réside dans le fait que les consonnes peuvent être géminées sans entraîner de différences de prononciation, comme [p] ou [b] dans *apporte* ou *abbé*. Les consonnes ne sont doublées que dans la partie médiane des mots et certaines le sont plus fréquemment que d'autres ([pp] par rapport à [bb]). En revanche, les voyelles, en français, ne sont jamais géminées, exception faite de quelques mots empruntés à d'autres langues (*tweed*, *foot*, *continuum*). Dans certains cas toutefois, la gémination des consonnes a une conséquence

sur le plan phonologique. Ainsi, les consonnes doublées modifient la prononciation de la voyelle [e] qui devient alors /ɛ/ (*dette* /dɛt/ *versus* *petit* /p(ə)ti/). De même [ss] correspond à /s/, alors que le [s] simple intervocalique se prononce /z/. La gémiation peut donc renvoyer à des régularités soit graphophonémiques, soit strictement orthographiques. Le traitement de ces graphèmes particuliers a été étudié principalement par Pacton en français, en adaptant la méthode mise au point par Cassar et Treiman (1997) pour l'anglais. L'objectif de ce type d'études est d'évaluer à partir de quel moment peut émerger une certaine forme de sensibilité à des régularités orthographiques non motivées phonologiquement.

Par exemple, Pacton, Fayol et Perruchet (1999) se sont intéressés à l'emploi des doubles lettres chez les enfants scolarisés. Dans leur tâche, les enfants devaient choisir entre deux pseudo-mots celui qui correspondait le plus à un vrai mot. Les paires proposées contenaient des pseudo-mots avec un doublet consonantique fréquent *vs.* un doublet consonantique « jamais réalisé » (*ullate-ujjate*) ainsi que des pseudo-mots avec un doublet vocalique illégal *vs.* un doublet consonantique légal (*tiilos-tillos*). Ainsi, dans chaque paire présentée, un seul pseudo-mot était conforme au patron régulier du français : par exemple, dans la paire *ullate-ujjate*, le premier pseudo-mot devrait être préférentiellement choisi par les enfants car la lettre [l] est fréquemment doublée en français *versus* la lettre [u] jamais doublée. Dès le CP, les enfants considèrent, à 85%, que l'item avec une double consonne existante en français (*tillos* ou *ullate*) ressemble plus à un vrai mot que l'item avec une double voyelle n'apparaissant pas en français (*tiilos* ou *ujjate*). Très précocement, les enfants choisissent donc significativement plus les items congruents avec les régularités de la langue française.

Une autre étude de Pacton, Perruchet, Fayol et Cleeremans (2001), qui examine les dimensions strictement orthographiques du doublage des lettres, confirme les résultats de Pacton *et al.* (1999). Ils ont présenté à des élèves de différents niveaux scolaires, du CP au CM2, des paires de pseudo-mots avec des consonnes ou des voyelles géminées. Les enfants devaient choisir l'item qui ressemble le plus à un mot. Les pseudo-mots contenaient des consonnes de même fréquence qui sont rarement ([c], [d]) ou fréquemment ([s], [m]) doublées en français. Les résultats montrent que les items avec une consonne fréquemment doublée sont plus souvent choisis que les autres. Par exemple, *issote* est préféré à *iddote*. Cette forme de sensibilité à des régularités orthographiques subtiles s'observe dès le CP et augmente avec le niveau scolaire (au-dessus de 70, 80 et 90 % en CP, CE1 et CE2 respectivement). Une expérience contrôle avec les mêmes consonnes non doublées (*usate vs. udate*) a permis de vérifier que ce résultat ne s'explique pas par la fréquence des lettres, mais par la fréquence de la gémiation. Cette expérience montre que les enfants développent ainsi très tôt une sensibilité aux fréquences de

suite de lettres composant la langue écrite et cette sensibilité s'accroît avec le temps de fréquentation de la langue écrite. Dans une autre expérience, les auteurs ont examiné l'effet de la localisation de la gémation. Les lettres doubles apparaissent en position légale (entre deux voyelles) ou non (en début ou en fin de mot). Dès le CP, les enfants sélectionnent les pseudo-mots avec des doubles consonnes en position légale, même lorsque la consonne n'est jamais doublée en français. Par exemple, *lurret* est préféré à *lluret* et *jukker* à *jjuker*. En dehors d'une évolution entre le CP et le CE1, cette sensibilité n'augmente pas avec le niveau scolaire. En CP, les choix préférentiels sont de l'ordre de 82 % pour les consonnes fréquemment gémées et de 67 % pour celles qui ne le sont jamais. Pour les trois autres niveaux scolaires, les scores se situent aux alentours de 90 et de 65 %. Le point le plus important est que la même tendance se retrouve non seulement pour des consonnes fréquemment gémées, mais également pour celles qui ne le sont jamais, ce qui indique que cette forme de sensibilité serait relativement abstraite. C'est également ce que suggère l'examen de l'effet de la fréquence de la gémation pour les consonnes comparativement aux voyelles. Les enfants choisissent les items avec une consonne gémée (*bannil* et non ceux avec une voyelle double *baanil*), dans pratiquement 100 % des cas et ce, dès le CE1. Ce choix préférentiel se maintient lorsque la consonne n'apparaît jamais dans cette forme en français. Ainsi, le mot inventé *tukke* est préféré à *tuuke* dans 80 % des cas, quel que soit le niveau scolaire, excepté en CP (60 %).

Une autre série d'expériences (Pacton & Fayol, 2000) a cette fois-ci porté sur les gémations qui ont une incidence phonologique, en l'occurrence, le doublage de [s] ou de [l] et [t] après [e]. Les participants devaient entourer, parmi deux pseudo-mots, celui prononcé par l'expérimentateur. Les enfants de CP sélectionnent le plus souvent les items avec un simple [l] ou un simple [t], le choix étant au niveau du hasard en CE1. Il faut attendre le CE2 pour qu'ils choisissent préférentiellement les pseudo-mots contenant un [l] ou un [t] gémé lorsqu'ils entendent /-ɛl/ ou /-ɛt/ et, même en CM2, ce choix n'est fait que dans 70 % des cas. En revanche, le fait que le /s/ en position intervocalique s'écrive généralement [ss] semble acquis très précocement (91 % de choix correct en CP), mais pas l'usage qui stipule que le [s] double ne se rencontre jamais après une consonne. En effet, dans une tâche strictement visuelle (pas de mot prononcé par l'expérimentateur), les enfants du CP au CE2 sélectionnent dans 38 % des cas en moyenne *lubssir* et non *lubsir* comme un mot possible en français.

Les expériences de Pacton, Fayol et Perruchet (1999) et de Pacton, Perruchet, Fayol et Cleeremans (2001) indiquent que, très tôt, les enfants sont sensibles à des aspects purement orthographiques de l'écriture du français. Cette sensibilité s'observe également pour des graphies

qu'ils n'ont jamais vues (*cf.* les études sur la place des consonnes et celles sur les voyelles géminées). Ces résultats ne peuvent donc s'expliquer exclusivement par des facteurs d'exposition à l'écrit. Ils témoignent de l'intériorisation de régularités orthographiques de façon implicite, avec un transfert d'un domaine de compétence à un autre qui s'opère implicitement. Ce point est crucial pour les théories de l'apprentissage. Il faut toutefois reconnaître que certaines des représentations orthographiques ne sont acquises que tardivement et, de plus, très partiellement (*cf.* les études sur les effets du doublage de la consonne après la lettre [e], Pacton & Fayol, 2000). Dans les autres cas, sauf pour la sensibilité à la fréquence des consonnes géminées, il n'y a pratiquement pas de progrès entre 8 et 11 ans. Si une certaine forme de sensibilité orthographique précoce existe bien, le mécanisme qui permet son instauration ne semble donc pas fortement génératif.

Les divers travaux réalisés en anglais comme en français mettent en évidence que les enfants sont sensibles à certaines régularités orthographiques de la langue bien plus tôt que ne le laissent supposer les modèles en stades du développement orthographique. En effet, les dernières expériences sur la gémination indiquent une certaine forme de sensibilité précoce à des variables strictement orthographiques, ce qui est contradictoire avec les modèles classiques à étapes (Frith, 1985) pour lesquels ce type de sensibilité caractérise la dernière phase de l'acquisition de la lecture : la phase orthographique. Dès le CP, les enfants possèdent des connaissances orthographiques, qui ne leur ont pas été enseignées : l'enfant commence à extraire des connaissances implicites à partir des caractéristiques structurales du système écrit dès qu'il y est exposé et sans doute avant qu'ait débuté un apprentissage formel.

Hormis le processus de gémination, d'autres types de régularités pourraient faire l'objet d'une certaine sensibilité chez l'enfant. En effet, dans la langue française, certains patrons orthographiques comme les bigrammes (*i.e.*, paires de lettres) sont plus fréquents que d'autres selon la position dans le mot. Cette particularité ne fait évidemment l'objet d'aucun enseignement explicite : la connaissance par l'enfant de ce type de régularités strictement orthographiques pourrait donc être exclusivement implicite.

3.2- Effet de fréquence de certains patrons orthographiques : l'exemple des bigrammes

Un bigramme correspond à une suite de deux lettres. Par exemple, dans le mot *pomme*, plusieurs bigrammes peuvent être distingués : [po], [om], [mm], [me]. Ces bigrammes sont de fréquence variable, des tables existent pour certaines langues et permettent d'obtenir la fréquence des différents bigrammes pour des suites de lettres et de calculer une fréquence moyenne des

bigrammes composant les mots et pseudo-mots (*cf.* Content & Radeau, 1988 pour le Français et pour plus de détails). Par exemple, pour le mot *pomme*, les fréquences des bigrammes sont les suivantes : [po] = 1632 (sur 1 million), [om] = 554, [mm] = 224, [me] = 2178, ce qui donne une fréquence moyenne des bigrammes de 1147. En outre, cette fréquence « bigrammique » dépend de la position du bigramme dans l’item. En position initiale ou finale, un bigramme n’aura pas la même fréquence d’apparition : si [ca] en position initiale a une fréquence de 368, sa fréquence en position finale n’est que de 2. Il existe également des tables pour les fréquences de suites de trois lettres dénommées trigrammes.

Au même titre qu’une sensibilité à certaines régularités orthographiques telles que le doublement des lettres apparaît chez les enfants (*cf.* les études du point précédent), il est envisageable que très tôt une forme de connaissance implicite sur les séquences de lettres fréquentes de la langue puisse se développer. Les enfants pourraient notamment s’appuyer sur le voisinage orthographique pour extraire et « reconnaître » les séquences de deux lettres fréquentes dans leur langue. Une mesure possible de la taille du voisinage orthographique a été proposée par Coltheart, Davelaar, Jonasson et Besner (1977) : la mesure N (N en référence à *Neighbor* qui signifie *voisin* en anglais). N se définit comme étant le nombre de mots différents dans une langue donnée qui peut être créé en changeant seulement une lettre dans le mot stimulus tout en préservant la position des autres lettres. Par exemple, le mot *foin* possède comme voisins orthographiques les mots *coin*, *soin*, *loin*, *foie* et *fois*.

Le voisinage orthographique peut être défini également en regard de la fréquence du voisinage. Ainsi, certains mots possèdent un ou plusieurs voisins orthographiques plus fréquents (il s’agit de la fréquence lexicale) qu’eux mêmes, d’autres au contraire ne possèdent pas de voisins plus fréquents. Le mot *foin* possède quatre voisins orthographiques plus fréquents que lui, à savoir *fois*, *loin*, *coin* et *soin*. Il est possible de distinguer d’autres types de voisins que les voisins orthographiques comme les voisins phonologiques ainsi que les voisins phonographiques d’un mot.

Un voisin phonologique d’un mot cible est un mot de longueur identique (en nombre de phonèmes) créé par substitution d’un seul phonème. Par exemple, en anglais, le mot *rack* (*casier* en français) est un voisin orthographique du mot *race* (*course*) mais il n’est pas un voisin phonologique de *race* car sa forme phonologique diffère par plus d’un phonème. En revanche, le mot *raise* (*lever*) est un voisin phonologique de *race* mais pas un voisin orthographique de *race* car son orthographe diffère par plus d’une lettre de *race*.

Les voisins phonographiques correspondent aux mots qui sont à la fois voisins orthographiques et phonologiques d’un mot cible. Par exemple, les mots *face* (*visage* en

français), *rice* (*riz*) et *rate* (*taux*) sont des voisins phonographiques de *race*. Des distinctions plus fines peuvent être envisagées relativement au voisinage phonographique. Les voisins phonographiques peuvent différer du mot cible par la consonne initiale, la voyelle ou la consonne finale. Ainsi, un mot cible et ses voisins phonographiques peuvent partager le squelette consonantique comme *rice* pour le mot *race*, la voyelle et la consonne finale comme *face* pour *race*, la consonne initiale et la voyelle comme *rate* pour *race*.

L'extraction par l'enfant de séquences de lettres *via* les différents types de voisinage n'est qu'une hypothèse, étant donné le peu d'études sur le sujet. Malgré tout, il apparaît concevable qu'indépendamment de l'acquisition par apprentissage explicite d'unités de lecture correspondant à des unités phonologiques, l'enfant développe implicitement une sensibilité aux unités orthographiques écrites fréquentes de la langue telles que les bigrammes. Cette sensibilité pourrait faciliter la reconnaissance de mots et non mots porteurs de bigrammes fréquents dans la langue : l'enfant par l'exposition fréquente à l'écrit et par le phénomène de voisinage peut être en mesure d'extraire des mots rencontrés des séquences de lettres plus fréquentes que d'autres et réutiliser ensuite ces séquences pour reconnaître et/ou lire des mots nouveaux. La sensibilité aux bigrammes pourrait donc constituer une aide en lecture.

En outre, la position du bigramme dans l'item peut aussi avoir un impact sur la facilité à reconnaître le mot. En effet, en position initiale, un bigramme consonantique, s'il est suivi d'une voyelle, correspond à l'attaque. Or, la prise en compte de l'attaque dans le traitement du langage écrit, comme pour le langage oral, est déterminante dans la mesure où la première syllabe d'un item occupe une place privilégiée dans le traitement de l'information écrite (Segui, 1984). Il est donc possible d'envisager que le bigramme en position initiale soit plus saillant pour l'enfant et facilite davantage la reconnaissance de mots et non mots écrits que le bigramme en position finale.

L'ensemble des éléments de ce chapitre portant sur les connaissances orthographiques précoces chez l'enfant indique que le développement orthographique est complexe : dans le cadre d'une conception intégrative (Fayol *et al.*, 1998), il apparaît nécessaire de prendre en compte les divers mécanismes explicatifs du traitement orthographique, tant implicite qu'explicite. Avant l'apprentissage de la lecture, l'enfant élabore un ensemble de connaissances implicites sur la langue écrite (connaissances visuelle et phonologique, connaissances des lettres et connaissances sur les régularités orthographiques). Ces connaissances constituent des précurseurs nécessaires à l'apprentissage ultérieur de la lecture et au développement des connaissances explicites. Les connaissances explicites fournies par l'enseignement (comme les règles de correspondances graphèmes-phonèmes) sont en effet intégrées par chaque enfant sur la

base de ces connaissances implicites spontanées. Il apparaît fondamental de ne pas nier l'importance de ces connaissances, notamment orthographiques, afin de mieux comprendre les mécanismes qui sous-tendent l'apprentissage de la lecture.

D'autres processus interviennent également comme ceux de type morphologique (*cf.* Colé & Fayol, 2000 ; Treiman & Cassar, 1997). Au même titre que les aspects orthographiques, les éléments morphologiques ont souvent été délaissés ou pris en compte de façon tardive dans les modèles développementaux en stades de l'apprentissage de la lecture. De même, la grande majorité des études sur l'acquisition de la lecture s'est focalisée sur les liens entre cet apprentissage et la conscience phonologique. Par opposition, la morphologie a reçu beaucoup moins d'attention dans les études sur l'acquisition et sur les troubles de la lecture. Pourtant, dans la mesure où l'orthographe de la langue française comporte à la fois des éléments phonologiques et morphologiques, il est probable qu'une conscience explicite de ces deux niveaux caractéristiques de la structure du langage soit requise à la réussite en lecture et orthographe. Ainsi, 80 % des mots français sont plurimorphémiques (*cf. Le Robert Méthodique*, Rey-Debove, 1984) : il est donc vraisemblable que cette dimension soit pertinente dans l'apprentissage de la lecture.

A partir de ces constats, l'objectif du chapitre 4 est de montrer qu'à l'image de ce qui a été montré en ce qui concerne les liens entre phonologie et lecture et entre orthographe et lecture, une sensibilité précoce à la morphologie pourrait exister avant l'apprentissage de lecture et être accrue par cet apprentissage.

CHAPITRE 4 – LECTURE ET MORPHOLOGIE

En dépit d'une littérature abondante sur le développement métalinguistique (voir Gombert, 1992), peu d'études ont jusqu'à présent été réalisées sur la conscience morphologique et les relations entre morphologie et lecture. Il semble pourtant indéniable que les aspects morphologiques constituent des éléments pertinents dans l'acquisition de la lecture dans les toutes premières années d'apprentissage.

En effet, comme les chapitres précédents l'ont démontré, la connaissance des correspondances entre graphèmes et phonèmes ne suffit pas à une maîtrise des mots écrits. Certains mots comportent notamment des irrégularités qui les rendent difficiles à décoder par une stricte application des correspondances graphophonologiques (*second, femme, etc.*) Par ailleurs, l'orthographe comporte fréquemment des lettres muettes, en particulier à la fin des mots (*part, pars, pare, etc.*). Plusieurs types de connaissances permettent de dépasser ce type de difficultés. Les connaissances orthographiques en relation avec la sensibilité à la consistance graphophonémique peuvent constituer une aide non négligeable. Celles-ci ont été développées dans le chapitre 2. L'enfant possède aussi des connaissances implicites strictement orthographiques qui peuvent également faciliter la reconnaissance de certains mots écrits. Surtout, un autre type de connaissance est utilisable. Il s'agit de la connaissance du code graphomorphologique car, outre la possibilité d'analyser les mots en unités marquant les phonèmes, les mots sont également décomposables en unités de sens, les morphèmes.

Malgré ce constat, les aspects morphologiques ont souvent été négligés dans les modèles classiques d'apprentissage de la lecture ou pris en compte tardivement. Ainsi, ce sont surtout les modèles de la lecture experte qui rendent compte de l'importance des traitements morphologiques dans la reconnaissance de mots écrits. Cependant, de plus en plus d'études se sont récemment centrées sur les connaissances morphologiques précoces des enfants et sur l'impact de celles-ci en lecture. Ce chapitre présentera ces études développementales afin de mettre en évidence la sensibilité précoce des enfants à la morphologie. Au préalable, une définition précise de la morphologie sera fournie ainsi qu'une présentation non exhaustive des principaux modèles traitant de la morphologie chez le lecteur expert.

1- Les connaissances morphologiques

1.1- Définition de la morphologie

La morphologie consiste en l'étude de la structure morphémique des mots. Un mot peut être composé d'unités phonologiques telles que la syllabe, la rime ou le phonème mais également d'unités morphologiques appelés morphèmes qui sont les plus petites unités de sens (Corbin, 1985, 1987). La morphologie renvoie ainsi aux unités minimales de première articulation du langage que sont les morphèmes.

Il existe des mots morphologiquement simples composés d'un seul morphème (par exemple, le mot *chat*) et des mots morphologiquement complexes composés d'au moins deux morphèmes, à savoir un radical et un ou plusieurs affixes (par exemple, le mot *chaton* est composé du radical *chat-* et de l'affixe diminutif *-on* ; *chaton* désigne ainsi « un petit chat »). Les mots morphologiquement complexes appelés aussi plurimorphémiques représentent 80 % des mots du *Robert Méthodique* (Rey-Debove, 1984). La combinaison des morphèmes constitue donc un mécanisme puissant de création de mots et les morphèmes apparaissent comme des unités régulières de la langue dans la mesure où ils sont représentés fréquemment.

Différents types de morphèmes existent dans la langue. Ceux-ci vont être présentés et définis dans le point suivant.

1.2- Les différents types de morphèmes

Les morphèmes peuvent appartenir à deux catégories : les morphèmes lexicaux ou lexèmes et les morphèmes grammaticaux (Corbin, 1985, 1987). Les morphèmes lexicaux ont une fonction sémantique : ils permettent aux mots d'avoir une signification en eux-mêmes. Ce sont des noms, des adjectifs, des verbes, des adverbes. Les morphèmes grammaticaux ont quant à eux une fonction syntaxique : ils insèrent le mot dans des séries et indiquent souvent ses relations avec d'autres éléments de la phrase. Ainsi, dans *chanteur*, *chant-* est un lexème tandis que *-eur* est un morphème grammatical.

Les morphèmes peuvent se rencontrer à l'état libre tandis que d'autres, ne le pouvant pas, doivent s'adjoindre à d'autres morphèmes. Par exemple, *chant-* est un morphème libre et *-eur* est un morphème non libre. Ces morphèmes non libres sont appelés liés quand ils ne constituent pas des mots en eux-mêmes mais qu'ils s'ajoutent aux morphèmes lexicaux. Les morphèmes libres constituent les bases et les morphèmes liés les affixes. Les affixes peuvent être flexionnels ou dérivationnels. Les affixes flexionnels marquent le genre et le nombre des noms et des adjectifs (exemple : *amie/amis/petite/petits*) ou le temps et le nombre des verbes (*je rentre / nous rentrons*). Les mots qui comportent ces affixes flexionnels sont dits « fléchis ». Les affixes

dérivationnels sont les préfixes et suffixes (exemple : *défaire* ; *chocolatier*). Les mots qui subissent le processus de dérivation sont appelés « dérivés ». En morphologie dérivationnelle, le champ d'étude est vaste et complexe dans la mesure où il existe également des mots comportant des pseudo-affixes ou des pseudo-racines. Le terme de « pseudo » est utilisé pour rendre compte du fait que les formes en question sont isomorphes à des affixes ou à des racines mais apparaissent en l'occurrence comme des structures non dérivationnelles (par exemple : *discuter* pour le pseudo-affixe ; *chantier* pour la pseudo-racine, voir Babin, 1998). Les pseudo-affixés entraînent souvent des problèmes particuliers car ils ressemblent à ceux qui sont réellement affixés. Généralement, les pseudo-affixés sont traités plus lentement que les mots affixés. En effet, les formes pseudo-affixées constituant une minorité de mots (en comparaison avec les formes affixées correspondantes), ils sont généralement plus difficiles à apprendre et leur accès est plus lent (Smith, 1995).

Les morphèmes donnent des informations importantes sur le sens et la catégorie syntaxique des mots : c'est pourquoi ces unités sont essentiellement étudiées dans les recherches sur la compréhension en lecture. Ainsi, le mot *danseur* composé de la base ou radical *danse-* et de l'affixe dérivationnel *-eur* (suffixe) signifie la « personne qui danse ». Le processus de dérivation est intéressant dans la mesure où les morphèmes véhiculent du sens et peuvent être utiles lors de la compréhension de textes.

Etant donné le nombre important de mots plurimorphémiques dans la langue française, l'enfant devrait être capable relativement tôt d'analyser les mots parlés comme écrits en morphèmes. Cette conscience morphologique pourrait, selon Colé et Fayol (2000), intervenir de façon importante et précoce dans l'apprentissage de la lecture. Les auteurs donnent quatre raisons à cela :

- (1) Le vocabulaire parlé et celui des manuels scolaires comprennent des mots majoritairement plurimorphémiques ;
- (2) les lecteurs experts reconnaissent les mots plurimorphémiques écrits de façon analytique en effectuant une analyse automatique de leur structure morphologique ;
- (3) les jeunes enfants, lors des activités de lecture, développent leurs connaissances morphologiques implicites ;
- (4) l'apprenti lecteur dispose d'un certain nombre d'indices de nature à attirer plus ou moins son attention sur la structure morphologique des mots (exemple : le morphème « t » du mot *lait* fait le lien entre ce mot et ses dérivés, c'est-à-dire le lien visuel entre les différents mots d'une même famille morphologique).

Pourtant, malgré ces constats, très peu de modèles développementaux d'apprentissage de la lecture ont envisagé l'éventualité de traitements morphologiques chez le jeune lecteur. Cette quasi-absence de la morphologie dans les modèles soulève une interrogation car les mots plurimorphémiques, du fait même de leur représentation dans la langue, peuvent être fréquemment rencontrés par l'enfant pré lecteur ou lecteur débutant. L'existence d'une analyse morphologique a surtout trouvé un écho dans les modélisations de la lecture experte. Les auteurs de tels modèles postulent que les processus engagés lors de l'identification des mots seraient sensibles à leur structure morphologique interne. Si ces auteurs s'accordent sur l'existence d'une analyse morphologique par le lecteur expert, ils se distinguent dans la façon dont ils modélisent l'intervention de la morphologie dans les procédures de reconnaissance des mots écrits. Le point suivant va décrire quelques-uns de ces modèles. L'objectif est de mieux comprendre les traitements morphologiques en jeu dans la lecture. Ensuite, il s'agira de présenter un certain nombre de données expérimentales qui montre que ces traitements morphologiques, contrairement à ce que prédisent les modèles développementaux, peuvent intervenir précocement chez l'apprenti lecteur.

2- Le rôle de la morphologie en lecture

2.1- Lecture experte : les modèles de décomposition morphologique

Les modèles présentés ci-après envisagent les procédures de reconnaissance des mots écrits en se référant à des architectures à double voie ou à des architectures interactives comparables à celles envisagées dans les modèles connexionnistes.

Taft et Forster (1975, 1976 ; et Taft, 1979, 1981, 1984, 1985) ont proposé un des premiers modèles psycholinguistiques décrivant les mécanismes sous-tendant la reconnaissance des mots morphologiquement complexes. Ce modèle de « décomposition prélexicale » postule que seul le composant racine du mot morphologiquement construit possède une entrée lexicale propre. Ainsi, le mot *revenir* serait représenté dans le lexique par sa racine (*venir*) et non sous une forme intégrée (*revenir*). Néanmoins, l'entrée-racine (*venir*) posséderait toutes les informations concernant les affixes avec lesquels elle peut se combiner (*re-*, *contre-*, *pro-*...). De fait, *revenir*, *contrevenir*, *provenir*, qui appartiennent à la même famille morphologique, partageraient l'entrée-racine *venir*. Cette hypothèse sur le format de stockage des unités lexicales conduit à envisager les procédures d'accès aux mots morphologiquement construits comme étant de nature décompositionnelle.

Avant l'accès au lexique (d'où la dénomination de « décomposition prélexicale »), *revenir* serait décomposé en ses éléments constitutifs (*re-* et *venir*). En effet, la recherche de la représentation correspondante à la racine est précédée d'une étape ayant pour objet d'extraire la racine qui servira d'entrée lexicale (procédure d'*affix stripping*, i.e., « dépouillage de l'affixe », étape 1). Lorsque la représentation de cette entrée est localisée (issue de l'étape 2), un test permet de vérifier la légalité de la combinaison entre l'affixe et la racine et de valider l'identification du stimulus en tant que mot (étape 3 ; voir Taft & Forster, 1975, 1976 ; Taft, Hambly, & Kinoshita, 1986 ; Taft, 1981, 1984 ; Smith & Sterling, 1982). Si l'une de ces étapes conduit à une réponse négative, un redémarrage du système est enclenché afin de localiser l'entrée du mot sous sa forme entière. La Figure 3 ci-dessous résume ces principales étapes de traitement.

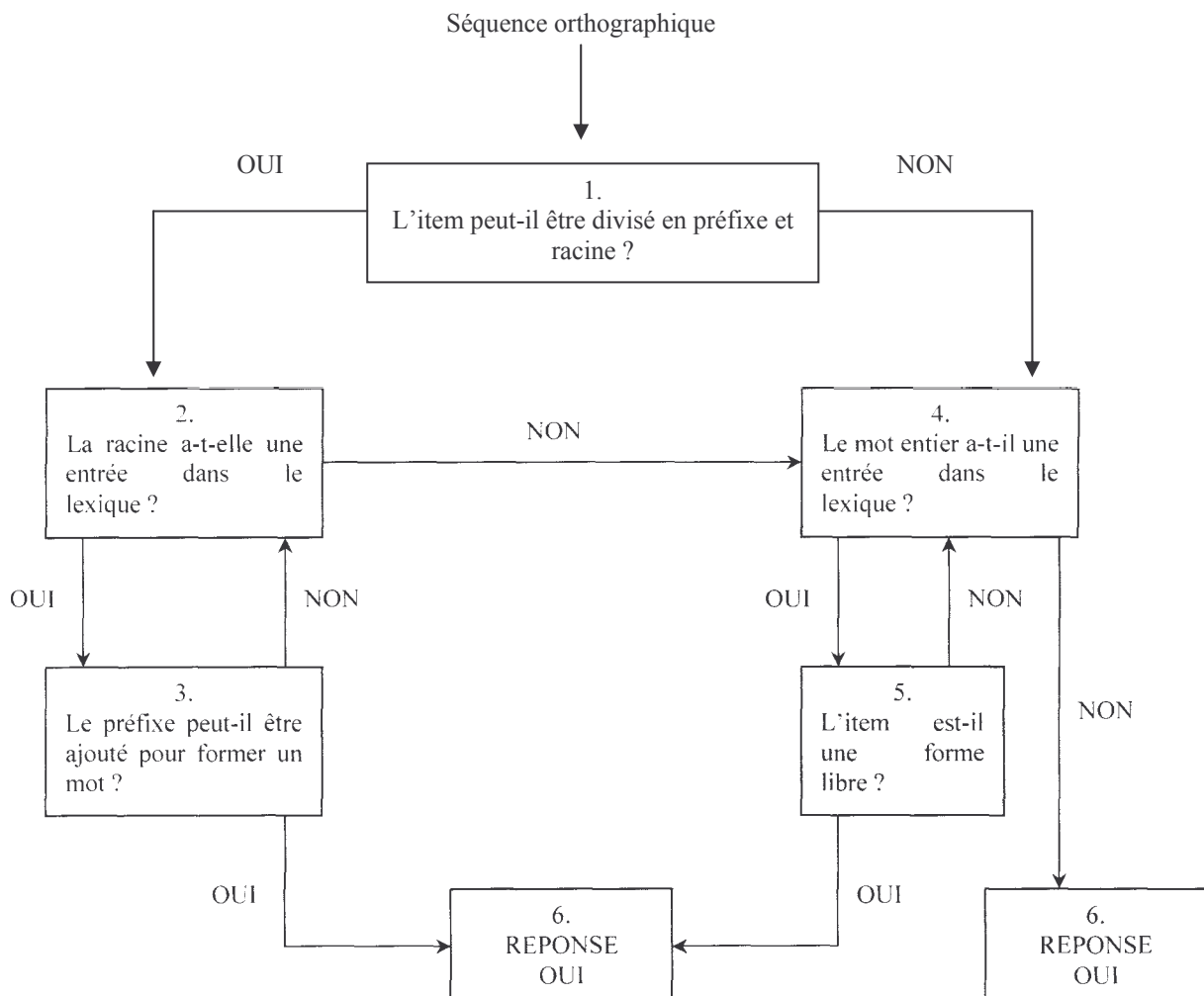


Figure 3. Modèle de reconnaissance de mots par « décomposition prélexicale » de Taft et Forster (1975)

Une des caractéristiques principales de l'hypothèse de décomposition énoncée par Taft et Forster est que la procédure de décomposition s'applique automatiquement et de façon aveugle pour toutes les formes potentiellement complexes. Par conséquent, le traitement peut, dans certains cas, conduire à un échec. Ainsi, l'application de la procédure d'*affix stripping* à un pseudo-affixe comme *décorer* conduit à isoler l'élément *-corer*. Cependant, comme cet élément n'est pas représenté dans le lexique, une recherche additionnelle, basée sur la forme entière du mot devra être entreprise. Pour le mot *démanger*, l'extraction de l'affixe conduit à isoler l'élément *manger*. *Manger* est une racine disponible dans la langue, néanmoins, l'application du test de compatibilité indique que les éléments *dé-* et *manger* ne peuvent être combinés par une règle de construction morphologique. De fait, dans ce cas aussi, la recherche d'une représentation sur la base du stimulus global est directement enclenchée.

En revanche, l'application d'une procédure décompositionnelle serait aussi efficace dans le traitement de mots comme *dégeler* (construit par une opération de dérivation en *dé-* s'appliquant sur la base libre *geler*) que dans le traitement d'items comme *dévoré* (construit par une opération de dérivation en *dé-* s'appliquant sur la base non autonome *-voré*). Les bases non autonomes seraient représentées dans le lexique de la même façon que les bases autonomes.

La contrepartie de cette procédure serait une augmentation de la complexité du traitement par rapport aux modèles stockant tous les mots dans le lexique et ne nécessitant pas une décomposition morphologique prélexicale. Ce type de modèle est donc finalement peu économique (Burani & Laudanna, 1992 ; Laudanna, Burani, & Cermele, 1994 ; Sandra, 1994). Taft et Forster (1975, 1976) pensaient que le fait de ne stocker que la base réduirait considérablement le nombre de mots dans le lexique mental et serait donc économique. Toutefois, certains auteurs comme Laudanna et Burani (1995) ont utilisé des données statistiques sur la langue afin d'évaluer la fréquence à laquelle les procédures de décomposition morphologique envisagées par Taft et Forster (1975) conduisent à un échec. Leurs résultats indiquent que le modèle de Taft et Forster (1975) est effectivement peu économique.

Le modèle AAM (*Augmented Addressed Morphology*) proposé par Caramazza, Laudanna et Romani (1988 ; voir aussi Burani & Laudanna, 1992 ; Caramazza, Miceli, Silveri, & Laudanna, 1985 ; Chialant & Caramazza, 1995 ; Laudanna, Badecker, & Caramazza, 1992 ; Laudanna & Burani, 1985 ; Laudanna *et al.*, 1994 ; Laudanna, Cermele, & Caramazza, 1997) est beaucoup plus flexible que celui de Taft et Forster et permet de réduire les désavantages liés à l'économie de stockage. Comme le montre la Figure 4 ci-après, son architecture comprend deux niveaux : le niveau des unités d'accès et le niveau des représentations lexicales. Les unités

d'accès peuvent être celles codant la forme entière des mots complexes (*revenir*) et celles se rapportant aux morphèmes qui les composent (*re-*, *venir*). Les représentations orthographiques lexicales sont, quant à elles, sous une forme décomposée. Les racines sont représentées séparément des affixes.

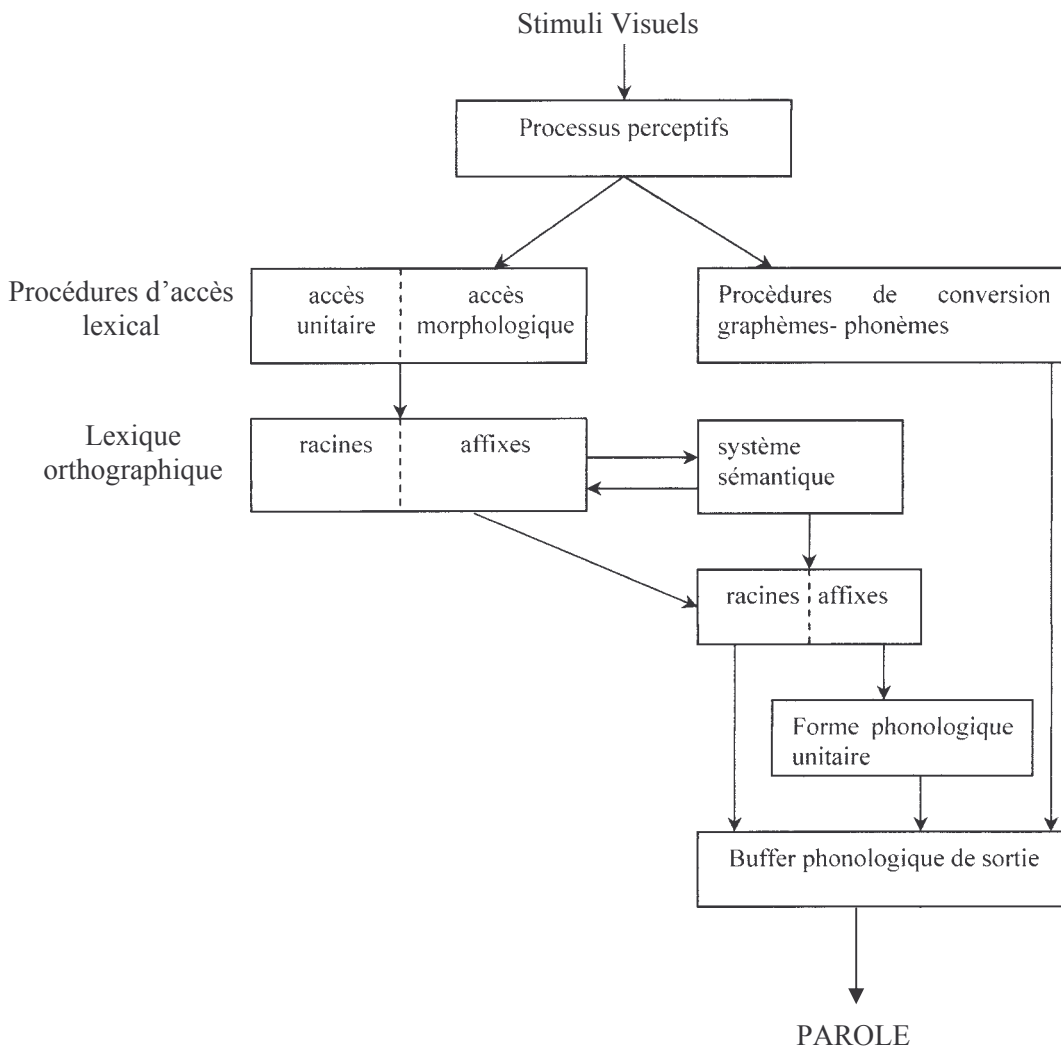


Figure 4. Modèle AAM de Laudanna et Burani (1985)

Deux procédures d'accès distinctes, mais fonctionnant en parallèle, sont postulées : l'une s'appuyant sur la forme globale des mots complexes et l'autre employant les représentations morphémiques. Chaque mot complexe active la représentation morphémique codant sa forme entière et celles associées à chaque morphème qu'il renferme. Le code d'accès atteignant le plus rapidement le seuil d'activation prédéfini est celui qui active les représentations du lexique proprement dit. Une des hypothèses centrales de ce modèle serait que, dans le cas des mots fréquents, le code d'adressage de la forme globale est le plus efficace pour accéder au lexique

d'entrées orthographiques. En revanche, pour les mots peu fréquents ou jamais rencontrés le code d'adressage de la forme décomposée atteindrait le seuil le premier. Selon ce modèle, les mots dérivés tels que *dégeler* et les mots pseudo-dérivés tels que *décorer* sont traités de manière identique jusqu'à l'accès lexical. De fait, s'ils sont de fréquence équivalente, ils devraient être reconnus aussi rapidement.

Taft en 1994 a proposé un second modèle afin de répondre aux critiques sur le problème de flexibilité de son modèle antérieur. Dans ce modèle à activation interactive, l'idée d'une décomposition morphologique non obligatoire est conservée. De fait, ce modèle est beaucoup plus flexible que celui proposé en 1975. La procédure de décomposition morphologique fait partie intégrante du processus d'activation et ne constitue donc plus une étape discrète du traitement. Les affixes et les bases seraient ici des unités d'activation indépendantes.

Taft envisage plusieurs niveaux de représentations : un niveau morphémique, un niveau lexical et un niveau conceptuel. Ces différents niveaux interagissent par un principe d'activation. Le niveau morphémique interagit avec le niveau lexical qui interagit avec le niveau conceptuel. Cependant, l'activation peut également se propager en retour du niveau sémantique aux niveaux inférieurs. Le niveau morphémique comprendrait les représentations des affixes (préfixes et suffixes) et des racines liées. Le niveau lexical comprend des représentations codant la forme globale des mots monomorphémiques ou plurimorphémiques. Les morphèmes libres sont, de fait, stockés à ce niveau.

Stocker les formes libres au niveau morphémique conduirait à une certaine redondance du système. Ainsi, la reconnaissance du mot *revenir* activerait au niveau morphémique les morphèmes *re-* et *venir*. L'activation de ce morphème se propagerait vers le niveau lexical aux représentations la contenant (*revenir*, *provenir*, ...). De la même façon, l'activation de *re-* activerait, au niveau lexical, les représentations contenant ce préfixe (*revenir*, *relire*, *redire*, ...). Taft propose, pour les mots comme *revenir* en français ou *reheat* en anglais, que seul le morphème *re-* soit représenté au niveau morphologique prélexical. Au niveau des représentations lexicales, les mots monomorphémiques (*venir* ou *heat*) seraient connectés avec les mots plurimorphémiques qui leur sont associés (*revenir* ou *reheat*).

Pour Taft, les mots comme *démanger* en français ou *relate* en anglais seraient traités différemment. En français, *démanger* ne signifie pas « arrêter d'être mangé ». De même, en anglais, *relate* ne signifie pas « encore en retard ». Les bases *manger* ou *late* seraient représentées au niveau morphémique et au niveau lexical tandis que *démanger* ou *relate* ne seraient représentés qu'au niveau lexical. En français, le préfixe *dé-* ne serait pas connecté à la

représentation lexicale *manger* et les représentations *démanger* et *manger* ne seraient pas interconnectées. Pour *relate*, le préfixe *re-* ne serait pas connecté à la représentation lexicale *late* et les représentations *relate* et *late* ne seraient pas interconnectées.

Par ailleurs, seules les représentations globales des mots comme *démanger* ou *relate*, mots sémantiquement opaques, sont directement reliées à une représentation du niveau conceptuel. Les morphèmes qui composent les mots transparents comme *revenir* ou *reheat* activent directement la signification qui leur est associée au niveau conceptuel. La signification du mot construit correspond à l'activation conjointe de la représentation sémantique de chacun de ses morphèmes (par exemple, pour *revenir* : *re-* : « encore » ; *venir* : « se déplacer »). La représentation globale *revenir* n'a pas de correspondant direct au niveau conceptuel. Enfin, Taft indique que les affixes qui auraient une signification récurrente dans plusieurs mots activeraient directement la représentation correspondante au niveau conceptuel, même si l'affixe en question n'apporte pas le sens attendu dans le mot en cours de traitement.

Les modèles AAM de Caramazza *et al.* (1988) et d'activation interactive de Taft (1994) envisagent tous les deux l'intervention des bases et des morphèmes dans les traitements mais ils se distinguent malgré tout au niveau des procédures mises en jeu. Selon Taft (1994, p. 292), « *both include a representation of the whole polymorphemic word and of its morphemes, but in the interactive-activation model, activation passes through the latter to get to the former, while the two types of representation are processed independently in the AAM model* » [les deux incluent une représentation du mot polymorphémique et de ses morphèmes, mais dans le modèle d'activation interactive, l'activation de la représentation lexicale du mot est médiatisée par l'activation préalable de ses constituants morphémiques, alors que les deux types de représentations sont traitées indépendamment dans le modèle AAM].

Ces modèles ne sont pas les seuls existants dans le domaine de la morphologie chez le lecteur expert (*cf.* le modèle hybride de Deutsh, Forst, & Forster, 1998 ; les modèles distribués de Seidenberg & Gonnerman, 2000 et de Plaut & Gonnerman, 2000 débattent également du statut et du rôle joué par la structure morphologique des mots lors de leur reconnaissance). Cependant, les travaux réalisés chez l'expert se réfèrent essentiellement à ces modèles. Les recherches ont notamment montré que le lecteur expert adulte effectue automatiquement une analyse morphologique des mots complexes qu'il lit (Alvarez, Carreiras, & Taft, 2001 ; Bertram & Hyona, 2003 ; Bertram, Schreuder, & Baayen, 2000 ; Colé *et al.*, 1997 ; Hyona, Vainio, & Laine, 2002). Ces recherches soutiennent ainsi une intervention de la morphologie dans les traitements effectués par le lecteur expert. Un des principaux effets dégagés chez l'expert est

celui de pseudo-affixation. Cet effet est d'ailleurs également retrouvé chez l'apprenti lecteur comme en témoigne les quelques travaux réalisés auprès des jeunes enfants et qui seront présentés dans le point 3.

2.2- Effets de la morphologie : la pseudo-affixation

Comme souligné précédemment, les pseudo-affixés peuvent entraîner des problèmes particuliers du fait même de leur construction : ce sont des mots qui ressemblent à de « vrais » affixés mais qui, en réalité, n'en sont pas. Taft et Forster (1975) avaient déjà remarqué que les mots pseudo-affixés posaient des difficultés de traitement car la pseudo-racine dégagée ne permettait pas un accès au lexique. Ainsi, le traitement cognitif des pseudo-affixés est plus lent que pour les affixés. Plusieurs travaux vont dans ce sens (Babin, 1996 ; Bergman, Hudson, & Eling, 1988 ; Henderson, Wallis, & Knight, 1984 ; Rubin, Becker, & Freeman, 1979 ; Smith, 1995 ; Taft, 1981 ; Pillon, 1998) et font apparaître que les temps de décision lexicale et de dénomination chez le lecteur expert sont plus longs pour les pseudo-affixés (*déguiser*) que pour les affixés (*déplier*).

Ces constats appuient l'hypothèse de décomposition morphologique. Les résultats de Pillon (1998) en français confirment d'ailleurs cette hypothèse. Cet auteur montre que des mots pseudo-préfixés tels que *dégoter* ou *recruter* sont lus plus vite que des mots monomorphémiques contrôles comme *capoter* ou *modérer*. Taft et Forster (1975) ont également mis en évidence que des pseudo-mots décomposables comme *dejuvenate* (*de-* et *juvenate* sont des morphèmes) étaient rejetés plus lentement que des pseudo-mots non décomposables du point de vue morphologique comme *deligion*.

Cependant, d'autres études réalisées dans différentes langues (en anglais : Henderson *et al.*, 1984 ; Manelis & Tharp, 1977 ; Taft & Forster, 1976 ; en français : Beauvillain, 1991 ; Beauvillain & Segui, 1992 ; Holmes & O'Regan, 1992 ; Segui & Zubizarreta, 1985 ; en hollandais : Bergman *et al.*, 1988) indiquent que la décomposition morphologique n'apparaîtrait pas dans le cas des pseudo-suffixés. Selon Taft (1985), la procédure d'*affix stripping* ne s'appliquerait pas aux items suffixés. L'analyse du mot de gauche à droite permettrait d'abord l'identification de la base puis le traitement se poursuivrait pour reconnaître le type de suffixe. Dans ce cas, la base des items suffixés serait donc traitée en priorité. Cependant, d'autres travaux indiquent une analyse morphologique spécifique lors du traitement des suffixés. Par exemple, dans une épreuve de lecture avec correction de fautes, Pavard (1983) remarque que les fautes situées dans des mots pseudo-suffixés (*chalour* pour *chaleur*) sont moins bien détectées que celles à l'intérieur de mots suffixés (*chassour* pour *chasseur*). Ces résultats indiquent un effet de pseudo-suffixation : le traitement des suffixés et pseudo-suffixés serait différent. Dans une autre

étude traitant de la sensibilité du traitement des mots affixés aux propriétés du contexte, Colé, Beauvillain, Pavard et Segui (1986) indiquent que les mots suffixés sont traités plus rapidement lorsqu'ils sont précédés d'autres mots suffixés. Cet effet du contexte n'est en revanche pas retrouvé avec des préfixés. Il y aurait donc une différence de traitement entre suffixés et préfixés. Cette différence pourrait s'expliquer par l'organisation morphologique des mots (préfixés : affixe + racine vs. suffixés : racine + affixe) et par le traitement séquentiel de gauche à droite qui « avantagerait » les items suffixés.

Ces travaux montrent que la structure morphologique intervient dans le traitement des mots écrits comme en témoignent l'observation d'effets de pseudo-affixation et le traitement différentiel des suffixés et préfixés. La morphologie tient donc une place importante dans les processus de reconnaissance de mots chez l'adulte expert. De ce fait, il semble pertinent de s'interroger sur la place accordée aux traitements morphologiques dans l'apprentissage de la lecture. En effet, depuis quelques années, de nombreuses recherches se sont focalisées sur les connaissances morphologiques de l'apprenti lecteur et sur la façon dont celles-ci pouvaient influencer la reconnaissance de mots et de non mots et ce, malgré le peu d'intérêt témoigné par la plupart des modèles d'apprentissage de la lecture pour les aspects morphologiques. Pourtant, et comme cela a été déjà souligné dans les chapitres précédents, il est tout à fait envisageable que le jeune enfant possède des connaissances morphologiques implicites sur lesquelles il pourra s'appuyer en lecture et qui permettront le développement ultérieur de connaissances explicites (Gombert, 2003).

3- Implicite versus explicite en morphologie

3.1- La morphologie dans l'apprentissage de la lecture

« L'apprentissage de la lecture nécessite avant tout la découverte explicite et consciente et la construction à partir de ce qui est déjà connu implicitement pour l'oral : la structure d'une langue et, particulièrement, la structure sonore de cette langue » (Gleitman & Rozin, 1977, p. 3). La conscience de la structure sonore et l'habileté à manipuler les sons dans les mots semblent faciliter la maîtrise du code alphabétique. C'est pour cette raison que la conscience phonologique est l'aspect du développement métalinguistique qui a été le plus souvent mis en relation avec la réussite en lecture et en orthographe et qui a reçu le plus d'attention au niveau des recherches et des modélisations. De son côté, la conscience morphologique s'appuie sur les connaissances phonologiques mais elle est également en relation avec d'autres types de connaissances chez l'enfant, comme celles sémantiques et syntaxiques. De ce fait, la conscience morphologique

fournit un indice plus général des capacités métalinguistiques que les autres domaines considérés de façon isolée.

Malgré l'importance des aspects morphologiques, leur place dans la plupart des modélisations de l'apprentissage de la lecture reste modeste (*cf.* Chapitre 1). Ainsi, si la structure morphologique prend une place importante dans la lecture experte, l'intérêt accordé à la morphologie n'est pas retrouvé dans les modèles de l'apprentissage de la lecture où il est au contraire minimisé. Dans le modèle de Frith (1985), le traitement morphologique des mots ne jouerait un rôle dans l'apprentissage de la lecture qu'après plusieurs années de lecture. Ce traitement morphologique serait alors un signe d'expertise qui n'apparaîtrait qu'au stade orthographique, au moment où l'enfant a acquis les correspondances graphèmes-phonèmes de façon explicite.

Dans la modélisation de Seymour (1997), pourtant plus interactive, la structure morphographique n'apparaît qu'à l'issue de l'achèvement de la structure orthographique. Cette structure morphographique ne participe même pas à l'élaboration initiale du lexique orthographique. Ainsi, la structure morphologique des mots ne serait utilisée dans la lecture qu'à partir du moment où l'élève maîtrise le décodage graphophonologique. Ce recours à la morphologie serait essentiellement destiné à pallier l'irrégularité de certains mots, qui ne peuvent pas être correctement lus par simple conversion des graphèmes en phonèmes, et à permettre une orthographe correcte. Pour Seymour (1997), comme pour Frith (1985), le traitement d'unités orthographiques larges comme le morphème n'interviendrait que tardivement dans les procédures d'identification des mots écrits. L'utilisation d'unités morphologiques pour reconnaître les mots écrits serait essentiellement un signe d'expertise en lecture.

Dans les modèles de lecture par analogies tels que celui développé par Gombert *et al.* (1997), les connaissances orthographiques apparaissent beaucoup plus tôt et peuvent aider l'apprenti lecteur à lire des mots jamais rencontrés auparavant. Des analogies entre les régularités orthographiques et les régularités phonologiques favoriseraient notamment la mise en place de capacités d'identification des mots écrits. Cependant, même si ce modèle a permis de faire évoluer la conception de l'apprentissage de la lecture en introduisant de façon précoce les connaissances orthographiques, il ne fait pas clairement mention de l'existence éventuelle de traitements morphologiques.

La conception interactive développée par Seidenberg et McClelland (1989) introduit elle aussi non seulement des unités phonologiques mais également des unités orthographiques et sémantiques. La reconnaissance d'un mot est issue de l'interaction entre les informations fournies par l'ensemble de ces unités. Les aspects sémantiques peuvent évidemment être

rapprochés d'informations de type morphologique mais malgré cela la dimension morphologique en tant que telle ne fait pas l'objet d'un développement conséquent dans le modèle de Seidenberg et McClelland (1989). De plus, bien que rendant compte d'un système qui évolue par la pratique, ce modèle n'est pas à proprement parler un modèle d'apprentissage.

Ainsi, la plupart des modèles d'apprentissage de la lecture délaissent le traitement morphologique des mots ou considèrent que celui-ci ne survient que tardivement. Dans ces modèles, les connaissances morphologiques du fait même de leur prise en compte tardive sont considérées comme faisant l'objet d'un apprentissage explicite. Pourtant, Gombert (2003 ; voir chapitre 1) considère que, très tôt avant même l'apprentissage de la lecture, l'enfant est capable d'extraire de son environnement linguistique des régularités de type morphologique. Ainsi, avant l'apprentissage explicite de la lecture, l'enfant posséderait une connaissance morphologique essentiellement implicite. Ce type de connaissances peut être mis en évidence dans les productions verbales spontanées et erronées des enfants. Par exemple, l'enfant âgé d'à peine trois ans produit parfois des erreurs sur des formes verbales qui révèlent bien souvent sa connaissance implicite des règles de flexion verbale (*j'ai appris*). Ces erreurs ne correspondent en fait qu'à une assimilation de règles, une extraction de régularités propres à la langue et renvoient au principe de conventionnalité ou phénomène de surgénéralisation (Corbin, 1987). Un tel phénomène se retrouve pour les processus de dérivation. Un enfant peut ainsi produire *lapinier* pour désigner l'endroit où se trouvent des lapins en ignorant le mot attesté *lapinière* car il aura entendu précédemment *pigeonnier* pour désigner l'endroit où se trouvent les pigeons. Corbin (1987) définit ce type d'activités langagières comme « des aptitudes à produire et à comprendre des mots construits non mémorisés auparavant, conformément à des règles qu'ils ont acquises non explicitement lors de leur apprentissage, et qui obéissent à des principes de construction des mots » (p. 68). Même l'adulte invente parfois des mots nouveaux en utilisant des affixes pour la signification qu'ils dégagent (par exemple, le suffixe *-ette* qui signifie souvent le mot petit dans *maisonnette*, *fillette*). Chez l'enfant, ce mécanisme de construction est inconscient, implicite car il utilise des suffixes alors même qu'il ne connaît pas la règle sous-jacente au mécanisme en question. La compétence dérivationnelle pourrait ainsi se construire dès l'enfance et les connaissances morphologiques qui en découlent seraient le reflet d'un apprentissage implicite.

Par exemple, lorsqu'un élève rencontre le mot *chaton*, il reconnaît le mot *chat* et le suffixe *-on*. En fait, lorsqu'il rencontre un mot morphologiquement construit, tous les mots connus qui lui ressemblent sont activés en mémoire. L'analyse morphologique est donc souvent inconsciente, son apprentissage est alors implicite.

Face à ces constats, il est difficile de ne pas soutenir que l'enfant possède de façon précoce des connaissances morphologiques sur les mots. Ces connaissances seraient « acquises » de façon implicite *via* la fréquentation d'écrits. Le point suivant présente un certain nombre d'études qui confirment non seulement la présence d'une sensibilité précoce aux aspects morphologiques de la langue, mais qui soutiennent également le rôle important des connaissances morphologiques dans la reconnaissance de mots écrits connus et/ou nouveaux.

3.2- Les traitements morphologiques chez l'apprenti lecteur

Dans le premier paragraphe de ce chapitre, Colé et Fayol (2000) avaient fourni quatre raisons permettant de justifier une intervention précoce de la morphologie dans l'apprentissage de la lecture en français. En langue anglaise, des raisons similaires amènent à considérer l'analyse morphologique comme une composante importante dans la réussite en lecture. Fowler et Liberman (1995) soulignent ainsi que :

- (1) Un corpus important de recherches indique que le vocabulaire anglais, en particulier le vocabulaire des ouvrages scolaires, est morphologiquement complexe (comme en français). Nagy et Anderson (1984) estiment notamment que 60 % des nouveaux mots acquis par les enfants d'âge scolaire sont morphologiquement complexes avec des structures morphologiques relativement transparentes, de telle sorte que la signification des différentes parties fournit suffisamment d'informations pour deviner la signification des mots entiers (connaître *dog* et *house* permet de deviner facilement la signification de *doghouse*) ;
- (2) les tâches de lecture permettent aux jeunes enfants de développer une connaissance de la morphologie dérivationnelle et flexionnelle qui au départ est implicite. Par exemple, Gordon (1989) a constaté chez des enfants de 5, 7 et 9 ans que la fréquence élevée du radical facilitait la reconnaissance des mots suffixés de basse fréquence dans une tâche de décision lexicale avec des mots présentés oralement. Une autre preuve de l'existence d'une morphologie productive chez les jeunes enfants est leur utilisation fréquente de dérivations pour créer des mots, comme *cooker* en référence à « celui qui cuisine » (*one who cooks*) ou encore *fix-man* signifiant *mechanic* (*mécanicien* en français) en référence à « celui qui répare » (*one who fixes*) (Bowerman, 1982 ; Clark & Berman, 1984, 1987 ; Clark & Hecht, 1982) ;
- (3) enfin, la troisième et principale raison selon laquelle l'analyse morphologique devrait faciliter la lecture et l'orthographe est que la majorité des conventions orthographiques en anglais ne peut être expliquée uniquement par une approche phonétique mais peut l'être aussi dans une perspective morphologique. En effet, l'orthographe anglaise est morphophonémique : en anglais, différentes orthographe servent souvent à désambiguïser

différents morphèmes avec des formes phonologiques identiques, comme dans *threw vs. through*.

Face à l'orthographe anglaise, l'analyse phonologique ne serait pas suffisante au lecteur débutant. Une analyse de type morphologique apparaît sans conteste comme fondamentale dans les procédures de reconnaissance de mots écrits. Le lecteur débutant doit donc posséder des connaissances morphologiques suffisamment établies afin de réaliser une telle analyse.

Concernant de telles connaissances, Gombert, Colé, Valdois, Goigoux, Mousty et Fayol (2002) ont fait le point sur les compétences morphologiques des apprentis lecteurs et des lecteurs débutants. Il apparaît que, dès l'âge de 6 ans, les enfants maîtrisent l'utilisation orale des flexions (marques de genre, de nombre, de temps). Lorsque les enfants commencent à lire et écrire, ils sont alors dans les dernières étapes de maîtrise des flexions (Berko, 1958 ; Brown, 1973 ; Clay, 1982 ; de Villiers & de Villiers, 1973). Ils apprennent également les principes dérivationnels comme le montrent les inventions spontanées de mots et leur utilisation de formes productives. (Clark & Cohen, 1984). Cependant, l'utilisation des règles dérivationnelles dépend de la complexité des suffixes et des préfixes, leur acquisition serait plus lente et plus tardive. Ainsi, les enfants en GSM et CP apprendraient d'abord les formes dérivées qui sont phonétiquement et sémantiquement transparentes (i.e., la base du mot est clairement représentée dans la forme dérivée), productive et proche de la signification (e.g., *danse* et *danseur* ; *lait* et *laitier*).

Plusieurs études ont mis en évidence la présence de telles connaissances en morphologie chez des enfants avant même l'apprentissage de la lecture. Par exemple, Colé, Marec-Breton, Royer et Gombert (2003 ; voir également Marec-Breton, 2003 dans une thèse consacrée à ce sujet) ont montré que des enfants prélecteurs de langue française pouvaient manifester des choix préférentiels à l'égard de pseudo-mots morphologiquement construits. Ces auteurs ont ainsi mis en évidence une sensibilité à la morphologie préalable à l'apprentissage de la lecture grâce à une tâche orale de préférence de mots. Pour cette tâche, des enfants âgés de 5 ans devaient choisir entre deux pseudo-mots celui ressemblant le plus à un vrai mot. Dans chaque paire, un pseudo-mot consistait en un préfixe et une base (*préfader*, *pré-* est un préfixe fréquent en français) tandis que l'autre était construit à partir des mêmes éléments phonologiques (*pradéfer*) mais n'avait pas de structure morphologique (*pra-* n'est pas un préfixe). Dès l'âge de 5 ans, les enfants prélecteurs considèrent que les préfixés ressemblent plus à des vrais mots que les autres items (67,25 % de réponses préférentielles à l'égard des items préfixés). Les auteurs en concluent qu'une sensibilité à la morphologie des mots oraux existe avant l'apprentissage de la lecture. Elle serait due à des connaissances implicites basées sur la rencontre de régularités de la langue orale.

En outre, ces auteurs ont constaté que l'apprentissage explicite de la lecture permettrait ensuite d'accroître cette sensibilité et favoriserait notamment la conscience de certaines caractéristiques morphologiques. Ils ont proposé aux mêmes enfants de 5 ans et à des lecteurs débutants (CP au CE2) une tâche explicite orale de jugement d'un intrus dans une liste de triplets (*dévoré* est un pseudo-préfixé car *dé-* ne joue pas le rôle de préfixe et constitue l'intrus parmi deux mots préfixés tels que *défaire* et *démonter*). Les résultats indiquent un choix aléatoire pour les prélecteurs, les enfants de CP et CE1 tandis que ceux de CE2 détectent l'intrus de façon significative. Non seulement les enfants témoignent de façon précoce de connaissances morphologiques implicites, mais ils possèdent également des connaissances explicites qui apparaissent plus tardivement car se développant sur la base des connaissances implicites préétablies.

L'évaluation du développement des connaissances explicites se focalise bien souvent sur la transition d'une sensibilité ou conscience intuitive (i.e., implicite) de la structure morphémique des mots à une conscience plus explicite. Pour comprendre cette transition d'une conscience implicite à explicite, il est nécessaire de se centrer sur les comportements qui signalent ce changement et sur les tâches qui fournissent la preuve d'une conscience explicite. Or, il faut bien reconnaître qu'un des problèmes majeurs est que les chercheurs dans ce domaine ont souvent différents points de vue sur les comportements signalant une conscience métalinguistique ou explicite. Malgré cela, une des façons les plus fréquentes d'évaluer la sensibilité explicite des enfants est d'utiliser des tâches avec des consignes détaillées qui nécessitent de la part des enfants une analyse du langage (i.e., manipulation ou jugement sur langage). En effet, les tâches qui nécessitent la manipulation de mots ou de phrases peuvent exploiter un certain niveau de conscience explicite, en indiquant notamment comment l'enfant analyse et utilise le langage. La tâche de jugement d'intrus utilisée par Colé *et al.* (2003) est ainsi une façon adéquate de mesurer les connaissances explicites possédées par les jeunes apprenants.

Concernant les connaissances morphologiques explicites, Rubin (1988) a réalisé une étude visant à explorer la conscience de la structure morphologique dans l'orthographe de formes fléchies chez des enfants de GSM et de CP. Les résultats obtenus indiquent que les enfants de CP avec de faibles connaissances morphologiques omettaient davantage de morphèmes fléchis en orthographe et étaient moins capables d'identifier les morphèmes de base dans les mots oraux que les GSM et les CP possédant de hauts niveaux de connaissances morphologiques implicites. Il y aurait donc bien une interaction entre la présence préalable de connaissances morphologiques implicites et l'établissement de connaissances explicites. Ces connaissances morphologiques explicites seraient peu développées à un niveau précoce comme chez les élèves

de niveau GSM selon Carlisle (1995). Néanmoins, les performances d'analyse morphologique augmentent de façon considérable de la GSM au CP. Une étude française menée par Lecocq, Casalis, Leuwers et Watteau (1996) confirme ces résultats. Des élèves de GSM au CP ont été soumis à deux épreuves. La première est une épreuve d'identification morphémique : l'enfant doit choisir parmi trois mots (exemples : *pommier*, *poire*, *pommade*) l'item qui constitue le dérivé d'une cible fournie (exemple : *pomme*). Dans la deuxième épreuve, l'expérimentateur prononce le début d'une phrase (exemple : *une petite trine est une ...*) et l'enfant doit la terminer (réponse : *trinette*). Les auteurs ont observé un progrès significatif de la GSM au CP : le taux de réussite, à la première épreuve, passe de 48 % à 61 %. En revanche, les connaissances morphologiques ne diffèrent pas significativement en CP (61 %) ou en CE1 (67 %).

Ces quelques études montrent que, de façon précoce, l'enfant possède des connaissances implicites non négligeables en morphologie flexionnelle et dérivationnelle. Selon Gombert (2002), le développement des connaissances morphologiques implicites est loin d'être achevé lorsque les élèves débutent l'apprentissage de la lecture. Même si d'importants progrès sont constatés dans les premières années de cet apprentissage (GSM-CP), le développement se poursuit bien au-delà du CM1. Les connaissances morphologiques qui se construisent sont celles que l'enfant maîtrise au cours de l'acquisition du langage parlé et qu'il utilise tous les jours de façon implicite. Peu à peu, il acquiert des connaissances explicites qui s'élaborent sur la base des connaissances implicites. Ces connaissances sont susceptibles d'être mises en relation avec la réussite en lecture. Plusieurs études ont en effet mis en évidence l'implication de la structure morphologique dans le processus d'identification de mots. Les paragraphes suivants décrivent ces études réalisées dans différentes langues et auprès de populations d'âges différents.

La majorité des recherches concernant l'impact de la morphologie sur l'apprentissage de la lecture s'est surtout intéressée à des niveaux avancés de l'apprentissage (niveau CE2 au niveau CM2). Par ailleurs, la plupart de ces études a été menée en langue anglaise (exceptées celles de Casalis & Louis-Alexandre, 2000 et de Colé *et al.*, 2003). Pour cette langue, Fowler et Liberman (1995) ont synthétisé trois principaux types de données expérimentales qui suggèrent que la connaissance morphologique contribue de façon importante et unique à la réussite en lecture, indépendamment de celle apportée par la conscience phonologique :

- (1) Les faibles lecteurs présentent des difficultés spécifiques dans l'analyse de formes morphologiquement complexes qui apparaissent notamment dans la lecture à haute voix de certains morphèmes comme les suffixes flexionnels (e.g., Henderson & Shores, 1982). De telles difficultés suggèrent une insensibilité globale à la structure interne des mots ;

(2) un certain nombre d'études expérimentales montre que les erreurs morphologiques en orthographe (écriture) ne peuvent être complètement attribuées à des difficultés au niveau des correspondances graphèmes-phonèmes, dans la mesure où les problèmes morphologiques persistent même après un apprentissage des correspondances graphèmes-phonèmes. Hanson, Shankweiler et Fisher (1983) ont notamment montré que des adolescents avec des difficultés en lecture orthographiaient des mots phonétiquement transparents comme *plastic* ou *splinter* aussi bien que des normo lecteurs du même âge. En revanche, les difficultés entre les deux groupes étaient très prononcées pour orthographier des mots nécessitant une sensibilité à la structure morphémique (comme *condemn/condemnation*) ou répondant à des conventions orthographiques spécifiques (e.g., *heroes* ou *galleries*) (voir également Fisher, Shankweiler, & Liberman, 1985 ; Templeton & Scarborough-Franks, 1985) ;

(3) les limitations morphologiques des lecteurs en difficulté sont évidentes non seulement en lecture et écriture mais également dans les tâches métalinguistiques impliquant des mots oraux. La tâche proposée par Berko (1958) en atteste. Cette tâche consiste à faire produire aux enfants des formes fléchies à partir de dessins d'animaux imaginaires appelés chimères. Le dessin d'une chimère est montré à l'enfant tandis que l'expérimentateur nomme le dessin en utilisant un pseudo-mot : « c'est un *wug* ». Puis, l'expérimentateur montre à l'enfant un second dessin avec deux exemplaires du même animal et lui dit : « là, il y a deux ». La réponse attendue de l'enfant est *wugs*. Or, il est observé depuis longtemps que les lecteurs en difficulté de tous âges sont moins capables que les bons lecteurs de produire le suffixe morphologique approprié pour ce type de mots sans signification.

Plusieurs études en langue anglaise vont ainsi dans le sens d'une implication de la structure morphologique dans les procédures de reconnaissance et de traitement de mots écrits. Pour le traitement des formes fléchies exclusivement, les données anglophones chez l'apprenti lecteur sont assez rares car la majorité des recherches a été conduite dans le domaine de l'acquisition de l'orthographe (Nunes, Bryant, & Bindman, 1998). Malgré tout, quelques recherches apportent des résultats intéressants sur les relations entre morphologie flexionnelle (et dérivationnelle) et lecture. Carlisle et Nomanbhoy (1993) montrent que les connaissances métamorphologiques flexionnelles (mais aussi dérivationnelles) sont impliquées dans la lecture de mots et ce, dès la première année. Deux tâches métamorphologiques ont été utilisées : une tâche de jugement de relation morphologique et une tâche de production de mots dérivés et fléchis en contexte de phrase. Dans la première, l'enfant devait décider si deux mots présentés oralement étaient reliés ou non (exemple : *teach/teacher* et *moth/mother*) et dans la seconde,

l'enfant devait compléter une phrase énoncée oralement en produisant le mot fléchi/dérivé qui convient, la base étant fournie préalablement (exemple : *help/Father tells me that I am a good.../helper*). Les résultats obtenus montrent que les connaissances morphologiques (évaluées par les deux tâches métamorphologiques) contribuent au niveau de lecture des enfants. Cette contribution est modeste (4 %) mais néanmoins significative.

Feldman et Andjelkovic (1992) ont également exploré les relations entre connaissances morphologiques et réussite en lecture mais chez des élèves plus âgés d'origine Serbo-croate. Ces enfants âgés de 7 à 8 ans avaient pour tâche de compléter des phrases impliquant différents types de mots morphologiquement complexes. Une épreuve de délétion de phonèmes leur était aussi proposée. Les analyses corrélationnelles indiquent une corrélation positive entre la tâche de délétion de phonèmes et celle d'utilisation des flexions. La corrélation entre l'épreuve phonémique et la tâche d'utilisation des dérivations est également positive. Bien que les auteurs n'aient pas évalué à proprement parler l'habileté en lecture, ils rappellent que les capacités métalinguistiques constituent un des facteurs les plus importants dans l'apprentissage de la lecture : notamment, la conscience d'unités linguistiques au dessus du niveau du phonème tel que le morphème contribuerait de façon significative à la capacité en lecture.

En français, Casalis et Louis-Alexandre (2000) se sont également intéressées aux connaissances morphologiques flexionnelles de façon longitudinale chez des enfants de la GSM au CE1. La morphologie flexionnelle était évaluée par une tâche de production de formes fléchies dans laquelle l'enfant devait donner la forme fléchie d'un nom, adjectif ou d'un verbe fourni oralement par l'expérimentateur. Deux types de flexions étaient exploités à la fois pour des mots et des pseudo-mots : le marquage du féminin (*un boulanger / une boulangère* et *un pomecteur / une pomectrice* pour les noms ; *bon / bonne* pour les adjectifs) et le marquage du temps des verbes (*je chante / je chantais* et *je plosse / je plossais* pour l'imparfait ; *il chante / il va chanter* et *tu blines / tu sais bliner* pour l'infinitif). Les auteurs ont réalisé une analyse de régression pas à pas afin de déterminer dans quelle mesure l'analyse morphologique effectuée en GSM pouvait expliquer les performances ultérieures en lecture en CP (mesurées par la *Bat-Elem*, Savigny, 1974) et en compréhension écrite en CE1 (évaluées par le test *ECOSSE*, Lecocq, 1996). Les données indiquent que les connaissances morphologiques flexionnelles en GSM expliquent de façon significative 5,7 % de la variance en lecture en CP et 22 % de la variance en compréhension en CE1. Les compétences précoces en morphologie dérivationnelle également mesurées n'auraient, quant à elles, qu'un impact tardif sur la lecture (8,2 % de la variance expliquée sur les performances en CE1 à *l'Alouette*, Levafrais, 1967). Ces résultats confirment

donc l'implication des connaissances morphologiques dans la lecture, mais surtout au niveau de la compréhension et tardivement.

Concernant les procédures de lecture des mots fléchis, Laxon, Rickard et Coltheart (1992) ont montré que des lecteurs de 7 à 10 ans effectuent un traitement morphologique des mots fléchis mais aussi des formes dérivées qu'ils lisent. Ils observent notamment que la lecture à haute voix de mots suffixés tels que *dancer* (en français : *danseur*) occasionne significativement moins d'erreurs que celle de mots pseudo-suffixés tels que *dinner* (dans lequel *-er* n'est pas un suffixe comme par exemple *-eur* dans le mot couleur). Un équivalent des situations avec des formes fléchies serait, en français *vivant* et *enfant*, et en anglais *wicked* et *locked*. Ces résultats sont obtenus à la fois avec des bons et moins bons lecteurs, les bons lecteurs produisant également moins d'erreurs que les moins bons. Néanmoins, quel que soit le niveau de lecture des enfants, les mots suffixés dérivés en *-er* sont lus avec plus de précision que ceux fléchis se terminant en *-ed*, alors que pour les adultes, une telle différence n'est pas observée. Surtout, les auteurs soulignent un élément important : les difficultés rencontrées par les enfants pourraient s'expliquer aussi par les règles phonologiques qui gouvernent la prononciation du suffixe *-ed* qui sont multiples et complexes, trois prononciations étant possibles en anglais. Les résultats suggèrent donc que l'analyse morphologique des mots lus se trouverait encore sous la dépendance de facteurs phonologiques.

Cependant, récemment, Feldman, Rueckl, DiLiberto, Pastizzo et Vellutino (2002) ont obtenu avec des enfants de niveau CM2 des résultats qui suggèrent un traitement morphologique des mots fléchis qui n'est plus dépendant de ces contraintes phonologiques. Les auteurs ont proposé une tâche de complétement sur une cible (exemple : T_N) qui est précédée par une amorce identique (exemple : TURN) ou une amorce reliée orthographiquement uniquement (exemple : TURNIP) ou morphologiquement (exemple : TURNED). Dans ce dernier cas, l'amorce peut également consister en un mot relié morphologiquement mais opaque d'un point de vue orthographique (exemple : SOLD et SELD avec pour contrôle orthographique SELF). Les résultats montrent un effet de facilitation morphologique dans le pourcentage de complétement de fragment par rapport à la condition orthographique et ceci que l'amorce soit transparente orthographiquement ou pas. Toutefois, l'effet de facilitation morphologique est plus important dans le cas des amorces morphologiques transparentes que dans le cas où elles ne le sont pas. Ceci suggère donc que le traitement morphologique reste encore à un âge relativement avancé sous l'influence de facteurs formels. Les études relatives au traitement des formes dérivationnelles donnent des résultats un peu différents.

Les études qui se sont intéressées à la morphologie dérivationnelle étayent l'hypothèse selon laquelle la structure morphologique joue un rôle dans les processus de reconnaissance de mots écrits. Ainsi, Carlisle (1988) a montré que des lecteurs en difficulté de langue anglaise de niveau CM1, 6^e et 4^e de collège étaient moins experts que les bons lecteurs dans la production de dérivés morphologiques et de formes de base lors d'une tâche de complétion de phrases présentées oralement (i.e., *Magician. I really enjoy.../magic*).

Cependant, à des niveaux plus précoces de l'apprentissage de la lecture, les résultats obtenus en anglais par Carlisle et Nomanbhoy (1993) et Carlisle (1995), et en français par Casalis et Louis-Alexandre (2000) apparaissent parfois contradictoires. En effet, une contribution significative des connaissances dérivationnelles dans les performances en lecture de mots est relevée en anglais, mais pas en français. Pour les enfants francophones, ce n'est qu'en CE1 que les compétences morphologiques expliquent une partie de la variance dans un test de lecture étalonné (*l'Alouette*, Lefavrais, 1967) ou dans un test de compréhension écrite (*l'ECOSSE*, Lecocq, 1996). Ces derniers résultats, également observés par Carlisle (1995) suggèrent que les connaissances métamorphologiques faciliteraient relativement précocement la lecture (au niveau CE1) et, surtout, la compréhension. Ce qui s'explique aisément compte tenu du fait que les morphèmes constitutifs des mots affixés véhiculent des informations impliquées de façon cruciale dans le calcul syntaxique et sémantique de la séquence dans laquelle ils sont insérés. En revanche, les résultats pour le CP sont contradictoires.

Les résultats obtenus en français par Marec-Breton, Gombert et Colé (2005) sont plus clairs. En effet, ces auteurs ont montré, *via* une tâche de lecture de mots et de non mots, que des lecteurs débutants lisent mieux (plus rapidement et avec moins d'erreurs) les mots préfixés (*retomber*) que les mots pseudo-préfixés (*renifler*). Les pseudo-mots préfixés (*remourir*) sont également mieux lus que ceux pseudo-préfixés (*recounir*). Ces résultats indiquent, que dès le début de l'apprentissage de la lecture, l'enfant prend en compte la structure morphologique des mots et que la reconnaissance d'une base ou un affixe au sein d'un item inconnu (pseudo-mot) facilite le traitement des items à lire. Les connaissances morphologiques contribueraient donc précocement à la lecture.

Les études anglophones conduites avec des élèves plus âgés rapportent également des résultats clairs. Ainsi, Fowler et Liberman (1995) observent, avec des enfants de niveau CE1 au CM1, des corrélations significatives entre les performances aux tâches morphologiques et celles de lecture de mots et de pseudo-mots (même après avoir contrôlé l'effet de l'âge et du vocabulaire). Plus précisément, une analyse de régression hiérarchique montre que les

performances aux tâches morphologiques expliquent entre 42 % et 34 % de la variance respectivement en lecture de mots et de pseudo-mots. D'autres recherches comme celles de Shankweiler *et al.* (1995) et de Mahony, Singson et Mann (2000) montrent que la contribution des connaissances métamorphologiques à la lecture de mots augmente du niveau CE2 à la 6^e de collège : 9 % de la variance en lecture de mots expliquée par certaines tâches métamorphologiques dans l'étude de Mahony *et al.* (2000), ce qui constitue le double du pourcentage de variance expliquée rapporté par Carlisle et Nomanbhoy (1993) pour les enfants de niveau CP. Dans le même temps, la contribution des connaissances métaphonologiques décroît mais reste néanmoins significative (entre 10 et 15 % d'après les études précitées). A ce propos, Singson, Mahony et Mann (2000) observent *via* des régressions hiérarchiques que la conscience phonologique perd sa contribution à la variance en lecture au fur et à mesure des niveaux scolaires jusqu'à une absence de contribution à partir du niveau CM1 (niveau CE2: 36 % ; niveau CM1 : 4 % ; niveau CM2 : 3,2 % et 6^e de collège : 1 %). Enfin, Carlisle (2000) observe que les connaissances morphologiques contribuent significativement à la variance en compréhension écrite au niveau CE2 (43 %), cette contribution étant encore plus importante au niveau CM2 (55 %).

Dans d'autres langues comme l'italien, les connaissances morphologiques dérivationnelles apparaissent également impliquées dans la reconnaissance des mots écrits. Burani, Marcolini et Stella (2002) montrent ainsi que dès le niveau CE2, les enfants italiens utilisent les informations morphologiques en lecture. Les auteurs ont comparé les temps de décision lexicale et de dénomination à des pseudo-mots constitués d'une base et d'un suffixe dérivationnel en combinaison illégale (exemple : *donn + ista* ; en français : *lait + ard = laitard*) par rapport à ceux obtenus pour des pseudo-mots ne contenant pas d'afixe. Ils observent des résultats identiques pour les deux types de tâches et pour tous les niveaux scolaires considérés (du niveau CE2 au niveau CM2) : les pseudo-mots « morphologiques » sont prononcés plus rapidement que les « non morphologiques ».

L'ensemble de ces résultats montre bien que les enfants utilisent des éléments morphémiques composant les mots afin de les reconnaître et en faciliter la lecture. Cependant, dans certaines orthographes très transparentes comme l'italien, il est possible que les enfants s'aident également des éléments orthographiques. La procédure morphologique pourrait alors davantage relever d'une procédure lexicale de type orthographique, notamment chez les enfants plus âgés qui possèdent déjà des connaissances orthographiques. Chez les enfants plus jeunes, en début d'apprentissage de la lecture, l'analyse morphologique pourrait plutôt s'appuyer sur des

aspects sémantiques. En français, Colé *et al.* (2003) observent des résultats qui confortent cette hypothèse.

Trois situations d'amorçage ont été proposées à des enfants de fin de CP et CE1 bons et moins bons lecteurs. La première consiste à présenter un mot cible suffixé précédé d'un mot amorce qui lui est relié morphologiquement (exemple : le mot *laitier* est précédé par sa racine *lait*) ou relié orthographiquement (exemple : *laitier* est précédé par *laitue*) ou qui n'entretient aucune relation avec lui (exemple : *pomme*). L'utilisation de la situation orthographiquement reliée permet de dissocier les effets du partage formel émanant de la relation morphologique. Cette situation permet de s'assurer que les effets potentiels de facilitation morphologique ne relèvent pas uniquement d'un partage d'une structure orthographique commune entre l'amorce et la cible. La similarité orthographique des situations orthographique et morphologique a été contrôlée avec le calcul d'un indice de similarité orthographique (Van Orden, Johnston, & Hale, 1988).

Les résultats indiquent un effet de facilitation de l'amorçage morphologique par rapport aux situations contrôle et orthographique : les temps observés pour la condition morphologique sont toujours significativement plus courts sauf chez les moins bons lecteurs de CP. La reconnaissance d'un mot suffixé dérivé est donc facilitée par la présentation préalable de sa base et cet effet de facilitation ne peut être uniquement attribué au partage de propriétés formelles qu'entretiennent base et mot dérivé.

Une deuxième expérience, utilisant le même paradigme et les mêmes conditions d'amorçage mais avec des mots dérivés dont la base subit une transformation phonologique (par exemple : *camion-camionneur*), montre que les enfants de CP sont insensibles, contrairement à ceux de CE1, à ces transformations formelles. Les enfants les plus jeunes utilisent donc fortement le lien sémantique qui unit deux membres d'une même famille morphologique.

Ces différentes données obtenues pour plusieurs langues spécifiques suggèrent que dès la fin de la première année de primaire, les bons lecteurs se servent de la structure interne des mots complexes pour les lire. Dès la fin de cette première année, l'identification des mots peut être tributaire de facteurs multiples qui ne se restreignent pas à ceux impliqués dans le décodage graphophonologique. L'analyse des mots écrits en leurs composants morphologiques assisterait donc l'accès au sens des mots lus et pourrait participer à la reconnaissance des mots écrits qui au CP n'est pas experte. Cette possibilité est d'autant plus probable que l'apprenti lecteur serait confronté à de nombreux mots complexes dans les livres scolaires (Colé & Fayol, 2000 ; Nagy & Anderson, 1984). Des traitements implicites spontanément effectués par l'enfant pourraient donc

expliquer les résultats rapportés. Ces traitements concerneraient un certain nombre de régularités graphomorphologiques, c'est-à-dire des séquences courtes de lettres fréquentes rapidement décodables telles que les préfixes et auxquelles il est possible d'associer un sens. Puis, l'apprentissage explicite de la morphologie écrite pourrait être important pour permettre au lecteur débutant de lire les mots irréguliers (ou incluant des lettres muettes) et au scripteur débutant d'orthographier correctement les mots tant que les habiletés implicites ne sont pas suffisantes. Plus tard, ces connaissances explicites seront disponibles pour l'application des règles orthographiques sophistiquées (morphosyntaxe) et pour l'activité de correction orthographique.

Ainsi, quand un enfant apprend à lire, le système cognitif ne peut ignorer ces régularités graphomorphologiques. Le schéma est ici le même que pour la phonologie et l'orthographe (*cf.* Chapitres 2 et 3) : une sensibilité précoce chez les prélecteurs (connaissances implicites), un rôle de la dimension linguistique chez le lecteur débutant, puis une conscience tardive (connaissances explicites) de cette dimension favorisée par l'apprentissage de la lecture et de l'écriture. Ces connaissances implicites et explicites de la structure morphologique pourraient participer relativement précocement à la lecture.

En résumé, les études présentées dans les différents chapitres théoriques ont mis en évidence que non seulement les connaissances graphophonologiques mais aussi celles orthographiques et morphologiques pouvaient se développer de façon précoce chez l'enfant et être utilisées pour faciliter la lecture de mots connus et de mots nouveaux. Le chapitre suivant s'appuie sur les principaux résultats de ces études pour présenter un certain nombre de prédictions concernant principalement la sensibilité précoce des enfants aux régularités graphophonologiques, orthographiques et morphologiques.

CHAPITRE 5 – PROBLEMATIQUE

La plupart des modèles développementaux d'apprentissage de la lecture ainsi que de nombreuses recherches ont mis en évidence que le principe alphabétique repose sur les correspondances graphèmes-phonèmes. Dès lors, l'influence d'autres types de traitement, notamment orthographique et morphologique, est réputée survenir tardivement, généralement en lien avec l'apprentissage de l'orthographe (Seymour, 1997). Pourtant, de nouvelles recherches se multiplient depuis quelques années sur l'importance des correspondances entre configurations orthographiques et morphèmes. Plusieurs modélisations interactives ont mis en exergue, d'une part, les interactions entre unités orthographiques et phonologiques et, d'autre part, le rôle des connaissances antérieures sur les procédures de reconnaissance de mots écrits (*cf. via* les analogies dans le modèle de lecture de Gombert *et al.*, 1997 ; voir également le modèle de Colé *et al.*, 1999).

De plus, que ce soit pour le code phonologique, orthographique et/ou morphographique, deux types de processus seraient impliqués dans l'apprentissage de la lecture. Dans plusieurs recherches et modélisations, ce sont les effets directs ou explicites de l'enseignement qui sont mis en évidence. Or, il apparaît nécessaire de tenir compte également de l'apprentissage implicite qui renvoie aux modifications du système cognitif sous l'influence de la répétition de la rencontre des patrons orthographiques fréquents et des cooccurrences entre orthographe et propriétés orales associées (Gombert, 2003). Cette dimension implicite s'élabore *via* l'exposition à l'écrit qui contribue au développement des processus orthographiques impliqués dans l'identification, la production de mots écrits et les compétences en lecture. Ce facteur expérientiel contribuerait ainsi au développement, à la précision et au stockage de représentations orthographiques et morphologiques. En effet, grâce aux contacts réguliers de l'enfant avec l'écrit (environnement familial et scolaire), l'enfant serait capable d'extraire certaines régularités visuo-orthographiques, phonologiques et morphologiques de l'écrit, qui jouent un rôle dans l'apprentissage de la lecture.

Fort de ces constats, l'objectif de cette thèse est de montrer que dès les premières étapes de l'apprentissage de la lecture, avant même que les associations entre graphèmes et phonèmes ne soient totalement installées, les enfants seraient en mesure de développer des connaissances implicites et explicites sur certaines régularités du système écrit. Ces régularités pourraient être de trois types : graphophonologiques, orthographiques et morphologiques. Dans le cas des régularités graphophonologiques, il s'agit de s'intéresser plus spécifiquement à la sensibilité des

enfants à la consistance des relations graphèmes-phonèmes. Pour l'orthographe et la morphologie, les régularités concernent respectivement la fréquence des bigrammes et la structure morphologique dérivationnelle, en particulier les phénomènes de pseudo-affixation. Il s'agit de montrer que ces régularités affectent les processus de reconnaissance de mots écrits lors de la lecture.

1- Prédications sur les aspects graphophonologiques : la consistance graphophonologique

En référence aux travaux de Lamarche (1986), mais également à ceux de Treiman (1993) et de Fayol *et al.* (1996), les enfants pourraient être sensibles à de nombreuses régularités dans les transcriptions graphèmes-phonèmes, dès le début de l'apprentissage de la lecture. Certaines transcriptions étant plus fréquentes que d'autres, elles seraient alors utilisées plus souvent, conformément aux régularités spécifiques de la langue. Ces régularités de type probabiliste ne peuvent être enseignées par des règles dans le contexte scolaire. L'enfant y serait donc sensible de façon implicite *via* la fréquentation d'écrits. Cette sensibilité pourrait se manifester par une préférence à l'égard d'items contenant des graphèmes consistants ou réguliers avant même l'apprentissage de la lecture au CP.

Ces connaissances implicites apparaissent précocement et n'arrêtent pas leur développement au début de l'apprentissage de la lecture. En revanche, avec l'enseignement explicite de la lecture et l'accroissement des connaissances implicites, les connaissances explicites émergent et se développent. L'apprentissage implicite se poursuivant tant que l'individu lit (Gombert, 2002, 2003), la sensibilité aux régularités des correspondances graphèmes-phonèmes devrait augmenter avec le niveau de lecture.

De plus, conformément aux résultats de Sprenger-Charolles, Siegel et Bonnet (1998) et de Sprenger-Charolles, Siegel, Béchenec et Serniclaes (2003) obtenus en français, les processus de lecture (de mots connus comme de mots nouveaux) pourraient être affectés par la régularité des correspondances graphèmes-phonèmes. Les items réguliers devraient être plus facilement lus que ceux irréguliers. Cet effet serait d'autant plus manifeste que le niveau de lecture augmente.

2- Prédications sur les aspects orthographiques : la fréquence des bigrammes

Les études de Cassar et Treiman (1997) en anglais ainsi que celles de Pacton, Fayol et Perruchet (1999), Pacton, Perruchet, Fayol et Cleeremans (2001) et Pacton et Fayol (2000) pour le français indiquent que les enfants possèdent de façon très précoce, dès le CP, des connaissances sur la

fréquence de certaines suites de lettres comme les lettres doubles. Ces connaissances se manifestent notamment par des choix préférentiels à l'égard de pseudo-mots contenant des lettres fréquemment doublées, en accord avec les régularités de la langue, par rapport à des pseudo-mots avec des lettres peu ou jamais doublées. Ici encore, la sensibilité à de telles régularités ne peut être attribuée à des règles apprises de façon explicite : très tôt, les connaissances orthographiques possédées par l'enfant seraient de nature implicite. Outre le processus de gémination, il existe dans la langue française des patrons orthographiques plus fréquents que d'autres. Ainsi, certains bigrammes sont plus ou moins fréquents en fonction de leur position (initiale et finale, par exemple) dans un item donné. Cette plus ou moins grande fréquence des bigrammes dans la langue est arbitraire et ne répond pas à des règles formelles. Au même titre que les enfants possèdent une sensibilité à la fréquence des lettres doublées en français, ils pourraient, avant même d'apprendre à lire, (i.e., avant l'entrée au CP) manifester une préférence pour des items contenant des bigrammes fréquents *vs.* rares. Cette préférence pourrait notamment s'appuyer sur les analogies orthographiques avec des mots connus contenant de tels bigrammes, c'est-à-dire par le biais du voisinage orthographique (Coltheart *et al.*, 1977). De plus, cette préférence pourrait être d'autant plus marquée que le bigramme fréquent est en position initiale dans l'item. En effet, comme l'a déjà souligné Segui (1984), l'attaque d'un mot de par sa position est traitée plus rapidement que la rime. Le bigramme en position initiale étant assimilable à l'attaque, son traitement serait facilité en comparaison de celui en position finale.

Conformément à ce qui a été développé précédemment sur les dimensions implicite et explicite en graphophonologie (Gombert, 2002, 2003), les connaissances implicites développées sur la fréquence des bigrammes devraient augmenter avec le niveau de lecture des enfants. Les connaissances explicites se développeraient sur la base de ces connaissances implicites : leur développement serait donc plus tardif et en lien direct avec l'apprentissage explicite de la lecture.

Si les enfants sont sensibles à la fréquence de certains bigrammes et qu'ils préfèrent les items construits avec des bigrammes fréquents, il est vraisemblable qu'ils utilisent ces connaissances en lecture dans la reconnaissance de mots connus et nouveaux. La structure orthographique pourrait ainsi être utilisée en lecture et faciliter la dénomination d'items contenant des bigrammes fréquents dans la langue au fur et à mesure que le niveau de lecture s'accroît.

3- Prédications sur les aspects morphologiques : l'effet de pseudo-affixation

Plusieurs études menées auprès de très jeunes enfants ont mis en évidence une sensibilité à la structure morphologique dérivationnelle du français (Colé *et al.*, 2003 ; Marec-Breton, 2003 ; Marec-Breton *et al.*, 2005). Une telle sensibilité a pu être mesurée *via* des tâches de plausibilité lexicale dans lesquelles l'enfant devait indiquer parmi des items affixés et pseudo-affixés celui ressemblant le plus à un vrai mot. Les enfants, de façon très précoce, préféraient les items affixés, c'est-à-dire morphologiquement construits, aux items pseudo-affixés. Cet effet prégnant de pseudo-affixation témoigne de la présence de connaissances implicites sur la structure morphologique. Les études menées jusqu'à présent sur cet effet ont surtout utilisé des tâches orales. L'objectif de cette thèse est de répliquer un tel effet à l'écrit et de vérifier ainsi la précocité des connaissances implicites morphologiques en modalité écrite. De plus, à l'étude de l'effet de pseudo-affixation sera ajoutée celle de l'effet du type d'affixe. En effet, chez l'adulte expert, Colé *et al.* (1986) avaient constaté que le traitement des suffixés était plus rapide que celui des préfixés de par la séquentialité du traitement de gauche à droite qui facilite la récupération des items suffixés. Cependant, chez l'apprenti lecteur, des résultats contradictoires ont été obtenus par Marec-Breton (2003) : la reconnaissance de préfixés serait plus facile en comparaison des suffixés, les préfixes (et par conséquent les items qui leur sont associés aussi) étant plus facilement isolables que les suffixes pour deux raisons principales. D'une part, la signification d'un suffixe est plus abstraite que celle d'un préfixe (Corbin, 1999). D'autre part, le préfixe ne modifie pas la structure phonétique ou syllabique de la base à laquelle il s'associe, contrairement au suffixe.

Il s'agit, en outre, de confirmer l'accroissement de telles connaissances implicites avec le niveau de lecture et de mettre en évidence la présence de connaissances explicites élaborées sur la base de celles implicites.

Enfin, et en accord avec les différents travaux menés en français (Colé *et al.*, 2003 ; Marec-Breton, 2003 ; Marec-Breton *et al.*, 2005), en anglais (Laxon *et al.*, 1992) ou encore en italien (Burani *et al.*, 2002), la sensibilité à la structure morphologique pourrait participer à la reconnaissance de mots écrits connus et de mots nouveaux dès le début de l'apprentissage de la lecture. La lecture d'items affixés devrait ainsi être facilitée par rapport à celle d'items pseudo-affixés.

4- Prédications sur les relations entre les connaissances graphophonologiques, orthographiques, morphologiques et la lecture

Dans les modèles développementaux (Frith, 1985), le développement des connaissances phonologiques, orthographiques et morphologiques est considéré de façon successive. Les connaissances orthographiques et morphologiques se développent tardivement et sont précédées des connaissances d'ordre phonologique. Dans les modèles interactifs comme ceux de Gombert *et al.* (1997) et de Colé *et al.* (1999), les connaissances notamment orthographiques et phonologiques sont considérées comme s'élaborant de façon interactive. Ces connaissances entretiennent alors entre elles des liens non négligeables.

Un des objectifs de cette thèse est de vérifier les liens pouvant être établis non seulement entre les divers types de connaissances – graphophonologiques, orthographiques et morphologiques – mais également entre ces connaissances et la lecture. L'hypothèse de liens entre les connaissances graphophonologiques, orthographiques, morphologiques et la lecture s'appuie sur les résultats déjà obtenus par plusieurs études expérimentales. Ainsi, concernant l'éventuelle interdépendance entre les connaissances morphologiques et phonologiques, Fowler et Liberman (1995) constatent que des enfants âgés de 7 ½ à 9 ½ ans présentant des difficultés au niveau phonologique ont également des difficultés importantes dans la détection de relations morphologiques entre items. Le même type de relations peut également être établi entre les unités phonologiques et les unités orthographiques. Ainsi, comme l'ont déjà souligné la plupart des modèles interactifs (Colé *et al.*, 1999 ; Gombert *et al.*, 1997), des connexions lexicales existent entre les représentations lexicales et phonologiques pendant le traitement des mots écrits. De tels liens ont aussi été mis en évidence par Martinet *et al.* (2004 ; voir paragraphe 2, chapitre 3). Ces auteurs considèrent que les connaissances orthographiques lexicales et les connaissances générales sur les correspondances graphèmes-phonèmes sont acquises simultanément et s'influencent mutuellement dès le début de l'apprentissage de la lecture. Ces mêmes connaissances orthographiques pourraient entretenir des liens avec les connaissances morphologiques dans la mesure où il est établi que la connaissance orthographique est une source puissante d'information concernant les relations morphologiques, en particulier celles de type dérivationnel (Fowler & Liberman, 1995). En outre, en se référant au schéma développemental de l'apprentissage de la lecture élaboré par Gombert (2003), il apparaît que les connaissances explicites s'élaborent à partir des connaissances implicites préétablies. Pour chaque aspect étudié – graphophonologique, orthographique et morphologique – des corrélations positives devraient être obtenues entre les dimensions implicites et explicites.

Concernant la lecture et ses relations spécifiques avec chaque type de connaissances, les différentes études présentées dans chacun des chapitres théoriques précédents ont mis en évidence des liens importants entre reconnaissance de mots écrits et capacités phonologiques, orthographiques et morphologiques. Au tout début de l'apprentissage de la lecture, il est vraisemblable que les compétences phonologiques soient les plus fortement reliées à l'apprentissage de la lecture (Duncan, Seymour, & Hill, 2000 ; Hulme, Mutter, & Snowling, 1998 ; Hulme *et al.*, 2002 ; Muter *et al.*, 1998). Sprenger-Charolles *et al.* (1998) relèvent ainsi en français des corrélations positives entre les réponses correctes en dénomination et la production d'erreurs phonologiques chez les lecteurs débutants. De telles corrélations devraient également être obtenues entre la lecture et les connaissances orthographiques. En effet, les relations entre lecture et orthographe sont établies depuis longtemps dans les modèles développementaux d'apprentissage (Frith, 1985 ; Ehri, 1987). De même, dans les modèles connexionnistes (Colé *et al.*, 1999 ; Gombert *et al.*, 1997 ; mais aussi Ans *et al.*, 1998 ; Harm & Seidenberg, 1999 ; Plaut *et al.*, 1996), les connaissances orthographiques contribuent aux procédures de reconnaissance de mots écrits. De fortes associations entre les capacités en lecture et les connaissances morphologiques sont aussi attendues en référence aux données obtenues par Colé *et al.* (2003) qui ont observé des corrélations entre le niveau de connaissances morphologiques d'élèves du CE2 au CM2 et leur niveau de lecture. Ce lien significatif entre les compétences morphologiques et le niveau de lecture a été obtenu dès le niveau CE1 par Carlisle et Nomanbhoy (1993), Casalis et Louis-Alexandre (2000), Fowler et Liberman (1995) ou encore Laxon *et al.* (1992). Il s'agit de vérifier ces liens entre lecture et morphologie de façon plus précoce, c'est-à-dire dès le CP.

Des analyses de régression permettent généralement de préciser les constats des analyses corrélationnelles concernant les relations entre les connaissances graphophonologiques, orthographiques, morphologiques et la lecture. Cependant, la plupart de ces analyses indique que ce sont avant tout les compétences graphophonologiques qui constituent les meilleurs prédicteurs du niveau de lecture. Par exemple, Bruck, Genesee et Caravolas (1997), dans une étude interlangues avec des enfants canadiens francophones et anglophones suivis de 6 à 7 ans, montrent qu'avant l'apprentissage de la lecture et quelle que soit la langue, les deux compétences qui contribuent le plus au niveau ultérieur en lecture sont les connaissances des lettres et la manipulation phonémique. D'autres études comme celles de Manis, Custodio et Szezulski (1993) pour l'anglais et de Sprenger-Charolles *et al.* (2003) pour le français ont mis en évidence que les compétences phonologiques précoces, telles que révélées par la lecture de pseudo-mots, capturent la majeure partie de la variance unique dans les compétences ultérieures en identification des mots écrits. Dans l'étude de Sprenger-Charolles *et al.* (2003), les auteurs se

sont intéressés au poids respectif des capacités phonologiques et orthographiques précoces (milieu et fin du CP) sur les habiletés orthographiques ultérieures (fin du CP et fin du CM1). Les analyses réalisées suggèrent que les habiletés phonologiques précoces possèdent un fort pouvoir génératif puisque ce sont les seules qui prédisent les compétences ultérieures orthographiques, sur le long terme. Toutefois, dans le court et le moyen terme, les habiletés orthographiques sont prédites à la fois par les compétences phonologiques et orthographiques antérieures. Les changements relevés dans le temps pour les prédicteurs pourraient suggérer l'existence de sauts qualitatifs au cours de l'apprentissage. Ainsi, au tout début de l'apprentissage de la lecture, la lecture serait prédite essentiellement par les compétences phonologiques. Les enfants utiliseraient d'abord les correspondances graphophonémiques mais ensuite des connexions pourraient progressivement se créer entre unités orthographiques et phonologiques permettant aux connaissances orthographiques de contribuer également aux compétences en lecture. Concernant la morphologie, elle contribuerait également aux capacités de lecture mais comme pour l'orthographe cette contribution serait plus tardive que celle de la phonologie. Ainsi, Carlisle et Nomanbhoy (1993) ont montré que la conscience phonologique prédisait 33 % de la variance en lecture chez les enfants de niveau CP tandis que la morphologie rendait compte de seulement 4 % de variance supplémentaire. Ainsi, la morphologie jouerait un rôle distinct de la phonologie en lecture mais ce rôle bien que significatif serait minime. Des résultats similaires ont été obtenus par Carlisle (1995) dans son étude longitudinale réalisée chez des enfants du niveau GSM au niveau CE1. Ses données montrent que les performances des enfants de niveau CP sur des tâches de conscience phonologique et morphologique rendent compte d'une part significative de la variance au niveau CE1 pour des tests d'analyse de mots et de compréhension en lecture : la tâche de production morphologique est le meilleur prédicteur de la compréhension en lecture tandis que la tâche phonologique est le meilleur prédicteur de l'analyse de mots. Casalis et Louis-Alexandre (2000) confortent également ces résultats : les connaissances morphologiques sont impliquées dans la lecture (surtout au niveau de la compréhension) mais plus tardivement que les connaissances graphophonologiques.

A partir de ces différents résultats issus d'analyses corrélationnelles et de régression, cette thèse tente de démontrer que les connaissances graphophonologiques, orthographiques et morphologiques interagissent entre elles et avec la lecture. Il s'agit en parallèle de mettre en évidence que les connaissances graphophonologiques contribuent précocement aux compétences en lecture, les connaissances orthographiques et morphologiques y participant également mais de façon plus tardive. Il faut ajouter que pour la morphologie, cette thèse s'est focalisée sur les aspects dérivationnels. Ce choix s'appuie sur les constats de Carlisle (1995). Selon elle, les

dérivations pourraient être de meilleurs prédicteurs de la réussite en lecture que les flexions dans la mesure où elles impliquent la compréhension des relations phonologiques, des rôles syntaxiques et des relations sémantiques.

5- Présentation des études expérimentales

L'objectif général de cette thèse est d'étudier le développement des connaissances graphophonologiques, orthographiques et morphologiques implicites et explicites et le lien que ce développement pourrait avoir avec l'apprentissage de la lecture. Il s'agit, en outre, de vérifier si très tôt, les apprentis lecteurs sont en mesure d'utiliser activement les informations graphophonologiques, orthographiques et morphologiques pour lire des mots écrits connus et nouveaux.

Les participants sont des enfants tout venants scolarisés de la GSM au CM2. Le choix d'un niveau aussi précoce que la GSM est justifié par deux raisons. D'une part, la plupart des recherches qui figurent dans les chapitres 2, 3 et 4 se sont rarement intéressées à des niveaux en deçà du CP. D'autre part, il suffit de pénétrer dans une classe de GSM pour s'apercevoir que dès ce niveau, les enfants sont sensibilisés à l'écrit *via* des activités ludiques mais qui constituent néanmoins une première exposition aux spécificités de leur langue. Le passage de la GSM au CP représente ainsi une période charnière pendant laquelle les enfants passent d'une conscience implicite à une conscience beaucoup plus explicite dans l'utilisation du langage et des structures de ce langage. La population est également constituée d'un groupe d'adultes. La participation de ce groupe répond à plusieurs objectifs. Tout d'abord, selon Gombert (2003), l'apprentissage implicite se développe bien au-delà de la scolarisation : il se poursuit tant que l'individu lit et s'accroît au fur et à mesure des lectures individuelles. Les connaissances explicites décuplées par les connaissances implicites devraient également poursuivre leur développement chez le lecteur expert. Ainsi, il s'agit de vérifier l'étendue des connaissances implicites et explicites en graphophonologie, orthographe et morphologie chez l'expert. L'utilisation de telles connaissances en lecture chez le lecteur expert est comparée à celle d'enfants de niveau CM2 (i.e., les plus efficaces en lecture) afin de déterminer si la nature des traitements opérés en lecture se modifie après ce niveau et ainsi, vérifier si les enfants de niveau CM2 sont toujours apprenants ou peuvent déjà être considérés comme des experts en lecture.

Concernant le matériel choisi, il vise à mettre en exergue chaque dimension étudiée – implicite, explicite et lecture – pour chacune des connaissances envisagées (graphophonologie, orthographe et morphologie). Pour la dimension implicite, des tâches de plausibilité lexicale portant sur des pseudo-mots sont utilisées conformément au type d'épreuve généralement choisi

pour mettre en évidence les aspects implicites dans la plupart des études décrites dans les chapitres 2, 3 et 4. L'usage de pseudo-mots permet d'écarter tout usage de règles spécifiques et rend ainsi compte de la dimension implicite. L'évaluation des connaissances explicites est réalisée à partir de mots ou de syllabes dans le cas spécifique de l'orthographe pour l'étude des bigrammes. Pour la dimension explicite, les tâches utilisées, de par les consignes détaillées et l'utilisation d'un matériel lexical concret, nécessitent une analyse du participant. Ces tâches permettent ainsi d'exploiter un certain niveau de conscience explicite. Pour chaque type de connaissances évalué et pour chaque dimension (i.e., implicite et explicite), le matériel a été construit en utilisant les spécificités étudiées dans les chapitres 2, 3 et 4 : les graphèmes consistants *vs.* inconsistants (graphophonologie), les bigrammes fréquents *vs.* rares en position initiale et finale (orthographe), les affixes *vs.* pseudo-affixes (morphologie). Enfin, pour vérifier si les participants se servent de la structure graphophonologique, orthographique et morphologique en lecture, des épreuves de lecture à haute voix de mots et pseudo-mots sont utilisées.

Le matériel est aussi constitué d'une épreuve de lecture rapide de mots permettant de déterminer au préalable le niveau de lecture des enfants et d'une épreuve de vocabulaire. Cette dernière épreuve est exploitée avant tout pour mettre en évidence d'éventuels liens avec les connaissances graphophonologiques, orthographiques, morphologiques et la lecture. En effet, depuis de nombreuses années, l'étendue du vocabulaire est régulièrement mise en lien avec la réussite en lecture (voir par exemple, Anderson & Freebody, 1985 ; Anglin, 1993 ; Nagy & Anderson, 1984 et plus récemment Colé *et al.*, 2003). De même, plusieurs auteurs ont établi des relations entre le vocabulaire et la sensibilité morphologique. Ainsi, Fowler et Liberman (1995) soulignent le rôle important de la conscience morphologique dans l'acquisition du vocabulaire durant les années scolaires. Carlisle (2000) obtient respectivement des corrélations positives chez des enfants de niveau CE2 et CM2, indiquant des liens importants entre connaissances morphologiques et vocabulaire. Quant aux connaissances graphophonologiques et orthographiques, elles nécessitent elles aussi pour se développer une fréquentation d'écrits et donc une confrontation au vocabulaire disponible dans la langue. L'utilisation des résultats obtenus à l'épreuve de vocabulaire devrait ainsi permettre d'évaluer les liens existants entre graphophonologie, orthographe, morphologie, lecture et vocabulaire et d'examiner si le vocabulaire constitue un prédicteur puissant de la lecture.

D'un point de vue expérimental, le chapitre 6 qui suit vise à mettre en évidence le développement des connaissances implicites et explicites en graphophonologique, orthographe et morphologie ainsi que l'utilisation de ces connaissances en lecture. Il s'agit également de

montrer que ces connaissances comme leur utilisation en lecture dépendent du niveau de lecture préalable. Les trois expériences qui constituent le chapitre 6 renvoient respectivement à chaque connaissance étudiée : graphophonologie puis orthographe et enfin morphologie. Le chapitre 7 s'intéresse aux connaissances implicites et explicites possédées par les lecteurs experts (adultes) et s'attache également à comparer les traitements effectués en lecture par ces lecteurs experts à ceux réalisés par les enfants les plus efficaces en lecture. Enfin, le chapitre 8 porte sur les relations entre les connaissances graphophonologiques, orthographiques, morphologiques, la lecture et le vocabulaire en fonction du niveau de lecture préalable des participants. L'objectif est ici non seulement d'étudier les liens (corrélations) entre les différents types de capacités mais également de déterminer laquelle de ces capacités contribue le plus à la lecture.

CHAPITRE 6 – GRAPHOPHONOLOGIE, ORTHOGRAPHE ET MORPHOLOGIE : ASPECTS DEVELOPPEMENTAUX

Les expériences présentées dans ce chapitre ont pour objectif de mettre en évidence l'existence de connaissances implicites et explicites en graphophonologie, orthographe et morphologie. Il s'agit également de montrer l'utilisation de la structure graphophonologique, orthographique et morphologique pour lire des mots connus et nouveaux. Ces expériences réalisées auprès d'enfants scolarisés de la GSM au CM2 permettront par ailleurs de tester l'effet de l'expertise en lecture sur l'étendue des connaissances implicites et explicites ainsi que sur le type de traitement (graphophonologique, orthographique et morphologique) opéré lors de la lecture.

1- Expérience 1 : graphophonologie

Cette première expérience porte sur les aspects graphophonologiques. La spécificité graphophonologique testée est ici l'effet de consistance graphème-phonème. Il s'agira d'étudier le développement des connaissances implicites et explicites sur la consistance des correspondances graphèmes-phonèmes chez les enfants de la GSM au CM2. Pour ce faire, des tâches de plausibilité lexicale portant sur des pseudo-mots (versant implicite) et des tâches d'analyse explicite sur des mots seront utilisées. L'analyse des traitements graphophonologiques en lecture sera réalisée à partir d'une épreuve de lecture à haute voix de mots et pseudo-mots en exploitant à la fois l'exactitude et la vitesse de traitement (mesure du temps de lecture). Cette épreuve ainsi que la tâche explicite ne seront pas proposées aux enfants de GSM car, à ce niveau scolaire, les enfants ne sont encore que très peu capables d'effectuer une analyse phonologique sur des mots et ne sont pas encore lecteurs.

1.1- Méthode

1.1.1- Population

Cent soixante dix-neuf enfants de langue française et scolarisés dans différentes écoles de Nantes et de son agglomération ont participé à cette expérience qui s'est déroulée en janvier 2004. Les caractéristiques de ces enfants figurent dans le Tableau 1 ci-après.

Tableau 1. Caractéristiques des enfants de la population initiale

Niveau scolaire	Effectif	Age chronologique moyen (en mois)	Ecart-type	Age du plus jeune	Age du plus âgé
GSM	30	5,5	0,3	5,0	6,0
CP	30	6,6	0,4	6,1	7,1
CE1	28	7,6	0,6	7,0	8,6
CE2	30	8,4	0,5	7,11	9,6
CM1	28	9,6	0,4	9,10	10,7
CM2	33	10,6	0,6	9,9	12,0

Deux épreuves contrôles ont été utilisées auprès de l'échantillon initial afin de déterminer le niveau d'expertise en lecture d'une part et les connaissances en vocabulaire d'autre part. L'expertise en lecture a été établie pour les enfants scolarisés du CP au CM2 à partir de l'épreuve *Lecture en Une Minute (LUM)* (voir Annexe 1, pp. 241-242) de Khomsi (1999). En GSM, cette épreuve a été proposée également mais aucun des enfants n'est parvenu à lire plus de deux mots. Pour ces enfants, le niveau de lecture n'a donc pu être établi. Cette épreuve évalue la lecture à haute voix : il s'agit de lire le plus rapidement possible, en une minute maximum, 105 mots de difficulté croissante. La note maximale, qui correspond au nombre de mots lus correctement en une minute, est de 105. L'étalonnage de cette épreuve n'est pas suffisamment précis, notamment pour les enfants scolarisés en CP. De ce fait, un niveau moyen de performances a été calculé pour chaque niveau scolaire de la population initiale (voir Tableau 2). Les enfants ont ensuite été répartis en cinq niveaux de lecture par rapport à ce niveau de référence. Les enfants avec des scores proches du niveau moyen observé à leur niveau scolaire sont placés dans le groupe de leur niveau scolaire tandis que ceux avec un score inférieur d'au moins un écart-type au niveau observé sont placés dans le groupe immédiatement inférieur. Les enfants avec un score supérieur d'au moins un écart-type au niveau observé sont placés dans le groupe immédiatement supérieur. Ainsi, un enfant de CM1 avec un score *LUM* de 54 est placé dans le groupe de niveau CE2, un enfant avec un score *LUM* de 62 est placé dans le groupe de niveau CM1, un enfant ayant un score *LUM* de 92 est placé dans le groupe de niveau CM2. Pour le niveau scolaire CM2, les enfants ayant un score *LUM* supérieur à un écart-type de la moyenne de référence (c'est-à-dire supérieur à 94) ont tout de même été inclus dans le groupe de niveau CM2 (sinon il aurait été nécessaire de créer un autre groupe pour un trop petit nombre d'enfants). Le même critère a été retenu pour les enfants ayant un score inférieur d'un écart-type de la moyenne de référence du niveau scolaire CP.

Tableau 2. Performances moyennes obtenues à *LUM* pour chaque niveau scolaire

Niveau scolaire	Moyenne	Ecart-type	Borne inférieure	Borne supérieure
CP	16,83	6,29	10,54	23,12
CE1	42,64	10,67	31,97	53,31
CE2	63,93	14,03	49,9	77,96
CM1	75,03	16,54	58,49	91,57
CM2	78,94	15,69	63,25	94,63

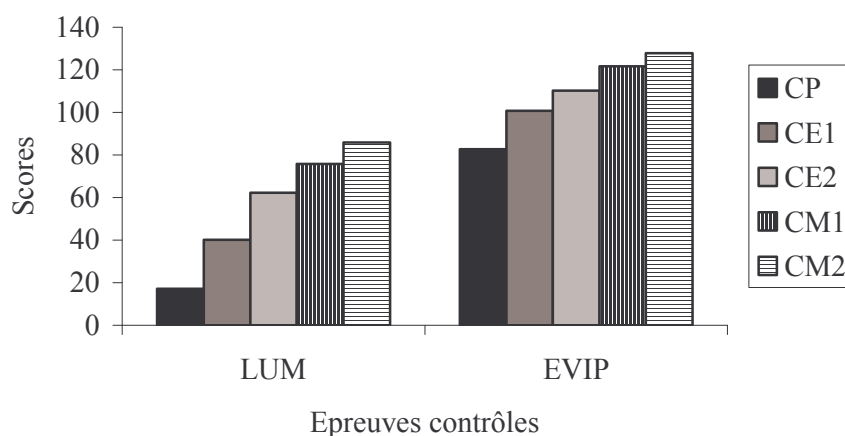
La répartition des enfants selon leur niveau d'expertise à *LUM* figure dans le Tableau 3 ci-dessous. C'est donc le niveau de lecture et non l'inscription officielle dans un niveau scolaire qui a présidé à la détermination des groupes. Toutefois, par commodité les cinq groupes de niveau de lecture *LUM* ainsi constitués seront dénommés : niveau CP, CE1, CE2, CM1 et CM2.

Tableau 3. Nouvelle répartition des enfants selon leur niveau de lecture

Niveau de lecture	Effectif	Age chronologique moyen (<i>Ecart-type</i>)	Score LUM moyen (<i>Ecart-type</i>)
CP	31	6,7 (0,6)	17,10 (6,44)
CE1	29	7,8 (1,2)	40,21 (6,87)
CE2	32	8,5 (0,7)	62,25 (8,54)
CM1	25	9,5 (0,9)	75,84 (10,93)
CM2	32	10,4 (0,6)	85,91 (12,94)

Les groupes de niveau de lecture se différencient sur les scores obtenus à l'épreuve *LUM* (cf. Figure 5). Ainsi, les scores de enfants de niveau CP se distinguent de manière significative de ceux des enfants des niveaux CE1 ($F(1, 58) = 80,99$; $p < .0001$), CE2 ($F(1, 61) = 392,45$; $p < .0001$), CM1 ($F(1, 54) = 888,71$; $p < .0001$) et CM2 ($F(1, 61) = 933,49$; $p < .0001$). Les performances des enfants de niveau CE1 se différencient significativement de celles des enfants de niveau CE2, CM1 et CM2 (respectivement $F(1, 59) = 109,20$; $p < .0001$; $F(1, 52) = 286,33$; $p < .0001$; $F(1, 59) = 383,94$; $p < .0001$). Les scores des enfants de niveau CE2 se distinguent également significativement de ceux des enfants de niveau CM1 ($F(1, 55) = 29,27$; $p < .0001$) et CM2 ($F(1, 62) = 73,45$; $p < .0001$). Enfin, les performances des enfants de niveau CM1 se différencient de façon significative de celles obtenues par les enfants de niveau CM2 ($F(1, 55) = 9,80$; $p < .01$).

Les connaissances en vocabulaire des enfants des niveaux CP à CM2) ont été évaluées grâce à l'épreuve d'*Evaluation du Vocabulaire en Images Peabody (EVIP, cf. Annexe 1, pp. 243-244)* de Dunn et Thiéroult-Whalen (1993). Cette tâche consiste pour le participant à choisir parmi quatre images celle qui illustre le mieux la signification d'un mot prononcé à haute voix par l'expérimentateur. Pour cette épreuve, les enfants de niveau CP (82,74 points) se différencient significativement des enfants de niveau CE1 (100,69 ; $F(1, 58) = 23,30$; $p < .0001$; voir Figure 5), CE2 (110,25 ; $F(1, 61) = 74,77$; $p < .0001$), CM1 (121,68 ; $F(1, 54) = 120,11$; $p < .0001$) et CM2 (127,75 ; $F(1, 61) = 163,17$; $p < .0001$). Les scores des enfants de niveau CE1 se distinguent de manière significative de ceux des enfants des niveaux CE2, CM1 et CM2 (respectivement $F(1, 59) = 8,30$; $p < .01$; $F(1, 62) = 32,06$; $p < .0001$; $F(1, 59) = 55,04$; $p < .0001$). Les performances des enfants de niveau CE2 diffèrent significativement de celles obtenues par les enfants de niveau CM1 ($F(1, 55) = 13,81$; $p < .001$) et de niveau CM2 ($F(1, 62) = 31,49$; $p < .0001$). En revanche, la différence entre les enfants des niveaux CM1 et CM2 n'est pas significative ($F(1, 55) = 3,04$; *ns*).



Notes maximales possibles à *LUM* : 105 ; à l'*EVIP* : 170.

Figure 5. Performances aux épreuves *LUM* et *EVIP* obtenues par les cinq groupes de niveau de lecture

1.1.2- Matériel

Epreuves graphophonologiques implicite, explicite et lecture à haute voix

Les trois épreuves ci-dessous consistent à évaluer les connaissances graphophonologiques implicites et explicites dont disposent l'apprenti lecteur et le lecteur débutant. Elles vont porter plus spécifiquement sur les relations graphophonologiques en exploitant le degré de consistance

des associations entre graphèmes et phonèmes (sens de la relation exploitée : graphème → phonème).

Tâche graphophonologique implicite de plausibilité lexicale écrite

Cette épreuve est constituée de 8 paires d'items expérimentaux et de 4 paires d'items distracteurs (cf. Annexe 2, pp. 245-246). Les items expérimentaux sont ici peu nombreux. En effet, le nombre de graphèmes inconsistants étant relativement peu important dans la langue française, le nombre d'items pouvant être élaborés était restreint. Chaque paire d'items expérimentaux contient un pseudo-mot consistant et un pseudo-mot inconsistant. Le pseudo-mot consistant est formé exclusivement de graphèmes consistants. Le pseudo-mot inconsistant est constitué des mêmes graphèmes que le pseudo-mot consistant qui lui est associé dans la paire sauf un graphème qui est inconsistant dans la langue (exemple : dans la paire *auvron-aivron*, les deux pseudo-mots sont constitués de 4 graphèmes chacun ; [au], [v], [r] et [on] sont des graphèmes consistants pour le premier mot : ils se prononcent /o/, /v/, /ʀ/ et /ɔ̃/; [ai] est un graphème inconsistant pour le second pseudo-mot : il peut se prononcer /e/ et /ɛ/).

Les items distracteurs sont constitués de pseudo-mots consistants. Dans chaque paire, les graphèmes sont identiques pour les deux pseudo-mots sauf un qui reste un graphème consistant mais peu fréquent dans la langue compte tenu de sa position dans l'item (exemple : dans la paire *bontra-bontro*, seuls les graphèmes [a] et [o] diffèrent entre les deux pseudo-mots, ces graphèmes sont consistants mais [-o] /o/ est moins fréquent que [-a] /a/ en position finale ; les graphèmes communs aux deux pseudo-mots [b], [on], [t] et [r] sont consistants : ils se prononcent /b/, /ɔ̃/, /t/ et /ʀ/). Ces items distracteurs ne seront pas pris en compte lors de l'analyse des réponses fournies.

Pour les items expérimentaux, le nombre de lettres est de 6 pour 4 paires et de 7 pour 4 paires. Le nombre de syllabes est de 2 pour 5 paires et de 3 pour 3 paires. La plupart des graphèmes retranscrivent les phonèmes de la langue. Cependant, quelques graphèmes ne sont pas phonocentrés : par exemple, le graphème [c] en position finale peut se prononcer /k/ (*sac* /sak/ ou *bac* /bak/) mais peut aussi se comporter comme une lettre muette (*tabac* /taba/ ou *estomac* /estoma/). Pour cette raison, il n'est pas possible de fournir le nombre exact de phonèmes car celui-ci varie selon la prononciation envisagée des graphèmes correspondants. La fréquence des bigrammes et des trigrammes composant ces items a été prise en compte : l'analyse statistique réalisée figure en annexe (cf. Annexe 2, p. 246). Quatre paires de pseudo-mots (*fita-gntv*, *ratu-bkse*, *nilo-mrzi* et *tale-xlpg*) constituent l'entraînement. Un seul pseudo-mot de chacune des paires peut constituer un mot français (i.e., est prononçable). Les paires expérimentales et les

paires d'items distracteurs sont placées selon un ordre aléatoire (randomisation à l'intérieur des paires et pour l'ensemble des paires), après les paires de pseudo-mots d'entraînement.

Pour les enfants de GSM, chacune des paires est écrite sur une feuille, en minuscules (Police Arial, Taille 28). A partir du niveau CP et jusqu'au CM2, les feuilles sur lesquelles figurent les paires sont agrafées de manière à constituer un carnet. Chaque page est numérotée et contient une seule paire de pseudo-mots afin d'éviter que les participants ne reviennent en arrière pour modifier leurs réponses précédentes.

Tâche graphophonologique explicite de jugement portant sur des mots irréguliers et réguliers fréquents de la langue (à partir du niveau CP)

Dix paires d'items expérimentaux constituent le matériel (*cf.* Annexe 2, pp. 246-247). Chaque paire d'items expérimentaux est constituée de mots fréquents dont l'un est irrégulier et l'autre régulier sur le plan grapho-phonologique. Les mots irréguliers de chaque paire sont des mots pour lesquels les règles habituelles de transcodage graphèmes-phonèmes ne fonctionnent pas (exemple : *ville* /vil/ et *août* /u(t)/ sont des mots fréquents et irréguliers de la langue française). Les mots réguliers associés dans la paire ont la caractéristique de comporter chacun un graphème identique aux mots de la première liste mais ce graphème a cette fois-ci une prononciation régulière. Par exemple, dans la paire *oignon-oiseau*, le mot *oignon* /ɔ̃ʁɔ̃/ est irrégulier car le graphème [oi] se prononce ici /ɔ̃/ tandis que dans *oiseau* /wazo/, ce même graphème va se prononcer de façon régulière : le phonème correspondant est alors /wa/. La fréquence lexicale des items choisis est fournie en annexe (base de données lexicales utilisée : *Manulex* de Lété, Sprenger-Charolles, & Colé, 2004 ; *cf.* Annexe 2, p. 247). L'analyse statistique comparative des fréquences lexicales des items réguliers et des items irréguliers figure en annexe (*cf.* Annexe 2, p. 247). Deux paires de pseudo-mots constituent l'entraînement : *seconde-raconter* /s(ə)gɔ̃d/-/ʁakɔ̃tə/ et *huit-nuit* /'ɥit/-/nɥi/. Ces items d'entraînement répondent aux mêmes critères de régularité que les items expérimentaux. Pour chaque paire d'items, c'est le mot irrégulier qui occupe la première position. L'ordre de présentation des paires est aléatoire. Chacune des paires est écrite sur une feuille, en minuscules (police Arial, taille 28) et les paires expérimentales sont placées après celles utilisées pour l'entraînement.

Tâche de lecture à haute voix (uniquement à partir du niveau CP)

Le matériel expérimental est constitué de 32 items à lire (*cf.* Annexe 2, pp. 247-248), soit :

- 8 mots consistants (*banane* /banan/, *lavabo* /lavabo/)
- 8 mots inconsistants (*monsieur* /mɛsjø/, *automne* /ɔ̃(o)tɔ̃n/)

- 8 pseudo-mots consistants (*bertalie* /bɛʁtali/, *janitou* /ʒanitu/),
- 8 pseudo-mots inconsistants (*baliden*, *lautil*).

Les items sont composés de 6 à 8 lettres et de 2 à 3 syllabes. Comme pour l'épreuve implicite, il n'est pas possible de fournir ici le nombre exact de phonèmes. La fréquence des bigrammes et des trigrammes composant les items a été prise en compte. Les analyses statistiques portant sur la fréquence des bigrammes, des trigrammes, sur le nombre de lettres et de syllabes figurent en annexe (cf. Annexe 2, p. 248). La fréquence lexicale des mots réguliers et des mots irréguliers a été prise en compte : l'analyse statistique comparative est fournie en annexe (cf. Annexe 2, p. 248). Quatre items neutres constituent l'entraînement afin de familiariser les participants au matériel et à la lecture à voix haute. Ces items sont : *faute*, *tarola*, *botte*, *fudin*. Les items sont écrits en minuscules et présentés un par un sur un écran d'ordinateur. Les 32 items sont présentés selon un ordre aléatoire et juste après les quatre items de familiarisation.

1.1.3- Procédure

Epreuves graphophonologiques implicite, explicite et lecture à haute voix

Tâche graphophonologique implicite de plausibilité lexicale écrite

La passation est individuelle pour les GSM et semi-collective (groupe de 6 à 8 participants) à partir du niveau CP. Avant de commencer, il est dit aux participants que leur tâche consiste à bien regarder chaque paire de pseudo-mots et à montrer (GSM) ou à entourer (du niveau CP au niveau CM2) le pseudo-mot de chaque paire qui leur semble le plus facile à lire pour un enfant qui commence tout juste à apprendre à lire. En GSM, les feuilles sur lesquelles sont inscrits les items sont présentées une à une, devant le participant, au fur et à mesure des réponses fournies. Des niveaux CP à CM2, les numéros des pages du carnet correspondant à chaque paire sont prononcés pour éviter les oublis d'items et favoriser la considération attentive de chaque paire de pseudo-mots. Pour la cotation, l'épreuve est notée sur 8 points pour les enfants de GSM et ceux des niveaux CP à CM2 (un point par réponse correcte).

La consigne donnée aux enfants de GSM est la suivante : « *Je vais te montrer des feuilles. Sur chaque feuille, deux mots sont écrits. Les mots que tu vas voir ne sont pas de vrais mots : ils n'existent pas, ils ont été inventés et ne veulent rien dire. On voudrait montrer ces mots à un enfant qui commence juste à lire pour savoir s'il y a des mots qui sont plus faciles à lire que d'autres. Alors, regarde bien les deux mots : je voudrais que tu me montres celui qui à ton avis serait le plus facile à lire pour cet enfant. Tu as bien compris ? Tu n'as pas de questions ?* ».

La consigne fournie aux enfants des niveaux CP à CM2 est : « *Je vais vous distribuer un petit carnet. Sur chaque page du carnet, deux mots sont écrits. Les mots que vous allez voir ne sont pas de vrais mots : ils n'existent pas, ils ont été inventés et ne veulent rien dire. On voudrait montrer ces mots à un enfant qui commence juste à lire pour savoir s'il y a des mots qui sont plus faciles à lire que d'autres. Alors, regardez bien les deux mots : je voudrais que vous entouriez avec un crayon celui qui à votre avis serait le plus facile à lire pour cet enfant. Quand vous avez fini d'entourer pour la première page, vous attendez que je vous dise de tourner la page pour continuer. Attention, vous n'avez pas le droit de revenir en arrière. Vous avez bien compris ? Vous n'avez pas de questions ?* ».

Tâche graphophonologique explicite de jugement portant sur des mots irréguliers et réguliers fréquents de la langue (à partir du niveau CP)

La passation est individuelle. Avant de commencer, il est signalé au participant qu'il va voir des mots qu'il a sûrement déjà vus et entendus. Les deux paires d'entraînement sont proposées puis débute la phase expérimentale. Les feuilles sur lesquelles sont inscrits les items sont présentées une à une, devant le participant, au fur et à mesure des réponses fournies. Aucun feedback correctif n'est fourni. Les réponses orales du participant ou l'absence de réponse sont notées par l'expérimentateur. Un point est attribué si les deux façons de lire les mots (correctes et incorrectes) sont données. Dix points sont ainsi attribués respectivement aux items irréguliers et réguliers. La note globale est sur 20.

La consigne est la suivante: « *Regarde bien ces deux mots. Ça, c'est le mot seconde (montrer seconde). Ça, c'est le mot raconter (montrer raconter). On va essayer de trouver la bonne façon et la mauvaise façon de lire ces mots. Pour seconde, la bonne façon c'est /s(ə)gõd/. La mauvaise façon, c'est de lire le [c] de seconde comme dans raconter. Si on lit seconde comme raconter, ça fait quoi à ton avis ?* ». (Le participant doit fournir sa réponse oralement. La réponse attendue est /s(ə)kõd/. Si le participant fournit une autre réponse ou ne fournit aucune réponse, un feedback correctif est fourni). « *Maintenant, pour le mot raconter, la bonne façon de le lire, c'est /rakõte/. La mauvaise façon de lire raconter, c'est de le lire comme seconde. Si on lit raconter comme seconde, ça fait quoi à ton avis ?* ». (Le participant doit également fournir sa réponse oralement. La réponse attendue est /ragõte/. Si le participant fournit une autre réponse ou ne fournit aucune réponse, un feedback correctif lui est fourni). « *Bon, alors maintenant, regarde ces deux mots. Ça, c'est le mot huit (montrer huit). Ça, c'est le mot nuit (montrer nuit). La bonne façon de lire huit, c'est /'uit/. La mauvaise façon, c'est de lire le [t] à la fin de huit comme dans nuit. Si on lit huit comme nuit, ça fait quoi à ton avis ?* ». (Le participant doit

fournir sa réponse oralement. La réponse attendue est /'ɥi/. Si le participant fournit une autre réponse ou ne fournit aucune réponse, un feed-back correctif est fourni). « *Maintenant, pour nuit, la bonne façon de le lire, c'est / nuʝi/. La mauvaise façon de lire nuit, c'est de le lire comme huit. Si on lit nuit comme huit, ça fait quoi à ton avis ?* ». (Le participant doit également fournir sa réponse oralement. La réponse attendue est /nuʝit/. Si le participant fournit une autre réponse ou ne fournit aucune réponse, un feed-back correctif lui est fourni). « *Tu as bien compris ? Tu n'as pas de questions ? Alors, maintenant regarde bien ces deux mots (montrer la première paire expérimentale). Est-ce que tu peux me donner la bonne et la mauvaise façon de lire pour chaque mot ?* ».

Tâche de lecture à haute voix (à partir du niveau CP)

La passation est individuelle. Il est précisé aux enfants qu'ils vont devoir lire à voix haute des mots apparaissant sur un écran d'ordinateur. Les participants sont prévenus avant de commencer la lecture que parmi tous les mots certains n'existent pas et ne veulent rien dire. Ils ont pour consigne de les lire comme de vrais mots. Les quatre items de familiarisation à la tâche sont proposés puis débute la phase expérimentale. Le passage d'un item au suivant est réalisé par l'expérimentateur en appuyant sur la barre d'espace. Une pause est effectuée tous les quatre items lus. Un point est attribué par mot lu correctement dans chacune des catégories (mots réguliers, pseudo-mots réguliers, mots irréguliers et pseudo-mots irréguliers).

La consigne précise est la suivante : « *Tu vas voir sur l'écran de l'ordinateur des mots. Il y a de vrais mots et des mots qui ont été inventés. Ces mots inventés n'existent pas et ne veulent rien dire. Tu dois lire à haute voix chaque mot qui apparaît sur l'écran* ».

L'ordre de passation a été aléatorisé pour l'ensemble des épreuves.

1.2- Hypothèses opérationnelles

Epreuve implicite : les pseudo-mots formés uniquement de graphèmes consistants (*bamiran*) devraient être jugés plus faciles à lire que ceux composés à la fois de graphèmes consistants et de graphèmes inconsistants (*bamirac*). Plus le niveau de lecture des enfants est avancé, plus les choix devraient s'orienter vers les items consistants.

Epreuve explicite : face à un item régulier (*chose*) et un item irrégulier (*chorale*), fournir les deux types de prononciation, c'est-à-dire la bonne – /ʃoz/ et /kɔʁal/ – et la mauvaise façon – /koz/ et /ʃɔʁal/ – devrait être plus facile dans le cas de l'item irrégulier. En effet, « irrégulariser » ou donner la mauvaise prononciation d'un item régulier est inhabituel pour un enfant.

« Régulariser » et « irrégulariser » devraient alors apparaître plus simples pour un item irrégulier que pour un item régulier. Cet effet de la régularité de l’item devrait s’accroître avec l’efficacité en lecture.

Lecture à haute voix : les mots (*banane*) devraient être mieux lus (plus rapidement et avec moins d’erreurs) que les pseudo-mots (*bertainie*). Les items consistants seront mieux lus que les inconsistants notamment dans le cas des mots (*camarade vs. aquarium*). Pour les pseudo-mots, les consistants (*falore*) devraient être lus plus rapidement que les inconsistants (*archérou*) mais aucune différence significative ne devrait être obtenue entre ces deux types d’items au niveau de la précision de la lecture. En effet, pour les pseudo-mots inconsistants, les prononciations régulières ou irrégulières sont acceptables du fait de la non lexicalité de l’item. Ces résultats seraient plus prononcés chez les enfants les plus efficaces en lecture.

1.3- Résultats

1.3.1- Tâches graphophonologiques implicite et explicite

Les performances obtenues par les enfants aux deux épreuves graphophonologiques figurent dans le Tableau 4 ci-dessous. Les enfants scolarisés en GSM n’ont passé que la tâche implicite. Les enfants à partir du niveau CP jusqu’au niveau CM2 ont réalisé les épreuves implicite et explicite. Pour la tâche explicite, une distinction des scores aux items réguliers et irréguliers a été effectuée.

Tableau 4. Performances exprimées en pourcentages moyens de réussite² (et *écart-types*) obtenues aux épreuves morphologiques selon le type d’affixe (préfixe et suffixe)

	Graphophonologie implicite	Graphophonologie explicite		
		Irréguliers	Réguliers	Moyenne
GSM	50,42 (16,24)			
Niveau CP	47,18 (19,82)	10,32 (13,78)	14,52 (13,12)	12,42
Niveau CE1	47,41 (16,50)	36,21 (23,21)	30 (22,99)	33,10
Niveau CE2	50 (17,68)	65 (22,58)	50,94 (25,70)	57,97
Niveau CM1	62,50 (13,01)	79,20 (20,40)	70 (23,27)	74,60
Niveau CM2	61,33 (16,61)	79,69 (20,24)	71,25 (23,93)	75,47
Moyenne	53,14	54,08	47,34	50,71

² Par convention, le choix de l’item consistant dans l’épreuve implicite est qualifié de « réussite » alors qu’en toute rigueur il n’y a pas de réussite dans cette épreuve, le choix exprimant une préférence.

Tâche graphophonologique implicite

En graphophonologie implicite, les enfants de la GSM au niveau de lecture CM2 doivent choisir entre deux pseudo-mots (l'un contenant un graphème irrégulier et des graphèmes irréguliers comme *bamirac* et l'autre ne contenant que des graphèmes réguliers tel que *bamiran*). Les écart-types sont assez importants pour cette tâche indiquant des résultats relativement dispersés (voir Tableau 4). L'écart-type le plus faible est celui des enfants de niveau CM1 (13,01).

Les analyses qui suivent portent sur un seul facteur : le niveau de lecture (six modalités : GSM, niveaux CP, CE1, CE2, CM1, CM2). Le plan d'analyse correspond à : $S_n < N_6 >$ où S est le facteur sujet et N correspond au niveau de lecture des enfants.

Pour cette tâche de plausibilité lexicale, seules les réponses des enfants des niveaux CM1 et CM2 se différencient du hasard ($p < .001$ pour les deux niveaux). Ces enfants choisissent donc préférentiellement les items réguliers comme ressemblant le plus à vrai mot. Pour les autres enfants (de la GSM au niveau de lecture CE2), les réponses fournies ne se distinguent pas du hasard.

Le facteur niveau de lecture est significatif ($F(5, 173) = 4,85 ; p < .001$) Les enfants de niveau CM2 (61,33 %) et CM1 (62,50 %) ont de meilleures performances que ceux de GSM (50,42 % ; respectivement $F(1, 60) = 6,83 ; p < .05 ; F(1, 53) = 9,01 ; p < .01$), de niveau CP (47,18 % ; $F(1, 63) = 9,46 ; p < .01 ; F(1, 54) = 11,07 ; p < .01$), CE1 (47,41 % ; $F(1, 59) = 10,75 ; p < .01 ; F(1, 52) = 13,60 ; p < .001$) et CE2 (50 % ; $F(1, 62) = 6,98 ; p < .05 ; F(1, 55) = 8,77 ; p < .01$). Les autres différences ne sont pas significatives.

Tâche graphophonologique explicite

Dans cette épreuve proposée aux enfants des niveaux CP à CM2, deux types d'item – l'un régulier, l'autre irrégulier – sont présentés. Les enfants doivent fournir la bonne et la mauvaise façon de prononcer chaque item (régulariser l'item irrégulier et « irrégulariser » l'item régulier). Les analyses réalisées prennent en compte deux facteurs : le niveau de lecture (cinq modalités : CP, CE1, CE2, CM1, CM2) et la régularité des items présentés (irréguliers vs. réguliers). Ainsi, le plan d'analyse est le suivant : $S_n < N_5 > * R_2$. Le facteur S correspond aux sujets, N correspond au niveau de lecture et R désigne la régularité des items présentés.

Sur l'ensemble de la population, la moitié des enfants (50,71 %) réussit à répondre correctement à cette consigne. Le niveau de lecture a un effet significatif ($F(4, 144) = 58,71 ; p < .0001$). Les enfants de niveau CM2 (75,47 %) et de niveau CM1 (74,60 %) ont de meilleures performances que ceux des niveaux CP à CE2 (CP : 12,42 % ; CE1 : 33,10 % et CE2 : 57,97 %). Les différences sont toutes significatives à .01, au minimum. L'effet du facteur régularité (items

irréguliers vs. items réguliers) est également significatif ($F(1, 144) = 24,05 ; p < .0001$). Fournir les deux types de prononciation pour les items irréguliers est plus facile que pour les items réguliers, tous niveaux de lecture confondus ($54,08 \% > 47,34 \%$). Enfin, les deux facteurs (niveau de lecture et régularité) interagissent de manière significative ($F(4, 144) = 5,08 ; p < .0001$). La Figure 6 montre que la réussite est effectivement meilleure pour les items irréguliers que pour les réguliers, mais seulement chez les enfants de niveau CE2 ($65 \% > 50,94 \% ; p = .0240$). Pour les autres niveaux de lecture, aucune différence significative entre les deux types d'items n'est observée.

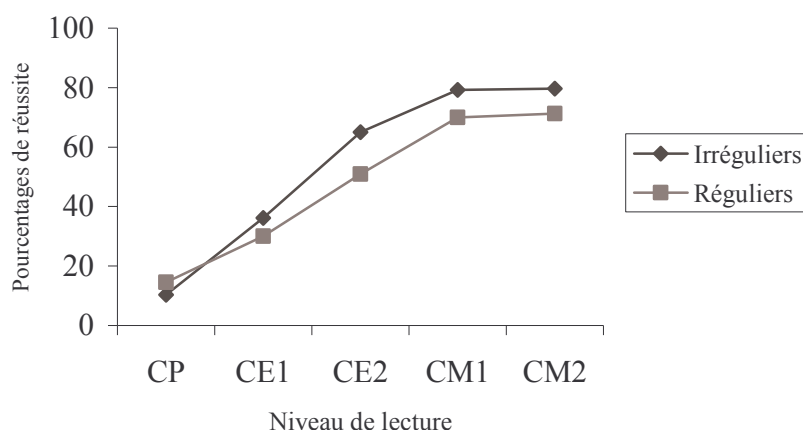


Figure 6. Performances aux items irréguliers et réguliers en fonction du niveau de lecture

1.3.2- Tâche de lecture à haute voix

Les analyses ont été réalisées sur l'exactitude en lecture (pourcentages d'items lus correctement) et sur les temps de lecture correcte (i.e., 2 V.D. ici). Trois facteurs principaux ont été pris en compte : le niveau de lecture (cinq modalités : CP, CE1, CE2, CM1, CM2), la lexicalité des items (mots vs. pseudo-mots) et la régularité des items (réguliers vs. irréguliers). Pour les analyses de l'exactitude et des temps de lecture correcte, le plan est le suivant : $S_n < N_5 > * L_2 * R_2$ (S : facteur sujet ; N : facteur niveau de lecture ; L : facteur lexicalité des items ; R : facteur régularité des items).

Exactitude

Le Tableau 5 rend compte des scores obtenus par les cinq groupes d'enfants aux items (mots et pseudo-mots) réguliers et irréguliers.

Tableau 5. Performances moyennes (et *écart-types*) en fonction du niveau de lecture

Niveau de lecture	Mots réguliers	Mots irréguliers	Pseudo-mots réguliers	Pseudo-mots irréguliers
CP	78,23 (16,13)	20,16 (13,95)	54,03 (21,01)	63,71 (19,99)
CE1	93,53 (7,92)	52,16 (19,50)	77,59 (19,30)	79,31 (19,56)
CE2	98,05 (4,61)	82,42 (18,47)	78,13 (20,08)	87,50 (13,84)
CM1	100 (0)	84,50 (13,64)	82 (15,34)	85,50 (13,35)
CM2	98,44 (4,20)	92,97 (9,49)	85,55 (12,74)	86,33 (13,23)
Moyenne	93,65	66,44	75,46	80,47

L'analyse des écart-types indique une hétérogénéité des résultats chez les enfants de niveau CP pour les pseudo-mots réguliers (21,01) et irréguliers (19,99) ainsi que chez les enfants de niveau CE1 pour ces mêmes items (respectivement 19,30 et 19,56) et pour les mots irréguliers (19,50). La dispersion est également importante pour les mots réguliers (18,47) et les pseudo-mots réguliers (20,08) chez ceux de niveau CE2.

L'analyse de variance réalisée sur les scores de lecture des enfants de niveau CP à CM2 indique un effet significatif du facteur niveau de lecture ($F(4, 144) = 64,38$; $p < .0001$). Les enfants de niveau CP (54,03 %) sont moins efficaces en lecture que les enfants des niveaux CE1 à CM2 (CE1 : 75,65 % ; CE2 : 86,52 % ; CM1 : 88 % ; CM2 : 90,82 % ; toutes les différences sont significatives à .0001). Les enfants de niveau CE1 sont également moins bons en lecture que ceux des niveaux CE2 à CM2 (respectivement $p = .0001$; $p = .002$ et $p = .0071$). Les différences ne sont pas significatives pour les autres niveaux. L'effet du facteur lexicalité est tendanciellement significatif ($F(1, 144) = 3,36$; $p = .069$). Tous groupes confondus, les mots (80,05 %) tendent à être mieux lus que les pseudo-mots (77,96 %). L'effet de la régularité des items est clairement significatif : les enfants lisent les items réguliers (84,55 %) en faisant moins d'erreurs que sur les items irréguliers (73,46 % ; $F(1, 144) = 122,82$; $p < .0001$).

Il existe des interactions significatives entre (a) le niveau de lecture et la lexicalité, (b) le niveau de lecture et la régularité des items et (c) la lexicalité et la régularité des items. Dans la mesure où ces trois facteurs interagissent également ($F(4, 144) = 26,90$; $p < .0001$; voir Figure 7), seule cette dernière interaction sera commentée. Pour l'ensemble des enfants, les mots réguliers sont mieux lus que les pseudo-mots réguliers (CP : 78,23 % vs. 54,03 % ; CE2 : 98,05 % vs. 78,13 % ; CM2 : 98,44 % vs. 85,55 % ; CE1 : 93,53 % vs. 77,59 % et CM1 : 100 % vs. 82

% : $p < .0001$ pour tous les niveaux de lecture). Les pseudo-mots irréguliers sont lus avec moins d'erreurs que les mots irréguliers mais par les enfants des niveaux CP (63,71 % > 20,16 % ; $p < .0001$) et CE1 (79,31 % > 52,16 % ; $p < .0001$) seulement. Pour les enfants des niveaux CM1 et CE2, les différences entre les deux catégories d'items irréguliers ne sont pas significatives et pour le niveau CM2, les mots irréguliers (92,97 %) sont lus avec moins d'erreurs que les pseudo-mots irréguliers (86,33 % ; $p = .0373$). Cette interaction confirme que les mots réguliers sont mieux lus que les mots irréguliers pour chaque niveau de lecture (CP : 78,23 % vs. 20,16 % ; CE1 : 93,53 % vs. 52,16 % ; CE2 : 98,05 % vs. 82,42 % ; $p < .0001$ pour les trois niveaux de lecture ; CM1 : 100 % vs. 84,50 % ; $p = .002$; CM2 : 98,44 % vs. 92,97 % ; $p = .0012$). Elle montre que les pseudo-mots irréguliers sont lus avec moins d'erreurs que les pseudo-mots réguliers seulement chez les enfants de niveau CE2 (87,50 % > 78,13 % ; $p = .0472$). La différence entre les deux types de pseudo-mots n'est que tendanciellement significative pour le niveau CP (irréguliers : 63,71 % ; réguliers : 54,03 % ; $p = .0596$) et non significative pour les autres groupes. L'interaction des trois facteurs montre bien que les enfants de niveau CM2 ont de meilleures performances que ceux de CM1 et CE2 pour les mots mais seulement pour les irréguliers (CM2 : 92,97 % ; CM1 : 84,50 % ; $p = .0176$ et CE2 : 82,42 % ; $p = .0106$). L'absence de différence significative entre les niveaux CM1 et CE1 sur les pseudo-mots s'observe tant pour les réguliers que pour les irréguliers alors qu'entre les niveaux CE2 et CE1, elle n'apparaît qu'avec les réguliers. Ainsi, pour les pseudo-mots irréguliers, les enfants de niveau CE2 lisent mieux que ceux de niveau CE1 (87,50 % > 79,31 % ; $p = .0465$). Enfin, alors que les enfants de niveau CE1 lisent les mots réguliers (93,53 %) et irréguliers (52,16 %) avec davantage d'erreurs que ceux de niveau CM2 (98,44 % ; $p = .0027$ et 92,44 % ; $p < .01$), les différences entre les deux groupes d'enfants ne sont que tendanciellement significatives pour les pseudo-mots réguliers et irréguliers ($p = .0754$ et $p = .0817$).

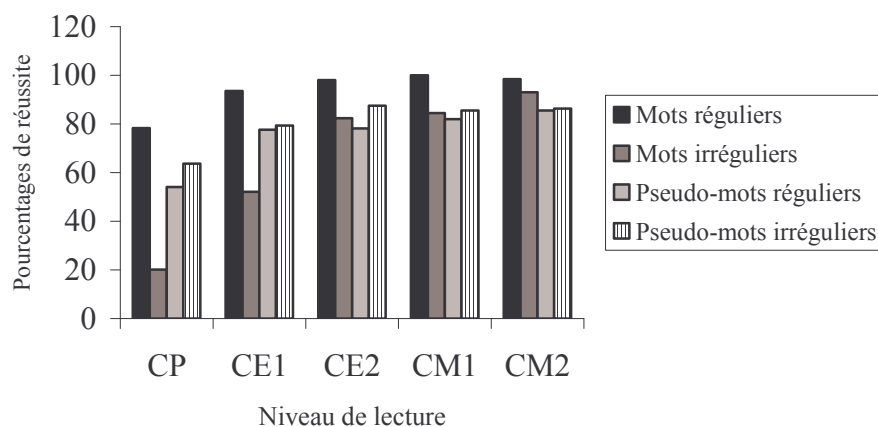


Figure 7. Performances à l'ensemble des items pour les cinq niveaux de lecture

Temps

Les temps obtenus pour les items incorrectement prononcés ont été écartés des analyses, conformément à ce qui est généralement observé dans la littérature portant sur la dénomination en lecture. Le temps moyen de lecture correcte pour chaque catégorie dépend du nombre d'items correctement lus dans cette catégorie. Seules les moyennes des enfants qui ont été capables de lire plus de deux items par catégorie ont été prises en compte. Ainsi, les analyses de temps de lecture correcte des items graphophonologiques concernent 13 enfants de niveau CP, 29 enfants de niveau CE1, 32 enfants de niveau CE2, 29 enfants de niveau CM1 et 32 enfants de niveau CM2.

Les temps moyens de lecture des items (exprimés en millisecondes, ms) figurent dans le Tableau ci-dessous.

Tableau 6. Temps moyens de lecture correcte (et *écart-types*) des items de l'épreuve graphophonologique en fonction du niveau de lecture

Niveau de lecture	Mots réguliers	Mots irréguliers	Pseudo-mots réguliers	Pseudo-mots irréguliers
CP	4499,18 (1872,24)	5393,47 (1761,89)	4811,00 (1715,85)	5419,76 (1884,58)
CE1	2422,02 (433,04)	2927,63 (726,18)	3948,25 (850,09)	3336,12 (1067,73)
CE2	1897,27 (252,42)	2145,83 (390,32)	2973,70 (840,79)	3460,17 (718,21)
CM1	1711,50 (217,86)	1788,55 (283,20)	2202,34 (372,70)	2339,86 (339,41)
CM2	1798,33 (254,62)	2163,66 (350,91)	2190,36 (449,56)	2418,54 (472,51)
Moyenne	2465,66	2883,83	3225,13	3394,89

Les temps de lecture correcte montrent une forte dispersion chez les enfants de niveau CP pour l'ensemble des items (mots réguliers : 1872,24 ms ; mots irréguliers : 1761,89 ms ; pseudo-mots réguliers : 1715,85 ms ; pseudo-mots pseudo-irréguliers : 1884,58 ms). L'hétérogénéité est également un peu plus importante pour les catégories pseudo-mots réguliers et pseudo-mots irréguliers par rapport aux autres catégories d'items chez les enfants des niveaux CE1 (850,09 ms et 1067,73 ms) et CE2 (840,79 ms et 718,21 ms). Pour les niveaux CM1 et CM2, les temps sont relativement homogènes.

Le niveau de lecture a un effet significatif sur les temps de lecture correcte des items ($F(4, 126) = 23,75 ; p < .0001$). Les enfants de niveau CP (5030,85 ms) lisent moins vite que les enfants des niveaux CE1 (3158,50 ms ; $p = .0034$), CE2 (2619,24 ms ; $p = .0001$), CM1 et CM2 (2010,56 ms et 2142,72 ms ; $p < .0001$). Les différences entre ces niveaux sont significatives à $.0001$. Les enfants de niveau CE1 sont également moins rapides en lecture que ceux des niveaux

CM1 et CM2 (respectivement $p = .0016$ et $p = .0030$). Enfin, les enfants de niveau CE2 lisent moins vite que ceux de niveau CM1 mais la différence entre les deux groupes n'est que tendanciellement significative ($p = .0638$). Concernant les deux autres facteurs, les pseudo-mots sont lus moins vite que les mots ($3310,01 \text{ ms} > 2674,74 \text{ ms}$; $F(1, 126) = 19,90$; $p < .0001$) et les enfants lisent plus rapidement les items réguliers que les irréguliers ($2845,40 \text{ ms} < 3139,36 \text{ ms}$; $F(1, 126) = 3,92$; $p < .05$). Aucune des interactions entre les différents facteurs n'est significative.

1.4- Discussion

Cette première expérience avait pour objectif d'étudier d'une part, le développement des connaissances graphophonologiques implicites (de la GSM au niveau CM2) et explicites (du niveau CP au niveau CM2) et d'autre part, l'évolution des traitements graphophonologiques réalisés par les lecteurs en fonction de leur niveau d'expertise en lecture.

1.4.1- Développement des connaissances graphophonologiques implicites et explicites

Deux tâches d'évaluation des connaissances graphophonologiques ont été utilisées dans l'expérience 1 : une tâche implicite de plausibilité lexicale écrite et une tâche d'analyse graphophonologique explicite.

Dans l'épreuve implicite, les enfants devaient indiquer parmi deux pseudo-mots, dont l'un était consistant (*piralue*) et l'autre inconsistant (*piralum*), lequel leur semblait le plus facile à lire. Il était supposé que les pseudo-mots consistants seraient considérés par les enfants comme plus faciles à lire que les inconsistants et que ces choix devraient augmenter avec le niveau de lecture. Les résultats indiquent que seuls les enfants les plus efficaces en lecture (niveaux de lecture CM1 et CM2) fournissent des réponses qui diffèrent du hasard : pour ces enfants, les items consistants sont considérés comme plus faciles à lire que ceux inconsistants. En revanche, les performances ne se différencient pas de façon significative du hasard pour les niveaux de lecture plus faibles (niveaux CP, CE1 et CE2) ainsi que pour les enfants en phase de pré-apprentissage de la lecture (GSM). Pour ces niveaux, les choix ne reposent pas sur une prise en compte de la consistance graphèmes-phonèmes tandis qu'à des niveaux de lecture plus avancés (CM1 et CM2), les enfants sont sensibles à la fréquence des associations entre graphèmes et phonèmes. Ces premiers résultats concernant l'éventualité d'une émergence précoce des connaissances implicites vont donc à l'encontre des résultats obtenus par Lamarche (1986) qui constatait que, très tôt, les enfants étaient sensibles à la fréquence des correspondances graphèmes-phonèmes. De même, Fayol *et al.* (1996) avaient mis en évidence que dès le CP et le

CE1, les enfants témoignaient d'une certaine sensibilité à des régularités de la langue concernant les transcriptions des phonèmes /o/ et /s/. Dans l'étude présentée ci-avant, la sensibilité graphophonologique existe, mais elle apparaît beaucoup plus tardivement. Ces résultats peuvent trouver une voie d'explication dans la diversité des graphèmes utilisés dans l'épreuve implicite de l'expérience 1. En effet, Fayol *et al.* (1996) ou encore Pacton *et al.* (1999) se sont centrés sur les phonèmes /o/ et /s/ et leurs transcriptions graphiques en fonction de la position dans un item donné. Ces auteurs se sont donc seulement intéressés à deux phonèmes spécifiques et à leurs transcriptions graphiques respectives. A l'opposé, l'étude réalisée dans cette thèse porte sur une diversité de graphèmes ([th-], [ch-], [au-], [ai-], [-an], [-ac], etc.) pouvant avoir une seule (item consistant) ou plusieurs réalisations phonémiques (items inconsistants). Cette diversité pourrait être à l'origine des performances chez les enfants de niveau de lecture inférieur au CM1.

Les résultats obtenus à l'épreuve graphophonologique explicite fournissent des éléments intéressants sur la façon dont les connaissances explicites se mettent en place. Dans cette tâche, l'enfant devait, face à un item régulier (*chose*) et un item irrégulier (*chorale*), fournir les deux types de prononciations, c'est-à-dire la bonne – /ʒoz/ et /kɔʁal/ – et la mauvaise façon – /koz/ et /ʒɔʁal/ –. L'hypothèse principale était qu'il devait être plus facile de « régulariser » l'item dans le cas de l'irrégulier dans la mesure où « irrégulariser », c'est-à-dire donner la mauvaise prononciation d'un item régulier, est inhabituel pour un enfant. Les données indiquent que, malgré des performances meilleures pour les enfants de niveau CM2 et CM1 par rapport aux autres niveaux pour l'ensemble des items, seuls les enfants de niveau CE2 réussissent mieux les items irréguliers. Pour les enfants de niveau CE1 et surtout de niveau CP, les performances sont faibles, indiquant que les enfants ne maîtrisent que très peu les correspondances graphèmes-phonèmes que celles-ci soient consistantes ou inconsistantes. Ces derniers résultats sont conformes au schéma développemental établi par Gombert (2003) : les connaissances explicites s'élaborent sur la base des connaissances implicites et se développent conjointement à l'apprentissage explicite de la lecture. Aux niveaux de lecture CP et CE1, il est envisageable que l'efficacité en lecture soit encore insuffisante pour permettre un développement des connaissances graphophonologiques explicites. Cependant, il faut souligner qu'à partir du niveau CE1, les écart-types sont relativement élevés pour les items réguliers de la tâche explicite comme pour les irréguliers : les différences interindividuelles sont donc fortes, indiquant des disparités chez les enfants quant à l'étendue de leurs connaissances graphophonologiques explicites. Au niveau de lecture CE2, l'amélioration des compétences en lecture ainsi que l'accroissement concomitant des connaissances orthographiques sur les mots irréguliers permettraient d'expliquer l'augmentation des performances sur les items irréguliers. Pour les enfants des

niveaux CM1 et CM2, l'automatisation des procédures de correspondances graphophonémiques et l'accroissement de leurs connaissances explicites élaborées sur la base de celles implicites (voir paragraphe précédent) leur permettraient de répondre aussi bien aux items réguliers qu'aux items irréguliers.

Outre l'étude du développement des connaissances graphophonologiques implicites et explicites, l'expérience 1 se donnait également pour objectif de rendre compte des traitements graphophonologiques réalisés au cours de la lecture.

1.4.2- Traitements graphophonologiques et niveau de lecture

Dans la tâche de lecture à haute voix proposée aux enfants des niveaux CP à CM2, quatre catégories d'items devaient être lues : des mots consistants (*banane, lavabo*), des mots inconsistants ou irréguliers (*monsieur, automne*), des pseudo-mots consistants (*bertalie, janitou*) et enfin des pseudo-mots inconsistants construits avec des graphèmes consistants et un graphème inconsistant (*baliden, lautil*). Les analyses ont porté sur l'exactitude et sur les temps de lecture correcte des items en fonction du niveau de lecture des enfants.

Ces analyses montrent que, d'une façon générale, les processus de lecture sont affectés par la lexicalité et la structure graphophonologique des items, c'est-à-dire la régularité des correspondances graphèmes-phonèmes, ceci conformément aux hypothèses. En effet, globalement, les mots (*banane, monsieur*) sont mieux lus (plus vite et avec moins d'erreurs) que les pseudo-mots (*janitou, lautil*). Les items consistants (*lavabo, bertalie*) sont également mieux lus que les inconsistants (*automne, baliden*) en termes d'exactitude et de temps de lecture. Logiquement, les enfants les moins experts en lecture (niveaux CP et CE1) lisent moins bien que les autres groupes de niveau. Ces mêmes enfants mettent également plus de temps à lire correctement les items (seulement dans le cas de la comparaison avec les niveaux CM1 et CM2 pour les enfants de niveau CE1). Les résultats concernant l'exactitude en lecture sont néanmoins à nuancer par l'interaction entre le niveau de lecture, la lexicalité et la régularité ou consistance des items. Les mots réguliers (*banane*) sont effectivement mieux lus que les pseudo-mots réguliers (*bertalie*) pour l'ensemble des enfants. Il en est de même pour les mots réguliers (*lavabo*) par rapport aux mots irréguliers (*monsieur*). Ce dernier résultat concorde avec les données obtenues par Sprenger-Charolles, Siegel et Bonnet (1998) et par Sprenger-Charolles, Siegel, Béchenec et Serniclaes (2003) qui avaient montré que la lecture des mots était affectée par la régularité. La consistance des correspondances graphèmes-phonèmes constitue donc un facteur qui influence la reconnaissance de mots connus et ce très précocement, dès le début de

l'apprentissage de la lecture. Cependant, dans le cas des irréguliers, la différence entre les pseudo-mots (*baliden*) et les mots (*monsieur*) n'est pas significative pour les enfants des niveaux CE2 et CM1. Cette absence de différence entre les deux catégories d'items peut s'expliquer par les bons résultats obtenus par ces enfants (plus de 80 % de réussite) mais aussi par la nature même des items en question. En effet, les prononciations régulières ou irrégulières sont acceptables du fait de la non lexicalité de l'item dans le cas des pseudo-mots irréguliers : à des niveaux de lecture CE2 ou CM1, les enfants peuvent facilement lire à haute voix ce type d'item. A partir du niveau de lecture CE2 et au-delà, les connaissances explicites sur la consistance des correspondances graphèmes-phonèmes tendent à se développer (voir paragraphe sur les performances des enfants de niveau CE2 à l'épreuve explicite de graphophonologie) et permettent d'améliorer l'exactitude en lecture des mots irréguliers. Ce constat est renforcé par la différence de performances entre les enfants de niveau CE1 (52,16 % de réussite) et de niveau CE2 (82,42 %) sur les mots irréguliers tandis que ces mêmes enfants ont des performances aux mots réguliers qui plafonnent et sont très proches (respectivement 93,53 % et 98,05 %). Chez les enfants de niveau CE1 mais aussi de CP, ce sont d'ailleurs les pseudo-mots irréguliers qui sont lus avec le moins d'erreurs par rapport aux mots irréguliers. L'écart de performance entre les deux types d'items est important, notamment pour les enfants de niveau de lecture CP dont les performances aux mots irréguliers sont très faibles (20,16 %). Ces résultats sont logiques au vu des performances de ces mêmes enfants aux épreuves implicite et explicite. Leur sensibilité et leurs connaissances de la consistance graphophonémique sont peu développées et encore difficiles à mobiliser face à des mots contenant des graphèmes irréguliers tandis que pour les enfants les plus experts en lecture (niveau CM2), les mots irréguliers sont directement accessibles dans le lexique mental grâce à l'appui des connaissances explicites bien établies. Ces mots sont ainsi lus avec plus d'exactitude que les pseudo-mots irréguliers. D'ailleurs, les enfants plus experts lisent beaucoup mieux ces items que ceux d'un niveau de lecture inférieur comme les enfants de niveau CM1 et CE2 confirmant une récupération automatique des mots irréguliers en mémoire chez les enfants de niveau CM2. Enfin et conformément aux hypothèses, la différence entre les pseudo-mots irréguliers et les pseudo-mots réguliers n'est pas significative sauf pour les enfants de niveau CE2 qui font moins d'erreurs en lisant *lautil* que *bertalie*. Ce résultat est surprenant mais peut sans doute s'expliquer par l'écart-type relativement élevé dans le cas des pseudo-mots réguliers : pour cette catégorie d'items, les différences interindividuelles sont plus importantes que pour les pseudo-mots irréguliers.

2- Expérience 2 : orthographe

Cette deuxième expérience traite des aspects orthographiques. L'intérêt est ici porté sur l'effet de fréquence des bigrammes en fonction de leur position dans un item donné (initiale ou finale). L'objectif est d'étudier le développement des connaissances implicites et explicites sur la fréquence bigrammique auprès de la même population que pour l'expérience 1. La dimension implicite est explorée *via* une tâche de plausibilité lexicale portant sur des pseudo-mots tandis que la dimension explicite est testée grâce à une épreuve d'analyse explicite de bigrammes isolés. Les traitements orthographiques effectués en lecture sont évalués par une épreuve de lecture à haute voix de mots et pseudo-mots. Comme pour la graphophonologie, les enfants de GSM ne passent que l'épreuve implicite.

2.1- Méthode

2.1.1- Population

Les participants à cette deuxième expérience sont les mêmes que ceux de l'expérience 1. Les caractéristiques de la population sont donc similaires (*cf.* paragraphe 1, p. 102).

2.1.2- Matériel

Epreuves orthographiques implicite, explicite et lecture à haute voix

Pour les trois types de tâches (jugement implicite, jugement explicite, lecture), les connaissances orthographiques sont évaluées à partir de mots et pseudo-mots contenant des bigrammes (couples de lettres) fréquents ou rares, en position initiale ou finale.

Les bigrammes retenus en position initiale sont [ca-] et [na-]. Les bigrammes situés en position finale sont [-te] et [-ve]. Ces bigrammes ont été choisis sur la base de leur fréquence d'apparition selon la position dans le mot dans la langue française :

- [ca-] en position initiale a une fréquence d'apparition de 368 (sur 12956 combinaisons de bigrammes possibles en tenant compte de toutes les lettres disponibles) : bigramme fréquent,
- [na-] en position initiale a une fréquence de 63 (sur 12956) : bigramme rare,
- [-te] en position finale a une fréquence de 679 sur (12956) : bigramme fréquent,
- [-ve] en position finale a une fréquence de 81 (sur 12956) : bigramme rare.

Ces bigrammes sont utilisés dans des pseudo-mots pour l'épreuve implicite, de manière isolée pour l'épreuve explicite et enfin dans des mots et pseudo-mots en lecture.

Tâche orthographique implicite de plausibilité lexicale écrite

Elle est constituée de 12 paires d'items expérimentaux et de 4 paires d'items distracteurs en GSM, de 16 paires d'items expérimentaux et de 8 paires d'items distracteurs du niveau CP au niveau CM2 (cf. Annexe 3, pp. 249-250). Chaque paire d'items expérimentaux contient un pseudo-mot à bigramme fréquent et un pseudo-mot à bigramme rare, en position initiale ou finale, soit :

- 6 paires en GSM et 8 paires des niveaux CP à CM2 exploitent les bigrammes [ca-] et [na-] pour la position initiale (*cavure-navure*),
- 6 paires (GSM) et 8 paires (niveaux CP-CM2) concernent les bigrammes [-te] et [-ve] en position finale (*latote-latove*).

Les items distracteurs sont constitués de pseudo-mots construits selon le même principe que les items expérimentaux : dans chaque paire d'items, la base est identique et seule la partie initiale (*latilo-fatilo*) ou finale (*jotase-jotale*) change. Ces items distracteurs ne sont pas pris en compte pour l'analyse des réponses fournies.

Pour les items expérimentaux, le nombre de lettres est de 6 pour les 12 paires proposées en GSM ainsi que pour les 16 paires fournies aux enfants des niveaux CP à CM2. Le nombre de phonèmes est de 6 pour 2 paires, de 5 pour 9 paires et de 4 pour une paire (GSM), de 6 pour 4 paires, de 5 pour 11 paires et de 4 pour 1 paire (niveaux CP-CM2). Le nombre de syllabes est de 2 pour 10 paires et de 3 pour 2 paires en GSM, de 2 pour 12 paires et de 3 pour 4 paires du niveau CP au niveau CM2. Deux paires de pseudo-mots en GSM (*balu-tlpz*, *nufe-mfqc*) et 4 paires pour les enfants des niveaux CP à CM2 (*ibidem*, *ulit-jpjl* et *dori-spkw*) constituent l'entraînement. Le mode de présentation est similaire à celui utilisé pour l'épreuve implicite de graphophonologie (voir Expérience 1).

Tâche orthographique explicite de jugement de patrons orthographiques (bigrammes isolés) (à partir du niveau CP)

Dix paires de bigrammes sont proposées (cf. Annexe 3, pp. 250-251). Une paire contient un bigramme fréquent en position initiale [ca-] (fréquence : 368) *versus* un bigramme rare en position initiale [na-] (fréquence : 63). La seconde paire comporte un bigramme fréquent en position finale [-te] (fréquence : 679) *versus* un bigramme rare en position finale [-ve] (fréquence : 81). Huit paires saillantes sont également proposées pour lesquelles la réponse est évidente (par exemple : *mi-my* pour la position initiale, très peu de mots de la langue française commencent par *my-* ; *ir-yr* pour la position finale, très peu de mots de la langue française se terminant par *-yr*). Deux paires de bigrammes constituent l'entraînement : *ma-mu* (pour la

position initiale : *ma-* fréquent : 284, *mu-* peu fréquent : 56) ; *fe-le* (pour la position finale : *-le* fréquent : 631, *-fe* peu fréquent : 19). Chacune des paires est écrite sur une feuille, en minuscules (police Arial, Taille 28). Les paires expérimentales sont placées après les deux paires d'entraînement.

Tâche de lecture à haute voix (uniquement à partir du niveau CP)

Le matériel expérimental est constitué de 32 items à lire (*cf.* Annexe 3, pp. 251-253), soit :

- 8 mots à bigramme fréquent dont 4 en position initiale [ca-] (fréquence : 368) et 4 en position finale [-te] (fréquence : 679) (*canard, tomate*),
- 8 mots à bigramme rare dont 4 en position initiale [na-] (fréquence : 63) et 4 en position finale [-ve] (fréquence : 81) (*nature, lessive*),
- 8 pseudo-mots à bigramme fréquent : 4 en position initiale [ca-] et 4 en position finale [-te] (*caloude, bofite*),
- 8 pseudo-mots à bigramme rare : 4 en position initiale [na-] et 4 en position finale [-ve] (*nalube, binouve*).

Les items sont composés de 6 à 8 lettres, 5 à 7 phonèmes et de 2 syllabes. Comme pour l'épreuve graphophonologique de lecture, la fréquence des bigrammes et des trigrammes composant ces items a été prise en compte ainsi que la fréquence lexicale des mots. Les analyses statistiques concernant les bigrammes, trigrammes, lettres, phonèmes, syllabes et fréquence lexicale apparaissent en annexe (*cf.* Annexe 3, pp. 252-253). Les quatre items d'entraînement sont : *cour, bilato, table, marode*. La présentation des items est similaire à celle utilisée pour l'épreuve de lecture en graphophonologie.

2.1.3- Procédure

Epreuves orthographiques implicite, explicite et lecture à haute voix

Tâche orthographique implicite de plausibilité lexicale écrite

La passation est individuelle pour les GSM et semi-collective du niveau CP au niveau CM2. Avant de commencer, il est précisé aux enfants que leur tâche consiste à bien regarder chaque paire de pseudo-mots et à montrer (GSM) ou à entourer (du niveau CP au niveau CM2) le pseudo-mot de chaque paire qui ressemble le plus à un vrai mot. En GSM, les feuilles sur lesquelles sont inscrits les items sont présentées une à une, devant le participant, au fur et à mesure des réponses fournies. Au niveau de la cotation pour les GSM, la note globale est de 12 (deux sous notes : bigrammes en position initiale sur 6, bigrammes en position finale sur 6). Du niveau CP au CM2, les numéros des pages du carnet correspondant à chaque paire sont

prononcés pour éviter les oublis d'items et favoriser la considération attentive de chaque paire de pseudo-mots. Pour ces cinq niveaux de lecture, la tâche est notée sur 16 (bigrammes en position initiale sur 8, bigrammes en position finale sur 8).

Pour les enfants de GSM, la consigne est la suivante : « *Je vais te montrer des feuilles. Sur chaque feuille, deux mots sont écrits. Ces mots ne sont pas de vrais mots : ils n'existent pas et ils ne veulent rien dire. Regarde bien les deux mots et montre-moi celui qui ressemble le plus à un vrai mot. Tu as bien compris ? Tu n'as pas de questions ?* ».

La consigne fournie aux enfants du niveau CP au CM2 est : « *Je vais vous distribuer un petit carnet. Sur chaque page du carnet, deux mots sont écrits. Ces mots ne sont pas de vrais mots : ils n'existent pas et ils ne veulent rien dire. Regardez bien les deux mots et entourez avec un crayon celui qui ressemble le plus à un vrai mot. Quand vous avez fini d'entourer pour la première page, vous attendez que je vous dise de tourner la page pour continuer. Attention, vous n'avez pas le droit de revenir en arrière. Vous avez bien compris ? Vous n'avez pas de questions ?* ».

Tâche orthographique explicite de jugement de patrons orthographiques

La passation est individuelle pour l'ensemble des participants. Les deux paires d'entraînement sont proposées puis débute la phase expérimentale. Dans un premier temps, les bigrammes en position initiale sont présentés selon un ordre aléatoire, puis c'est au tour des bigrammes en position finale. Les feuilles sur lesquelles sont inscrites les paires sont présentées une à une, devant le participant, au fur et à mesure des réponses fournies. Le participant ne reçoit aucun feedback suite à ses réponses. L'enfant peut fournir sa réponse oralement ou la désigner du doigt. Un point est attribué par réponse correcte. Cette épreuve est ainsi notée sur deux points (un point pour les bigrammes en position initiale et un point pour ceux en position finale).

La consigne proposée aux enfants est la suivante: « *Regarde bien ces petits morceaux de mots. Je voudrais que tu me dises s'il y a plus de mots qui commencent (pour les bigrammes en position initiale) / qui se terminent (pour les bigrammes en position finale) par ce morceau là « ... » (montrer au participant le premier bigramme de la paire) ou par celui-ci « ... » (montrer au participant le second bigramme de la paire). Si tu ne sais pas, ce n'est pas grave : il suffit que tu dises « je ne sais pas » (la réponse « je ne sais pas » est acceptée).*

Tâche de lecture à haute voix

La procédure et la consigne fournie aux participants sont identiques à celles de l'épreuve graphophonologique. Un point est attribué par mot lu correctement dans chacune des catégories, (mots à bigramme fréquent, pseudo-mots à bigramme fréquent, mots à bigramme rare et pseudo-mots à bigramme rare).

Pour toutes ces tâches orthographiques, l'ordre de passation a été également aléatorisé.

2.2- Hypothèses opérationnelles

Epreuve implicite : les items construits avec un bigramme fréquent, en position initiale ou finale dans l'item (*calufe* ; *latote*), devraient être plus souvent choisis comme ressemblant à de vrais mots que ceux construits avec un bigramme rare (*nalufe* ; *latove*). Les bigrammes en position initiale pourraient être plus faciles à repérer par les enfants par rapport à ceux en position finale. Les enfants pourraient alors être plus performants face aux paires à bigramme en position initiale (*calufe-nalufe*) vs. les paires à bigramme en position finale (*latote-latove*). Les préférences pour les items à bigrammes fréquents devraient s'accroître avec le niveau de lecture des enfants.

Epreuve explicite : les bigrammes fréquents (*ca* ; *te*) devraient être choisis de manière préférentielle aux bigrammes rares (*na* ; *ve*). Ce choix préférentiel serait accentué dans le cas du bigramme en position initiale (*ca*). Ces résultats pourraient être d'autant plus marqués que les enfants sont efficaces en lecture.

Les connaissances orthographiques explicites sont fortement dépendantes de l'apprentissage de la lecture. Elles devraient donc se développer plus tardivement que les connaissances implicites. Ainsi, l'épreuve explicite devrait être moins bien réussie que celle implicite.

Lecture à haute voix : les pseudo-mots (*narufe*) seraient lus moins vite et avec davantage d'erreurs que les mots (*conserve*). Les items à bigrammes fréquents (*camion*, *tomate*) devraient être mieux lus que ceux à bigrammes rares (*nalube*, *firanve*). Les effets de lexicalité et de fréquence du bigramme augmenteraient avec le niveau de lecture des enfants.

2.3- Résultats

2.3.1- Tâches orthographiques implicite et explicite

Les performances aux épreuves implicite et explicite d'orthographe sont analysées ci-après. Les scores obtenus à ces épreuves sont regroupés dans le Tableau 7.

Tableau 7. Scores (pourcentages de réussite³) obtenus aux deux épreuves orthographiques (et *écart-types*) par l'ensemble des enfants selon la position du bigramme

	Orthographe implicite			Orthographe explicite		
	Initiale	Finale	Moyenne	Initiale	Finale	Moyenne
GSM	48,89 (18,01)	51,11 (19,54)	50			
Niveau CP	55,24 (21,11)	54,84 (18,45)	55,04	67,74 (47,52)	54,84 (50,59)	61,29
Niveau CE1	63,79 (19,58)	53,45 (18,87)	58,62	68,97 (47,08)	51,72 (50,85)	60,34
Niveau CE2	66,80 (17,85)	68,75 (21,53)	67,77	53,13 (50,70)	81,25 (39,66)	67,19
Niveau CM1	58,50 (20,95)	52 (22,73)	55,25	52 (50,99)	80 (40,82)	66
Niveau CM2	69,14 (17,67)	58,98 (18,58)	64,06	56,25 (50,40)	65,63 (48,26)	60,94
Moyenne	60,39	56,52	58,46	59,62	66,69	63,15

La comparaison des deux épreuves orthographiques (implicite *vs.* explicite) est envisageable ici car dans les deux cas, le hasard en termes de réponses fournies est de 50 %. Comme pour l'épreuve graphophonologique explicite, les enfants de GSM n'ont pas passé la tâche explicite d'orthographe. La comparaison des scores aux deux épreuves ne porte donc que sur les enfants des niveaux CP à CM2. Les analyses suivantes mettent en jeu deux facteurs : le niveau de lecture (cinq modalités : CP, CE1, CE2, CM1, CM2) et la nature de l'épreuve (implicite *vs.* explicite). Le plan est : $S_n < N_5 > * E_2$ où S désigne le facteur sujets, N le facteur niveau de lecture des enfants et E le type ou la nature de l'épreuve utilisée.

Les analyses indiquent que l'effet du niveau de lecture n'est pas significatif ($F(4, 144) = 1,09$; *ns*). Les enfants des différents niveaux de lecture ne se distinguent pas au niveau de leurs performances aux épreuves orthographiques (voir Tableau 7 pour les différents scores). L'effet du facteur nature de l'épreuve (implicite *vs.* explicite) n'est pas non plus significatif ($F(1, 144) = 1,01$; *ns*). L'épreuve implicite (60,15 %) n'est pas mieux réussie que l'épreuve explicite (63,15 %) par les enfants. L'interaction entre ces deux facteurs n'est pas significative ($F(4, 144) < 1$).

Tâche orthographique implicite

Pour cette épreuve implicite, chaque paire de pseudo-mots est constituée d'un item contenant un bigramme fréquent et d'un item à bigramme rare, en position initiale pour huit paires et en position finale pour les huit paires restantes. De façon similaire à la

³ Par convention, le choix de l'item incluant le bigramme fréquent dans l'épreuve implicite est qualifié de « réussite » alors qu'en toute rigueur il n'y a pas de réussite dans cette épreuve, le choix exprimant une préférence.

graphophonologie, il est demandé aux enfants de choisir entre les deux pseudo-mots celui qui ressemble le plus à un vrai mot. Pour cette tâche implicite, les écart-types sont assez importants (cf. Tableau 7). Pour chaque niveau de lecture, il existe de fortes différences interindividuelles pour les items à bigramme en position initiale ou finale.

Deux facteurs principaux sont testés ici : le niveau de lecture (six modalités : GSM, niveaux de lecture CP, CE1, CE2, CM1, CM2), et la position du bigramme dans le mot (initiale vs. finale). Le plan d'analyse est alors $S_n < N_6 > * P_2$. Le facteur S correspond aux sujets, le facteur N correspond au niveau de lecture et enfin le facteur P désigne la position du bigramme.

Les enfants de GSM fournissent des réponses qui ne se distinguent pas du hasard. Pour ce groupe, les items à bigramme fréquent ne sont donc pas choisis préférentiellement aux items à bigramme rare, en position initiale comme en finale. Le même profil est obtenu chez les enfants de niveau CP pour les items à bigramme en position initiale et en position finale. Cependant, leurs performances totales (positions initiale et finale confondues) se distinguent du hasard ($p < .05$). Pour les enfants de niveau CE1, les réponses sont différentes du hasard pour les items à bigramme en position initiale ($p < .001$). Bien que leurs réponses ne se distinguent pas du hasard pour les items à bigramme en position finale, la différence est significative pour le total de leurs réponses ($p < .01$). Pour le niveau CM1, seuls les scores aux items à bigramme en position initiale diffèrent du hasard ($p < .05$). Enfin, les enfants de niveau CE2 comme ceux de niveau CM2 choisissent préférentiellement les items contenant un bigramme fréquent⁴ (CE2 : $p < .0001$; CM2 : $p < .0001$). Globalement dans 58,46 % des cas seulement, les items avec un bigramme fréquent (*cavure* pour la position initiale ; *latote* pour la position finale) sont choisis préférentiellement aux items avec un bigramme rare (*navure* ; *latove*).

Les analyses réalisées sur les scores obtenus à cette épreuve implicite indiquent un effet significatif du facteur niveau de lecture ($F(5, 173) = 6,71$; $p < .0001$; voir Figure 8). Plus les enfants sont d'un niveau de lecture élevé, plus ils choisissent l'item comportant un bigramme fréquent. Ce résultat doit cependant être nuancé car, en fonction du niveau, le choix préférentiel pour l'item avec bigramme fréquent diffère. Ainsi, les GSM (50 %) ont des performances significativement inférieures aux enfants des niveaux CE1 (58,62 % ; $F(1, 57) = 5,73$; $p < .05$), CE2 (67,77 % ; $F(1, 60) = 25,71$; $p < .0001$) et CM2 (64,06 % ; $F(1, 60) = 18,09$; $p < .0001$). Les enfants de niveau CP (55,04 %) ont également des scores significativement inférieurs aux enfants des niveaux CE2 ($F(1, 61) = 13,15$; $p < .001$) et CM2 ($F(1, 61) = 7,39$; $p < .01$), mais ne se différencient pas des enfants de GSM ($F(1, 59) = 2,19$; *ns*) et des enfants de niveau CE1 ($F(1, 58) < 1$). Les enfants de niveau CE1 réussissent moins bien l'épreuve que les enfants de niveau

⁴ Les réponses sont différentes du hasard que le bigramme soit en position initiale ou finale.

CE2 ($F(1, 69) = 6,11 ; p < .05$). Les enfants de niveau CM1 (55,25 %) ne sont pas aussi performants à l'épreuve implicite que les enfants des niveaux CE2 ($F(1, 55) = 9,88 ; p < .01$) et CM2 ($F(1, 55) = 5,45 ; p < .05$). Ces mêmes enfants ne se distinguent pas des enfants des niveaux CE1 ($F(1, 52) < 1$), CP ($F(1, 54) < 1$) et de GSM (50 % ; $F(1, 53) = 1,84 ; ns$). Enfin, les enfants de niveau CM2 ne se distinguent pas des enfants des niveaux CE2 ($F(1, 62) = 1,18 ; ns$) et CE1 ($F(1, 59) = 2,40 ; ns$).

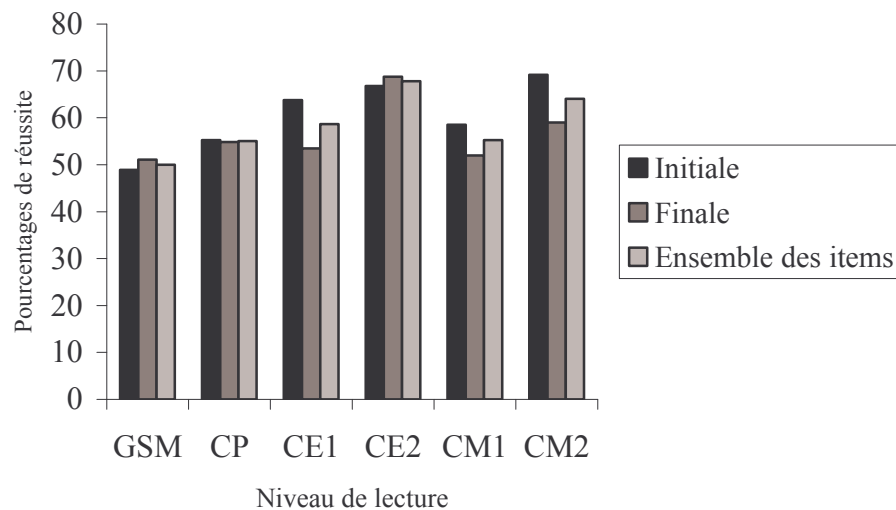


Figure 8. Pourcentages de choix de l'item à bigramme fréquent (positions initiale et finale) chez les enfants de GSM au niveau de lecture CM2

Les items utilisés dans cette épreuve contiennent des bigrammes fréquents et rares en position initiale ou finale. L'effet de la position du bigramme est tendanciellement significatif ($F(1, 173) = 3,55 ; p = .061$). Les items avec un bigramme fréquent en position initiale tendent à être plus souvent choisis que ceux avec un bigramme fréquent en position finale (60,39 % > 56,52 %). Les deux facteurs (niveau de lecture et position du bigramme) n'interagissent pas de manière significative ($F(5, 173) = 1,39 ; ns$).

Tâche orthographique explicite

L'ensemble des enfants (sauf ceux scolarisés en GSM) doit ici choisir entre deux bigrammes en fonction de sa fréquence ou de sa rareté en début ou en fin de mot. Comparativement à l'épreuve implicite, les écart-types sont ici très élevés. Les résultats obtenus au test explicite sont très hétérogènes et ce, pour chacun des niveaux de lecture (CP, bigramme en position initiale : 47,52 ; en position finale : 50,59 ; CE1 : 47,08 et 50,85 ; CE2 : 50,70 et 39,66 ; CM1 : 50,99 et 40,82 ; CM2 : 50,40 et 48,26).

Les analyses réalisées sur les performances concernent deux facteurs principaux : le niveau de lecture (cinq modalités : CP, CE1, CE2, CM1, CM2) et la position du bigramme dans le mot (initiale vs. finale). Le plan d'analyse est alors $S_n < N_5 > * P_2$ (S : facteur sujet ; N : facteur niveau de lecture ; P : facteur position du bigramme).

Cette épreuve est mieux réussie que celle implicite : 63,15 % de bigrammes fréquents, quelle que soit la position, sont correctement détectés (cf. Tableau 7). Cette fois-ci, les réponses des enfants des niveaux CP et CE1 se distinguent du hasard seulement pour les bigrammes en position initiale ($p < .05$ pour les deux niveaux). Pour les niveaux du CE2 au CM2, les scores concernant ces bigrammes spécifiques ne se distinguent pas du hasard. Seules les réponses fournies pour les bigrammes en position finale diffèrent du hasard ($p < .05$) chez les enfants de niveau CM2. Quant aux enfants des niveaux CM1 et CE2, ils choisissent préférentiellement les bigrammes fréquents lorsqu'ils sont en position finale ($p < .001$ dans les deux cas). Le total des réponses pour ces mêmes enfants se distinguent du hasard également (respectivement $p < .05$ et $p < .01$).

En outre, les facteurs niveau de lecture et position du bigramme n'ont pas d'effet significatif sur les scores obtenus ($F(4, 144) < 1$ et $F(4, 144) = 1,72$; *ns*). L'interaction entre ces deux facteurs est significative ($F(4, 144) = 3,19$; $p < .02$). L'effet de position du bigramme est significatif pour les enfants des niveaux CM1 ($p = .0263$) et CE2 ($p = .0111$). Pour ces deux niveaux, les bigrammes en position initiale sont nettement moins bien réussis que ceux en position finale (CM1 : 52 % < 80 % ; CE2 : 53,13 % < 81,25 % ; voir Figure 9). Pour les autres niveaux, les différences ne sont pas significatives.

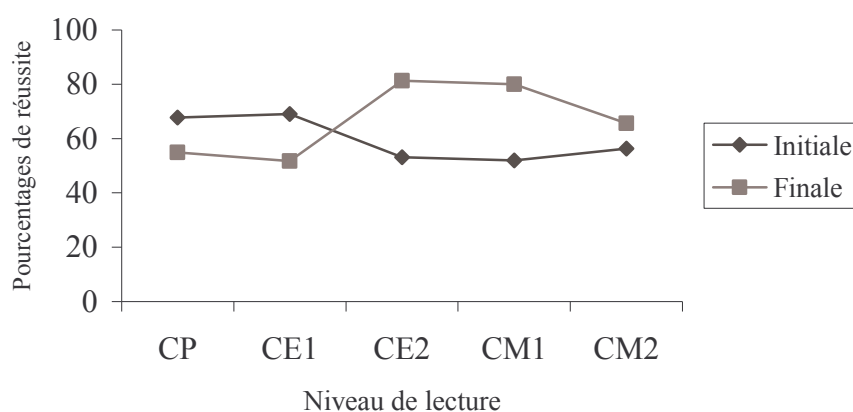


Figure 9. Pourcentages de bigrammes fréquents correctement détectés en fonction de la position dans un mot et du niveau de lecture

2.3.2- Tâche de lecture à haute voix

Les items proposés en lecture sont de deux types : bigramme fréquent ou rare, en position initiale ou position finale. Les résultats sont présentés en tenant compte de cette distinction. Ils concernent l'exactitude en lecture et les temps de lecture correcte en fonction de trois facteurs : le niveau de lecture (cinq modalités : CP, CE1, CE2, CM1, CM2), la lexicalité des items (mots *vs.* pseudo-mots) et la fréquence du bigramme (fréquent *vs.* rare). Le plan d'analyse pour les scores et les temps de lecture (bigrammes en position initiale et en position finale) est le suivant : $S_n < N_5 > * L_2 * F_2$. Le facteur S désigne les sujets, le facteur N correspond au niveau de lecture, L correspond à la lexicalité des items et F désigne la fréquence bigrammique.

2.3.2.1- Position initiale

Exactitude

Les performances obtenues par l'ensemble des enfants aux mots et pseudo-mots à bigramme fréquent et rare figurent dans le Tableau 8 ci-après.

Tableau 8. Scores (pourcentages de réussite) (et *écart-types*) obtenus aux items à bigramme en position initiale selon le niveau de lecture des enfants

Niveau de lecture	Mots bigramme fréquent	Mots bigramme rare	Pseudo-mots bigramme fréquent	Pseudo-mots bigramme rare
CP	57,26 (23,45)	66,94 (23,62)	62,10 (27,29)	58,87 (27,80)
CE1	93,97 (10,89)	90,52 (14,04)	80,17 (19,34)	75,86 (26,29)
CE2	96,09 (11,20)	95,31 (9,91)	86,72 (21,05)	88,28 (16,78)
CM1	99 (5,00)	96 (9,35)	90 (12,50)	87 (17,85)
CM2	99,22 (4,42)	94,53 (12,27)	88,28 (17,94)	92,19 (16,11)
Moyenne	89,11	88,66	81,45	80,44

La dispersion des résultats est importante pour les lecteurs de niveau CP pour chacun des items (mots à bigramme fréquent : 23,45 ; à bigramme rare : 23,62 ; pseudo-mots à bigramme fréquent : 27,29 ; à bigramme rare : 27,80). Pour les autres niveaux de lecture, les écart-types sont plus faibles, notamment pour la lecture des mots à bigrammes fréquent et rare. Pour les pseudo-mots à bigrammes fréquent et rare, les résultats sont assez hétérogènes chez les enfants de niveau CE2 et CM2 (respectivement 21,05 et 16,78 pour le CE2 ; 17,94 et 16,11 pour le CM2). Chez les enfants de niveau CM1, les performances en lecture de pseudo-mots à bigramme rare (17,85) sont hétérogènes par rapport aux autres catégories d'items.

Les analyses effectuées sur les items à bigramme en position initiale indiquent un effet significatif des facteurs niveau de lecture ($F(4, 144) = 44,73 ; p < .0001$) et lexicalité ($F(1, 144)$

= 26,70 ; $p < .0001$). Comme le montre la Figure 10, les enfants de niveau CP (61,29 %) lisent correctement moins d'items que les enfants des autres niveaux (CE1 : 85,13 % ; $p = .0032$, CE2 : 91,60 % ; $p < .00$, CM1 : 93 % ; $p < .00$, CM2 : 93,55 % ; $p < .0001$).

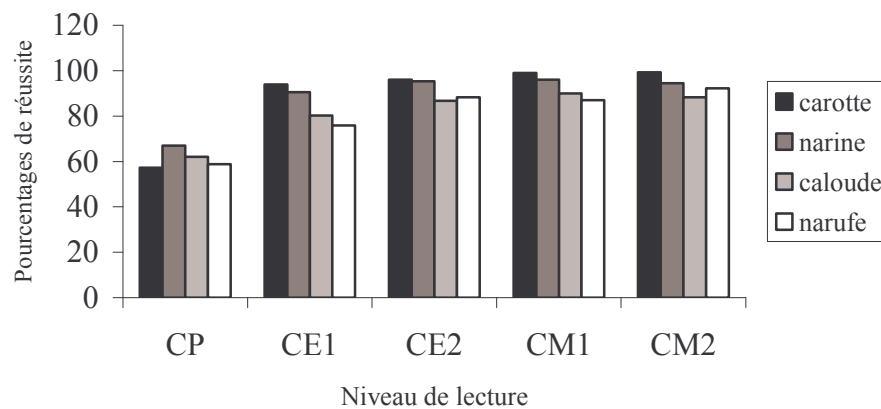


Figure 10. Items à bigrammes en position initiale lus correctement en fonction du niveau de lecture (pourcentages de réussite)

De plus, les enfants des niveaux CM2 et CM1 lisent mieux que les enfants de niveau CE1. La différence entre niveau CM2 et niveau CE1 est significative ($p = .0284$) mais elle n'est que tendanciellement significative pour les niveaux CM1 et CE1 ($p = .0538$). Les performances des enfants des autres niveaux ne se distinguent pas de façon significative. En outre, les analyses montrent que les enfants, tous groupes confondus, lisent moins bien les pseudo-mots comme *caloude* ou *narufe* (80,94 %) que les mots tels que *carotte* ou *narine* (87,88 %). Le facteur fréquence du bigramme n'a pas d'effet significatif sur les scores en lecture ($F(1, 144) < 1$) : les items à bigramme fréquent (*carotte*, *caloude*) ne sont pas mieux lus que ceux à bigramme rare en position initiale (*narine*, *narufe*). Enfin, aucune interaction n'est significative.

Temps

De façon similaire à la graphophonologie, seuls les temps de lecture correcte ont fait l'objet d'une analyse. L'effectif est donc réduit de 21 enfants par rapport à l'effectif initial. Les temps de lecture correcte de 17 enfants de niveau CP, 26 enfants de niveau CE1, 30 enfants de niveau CE2, 25 enfants de niveau CM1 et 30 enfants de niveau CM2 ont ainsi été analysés.

Ces temps moyens de lecture des items à bigramme en position initiale sont regroupés dans le Tableau 9. Ils sont exprimés en millisecondes (ms).

Tableau 9. Temps moyens de lecture correcte (et *écart-types*) des items à bigramme en position initiale selon le niveau de lecture

Niveau de lecture	Mots bigramme fréquent	Mots bigramme rare	Pseudo-mots bigramme fréquent	Pseudo-mots bigramme rare
CP	4012,70 (1954,77)	3678,33 (1167,91)	4261,29 (1470)	4309,76 (1659,59)
CE1	2457,97 (628,17)	2624,50 (505,62)	3297,53 (976,39)	3004,64 (762,04)
CE2	1837,72 (284,09)	1929,23 (299,02)	2454,06 (673,18)	2459,22 (587,76)
CM1	1598,90 (225,66)	1693,88 (263,47)	2004,11 (472,58)	1944,18 (425,60)
CM2	1674,72 (305,11)	1728,22 (317,76)	1956,12 (410,44)	1971,80 (373,35)
Moyenne	2316,40	2330,83	2794,62	2737,92

D'une manière générale, les écart-types les plus élevés sont ceux des enfants de niveau CP. Pour chaque catégorie d'items, les temps de lecture correcte sont très hétérogènes (mots à bigramme fréquent : 1954,77 ms ; mots à bigramme rare : 1167,91 ms ; pseudo-mots à bigramme fréquent : 1470 ms ; pseudo-mots à bigramme rare : 1659,59 ms).

Le niveau de lecture a un effet significatif sur les temps de lecture correcte des items ($F(4, 123) = 42,87 ; p < .0001$). Les enfants de niveau CP (4065,52 ms) sont plus lents que ceux des niveaux CE1 (2846,16 ms), CE2 (2170,06 ms), CM1 (1810,27 ms) et CM2 (1832,72 ms) ; les différences sont alors toutes significatives à .001). Les enfants de niveau CE2 lisent plus lentement que les niveaux CM1 ($p = .0022$) et CM2 ($p = .0023$) tandis que ces deux derniers groupes ne diffèrent pas significativement. L'effet du facteur lexicalité est également significatif : les mots sont lus plus rapidement que les pseudo-mots (2323,62 ms < 2766,27 ms ; $F(1, 123) = 133,01 ; p < .0001$). En revanche, la différence entre les items à bigramme fréquent (2555,51 ms) et à bigramme rare (2534,38 ms) n'est pas significative ($F(1, 123) < 1$).

L'interaction entre les facteurs niveau de lecture et lexicalité est significative ($F(1, 123) = 3,52 ; p < .01$). La Figure 11 révèle que pour le niveau CP la différence de temps de lecture entre les mots (3845,52 ms) et les pseudo-mots (4285,53 ms) n'est pas significative ($p = .3820$). Pour tous les autres niveaux, les mots sont lus plus rapidement que les pseudo-mots (CM2 : $p = .0038$; CM1 : $p = .0014$; CE2 : $p < .0001$; CE1 : $p = .0027$). Cette figure montre aussi que si les enfants de niveau CM2 lisent effectivement les pseudo-mots plus vite que les enfants de niveau CE2 (1963,96 ms < 2456,64 ms ; $p = .0002$), la différence entre ces deux niveaux n'est en revanche que tendanciellement significative ($p = .0640$) sur les mots (respectivement : 1701,47 ms et 1883,48 ms).

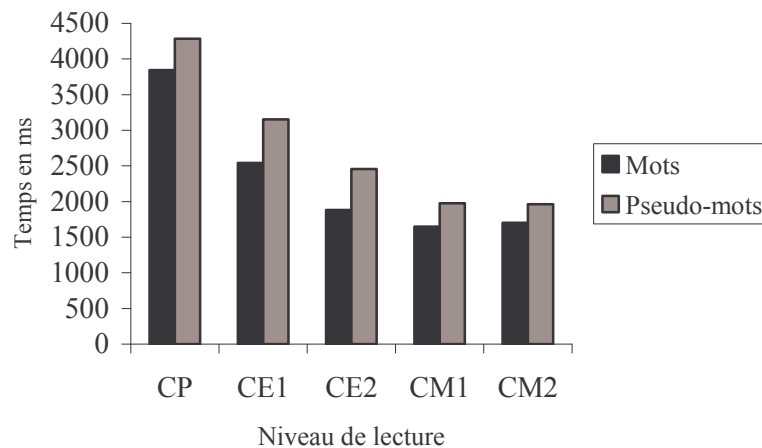


Figure 11. Temps de lecture correcte des mots et pseudo-mots à bigramme en position initiale selon le niveau de lecture

Les facteurs niveau de lecture, lexicalité et fréquence du bigramme interagissent ($F(4, 123) = 3,49$; $p < .01$; cf. Figure 12). L'analyse de cette interaction confirme pourtant que l'effet de la lexicalité existe à chaque niveau de lecture à la fois pour les items à bigramme rare (CM2 : $p = .0086$; CM1 : $p = .0159$; CE2 : $p < .0001$; CE1 : $p = .0390$) et ceux à bigramme fréquent (CM2 : $p = .0038$; CM1 : $p = .0003$; CE2 : $p < .0001$; CE1 : $p = .0006$) sauf au niveau CP. Elle témoigne également de l'absence d'effet de la fréquence du bigramme pour les mots comme pour les pseudo-mots et ce pour l'ensemble des enfants. L'interaction réaffirme l'effet du niveau de lecture pour toutes les catégories d'items : pour les mots et les pseudo-mots à bigramme fréquent ou rare, les différences entre les niveaux de lecture sont significatives sauf entre le CM2 et le CM1 ($p < .05$ entre le CM2 et le CE2 ; $p < .01$ pour le reste).

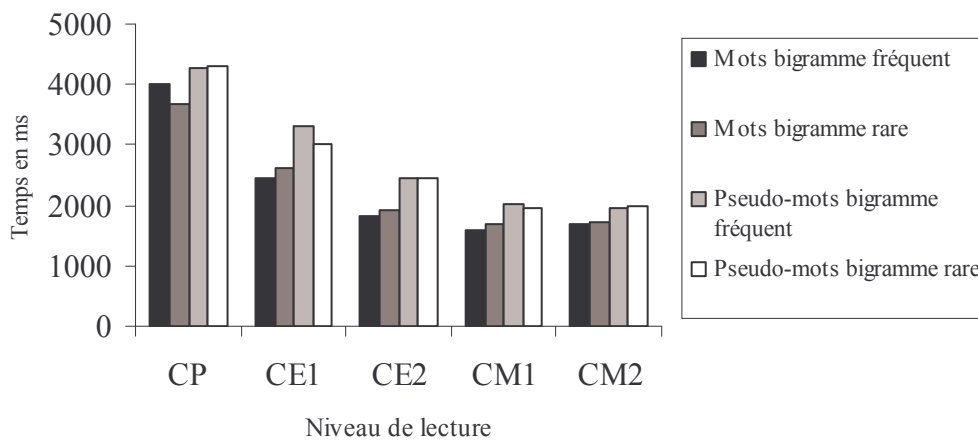


Figure 12. Temps de lecture correcte des items pour les cinq groupes d'enfants en fonction de la lexicalité et de la fréquence du bigramme

2.3.2.2- Position finale

Exactitude

Les scores concernant les items à bigramme fréquent et rare en position finale obtenus par les enfants du niveau CP au niveau CM2 sont regroupés dans le Tableau 10.

Tableau 10. Pourcentages de réussite (et *écart-types*) obtenus aux items à bigramme en position finale par les enfants des cinq groupes de niveau de lecture

Niveau de lecture	Mots bigramme fréquent	Mots bigramme rare	Pseudo-mots bigramme fréquent	Pseudo-mots bigramme rare
CP	58,06 (26,13)	59,68 (26,36)	52,42 (25,29)	66,94 (28,42)
CE1	87,93 (17,19)	82,76 (19,02)	75 (24,09)	75,86 (17,01)
CE2	94,53 (12,27)	94,53 (12,27)	82,03 (22,21)	78,13 (20,82)
CM1	97 (8,29)	97 (8,29)	83 (23,63)	78 (20,82)
CM2	98,44 (6,15)	98,44 (6,15)	86,72 (21,05)	85,94 (15,47)
Moyenne	87,19	86,48	75,83	76,97

Le profil de dispersion des performances pour les items à bigramme en position finale est très similaire à celui obtenu avec les items à bigramme en position initiale. En effet, les résultats sont hétérogènes chez les enfants de niveau CP pour l'ensemble des items (mots à bigramme fréquent : 26,13 ; à bigramme rare : 26,36 ; pseudo-mots à bigramme fréquent : 25,29 ; à bigramme rare : 28,42). Pour les autres niveaux de lecture, la dispersion des performances est moins importante sauf pour les pseudo-mots à bigramme fréquent (CE1 : 24,09 ; CE2 : 22,21 ; CM1 : 23,63 et CM2 : 21,05) et les pseudo-mots à bigramme rare pour les enfants de niveau CE2 et CM1 (20,82 dans les deux cas).

Pour les items à bigramme en position finale, le niveau de lecture a un effet significatif sur les performances ($F(4, 144) = 44,73$; $p < .0001$). Les enfants de niveau CP sont moins performants en lecture que les enfants des autres niveaux (CP : 59,27 % vs. CE1 : 80,39 % ; $p = .0011$, CE2 : 87,30 % ; $p < .0001$, CM1 : 88,75 % ; $p < .0001$, CM2 : 92,38 % ; $p < .0001$). Les enfants de niveau CE1 lisent également moins bien que ceux du niveau CM2 ($p = .0054$). Pour les autres niveaux, aucune différence n'est significative. Le facteur lexicalité est significatif ($F(1, 144) = 55,04$; $p < .0001$). Les enfants font davantage d'erreurs sur les pseudo-mots que sur les mots (76,40 % vs. 86,84 %).

L'effet du facteur lexicalité varie en fonction du niveau de lecture ($F(4, 144) = 4,65$; $p < .01$). En effet, comme le montre l'interaction (voir Figure 13), les mots (58,87 %) ne sont pas mieux lus que les pseudo-mots (59,68 %) chez les enfants de niveau CP. Cependant, pour les autres niveaux, les mots sont lus avec moins d'erreurs que les pseudo-mots (CM2 : 98,44 % >

86,33 % ; $p < .0001$, CM1 : 97 % > 80,50 % ; $p < .0001$, CE2 : 94,53 % > 80,08 % ; $p < .0001$, CE1 : 85,34 % > 80,08 % ; $p = .0074$).

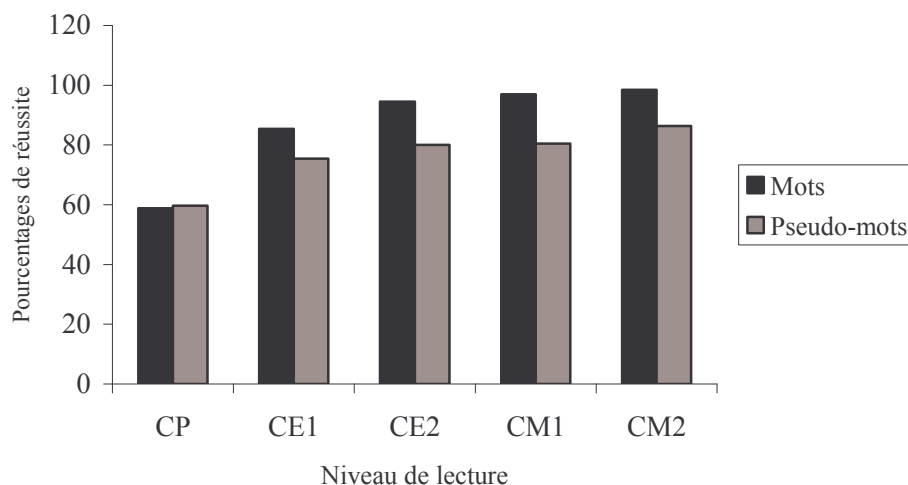


Figure 13. Performances obtenues aux items selon la lexicalité et le niveau de lecture

De plus, l'interaction entre les facteurs niveau de lecture et lexicalité montre que si, pour les pseudo-mots, la différence entre les scores obtenus par les enfants des niveaux CM2 (86,33 %) et CE2 (80,08 %) n'est pas significative, elle l'est avec les mots ($p = .0350$). Dans ce cas, le niveau CM2 (98,44 %) lit mieux que le niveau CE2 (94,53 %). Les enfants des niveaux CM1 ($p < .0001$) et CE2 ($p = .0003$) lisent également en faisant moins d'erreurs sur les mots que les enfants de niveau CE1. Concernant le facteur fréquence du bigramme, les scores aux items à bigramme fréquent (81,51 %) et ceux aux items à bigramme rare en position finale (81,73 %) ne diffèrent pas significativement ($F(1, 144) < 1$). Ce facteur n'interagit ni avec le niveau de lecture ($F(4, 144) = 1,46$; ns), ni avec la lexicalité ($F(1, 144) < 1$). L'interaction triple des facteurs n'est pas non plus significative ($F(4, 144) = 1,28$; ns).

Temps

L'analyse qui suit a été réalisée sur les temps de lecture correcte pour les items à bigramme en position finale obtenus par 123 enfants (CM2 : 31 enfants ; CM1 : 23 ; CE2 : 29 ; CE1 : 25 ; CP : 15).

Le Tableau 11 reprend ces différents temps moyens exprimés en ms pour les cinq groupes de niveaux de lecture.

Tableau 11. Temps moyens de lecture correcte (et *écart-types*) des items à bigramme en position finale obtenus par les enfants des niveaux CP à CM2

Niveau de lecture	Mots bigramme fréquent	Mots bigramme rare	Pseudo-mots bigramme fréquent	Pseudo-mots bigramme rare
CP	4546,29 (1827,35)	4447,61 (1798,63)	4277,76 (1647,68)	4339,48 (1606,59)
CE1	2643,88 (581,82)	2820,20 (849,17)	3278,07 (1518,80)	3187,22 (739,58)
CE2	2115,20 (388,15)	2086,96 (332,04)	2568,30 (648,66)	2753,67 (653,42)
CM1	1679,80 (285,32)	1736,24 (268,01)	1966,85 (323,15)	2096,81 (398,16)
CM2	1731,21 (332,39)	1719,73 (317,76)	2011,98 (410,44)	2153,28 (373,35)
Moyenne	2543,27	2562,15	2820,59	2906,09

Comme pour les items à bigramme en position initiale, les temps de lecture les plus hétérogènes concernent les enfants de niveau CP (mots à bigramme fréquent : 1827,35 ms ; à bigramme rare : 1798,63 ms ; pseudo-mots à bigramme fréquent : 1647,68 ms ; à bigramme rare : 1606,59 ms). Les enfants de niveau CE1 présentent également des temps de lecture avec une forte dispersion pour les pseudo-mots à bigramme fréquent (1518,80 ms).

Les analyses indiquent à nouveau un effet du niveau de lecture ($F(4, 118) = 48,78 ; p < .0001$). Le niveau CP (4402,78 ms) met davantage de temps que le niveau CE1 (2982,34 ms) à lire correctement un item ($p = .0016$). Comme le montre la Figure 14, les enfants de niveau CE2 (2381,03 ms) sont plus rapides que ceux des niveaux CE1 ($p = .0039$) et CP ($p < .0001$).

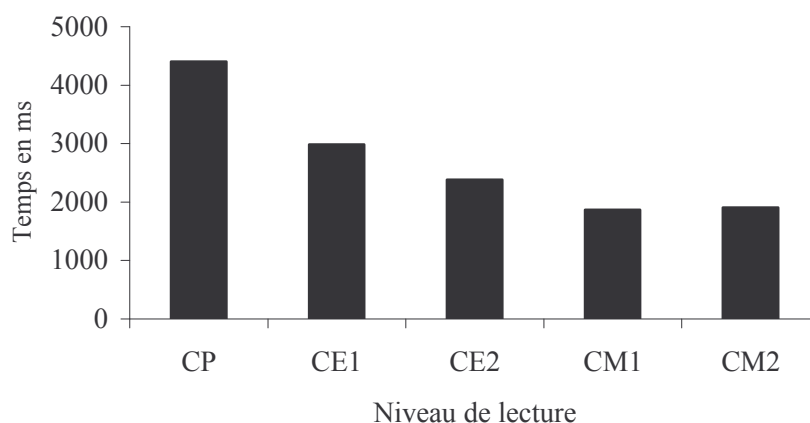


Figure 14. Temps de lecture correcte pour l'ensemble des items selon le niveau de lecture

Les enfants des niveaux CM2 (1904,05 ms) et CM1 (1869,92 ms) lisent plus rapidement que les enfants des autres groupes (CE2 : $p = .0001$ avec les niveaux CM2 et CM1 ; CE1 : $p < .0001$ avec les deux groupes ; CP : $p < .0001$ avec les deux niveaux). Les enfants lisent plus vite les mots (comme *tomate*) que les pseudo-mots (tel que *bofite*). Les mots sont lus en moyenne en

2552,71 ms contre 2863,34 ms pour les pseudo-mots ($F(1, 118) = 24,11 ; p < .0001$). Les items à bigramme fréquent en position finale (*cravate, possute* : 2681,83 ms) ne sont pas lus avec une plus grande rapidité que les items à bigramme rare (*lessive, toruve* : 2734,12 ms ; $F(1, 118) = 1,17 ; ns$).

Les facteurs niveau de lecture et lexicalité interagissent significativement ($F(4, 118) = 3,36 ; p < .05$, cf. Figure 15).

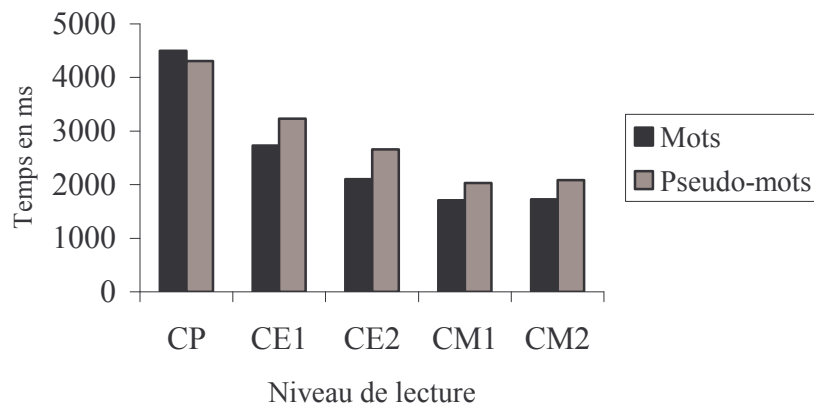


Figure 15. Temps moyens mis à lire les mots et les pseudo-mots en fonction du niveau de lecture

Cette interaction confirme que les pseudo-mots sont lus moins vite que les mots chez tous les enfants sauf pour ceux de niveau CP. Chez ces enfants, la différence entre les mots (4496,95 ms) et les pseudo-mots (4308,62 ms) n'est pas significative (voir Figure 15).

2.4- Discussion

Cette deuxième expérience qui traitait des aspects orthographiques fournit des résultats quelque peu différents de ceux obtenus en graphophonologie. Pour l'orthographe, il s'agissait de mettre en évidence la sensibilité des enfants à la fréquence des bigrammes en fonction de leur position initiale ou finale dans un item donné. Dans un premier temps, l'étude des connaissances orthographiques implicites et explicites et de leur évolution en fonction du niveau de lecture des enfants a été réalisée. Dans un deuxième temps, l'influence de la structure orthographique sur les processus de reconnaissance de mots écrits a été évaluée chez les enfants de niveau de lecture CP à CM2.

2.4.1- Développement des connaissances orthographiques implicites et explicites

Comme pour la graphophonologie, les connaissances orthographiques ont été évaluées par une épreuve de plausibilité lexicale (versant implicite) et un test de jugement de bigrammes isolés (versant explicite).

Afin d'estimer les connaissances orthographiques implicites des enfants de la GSM au niveau CM2, des pseudo-mots contenant un bigramme fréquent (en position initiale ou finale : *calufe* ; *latote*) ou un bigramme rare (*nalufe* ; *latove*) étaient proposés. Les enfants devaient indiquer celui qui leur semblait ressembler le plus à un vrai mot. Il était supposé ici que les enfants devaient manifester une préférence pour les items à bigrammes fréquents. Or, les résultats indiquent que les enfants de GSM ne manifestent pas de choix préférentiel pour les pseudo-mots contenant un bigramme fréquent, quelle que soit sa position tandis que les enfants de niveau CP sélectionnent le plus souvent les items avec un bigramme fréquent, toutes positions confondues (initiale ou finale), le choix étant cependant au niveau du hasard pour chacune des positions prise individuellement. Il faut donc attendre que les enfants possèdent un certain niveau de lecture pour qu'une forme de sensibilité à la fréquence des patrons de lettres apparaisse. Cette sensibilité est confirmée pour les niveaux de lecture supérieurs. Par exemple, au niveau CE1, les enfants préfèrent les pseudo-mots avec un bigramme fréquent en position initiale et toutes positions confondues, le choix se situant au niveau du hasard pour la position finale. Toutefois, au niveau CM1, la sensibilité à la fréquence bigrammique n'apparaît que pour les bigrammes en position initiale alors que pour les niveaux CE2 et CM2, les items avec un bigramme fréquent sont préférés quelle que soit la position du bigramme (initiale, finale et confondues). Ainsi, les enfants des niveaux CE2 et CM2 sélectionnent dans 66 % des cas en moyenne *calufe* et non *nalufe* ou encore *latote* et non *latove* comme des mots possibles en français. Ces premiers résultats sur l'émergence d'une sensibilité précoce aux régularités orthographiques de la langue non motivées phonologiquement peuvent être rapprochés en partie de ceux obtenus par Cassar et Treiman (1997) en anglais, Pacton, Fayol et Perruchet (1999) et Pacton, Perruchet, Fayol et Cleeremans (2001) en français. En effet, ces auteurs ont mis en évidence que les enfants possèdent de façon très précoce, dès le CP, des connaissances sur la fréquence de certaines suites de lettres. En français, les données portent essentiellement sur le processus de gémination et indiquent que les items avec une consonne fréquemment doublée sont plus souvent choisis que les autres. Cette forme de sensibilité aux régularités orthographiques s'observe dès le CP. Elle peut être rapprochée de celle obtenue chez les enfants de niveau de lecture CP à CM2 dans la présente étude. Très tôt (mais pas dès la GSM, contrairement aux hypothèses), les enfants sont sensibles à des aspects purement orthographiques tels que la fréquence de certains bigrammes

dans la langue. Cette sensibilité témoigne de l'intériorisation de régularités orthographiques de façon implicite dans la mesure où ces régularités ne renvoient pas à des règles apprises explicitement. En outre, et conformément aux conclusions des études de Pacton, Fayol et Perruchet (1999) ainsi que de Pacton, Perruchet, Fayol et Cleeremans (2001), cette sensibilité s'accroît avec le niveau de lecture, c'est-à-dire avec le temps de fréquentation de la langue écrite. En effet, les résultats indiquent un effet significatif du niveau d'expertise en lecture bien que les enfants de niveau CM1 (55,25 % de choix préférentiels) obtiennent des résultats assez surprenants : ils ne se différencient que des enfants de niveau CE2 (67,77 %) et CM2 (64,06 %). Ce résultat peut sans doute s'expliquer par les différences interindividuelles fortes observées pour le niveau CM1, notamment dans le cas des items contenant un bigramme fréquent ou rare en position initiale. Les disparités entre enfants sont donc importantes pour ce niveau de lecture. Une certaine forme de sensibilité précoce à la fréquence des bigrammes existe bien et se développe avec l'efficacité en lecture. De plus, cette sensibilité tend à être dépendante de la position du bigramme : d'une façon générale, les enfants apparaissent plus sensibles à la fréquence du bigramme lorsque celui-ci constitue l'attaque du pseudo-mot. Ainsi, dans plus de 60 % des cas, les enfants préfèrent les items avec un bigramme fréquent en position initiale contre 56 % pour les items avec un bigramme fréquent en position finale. Comme l'avait déjà suggéré Segui (1984), l'attaque, qui peut être ici comparable au bigramme initial, serait traitée plus facilement que la rime (assimilable au bigramme final).

Les données obtenues sont différentes au test évaluant l'étendue des connaissances explicites en orthographe chez les enfants des niveaux CP à CM2. Cette épreuve exploitait également les bigrammes fréquents et rares selon la position dans l'item mais de manière isolée. Ainsi, les enfants devaient indiquer s'il y avait davantage de mots dans la langue qui commençaient par [ca-] ou [na-] et qui se terminaient par [-te] ou [-ve]. De façon similaire à l'épreuve implicite, il était prédit que les enfants devaient choisir préférentiellement les bigrammes fréquents. Or, les résultats indiquent que les choix pour le test explicite se situent davantage au niveau du hasard que dans le cas de l'épreuve implicite et ce pour l'ensemble des enfants. Ainsi, les enfants des niveaux CP et CE1 considèrent dans plus de 68 % des cas en moyenne qu'il y a davantage de mots commençant par [ca-] que par [na-] mais leurs choix sont au niveau du hasard pour les bigrammes en position finale et toutes positions confondues. Pour les niveaux CE2 et CM1, le profil inverse est obtenu : ces enfants manifestent une préférence pour le bigramme [-te] au détriment de [-ve] en position finale et leurs performances diffèrent du hasard toutes positions confondues (respectivement 67,19 % et 66 %) mais les choix sont au hasard pour les bigrammes en position initiale. Enfin, les enfants de niveau CM2 considèrent que

davantage de mots se terminent par [-te] que par [-ve] mais ne manifestent pas de choix préférentiel dans le cas des bigrammes en position initiale et toutes positions confondues. Ces résultats indiquent que si les enfants à partir du niveau CP jusqu'au niveau CM2 ont une sensibilité à la fréquence des bigrammes et développent précocement des connaissances implicites (voir paragraphe précédent), leurs connaissances explicites semblent, quant à elles, moins étendues. Les données obtenues ne permettent pas de conclure que les connaissances explicites en orthographe s'élaborent sur la base des connaissances implicites, contrairement au schéma développemental de Gombert (2003). A partir de ce même modèle, une des hypothèses de cette expérience était que les connaissances explicites étant fortement dépendantes de l'apprentissage de la lecture, elles devaient se développer plus tardivement que les connaissances implicites. L'épreuve explicite devait alors être moins bien réussie que celle implicite. En réalité, la différence de performances entre les deux épreuves n'est pas significative. De même, les connaissances explicites sur la fréquence des bigrammes ne se développent pas avec le niveau d'expertise en lecture comme le prédisaient les hypothèses s'appuyant sur la modélisation de Gombert (2003). Aucun des différents niveaux de lecture ne se différencie au niveau des performances obtenues et contrairement à ce qui était obtenu pour l'épreuve implicite, la position du bigramme (initiale ou finale) n'a pas d'effet, de façon globale, sur le choix des enfants sauf pour les enfants de niveau CE2 et CM1. Pour ces deux niveaux de lecture, les enfants réussissent plus facilement à fournir des réponses correctes pour les bigrammes en position finale (i.e., davantage de mots se terminent par [-te] que par [-ve]) que pour ceux en position initiale. Ce dernier résultat va à l'encontre de ce qui était obtenu à l'épreuve implicite où les choix préférentiels étaient accentués dans le cas du bigramme en position initiale. D'une façon générale, les résultats obtenus à cette épreuve orthographique ne vont pas dans le sens des hypothèses formulées sur le développement des connaissances explicites sur la fréquence des bigrammes. Ces résultats peuvent trouver une voie d'explication dans la nature même de l'épreuve. En effet, contrairement à l'épreuve explicite utilisée pour évaluer les connaissances graphophonologiques des enfants sur la consistance graphèmes-phonèmes, le test orthographique portait sur des paires de lettres isolées (i.e., les bigrammes) et non sur des mots. Ce type de matériel est plus abstrait que des mots ou même des pseudo-mots. Dès lors, le jugement à effectuer par les enfants est sans doute plus complexe dans ce cas que face à des mots que les enfants ont l'habitude de traiter dans le contexte scolaire. D'ailleurs, les écarts-types très élevés observés pour chaque niveau de lecture et pour chacun des bigrammes en fonction de leur position indiquent une forte disparité dans les performances obtenues. Pour de nombreux enfants, cette épreuve a sans doute été difficile à réaliser : le matériel a pu être jugé comme trop

abstrait et ne permettant pas aux enfants de se servir de leurs connaissances implicites afin de pouvoir répondre aux consignes.

L'étude des traitements orthographiques opérés durant la lecture va permettre d'obtenir des informations complémentaires sur la façon dont les enfants utilisent la structure orthographique, en l'occurrence la fréquence des bigrammes, pour lire des mots connus et nouveaux (pseudo-mots) en fonction du niveau d'expertise en lecture.

2.4.2- Traitements orthographiques et niveau de lecture

Sur le même mode que l'épreuve de lecture à haute voix proposée en graphophonologie, les enfants se voyaient proposer ici quatre types d'items : des mots à bigramme fréquent en position initiale et finale (canard, tomate), des mots à bigramme rare en position initiale et finale (nature, lessive), des pseudo-mots à bigramme fréquent (caloude, bofite) et des pseudo-mots à bigramme rare (naluve, binouve). Les données obtenues ont été analysées en prenant en compte l'exactitude et le temps de lecture correcte des items selon l'expertise en lecture des enfants. Une analyse différentielle des bigrammes en fonction de la position dans l'item (initiale ou finale) a été réalisée.

Pour les items à bigramme en position initiale, l'analyse des scores et des temps de lecture ne permet pas de vérifier de façon générale l'hypothèse selon laquelle la structure orthographique aide la reconnaissance des mots écrits dès le début de l'apprentissage de la lecture et même au-delà. Concernant l'influence de la structure orthographique sur l'identification de mots écrits, les hypothèses s'appuyaient notamment sur le modèle de lecture par analogies de Gombert *et al.* (1997) et celui de Colé *et al.* (1999). Dans de telles conceptions, les représentations orthographiques sont susceptibles de se développer avant même l'apprentissage explicite de la lecture : la procédure de lecture, en particulier de mots nouveaux, pourrait s'effectuer sur la base d'analogies avec des mots dont l'enfant connaît la forme orthographique (Goswami, 1986, 1988b, 1990 ; Gombert *et al.*, 1997). Le voisinage orthographique pourrait alors constituer une aide en lecture facilitant la reconnaissance de mots connus et de mots nouveaux. Dans le cadre de la présente étude, l'idée sous-jacente était que la sensibilité des enfants aux unités orthographiques fréquentes comme les bigrammes pouvait faciliter la reconnaissance de mots et pseudo-mots contenant de tels bigrammes. L'enfant, notamment par le biais du voisinage orthographique, serait en mesure d'extraire, des mots et non mots rencontrés, les séquences de lettres les plus fréquentes (i.e., bigrammes fréquents). Par analogie avec des mots connus contenant ces séquences, les processus de reconnaissance de mots

et non mots écrits pourraient alors être facilités. Or, les résultats obtenus indiquent que les items (mots et pseudo-mots) à bigramme fréquent ne sont pas mieux lus (exactitude et temps de lecture) que ceux à bigramme rare. Ces résultats s'observent quel que soit le niveau en lecture. Cependant, d'une façon générale, les enfants les moins experts en lecture (niveau CP) lisent moins vite et font davantage d'erreurs que ceux plus efficaces (niveau CE1 à CM2). Les enfants de niveau CE1 lisent aussi moins bien que ceux de niveau CM2 tandis que les enfants de niveau CE2 mettent plus de temps à lire les items que les niveaux CM1 et CM2. Si l'indice de fréquence bigrammique (en position initiale dans l'item) n'est pas utilisé pour lire des mots connus et des pseudo-mots, la lexicalité des items est en revanche un facteur qui affecte l'exactitude et le temps de lecture. Ainsi, chez tous les enfants, les mots sont lus avec moins d'erreurs que les pseudo-mots. Les mots sont également lus plus rapidement que les pseudo-mots hormis chez les enfants de CP où la différence entre les deux catégories d'items n'est pas significative. Pour ces enfants dont l'apprentissage de la lecture n'en est qu'à ses prémices, l'expérience 1 sur les effets de consistance graphophonémique avait montré que la structure graphophonologique pouvait constituer une aide dans la reconnaissance de mots : les mots réguliers y étaient lus plus rapidement que les mots irréguliers. Les enfants de niveau CP qui apprennent à cette période les correspondances graphèmes-phonèmes de façon explicite pouvaient ainsi se servir de leurs connaissances sur la consistance graphophonémique pour traiter plus rapidement les mots écrits. En revanche, leurs représentations orthographiques seraient encore insuffisamment développées pour permettre une reconnaissance des mots et une différenciation par rapport aux pseudo-mots. Pour les autres niveaux de lecture, les résultats sont plus surprenants quant à l'absence d'effet de la fréquence des bigrammes. Ils indiquent cependant qu'à partir du niveau CE1, les performances en termes d'exactitude en lecture sont assez élevées : les enfants pourraient se servir non pas de la structure orthographique mais de leurs connaissances graphophonologiques pour lire les mots connus et nouveaux. En effet, les bigrammes utilisés en position initiale [ca-] et [na-] diffèrent en termes de fréquence bigrammique mais pas en termes de consistance graphophonémique : [ca-] se prononce toujours /ka/ en position initiale et [na-] se prononce toujours /na/. Ces deux couples de lettres ne présentent pas de difficulté particulière du point de vue de la prononciation en position initiale et la récupération des items contenant de tels bigrammes pourrait donc se faire *via* les représentations graphophonologiques et non simplement orthographiques.

Les résultats obtenus pour les items à bigrammes en position finale sont globalement similaires et semblent conforter cette hypothèse. En effet, ici encore, la structure orthographique n'influence pas l'exactitude et la rapidité en lecture : les items à bigramme fréquent ne sont pas mieux lus (plus rapidement et avec moins d'erreurs) que ceux à bigramme rare. Comme

précédemment, les mots restent lus plus vite et avec moins d'erreurs que les pseudo-mots sauf chez les enfants de niveau CP où la différence entre les deux catégories d'items n'est pas significative. Logiquement, les enfants les moins efficaces en lecture (niveau CP) lisent moins bien que les autres enfants. Les enfants de niveau CE1 sont moins rapides que les autres niveaux de lecture et font davantage d'erreurs que les niveaux CE2 et CM1 pour les mots et pour toutes les catégories d'items par rapport aux enfants de niveau CM2. Au niveau CE1, les procédures de lecture ne sont pas encore bien assurées pour un grand nombre d'enfants : face à des items dont la structure peut être identifiée comme ressemblant à un mot existant réellement (i.e., les mots), les enfants de ce niveau sont susceptibles d'être encore hésitants et de produire des erreurs en lecture. Les enfants d'un niveau supérieur, niveau CE2, mais néanmoins proche font d'ailleurs également davantage d'erreurs sur les mots que ceux de niveau CM2.

3- Expérience 3 : morphologie

L'expérience 3 porte sur les aspects morphologiques dérivationnels, et plus particulièrement sur l'effet d'affixation. Comme pour les deux expériences précédentes, les connaissances implicites et explicites des procédures d'affixation possédées par les enfants de la GSM au niveau CM2 vont être évaluées grâce à une épreuve de plausibilité lexicale portant sur des pseudo-mots (versant implicite) et *via* un test d'analyse explicite de détection d'intrus morphologiques constitué de mots. Une épreuve de lecture à haute voix de mots et pseudo-mots permettra l'analyse des traitements morphologiques en lecture. A nouveau, les enfants de GSM ne sont soumis ni à l'épreuve explicite ni à celle de lecture à haute voix.

3.1- Méthode

3.1.1- Population

Les enfants ayant participé à cette troisième expérience sont les mêmes que ceux des deux expériences précédentes.

3.1.2- Matériel

Epreuves morphologiques implicite, explicite et lecture à haute voix

L'objectif de ces trois épreuves morphologiques est d'évaluer les connaissances morphologiques implicites et explicites dont disposent l'apprenti lecteur, le lecteur débutant et le lecteur expert ainsi que d'analyser les liens entre ces connaissances et la lecture. Pour les trois types de tâches (jugement implicite, jugement explicite, lecture), les morphèmes utilisés sont des affixes dérivationnels : préfixes et suffixes. Les préfixes retenus sont [re-] et [dé-]. Les suffixes sont [-age] et [-eur]. Ces affixes ont été retenus car ils font partie des affixes les plus employés de la langue française. Sur 35740 mots de la base de données lexicales *Brulex* (Content *et al.*, 1990) :

- 695 mots commencent par [re-] (dont 535 préfixés),
- 833 mots commencent par [dé-] (dont 410 préfixés),
- 528 mots finissent par le trigramme [-age] (dont 498 suffixés),
- et 823 mots finissent par [-eur] (dont 583 suffixés).

Ces affixes sont utilisés dans des pseudo-mots pour l'épreuve implicite, des mots pour l'épreuve explicite, des mots et pseudo-mots en lecture.

Tâche morphologique implicite de plausibilité lexicale écrite

Cette épreuve est constituée de 12 paires d'items expérimentaux et de 4 paires d'items distracteurs en GSM, de 16 paires d'items expérimentaux et de 8 paires d'items distracteurs du niveau CP au niveau CM2 (*cf.* Annexe 4, pp. 254-255).

Chaque paire d'items expérimentaux contient un pseudo-mot affixé et un pseudo-mot non affixé, soit :

- 3 paires en GSM et 4 paires du niveau CP au niveau CM2 exploitent le préfixe [re-] (*resibrer-risebrer*),
- 3 paires (GSM) et 4 paires (niveaux CP-CM2) concernent le préfixe [dé-] (*détiner-diténer*),
- 3 paires (GSM) et 4 paires (niveaux CP-CM2) pour le suffixe [-age] (*vaminage-nagivame*),
- 3 paires (GSM) et 4 paires (niveaux CP-CM2) pour le suffixe [-eur] (*donéveur-veurédon*).

Les items distracteurs sont constitués de pseudo-mots construits selon le même principe que les items expérimentaux. Les préfixes utilisés sont [en-] (*enroufir-ourenfir*) et [pré-] (*prédafer-pradéfer*). Les suffixes sont [-oir] (*talomoir-moirotal*) et [-esse] (*argesse-gessare*). Ces items distracteurs ne sont pas utilisés lors de l'analyse des réponses fournies.

Pour les items expérimentaux, le nombre de lettres est de 7 pour 6 paires et de 8 pour 6 paires (GSM), de 7 pour 7 paires et de 8 pour 9 paires (niveaux CP-CM2). Le nombre de phonèmes est de 6 pour 5 paires, de 7 pour 4 paires, de 6 à 7 pour 2 paires utilisant le suffixe [-eur] (exemple : dans la paire *donéveur-veurédon*, *donéveur* est composé de 7 phonèmes, *veurédon* de 6 phonèmes) pour les GSM, de 6 pour 6 paires, de 7 pour 7 paires et de 6 à 7 pour 3 paires utilisant le suffixe [-eur] (exemple : dans la paire *biniteur-teuribin*, *biniteur* est composé de 7 phonèmes, *teuribin* de 6 phonèmes) du niveau CP au niveau CM2. Pour chaque paire, le nombre de syllabes est identique (3 pour l'ensemble des paires). Deux paires de pseudo-mots en GSM (*rfhg-vosa*, *sjbt-moti*) et 4 paires pour les enfants des niveaux CP-CM2 (*ibidem*, *tipa-yyfc*, *arlu-txnc*) constituent l'entraînement. Un seul pseudo-mot de chacune des paires peut constituer un mot français. La présentation du matériel est la même que celle utilisée pour l'épreuve implicite graphophonologique.

Tâche morphologique explicite de jugement d'intrus (uniquement à partir du niveau CP)

Seize triplets de mots constituent le matériel expérimental (*cf.* Annexe 4, pp. 255-256). Chaque triplet contient deux mots affixés et un mot pseudo-affixé jouant le rôle d'intrus. Les mots affixés sont formés d'un affixe (préfixe et suffixe) et d'une base autonome, c'est-à-dire possédant une signification : par exemple *démêler* est formé du préfixe [dé-] et de la base [mêler] qui constitue un verbe signifiant. Les mots pseudo-affixés comportent un affixe et une base non

autonome : par exemple dans *repérer*, [pérer] n'a pas d'existence sémantique propre, [re-] ne joue plus alors le rôle de préfixe car il est associé à un élément sans propriété sémantique, d'où le terme de pseudo-affixe :

- 4 triplets sont formés avec le préfixe [re-] (*regagner-redonner-refléter*),
- 4 triplets sont formés avec le préfixe [dé-] (*dévisser-décoller-déchirer*),
- 4 triplets sont formés avec le suffixe [-age] (*pliage-rasage-ménage*),
- 4 triplets sont formés avec le suffixe [-eur] (*parleur-dormeur-pasteur*).

Pour les triplets préfixés, le nombre de lettres est identique (7 pour 4 triplets, 8 pour 4 triplets) ainsi que le nombre de syllabes (3 pour les 8 triplets). Pour les triplets suffixés, le nombre de lettres est le même dans chaque triplet (6 pour 4 triplets, 7 pour 4 triplets), et le nombre de syllabes est de 2 pour les 8 triplets. Le nombre de phonèmes est de 4 à 6. L'appartenance à la même catégorie sémantique est contrôlée : tous les préfixés sont des verbes et tous les suffixés sont des noms. La fréquence lexicale des mots a aussi été contrôlée (cf. Annexe 4, p. 256). Deux triplets constituent l'entraînement : un triplet avec le préfixe [in-] (*invisible-incapable-insecte*) et un triplet avec le suffixe [-ette] (*fillette-jupette-recette*). L'ordre de présentation des items dans chaque triplet et celui de l'ensemble des triplets est aléatoire. Le mode de présentation est identique à celui utilisé pour l'épreuve implicite.

Tâche de lecture à haute voix (uniquement à partir du niveau CP)

Le matériel expérimental est constitué de 32 items à lire (cf. Annexe 4, pp. 257-258), soit :

- 8 mots affixés dont 2 préfixés [re-], 2 préfixés [dé-], 2 suffixés [-age], et 2 suffixés [-eur] : ces mots sont constitués d'un affixe et d'une base signifiante (*dépasser, nageur*),
- 8 mots pseudo-affixés : ces mots ne peuvent être décomposés car la base n'a pas d'existence sémantique et l'affixe ne peut alors être considéré comme un morphème (*renifler, bagage*),
- 8 pseudo-mots affixés : les items sont construits à partir d'une base existante et d'un affixe mais l'item élaboré n'existe pas (*décasser* : construit à partir de [dé-] et [-casser]),
- 8 pseudo-mots non affixés élaborés avec une fausse base et un affixe. Les items ainsi formés n'existent pas et n'ont aucune signification (*rebaner, bouveur*).

Les items suffixés sont composés de 6 à 8 lettres, 5 à 6 phonèmes et de 2 à 3 syllabes. Les items préfixés sont composés de 7 à 9 lettres, 6 à 7 phonèmes et de 3 syllabes. La fréquence des bigrammes et des trigrammes composant ces items a également été prise en compte. La différence entre l'ensemble des items expérimentaux en terme de fréquence des bigrammes, de fréquence des trigrammes, de nombre de lettres, de nombre de phonèmes, de nombre de syllabes a été analysée du point de vue statistique. La fréquence lexicale des mots a également été

contrôlée. Ces analyses figurent en annexe (Annexe 4, pp. 257-258). Les items proposés pour l'entraînement sont : *valise, pamoti, fleur, lipaber*. Les items sont présentés de la même façon que dans l'épreuve graphophonologique de lecture.

3.1.3- Procédure

Epreuves morphologiques implicite, explicite et lecture à haute voix

Tâche morphologique implicite de plausibilité lexicale écrite

La procédure et les consignes sont les mêmes que celles adoptées pour l'épreuve orthographique implicite. Cette épreuve est notée sur 12 points pour les GSM (deux sous notes : préfixés sur 6, suffixés sur 6). Des niveaux CP à CM2, la note globale est de 16 (préfixés sur 8 et suffixés sur 8).

Tâche morphologique explicite de jugement d'intrus

La passation est semi-collective du niveau CP au niveau CM2. Avant de commencer, il est expliqué que dans la langue, il existe des mots construits grâce à d'autres mots de la même famille et des mots qui ne sont pas construits. L'exemple d'un mot construit morphologiquement est alors proposé au participant : *sonnette* se construit grâce à *sonne*. Cet exemple est fourni oralement aux enfants et par écrit pour les adultes. Pour les enfants, il leur est ensuite précisé de regarder les trois mots figurant sur chaque page et d'entourer l'intrus, c'est-à-dire le mot qui ne va pas avec les deux autres, qui n'est pas construit. La phase expérimentale commence après avoir fourni les deux triplets d'entraînement. Le déroulement de la passation ainsi que les critères de cotation sont similaires à ceux de l'épreuve implicite (note globale sur 16, préfixés sur 8 et suffixés sur 8 également).

La consigne donnée aux enfants est la suivante : « *Dans tous les mots qui existent, il y a des mots qui sont construits grâce à d'autres mots de la même famille et des mots qui ne sont pas construits comme ça. Par exemple, sonnette se construit grâce à sonne. Je vais vous distribuer un petit carnet. Sur chaque page du carnet, il y a trois mots qui sont écrits. Parmi ces trois mots, il y en a un qui n'est pas construit avec un mot de la même famille. Il faut trouver ce mot qui ne va pas avec les autres et l'entourer. Quand vous avez fini d'entourer pour la première page, vous attendez que je vous dise de tourner la page pour continuer. Attention, vous n'avez pas le droit de revenir en arrière. Vous avez bien compris ? Vous n'avez pas de questions ?* ».

Tâche de lecture à haute voix

La procédure et la consigne fournie aux participants sont identiques à celles de l'épreuve graphophonologique. Pour chaque catégorie, (mots affixés, pseudo-mots affixés, mots pseudo-affixés et pseudo-mots pseudo-affixés), un point est attribué par mot lu correctement.

Pour l'ensemble des épreuves, l'ordre de passation a été aléatorisé. Le nombre important d'épreuves a nécessité deux sessions qui se sont déroulées en janvier et début février 2004.

3.2- Hypothèses opérationnelles

Epreuve implicite : les pseudo-mots construits morphologiquement (c'est-à-dire possédant une structure morphologique du type affixe + pseudo-racine comme *resibrer*) devraient être plus souvent sélectionnés comme ressemblant à de vrais mots que les pseudo-mots ne possédant pas de structure morphologique (comme *risebrer*). Les préfixes étant plus faciles à isoler que les suffixes, les enfants devraient être plus performants dans le cas des paires à préfixe (*resibrer-risebrer*) par rapport aux paires à suffixe (*vaminage-nagivame*). Les enfants les plus efficaces en lecture devraient choisir majoritairement les items construits morphologiquement.

Epreuve explicite : dans chaque triplet d'items contenant deux mots affixés (*dévisser-décoller*) et un mot non affixé (*déchirer*), les enfants devraient être capables de détecter l'item non affixé comme intrus morphologique. Ce choix devrait être facilité dans le cas des triplets préfixés par rapport à ceux suffixés (*parleur-dormeur-pasteur*). La réussite à cette épreuve devrait s'accroître avec le niveau de lecture.

Lecture à haute voix : les mots (*déplier*) devraient être mieux lus (plus rapidement et avec moins d'erreurs) que les pseudo-mots (*décasser*). Les items construits morphologiquement ou affixés (comme *dépasser* ou *chasseur*) seront mieux lus que les pseudo-affixés (tels que *démouder* ou *bouveur*). Ces effets de lexicalité et d'affixation devraient être d'autant plus marqués que le niveau de lecture des enfants augmente.

3.3- Résultats

3.3.1- Tâches morphologiques implicite et explicite

Le Tableau 12 ci-après regroupe les performances obtenues par les enfants de GSM et par les cinq groupes de niveau de lecture aux deux tâches d'évaluation des connaissances morphologiques.

Tableau 12. Performances exprimées en pourcentages de choix de l’item affixé (et *écart-types*) obtenues aux épreuves morphologiques selon le type d’affixe (préfixe et suffixe)

	Morphologie implicite			Morphologie explicite		
	Préfixés	Suffixés	Moyenne	Préfixés	Suffixés	Moyenne
GSM	45,55 (16,17)	47,22 (18,08)	46,39			
Niveau CP	51,21 (17,04)	52,02 (15,53)	51,61	32,66 (27,99)	32,66 (25,49)	32,66
Niveau CE1	59,05 (18,23)	63,36 (21,06)	61,21	39,22 (26,92)	44,83 (29,82)	42,03
Niveau CE2	69,14 (19,74)	59,37 (21,11)	64,26	49,02 (36,28)	75 (21,77)	62,01
Niveau CM1	73,50 (20,76)	66 (22,15)	69,75	52 (25,16)	53 (22,78)	52,50
Niveau CM2	74,22 (16,33)	66,41 (19,12)	70,31	67,19 (20,07)	52,34 (15,70)	59,77
Moyenne	62,11	59,06	60,59	48,02	51,57	49,79

Tâche morphologique implicite

A cette tâche d’évaluation des connaissances morphologiques implicites (plausibilité lexicale), les enfants doivent choisir entre deux pseudo-mots celui qui ressemble le plus à un vrai mot. Les items sont de type préfixé ou de type suffixé. Les écart-types à cette épreuve sont relativement élevés et augmentent avec le niveau de lecture (jusqu’au niveau CM1 notamment). Plus les lecteurs sont efficaces, plus la dispersion est importante, notamment dans le cas des items suffixés où les écarts-types sont plus importants que pour les préfixés (sauf pour les enfants de niveau CP ; voir Tableau 12 pour les différentes valeurs).

Les analyses qui suivent portent sur deux facteurs principaux : le niveau de lecture (six modalités : GSM, niveaux CP, CE1, CE2, CM1, CM2) et le type d’affixe (préfixé vs. suffixé). Le plan d’analyse est le suivant : $S_n < N_6 > * A_2$. Le facteur S correspond aux sujets (participants), le facteur N au niveau de lecture des enfants et le facteur A au type d’affixation des items.

Les réponses des enfants de GSM et de niveau de lecture CP ne se distinguent pas du hasard (hasard = 50 % de réponses correctes). Ces enfants ne choisissent donc pas préférentiellement les items affixés, que ce soit pour les préfixés ou les suffixés. En revanche, les enfants des niveaux CE1 à CM2 préfèrent l’item construit morphologiquement à l’autre item⁵ (CE1 : $p < .001$; CE2 : $p < .0001$; CM1 : $p < .0001$; CM2 : $p < .0001$). Ainsi, les enfants de ces niveaux considèrent dans plus de 66 % des cas (*cf.* Tableau 12) que des items morphologiquement construits tels que *remobir* ou *vaminage* ressemblent davantage à des mots que ceux qui ne contiennent pas un affixe identifiable comme *romebir* ou *nagivame*.

⁵ Les réponses des enfants sont différentes du hasard pour les items préfixés comme pour les suffixés.

En outre, l'effet du niveau de lecture est significatif ($F(5, 173) = 14,93 ; p < .0001$) (cf. Figure 16). Globalement, les enfants les plus efficaces en lecture sont ceux qui manifestent le plus de choix préférentiels à l'égard des items affixés ou morphologiquement construits.

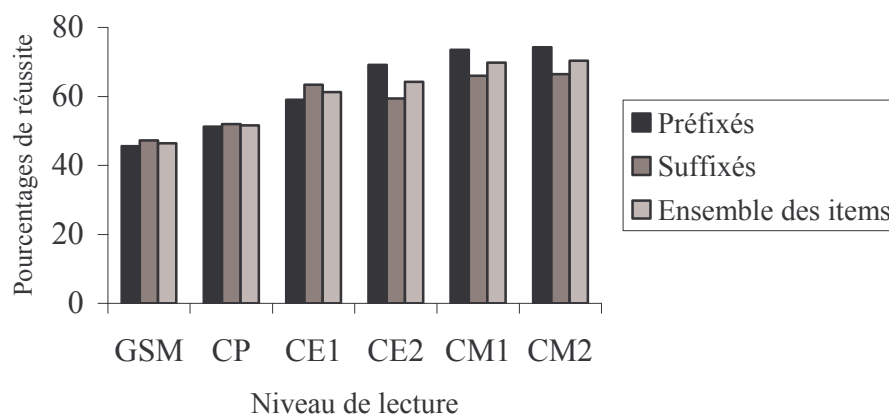


Figure 16. Pourcentages de choix de l'item affixé (préfixé et suffixé) chez les enfants de GSM au niveau de lecture CM2

Cependant, la préférence pour l'item affixé varie en fonction du niveau de lecture. En effet, les analyses ne montrent pas de différence significative entre les scores des enfants de GSM (46,39 %) et ceux des enfants de niveau CP (51,61 % ; $F(1, 59) = 2,36 ; ns$), ni entre ceux des enfants de niveaux CE1 (61,21 %) et CE2 (64,26 % ; $F(1, 59) < 1$), ni entre ceux des enfants de niveau CE2 et ceux des enfants des niveaux CM1 (69,75 % ; $F(1, 55) = 2,28 ; ns$) et CM2 (70,31 % ; $F(1, 62) = 3,37 ; ns$), ni encore entre les enfants des niveaux CM1 et CM2 ($F(1, 55) = < 1$). En revanche, les différences sont significatives entre les enfants de GSM et les enfants des niveaux CE1 ($F(1, 57) = 18,64 ; p < .0001$), CE2 ($F(1, 60) = 34,34 ; p < .0001$), CM1 ($F(1, 53) = 49,98 ; p < .0001$) et CM2 ($F(1, 60) = 62,83 ; p < .0001$). De plus, le niveau CP obtient des scores significativement inférieurs (51,61 %) aux niveaux CE1 ($F(1, 58) = 5,71 ; p < .05$), CE2 ($F(1, 61) = 12,05 ; p < .001$), CM1 ($F(1, 54) = 20,48 ; p < .0001$) et CM2 ($F(1, 61) = 26,71 ; p < .0001$). Les enfants de niveau CE1 obtiennent également des scores inférieurs aux enfants des niveaux CM1 ($F(1, 52) = 4,44 ; p < .05$) et CM2 ($F(1, 59) = 6,17 ; p < .05$). Enfin, les scores aux items préfixés (62,11 %) ne se distinguent pas de ceux obtenus aux items suffixés (59,06 % ; $F(1, 173) = 2,42 ; ns$) et l'interaction entre les deux facteurs (niveau de lecture et affixation) n'est pas significative ($F(5, 173) = 1,57 ; ns$).

Tâche morphologique explicite

Dans la tâche morphologique explicite, les enfants des niveaux CP, CE1, CE2, CM1 et CM2 doivent détecter les mots non construits morphologiquement (préfixés et suffixés). Comme pour l'épreuve de morphologie implicite, les différences interindividuelles sont importantes à l'épreuve explicite. En effet, les écart-types sont importants et plus élevés pour les items préfixés que pour les suffixés (CP : 27,99 > 25,49 ; CE2 : 36,28 > 21,77 ; CM1 : 25,16 > 22,78 ; CM2 : 20,07 > 15,70) sauf pour les enfants de niveau CE1 (26,92 < 29,82).

Dans les analyses suivantes, deux facteurs sont mis en jeu : le niveau de lecture (cinq modalités : CP, CE1, CE2, CM1, CM2) et le type d'affixe (préfixé vs. suffixé). Ainsi, le plan d'analyse est : $S_n < N_5 > * A_2$. Le facteur S correspond aux sujets (participants), le facteur N représente le niveau de lecture des enfants et enfin le facteur A désigne le type d'affixation des items.

Comme le montre le Tableau 12, le pourcentage d'intrus correctement détectés est de 49,79 % pour l'ensemble des enfants. Cette tâche est donc très moyennement réussie. Comme pour l'épreuve implicite, les réponses des enfants de niveau de lecture CP ne se distinguent pas du hasard (33,3 %) à la fois pour les préfixés et les suffixés. Les enfants de niveau CE1 détectent correctement les intrus de type suffixé ($p < .01$) mais pas les préfixés. Enfin, pour les enfants des niveaux CE2, CM1 et CM2, les réponses fournies diffèrent du hasard pour les deux catégories d'items (préfixés et suffixés ; CE2 : $p < .0001$; CM1 : $p < .0001$; CM2 : $p < .0001$).

Par ailleurs, le niveau de lecture a un effet significatif sur les performances ($F(4, 144) = 26,71$; $p < .0001$). Néanmoins, comme à l'épreuve implicite, le choix de l'intrus diffère selon le niveau de lecture. En effet, les performances des enfants de niveau CE2 (62,01 %) ne se différencient pas de celles obtenues par le niveau CM2 (59,77 % ; $F(1, 62) < 1$). De la même façon, les scores obtenus par les enfants de niveau CM1 (52,50 %) ne se distinguent pas de ceux des enfants de niveau CM2 ($F(1, 55) = 1,80$; *ns*). En revanche, la différence entre les niveaux CE2 et CM1 est significative ($F(1, 55) = 2,95$; $p < .05$) : le niveau CE2 réussit mieux l'épreuve explicite que le niveau CM1. Pour les autres niveaux de lecture, toutes les différences sont significatives (CP : 32,66 % vs. CE1 : 42,03 % ; $F(1, 58) = 7,83$; $p < .01$; CP-CE2 : $F(1, 61) = 50,35$; $p < .0001$; CP-CM1 : $F(1, 54) = 20,06$; $p < .0001$; CP-CM2 : $F(1, 58) = 7,83$; $p < .0001$; CE1-CE2 : $F(1, 59) = 19,15$; $p < .0001$; CE1-CM1 : $F(1, 52) = 4,52$; $p < .05$; CE1-CM2 : $F(1, 59) = 16,01$; $p < .0001$). En outre, les intrus de type préfixé (48,02 %) ne sont pas mieux détectés que ceux de type suffixé (51,57 % ; $F(1, 144) = 1,33$; *ns*). Enfin, comme le montre la Figure 17, l'interaction entre le niveau de lecture et le type d'affixe est significative ($F(4, 144) = 4,94$; $p < .001$). Au niveau CE2, les intrus suffixés (75 %) sont bien mieux détectés

que les préfixés (49,02 % ; $p < .001$). Le profil inverse est obtenu pour les enfants de niveau CM2 : ils réussissent mieux à trouver les intrus de type préfixé que ceux de type suffixé (67,19 % > 52,34 %, $p = .0286$). Pour les autres niveaux de lecture, les différences entre les deux types d'affixe restent non significatives.

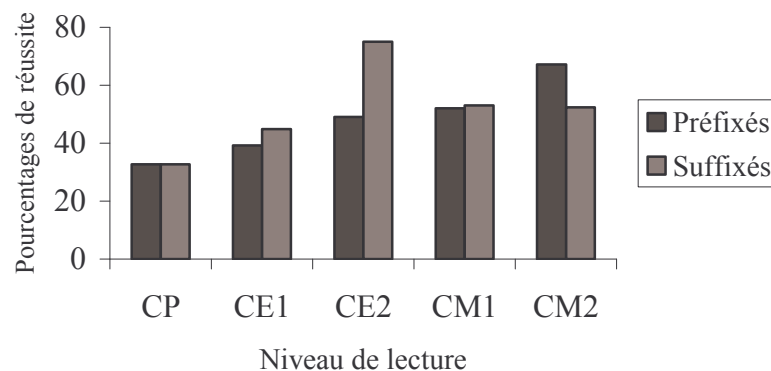


Figure 17. Pourcentages d'intrus morphologiques correctement détectés en fonction du niveau de lecture

3.3.2- Tâche morphologique de lecture à haute voix

Les items de type préfixé et suffixé ont été distingués au niveau de la présentation des résultats. Les analyses ont été réalisées sur les scores (pourcentages de réussite) de lecture (exactitude) ainsi que sur les temps de lecture correcte en prenant en compte trois facteurs principaux : le niveau de lecture (cinq modalités : CP, CE1, CE2, CM1, CM2), la lexicalité des items (mots *vs.* pseudo-mots) et le caractère plurimorphémique (affixés *vs.* pseudo-affixés). Pour les analyses sur les scores, le plan est le suivant : $S_n < N_5 > * L_2 * M_2$, avec S le facteur sujets, N le facteur niveau de lecture des enfants, L désigne la lexicalité des items et enfin M correspond au facteur caractère plurimorphémique des items.

3.3.2.1- Préfixés

Exactitude

Le Tableau 13 rend compte des scores obtenus par les cinq groupes d'enfants aux items (mots et pseudo-mots) affixés et pseudo-affixés.

Tableau 13. Performances moyennes (et *écart-types*) obtenues aux items préfixés en fonction du niveau de lecture

Niveau de lecture	Mots affixés	Mots pseudo-affixés	Pseudo-mots affixés	Pseudo-mots pseudo-affixés
CP	20,97 (30,34)	16,94 (23,62)	21,77 (30,10)	20,97 (31,02)
CE1	85,34 (19,50)	74,14 (23,61)	73,28 (19,97)	71,55 (19,74)
CE2	92,97 (15,86)	87,50 (16,80)	89,06 (16,73)	72,66 (24,06)
CM1	95 (10,21)	87 (20,56)	89 (14,58)	81 (14,93)
CM2	97,66 (7,40)	96,09 (11,20)	92,97 (11,42)	85,16 (17,80)
Moyenne	78,39	72,33	73,22	66,27

Les écart-types sont très élevés chez les enfants de niveau CP : pour tous les items, ils sont supérieurs aux moyennes (voir Tableau 13 ci-dessus). Pour le niveau CE1, les écart-types restent également importants, notamment dans le cas des mots pseudo-affixés (23,61). Les écart-types tendent ensuite à diminuer avec le niveau de lecture, sauf pour les pseudo-mots pseudo-affixés chez les enfants de niveau CE2 (24,06) et les mots pseudo-affixés chez ceux de niveau CM1 (20,56) pour lesquels les différences interindividuelles sont fortes.

L'effet du facteur niveau de lecture est significatif ($F(4, 144) = 120,57 ; p < .0001$). Les enfants de niveau CP font plus d'erreurs de lecture que les enfants des autres groupes de niveau (CP : 20,16 % de réussite, CE1 : 76,08 % ; CE2 : 85,54 % ; CM1 : 88 % ; CM2 : 92,97 % ; toutes les différences sont significatives à .0001). Les enfants de niveau CE1 réussissent moins bien que ceux de niveau CM1 ($p < .05$) et de niveau CM2 ($p < .0004$). La différence entre les enfants de niveau CE2 et ceux de niveau CE1 est tendanciellement significative ($p = .0690$) et les autres différences sont non significatives. De plus, les cinq groupes lisent globalement plus de mots avec exactitude que de pseudo-mots (75,36 % vs. 69,74 %, $F(1, 144) = 17,22 ; p < .0001$) et davantage d'affixés que de pseudo-affixés (75,80 % vs. 69,30 %, $F(1, 144) = 28,10 ; p < .0001$).

Enfin, comme le montre la Figure 18, l'interaction entre les facteurs niveau de lecture et lexicalité est significative ($F(4, 144) = 2,49 ; p < .05$).

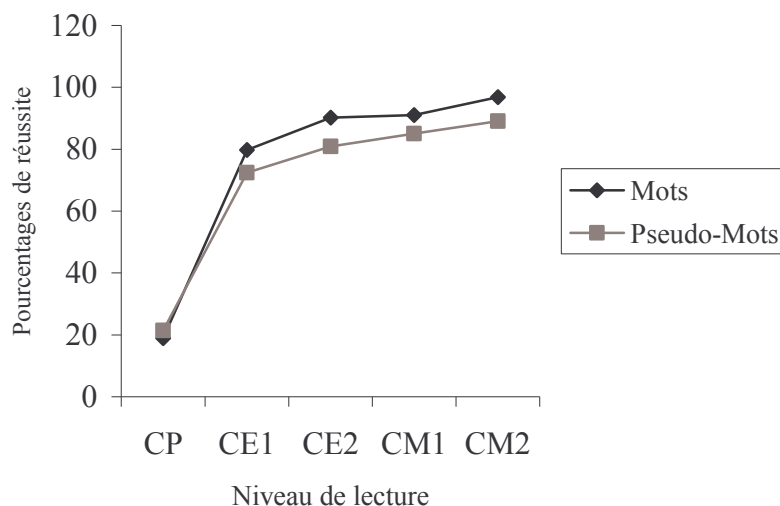


Figure 18. Performances moyennes obtenues pour la lecture de mots et de pseudo-mots par les cinq groupes d'enfants

La différence entre les mots et les pseudo-mots en termes d'exactitude n'est pas significative chez les enfants de niveau CP, CE1 et CM1. Seuls les enfants des niveaux CE2 ($p = .0198$) et CM2 ($p = .0014$) lisent mieux les mots que les pseudo-mots (respectivement : 90,23 % vs. 80,86 % ; 96,87 % vs. 89,06 %). L'interaction révèle aussi que les enfants de niveau CM2 (96,88 %) ont de meilleurs résultats que les enfants des niveaux CM1 (91 % ; $p = .0185$) et CE2 (90,23 % ; $p = .0035$) pour les mots. Cette interaction confirme clairement que les enfants de niveau CE2 lisent mieux que les enfants de niveau CE1, les mots (90,23 % > 79,24 % ; $p = .0130$) comme les pseudo-mots (80,86 % > 72,41 % ; $p = .0265$).

Temps

Les analyses ne portent que sur les temps de lecture correcte. De ce fait, les résultats de 39 enfants du niveau CM2 au niveau CE1 ont été retirés des analyses des temps de lecture correcte (1 pour le groupe de niveau CM2, 1 pour le groupe de niveau CM1, 3 pour le niveau CE2 et 4 pour le niveau CE1). Les enfants de niveau CP ont été retirés des analyses. En effet, l'effectif restant (4) était trop faible pour faire l'objet d'analyses statistiques. Les analyses suivantes concernent donc 25 enfants de niveau CE1, 29 enfants de niveau CE2, 24 enfants de niveau CM1 et 31 enfants de niveau CM2. Ainsi, le plan d'analyse est ici le suivant : $S_n < N_4 > * L_2 * M_2$ (S : facteur sujets ; N : facteur niveau de lecture ; L : facteur lexicalité des items ; M : facteur caractère plurimorphémique des items).

Les temps moyens de lecture des items de type préfixé (exprimés en millisecondes, ms) figurent dans le Tableau ci-dessous.

Tableau 14. Temps moyens de lecture correcte (et *écart-types*) des items préfixés selon du niveau de lecture

Niveau de lecture	Mots affixés	Mots pseudo-affixés	Pseudo-mots affixés	Pseudo-mots pseudo-affixés
CE1	2764,23 (547,63)	3162,46 (859,47)	3057,43 (686,06)	3233,54 (991,74)
CE2	2171,22 (393,99)	2355,86 (502,50)	2610,05 (644,90)	2628,50 (591,74)
CM1	1859,09 (371,88)	1922,23 (352,68)	2126,12 (307,12)	2256,05 (323,13)
CM2	1957,47 (412,02)	1990,36 (399,58)	2088,59 (390,93)	2350,42 (407,43)
Moyenne	2188	2357,73	2470,65	2617,13

La lecture des mots pseudo-affixés et des pseudo-mots pseudo-affixés fait apparaître des différences interindividuelles plus importantes au niveau des temps de lecture pour les enfants de niveau CE1 (respectivement 859,47 ms et 991,74 ms) par rapport aux autres catégories d'items.

Les effets des facteurs niveau de lecture et caractère plurimorphémique sont significatifs (respectivement $F(3, 105) = 23,16$; $p < .0001$ et $F(1, 105) = 18,11$; $p < .0001$). Les enfants de niveau CM2 (2096,71 ms) lisent plus rapidement que les enfants des niveaux CE2 (2441,41 ms ; $p = .0153$) et CE1 (3054,42 ms ; $p < .0001$). Les enfants de niveau CM1 (2040,87 ms) lisent plus vite que ceux de niveau CE2 ($p = .0025$) et de niveau CE1 ($p < .0001$). Les enfants de niveau CE2 sont plus rapides que ceux de niveau CE1 ($p = .0012$). En revanche, les autres différences ne sont pas significatives. Les affixés sont lus plus vite que les pseudo-affixés (2329,27 ms < 2487,43 ms). La différence entre les temps de lecture correcte des mots et des pseudo-mots est significative ($F(1, 105) = 66,54$; $p < .0001$) : les mots (2272,86 ms) sont lus plus rapidement que les pseudo-mots (2543,84 ms). Aucune interaction n'est significative.

3.3.2.2- Suffixés

Exactitude

Les performances en termes d'exactitude obtenues par les enfants aux items suffixés figurent dans le Tableau 15 ci-après.

Tableau 15. Pourcentages de réussite (et *écart-types*) obtenus aux items suffixés selon le niveau de lecture

Niveau de lecture	Mots affixés	Mots pseudo-affixés	Pseudo-mots affixés	Pseudo-mots pseudo-affixés
CP	26,61 (32,87)	42,74 (34,88)	25,81 (28,49)	28,23 (31,45)
CE1	83,62 (16,74)	88,79 (17,15)	89,66 (14,20)	92,24 (11,77)
CE2	91,41 (13,63)	95,31 (9,91)	92,97 (14,53)	87,50 (22)
CM1	94 (13,07)	94 (10,90)	96 (9,35)	91 (17,50)
CM2	96,09 (9,22)	99,22 (4,42)	96,09 (9,22)	93,75 (16,80)
Moyenne	78,35	84,01	80,10	78,54

Le profil de dispersion des performances chez les enfants de niveau CP est similaire à celui observé avec les items préfixés. En effet, les écart-types sont supérieurs aux moyennes pour tous les items (mots affixés : 32,87 ; pseudo-mots affixés : 28,49 ; pseudo-mots pseudo-affixés : 31,45) sauf pour les mots pseudo-affixés (34,88). La dispersion des résultats est ensuite moins importante pour les autres niveaux de lecture bien qu'elle reste forte dans le cas des pseudo-mots pseudo-affixés pour les enfants des niveaux CE2 à CM2 (respectivement 22 ; 17,50 et 16,80).

Il existe un effet du niveau de lecture ($F(4, 144) = 111,80 ; p < .0001$). Les enfants de niveau CP (30,85 %) lisent correctement moins d'items que les enfants des autres niveaux de lecture (CE1 : 88,58 % ; CE2 : 91,80 % ; CM1 : 93,75 % et CM2 : 96,29 % ; ces différences sont toutes significatives à .0001). Les performances des enfants de niveau CE1 diffèrent également de celles du niveau CM2 ($p = .0240$). Les autres différences sont non significatives. L'effet du facteur lexicalité n'est pas significatif ($F(1, 144) = 2,48 ; ns$). Les mots (81,18 %) ne sont pas mieux lus que les pseudo-mots (79,32 %). Les pseudo-affixés sont tendanciellement mieux lus (81,28 %) que les affixés (79,23 % ; $F(1, 144) = 3,28 ; p < .07$).

Bien que l'interaction entre les facteurs niveau de lecture et lexicalité soit significative ($F(4, 144) = 2,97 ; p < .05$), les différences en termes d'exactitude entre les mots et les pseudo-mots pour chaque groupe sont très faibles (*cf.* Figure 19) et aucune n'est significative. En revanche, les items sont effectivement mieux lus par les enfants de niveau CM2 que par ceux de niveau CE1 mais uniquement pour les mots (97,66 % > 86,21 % ; $p < .0001$). Les enfants de niveau CM2 lisent également mieux les mots que ceux de niveau CE2 (93,36 % ; $p = .0423$). Les résultats des enfants de niveau CE1 sont moins bons que ceux des niveaux CM1 (94 % ; $p = .0278$) et CE2 (93,36 %, $p = .0094$) en lecture de mots.

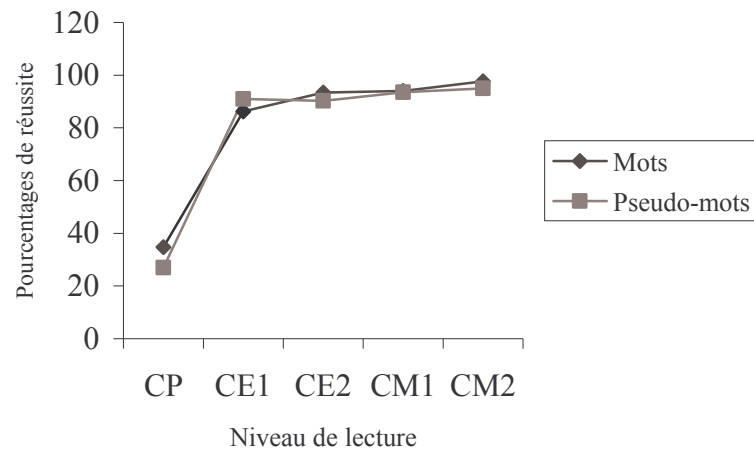


Figure 19. Performances moyennes en lecture de mots et pseudo-mots selon le niveau de lecture

Les facteurs niveau de lecture et caractère plurimorphémique interagissent significativement ($F(4, 144) = 3,39$; $p < .05$; voir Figure 20). Pourtant, comme pour l'interaction précédente, les performances aux items suffixés ne se distinguent significativement pas de celles obtenues aux items pseudo-suffixés. La Figure 20 montre bien que ce résultat s'observe pour chaque niveau de lecture. Cependant, alors que les différences entre les niveaux CM2 et CE2 étaient non significatives en termes d'exactitude de la lecture, les affixés sont mieux lus par les enfants de niveau CM2 (96,09 % > 92,19 % ; $p = .0694$). Cette interaction indique également que les enfants de niveau CE1 (86,64 %) lisent moins bien les affixés que les niveaux CE2 ($p = .0301$) et CM1 (95 % ; $p = .0041$).

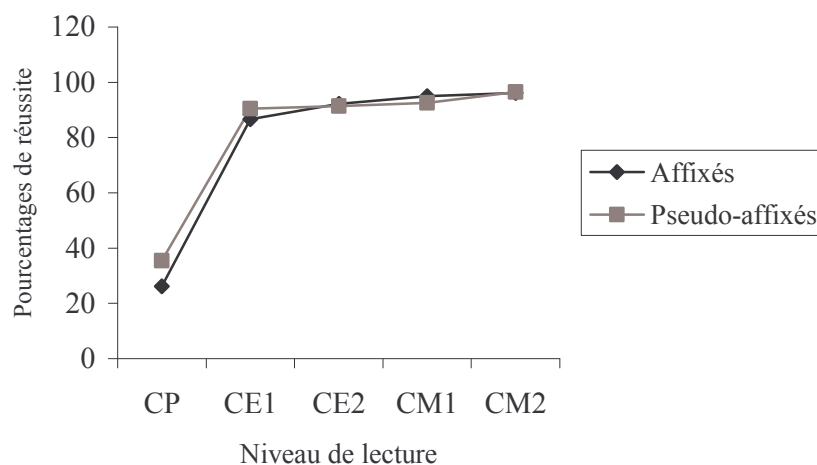


Figure 20. Réussite en lecture aux items affixés et pseudo-affixés pour les enfants des niveaux CP à CM2

Enfin, la lexicalité et le caractère plurimorphémique interagissent significativement ($F(1, 144) = 8,92$; $p < .01$; cf. Figure 21).

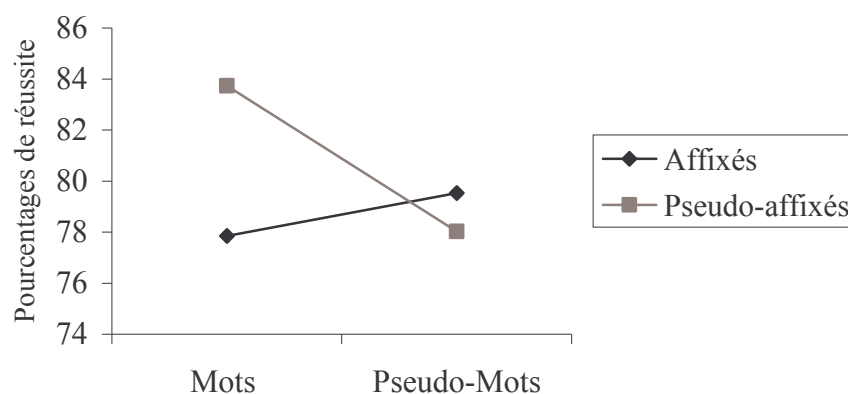


Figure 21. Performances moyennes aux items suffixés selon la lexicalité et le caractère plurimorphémique

Cette interaction confirme l'absence d'effet de lexicalité dans le cas des affixés : les mots (78,35 %) ne sont pas mieux lus que les pseudo-mots (80,10 %). En revanche, les mots pseudo-affixés sont lus avec moins d'erreurs que les pseudo-mots pseudo-affixés (84,01 % > 78,54 %, $p = .0041$). Alors que l'effet du caractère plurimorphémique était tendanciellement significatif, les pseudo-mots affixés ne diffèrent pas des pseudo-affixés au niveau de l'exactitude en lecture. Cet effet d'affixation est confirmé avec les mots : les mots suffixés sont moins bien lus que les pseudo-suffixés (78,35 % < 84,01 % ; $p = .0018$). L'interaction entre les facteurs niveau de lecture, lexicalité et caractère plurimorphémique n'est pas significative ($F < 1$).

Temps

L'analyse de variance réalisée ici sur les suffixés ne porte que sur les temps de lecture correcte comme dans le cas des préfixés. Cette fois-ci les résultats de 28 enfants du niveau CE1 au niveau CM2 ont été retirés des analyses. Comme pour les préfixés, les temps de lecture des enfants de niveau CP n'ont pas été analysés car l'effectif pour ce niveau était trop faible (6). La population est ainsi constituée de 29 enfants de niveau CE1, 31 enfants de niveau CE2, 24 enfants de niveau CM1 et 31 enfants de niveau CM2. Le plan d'analyse est similaire à celui établi pour les items préfixés.

Le Tableau 16 regroupe l'ensemble des temps (en millisecondes, ms) mis à lire correctement les items de type suffixé.

Tableau 16. Temps moyens de lecture correcte (et *écart-types*) des items suffixés en fonction du niveau de lecture

Niveau de lecture	Mots affixés	Mots pseudo-affixés	Pseudo-mots affixés	Pseudo-mots pseudo-affixés
CE1	3252,92 (1147,75)	2710,87 (813,74)	2957,54 (847,34)	3329,14 (997,85)
CE2	2203,98 (366,79)	1929,90 (261,42)	2362,67 (508,32)	2480,29 (722,30)
CM1	1838,23 (326,04)	1638,49 (234,34)	1919,81 (302,45)	2039,35 (307,12)
CM2	1892,27 (385,33)	1730,52 (277,41)	1996,38 (440,46)	2134,43 (411,25)
Moyenne	2296,85	2002,44	2309,10	2495,80

Chez les enfants de niveau CE1, les résultats sont très hétérogènes pour les mots pseudo-affixés (1147,75 ms). Les écart-types sont moins importants pour les autres catégories d'items mais restent néanmoins élevés (mots pseudo-affixés : 813,74 ms ; pseudo-mots affixés : 847,34 ms ; pseudo-mots pseudo-affixés : 997,85 ms). Avec l'efficiency en lecture, cette hétérogénéité diminue mais reste cependant importante chez les enfants de niveau CE2 pour les pseudo-mots affixés (508,32 ms) et les pseudo-affixés (722,30 ms).

Le niveau de lecture a un effet significatif sur la rapidité des enfants à lire les items ($F(3, 111) = 33,83$; $p < .0001$). Les lecteurs les plus efficaces sont ceux qui lisent le plus rapidement. La lecture d'un item nécessite en moyenne 3062,62 ms pour un enfant de niveau CE1, 2244,21 ms pour un enfant de niveau CE2, 1858,97 ms pour un enfant de niveau CM1 et enfin 1938,40 ms pour un enfant de niveau CM2. Les différences entre ces temps de lecture correcte sont toutes significatives sauf entre les niveaux CM2 et CM1 ($p < .01$ pour toutes les comparaisons). De plus, les mots sont lus plus vite que les pseudo-mots (2149,65 ms < 2402,45 ms ; $F(1, 111) = 44,73$; $p < .0001$) tandis que la différence entre les suffixés (2302,97 ms) et les pseudo-suffixés n'est pas significative (2249,12 ms ; $F(1, 111) = 2,91$; *ns*).

L'analyse de variance montre une interaction significative entre les facteurs lexicalité et caractère plurimorphémique ($F(1, 111) = 46,71$; $p < .0001$; voir Figure 22).

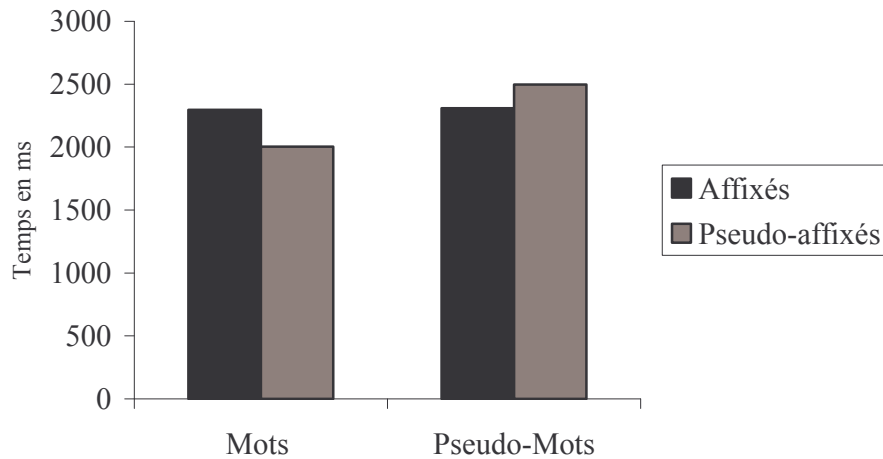


Figure 22. Temps de lecture des items suffixés selon la lexicalité et le caractère plurimorphémique

Cette interaction permet de voir que l'effet de la lexicalité est significatif sur les pseudo-affixés ($p = .0001$) : les mots pseudo-affixés (2002,44 ms) sont lus plus vite que les pseudo-mots pseudo-affixés (2495,80 ms). En revanche, il n'est pas significatif sur les affixés pour lesquels la différence entre les pseudo-mots (2309,10 ms) et les mots (2296,85 ms) n'est pas significative. Cette interaction montre également que les mots suffixés sont lus moins rapidement que les mots pseudo-suffixés (2296,85 ms > 2002,44 ms ; $p = .0033$) tandis que la différence entre les pseudo-mots suffixés (2309,10 ms) et ceux pseudo-suffixés (2495,80 ms) est tendanciellement significative ($p = .0668$).

Enfin, l'interaction entre les trois facteurs, niveau de lecture, lexicalité et caractère plurimorphémique, est significative ($F(3, 111) = 4,33$; $p = .0063$; voir Figure 23).

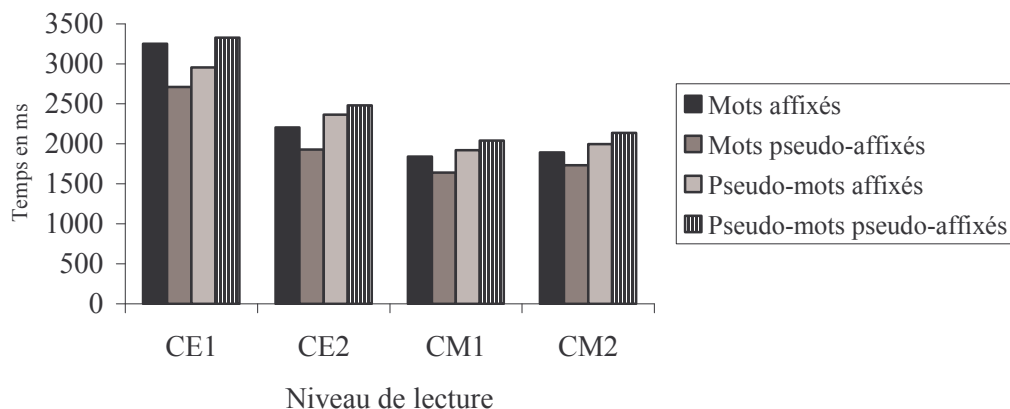


Figure 23. Temps de lecture correcte des items pour les cinq groupes d'enfants en fonction de la lexicalité et du caractère plurimorphémique

Cette interaction indique que les enfants de niveau CM1 ($p = .0187$), CE2 ($p = .0012$) et CE1 ($p = .0426$) lisent les mots suffixés moins rapidement que les mots pseudo-suffixés alors que la différence entre les deux catégories d'items n'est que tendanciellement significative pour les enfants de niveau CM2 ($p = .0627$). En revanche, pour tous les enfants, l'effet du caractère plurimorphémique n'est pas significatif dans le cas des pseudo-mots. L'interaction des trois facteurs confirme la plus grande rapidité des enfants à lire les mots pseudo-suffixés par rapport aux pseudo-mots pseudo-suffixés (CM2 : $p < .0001$; CM1 : $p < .0001$; CE2 : $p < .001$; CE1 : $p = .0123$) et l'absence de différence significative entre mots suffixés et pseudo-mots suffixés au niveau des temps de lecture.

3.4- Discussion

Les différentes expériences réalisées sur la morphologie avaient pour objectif de tester d'une part, l'hypothèse selon laquelle, très précocement (dès la GSM), les enfants développaient une sensibilité aux régularités morphologiques de la langue et possédaient des connaissances implicites et explicites sur les phénomènes de préfixation et de suffixation. D'autre part, il s'agissait d'éprouver l'hypothèse d'une utilisation de la morphologie par les apprentis lecteurs et les lecteurs plus efficaces lors de la reconnaissance de mots familiers et de mots nouveaux.

3.4.1- Développement des connaissances morphologiques implicites et explicites

Pour mettre en évidence la présence de connaissances implicites morphologiques chez l'enfant avant même l'apprentissage de la lecture, une tâche de plausibilité lexicale écrite a été utilisée auprès des enfants de GSM jusqu'au niveau de lecture CM2. Dans cette épreuve implicite, les enfants étaient invités à manifester des choix préférentiels à l'égard de pseudo-mots construits ou non morphologiquement. Ils devaient ainsi indiquer parmi deux pseudo-mots, dont l'un était affixé ou construit morphologiquement (*resibrer*) et l'autre était pseudo-affixé (*risebrer*), lequel ressemblait le plus à un vrai mot. Deux types de préfixes et deux types de suffixes ont été utilisés : [re-] (*cf.* exemple précédent) et [dé-] (*détiner-diténer*) pour les préfixes, [-age] (*vaminage-nagivame*) et [-eur] (*donéveur-veurédon*) pour les suffixes.

Les résultats obtenus indiquent un choix aléatoire pour les prélecteurs de GSM et les apprentis lecteurs de niveau CP tandis qu'à partir du niveau CE1 et jusqu'au niveau CM2, les enfants choisissent de manière préférentielle les pseudo-mots affixés. Ce choix préférentiel s'opère pour les préfixés (*resibrer, détiner*) comme pour les suffixés (*vaminage, donéveur*). Ces résultats confirment en partie ceux déjà obtenus par Colé *et al.* (2003) ainsi que Marec-Breton (2003). En effet, ces auteurs en utilisant une tâche similaire à l'oral avaient constaté que dès 5

ans, les enfants considéraient que les préfixés ressemblaient plus à de vrais mots que les autres items. Une aussi grande précocité n'est pas retrouvée dans la présente étude puisque ce n'est qu'à partir du niveau de lecture CE1 que les enfants témoignent d'une sensibilité à la morphologie des mots écrits. La modalité écrite de l'épreuve implicite utilisée ici peut expliquer cette apparition plus tardive des connaissances morphologiques implicites par rapport aux données de Colé *et al.* (2003) et de Marec-Breton (2003). En effet, malgré une exposition de plus en plus fréquente et précoce aux mots écrits dans le contexte scolaire, les enfants de GSM et de niveau de lecture CP ont pu éprouver une certaine gêne face à des pseudo-mots écrits qui ont le statut de mots inconnus. Cette difficulté liée à la modalité écrite de la tâche pourrait expliquer la non prise en compte de la structure morphologique dans les choix effectués par ces enfants. Il est important de souligner ici que les caractéristiques de la tâche implicite interdisent d'interpréter les scores comme traduisant un niveau, tout juste la préférence pour un type d'item (comme ceux morphologiquement construits) dénote-t-elle la sensibilité à une dimension (morphologique). Pour les enfants de GSM et de niveau CP, d'autres variables ou dimensions interférentes ont pu faire porter le choix sur l'item alternatif (i.e, non construits morphologiquement). En revanche, dès le niveau CE1, une sensibilité à la morphologie écrite existe et l'apprentissage explicite de la lecture semble accroître cette sensibilité comme en témoigne l'augmentation significative des performances avec le niveau de lecture préalable. La sensibilité à la morphologie écrite est indépendante du type d'affixe : les enfants, dès le niveau CE1, préfèrent les items affixés que l'affixe soit un préfixe ou un suffixe. Ce résultat peut néanmoins être nuancé par l'étude des écart-types obtenus pour les suffixés : les différences interindividuelles pour cette catégorie d'items sont importantes pour tous les enfants sauf pour ceux de GSM et de niveau CP. De façon générale, dans une tâche de plausibilité lexicale écrite utilisant des pseudo-mots, la reconnaissance de préfixés ne semble pas facilitée par rapport aux suffixés contrairement aux prédictions de Marec-Breton (2003).

Les résultats obtenus à la seconde épreuve morphologique d'évaluation des connaissances explicites confortent l'hypothèse selon laquelle l'apprentissage de la lecture permet d'amplifier la sensibilité morphologique et favorise notamment la conscience de certaines caractéristiques morphologiques, c'est-à-dire le développement de connaissances morphologiques explicites. Pour évaluer ces connaissances explicites, une tâche explicite écrite de jugement d'intrus dans une liste de triplets a été proposée aux enfants des niveaux CP à CM2. Par exemple, dans le triplet *regagner-redonner-refléter*, *refléter* est un pseudo-préfixé car [re-] ne joue pas le rôle de préfixe et constitue donc l'intrus. Les triplets élaborés exploitaient les préfixes [re-] et [dé-] ainsi que les suffixes [-age] et [-eur]. Les résultats peuvent être rapprochés

de ceux observés pour l'épreuve implicite : le choix reste aléatoire au niveau CP mais dès le CE1, les enfants sont en mesure de détecter l'intrus morphologique pseudo-suffixé et aux niveaux CE2 à CM2, l'intrus est détecté qu'il soit pseudo-suffixé ou pseudo-préfixé. Colé *et al.* (2003) avaient obtenu des résultats proches avec une épreuve similaire en modalité orale mais dans ce cas ce n'était qu'à partir du niveau CE2 que les enfants détectaient l'intrus de façon significative. A nouveau, les enfants les moins efficaces en lecture (niveaux CP et CE1) réussissent moins bien que ceux plus « experts » en lecture. Un résultat surprenant doit néanmoins être noté pour les enfants de niveau CE2 qui ont de meilleures performances que ceux de niveau CM1. Ce résultat s'explique sans doute par la bonne performance des enfants de niveau CE2 aux items de type suffixé (75 % de réussite contre 53 % pour le niveau CM1). Ces mêmes enfants détectent d'ailleurs mieux les intrus de type suffixé que de type préfixé. Le profil inverse est obtenu pour les enfants de niveau CM2 alors que pour les autres enfants la différence entre les deux catégories d'affixes n'est pas significative comme dans l'épreuve implicite. Ici encore, il est difficile d'affirmer que la reconnaissance des préfixés soit facilitée par rapport aux suffixés. Les données obtenues à l'épreuve implicite témoignent en revanche de l'existence dès le niveau CE1 d'une sensibilité à la structure morphologique : ces mêmes enfants possèdent également des connaissances explicites qui apparaissent conjointement et se développent sur la base des connaissances implicites préétablies, conformément au schéma d'apprentissage de Gombert (2003).

L'analyse de la lecture des enfants des niveaux CP à CM2, en prenant en compte l'exactitude et le temps de lecture, montre la présence d'effets de la structure morphologique bien que ces effets soient très localisés.

3.4.2- Traitements morphologiques et niveau de lecture

Dans l'épreuve de lecture à haute voix administrée aux enfants des niveaux CP à CM2, des mots et pseudo-mots affixés et pseudo-affixés étaient proposés. Comme pour les épreuves implicite et explicite, les préfixes [dé-] et [re-] (*dépasser, renifler*) et les suffixes [-age] et [-eur] étaient utilisés (*bagage, bouveur*). Une analyse différentielle des préfixes et des suffixes a été réalisée. Pour l'analyse des temps de lecture correcte, les enfants de niveau CP ont été retirés (*cf.* p. 153).

Pour les items élaborés avec les préfixes [re-] et [dé-], la structure morphologique des items intervient à la fois dans la reconnaissance de mots connus et d'items non familiers (pseudo-mots) pour l'ensemble des enfants. Cet effet de la structure morphologique est prégnant tant pour l'exactitude (enfants de niveau CP à CM2) que pour les temps de lecture (niveau CE1 à CM2). Ces résultats vont dans le sens de ceux obtenus par Marec-Breton *et al.* (2005) qui avaient

constaté que des lecteurs débutants lisaient avec moins d'erreurs des mots préfixés comme *retomber* que des mots pseudo-préfixés tels que *renifler* et que les pseudo-mots préfixés (*remourir*) étaient également mieux lus que ceux pseudo-préfixés (*recounir*). La structure morphologique constituerait donc bien, de manière précoce, une aide notable à la lecture chez les enfants comme l'avaient également suggéré Burani *et al.* (2002). D'une manière générale, les enfants apparaissent également sensibles à la lexicalité des items : les mots sont lus plus rapidement que les pseudo-mots chez tous les enfants, toutefois cet effet varie en fonction du niveau de lecture pour l'exactitude. Chez les enfants de niveau CE2 et CM2, les pseudo-mots sont lus avec davantage d'erreurs que les mots alors que chez les enfants moins efficaces mais aussi pour ceux de niveau CM1, les erreurs de lecture sont aussi nombreuses pour les mots que pour les pseudo-mots. Pour les niveaux les plus avancés en lecture (CE2 et CM2), ce résultat semble logique : ces enfants sont sensibles à la structure morphologique et ont des connaissances explicites bien établies (voir plus haut) et surtout leur niveau d'expertise en lecture permet une récupération rapide des mots connus dans le lexique mental. Pour le niveau CM1, le résultat est en revanche surprenant : il peut s'expliquer notamment par les résultats proches obtenus entre mots pseudo-affixés (87 %) et pseudo-mots affixés (89 %) mais aussi par l'absence de hausse développementale entre les enfants de niveau CE2 (87,50 %) et ceux de niveau CM1 pour les mots pseudo-affixés. Ce résultat peut aussi être mis en parallèle des meilleures performances des enfants de niveau CE2 par rapport au niveau CM1 à la tâche explicite. Les enfants de niveau CM1 semblent éprouver davantage de difficultés que les enfants de niveau CE2 à utiliser leurs connaissances lexicales face à des mots familiers. Pour les autres niveaux de lecture, les enfants les moins efficaces en lecture font logiquement plus d'erreurs par rapport aux enfants des niveaux supérieurs bien que les enfants de niveau CM2 n'aient de meilleures performances par rapport à ceux des niveaux CM1 et CE2 que dans le cas des mots. Pour la vitesse de traitement, les enfants dont le niveau de lecture est le plus faible sont les plus lents pour l'ensemble des items lus.

Pour les items de type préfixé, l'effet de la structure morphologique est soutenu chez tous les enfants à la fois sur l'exactitude à lire les mots connus et les mots nouveaux et sur la rapidité de traitement de ces items. L'analyse des résultats pour les items de type suffixé fournit des résultats différents sans toutefois pouvoir affirmer, comme cela a été fait pour les lecteurs experts (*cf.* Colé *et al.*, 1986), que les deux catégories d'items ne seraient pas traitées de façon équivalente. En effet, dans le cas des suffixes [-age] et [-eur], globalement les items pseudo-suffixés ne sont pas mieux lus que les suffixés et ce pour les enfants des niveaux CP à CM2. Il

n'y aurait donc pas d'effet simple de la structure morphologique. Cependant, la structure morphologique interagit avec la lexicalité des items. Cette interaction montre qu'il n'y a pas d'effet de la structure morphologique dans le cas des mots nouveaux (représentés par les pseudo-mots) sur l'exactitude et les temps de lecture. Le caractère plurimorphémique existe en revanche pour les mots mais cet effet est inverse à celui prédit par l'hypothèse de la présente étude : les mots suffixés comme *gommage* sont lus avec davantage d'erreurs et moins vite que les mots pseudo-suffixés tels que *message*, hormis chez les enfants de niveau CM2 où la différence de vitesse de traitement entre les deux types d'items n'est que tendanciellement significative. Ce résultat est surprenant et ne montre pas comme ceux d'autres études portant sur des populations de lecteurs plus âgés (par exemple, les travaux de Caramazza *et al.*, 1988) une facilitation de la lecture pour tous les items construits morphologiquement. Néanmoins, Marec-Breton *et al.* (2005) ont obtenu chez des enfants de CP et CE1 des résultats très proches pour les items de type suffixé : dans leur étude, les apprentis lecteurs ne lisaient pas mieux les mots suffixés (*danseur*) que les mots pseudo-suffixés (*douleur*). Ainsi, ces résultats comparés à ceux obtenus pour les items de type préfixé suggèrent que le recours à une analyse morphologique pourrait être plus systématique pour les items préfixés que pour les items suffixés : la structure morphologique de type préfixé constitue une aide à la lecture des mots connus et peu familiers alors qu'elle a peu d'impact dans le cas des items de type suffixé. Concernant le facteur de lexicalité des items, les analyses de résultats indiquent que les mots pseudo-suffixés sont lus à la fois avec moins d'erreurs et plus rapidement que les pseudo-mots pseudo-suffixés conformément aux hypothèses. Toutefois, pour les items construits morphologiquement, aucune différence significative en termes d'exactitude et de temps de lecture n'apparaît entre mots et pseudo-mots. Enfin, pour cette épreuve de lecture à haute voix, les enfants les moins efficaces en lecture (niveau CP) font plus d'erreurs que les autres enfants (niveaux CE1 à CM2). Ceux de niveau CE1 lisent moins bien que les enfants de niveau CE2 et CM1 mais seulement dans le cas des affixés et dans tous les cas comparativement aux enfants de niveau CM2. Les enfants de niveau CE2 font également plus d'erreurs sur les affixés par rapport à ceux de niveau CM2. Ces deux résultats concernant les enfants les moins efficaces en lecture semblent indiquer que pour les suffixés la structure morphologique est plus contraignante que facilitatrice pour l'exactitude en lecture. Pour la rapidité de traitement, les enfants des niveaux CE1 et CE2 lisent moins vite que les enfants plus experts en lecture, conformément aux hypothèses.

CHAPITRE 7 – EVOLUTION DES CONNAISSANCES GRAPHOPHONOLOGIQUES, ORTHOGRAPHIQUES ET MORPHOLOGIQUES

1- Expérience 4 : adultes et enfants de niveau CM2

Cette quatrième expérience vise dans un premier temps à mesurer l'étendue des connaissances implicites et explicites possédées par le lecteur expert en graphophonologie (effet de consistance graphèmes-phonèmes), orthographe (effet de fréquence du bigramme) et en morphologie (effet d'affixation). Pour ce faire, les tâches utilisées sont similaires à celles des enfants. L'analyse des traitements réalisés en lecture est effectuée également par une tâche de lecture à haute voix. Dans un deuxième temps, une comparaison des vitesses de traitement en lecture à haute voix adultes-enfants de niveau CM2 sera réalisée afin d'étudier l'évolution de ces traitements entre ces deux types de population.

1.1- Méthode

1.1.1- Population

Trente adultes (âge moyen : 24 ans et 8 mois, écart-type : 3 ans et 3 mois, âge du plus jeune : 20 ans et 3 mois, âge du plus âgé : 30 ans et 10 mois) ont participé à cette expérience en décembre 2003, janvier et février 2004. Aucun des participants adultes ne présentait de troubles de la lecture avérés.

Les enfants de niveau CM2 sont les mêmes que ceux ayant participé aux expériences précédentes.

1.1.2- Matériel et procédure

Les 20 participants adultes ont passé comme les enfants des enfants de niveau CP à CM2 les épreuves contrôles *LUM* et *EVIP* (ces deux épreuves ainsi que la procédure utilisée avec les adultes sont décrites en Annexe 1, pp. 241-244). Pour ces deux tâches, les scores moyens sont élevés. Ainsi, le pourcentage de réussite à *LUM* est de 98,3 %. La moyenne des mots correctement lus est de 98,25 items sur 105 (écart-type : 2,75) et le temps moyen pour lire un mot est de 0,53 secondes. A l'*EVIP*, le score moyen est de 159,57 sur 170 (écart-type : 7,74). La réussite à ce test est de 93,86 %.

Le matériel et la procédure utilisés avec les enfants de niveau CM2 ont déjà été décrits dans le chapitre précédent. Pour les adultes, le matériel utilisé dans le cas des tâches implicites de graphophonologie, orthographe et de morphologie est identique à celui des enfants des niveaux CP à CM2. Seul le mode de présentation diffère : les items sont regroupés dans un tableau (police Arial, Taille 12). Chaque paire correspond ainsi à une ligne de tableau (cf. Annexes 2, 3 et 4, p. 246, p. 250 et p. 255). La passation est collective, les critères de cotation similaires à ceux des enfants et la consigne est la suivante : « *Pour chaque ligne, entourez l’item qui ressemble le plus à un mot* ». Dans le cas de la tâche explicite de graphophonologie, le matériel et la procédure sont strictement identiques à ceux des enfants. A l’épreuve orthographique explicite, le matériel utilisé avec les adultes est exactement similaire à celui des enfants. La passation est individuelle et la consigne fournie est la suivante : « *Regardez bien ces syllabes. Je voudrais que vous m’indiquiez s’il y a plus de mots qui commencent (pour les bigrammes en position initiale) / qui se terminent (pour les bigrammes en position finale) par cette syllabe là « ... » (montrer au participant le premier bigramme de la paire) ou par celle-ci « ... » (montrer au participant le second bigramme de la paire)* ». Pour la tâche morphologique explicite, le contenu est similaire à celui des enfants mais les triplets sont présentés en ligne dans un tableau (cf. Annexe 4, p. 256). La passation est également collective et la cotation identique à celle utilisée avec les enfants. La consigne est : « *Pour chaque ligne, entourez le mot qui ne va pas avec les autres* ». Enfin, pour les trois épreuves de lecture à haute voix sur ordinateur, le matériel utilisé et la procédure sont identiques à ceux des enfants.

Comme pour les enfants, l’ordre de passation a été aléatorisé. Les adultes ont réalisé les différentes épreuves au mois de décembre 2003, janvier et février 2004 en une seule session.

Les hypothèses opérationnelles sont similaires à celles formulées pour les enfants (voir chapitre précédent, expériences 1, 2 et 3). Il est cependant nécessaire d’ajouter que même si les apprentissages implicite et explicite se poursuivent chez l’adulte tant que celui-ci lit, des scores plafonds sont attendus pour les trois types d’épreuves (notamment pour les épreuves explicite et de lecture à haute voix, l’épreuve implicite visant l’expression d’une préférence et non la réussite à une tâche).

1.2- Résultats du groupe adulte

1.2.1- Tâches graphophonologiques

Tâches implicite et explicite

Les performances moyennes (exprimées en pourcentages de réussite) obtenues par le groupe adulte aux tests graphomorphologiques (implicite et explicite) sont regroupées dans le Tableau 17 ci-dessous.

Tableau 17. Performances (et *écart-types*) obtenues aux épreuves graphophonologiques

Graphophonologie implicite	Graphophonologie explicite		
	Irréguliers	Réguliers	Moyenne
51,25 (18,96)	91,67 (8,34)	77 (19,15)	84,33

L'analyse des réponses dans l'épreuve implicite révèle que celles-ci ne se différencient pas du hasard. Ainsi, pour cette tâche de plausibilité lexicale, les participants ne choisissent pas préférentiellement les items ne contenant que des graphèmes réguliers.

Pour l'épreuve explicite, le plan d'analyse est : $S_{30} * R_2$. S est le facteur sujet et R désigne le facteur régularité des items (régulier *vs.* irrégulier). Les résultats sont plus hétérogènes dans le cas des items réguliers (écart-type : 19,15) que pour les items irréguliers (8,34). La régularité de l'item a un effet significatif sur les réponses données ($F(1, 29) = 28,59 ; p < .0001$). Les adultes parviennent plus facilement à fournir les deux types de prononciation (correcte et incorrecte) pour les items irréguliers (91,57 % de réussite) que pour les items réguliers (77 %).

Tâche graphophonologique de lecture à haute voix

Les résultats obtenus au niveau de l'exactitude en lecture atteignent un seuil plafond (95,94 %). Seuls les temps de lecture correcte ont donc fait l'objet d'une analyse. L'analyse des écart-types indique que pour les pseudo-mots irréguliers, la dispersion des temps de lecture correcte est plus forte (245,81 ms) que pour les autres catégories d'items (mots réguliers : 155,74 ms ; mots irréguliers : 125,10 ms ; pseudo-mots réguliers : 145,42 ms).

Les facteurs considérés dans les analyses suivantes sont la lexicalité des items (mots *vs.* pseudo-mots) et la régularité des items (réguliers *vs.* irréguliers). Le plan d'analyse est : $S_{30} * L_2 * R_2$ où S correspond au facteur sujet, L désigne la lexicalité des items et R correspond à la régularité des items.

L'effet de la lexicalité est significatif ($F(1, 29) = 245,18 ; p < .0001$). Les participants lisent les mots (1592,09 ms) plus rapidement que les pseudo-mots (1880,46 ms). Ils lisent également plus vite les items réguliers que ceux irréguliers (1680,34 ms < 1792,21 ms ; $F(1, 126) = 63,48 ; p < .0001$).

La plus grande rapidité constatée pour les réguliers par rapport aux irréguliers n'existe que dans le cas des pseudo-mots (1779,50 ms < 1981,42 ms ; $p = .0003$) comme en témoigne l'interaction significative des facteurs lexicalité et régularité (voir Figure 24 ; $F(1, 29) = 27,16$; $p < .0001$). Cette interaction confirme l'effet de la lexicalité tant pour les réguliers (mots : 1581,19 ms vs. pseudo-mots : 1779, 50 ms) que pour les irréguliers (1602,99 ms vs. 1981,42 ms ; $p < .0001$ dans les deux cas) mais révèle un empan de l'effet bien supérieur pour les irréguliers.

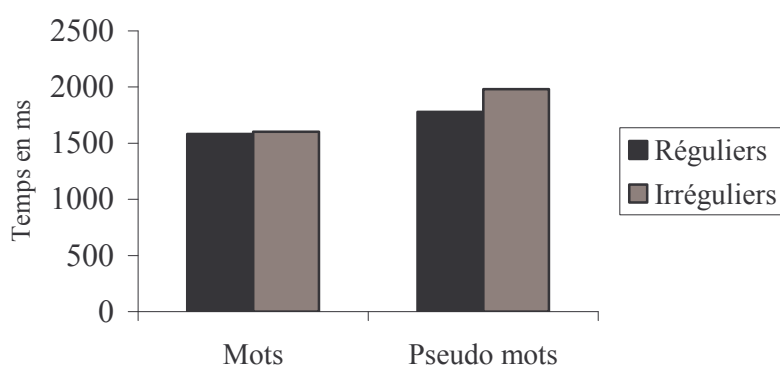


Figure 24. Temps de lecture en fonction de la lexicalité et de la régularité des items.

1.2.2- Tâches orthographiques

Tâches implicite et explicite

Pour la comparaison des résultats obtenus aux deux épreuves (implicite vs. explicite), le plan d'analyse est $S_{30} * E_2$ (S est le facteur sujet et E désigne le type d'épreuve). Pour l'analyse épreuve par épreuve, le plan d'analyse est $S_{30} * P_2$ où S désigne le facteur sujet et P correspond à la position du bigramme (initiale vs. finale).

Le Tableau ci-dessous rend compte des performances obtenues par les adultes aux deux épreuves orthographiques.

Tableau 18. Performances (et *écart-types*) obtenues aux tâches orthographiques

Orthographe implicite			Orthographe explicite		
Initiale	Finale	Moyenne	Initiale	Finale	Moyenne
75 (24,12)	69,17 (23,61)	72,08	83,33 (37,90)	93,33 (25,37)	88,33

La tâche orthographique implicite (72,08 %) est moins bien réussie⁶ que celle explicite (88,33 % ; $F(1, 29) = 12,28$; $p < .01$).

⁶ Par convention, le choix de l'item incluant le bigramme fréquent dans l'épreuve implicite est qualifié de « réussite » alors qu'en toute rigueur il n'y a pas de réussite dans cette épreuve, le choix exprimant une préférence.

Au test implicite, les écart-types sont relativement élevés pour les items à bigrammes en position initiale (24,12) et finale (23,61). Le groupe des adultes préfère significativement les items contenant un bigramme fréquent à ceux avec un bigramme rare que le bigramme soit en position initiale ($p < .0001$) ou en position finale ($p < .0001$). Le choix préférentiel du groupe ne dépend pas de la position des bigrammes ($F(1, 29) = 1,28$; *ns*). Les adultes préfèrent les items à bigramme fréquent que ce bigramme soit en position initiale (75 %) ou en finale (69,17 %).

A l'épreuve explicite, les écart-types indiquent une hétérogénéité des résultats notamment pour les bigrammes en position initiale (37,90). Les adultes considèrent qu'il y a davantage de mots commençant (*ca* vs. *na*) ou se terminant (*te* vs. *ve*) par un bigramme fréquent dans la langue ($p < .0001$ dans les deux cas). Les adultes réussissent aussi bien pour les bigrammes en position initiale (88,33 %) que pour ceux en position finale (93,33 % ; $F(1, 29) = 1,85$; *ns*).

Tâche orthographique de lecture à haute voix

De nouveau, l'épreuve orthographique de lecture à haute voix a été très bien réussie par les adultes en termes d'exactitude : 97,71 % des items dont le bigramme fréquent ou rare est en position initiale et 97,92 % des items bigrammes en position finale sont lus correctement. Ainsi, comme pour la graphophonologie, seuls les temps de lecture correcte font l'objet d'une analyse avec pour facteurs la lexicalité des items (mots vs. pseudo-mots) et la fréquence du bigramme (fréquent vs. rare). Le plan d'analyse est : $S_{30} * L_2 * F_2$ (S : facteur sujet ; L : facteur la lexicalité des items ; F : facteur fréquence du bigramme).

Position initiale

La dispersion des temps de lecture correcte pour les pseudo-mots à bigramme fréquent (145,44 ms) et à bigramme rare (149,35 ms) est plus importante que celle des temps de lecture pour les mots (bigramme fréquent : 110,03 ms ; bigramme rare : 125,74 ms).

Les deux facteurs principaux ont des effets significatifs sur les performances (lexicalité : $F(1, 29) = 50,69$; $p < .0001$; fréquence du bigramme : $F(1, 29) = 21,77$; $p < .0001$). Les mots (comme *camion* ou *narine* : 1533,79 ms) sont lus plus vite que les pseudo-mots (tels que *caloude* ou *narufe* : 1644,48 ms). Les items contenant un bigramme fréquent (*carotte* ou *cadute* : 1567,15 ms) sont également lus plus rapidement que ceux avec un bigramme rare (*nature* ou *nalube* : 1631,12 ms). L'interaction entre ces facteurs n'est pas significative ($F(1, 29) < 1$).

Position finale

De façon similaire aux items à bigramme en position initiale, les temps de lecture correcte sont hétérogènes pour les pseudo-mots (bigramme fréquent : 167,76 ms ; rare : 155,33 ms) comparativement aux mots (fréquent : 121,74 ms ; rare : 116,95 ms).

Pour les items dont le bigramme rare ou fréquent est en position finale, les analyses montrent également un effet significatif du facteur lexicalité ($F(1, 29) = 93,34 ; p < .0001$). Les pseudo-mots sont lus moins vite que les mots (1723,01 ms > 1567,12 ms). En revanche, la différence entre les items à bigramme fréquent (1646,18 ms) et les items à bigramme rare (1643,94 ms) n'est pas significative ($F(1, 29) < 1$).

Le facteur fréquence du bigramme interagit de façon significative avec la lexicalité ($F(1, 29) = 4,29 ; p < .05$). La lecture est plus rapide pour les mots que pour les pseudo-mots que le bigramme en position initiale soit fréquent ($p < .0001$) ou rare ($p = .0004$) mais l'empan de cet effet est plus important pour les bigrammes fréquents (1557,70 ms vs. 1734,66 ms) que pour les bigrammes rares (1576,53 ms vs. 1711,35 ms), comme le montre la Figure 25.

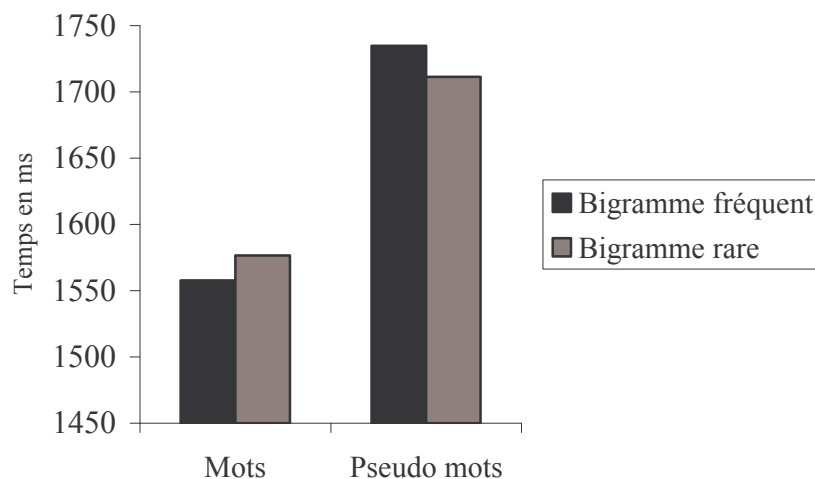


Figure 25. Temps de lecture selon la lexicalité des items et la fréquence du bigramme en position finale

1.2.3- Tâches morphologiques

Tâches implicites et explicites

Le plan d'analyse pour ces deux tâches est $S_{30} * A_2$. Le facteur S correspond aux sujets (participants) et le facteur A correspond au type d'affixation des items (préfixés vs. suffixés).

Les performances moyennes (exprimées en pourcentages de réussite) obtenues par le groupe adulte aux épreuves morphologiques (implicite et explicite) figurent dans le Tableau 19 ci-après.

Tableau 19. Performances (et *écart-types*) obtenues aux deux tests morphologiques en fonction du type d’affixe

Morphologie implicite			Morphologie explicite		
Préfixés	Suffixés	Moyenne	Préfixés	Suffixés	Moyenne
87,92 (14,11)	79,58 (14,86)	83,75	70,83 (25,29)	53,75 (26,29)	62,29

A la tâche implicite de plausibilité lexicale, les réponses des participants se distinguent du hasard pour les paires de type préfixé (*détiner vs. diténer* ; $p < .00$) ou de type suffixé (*donéveur vs. veurédon* ; $p < .00$). Les adultes choisissent ainsi majoritairement les items morphologiquement construits : ils préfèrent l’item morphologique dans 87,92 % des cas pour les préfixés et dans 79,58 % des cas pour les suffixés. Ces résultats indiquent que les choix préférentiels sont facilités dans le cas des paires préfixées par rapport à celles suffixées ($F(1, 29) = 7,34$; $p = .011$).

Dans la tâche explicite de détection d’intrus morphologiques, le pourcentage d’intrus correctement détectés n’est que de 62,29 %. Les résultats à cette épreuve explicite sont plus hétérogènes que ceux obtenus à l’épreuve implicite comme en témoignent les écart-types (*cf.* Tableau 19). Les réponses fournies par les participants se distinguent du hasard tant pour les suffixés que les préfixés ($p < .00$ dans les deux cas). Le type d’affixe a un effet significatif sur les performances ($F(1, 29) = 13$; $p < .01$) : les intrus de type préfixé (70,83 %) sont bien mieux détectés que ceux de type suffixé (53,75 %).

Tâche morphologique de lecture à haute voix

Cette épreuve a été globalement très bien réussie par les adultes. En effet, ils lisent sans faire d’erreurs 98,75 % des items morphologiques (99,17 % pour les items de type suffixé ; 98,33 % pour ceux de type préfixé). A nouveau, les analyses suivantes ne porteront donc que sur les temps de lecture correcte avec deux facteurs principaux : la lexicalité des items (mots *vs.* pseudo-mots) et le caractère plurimorphémique (affixés *vs.* pseudo-affixés). Le plan d’analyse est le suivant : $S_n * L_2 * M_2$ où S est le facteur sujets, L correspond au facteur lexicalité des items et M désigne le facteur caractère plurimorphémique des items.

Préfixés

Les analyses suivantes concernent 29 adultes. En effet, les résultats obtenus par un adulte ont été retirés car pour une catégorie moins de deux items ont pu être lus correctement.

La Figure 26 indique les temps de lecture moyens (en ms) pour chaque catégorie d'items (mot affixé : *déplier*, mot pseudo-affixé : *renifler*, pseudo-mot affixé : *débriser*, pseudo-mot pseudo-affixé : *rebaner*).

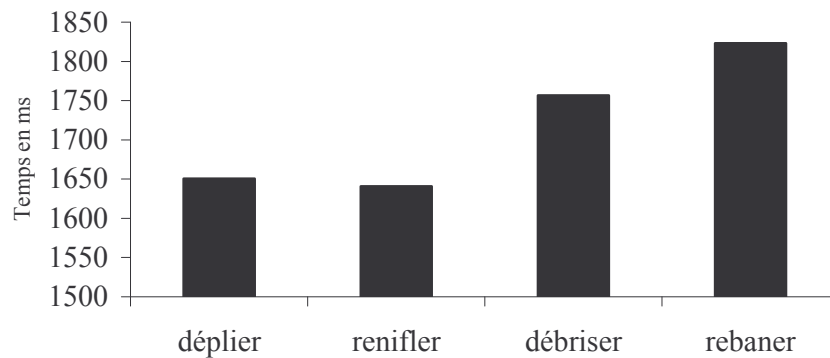


Figure 26. Temps de lecture correcte des items préfixés

Les performances les plus hétérogènes concernent les pseudo-mots pseudo-affixés (écart-type : 220,39 ms) comparativement aux autres catégories d'items (mots affixés : 155,96 ms ; mots pseudo-affixés : 184 ms ; pseudo-mots affixés : 185,13 ms).

Les adultes experts lisent plus rapidement les mots (1645,54 ms) que les pseudo-mots (1789,82 ms ; $F(1, 28) = 59,29$; $p < .0001$). L'effet du caractère plurimorphémique n'est pas significatif ($F(1, 28) = 3$; *ns*) : les items préfixés (1703,52 ms) ne sont pas lus plus vite que les pseudo-préfixés (1731,84 ms). L'interaction entre la lexicalité et l'affixation n'est pas non plus significative ($F(1, 28) = 3,05$; *ns*).

Suffixés

Comme pour les items préfixés, la dispersion des temps de lecture correcte est plus importante pour les pseudo-mots pseudo-affixés (214,59 ms) que pour le reste des items (mots affixés : 165,89 ms ; mots pseudo-affixés : 138,07 ms ; pseudo-mots affixés : 175,33 ms).

L'Anova suivante est réalisée sur les temps de lecture correcte des 30 adultes participants. La différence entre les suffixés (1620,49 ms) et les pseudo-suffixés (1646,25 ms) n'est pas significative ($F(1, 29) = 2,27$; *ns*). Les mots sont lus plus rapidement que les pseudo-mots (1572,39 ms < 1694,36 ms ; $F(1, 29) = 48,36$; $p < .0001$).

Les facteurs lexicalité et caractère plurimorphémique interagissent significativement ($F(1, 29) = 10,71 ; p < .01$; voir Figure 27).

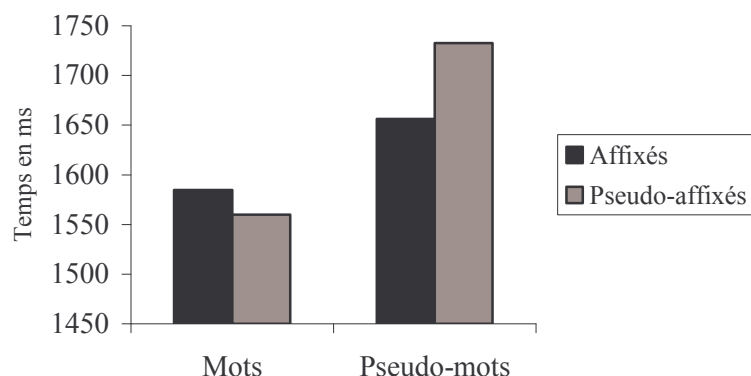


Figure 27. Temps de lecture des items suffixés en fonction de la lexicalité et de l’affixation

Cette interaction montre que seuls les pseudo-mots pseudo-affixés (1732,54 ms) sont lus plus lentement que les mots pseudo-affixés (1559,96 ms ; $p = .0004$). Malgré l’absence d’effet significatif du caractère plurimorphémique pour les mots comme pour les pseudo-mots, l’analyse fait donc apparaître que l’effet de lexicalité n’affecte que les items pseudo-affixés.

1.3- Discussion des résultats du groupe adulte

Les différentes expériences effectuées dans le chapitre précédent indiquent que le développement des connaissances en graphophonologie, orthographe mais aussi en morphologie n’est pas achevé. Ainsi, les résultats obtenus aux tâches d’évaluation des connaissances implicites et explicites s’améliorent avec le niveau d’expertise en lecture (hormis pour les connaissances orthographiques explicites) mais le développement de ces connaissances est loin d’être terminé. Par exemple, en graphophonologie, à la tâche implicite de plausibilité lexicale écrite, les items consistants ne sont choisis de manière préférentielle que dans 61,33 % des cas au niveau CM2 et à l’épreuve d’analyse graphophonologique explicite le pourcentage de réussite est seulement de 75,47 % pour ces mêmes enfants.

Cette quatrième expérience visait donc à voir ce qu’il en est chez des lecteurs experts : il s’agissait essentiellement d’étudier l’évolution des connaissances implicites et explicites en graphophonologie, orthographe et morphologie.

1.3.1- Graphophonologie

1.3.1.1- Développement des connaissances graphophonologiques implicites et explicites

Pour évaluer les connaissances graphophonologiques implicites et explicites des lecteurs experts, les tâches utilisées étaient similaires à celles des enfants : une tâche de plausibilité lexicale écrite et un test d'analyse graphophonologique explicite.

Les performances obtenues à l'épreuve implicite montrent que les choix des participants ne se différencient pas du hasard. Ainsi, les adultes ne choisissent pas préférentiellement les items ne contenant que des graphèmes réguliers comme ressemblant le plus à de vrais mots. Ces résultats indiqueraient que les choix des adultes ne reposeraient pas sur la prise en compte de la structure graphophonologiques des items mais que les performances pourraient être déterminées (de façon implicite) par d'autres dimensions que celles testées initialement. Cette absence de sensibilité à la consistance graphophonologique pourrait également s'expliquer par la nature même de la tâche. Les adultes sont peu habitués à traiter des mots non familiers du fait même de leur expertise en lecture. Les pseudo-mots proposés ont pu les déstabiliser et favoriser des choix aléatoires. Les résultats observés au test explicite confortent cette hypothèse : sur un matériel familier ne contenant que des mots, les performances sont bien plus élevées (84,33 % de réussite). Face à un item régulier (*chose*) et un item irrégulier (*chorale*), fournir les deux types de prononciation, c'est-à-dire la bonne – /ʃoz/ et /kɔʁal/ – et la mauvaise façon – /koz/ et /ʃɔʁal/ –, est jugé plus facile pour les adultes dans le cas de l'irrégulier, conformément aux hypothèses. « Régulariser » un item irrégulier est considéré comme plus simple à effectuer pour les adultes qu' « irrégulariser » un item régulier. Cependant, les scores des adultes ne plafonnent pas : tous les lecteurs ne disposent pas de connaissances suffisantes sur la consistance graphèmes-phonèmes pour ne pas faire d'erreurs. De plus, l'écart-type important obtenu pour les mots réguliers indique de fortes disparités au niveau des performances. Pour certains adultes, « irrégulariser » un mot régulier a pu être déstabilisant en raison justement de l'expertise en lecture : les correspondances graphèmes-phonèmes dominantes sont appliquées de manière automatique et il est dès lors difficile pour un lecteur expert d'aller à leur rencontre face à un mot régulier.

1.3.1.2- Traitements graphophonologiques et expertise en lecture

L'analyse des résultats obtenus en lecture à haute voix permet de vérifier si chez les adultes experts la structure graphophonologique est utilisée et facilite le traitement des mots familiers et celui de mots rencontrés pour la première fois (i.e., les pseudo-mots). L'épreuve de lecture est identique à celle utilisée avec les enfants mais pour les adultes, seuls les temps de lecture ont été

utilisés car les performances en termes d'exactitude atteignaient un seuil plafond. Conformément aux hypothèses, la structure graphophonologique est utilisée par les lecteurs experts mais uniquement dans le cas des mots non familiers représentés par les pseudo-mots. Ainsi, pour les mots, les réguliers sont lus aussi rapidement que les irréguliers. Cet effet de consistance ou de régularité observé seulement sur les mots inconnus indique qu'à un certain niveau d'expertise les mots réguliers et les mots irréguliers sont traités de façon identique : à la régularité s'ajoute la fréquence d'exposition, déjà mise en évidence par Sprenger-Charolles *et al.* (2003) chez des enfants de fin niveau CM1, qui permettrait une récupération automatique des items (réguliers ou irréguliers) dans le lexique mental. Cette fréquence d'exposition favoriserait l'augmentation de la familiarité aux mots chez les lecteurs experts qui est d'ailleurs remarquable *via* l'effet de lexicalité observé : les mots, qu'ils soient réguliers ou irréguliers, sont lus bien plus vite que les pseudo-mots.

1.3.2- Orthographe

1.3.2.1- Développement des connaissances orthographiques implicites et explicites

Afin de vérifier la sensibilité à la structure orthographique et le niveau de connaissances orthographiques explicites des lecteurs experts, les tests utilisés avec les enfants ont été répliqués avec les adultes. Les lecteurs experts se sont donc vus proposer l'épreuve implicite de plausibilité lexicale écrite ainsi que la tâche explicite de jugement de bigrammes présentés de manière isolée.

Comme pour les enfants de niveau CM2, qui peuvent être rapprochés des adultes en termes d'expertise en lecture, les lecteurs experts choisissent de manière préférentielle les items contenant un bigramme fréquent comme calufe ou latote que ceux avec un bigramme rare tels que nalufe ou latove. Les connaissances implicites des lecteurs experts sur la fréquence de certaines séquences de lettres semblent donc bien développées. Cependant, leurs performances ne plafonnent pas et ne sont pas si éloignées de celles obtenues par les enfants de niveau CM2 au même test (respectivement 72,08 % vs. 64,06 %). Contrairement aux enfants, la position du bigramme dans le pseudo-mot (initiale ou finale) n'a pas d'effet sur le choix préférentiel des adultes. Cette absence d'effet de la position du bigramme peut trouver une explication potentielle dans la manière dont les items sont traités par les lecteurs experts : l'identification s'effectue de façon globale et non de manière décompositionnelle sous forme d'attaque et de rime. Dès lors, l'attaque ne serait pas perçue comme un élément plus saillant que la rime (respectivement assimilables ici aux bigrammes en position initiale et finale). Il faut noter cependant que pour les

items à bigramme en position initiale comme en position finale, les écart-types sont relativement élevés indiquant que si la majorité des adultes choisissent préférentiellement les items à bigramme fréquent, pour de nombreux lecteurs experts le choix reste aléatoire.

Concernant les connaissances explicites évaluées par la tâche de jugement de bigrammes isolés, les performances des adultes sont bien plus élevées que celles des enfants de niveau CM2 (88,33 % vs. 60,94 %). Les résultats des lecteurs experts ne présentent pas de scores plafonds mais ils considèrent tout de même que davantage de mots commencent par [ca-] que par [na-] et qu'il existe également plus de mots se terminant par [-te] que par [-ve]. Comme pour l'épreuve implicite, la position du bigramme n'a pas d'effet sur les performances. Enfin, une des hypothèses de la présente étude prédisait que les connaissances explicites, fortement dépendantes de l'apprentissage de la lecture, devaient se développer plus tardivement que les connaissances implicites (Gombert, 2003). Chez les adultes, de par leur expertise en lecture, il serait envisageable de considérer que connaissances implicites et explicites pourraient être au même niveau de développement. Or, les résultats indiquent que les lecteurs experts ont des scores plus bas au test implicite qu'à la tâche explicite. Le moindre score au test implicite peut s'expliquer, comme cela a été fait précédemment pour les connaissances graphophonologiques implicites, par le caractère déstabilisant des pseudo-mots pour les adultes. De plus, malgré une sensibilité à la structure orthographique attestée par les choix préférentiels des adultes pour les items à bigrammes fréquents, certaines préférences ont pu parfois être déterminées (implicitement) par d'autres facteurs que ceux purement orthographiques.

1.3.1.2- Traitements orthographiques et expertise de lecture

Si la fréquence des bigrammes n'avait pas d'effet sur l'exactitude et le temps de lecture chez les enfants, son effet est en revanche significatif chez les lecteurs experts. Les mots et les pseudo-mots contenant un bigramme fréquent en position initiale sont lus plus rapidement que ceux à bigramme rare. Chez les enfants, face à l'absence d'effet de la structure orthographique, il avait été supposé que leurs représentations orthographiques pouvaient être insuffisamment développées et/ou que les correspondances phonologiques des bigrammes rares ou fréquents étant régulières, les enfants se servaient plutôt de leurs connaissances graphophonologiques pour lire les items. Les résultats obtenus avec les lecteurs experts rendent cette hypothèse plausible : l'effet de la structure orthographique sur le temps de lecture semble indiquer que les adultes mobiliseraient leurs représentations orthographiques et notamment leurs connaissances explicites, bien développées à ce niveau, pour lire les items. Ces connaissances faciliteraient alors le traitement des items contenant un bigramme fréquent dans la langue.

Lors de l'évaluation des connaissances implicites et explicites, la position du bigramme n'influçait pas les performances des adultes. Toutefois, en lecture, cette position semble avoir un effet dans la mesure où les résultats obtenus pour les items à bigramme en position finale diffèrent de ceux à bigramme en position initiale. Ainsi, dans l'épreuve de lecture à haute voix d'items à bigramme en position finale, seule la lexicalité des items a une influence sur la rapidité de lecture : l'empan de cet effet est plus important pour les items à bigramme fréquent mais de façon générale la fréquence des bigrammes en position finale ne modifie pas les temps de lecture.

Les derniers résultats concernant les adultes portent sur le niveau de sensibilité morphologique de ces lecteurs experts ainsi que sur les traitements morphologiques opérés lors de la lecture à haute voix.

1.3.3- Morphologie

1.3.3.1- Développement des connaissances morphologiques implicites et explicites

Comme dans l'expérience réalisée auprès des enfants, deux tâches étaient proposées : une épreuve implicite de plausibilité lexicale écrite et une tâche explicite de détection d'intrus morphologiques.

Les résultats au test implicite sont proches de ceux obtenus par les enfants et indiquent une sensibilité des adultes à la morphologie. Ainsi, les lecteurs experts préfèrent les items affixés aux pseudo-affixés. Néanmoins, les scores obtenus montrent que le développement de ces connaissances morphologiques n'est pas achevé avec 83,75 % de choix préférentiels. Ce développement inachevé est d'ailleurs largement confirmé par les performances obtenues à l'épreuve explicite avec seulement 62,29 % de bonnes réponses. Peu de progrès sont constatés par rapport aux enfants de niveau CM2 qui détectaient correctement l'intrus morphologique dans 59,77 % des cas. Sur cette épreuve explicite, les écart-types sont élevés comparativement à ceux du test implicite. *A posteriori*, en interrogeant les adultes sur la manière dont ils s'y étaient pris pour trouver l'intrus, certains ont indiqué qu'ils recherchaient celui qui d'un point de vue sémantique n'allait pas avec les autres alors que les principes de dérivation nécessaires à la réussite de cette tâche étaient expliqués au préalable. La recherche d'indices sémantiques et non purement morphologiques pourrait expliquer ces « faibles » résultats. Concernant un éventuel traitement différent entre préfixés et suffixés, alors que pour les enfants aucune différence entre les deux catégories d'affixes n'apparaissait et qu'il fallait attendre le niveau CM2 pour que les préfixés soient traités plus facilement que les suffixés (épreuve explicite), les adultes ne traitent

pas de façon équivalente préfixés et suffixés. En effet, les items de type préfixé sont mieux réussis que ceux de type suffixé, contrairement aux résultats obtenus par Colé *et al.* (1986) chez des lecteurs experts.

1.3.3.2- Traitements morphologiques et expertise en lecture

Suite aux travaux et à la modélisation de Taft et Forster (1975, 1976), de nombreux chercheurs se sont intéressés à la façon dont l'adulte traitait les mots complexes lus. La plupart des ces chercheurs ont ainsi montré que le lecteur expert adulte effectue automatiquement une analyse morphologique des mots complexes qu'il lit (Alvarez *et al.*, 2001 ; Bertram & Hyona, 2003 ; Bertram *et al.*, 2000 ; Colé *et al.*, 1997 ; Hyona *et al.*, 2002). Le lecteur expert, face à des mots morphologiquement complexes utiliserait une procédure de décomposition morphologique, conformément au modèle de Taft et Forster.

Pourtant, les résultats obtenus dans la présente étude ne confirment pas l'existence de traitements morphologiques au cours de la lecture chez les lecteurs experts. Pour les mots comme pour les pseudo-mots, les items affixés ne sont pas lus plus rapidement que les pseudo-affixés. Seule la lexicalité affecte le traitement des items : les mots sont lus plus vite que les pseudo-mots pour la catégorie préfixe. Pour la catégorie suffixe, seuls les mots pseudo-affixés sont lus plus rapidement que les pseudo-mots pseudo-affixés. Ce résultat en lecture à voix haute est surprenant et ne conforte pas les données obtenues par Marec-Breton (2003) : pour une population similaire en termes d'expertise, elle avait constaté que les lecteurs lisaient plus rapidement les items affixés que ceux non affixés pour les mots inconnus (pseudo-mots) mais pas pour les mots familiers.

Les résultats discutés ici concernaient les adultes. Le point suivant vise à comparer les procédures de lecture à haute voix des adultes à celles des enfants de niveau CM2 afin de déterminer si les traitements effectués entre ces deux groupes de lecteurs présentent des similarités.

1.4- Analyses comparatives adultes-enfants de niveau CM2

Ces analyses comparent les résultats en lecture à haute voix des enfants de niveau de lecture CM2 et des adultes pour mettre en évidence une éventuelle évolution dans les traitements graphophonologiques, orthographiques et morphologiques entre les enfants considérés encore comme apprentis lecteurs et les lecteurs experts.

1.4.1- Graphophonologie

Comme pour les adultes précédemment, les analyses ne portent ici que sur les temps de lecture correcte, le pourcentage de mots lus correctement étant relativement élevé pour les deux groupes confondus (93,38 %). Ces analyses prennent en compte trois principaux facteurs : le groupe (deux modalités : niveau CM2, adultes), la lexicalité des items (mots *vs.* pseudo-mots) et la régularité des items (réguliers *vs.* irréguliers). Le plan d'analyse est le suivant : $S_n < G_2 > * L_2 * R_2$ (S : facteur sujet ; G : facteur groupe ; L : lexicalité des items ; R : régularité des items). Les Anova sur les temps de lecture correcte concernent 30 adultes et 32 enfants de niveau CM2.

Les temps de lecture correcte exprimés en millisecondes figurent dans le Tableau 20.

Tableau 20. Temps moyens de lecture (et *écart-types*) pour les différentes catégories d'items

	Mots réguliers	Mots irréguliers	Pseudo-mots réguliers	Pseudo-mots irréguliers
CM2	1798,33 (254,62)	2163,66 (350,91)	2190,36 (449,56)	2418,54 (472,51)
Adultes	1581,19 (156,74)	1602,99 (125,10)	1779,50 (145,42)	1981,42 (245,81)
Moyenne	1689,76	1883,33	1984,93	2199,98

Les effets des trois facteurs sur les temps de lecture des participants sont significatifs (groupe : $F(1, 60) = 15,89$; $p < .001$; lexicalité : $F(1, 60) = 13,36$; $p < .001$; régularité des items : $F(1, 60) = 6,97$; $p < .05$). Ainsi, les enfants de niveau CM2 sont plus lents que les adultes (2142,72 ms > 1736,28 ms). Tous groupes confondus, les mots (1786,64 ms) sont lus plus vite que les pseudo-mots (2092,46 ms) et les items réguliers sont lus avec une plus grande rapidité que les irréguliers (1837,34 ms *vs.* 2041,65 ms). Aucune interaction n'est significative.

1.4.2- Orthographe

Les participants ont également bien réussi cette épreuve de lecture à haute voix avec 95,31 % de réussite. Les Anova suivantes concernent à nouveau les temps de lecture correcte avec trois facteurs : le groupe (deux modalités : niveau CM2, adultes), la lexicalité des items (mots *vs.* pseudo-mots) et la fréquence du bigramme (fréquent *vs.* rare). Le plan est donc : $S_{30} < G_2 > * L_2 * F_2$ (S : facteur sujet ; G : facteur groupe ; L : lexicalité des items ; F : facteur fréquence bigrammique).

Position initiale

L'échantillon est ici constitué de 30 adultes et 30 enfants de niveau CM2. Le Tableau 21 rend compte des temps de lecture correcte réalisés par ces deux groupes.

Tableau 21. Temps moyens de lecture en ms (et *écart-types*) des items à bigramme en position initiale chez les enfants de niveau CM2 et les lecteurs experts

	cabane	navire	cadute	nalube
CM2	1674,72 (305,11)	1728,22 (317,76)	1956,12 (410,44)	1971,80 (373,35)
Adultes	1494,82 (110,03)	1572,77 (125,74)	1639,48 (145,44)	1689,48 (149,35)
Moyenne	1584,77	1650,50	1797,80	1830,64

Les enfants de niveau CM2 lisent plus lentement que les lecteurs experts (1832,72 ms > 1599,14 ms ; $F(1, 58) = 14,50$; $p < .001$). Les effets des facteurs lexicalité et fréquence du bigramme sur les temps de lecture correcte sont significatifs ($F(1, 58) = 76,24$; $p < .0001$ et $F(1, 58) = 12,07$; $p < .001$). Les pseudo-mots sont lus plus lentement que les mots (1814,22 ms vs. 1617,63 ms). Tous groupes confondus, les items à bigramme fréquent (*cabane*, *cadute* : 1691,28 ms) sont lus avec une plus grande rapidité que ceux à bigramme rare en position initiale (*navire*, *nalube* : 1740,57 ms). L'interaction entre le groupe et la lexicalité est significative ($F(1, 58) = 8,57$; $p < .001$). La Figure 28 montre pourtant que les adultes lisent plus vite les mots (1533,79 ms) comme les pseudo-mots (1664,48 ms) que les enfants (respectivement 1701,47 ms ; $p = .0078$ et 1963,96 ms ; $p = .0001$). Néanmoins, l'écart de temps entre les lecteurs experts et les apprentis lecteurs est plus important pour les pseudo-mots (299,48 ms) que pour les mots (167,68 ms). De plus, l'écart entre les deux types d'items est plus faible pour les adultes (130,69 ms) que pour les enfants (262,49 ms). Les autres interactions ne sont pas significatives.

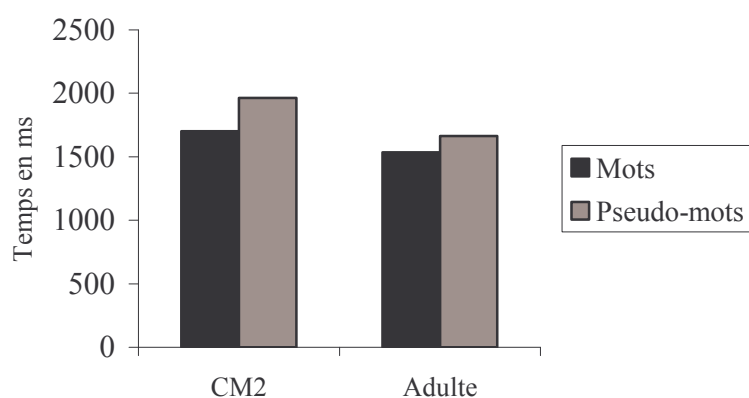


Figure 28. Temps de lecture des mots et pseudo-mots pour les deux groupes

Position finale

Le Tableau ci-dessous regroupe les temps de lecture obtenus par les participants.

Tableau 22. Temps moyens de lecture en ms (et *écart-types*) des items à bigramme en position finale chez les enfants de niveau CM2 et les lecteurs experts

	tomate	lessive	bofite	lutive
CM2	1731,21 (332,39)	1719,73 (244,35)	2011,98 (363,31)	2153,28 (502,52)
Adultes	1557,70 (121,74)	1576,53 (118,95)	1734,66 (167,76)	1711,35 (155,33)
Moyenne	1644,45	1648,13	1873,32	1932,31

L'effet de la fréquence du bigramme (fréquent *vs.* rare) n'est pas significatif ($F(1,59) = 2,69$; *ns*) : les items avec un bigramme fréquent comme *tomate* et *bofite* (1758,89 ms) ne sont pas lus plus rapidement que les items avec un bigramme rare en position finale tels que *lessive* et *lutive* (1790,22 ms). Les effets des facteurs groupe et lexicalité sont significatifs ($F(1, 59) = 17,32$; $p < .001$ et $F(1, 59) = 81,98$; $p < .0001$). Les adultes sont plus rapides que les enfants de niveau CM2, toutes catégories d'items confondues (1645,06 ms < 1904,05 ms) et les mots sont lus plus vite que les pseudo-mots (1646,29 ms *vs.* 1902,82 ms). Ces deux facteurs interagissent significativement ($F(1, 59) = 12,62$; $p < .001$) ainsi que les facteurs lexicalité et la fréquence bigrammique de façon tendancielle ($F(1, 59) = 3,20$; $p = .079$). L'interaction entre le groupe et la fréquence du bigramme n'est pas significative ($F(1, 59) = 3,09$; *ns*) mais l'interaction des trois facteurs est significative ($F(1, 59) = 9,89$; $p = .01$). La Figure 29 révèle que pour chaque catégorie d'items, les enfants de niveau CM2 restent plus lents que les adultes (mots à bigramme fréquent : 1731,21 ms > 1557,70 ms ; $p = .0145$; mots à bigramme rare : 1719,73 ms > 1576,53 ms ; $p = .0053$; pseudo-mots à bigramme fréquent : 2011,98 ms > 1734,66 ms ; $p = .0003$; pseudo-mots à bigramme rare : 2153,28 ms > 1711,35 ms ; $p < .0001$). Cependant, l'écart entre les mots et pseudo-mots, quelle que soit la fréquence du bigramme, est bien plus important dans les cas des lecteurs de niveau CM2 (357,16 ms) que dans celui des lecteurs adultes (155,89 ms). La Figure 29 montre aussi que les adultes restent plus rapides que le niveau CM2 dans la lecture de mots ($p = .0072$) et de pseudo-mots ($p = .0001$) mais l'écart de temps entre lecteurs experts et apprentis lecteurs est plus faible pour les mots (158,35 ms) que pour les pseudo-mots (359,62 ms).

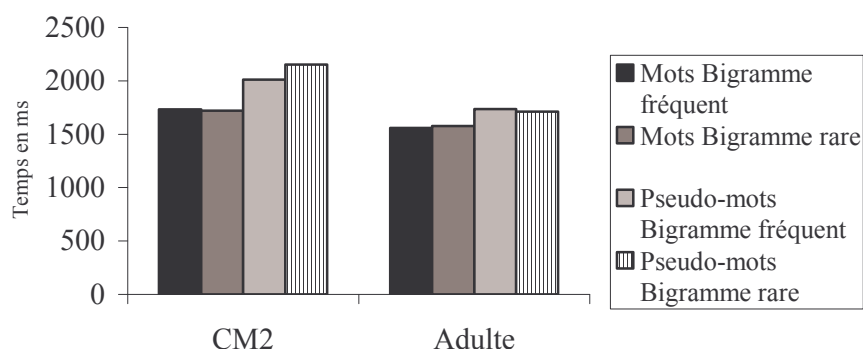


Figure 29. Temps de lecture correcte de l'ensemble des items par le niveau CM2 et le groupe adulte

1.4.3- Morphologie

Cette épreuve est également bien réussie par les participants (96,59 %). Trois facteurs sont mis en jeu : le groupe (deux modalités : niveau CM2, adultes), la lexicalité des items (mots vs. pseudo-mots) et le caractère plurimorphémique (affixés vs. pseudo-affixés). Le plan pour ces analyses est le suivant : $S_n < G_2 > * L_2 * M_2$, avec S le facteur sujets, G le facteur groupe, L correspond à la lexicalité des items et enfin M désigne le facteur caractère plurimorphémique des items.

Préfixés

Les temps de lecture correcte des items préfixés sont regroupés dans le Tableau ci-dessous. Ces temps concernent 29 adultes et 31 enfants de niveau CM2.

Tableau 23. Temps de lecture en ms (et *écart-types*) des items préfixés chez les enfants de niveau CM2 et les lecteurs experts

	dépasser	déguiser	décasser	dénaller
CM2	1957,47 (412,02)	1990,36 (399,58)	2088,59 (390,93)	2350,42 (407,43)
Adultes	1650,53 (155,96)	1640,55 (184)	1756,51 (185,13)	1823,12 (220,39)
Moyenne	1809,12	1821,28	1928,09	2095,56

Les adultes (1717,68 ms) lisent plus vite que les enfants de niveau CM2 (2096,71 ms ; $F(1, 58) = 21,12$; $p < .0001$). Tous groupes confondus, les pseudo-mots sont lus plus lentement que les mots (2004,66 ms > 1809,73 ms ; $F(1, 58) = 20,06$; $p < .0001$). En revanche, la différence entre les préfixés comme *dépasser* ou *décasser* (1863,28 ms) et les pseudo-préfixés

tels que *déguiser* et *dénaller* (1951,11 ms) n'est que tendanciellement significative ($F(1, 58) = 3,87 ; p = .054$). Aucune interaction n'est significative.

Suffixés

L'effectif est ici constitué de 30 adultes et 31 enfants de niveau CM2. Le Tableau 24 rend compte des différents temps obtenus par ces participants pour les items suffixés.

Tableau 24. Temps moyens de lecture en ms (et *écart-types*) des items préfixés chez les enfants de niveau CM2 et les lecteurs experts

	gommage	message	mordage	lorvage
CM2	1892,27 (385,33)	1730,52 (277,41)	1996,38 (440,46)	2134,43 (411,25)
Adultes	1584,81 (165,89)	1559,96 (138,07)	1656,18 (175,33)	1732,54 (214,59)
Moyenne	1738,54	1645,24	1826,28	1933,48

Les effets des facteurs groupe et lexicalité sur les temps de lecture correcte sont significatifs ($F(1, 59) = 19,23 ; p < .0001 ; F(1, 59) = 76,56 ; p < .0001$). Les enfants de niveau CM2 sont moins rapides que les adultes (1938,40 ms > 1633,37 ms) et les mots (1691,89 ms) sont lus plus rapidement que les pseudo-mots (1879,88 ms). La différence entre affixés (*gommage* ou *mordage* : 1782,41 ms) et pseudo-affixés (*message* ou *lorvage* : 1789,36 ms) n'est pas significative ($F(1, 59) < 1$). Les facteurs groupe et caractère plurimorphémique n'interagissent pas ensemble ($F(1, 59) = 1,08 ; ns$) tandis que les interactions entre lexicalité et affixation d'une part et entre groupe et lexicalité d'autre part, sont significatives ($F(1, 59) = 36,38 ; p < .0001$ et $F(1, 59) = 76,56 ; p < .0001$). Les trois facteurs interagissent également de façon significative ($F(1, 59) = 8,92 ; p < .01$, voir Figure 30). Seule cette interaction sera commentée.

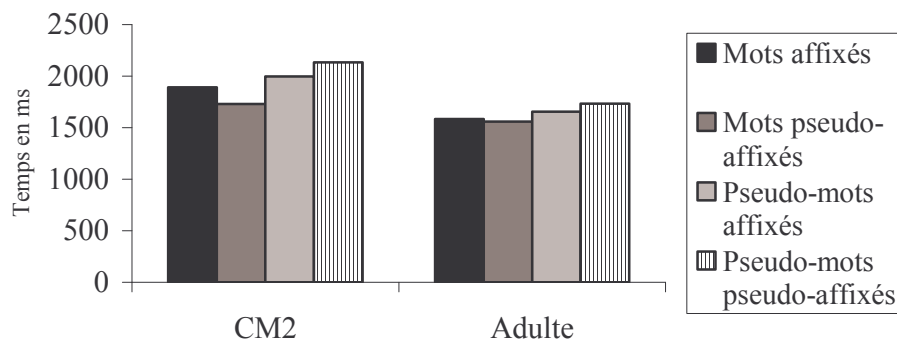


Figure 30. Temps mis par les adultes et les enfants de niveau CM2 pour lire les mots et pseudo-mots suffixés et pseudo-suffixés

Celle-ci est liée à la plus grande rapidité des adultes par rapport aux lecteurs de niveau CM2 pour les mots suffixés (1584,81 ms < 1892,27 ms ; $p = .001$), les mots pseudo-suffixés (1559,96 ms < 1730,52 ms ; $p = .0037$), les pseudo-mots suffixés (1656,18 ms < 1996,38 ms ; $p = .0002$) et les pseudo-mots pseudo-suffixés (1732,54 ms < 2134,43 ms ; $p < .0001$). Cette interaction indique que la différence entre mots affixés et pseudo-mots affixés n'est significative ni chez les enfants de niveau CM2 ($p = .3259$), ni chez les adultes ($p = .1108$). Il en est de même pour la différence entre les pseudo-mots affixés et les pseudo-mots pseudo-affixés ($p = .1367$ pour les lecteurs experts et $p = .2070$ pour les enfants de niveau CM2). En revanche, les pseudo-mots pseudo-affixés sont lus plus lentement que les mots pseudo-affixés (adultes : $p = .0005$; CM2 : $p < .0001$). Enfin, les enfants de niveau CM2 tendent à lire les mots pseudo-affixés plus rapidement que les mots affixés ($p = .0627$) alors que la différence entre ces deux catégories d'items n'est pas significative pour les adultes.

1.5- Discussion des résultats issus des comparaisons adultes-enfants de niveau CM2

L'objectif principal des comparaisons entre lecteurs experts et lecteurs de niveau CM2 est de voir si les traitements opérés entre ces deux niveaux de lecture se modifient sous l'effet de l'expertise en lecture. Les enfants de niveau CM2 peuvent être considérés encore en phase d'apprentissage du fait de leur âge et de leurs années de scolarisation. Toutefois, les résultats obtenus en lecture à haute voix en termes d'exactitude atteignent des scores plafonds, proches de ceux des lecteurs experts. Il est donc légitime de se demander si les enfants de niveau CM2 sont toujours en phase d'apprentissage ou s'ils peuvent être considérés comme des lecteurs experts. L'étude spécifique de l'évolution des traitements graphophonologiques, orthographiques et morphologiques en termes de rapidité va permettre de déterminer le degré de proximité entre le niveau de lecture des enfants et celui des adultes.

1.5.1- Evolution des traitements graphophonologiques entre le niveau CM2 et l'expertise en lecture

Les analyses de variance réalisées sur les temps de lecture des enfants de niveau CM2 et des lecteurs experts indiquent des traitements graphophonologiques très proches. En effet, globalement, les mots (*camarade*) sont lus plus vite que les pseudo-mots (*falore*) et les réguliers (*maladie, janitou*) sont lus également plus rapidement que les irréguliers (*parfum, lautil*). Il existe cependant un effet simple du groupe : si la structure graphophonologique facilite le traitement en lecture à la fois chez les enfants de niveau CM2 et les lecteurs experts, les enfants sont cependant plus lents que les lecteurs experts. Ainsi, les enfants de niveau CM2 se

comportent comme les experts quant aux traitements graphophonologiques effectués mais leur récupération des items en mémoire est moins rapide que celle des adultes.

1.5.2- Evolution des traitements orthographiques entre le niveau CM2 et l'expertise en lecture

Les analyses réalisées dans l'expérience 2 (voir chapitre 6, pp. 130-132) indiquaient que les enfants de niveau de lecture CM2 lisaient de la même façon les items à bigramme fréquent en position initiale (*caloude*) que ceux à bigramme rare (*narufé*) alors que chez les lecteurs experts la fréquence du bigramme avait un effet facilitateur sur la rapidité de traitement des items. Les analyses de variance effectuées sur les deux groupes montrent pourtant une similarité de traitement : globalement, les items à bigramme fréquent en initiale sont lus plus vite que ceux à bigramme rare. De plus, cet effet simple de la fréquence bigrammique n'interagit pas avec le niveau de lecture. Il semble donc possible d'envisager que l'effet de la fréquence du bigramme observé chez les adultes soit suffisamment important pour « masquer » l'absence de cet effet chez les enfants de niveau CM2. Bien que de façon globale les pseudo-mots soient lus moins rapidement que les mots, la différence de traitement entre les deux types d'items est plus marquée chez les enfants de niveau CM2. Par rapport aux lecteurs experts, la différence de traitement se situe essentiellement au niveau des pseudo-mots traités moins rapidement chez les enfants de niveau CM2. Ainsi, d'une manière générale, il est difficile de conclure à une similarité des traitements orthographiques dans le cas des bigrammes en position initiale entre les enfants de niveau CM2 et les lecteurs experts.

Pour les bigrammes fréquents et rares en position finale, des résultats similaires sont obtenus concernant l'écart de traitement entre adultes et enfants pour les pseudo-mots. En revanche, les analyses confirment que la fréquence des bigrammes n'a d'effet ni chez les experts ni chez les enfants de niveau CM2. Ainsi, dans le cas des bigrammes en position finale, les traitements orthographiques intervenant dans les procédures de reconnaissance de mots écrits évoluent très peu après l'école primaire.

1.5.3- Evolution des traitements morphologiques entre le niveau CM2 et les lecteurs experts

Les analyses de variance réalisées sur les temps de lecture montrent que même si les adultes lisent plus vite que les enfants de niveau CM2, les traitements morphologiques semblent similaires à la fois pour les préfixés et les suffixés : globalement, les mots sont lus plus rapidement que les pseudo-mots, les préfixés tendent à être lus plus vite que les pseudo-préfixés et aucune différence n'apparaît entre suffixés et pseudo-suffixés. Ces résultats, notamment pour la catégorie préfixe, sont assez surprenants. En effet, dans l'expérience 3 du chapitre 6, les

enfants de niveau CM2 lisaient effectivement plus rapidement les préfixés que les pseudo-préfixés mais dans l'expérience 4 les adultes traitaient de façon similaire les deux catégories d'items. Comme cela a été supposé précédemment pour les traitements orthographiques, l'effet prégnant de la structure morphologique chez les enfants pourrait masquer l'absence d'effet chez les adultes et expliquer alors que seule une tendance est observée pour les analyses portant sur les deux groupes.

Pour la catégorie suffixe, l'interaction triple entre le niveau de lecture, la lexicalité et le caractère plurimorphémique confirme que les enfants de niveau CM2 tendent à lire les mots suffixés (*chasseur*) plus lentement que les mots pseudo-suffixés (*douleur*). Cet effet contraire aux hypothèses sur la facilitation en lecture pour les items affixés n'est pas retrouvé chez les lecteurs experts : les adultes lisent aussi bien les mots pseudo-suffixés que les suffixés. Il y aurait donc une évolution dans les traitements morphologiques réalisés sur les mots entre le primaire et l'âge adulte mais pas pour les pseudo-mots : pour cette catégorie d'items, suffixés et pseudo-suffixés sont traités de façon similaire en termes de vitesse chez les enfants de niveau CM2 et les lecteurs experts.

En conclusion, la similarité des traitements opérés par les enfants de niveau CM2 et les lecteurs experts dépend de la structure étudiée. Pour les structures graphophonologique et orthographique – items à bigramme en position finale –, les traitements observés chez les lecteurs experts sont proches de ceux observés chez les enfants de niveau CM2. Les traitements morphologiques n'évoluent pas non plus entre adultes et enfants dans le cas spécifique des pseudo-mots de type suffixe. En revanche, les mots de type suffixe, les items de type préfixe et ceux à bigrammes en position initiale ne sont pas traités de façon similaire par les deux groupes de niveau.

CHAPITRE 8 – RELATIONS ENTRE LES CONNAISSANCES GRAPHOPHONOLOGIQUES, ORTHOGRAPHIQUES, MORPHOLOGIQUES, LES NIVEAUX DE LECTURE ET DE VOCABULAIRE

Les analyses suivantes visent à mettre en évidence les relations existantes entre les trois grands types de connaissances chez les enfants (GSM et niveaux de lecture CP à CM2) et les adultes *via* des corrélations et des régressions multiples. Dans le cas des régressions multiples, il s'agit de mettre en évidence la ou les compétences qui contribuent le mieux à expliquer le niveau de lecture.

1- Corrélations entre les notes globales pour les différents groupes

Des analyses de corrélations ont été réalisées sur les notes globales obtenues aux épreuves implicite et explicite de graphophonologie, orthographe, morphologie ainsi qu'aux épreuves *LUM* et *EVIP*. Ces analyses concernent les enfants scolarisés en GSM, les niveaux de lecture du CP au CM2 ainsi que le groupe d'adulte. En GSM, les enfants n'ayant pas réalisé les épreuves explicites, les corrélations ne portent que sur les scores implicites obtenus aux tâches de graphophonologie, d'orthographe et de morphologie ainsi que sur les performances à l'*EVIP*.

Les différents tableaux présentés ci-après récapitulent les principales corrélations positives significatives à .05. Les codes utilisés dans ces tableaux correspondent aux différentes épreuves : GPI pour Graphophonologie Implicite, GPE pour Graphophonologie Explicite, OI pour Orthographe Implicite, OE pour Orthographe Explicite, MI pour Morphologie Implicite, ME pour Morphologie Explicite, *LUM* pour la *Lecture en Une Minute* et enfin *EVIP* pour l'épreuve de vocabulaire.

1.1- Corrélations en GSM

A ce niveau scolaire, aucune corrélation significative entre les différentes connaissances n'a été obtenue.

1.2- Corrélations au niveau CP

Pour ce niveau de lecture, les corrélations positives et significatives ($p < .05$) sont reportées dans le Tableau 25.

Tableau 25. Corrélations entre les différentes tâches d'évaluation des connaissances, du vocabulaire et *LUM*

	GPI	GPE	OI	OE	MI	ME	LUM	EVIP
GPI	1							
GPE	ns	1						
OI	ns	ns	1					
OE	ns	ns	ns	1				
MI	ns	.44	ns	ns	1			
ME	.49	ns	ns	ns	ns	1		
LUM	ns	.48	ns	ns	.49	ns	1	
EVIP	ns	ns	ns	ns	ns	ns	.55	1

Les analyses au niveau CP indiquent l'existence de corrélations significatives entre le score obtenu en lecture (et mesuré par *LUM*) et le niveau en vocabulaire d'une part ($r = .55$) et entre le score *LUM* et les performances à la tâche de graphophonologie explicite d'autre part ($r = .48$). Elles montrent également que les enfants de ce groupe qui sont sensibles à la structure morphologique obtiennent les meilleurs scores en graphophonologie explicite ($r = .44$) et sont aussi les lecteurs les plus efficaces ($r = .49$). Enfin, les enfants de niveau CP qui détectent le mieux les intrus morphologiques dans la tâche explicite sont sensibles à la consistance des items de l'épreuve graphophonologique implicite.

1.3- Corrélations au niveau CE1

Pour ce niveau de lecture, les analyses mettent en évidence, comme pour le niveau CP, un lien entre les performances en lecture et le niveau de vocabulaire ($r = .40$). De bonnes performances à *LUM* sont aussi associées à des scores élevés à l'épreuve explicite de graphophonologie ($r = .55$). Ce résultat conforte l'idée d'un développement des connaissances phonologiques explicites parallèle à l'apprentissage de la lecture. De même, plus le lexique est développé, plus les enfants sont performants à la tâche impliquant les connaissances graphophonologiques explicites ($r = .41$). Les deux épreuves orthographiques – implicite et explicite – sont corrélées positivement ($r = .47$).

1.4- Corrélations au niveau CE2

Ces analyses corrélationnelles figurent dans le Tableau 26 ci-après.

Tableau 26. Corrélations entre les différentes tâches d'évaluation des connaissances, du vocabulaire et *LUM*

	GPI	GPE	OI	OE	MI	ME	LUM	EVIP
GPI	1							
GPE	ns	1						
OI	.40	ns	1					
OE	ns	ns	-.37	1				
MI	ns	ns	ns	ns	1			
ME	ns	ns	ns	ns	ns	1		
LUM	ns	.35	ns	ns	ns	.40	1	
EVIP	ns	ns	ns	ns	ns	.43	ns	1

Les enfants de niveau CE2 qui détectent facilement les intrus morphologiques sont aussi ceux qui obtiennent les meilleurs scores en lecture et en vocabulaire. Un bon niveau en lecture est également associé à de bonnes connaissances explicites en graphophonologie ($r = .35$) tandis que les scores graphophonologiques implicites sont liés aux performances à l'épreuve orthographique implicite. En revanche, les performances aux deux tâches orthographiques sont corrélées négativement ($r = -.37$) : des scores élevés à l'épreuve implicite ne sont donc pas associés à la réussite au test orthographique explicite, ce qui suggère qu'à ce niveau ce sont d'autres types d'indices qui déterminent les préférences.

1.5- Corrélations au niveau CM1

A ce niveau de lecture, seules les performances à l'*EVIP* sont corrélées significativement à celles obtenues à la tâche orthographique implicite ($r = .53$). Les enfants possédant un bon niveau en vocabulaire sont sensibles à la structure orthographique des items de l'épreuve implicite.

1.6- Corrélations au niveau CM2

Comme pour les niveaux CP et CE1, le niveau en vocabulaire est corrélé au niveau *LUM* ($r = .35$). Les lecteurs les plus efficaces de ce niveau ont un vocabulaire important. Ce bon niveau en vocabulaire est associé à des connaissances explicites importantes en graphophonologie ($r = .46$) et en morphologie ($r = .43$). Enfin, les deux tests orthographiques sont corrélés ($r = .51$). Ainsi, les enfants de niveau CM2 qui choisissent de manière préférentielle les items à bigrammes fréquents (épreuve implicite) considèrent également qu'il y a davantage de mots commençant par [ca-] que par [na-] et plus de mots finissant par [-te] que par [-ve] (tâche explicite).

1.7- Corrélations chez le groupe adulte

L'efficacité en lecture mesurée par le temps moyen pour lire un mot dans l'épreuve *LUM* est corrélée significativement aux performances obtenues à l'*EVIP*. Les adultes qui lisent le plus rapidement sont ceux dont le vocabulaire est le plus étendu. Des scores importants à l'*EVIP* sont également associés à une sensibilité élevée à la structure morphologique. Par ailleurs, l'épreuve orthographique implicite est corrélée aux épreuves orthographique explicite ($r = .37$) et graphophonologique implicite ($r = .38$).

2- Régressions multiples

Afin de déterminer dans quelle mesure les différents types de connaissances – graphophonologiques, orthographiques et morphologiques – peuvent prédire la capacité de décodage (mesuré par le score obtenu à l'épreuve *LUM*), des analyses de régression pas à pas ont été conduites pour les enfants des groupes de niveau CP, CE1 et CE2. Aucune analyse de régression n'a été menée avec les enfants de GSM dans la mesure où les corrélations entre les différentes variables mesurées n'étaient pas significatives (*cf.* Analyse corrélationnelle décrite précédemment). Au niveau CM1, aucune des variables n'est corrélée avec *LUM*. Au niveau CM2 et pour le groupe des adultes, seul le niveau de vocabulaire est corrélé positivement avec *LUM* (*cf.* Analyse des corrélations précédente). Pour ces trois derniers groupes, aucune analyse de régression multiple n'a donc pu être menée.

Au niveau CP, seules les performances en morphologie implicite et en graphophonologie explicite ainsi que le niveau de vocabulaire ont été entrés dans le modèle (corrélations les plus élevées avec *LUM*). Comme le montre le Tableau 27, seul le modèle introduisant le niveau de vocabulaire dans un premier temps et les connaissances morphologiques implicites dans un second temps est significatif. Ainsi, dès le niveau CP, les connaissances morphologiques implicites contribuent de façon indépendante à la lecture (13 % de la variance expliquée).

Tableau 27. Régressions multiples au niveau CP

	R	R ²	R ² modifié	p
Pas 1. EVIP	.55	.31	.28	<.002
Pas 2. Morphologie implicite	.66	.44	.13	<.001

Au niveau CE1, seules les variables niveau de vocabulaire et performances en graphophonologie explicite ont été entrées dans le modèle. Le Tableau 28 indique que seules les connaissances en graphophonologie explicite contribuent de façon indépendante à la lecture avec 28 % de variance expliquée.

Tableau 28. Régressions multiples au niveau CE1

	R	R ²	R ² modifié	p
Pas 1. Graphophonologie explicite	.55	.31	.28	<.002
Pas 2. EVIP	.58	.34	.03	ns

En revanche, au niveau CE2, aucune des variables testées (morphologie explicite et graphophonologie explicite) ne rend compte significativement de l'habileté en lecture.

3- Discussion

L'analyse des corrélations niveau par niveau montre que chez les prélecteurs aucune corrélation significative n'est obtenue. A un niveau scolaire aussi précoce (GSM), il est probable que la difficulté de réalisation des différentes tâches en modalité écrite explique cette absence de liens entre les différentes connaissances. Dans le cadre de la présente étude, le choix d'un niveau scolaire aussi précoce répondait à la quasi absence d'enfants prélecteurs dans la plupart des études sur un même sujet. Dans les recherches intégrant des enfants de GSM, les tâches utilisées étaient le plus souvent orales (Colé *et al.*, 2003) : l'objectif était alors de répliquer les résultats obtenus dans le cas de la modalité écrite partant de l'hypothèse que, dès la GSM, les enfants peuvent être fréquemment exposés à l'écrit en contexte scolaire. L'absence de corrélations s'explique sans doute par les choix aléatoires des enfants prélecteurs face à des épreuves trop complexes à réaliser.

Chez les apprentis lecteurs de niveau CP, la sensibilité morphologique ainsi que le fait de disposer de connaissances explicites sur la consistance graphèmes-phonèmes et d'un niveau de vocabulaire conséquent permettent d'accroître les performances en lecture. Ces résultats confortent ainsi les relations existantes entre graphophonologie et lecture déjà mises en évidence dans les modèles développementaux d'apprentissage de la lecture (Frith, 1985 ; Seidenberg & McClelland, 1989 ; Seymour, 1997) et par de nombreux chercheurs (Duncan *et al.*, 2000 ; Hulme *et al.*, 1998 ; Hulme *et al.*, 2002 ; Muter *et al.*, 1998 ; Sprenger-Charolles *et al.*, 1998). Surtout, ces corrélations montrent que les connaissances graphophonologiques explicites ne sont pas les seules à intervenir dans les procédures de reconnaissance de mots connus et/ou peu familiers. Les connaissances morphologiques implicites participent également de façon précoce à ces

procédures. De plus, les relations entre niveau phonologique et niveau morphologique établies par Fowler et Liberman (1995) sont confortées pour le niveau CP par les liens entre sensibilité graphophonologique et connaissances morphologiques explicites ainsi que par les corrélations entre les connaissances graphophonologiques explicites et la sensibilité morphologique. Un autre résultat intéressant provient des analyses de régressions qui montrent qu'à ce niveau de lecture, le vocabulaire contribue de façon importante à l'efficacité en lecture. Ces régressions confirment ainsi les relations, déjà établies par plusieurs auteurs, entre acquisition du vocabulaire et apprentissage de la lecture (Anderson & Freebody, 1985 ; Anglin, 1993 ; Nagy & Anderson, 1984 ; Colé *et al.*, 2003). Au tout début de l'apprentissage de la lecture, les enfants s'appuient sur les mots connus pour pouvoir lire des mots nouveaux. Ce constat peut être mis en parallèle du rôle des analogies en lecture souligné par Gombert *et al.* (1997) : de façon précoce, l'enfant serait capable de mettre en œuvre des procédures de lecture de mots nouveaux sur la base d'analogies avec des mots dont il connaît la forme orthographique. Les analyses de régression indiquent également que la sensibilité morphologique a une contribution indépendante de 13 % au niveau de lecture. Cette contribution est loin d'être négligeable et est plus importante que celle obtenue par Carlisle et Nomanbhoy (1993) qui avaient montré que la conscience phonologique prédisait 33 % de la variance en lecture chez les enfants de niveau CP tandis que la morphologie rendait compte de seulement 4 % de variance supplémentaire. De façon relativement précoce, les enfants semblent s'appuyer sur les régularités morphologiques des mots pour lire : de telles relations avaient déjà été mises en évidence par plusieurs auteurs mais à des niveaux plus élevés comme le CE1 (Casalis & Louis-Alexandre, 2000 ; Laxon *et al.*, 1992) ou encore le CE2 (Colé *et al.*, 2003). La présente étude met en évidence une précocité plus importante au niveau de l'émergence du rôle de la morphologie en lecture. Au niveau CP, les enfants s'appuient sur leurs connaissances implicites et explicites en interaction avec leur sensibilité aux correspondances graphèmes-phonèmes apprises explicitement.

Une évolution apparaît au niveau de lecture supérieur : au niveau CE1, l'efficacité en lecture s'appuie essentiellement sur le vocabulaire et sur les connaissances explicites de la correspondance graphèmes-phonèmes. Ce sont les connaissances graphophonologiques explicites qui contribuent le plus au niveau de lecture (28 % contre 3 % pour le vocabulaire). A ce niveau de lecture, il est possible d'envisager que l'apprentissage explicite des correspondances graphèmes-phonèmes constitue une part importante de l'enseignement scolaire. Les traitements attentionnels de ces lecteurs peu expérimentés pourraient se focaliser sur les aspects graphophonologiques des mots écrits. Les résultats pour le niveau CE1 se rapprocheraient alors de ceux obtenus par Manis *et al.* (1993) pour l'anglais et de Sprenger-Charolles *et al.* (2003) pour le français montrant que

les compétences phonologiques capturent la majeure partie de la variance unique dans les compétences en identification des mots écrits. Une autre donnée intéressante pour le niveau de lecture CE1 est l'apparition dans l'étude des corrélations de liens entre sensibilité à la fréquence des bigrammes et connaissances explicites orthographiques. A ce niveau de lecture, la sensibilité précoce à la fréquence de paires de lettres favorise l'émergence d'une base de connaissances explicites (Gombert, 2003).

Au niveau de lecture CE2, les enfants qui réussissent le mieux le test explicite de graphophonologie sont également les plus efficaces en lecture. Ce résultat est sans doute à mettre directement en parallèle des performances obtenues par les enfants de ce niveau à l'épreuve graphophonologique explicite. Pour ces enfants, « régulariser » un item irrégulier était plus simple à effectuer qu'« irrégulariser » un item régulier. Pour expliquer ces résultats, il était supposé que l'amélioration des compétences en lecture ainsi que l'accroissement concomitant des connaissances orthographiques sur les mots irréguliers favorisait la prise de conscience explicite de la consistance et de l'inconsistance des correspondances graphèmes-phonèmes. Les liens observés entre sensibilité graphophonologique et sensibilité orthographique soutiendraient d'ailleurs cette hypothèse : les interactions entre informations graphophonologiques et informations orthographiques favoriseraient la reconnaissance des mots écrits, conformément aux prédictions des modèles interactifs de Colé et *al.* (1999) et de Gombert *et al.* (1997). Les connaissances graphophonologiques explicites ne sont pas les seules utilisées en lecture. Les connaissances morphologiques explicites sont également corrélées à l'efficacité en lecture (et aussi au niveau de vocabulaire, comme observé par Fowler & Liberman, 1995), conformément aux résultats déjà obtenus par Colé *et al.* (2003), Carlisle et Nomanbhoy (1993), Casalis et Louis-Alexandre (2000) ou encore Laxon *et al.* (1992) alors qu'au niveau CE1 ce type de lien était inexistant et qu'au niveau CP le versant implicite de la morphologie contribuait au niveau de lecture. Ces changements relevés dans le temps pourraient suggérer l'existence de sauts qualitatifs au cours de l'apprentissage. Cette notion de saut qualitatif a déjà été utilisée par Sprenger-Charolles *et al.* (2003) pour désigner l'évolution dans l'utilisation des connaissances contribuant à la lecture. Ainsi, au tout début de l'apprentissage de la lecture (niveau CP), la lecture serait prédite essentiellement par le niveau de vocabulaire *via* le recours aux analogies et également par la sensibilité morphologique émergente. Au niveau CE1, la forte influence de l'apprentissage explicite des correspondances graphèmes-phonèmes favoriserait un recours important aux connaissances explicites graphophonologiques en lecture. Ensuite, au niveau de lecture CE2, des connexions pourraient progressivement se créer entre unités orthographiques et graphophonologiques (*cf.* les liens entre sensibilité graphophonologique et sensibilité

orthographique) mais aussi entre unités morphologiques et graphophonologiques. Ces connexions participeraient aux procédures de reconnaissance des mots écrits.

Le niveau CM1 se caractérise par des résultats assez étonnants : à ce niveau de lecture, la seule corrélation significative est celle existant entre le niveau de vocabulaire et la sensibilité orthographique. Au niveau CM2 mais également chez les lecteurs experts, seul le niveau de vocabulaire est lié à l'efficacité en lecture. Il semble que les tâches implicites et explicites écrites soient plus liées aux variations de performances de lecture des débutants (i.e., niveaux CP, CE1 et CE2) qu'à celles des performances de la lecture plus automatique des lecteurs expérimentés à partir du niveau CM1⁷. Chez les lecteurs débutants, la lecture reste influencée par l'attention portée aux constituants linguistiques des mots écrits (aspects graphophonologiques, orthographiques et morphologiques). A mesure que l'expertise augmente, l'efficacité en lecture dépend essentiellement du vocabulaire disponible qui lui-même s'appuie pour se développer sur les connaissances explicites morphologiques (adultes) mais aussi graphophonologiques (niveau de lecture CM2).

⁷ Il faut toutefois prendre en compte que la division par niveau diminue fortement la dispersion des niveaux de lecture, les corrélations avec la lecture sont donc plus difficiles à obtenir.

CHAPITRE 9 – DISCUSSION GENERALE

Les principaux modèles développementaux de l'apprentissage de la lecture, comme celui de Frith, (1985) tendent à réduire cet apprentissage au décodage et à l'assemblage phonologiques. Des modèles plus interactifs (*cf.* Chapitre 1) ont pris en compte d'autres types de connaissances comme celles orthographiques ou sémantiques (Seidenberg & McClelland, 1989). Par exemple, dans le modèle à double fondation de Seymour (1997), l'habileté à reconnaître des mots écrits dépend de l'interaction entre les structures phonologiques, orthographiques et morphologiques. L'intérêt de ce type de modèle est de montrer que des connaissances autres que strictement phonologiques peuvent être nécessaires au traitement des mots et non mots écrits. Cependant, le modèle à fondation duale ne fait intervenir que tardivement la morphologie, une fois que le développement de la structure orthographique est achevé. Dans le cas des connaissances orthographiques, plusieurs modèles d'acquisition se sont intéressés à leur développement (voir chapitre 3 avec les modèles en stades de Gentry, 1982 et de Henderson, 1985). Les connaissances orthographiques sont alors prises en compte de façon relativement précoce mais dans le contexte de l'apprentissage de l'écriture et non exclusivement de la lecture. Lorsque des liens sont établis entre orthographe et lecture, comme dans les modèles de Frith (1985) ou de Ehri (1987), ce n'est qu'à des niveaux tardifs de l'apprentissage.

Cependant, d'autres conceptions de l'apprentissage de la lecture se sont développées ces dernières années mettant en avant l'intervention plus précoce des unités orthographiques en lecture ainsi que l'interaction entre orthographe et phonologie (*cf.* Chapitre 1 avec le modèle de lecture par analogies de Gombert *et al.*, 1997 et le modèle d'activation des niveaux orthographiques et phonologiques de Colé *et al.*, 1999). Concernant les aspects morphologiques, des travaux récents réalisés notamment en anglais et en français ont montré que l'unité morphologique était présente dans la langue et pouvait être utilisée très tôt et de façon pertinente par les enfants en lecture. Cette intervention précoce de la morphologie et de l'orthographe en lecture est suggérée en particulier par Gombert (2003). Selon cet auteur, dès les premiers contacts avec l'écrit à l'école maternelle débute un travail de découverte et d'hypothèses sur les lettres et sur l'écrit qui va nourrir et développer les premières connaissances morphologiques et orthographiques conscientes de l'enfant sur le langage écrit. Cette exposition précoce à l'écrit renvoie au contexte scolaire mais aussi aux pratiques familiales de lecture et à leurs effets sur les compétences langagières précoces (Ecalte & Magnan, 2002a). En effet, les parents peuvent avoir différentes pratiques centrées autour de l'écrit avec leur enfant et ces pratiques associées aux

interactions parents-enfant sont susceptibles d’avoir un effet sur le développement des connaissances relatives à l’écrit. Il est acquis depuis quelques années que la lecture partagée d’histoires entre adulte et enfant constitue une routine interactive idéale à l’acquisition du langage oral (*cf.* Hargrave & Sénéchal, 2000) mais cette pratique peut également favoriser l’intégration d’un certain nombre de conventions régissant l’écrit (Lanoë, 2000). Ainsi, l’enfant prend conscience des formes syntaxiques et grammaticales propres à l’écrit, souvent plus complexes, plus riches que celles utilisées à l’oral. Le vocabulaire employé dans les livres est généralement plus étendu (Sénéchal, 2000a). En se distanciant du langage oral, outil de communication quotidien fonctionnel, l’enfant au cours de lectures conjointes de livres peut être amené à engager un travail de réflexion sur l’écrit (Ecalte & Magnan, 2002a). La répétition des lectures de mêmes livres conduirait peu à peu les enfants à se focaliser sur les formes écrites.

L’exposition à l’écrit pourrait ainsi contribuer pour une part importante au développement des processus orthographiques impliqués en identification de mots écrits. L’exposition à l’écrit serait donc un facteur expérientiel contribuant au développement, à la précision et au stockage des représentations orthographiques. Au cours de ses contacts réguliers avec l’écrit, l’enfant est capable d’en extraire certaines propriétés orthographiques, phonologiques et morphologiques nécessaires à l’apprentissage ultérieur de la lecture. Cette extraction permettrait alors le développement de connaissances implicites sur certaines régularités de la langue qui ne répondent pas à des règles apprises explicitement (Perruchet & Pacton, 2004). De la sorte, avant même le début de l’école primaire, l’enfant dispose d’habiletés implicites qui constitueront le socle sur lequel va s’appuyer l’enseignement de la lecture et de l’écriture. L’enseignement, notamment celui des codes et des règles de conversion, joue également un rôle primordial dans l’apprentissage. D’une part, il va permettre l’installation d’un système de contrôle, c’est-à-dire de connaissances conscientes ou explicites que l’enfant pourra utiliser de façon intentionnelle lorsqu’il en aura besoin. D’autre part, l’enseignement a également pour effet d’accélérer les apprentissages implicites. En conduisant l’enfant à multiplier, varier, et répéter la manipulation de l’écrit, l’apprentissage explicite décuple les apprentissages implicites. Ainsi, ces deux types d’apprentissage ne peuvent s’opérer l’un sans l’autre : ce sont les apprentissages implicites qui sont responsables des automatismes de lecture, mais les connaissances conscientes sont indispensables au déploiement de ces apprentissages et sont nécessaires à chaque fois que le lecteur doit prendre des décisions dans l’activité de lecture-écriture (Gombert, 2003). Il n’y a pas substitution mais interaction entre les deux apprentissages.

Ainsi, avant même l’apprentissage de la lecture, l’enfant serait en mesure d’extraire de son environnement linguistique des régularités de type graphophonologique, orthographique et

morphologique. Ces connaissances implicites sont susceptibles de se développer avec l'apprentissage de la lecture et de favoriser l'émergence de connaissances explicites. Dans ce cadre, l'objectif de cette thèse était de mettre en évidence l'apparition précoce (dès la GSM) de connaissances sur les régularités graphophonologiques, orthographiques et morphologiques, et d'étudier le développement de ces connaissances en lien avec la lecture. Il s'agissait également de mettre en évidence les interactions possibles entre les trois types de connaissances (graphophonologiques, orthographiques et morphologiques) et de vérifier l'utilisation de la structure graphophonologique, orthographique et morphologique lors de la lecture dès le début de cet apprentissage mais aussi à des niveaux plus avancés (lecture experte).

1- Connaissances graphophonologiques, orthographiques, morphologiques et apprentissage de la lecture

1.1- Les connaissances graphophonologiques, orthographiques et morphologiques

Concernant les connaissances graphophonologiques, de nombreux travaux ont établi depuis longtemps des relations fortes entre l'apprentissage des correspondances graphèmes-phonèmes et les capacités de reconnaissance de mots (pour une revue non exhaustive de ces travaux, voir chapitre 2). Bien souvent, ce sont les connaissances graphophonologiques explicites qui sont évoquées pour rendre compte des capacités d'identification de mots écrits. Sans remettre en cause l'importance de l'apprentissage explicite des correspondances graphèmes-phonèmes dans les procédures de reconnaissance des mots, un des objectifs de la présente étude était de montrer que les enfants possédaient de manière précoce des connaissances sur la régularité ou consistance de ces correspondances. Cette hypothèse s'appuyait sur le fait que dans la langue française, de nombreuses relations graphophonémiques ne sont pas transparentes : pour un même phonème, plusieurs réalisations graphémiques peuvent coexister (Catach *et al.*, 1986). L'idée sous jacente était que les enfants pouvaient être sensibles relativement tôt aux régularités graphophonologiques de leur langue et donc à la fréquence de certaines associations graphèmes-phonèmes, plus consistantes que d'autres (voir les résultats de Lamarche, 1986 ; chapitre 2). Les résultats obtenus auprès d'enfants de la GSM au niveau de lecture CM2 ne confortent qu'en partie cette hypothèse. En effet, il faut attendre une certaine efficacité en lecture (niveaux CM1 et CM2) pour qu'une sensibilité à la structure graphophonologique (i.e., consistance graphèmes-phonèmes) se manifeste. A ce niveau de lecture, les enfants deviennent capables de distinguer, en termes de facilité de lecture, deux items dont l'un ne contient que des graphèmes consistants (*piralue*) et l'autre des graphèmes consistants et un graphème inconsistant (*piralum*). La sensibilité à la consistance graphophonémique mise en évidence ici émergerait plus tardivement

que celle constatée par Fayol *et al.* (1996). Ces auteurs s'étaient intéressés aux transcriptions des phonèmes /o/ et /s/ en fonction des différents graphèmes possibles et de la position dans un item donné. Ils avaient montré que dès le CP, l'enfant était capable d'établir des associations entre certaines positions dans les items (début/milieu/fin) et certains graphèmes. Les auteurs en concluaient que, de façon précoce, les enfants étaient sensibles aux régularités graphophonologiques apprises de manière implicite. Pacton *et al.* (1999) avaient également obtenu des données similaires dans les productions orthographiques de pseudo-mots contenant le phonème /o/ et réalisées par des enfants de CP. La mise en évidence plus tardive d'une telle sensibilité dans la présente étude peut trouver une explication dans la nature même de la tâche proposée. Dans les expériences de Fayol *et al.* (1996), de Pacton *et al.* (1999) mais également dans cette thèse, la tâche implicite exploite des pseudo-mots, ce qui est fréquent pour mettre en évidence des connaissances acquises implicitement. En revanche, les études de Fayol *et al.* (1996) ainsi que de Pacton *et al.* (1999) se sont centrées sur un seul ou deux phonèmes (le /o/ et/ou le /s/) et les transcriptions graphémiques envisageables. Ces deux phonèmes et leurs réalisations graphémiques sont relativement simples ([o], [eau], [au] pour /o/ ; [s], [ss], [c], [ç] pour /s/) et pourraient être intégrés par l'enfant bien plus tôt que les graphèmes utilisés dans le test implicite de l'expérience 1. Les graphèmes de ce test sont en effet plus nombreux et plus complexes ([th-], [ch-], [au-], [ai-], [-an], [-ac], etc.). Dès lors, la sensibilité des enfants à la plus ou moins grande consistance de ces graphèmes pourrait apparaître plus tardivement et être en lien direct avec la durée d'exposition à l'écrit et donc avec l'efficacité en lecture. L'utilisation de graphèmes plus restreints et plus accessibles aurait pu permettre une mise en évidence plus précoce d'une sensibilité à la consistance graphophonémique. La consigne proposée a pu également déconcerter les enfants les moins experts en lecture. En effet, il était proposé aux enfants de bien regarder chaque pseudo-mot d'une paire et de montrer (GSM) ou d'entourer (niveau CP à CM2) celui qui semblait le plus facile à lire pour un lecteur débutant. Au moment de la passation (fin janvier), cette expression « le plus facile à lire » a pu être déstabilisante pour des enfants prélecteurs (en GSM) mais aussi pour les apprentis lecteurs de niveau CP ou CE1 qui ne sont pas encore complètement assurés dans leur statut de lecteurs. Les fortes différences interindividuelles obtenues pour les enfants de niveau CP iraient dans le sens de cette hypothèse : pour ce niveau de lecture, certains enfants parviennent à répondre à la consigne et à considérer l'item consistant comme le plus facile à lire tandis que d'autres font des choix hasardeux. La sensibilité à la consistance graphophonémique est en revanche bien présente chez les enfants des niveaux CM1 et CM2 qui considèrent les items consistants comme plus faciles à lire que ceux inconsistants. Un résultat plus surprenant provient des données recueillies auprès des lecteurs

experts (adultes) dont les réponses à l'épreuve implicite ne diffèrent pas du hasard. Les résultats relativement moyens et loin d'atteindre un seuil plafond (seulement 51,25 % de choix préférentiels pour les items consistants) laisseraient penser que les préférences ont pu être déterminées (implicitement) par d'autres facteurs que ceux testés initialement. Ces performances obtenues chez les adultes peuvent aussi être attribuées à la nature des items, à savoir des pseudo-mots qui ont pu être déroutants pour des lecteurs experts habitués à ne lire généralement que des mots connus dans leur langue.

Les connaissances explicites qui sont l'objet de l'enseignement formel de la lecture ne peuvent s'intégrer qu'à partir des précurseurs, c'est-à-dire des connaissances implicites, que l'enfant a spontanément construits à partir de ses processus de traitements épilinguistiques (Gombert, 2003). Cette interaction développementale entre implicite et explicite est confortée par les résultats obtenus au test graphophonologique explicite. Les enfants des niveaux CM1 et CM2 qui sont sensibles à la consistance graphophonémique sont capables de fournir les deux types de prononciation (régulière et irrégulière) pour les mots consistants et inconsistants de l'épreuve explicite. Au niveau CE2, un résultat intéressant est à noter : à ce niveau de lecture, une amélioration significative des performances apparaît, notamment pour les items irréguliers, par rapport au niveau de lecture CE1. Il est possible d'envisager qu'à ce niveau de lecture correspondent des acquisitions nouvelles et une réorganisation des connaissances (Marouby-Terriou & Denhière, 2002). Les enfants de niveau CE2 pourraient être ainsi à une période charnière d'apprentissage au cours de laquelle ils se détachent du processus purement phonologique pour accéder à un déchiffrage alphabétique devant des mots inconnus, en l'occurrence des pseudo-mots. Le décalage est relativement important entre les enfants de niveau CE1 et ceux de niveau CE2, décalage qui fait apparaître une réelle différence dans le processus de traitement de l'information. Comme il a été suggéré dans le chapitre 6, à ce stade d'acquisition (niveau de lecture CE2), les connaissances d'ordre orthographique commencent à s'élaborer et constituent sans doute une aide, parallèlement aux connaissances graphophonologiques, dans le traitement des items réguliers et surtout irréguliers. Il est important de noter que pour l'ensemble des enfants, des différences interindividuelles importantes sont constatées indiquant que les processus de traitement, dans leurs réalisations, diffèrent concrètement d'un enfant à l'autre pour un même niveau de lecture. Ces différences interindividuelles sont fréquemment soulignées dans le cas des apprentissages implicites (Bastien-Toniazzo, 1992 ; Ecalle & Magnan, 2002a) mais elles semblent également non négligeables dans le cas des apprentissages explicites. Enfin, le développement des connaissances graphophonologiques explicites se poursuit au-delà de l'apprentissage de la

lecture mais n'est pas terminé avec l'expertise en lecture, les performances augmentant entre le niveau de lecture CM2 et les adultes sans pour autant atteindre des scores plafonds.

Un des objectifs de la deuxième expérience sur l'orthographe était de montrer, contrairement aux modèles développementaux d'apprentissage de la lecture (Frith, 1985) et aux modèles d'acquisition en stades de l'orthographe (Frith, 1985 ; Gentry, 1982 ; Henderson, 1985), une émergence précoce des connaissances orthographiques. Cet objectif reposait notamment sur les expériences réalisées par Cassar et Treiman (1997) en anglais et celles de Pacton, Fayol et Perruchet (1999) et de Pacton, Perruchet, Fayol et Cleeremans (2001) en français sur la sensibilité précoce des enfants au processus de gémination. Ces auteurs avaient montré que, dès le CP, les enfants manifestaient des préférences à l'égard de pseudo-mots contenant des lettres fréquemment doublées dans la langue. Ce choix préférentiel d'items congruents avec les régularités propres à chaque langue démontrait de façon claire le développement d'une sensibilité précoce aux fréquences de suite de lettres composant la langue écrite. De plus, cette sensibilité augmentait avec le degré d'apprentissage de la lecture. L'expérience réalisée dans la présente étude exploitait non pas les lettres fréquemment (ou non) doublées dans la langue française mais la fréquence de paires de lettres (i.e., bigrammes) en fonction de leur position (initiale ou finale) dans un item donné. Les résultats indiquent qu'à partir du niveau CP, les enfants préfèrent les pseudo-mots contenant un bigramme fréquent, notamment lorsque celui-ci est en position initiale. Cette sensibilité augmente parallèlement à l'efficacité en lecture (excepté pour les enfants de niveau CE1 qui ne se distinguent ni des enfants de niveau CE2 ni de ceux de niveau CM2). Les données de l'expérience 2 concernant les connaissances orthographiques implicites confirment donc les résultats des recherches sur la gémination. Les enfants sont sensibles précocement à la fréquence de séquences de lettres et intériorisent de manière implicite certaines régularités orthographiques de la langue. Il est envisageable de supposer, en s'appuyant sur les résultats de Martinet *et al.* (2004), que les enfants s'appuient sur les analogies pour traiter et déterminer les items ressemblant le plus à de vrais mots en fonction de la plus ou moins grande fréquence des bigrammes. En effet, Martinet *et al.* (2004), par le biais d'expériences visant à mettre en évidence l'utilisation précoce de connaissances orthographiques chez l'apprenti lecteur et orthographeur, ont montré que dès le CP, des effets précoces de l'analogie apparaissaient en production orthographique de pseudo-mots. Dans l'épreuve orthographique implicite, les enfants pourraient s'appuyer sur ces phénomènes d'analogie. En effet, dans la mesure où [ca-] en position initiale est plus fréquent que [na-], il est vraisemblable que davantage de mots de la langue commencent par [ca-] que par [na-] et que ces mots commençant

par [ca-] aient davantage de voisins orthographiques que ceux débutant par [na-]. Face à des pseudo-mots contenant un bigramme fréquent, les enfants pourraient avoir recours aux analogies avec des mots connus (commençant par *ca-*) et au voisinage orthographique pour déterminer leur préférence. Cette idée irait dans le sens du modèle de lecture par analogies de Gombert *et al.* (1997). Cette hypothèse semble également valable chez les adultes qui manifestent aussi une sensibilité à la fréquence des bigrammes mais indépendamment de la position dans l’item. Cependant, comme dans le cas des performances graphophonologiques implicites, les scores obtenus à la tâche orthographique implicite ne plafonnent pas : les choix d’items à bigramme fréquent sont majoritaires et reposent essentiellement sur la prise en compte de la structure orthographique mais les préférences manifestées pour les items alternatifs (i.e., items à bigrammes rares) indiqueraient l’intervention d’autres facteurs que ceux strictement orthographiques.

Contrairement à ce qui a pu être observé pour les connaissances graphophonologiques, le schéma développemental décrit par Gombert (2003) selon lequel les connaissances explicites s’élaborent sur la base des connaissances implicites n’est pas validé pour l’orthographe. Dans l’épreuve orthographique explicite, les enfants éprouvent des difficultés à déterminer si davantage de mots commencent par [ca-] que par [na-] ou se terminent par [-te] ou [-ve]. D’une manière générale, les enfants semblent plus en difficulté à traiter de la fréquence des bigrammes lorsque ceux-ci sont présentés de façon isolée plutôt qu’insérés dans des pseudo-mots. Ainsi, l’épreuve explicite ne se différencie pas de la tâche implicite et les performances n’évoluent pas avec le niveau de lecture sauf à un niveau d’expertise élevé où les scores augmentent sensiblement. La faiblesse relative des scores à l’épreuve orthographique explicite et l’absence d’évolution des performances avec le niveau de lecture chez les enfants peuvent être expliquées par la difficulté à réaliser un traitement orthographique sur des bigrammes isolés, moins concrets que des mots ou même des pseudo-mots contenant ces bigrammes et sur lesquels un recours aux analogies était envisageable. La nature même de la tâche pourrait donc être un facteur explicatif aux performances obtenues au test orthographique explicite car comme le souligne Carlisle (1995) des tâches peu familières peuvent parfois poser problème à des enfants qui commencent tout juste à développer une conscience explicite sur la structure des mots.

Concernant l’orthographe, une des principales critiques adressées aux modèles développementaux d’apprentissage de la lecture (voir chapitre 1) et d’acquisition de l’orthographe (*cf.* Chapitre 3) était la prise en compte tardive des traitements orthographiques dans les procédures de reconnaissance de mots. Une critique similaire peut être adressée dans le

cas de la morphologie. La morphologie est censée n'intervenir que tardivement dans la plupart des modélisations (voir chapitre 1) et bien souvent les chercheurs ont justifié cette intervention tardive en considérant les connaissances morphologiques des apprentis lecteurs insuffisamment développées. Pourtant, depuis une dizaine d'années, des recherches de plus en plus nombreuses montrent que les enfants sont sensibles à certains aspects morphologiques de la langue (comme les phénomènes de préfixation et de suffixation) et développeraient avant même l'apprentissage de la lecture des connaissances implicites en morphologie (Colé *et al.*, 2003 ; Marec-Breton, 2003). Sur la base de ces connaissances implicites, une conscience plus explicite de la structure morphologique pourrait même émerger dès le début de l'apprentissage de la lecture (*ibidem* ; Casalis *et al.*, 1996 ; Rubin, 1988). La majeure partie des tâches utilisées dans ces recherches a été proposée à l'oral. L'expérience 3 réalisée dans la présente étude visait à vérifier la sensibilité morphologique précoce mais en utilisant des tâches écrites.

Les résultats obtenus indiquent une sensibilité à la morphologie écrite à partir du niveau de lecture CE1, quel que soit le type d'affixe en jeu (préfixe ou suffixe). Cette sensibilité augmente avec l'apprentissage de la lecture mais contrairement aux expériences réalisées à l'oral par Colé *et al.* (2003) ou Marec-Breton (2003), elle n'apparaît pas dès la GSM ni même au CP. Chez les prélecteurs et les lecteurs débutants, il est difficile de conclure à une sensibilité à la structure morphologique des mots écrits en lien avec le niveau de lecture. Ces résultats contradictoires peuvent être expliqués par la relative complexité de la tâche écrite par rapport aux épreuves orales habituellement proposées par les chercheurs dans ce domaine. En GSM et au tout début de l'apprentissage de la lecture, malgré une exposition de plus en plus fréquente à l'écrit, la confrontation à des mots écrits inconnus et le jugement demandé aux enfants sur ces items ont pu les déstabiliser. Cette explication semble la plus plausible compte tenu des résultats obtenus sur le même thème par les chercheurs cités précédemment ou encore ceux de Gaux (2002) qui a utilisé également des épreuves morphologiques orales mais dont la consigne diffère. Cet auteur a réalisé une première étude transversale de la GSM au CE1 dans laquelle les enfants devaient choisir entre deux mots (*lusseur-chassier*) celui qui allait avec le mot cible dérivé proposé (*chasseur*). Ses résultats montrent que les enfants, dès la GSM, choisissent davantage l'item ayant la même base que le mot cible, les performances augmentant de la GSM au CP et plafonnant en CE1. Une seconde étude réalisée de la GSM au CP consistait pour l'enfant à choisir entre deux mots (*chanaute-chassaise*) celui allant avec le mot cible dérivé proposé (*chasseur*). A nouveau, les enfants préféraient l'item ayant la même base que le mot cible et cette préférence augmentait avec le niveau de lecture. Ces derniers résultats tendent à montrer qu'une sensibilité morphémique existe dès la GSM et que son développement est influencé par

l'apprentissage de la lecture. À l'oral, une sensibilité morphologique pourrait donc apparaître précocement tandis qu'à l'écrit il faut attendre un certain niveau d'efficacité en lecture (niveau CE1) pour que celle-ci puisse s'exprimer et s'accroître avec l'apprentissage.

Les données obtenues à l'épreuve explicite confortent cette hypothèse. La sensibilité à la structure morphologique développée au niveau CE1 sert d'appui au développement des connaissances explicites élaborées à ce même niveau de lecture. De plus, ces connaissances explicites augmentent avec l'efficacité. La tâche de détection d'intrus morphologiques servant à évaluer les connaissances explicites est similaire dans son principe à celle utilisée par Colé *et al.* (2003), excepté le fait qu'elle soit présentée à l'écrit dans la présente étude. Or, les résultats obtenus indiquent cette fois-ci une apparition plus précoce des capacités d'analyse morphologique par rapport aux données de Colé *et al.* (2003) qui indiquaient un choix correct de l'intrus seulement à partir du CE2. Un autre résultat intéressant concerne les enfants de niveau CE2 qui obtiennent de meilleurs résultats que ceux de niveau CM1. Ces résultats peuvent être rapprochés de ceux constatés précédemment pour l'épreuve graphophonologique explicite pour le niveau de lecture CE2. Pour ces enfants, il semble qu'il y ait un véritable saut qualitatif et quantitatif quant aux connaissances possédées en lien avec l'apprentissage de la lecture. Comme suggéré par Marouby-Terriou et Denhière (2002), ce niveau de lecture pourrait être caractérisé par une réorganisation des connaissances qui permettrait notamment d'expliquer le « bond » dans les performances obtenues aux items suffixés, en particulier, entre le niveau CE1 (44,83 % d'intrus de type suffixé correctement détectés) et le niveau CE2 (75 %). C'est d'ailleurs pour ce niveau de lecture CE2 qu'une différence apparaît entre préfixés et suffixés à l'avantage de ces derniers alors que pour les niveaux CP, CE1 et CM1 aucune différence n'est observée entre les deux catégories d'affixes. En revanche, chez les enfants de niveau CM2 et les lecteurs experts, le principe de préfixation semble bien mieux maîtrisé que celui de suffixation. Chez les adultes experts, les connaissances morphologiques poursuivent leur développement. Conformément aux résultats déjà obtenus par Marec-Breton (2003), ce développement semble étalé dans le temps et n'est pas terminé, notamment pour le versant explicite comme en témoignent les scores relativement proches des adultes par rapport à ceux des enfants de niveau CM2. La différence entre préfixés et suffixés observée chez les adultes ne conforte pas l'asymétrie de traitement en faveur des suffixés défendue par Colé *et al.* (1986). Pour ces auteurs, cette asymétrie entre les mots préfixés et suffixés relèverait directement de la différence d'organisation morphologique de ces deux types de mots (affixe + racine *vs.* racine + affixe) qui impliquerait que les mots suffixés et non pas les mots préfixés sont accédés *via* la représentation de la racine. Ainsi, le traitement des mots suffixés serait plutôt guidé par la représentation de la racine (située en début de mot, la

racine étant encore plus prédominante en modalité auditive) alors que celui des mots préfixés pourrait être plutôt guidé par la représentation de leur préfixe puis de la racine, ce qui ralentirait le traitement pour les préfixés. Les résultats obtenus dans la présente étude avec les adultes ne vont pas dans le sens de cette hypothèse mais concorderaient plutôt avec l'interprétation des linguistes en se focalisant sur la description fonctionnelle des préfixes par rapport aux suffixes (e.g., Gardes-Tamine, 1988). En effet, la différenciation des phénomènes de préfixation et de suffixation peut être établie du point de vue des changements sémantiques et morphosyntaxiques que l'adjonction des deux types d'affixes produit sur la racine. Ainsi, l'ajout d'un préfixe à une racine n'entraîne jamais la création d'un mot nouveau dont la classe morphosyntaxique diffère de la racine. De plus, les préfixes n'ont pas de fonction grammaticale, ils introduisent seulement un changement de sens. Leur fonction est donc exclusivement sémantique. Les suffixes, à l'inverse, ont des fonctions linguistiques beaucoup plus diversifiées, ils ont des valeurs grammaticales et contrairement aux préfixes, l'ajout d'un suffixe peut entraîner des allomorphes de la racine (tels que *plomb* dans *plombier*). Dès lors, en français, les préfixes apparaissent beaucoup plus transparents sémantiquement que les suffixes et pourraient alors être plus facilement accessibles.

1.2- Liens avec la lecture et le vocabulaire

De nombreuses études établissent des liens entre les connaissances graphophonologiques, orthographiques, morphologiques et la lecture (*cf.* Chapitre 5). Avec l'apprentissage de la lecture, le vocabulaire disponible tend à s'accroître : il est donc vraisemblable que le niveau de vocabulaire soit également corrélé à la lecture et aux différents types de connaissances implicites et explicites. Les résultats obtenus pour chaque niveau de lecture confortent en partie les hypothèses sur l'existence de tels liens.

Ainsi, au niveau de lecture CP, ce sont les enfants qui obtiennent les scores les plus élevés à la tâche graphophonologique explicite, à la tâche morphologique de plausibilité lexicale (morphologie implicite) et à l'épreuve de vocabulaire qui ont les scores les plus élevés en lecture. Les analyses de régression ne confirment que la participation du vocabulaire et des connaissances morphologiques implicites à la lecture. Pour le niveau CP, Carlisle et Nomanbhoy (1993) constataient que la conscience phonologique était le meilleur prédicteur des scores en lecture tandis que la morphologie rendait compte de seulement 4 % de variance supplémentaire. Les résultats de la présente étude indiquent une plus grande part de la morphologie en lecture dès le CP tandis que le vocabulaire disponible et non les capacités d'analyse graphophonologique constitue le prédicteur le plus puissant des compétences en identification de mots. En revanche, à

partir du niveau CE1, le profil est en partie différent : le niveau de lecture est corrélé aux performances en vocabulaire et aux connaissances graphophonologiques explicites mais ce sont ces dernières qui contribuent le plus à expliquer les scores en lecture. Cette contribution est moins importante que celle obtenue par Carlisle et Nomanbhoy (1993) pour le niveau CP (33 %) mais elle reste non négligeable (28 %). Ce résultat conforte de nombreuses études comme celles de Bruck *et al.* (1997), Manis *et al.* (1993) ou encore Sprenger-Charolles *et al.* (2003) qui montrent que les compétences phonologiques ont un fort pouvoir prédictif sur les capacités d'identification de mots. A partir du niveau CE2, les recherches de Carlisle (2000) et de Mahony *et al.* (2000) montraient que la contribution des connaissances métamorphologiques (i.e., morphologie explicite) à la lecture augmentait. Or, si effectivement les enfants de niveau CE2 les plus performants à la tâche de détection d'intrus morphologiques (versant explicite) sont aussi ceux qui ont les scores les plus élevés en lecture (et en vocabulaire), la contribution des connaissances morphologiques explicites à la lecture n'est pas significative. Les scores obtenus à l'épreuve graphophonologique explicite restent liés à la lecture comme au niveau CE1 mais ne contribuent plus de façon significative à expliquer les performances en lecture. A des niveaux de lecture plus élevés, CM1, CM2 et lecteurs experts, les scores de *Lecture en Une Minute* ne sont corrélés à aucun des scores obtenus aux différents types de tâches implicites et explicites. Pour les lecteurs les plus experts, notamment le niveau CM2 et les adultes, le vocabulaire disponible est la seule variable qui semble constituer une aide en lecture mais aussi dans la réalisation des épreuves d'analyse morphologique (adultes et niveau CM2) et graphophonologique (niveau CM2). A mesure que l'expertise en lecture augmente, l'étendue du vocabulaire permettrait d'accroître la réussite en lecture, conformément aux liens obtenus entre ces deux types de compétences par Anderson et Freebody (1985), Anglin (1993), Colé *et al.* (2003) ainsi que Nagy et Anderson (1984).

Globalement, les résultats de la présente étude ne vont pas dans le sens des prédictions de Carlisle et Nomanbhoy (1993), Carlisle (1995) ou encore de Casalis et Louis-Alexandre (2000) indiquant que les connaissances morphologiques étaient impliquées dans la lecture mais plus tardivement que les compétences phonologiques. Dès le niveau de lecture CP, les connaissances morphologiques implicites contribuent aux performances en lecture de manière indépendante même si au niveau CE1, l'impact de l'apprentissage explicite des correspondances graphèmes-phonèmes favorise une contribution majoritaire des capacités métaphonologiques à l'efficacité en lecture. L'absence de liens entre lecture et connaissances graphophonologiques, orthographiques et morphologiques dès le niveau CM1 pourrait s'expliquer par la nature de l'épreuve (*LUM*, Khomsi, 1999) utilisée pour évaluer les capacités de lecture (i.e., exactitude

pour les enfants et temps de lecture moyen par mot pour les lecteurs experts). Cette épreuve est sans doute pertinente pour évaluer le niveau de lecture au début de l'apprentissage mais est sans doute moins adaptée au fur et à mesure que l'expertise augmente.

1.3- Liens entre les connaissances graphophonologiques, orthographiques et morphologiques

En se référant au modèle de lecture par analogies de Gombert *et al.* (1997) ainsi qu'au modèle de Colé *et al.* (1999), des interactions entre unités orthographiques et unités phonologiques apparaissent relativement précocement. Dans de tels modèles, la reconnaissance de mots écrits nécessite l'activation de représentations phonologiques qui elles-mêmes sont activées par les représentations orthographiques lexicales et par les représentations phonologiques lexicales. Le développement des connaissances graphophonologiques et orthographiques ne s'effectuerait donc pas de façon successive, contrairement aux prédictions des modèles développementaux de l'apprentissage de la lecture, mais plutôt de façon interactive. Ce lien éventuel entre connaissances orthographiques et connaissances graphophonologiques est vérifié dans la présente étude mais uniquement chez les enfants de niveau CE2 et les lecteurs experts pour le versant implicite. Pour les enfants de niveau CE2, la présence de ce lien conforte l'hypothèse suggérée dans le chapitre 6 pour expliquer l'amélioration importante des performances aux items irréguliers de l'épreuve graphophonologique explicite entre le niveau CE1 et le niveau CE2. Au niveau CE2, les enfants utiliseraient leurs connaissances orthographiques parallèlement aux connaissances graphophonologiques, notamment face à des mots irréguliers. En effet, l'irrégularité d'un mot renvoie à la notion de consistance graphophonologique mais également aux représentations orthographiques en mémoire possédées par l'enfant sur les mots ne pouvant être lus par application directe des correspondances graphèmes-phonèmes. En revanche, à des niveaux de lecture plus faibles comme le niveau CP, les connaissances graphophonologiques interagissent plutôt avec les connaissances morphologiques.

L'existence d'un lien entre connaissances morphologiques et connaissances phonologiques a déjà été mis en évidence dans plusieurs études (Carlisle, 1995 ; Fowler & Liberman, 1995 ; Mahony *et al.*, 2000) mais la nature de ce lien reste difficile à déterminer. Par exemple, pour Carlisle et Nomanbhoy (1993), les morphèmes sont identifiés de façon globale tant que l'enfant n'a pas conscience des syllabes et des phonèmes. Puis, lorsque cette conscience est établie, les enfants sont en mesure d'analyser les mots et de découvrir les rôles sémantiques et syntaxiques des suffixes. Enfin, les enfants parviennent au fur et à mesure à apprendre les règles morphophonémiques et phonotactiques de leur langue. Ici, le développement des connaissances

morphologiques semble être sous la dépendance de la conscience phonologique. Pour Shankweiler *et al.* (1995), la relation entre morphologie et phonologie est plus une relation d'interaction que de dépendance. Ces auteurs ont voulu comparer les capacités langagières d'enfants mauvais lecteurs, d'enfants normo-lecteurs et d'enfants présentant un autre déficit cognitif, à partir de tâches phonologiques, morphologiques et syntaxiques. Il ressort de leur recherche que le déficit de la conscience phonologique est hautement corrélé au déficit morphologique chez les participants mauvais lecteurs. Ils ont notamment constaté qu'ils éprouvaient plus de difficulté à donner la forme dérivée d'un mot lorsque celui-ci subit un changement phonologique (par exemple, en français : /il vo/-/il val/). Ils interprètent ce déficit morphologique comme l'expression d'une limite phonologique. En effet, dans la mesure où dans une langue donnée le phonème est à la base, à la fois, de l'organisation du lexique et de l'organisation de la grammaire (et par conséquent de la morphologie), il ne serait pas surprenant que des enfants dyslexiques, qui présenteraient un déficit phonologique, montrent également un déficit morphologique, puisque les deux modules (phonologique et morphologique) s'étayaient mutuellement. Une erreur morphologique pourrait alors expliquer une erreur phonologique et réciproquement. Par exemple, une confusion phonologique /t/-/d/ risque d'engendrer une confusion lexicale (exemple : /petal/-/pedal/) ou encore une confusion phonologique /r/-/l/ peut entraîner une confusion grammaticale dans l'appréhension du futur. Par conséquent, un déficit phonologique peut s'accompagner d'un déficit morphologique et tout comme il y aurait une conscience phonologique, il existerait une conscience morphologique. Dans une telle recherche, le développement de la conscience phonologique et de la conscience morphologique apparaît réciproque. Fowler et Liberman (1995) nuancent cependant un tel point de vue : selon eux, la distinction entre les deux types de conscience (phonologique et morphologique) est parfois difficile à établir dans la mesure où une insensibilité aux phonèmes et aux morphèmes peut tout à fait provenir d'un seul et même déficit phonologique. Récemment, un ensemble de travaux ayant utilisé le paradigme d'amorçage a proposé que le traitement des mots morphologiquement complexes impliquait des unités morphémiques bien distinctes des représentations formelles (orthographiques/phonologiques) et sémantiques des mots (Colé *et al.*, 1997 ; Deutsch *et al.*, 1998 ; Drews & Zwitserlood, 1995 ; Frost, Deutsch, & Forster, 2000 ; Frost, Forster, & Deutsch, 1997 ; Giraud & Grainger, 2001 ; Grainger, Colé, & Segui, 1991 ; Longtin, Segui, & Hallé, 2003 ; Pastizzo & Feldman, 2002 ; Schriefers, Zwitserlood, & Roelofs, 1991 ; Stolz & Besner, 1998 ; Taft, 1994 ; Zwitserlood, 1994). Il est possible d'en déduire que les unités morphologiques sont représentées de façon explicite et indépendante au sein de l'architecture du lexique mental. Ces différents résultats montrent que la question des liens entre morphologie et

phonologie est loin d'être tranchée. Dans la présente étude, au-delà du niveau de lecture CP, les tâches morphologiques ne sont plus corrélées aux épreuves graphophonologiques. Par ailleurs, au niveau de lecture CP, les connaissances morphologiques implicites contribuent à la lecture de façon indépendante des connaissances phonologiques indiquant que la morphologie n'est pas le simple reflet de l'émergence d'une conscience phonologique explicite.

Une des hypothèses de cette étude était également l'existence d'un lien entre connaissances orthographiques et morphologiques. Pour Fowler et Liberman (1995), un tel lien est envisageable dans la mesure où la connaissance orthographique est une source puissante d'information concernant les relations morphologiques, en particulier celles de type dérivationnel. Or, dans la présente étude, aucune corrélation entre connaissances orthographiques et connaissances morphologiques n'est obtenue niveau par niveau. La nature du lien entre orthographe et morphologie reste donc à déterminer. Il existe des études montrant l'existence d'un lien entre connaissances morphologiques flexionnelles et orthographe (Pacton & Fayol, 2003 ; Sénéchal, 2000b ; Totereau, Thévenin, & Fayol, 1997...) et entre connaissances dérivationnelles et orthographe (Marec-Breton, 2003 ; Pacton, 2000). Le fait de ne pas retrouver de tels liens ici est surprenant et peut être attribué en particulier à la complexité de l'épreuve orthographique explicite comme cela a déjà été souligné précédemment.

Enfin, en référence au schéma développemental d'apprentissage de Gombert (2003), il était supposé que des liens devaient être observés entre les versants implicite et explicite pour chacune des dimensions étudiées (graphophonologique, orthographe et morphologie). Niveau par niveau, ce lien n'est pas observé en morphologie, contrairement aux résultats obtenus par Marec-Breton (2003) indiquant que les connaissances implicites contribuaient à l'installation des connaissances explicites. Le même type de résultat est obtenu dans le cas des connaissances graphophonologiques. La seule corrélation significative entre implicite et explicite concerne les connaissances orthographiques. Ainsi, chez les enfants de niveau CE1, CM2 et les lecteurs experts, la sensibilité à la structure orthographique est liée à de bonnes performances à l'épreuve explicite. Le schéma développemental de Gombert (2003) sur l'articulation entre implicite et explicite en termes de complémentarité n'est donc confirmé que pour les connaissances orthographiques. Ce résultat incite à revoir la construction des épreuves afin que celles-ci mesurent bien de façon distincte les aspects implicites et explicites de chaque dimension observée. Il est sans doute important de préciser que peu de commentaires ont été faits dans cette partie sur les enfants de GSM dans la mesure où aucune corrélation significative n'a pu être obtenue à ce niveau scolaire. Ces enfants, malgré une exposition à l'écrit de plus en plus importante dans le cadre du contexte scolaire, ont pu éprouver des difficultés face à des tâches

écrites auxquelles ils sont peu habitués. Néanmoins, il est envisageable de supposer qu'un apprentissage implicite de certaines régularités de la langue puisse se réaliser dès ce niveau scolaire et que les enfants soient dans une phase de transition, de passage d'une conscience implicite à une conscience plus explicite s'exprimant en CP. Surtout, les connaissances implicites éventuellement possédées par les enfants de GSM s'exprimeraient essentiellement à l'oral, mode de communication privilégié à cet âge, comme en attestent les données de Colé *et al.* (2003) et de Marec-Breton (2003) pour ce niveau scolaire.

2- Influence des traitements graphophonologiques, orthographiques et morphologiques sur la reconnaissance de mots écrits

2.1- Traitements graphophonologiques, orthographiques et morphologiques en lecture

L'objectif des expériences de lecture à haute voix de mots et de pseudo-mots était de montrer que les enfants pouvaient utiliser la structure graphophonologique (effet de consistance graphophonémique), orthographique (effet de fréquence des bigrammes) ou morphologique (effet d'affixation) des mots et que cette utilisation facilitait alors la lecture.

Dans le cas de l'étude de la structure graphophonologique, les enfants mais également les adultes devaient lire à voix haute quatre types d'items : des mots réguliers ou consistants (*banane, lavabo*), des mots inconsistants (*monsieur, automne*), des pseudo-mots consistants (*bertalie, janitou*) et enfin des pseudo-mots inconsistants (*baliden, lautil*). Une étude relativement proche a été réalisée par Sprenger-Charolles *et al.* (1998) sur les effets de régularité en lecture à haute voix du milieu du CP à la fin du niveau CM1. Les résultats obtenus montraient dès le CP un effet important de la régularité sur les performances en lecture : les mots réguliers étaient lus avec moins d'erreurs que les mots irréguliers. Les connaissances sur la consistance graphophonologique interviendraient donc précocement dans les processus de reconnaissance de mots écrits. Les données issues de la présente étude confortent ce constat : à partir du niveau de lecture CP jusqu'au niveau CM2, les mots réguliers sont lus plus rapidement et avec moins d'erreurs que les mots irréguliers. La force et la fréquence des associations entre unités graphiques et phonologiques auraient donc un impact sur la facilité à lire des mots. Cet effet de la consistance graphophonémique est prégnant pour l'ensemble des items (mots et pseudo-mots) au niveau des temps de lecture mais seulement dans le cas des mots familiers pour l'exactitude. Exceptés chez les enfants de niveau CE2 qui lisent mieux les pseudo-mots irréguliers que les pseudo-mots réguliers, la différence entre les deux catégories d'items en termes d'exactitude n'est pas significative pour les autres niveaux de lecture. Ce résultat s'explique aisément par la construction des pseudo-mots : pour les irréguliers, une prononciation irrégulière ou régulière est

acceptable en raison de la non lexicalité des items. Dès lors, le nombre d’erreurs de prononciation sur les pseudo-mots irréguliers n’est pas plus élevé que sur les pseudo-mots réguliers. Ainsi, conformément aux résultats des nombreuses comparaisons interlangues (Goswami *et al.*, 1997 ; Jansen, 1995 ; Jorm *et al.*, 1984 ; Juel *et al.*, 1986 ; Oney & Goldman, 1984 ; Porpodas, 1999 ; Seymour *et al.*, 2003 ; Treiman *et al.*, 1990 ; Wimmer & Goswami, 1994), il semble bien que la plus ou moins forte consistance des correspondances graphèmes-phonèmes ait un impact réel sur la lecture de mots connus chez les enfants scolarisés. A un niveau d’expertise en lecture plus élevé, cet effet de la régularité persiste : les adultes lisent plus vite les items réguliers (*janitou*) que les irréguliers (*baliden*) surtout lorsqu’ils rencontrent ces mots pour la première fois (i.e., statut des pseudo-mots).

Dans le cas des relations entre orthographe et lecture, il n’existe pas, *a priori*, de recherches ayant étudié spécifiquement l’impact de la fréquence bigrammique sur la lecture de mots et pseudo-mots. Néanmoins, les expériences réalisées sur le développement des connaissances orthographiques implicites et explicites (voir chapitre 6, expérience 2) indiquent une sensibilité précoce à la fréquence des bigrammes. De plus, en se référant au modèle de lecture de Gombert *et al.* (1997), il est envisageable de considérer que les enfants (et les adultes) puissent s’appuyer sur des analogies avec des mots connus contenant des bigrammes fréquents dans la langue pour lire les mots et pseudo-mots de l’épreuve de lecture à haute voix. Pour cette épreuve, quatre catégories d’items étaient proposées : des mots et des pseudo-mots avec des bigrammes fréquents en position initiale (*camion*, *caloude*) ou finale (*tomate*, *bofite*) ainsi que des mots et des pseudo-mots contenant des bigrammes rares en position initiale (*narine*, *nalube*) ou finale (*lessive*, *toruue*). L’hypothèse d’une facilitation de la lecture liée à la fréquence du bigramme n’est pas validée chez les enfants : ni l’exactitude, ni le temps de lecture des mots et pseudo-mots ne sont influencés par la fréquence des bigrammes et ce quelle que soit la position du bigramme dans l’item. Une explication à ce résultat pourrait être l’utilisation non pas des connaissances sur la fréquence bigrammique mais l’utilisation directe des correspondances graphèmes-phonèmes pour lire les items. En effet, les bigrammes utilisés dans l’épreuve de lecture sont différents en termes de fréquence mais, quelle que soit cette fréquence, leurs codes articulatoires sont facilement récupérables (en position initiale et en position finale). Ainsi, les prononciations de [ca-] et [na-] sont toutes deux régulières alors que les fréquences des deux bigrammes sont différentes. Il en est de même pour les bigrammes [-te] et [-ve] en position finale. En considérant les correspondances phonémiques de ces graphèmes, il est alors peu étonnant que les items contenant un bigramme fréquent soient lus de façon similaire à ceux contenant un bigramme rare, quelle que soit la position dans l’item. Cependant, chez les adultes,

l'hypothèse d'une lecture facilitée par la fréquence bigrammique est validée lorsque le bigramme est en position initiale. A un niveau d'expertise en lecture plus élevé, il semble que la structure orthographique puisse faciliter le traitement des mots connus comme des mots non familiers.

La structure graphophonologique mais pas celle orthographique constitue une aide lors de la lecture à haute voix. Les résultats obtenus en lecture de mots et pseudo-mots construits ou non morphologiquement (*cf.* Expérience 3, chapitre 6) apportent des informations complémentaires sur les traitements effectués par les enfants et les adultes. Depuis plusieurs années, de nombreux chercheurs se sont intéressés à l'analyse morphologique des mots au cours de la lecture. Une des premières études réalisées à ce sujet était celle de Laxon *et al.* (1992). Ces auteurs avaient constaté qu'un traitement morphologique des mots fléchis mais également des mots dérivés était réalisé dès l'âge de 7 ans. Ainsi, les enfants lisaient des mots suffixés tels que *dancer* (*danseur*) avec moins d'erreurs que des mots pseudo-suffixés comme *dinner* (dans lequel *-er* n'est pas un suffixe). Laxon *et al.* (1992) en concluaient que les enfants décomposaient les mots complexes et que cette décomposition morphologique facilitait la lecture. Un tel traitement morphologique a été confirmé dans le cas des mots fléchis chez les enfants de CM2 par Feldman *et al.* (2002). Burani *et al.* (2002) ont également constaté que dès le CE2, les enfants utilisaient les informations morphologiques contenues dans des mots nouveaux pour les lire. Plus récemment, Marec-Breton (2003) et Marec-Breton *et al.* (2005) ont mis en évidence de façon plus précoce cet impact de la structure morphologique en lecture. Leurs recherches ont montré que les lecteurs débutants lisaient mieux (plus vite et avec moins d'erreurs) les mots et pseudo-mots préfixés que les pseudo-préfixés. La structure morphologique interviendrait donc dans la lecture mais pas pour tous les items construits morphologiquement. Ainsi, Marec-Breton *et al.* (2005) constatent que les mots suffixés ne sont pas mieux lus que les mots pseudo-suffixés. L'effet de la structure morphologique pourrait donc être dépendant du type d'affixe contenu dans l'item. L'expérience de lecture à haute voix effectuée dans la présente étude visait à confirmer un tel traitement morphologique ainsi qu'à vérifier l'éventuelle asymétrie entre préfixé et suffixé. Pour cela, quatre types d'items ont été proposés aux enfants du niveau de lecture CP au niveau CM2 ainsi qu'aux adultes : des mots affixés (en *re-* et *dé-* pour les préfixes et en *-eur* et *-age* pour les suffixes), des mots pseudo-affixés (*désirer*, *message*), des pseudo-mots affixés (*reranger*, *gareur*) et des pseudo-mots pseudo-affixés (*dénaller*, *tarmeur*). Les résultats obtenus confortent en partie les données déjà obtenues par d'autres auteurs dans le même domaine. Dès le niveau CP, les connaissances morphologiques sur le phénomène de préfixation interviennent dans les procédures de reconnaissance des mots familiers mais aussi des mots nouveaux. Du niveau de lecture CP au niveau CM2, les items préfixés sont lus avec moins d'erreurs que les pseudo-

préfixés, conformément aux résultats de Marec-Breton (2003) et de Marec-Breton *et al.* (2005). A partir du niveau CE1 et jusqu'au niveau CM2, ces items préfixés sont également traités plus rapidement que les pseudo-préfixés. L'effet de la structure morphologique en lecture est donc réel et notablement plus précoce que celui constaté par Burani *et al.* (2002). Cette facilitation de la lecture pour les items construits morphologiquement semble acquise mais uniquement dans le cas des préfixés. En effet, dans le cas des suffixés, la structure morphologique ne facilite pas la lecture des items affixés dans la mesure où les pseudo-suffixés sont mieux lus (plus rapidement et avec moins d'erreurs) que les suffixés dans le cas des mots. Pour les pseudo-mots, aucune différence significative n'est obtenue entre suffixés et pseudo-suffixés. Ce résultat est très surprenant au vu des résultats de Marec-Breton (2003) qui observait un effet de la structure morphologique sur la lecture des préfixés et des suffixés pour les enfants du CP au CM2. L'étude réalisée en 2005 par Marec-Breton *et al.* (2005) indiquait toutefois que pour les items suffixés, l'influence de la structure morphologique ne se manifestait qu'au niveau CE1 sur les scores de lecture des pseudo-mots. Ce dernier résultat tend à montrer que l'analyse morphologique pourrait être plus précoce et systématique pour les préfixés que pour les suffixés mais s'oppose malgré tout aux données obtenues pour les suffixés dans la présente étude. Le résultat pour les items suffixés semble indiquer que pour les enfants les mots pseudo-suffixés seraient plus familiers que les mots suffixés. Ce résultat est contradictoire avec les recherches réalisées jusqu'à présent et s'avère difficile à interpréter ici : il suggère tout au plus que le traitement des préfixés et suffixés est asymétrique, plutôt au bénéfice des préfixes, contrairement aux propos de Colé *et al.* (1986). Un autre résultat étonnant concerne les adultes. Pour l'épreuve de lecture à haute voix, le caractère plurimorphémique n'a pas d'impact sur les temps de lecture correcte que l'affixe testé soit un préfixe ou un suffixe. Ainsi, les lecteurs experts lisent aussi rapidement les items affixés que les pseudo-affixés. Auprès d'une même population, Marec-Breton (2003) avait constaté que les adultes lisaient mieux les mots pseudo-préfixés que les mots préfixés. Le fait que les adultes traitent de manière similaire les items construits morphologiquement et ceux non construits semble indiquer que ces lecteurs ont recours à un traitement global des items et non décompositionnel. Ainsi, pour les lecteurs experts, conformément au modèle de Caramazza *et al.* (1988), les traitements réalisés en lecture impliqueraient une intervention conjointe des affixes et des bases favorisant un traitement identique des dérivés et des pseudo-dérivés qui seraient alors reconnus aussi rapidement. Chez les enfants, une analyse morphologique des mots connus et des mots peu familiers serait effectuée, en particulier dans le cas des préfixés (voir Alvarez *et al.*, 2001 ; Bertram & Hyona, 2003 ; Bertram *et al.*, 2000 ; Colé *et al.*, 1997 ; Hyona *et al.*, 2002 qui soutiennent cette analyse systématique mais chez le lecteur expert).

2.2- Evolution des traitements graphophonologiques, orthographiques et morphologiques

Pour la lecture, l'échantillon de la présente étude constitué d'enfants de niveau de lecture CP à CM2 ainsi que de lecteurs adultes experts a permis de rendre compte de l'évolution des traitements graphophonologiques, orthographiques et morphologiques réalisés par les lecteurs. En outre, la comparaison spécifique des enfants de niveau de lecture CM2 et des adultes avait pour objectif de voir si au niveau CM2, les enfants sont toujours apprenants ou si les processus de reconnaissance de mots connus et nouveaux sont proches de l'expertise.

En graphophonologie, l'expérience 1 réalisée sur les enfants de niveau de lecture CP à CM2 et l'expérience 4 effectuée sur les lecteurs experts indiquent des variations dans les traitements selon le niveau d'efficacité. L'effet de consistance graphophonologique sur la lecture des mots existe pour tous les enfants mais les enfants de niveau CP et CE1, moins experts en lecture, font davantage d'erreurs et sont moins rapides que ceux de niveau CE2 à CM2. Un tel profil était déjà relevé par Sprenger-Charolles *et al.* (2003) : au fur et à mesure de l'apprentissage de la lecture, les scores relevés pour les mots réguliers en lecture augmentent. Avec l'accroissement du niveau de lecture, les items bénéficient à la fois de la régularité et de la fréquence d'exposition permettant une meilleure précision de la lecture. Le traitement graphophonologique opéré par les adultes est toutefois différent de celui des enfants : l'expertise en lecture permet de lire de la même façon les mots réguliers et les mots irréguliers.

Dans le cas des traitements orthographiques, peu de variations sont constatées entre les enfants des différents niveaux de lecture. De façon générale, la fréquence bigrammique n'a pas d'effet sur la lecture (exactitude et temps de lecture correcte), que le bigramme soit en position initiale ou finale dans les items lus. Les principales différences relevées concernent les enfants de niveau CP et CE1 qui font plus d'erreurs de lecture sur les items contenant un bigramme en position initiale. Les enfants des niveaux CP et CE2 sont également plus lents dans le traitement de ces items. L'exactitude des enfants les moins efficaces en lecture est également en deçà des autres enfants pour les items à bigramme en position finale. Ces mêmes enfants et ceux de niveau CE1 et CE2 traitent aussi moins rapidement ces items. Par rapport aux lecteurs experts, aucune différence de traitement n'est constatée pour les items à bigramme en position finale : comme pour les enfants, la structure orthographique n'a pas d'effet sur la vitesse de traitement. En revanche, un traitement différentiel est constaté pour les items contenant un bigramme en position initiale : les lecteurs experts lisent plus vite les items à bigramme fréquent que ceux à bigramme rare alors que l'effet de la structure orthographique restait sans effet pour les enfants.

Les différences les plus importantes en fonction du niveau d'expertise en lecture semblent concerner les traitements morphologiques des préfixés et suffixés. En effet, pour les préfixés, l'effet de la structure morphologique est significatif sur la précision en lecture de mots et pseudo-mots pour tous les enfants même si les enfants de niveau CP et CE1 font plus d'erreurs. La structure morphologique a également un impact sur la vitesse de traitement des mots et des pseudo-mots : les enfants de niveau CE1 et CE2 traitent moins rapidement l'ensemble des items mais ils lisent les préfixés plus vite que les pseudo-préfixés. Il en est de même pour les enfants des niveaux CM1 et CM2. Enfin, les adultes ne semblent pas se servir de la structure morphologique, contrairement aux enfants, pour lire les items de type préfixé. Pour les suffixés, des différences notables de traitement apparaissent. Les enfants de niveau CP font plus d'erreurs que les autres enfants. Les enfants de niveau CE1 font également plus d'erreurs que ceux de niveau CE2 et CM1 mais seulement sur les items construits morphologiquement. Les erreurs sur ces items sont aussi plus nombreuses chez les enfants de niveau CE2 par rapport à ceux de niveau CM2. Surtout, la variation la plus importante par rapport aux préfixés concerne l'effet de la structure morphologique : cet effet est absent pour les pseudo-mots (scores et temps de lecture) alors que pour les mots, ce sont les pseudo-suffixés qui sont lus avec moins d'erreurs et plus rapidement que les suffixés. Pour les suffixés, la décomposition morphologique ne faciliterait donc pas la lecture des mots, contrairement aux résultats de l'expérience princeps de Laxon *et al.* (1992). Le traitement morphologique des lecteurs experts est différent de celui des enfants sur les mots : comme pour les préfixés, l'effet de la structure morphologique n'est pas significatif sur la vitesse de traitement. Ces résultats obtenus avec les adultes contredisent les données de Burani *et al.* (2002) et celles de Marec-Breton (2003) qui constataient peu de différences entre les enfants de niveau CE2 et les experts.

Exceptés pour les traitements orthographiques (items à bigramme fréquent en position finale), les adultes et les enfants, tous niveaux de lecture confondus, ne semblent pas traiter de manière similaire les items. L'étude spécifique des traitements réalisés par les enfants de niveau CM2 et les adultes visait à déterminer si les enfants de niveau CM2 pouvaient être considérés comme experts en lecture. Les résultats issus de l'expérience 4 (voir chapitre 7) indiquent que d'une manière générale, les enfants de niveau CM2 lisent moins vite que les adultes, quelle que soit la structure des items lus. Cette différence de vitesse de traitement semble logique : entre le niveau CM2 et l'âge adulte, les années de scolarité concourent essentiellement à « gagner » en rapidité de lecture. Si adultes et enfants sont sensibles à la structure graphophonologique, cette sensibilité ne s'exprime pas de la même façon en fonction de la lexicalité des items : elle

s'exerce surtout pour les pseudo-mots chez les experts et sur les mots chez les enfants. Les adultes et les enfants de niveau CM2 se différencient dans les traitements des items à bigramme en position initiale : les adultes sont sensibles à la fréquence bigrammique mais pas les enfants. Le traitement des items préfixés diffère pareillement : les enfants de niveau CM2 se servent de la structure morphologique pour lire mais pas les adultes. En revanche, pour les pseudo-mots suffixés, les enfants de niveau CM2 comme les adultes ne sont pas sensibles à la structure morphologique. Pour cette catégorie d'items, il est donc possible de conclure à une similarité de traitement entre experts et enfants de niveau CM2 tandis que pour les mots suffixés le traitement reste asymétrique.

CONCLUSION

L'objectif de cette thèse n'était pas de remettre en cause les modèles développementaux d'apprentissage de la lecture qui mettent en avant l'importance de l'apprentissage explicite des correspondances graphèmes-phonèmes dans les processus de reconnaissance de mots écrits. Il s'agissait plutôt de montrer que ces modèles donnent une vision trop statique de l'apprentissage de la lecture : les connaissances orthographiques et morphologiques y sont peu prises en compte ou de façon tardive et succédant aux connaissances graphophonologiques sous l'influence du strict enseignement scolaire. Or, pour mieux comprendre l'apprentissage de la lecture, il faut tenir compte de la structure statistique, c'est-à-dire des régularités spécifiques d'une langue donnée. Gombert (2003) considère ainsi que, très tôt, grâce notamment à l'exposition précoce à l'écrit, l'enfant est en mesure d'extraire de son environnement linguistique des régularités phonologiques mais aussi orthographiques et morphologiques. L'enfant n'attendrait pas l'apprentissage formel de la lecture pour commencer à acquérir des connaissances sur le langage écrit. Ce sont ces connaissances dites implicites qui constitueraient des précurseurs, lesquels vont rendre possible l'apprentissage scolaire (explicite) de la lecture. La compréhension de l'apprentissage de la lecture nécessiterait donc de tenir compte de ce qui est appris mais également de l'environnement spécifique d'apprentissage, c'est-à-dire les versants implicite et explicite.

Dans cette perspective, plusieurs expériences ont été réalisées et présentées dans les chapitres 6, 7 et 8. Ces expériences confortent en partie les hypothèses formulées dans le chapitre 5 exposant la problématique de cette thèse. Ainsi, les enfants sont sensibles à la consistance graphophonémique mais seulement à partir du niveau de lecture CM1. Avant et au tout début de l'apprentissage de la lecture, la dimension de consistance graphophonémique serait peu prise en compte par les enfants. Ce premier résultat ne va pas dans le sens des résultats obtenus par Fayol *et al.* (1996) ou encore Pacton *et al.* (1999) qui montraient que dès le CP, les enfants étaient sensibles aux contraintes graphotactiques de la langue sans que ces contraintes ne renvoient à des règles apprises explicitement. Cependant, bien que la sensibilité à la consistance graphophonologique évaluée dans cette thèse émerge plus tardivement, elle participe à l'installation des connaissances explicites, en accord avec le schéma développemental de Gombert (2003). Le développement de ces connaissances explicites est conjoint au développement de l'expertise en lecture et même à l'âge adulte, ce développement se poursuit. La mise en évidence tardive de la sensibilité graphophonologique et des connaissances explicites

peut être expliquée en partie par les tâches utilisées dont la construction expérimentale s'est avérée difficile. En effet, l'objectif était d'utiliser pour l'épreuve implicite des graphèmes diversifiés pouvant être contenus dans des pseudo-mots cohérents. Dans l'épreuve explicite, la consistance graphophonologique a été testée à l'aide de mots réguliers et irréguliers sur lesquels enfants et adultes devaient fournir les prononciations régulières et irrégulières. Bien que ces deux épreuves aient été pré-testées, les comportements des adultes et des enfants relevés au cours de la passation montrent d'une part, que l'utilisation d'une tâche écrite auprès de prélecteurs et de lecteurs débutants (GSM et niveau de lecture CP) est ardue et d'autre part, que les pseudo-mots constituent un matériel déstabilisant pour les adultes. Surtout, malgré une consigne claire et des exemples fournis pour l'épreuve explicite, les enfants comme les adultes ont montré une certaine difficulté à comprendre le principe de ce test, notamment pour les premiers items. Cependant, la consistance graphophonologique semble être un élément pris en compte par les enfants et les adultes lors de la lecture. Ainsi, dès le CP, la structure graphophonologique constitue une aide dans la lecture de mots. Chez l'adulte, cette structure favorise plutôt la lecture de mots nouveaux. Ces derniers résultats montrent que même lorsque l'apprentissage explicite des correspondances graphèmes-phonèmes n'est pas encore totalement établi, en CP, les lecteurs débutants tirent parti des régularités graphophonologiques en lecture.

Un autre résultat intéressant concerne l'émergence précoce d'une sensibilité orthographique. Une telle sensibilité avait déjà été mise en exergue pour les processus de gémination (*cf.* les expériences exposées dans le chapitre 3 de Cassar & Treiman, 1997 ; Pacton, Fayol, & Perruchet, 1999 ; Pacton, Perruchet, Fayol, & Cleeremans, 2001). Les résultats obtenus à l'expérience 2 indiquent que dès le niveau de lecture CP, les enfants témoignent de connaissances non négligeables sur la fréquence de certains bigrammes de la langue. Très tôt, mais pas avant le début de l'apprentissage de la lecture, les enfants sont capables d'intérioriser des séquences de lettres en fonction de leur fréquence d'apparition dans une langue donnée. Le développement de cette sensibilité se poursuit avec l'expertise en lecture. De plus, la position du bigramme dans l'item n'est pas anodine chez les enfants : la sensibilité à la fréquence bigrammique est plus importante si le bigramme est en position initiale. Comme le constatait déjà Segui (1984), le traitement de l'attaque est plus facile à effectuer que celui de la rime chez les apprenants. A mesure que l'expertise augmente, ce traitement différentiel s'atténue. En revanche, l'émergence précoce d'une sensibilité à la structure orthographique ne s'accompagne pas d'un développement concomitant des connaissances explicites sauf chez les lecteurs experts. Ce résultat inattendu incite à revoir la construction de l'épreuve orthographique explicite. Il

serait bon notamment de multiplier le nombre d'items expérimentaux et de les exploiter dans des mots et non de manière isolée. En effet, la nature de la tâche explicite et la consigne impliquaient de la part du participant une réflexion à la fois sur le bigramme pris isolément et sur le bigramme inséré dans un mot potentiel : cette double réflexion a pu déstabiliser les enfants, notamment les plus jeunes. Un autre résultat surprenant est l'absence d'effet de la fréquence bigrammique dans les procédures de reconnaissance de mots dans la mesure où les enfants apparaissaient sensibles à cette fréquence dans le test implicite. Seuls les adultes semblent tenir compte de la structure orthographique pour lire mais seulement pour les items à bigrammes en position initiale. Une explication à ce résultat pourrait être qu'au même titre que les enfants peuvent être amenés à utiliser leurs connaissances orthographiques pour lire des mots irréguliers parallèlement aux connaissances sur la consistance graphophonologique, la lecture des items proposés dans l'épreuve orthographique pourrait être facilitée par les connaissances graphophonologiques. Malgré les données des épreuves explicite et de lecture à haute voix qui vont à l'encontre des hypothèses formulées, les résultats obtenus pour l'orthographe restent encourageants. Ils montrent en particulier que les connaissances orthographiques émergent beaucoup plus tôt que ne le prédisent la plupart des modèles d'apprentissage de la lecture. Ces résultats se rapprochent plutôt des modèles de Gombert *et al.* (1997) et de Colé *et al.* (1999) qui défendent l'idée d'une interaction entre unités orthographiques et unités phonologiques.

Enfin, cette recherche a permis de confirmer l'existence d'une sensibilité à la structure morphologique et la mise en place de connaissances explicites sur la base de celles implicites mais seulement à partir du niveau CE1. A l'écrit, l'installation de telles connaissances serait donc plus tardive qu'à l'oral (voir les résultats obtenus en modalité orale par Colé *et al.*, 2003 ; Marec-Breton, 2003). De plus, les résultats obtenus confortent l'idée selon laquelle l'information morphologique contenue dans les mots morphologiquement complexes est utilisée par le lecteur lors de leur identification. Plusieurs travaux antérieurs ont déjà mis clairement en évidence que l'information morphologique était utilisée par le lecteur lors de la reconnaissance de mots isolés (e.g., Beauvillain, 1996 ; Boudelaa & Marslen-Wilson, 2001 ; Colé *et al.*, 2003 ; Colé *et al.*, 1997 ; Drews & Zwistserlood, 1995 ; Deutsch, Frost, Pollatsek, & Rayner, 2000 ; Feldman, 1991 ; Grainger *et al.*, 1991 ; Holmes & O'Reagan, 1992 ; Marec-Breton, 2003 ; Marec-Breton *et al.*, 2005 ; Meunier & Segui, 1999 ; Marslen-Wilson, Komisarjevski-Tyler, Waksler, & Older, 1994 ; Taft & Forster, 1975). Dans la présente étude, l'effet de la structure morphologique apparaît dès le CP au niveau de la précision à lire les mots et pseudo-mots de type préfixé et à partir du niveau CE1 pour la vitesse de traitement de ces mêmes items. La morphologie constitue

donc une information non négligeable dans l'apprentissage de la lecture. Au tout début de l'apprentissage de la lecture (niveau CP), elle contribue de façon indépendante à la lecture (*cf.* Analyses de régressions réalisées dans le chapitre 8). Cependant, il semble important de nuancer ce résultat en soulignant, comme l'a fait Marec-Breton (2003), que les connaissances morphologiques dépendent des connaissances graphophonologiques et du degré de transparence phonologique et sémantique des mots à lire. Au niveau de lecture CE1, ce sont d'ailleurs les connaissances graphophonologiques explicites qui expliquent le mieux les variations en lecture. De plus, il est également essentiel de préciser qu'étant donné le nombre important d'épreuves et d'expériences effectués dans cette thèse, certains contrôles n'ont pu être réalisés. Par exemple, Marec-Breton (2003) a vérifié si l'effet de la structure morphologique ne pouvait être attribué à l'identification d'un mot au sein d'un autre (*bouder* dans *débouder*). Pour cela, elle a comparé la lecture de pseudo-mots affixés (*dégarer*) à celle de pseudo-mots non affixés contenant les mêmes séquences orthographiques (*cagarer, démaner*). Un autre contrôle a été effectué par cet auteur : il s'agissait de vérifier l'influence de la recevabilité de la structure morphémique sur la lecture à l'aide d'items construits par l'association légale de deux morphèmes (*dégarer, bougeur*), d'items construits par l'association illégale de deux morphèmes (*dépapier, jambeur*), d'items construits par l'association d'une base et d'un pseudo-affixe (*cagarer, bougine*) et enfin d'items construits par l'association d'une pseudo-base et d'un affixe (*démaner, sanneur*). De tels contrôles n'ont pas pu être réalisés dans la présente étude mais les résultats de Marec-Breton (2003) indiquent que la facilitation de la lecture des items affixés serait de nature sémantique. De même, Colé *et al.* (2003) en utilisant plusieurs situations d'amorçage chez de jeunes enfants montrent qu'ils utilisent fortement le lien sémantique qui unit deux membres d'une même famille morphologique : l'analyse morphologique s'appuierait donc plutôt sur des aspects sémantiques et non orthographiques.

Ces différents résultats incitent à poursuivre l'étude des connaissances graphophonologiques, orthographiques et morphologiques implicites et explicites ainsi que leur impact sur les procédures d'identification de mots, notamment en accentuant les contrôles pour déterminer la nature exacte des traitements réalisés et en affinant davantage les épreuves afin de les rendre plus accessibles à des niveaux précoces d'apprentissage comme la GSM. De plus, l'étude des traitements graphophonologiques, orthographiques et morphologiques auprès de populations à risque constitue également un champ d'étude intéressant à développer. Ainsi,

Elbro et Arnbak (1996) ont constaté que les enfants dyslexiques phonologiques ne présentaient pas de déficit au niveau des connaissances morphologiques et élaboreraient « naturellement » une stratégie compensatoire de lecture de nature morphologique. Ce type de résultat a des implications au niveau pédagogique : chez des enfants présentant des difficultés phonologiques, une remédiation basée sur l'utilisation de la morphologie permettrait de compenser le déficit phonologique. Concernant l'apprentissage implicite, plusieurs études ont montré qu'une des spécificités de cet apprentissage est sa robustesse face à un grand nombre de variables, telles que l'âge, le niveau intellectuel, les troubles psychologiques (Reber, 1993 ; Reber, Walkanfeld, & Hernstadt, 1991). Toutefois, peu d'études se sont intéressées au maintien de l'apprentissage implicite chez des individus présentant divers niveaux intellectuels, et notamment un niveau intellectuel amoindri hormis une recherche récente de Detable et Vinter (2004). Or, ces chercheuses montrent que l'apprentissage implicite est relativement bien préservé chez des enfants déficients alors que les connaissances acquises explicitement sont significativement moins efficaces à court et à long terme, en comparaison aux performances de personnes montrant un développement ordinaire. Ce type de résultat offre des perspectives intéressantes aux professionnels de l'éducation adaptée travaillant auprès de jeunes souffrant de déficience intellectuelle afin de permettre à ces derniers d'acquérir certaines compétences orales et écrites de manière implicite.

REFERENCES BIBLIOGRAPHIQUES

- Adams, M.J. (1990). *Beginning to read*. Cambridge, MA: MIT Press.
- Alegria, J., Leybaert, J. & Mousty, P. (1994). Acquisition du langage écrit et troubles associés : évaluation, remédiation et théorie. In J. Grégoire & B. Piérart (Eds.), *Evaluer les troubles de la lecture : les nouveaux modèles théoriques et leurs implications diagnostiques* (pp. 105-123). Bruxelles: De Boeck.
- Alegria, J., & Morais, J. (1979). Le développement de l'habileté d'analyse phonétique conscience de la parole et l'apprentissage de la lecture. *Archives de Psychologie*, 183, 251-270.
- Alegria J., & Morais J. (1996). Métaphonologie, acquisition du langage écrit et troubles associés. In S.Carbonnel, P.Gillet, M.D. Martory, & S.Valdois (Eds), *Approche cognitive des troubles de la lecture et de l'écriture chez l'enfant et l'adulte* (pp.81-96). Marseille : Solal.
- Alvarez, C.J., Carreiras, M., & Taft, M. (2001). Syllables and morphemes: Contrasting frequency effects in Spanish. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27(2), 545-555.
- Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, R.C., & Freebody, P. (1985). Vocabulary knowledge. In H. Singer, & R.B. Ruddell (Eds.), *Theoretical models and processes of reading* (3rd Ed.) (pp. 343-371). Newark, DE: International Reading Association.
- Anglin, J.M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child development*, 58(238). Chicago: University of Chicago Press.
- Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace model for polysyllabic word reading. *Psychological Review*, 105, 678-723.
- Babin, J.P. (1996). Morphologie dérivationnelle et accès au lexique : le cas des verbes préfixés et pseudo-préfixés. *Canadian Journal of Experimental Psychology*, 50(4), 371-384.
- Babin, J.P. (1998). *Lexique mental et morphologie lexicale*. Berne: Peter Lang SA.
- Bastien, C. (1997). *Les connaissances de l'enfant à l'adulte*. Paris : Armand Colin.
- Bastien-Toniazzo, M. (1992). La représentation du mot écrit aux débuts de la lecture. *L'Année Psychologique*, 92(4), 489-509.
- Beauvillain, C. (1991). Le rôle de la structure morphologique dans la reconnaissance visuelle des mots. In R. Kolinski, J. Morais, & J. Segui (Eds.), *La reconnaissance des mots dans leurs différentes modalités sensorielles* (pp. 187-205). Paris : P.U.F.

- Beauvillain, C. (1996). The integration of morphological information and whole-word form information during eye fixations on prefixed and suffixed words. *Journal of Memory and Language*, 35, 801-820.
- Beauvillain, C., & Segui, J. (1992). Representation and processing of morphological information. In R. Frost, & L. Katz (Eds.), *Orthography, phonology and meaning: An overview* (pp. 361-376). Amsterdam: Elsevier.
- Bergman, M.W., Hudson, P.T.W., & Eling, P.A.T.M. (1988). How simple complex words can be: Morphological processing and word representations. *The Quarterly Journal of Experimental Psychology*, 40A(1), 41-72.
- Berko, J. (1958). The child's learning of English morphology. *Word*, 14, 150-177.
- Berry, D.C., & Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bertelson, P., & De Gelder, B. (1990). The emergence of phonological awareness: Comparative approaches. In G. Mattingly, & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception* (pp. 393-412). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bertram, R., & Hyona, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eyes movements when reading short and long Finnish compounds. *Journal of Memory and Language*, 48, 615-634.
- Bertram, R., Schreuder, R., & Baayen, R.H. (2000). The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 489-511.
- Bonin, P., Pacton, S., & Fayol, M. (2001). La production verbale écrite: évidences en faveur d'une (relative) autonomie de l'écrit. *Psychologie Française*, 46(1), 77-88.
- Bosse, M.L., Valdois, S., & Tainturier, M.J. (2003). Analogy without priming in early spelling development. *Reading and Writing: An Interdisciplinary Journal*, 16(7), 693-716.
- Boudelaa, S., & Marlsen-Wilson, W.D. (2001). Morphological units in Arabic mental lexicon. *Cognition*, 81(1), 65-92.
- Bowerman, M. (1982). Reorganizational processes in lexical and syntactic development. In E. Wanner, & L.R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 319-346). Cambridge: Cambridge University Press.
- Bowey, J.A., Cain, M., & Ryan, S. (1992). Reading-level design study of phonological skills underlying fourth-grade children word reading difficulties. *Child Development*, 63, 999-1011.
- Brady, S.A., & Shankweiler, D.P. (1991). *Phonological processes in literacy: A tribute to Isabelle Liberman*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brown, R. (1973). *A first language*. Cambridge, MA: M.I.T Press.

- Bruck, M., Genesee, F., & Caravolas, M. (1997). A cross linguistic study of early literacy acquisition. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 145-162). Mahwah, NJ: Lawrence Erlbaum Associates.
- Bruck, M., & Treiman, R. (1992). Learning to pronounce words: The limitations of analogies. *Reading Research Quarterly*, 27, 375-388.
- Brun, I., Faure, J., Ecalle, J., & Magnan, A. (2000). *The development of analogical processes in learning to read in French*. Paper presented at 10th Meetings of Society for Text and Discourse. Université de Lyon 2, 19-21 juillet.
- Bryant, P. (1993). Conscience phonologique et apprentissage de la lecture. In J.P. Jaffré, L. Sprenger-Charolles & M. Fayol (Eds.), *Les actes de la Villette* (pp. 176-192). Paris : Nathan.
- Bryant, P., & Bradley, L. (1985). *Children's reading problems*. Oxford: Blackwell.
- Burani, C., & Laudanna, A. (1992). Units of representation for derived words in the lexicon. In R. Frotz, & L. Katz (Eds.), *Orthography, phonology and meaning: An overview* (pp. 361-376). Amsterdam: Elsevier.
- Burani, C., Marcolini, S., & Stella, G. (2002). How early does morpholexical reading develop in readers of a shallow orthography? *Brain and Language*, 81, 568-586.
- Campbell, R. (1985). When children write nonwords to dictation. *Journal of Experimental Child Psychology*, 57, 133-151.
- Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, 28, 297-332.
- Caramazza, A., Miceli, G., Silveri, M., & Laudanna, A. (1985). Reaching mechanisms and the organisation of the mental lexicon: Evidence from acquired dyslexia. *Cognitive Neuropsychology*, 2, 81-114.
- Carlisle, J.F. (1988). Knowledge of derivational morphology and spelling ability in fourth, sixth and eight graders. *Applied Psycholinguistics*, 9(3), 247-266.
- Carlisle, J.F. (1995). Morphological awareness and early reading achievement. In L.B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 189-209). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Carlisle, J.F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing: An Interdisciplinary Journal*, 12, 169-190.
- Carlisle, J.F., & Nomanbhoy, D.M. (1993). Phonological and morphological awareness in first graders. *Applied Psycholinguistics*, 14, 177-195.

- Casalis, S., & Louis-Alexandre, M.F. (2000). Morphological analysis, phonological analysis and learning to read French: A longitudinal study. *Reading & Writing: An Interdisciplinary Journal*, 12, 303-335.
- Cassar, M., & Treiman, R. (1997). The beginning of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology*, 89, 631-644.
- Catach, N. (1980). *L'orthographe française. Traité théorique et pratique*. Paris : Nathan.
- Catach, N., Gruez, C., & Duprez, D. (1986). *L'orthographe française*. Paris : Nathan.
- Chialant, D., & Caramazza, A. (1995). Where is morphology and how is it processed. In L.B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 55-78). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Clark, E.V., & Berman, R. (1984). Structure and use in the acquisition of word formation. *Language*, 60, 542-591.
- Clark, E.V., & Berman, R. (1987). Types of linguistic knowledge: Interpreting and producing compound nouns. *Journal of Child Language*, 14, 547-567.
- Clark, E., & Cohen, S. (1984). Productivity and memory for newly formed words. *Journal of Child Language*, 11, 611-625.
- Clark, E.V., & Hecht, B.F. (1982). Learning to coin agent and instrument nouns. *Cognition*, 12, 1-24.
- Clay, M.M. (1982). The development of morphological rules in children of differing language backgrounds. In *Observing young readers: Selected papers* (pp. 103-112). Portsmouth, NH: Heineman.
- Colé, P. (1999). *Le rôle de la morphologie dans la reconnaissance des mots écrits chez le lecteur expert et l'apprenti lecteur*. Unpublished Habilitation à diriger des recherches, Université René Descartes - Paris 5, Paris.
- Colé, P., Beauvillain, C., Pavard, B., & Sehui, J. (1986). Organisation morphologique et accès au lexique. *L'Année Psychologique*, 86, 349-365.
- Colé, P. & Fayol, M. (2000). Reconnaissance de mots écrits et apprentissage de la lecture : rôle des connaissances morphologiques. In M. Kail, & M. Fayol (Eds.), *L'acquisition du langage : le langage en développement au-delà de trois ans* (pp. 151-181). Paris : P.U.F.
- Colé, P., Magnan, A., & Grainger, J. (1999). Syllable-sized units in visual word recognition: Evidence from skilled and beginning readers. *Applied Psycholinguistics*, 20, 507-532.
- Colé, P., Marec-Breton, N., Royer, C., & Gombert, J.E. (2003). Morphologie des mots et apprentissage de la lecture. *Rééducation Orthophonique*, 213, 57-76.
- Colé, P., Segui, J., & Taft, M. (1997). Words and morphemes as units of lexical access. *Journal of Memory and Language*, 37, 312-330.

- Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151-216). London: Academic Press.
- Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance 6* (pp.535-555). London: Academic Press.
- Content, A. (1985). Le développement de l'habileté d'analyse phonétique de la parole. *L'Année Psychologique*, 85, 73-99.
- Content, A. (1990). L'acquisition de la lecture : approche cognitive. *Cahiers du DLSL*, 9, 15-41.
- Content, A. (1993). Le rôle de la médiation phonologique dans l'apprentissage de la lecture. In J.P. Jaffré, L. Sprenger-Charolles, & M. Fayol (Eds.), *Lecture-écriture : acquisition* (pp.80-96). Paris : Nathan.
- Content, A., Mousty, P. & Radeau, M. (1990). BRULEX. Une base de données lexicales informatisée pour le français parlé et écrit. *L'Année Psychologique*, 90, 551-566.
- Content, A., & Radeau, M. (1988). Données statistiques sur la structure orthographique du français. *Cahiers de Psychologie Cognitive*, 8(4), 399-404.
- Corbin, D. (1985). Les bases non autonomes en Français ou comment intégrer l'exception dans un modèle lexical. *Langue Française*, 66, 54-76.
- Corbin, D. (1987). *Morphologie dérivationnelle et structuration du lexique*. Lille : Niemeyer.
- Corbin, D. (1999). Pour une théorie sémantique de la catégorisation affixale. *Faits de Langues*, 14, 65-77.
- De Villiers, J.G., & de Villiers, P.A. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research*, 2(3), 267-278.
- Deavers, R.P., & Brown, G.D.A. (1997). Rules versus analogies in children's spelling: Evidence for task dependence. *Reading and Writing: An Interdisciplinary Journal*, 9, 339-361.
- Detable, C., & Vinter, A. (2004). Le maintien dans le temps des effets d'un apprentissage implicite chez des enfants et adolescents avec un retard mental ? *L'Année Psychologique*, 104, 751-770.
- Deutsch, A., Forst, R., & Forster, K.I. (1998). Verbs and nouns are organized and accessed differently in the mental lexicon: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1238-1255.
- Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2000). Early morphological effects in word recognition in Hebrew: Evidence from parafoveal preview benefit. *Language and Cognitive Processes*, 15(4-5), 487-506.
- Drews, E., & Zwistserlood, P. (1995). Morphological and orthographic similarity in visual word recognition. *Journal of Experimental Psychology: Human, Perception and Performance*, 21, 1098-1116.

- Ducard, D., Honvault, R., & Jaffré, J.P. (1995). *L'orthographe en trois dimensions*. Paris : Nathan.
- Duncan, L.G., Seymour, P.H.K., & Hill, S. (2000). A small-to-large unit progression in metaphonological awareness and reading? *The Quarterly Journal of Experimental Psychology*, 53, 1081-1104.
- Dunn, L.M. & Thiéroult-Whalen, C.M. (1993). *Echelle de vocabulaire en images Peabody*. Toronto: Psycan.
- Ecalte, J. (2000). Lire et écrire des mots au CP : aspects prédictifs. *Revue de Psychologie de l'Education*, 1, 42-58.
- Ecalte, J., & Magnan, A. (2000). *La mise en œuvre précoce ou tardive des analogies en lecture de mots : vrai ou faux débat ?* Communication orale, 27^e Symposium de l'Association de Psychologie Scientifique de Langue Française APSLF, Université de Nantes, 14-16 Septembre.
- Ecalte, J., & Magnan, A. (2002a). *L'apprentissage de la lecture. Fonctionnement et développement cognitifs*. Paris : Armand Colin.
- Ecalte, J., & Magnan, A. (2002b). The development of epiphonological and metaphonological processing at the start of reading: A longitudinal study. *European Journal of Psychology of Education*, 17(1), 47-62.
- 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. (1987). Learning to read and to spell words. *Journal of Reading Behaviour*, 19, 5-31.
- Ehri, L.C. (1989). Apprendre à lire et à écrire des mots. In L. Rieben, & C.A. Perfetti (Eds.), *L'Apprenti Lecteur* (pp. 108- 128). Neuchâtel: Delachaux & Niestlé.
- 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). Apprendre à lire et apprendre à orthographier, c'est la même chose ou pratiquement la même chose. In L. Rieben, & C.A. Perfetti (Eds.), *L'Apprenti Lecteur*. (pp. 231-267). Neuchâtel: Delachaux et Niestlé.
- Elbro, C., & Arnbak, E. (1996). The role of morpheme recognition and morphological awareness in dyslexia. *Annals of Dyslexia*, 46, 209-240.
- Elbro, C., Borstrom, I., & Peterson, D.K. (1998). Predicting dyslexia from kindergarten: The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly*, 33(1), 36-60.

- Ellis, N.C. (1994). Longitudinal studies of spelling development. In G.D.A. Brown, & N.C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 155-178). Chichester: John Wiley & Sons Ltd.
- Ellis, N.C. (1997). Acquisition interactive de la lecture et de l'orthographe : étapes, stratégies et échanges de connaissances. In L. Rieben, M. Fayol, & C.A. Perfetti (Eds.), *Des orthographes et leur acquisition* (pp. 267-292). Neuchâtel: Delachaux et Niestlé.
- Englert, C.S., Hiebert, E.H., & Stewart, S.R. (1985). Spelling unfamiliar words by an analogy strategy. *Journal of Special Education, 19*, 291-306.
- Fayol, M., & Gombert, J.E. (1999). Apprentissage de la lecture et de l'écriture. In J.A. Rondal, & E. Esperet (Eds.), *Manuel de Psychologie de l'Enfant* (pp. 565-594). Bruxelles: Mardaga.
- Fayol, M., Lété, B. & Gabriel, M.A. (1996). Du développement de la correspondance (un) phonème - (plusieurs) graphèmes chez les enfants de 6 à 7 ans. *LIDIL, 13*, 67-85.
- Fayol, M., Pacton, S., & Totereau, C. (1998). Règles et régularités dans l'apprentissage et la mise en œuvre de l'orthographe du français. *Le Français Aujourd'hui, 122*, 54-61.
- Feldman, L.B. (1991). The contribution of morphology to word recognition. *Psychological Research, 53*, 33-41.
- Feldman, L.B., & Andjelkovic, D. (1992). Morphological analysis in word recognition. In R. Frost, & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 343-360). Amsterdam: Elsevier.
- Feldman, L.B., Rueckl, J., DiLiberto, K., Pastizzo, M., Vellutino, F. (2002). Morphological analysis by child readers as revealed by the fragment completion task. *Psychonomic Bulletin, 9*, 529-535.
- Fijalkow, E., & Fijalkow, J. (1994). Enseigner à lire-écrire au CP : état des lieux. *Revue Française de Pédagogie, 107*, 63-79.
- Fisher, F.W., Shankweiler, D., & Liberman, I.Y. (1985). Spelling proficiency and sensitivity to word structure. *Journal of Memory and Language, 24*, 282-295.
- Fowler, A.F., & Liberman, I.Y. (1995). The role of phonology and orthography in morphological awareness. In L.B. Feldman (Ed.), *Morphological aspects of language processing* (pp.157-188). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In K.E. Patterson, J.C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Cognitive and neuropsychological studies of phonological reading* (pp. 301-330): London: Lawrence Erlbaum Associates.
- Frith, U., Wimmer, W., & Landerl, K. (1998). Differences in phonological recoding in German- and English- speaking children. *Scientific Studies of Reading, 2*, 31-54.
- Frost, R., Deutsch, A., & Forster, K.I. (2000). Decomposing morphologically complex words in non linearmorphology. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*, 751-765.

- Frost, R., Forster, K.I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23(4), 1-28.
- Gardes-Tamine, J. (1988). *La Grammaire*. Paris : Armand Colin.
- Gaux, C. (2002). *Connaissances implicites et explicites des unités morpho-phonologiques et apprentissage de la lecture chez les jeunes enfants*. Communication orale au Séminaire du Labecd (Laboratoire de Psychologie, Education, Cognition et Développement), Université de Nantes, 7 mars.
- Gentry, J.R. (1982). An analysis of developmental spelling in GNYS AT WRK. *The Reading Teacher*, 36, 192-200.
- Gillon, G., & Dodd, B.J. (1994). A prospective study of the relationship between phonological, semantic and syntactic skills and specific reading disability. *Reading and Writing: An interdisciplinary Journal*, 6, 321-345.
- Gineste, M.D. (1997). *Analogie et cognition. Etude expérimentale et simulation informatique*. Paris : P.U.F.
- Giraud, H., & Grainger, J. (2001). Priming complex words: Evidence for supralexicale representations of morphology. *Psychonomic Bulletin and Review*, 8(1), 127-131.
- Gleitman, L.R., & Rozin, P. (1977). The structure and acquisition of reading: Relation between orthographies and the structure of languages. In A.S. Reber, & D.L. Scarborough (Eds.), *In towards a psychology of reading* (pp. 1-53). New York: Wiley.
- Gombert, J.E. (1990). *Le développement métalinguistique*. Paris : P.U.F.
- Gombert, J. E. (1992). Activités de lecture et activités associées. In M. Fayol, J.E. Gombert, P. Lécocq, L. Sprenger-Charolles, & D. Zagar (Eds.), *Psychologie cognitive de la lecture* (pp. 107-131). Paris : P.U.F.
- Gombert, J.E. (2000). La construction des connaissances phonologiques chez l'enfant. *Revue Parole*, 9/10, 89-100.
- Gombert, J.E. (2002). L'apprentissage de la lecture : processus et stratégie. In A. Florin & J. Morais (Eds.), *La maîtrise du langage* (pp. 79-94). Rennes : Presses Universitaires de Rennes.
- Gombert, J.E. (2003). Implicit and explicit learning to read: Implication as for subtypes of Dyslexia. *Current Psychology Letters. Behaviour, Brain and Cognition*, 10(1).
- Gombert, J. E., Bryant, P., & Warrick, N. (1997). Children's use of analogy in learning to read and to spell. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: research, theory, and practice across languages* (pp. 221-235). Mahwah, NJ: Lawrence Erlbaum Associates.

- Gombert, J.E., & Colé, P. (2000). Activités métalinguistiques, lecture et illettrisme. In M. Kail, & M. Fayol (Eds.), *L'acquisition du langage : le langage en développement au-delà de trois ans* (pp. 117-150). Paris : PUF.
- Gombert, J.E., Colé, P., Valdois, S., Goigoux, R., Mousty, P., & Fayol, M. (2002). *Enseigner la lecture au cycle 2*. Paris : Nathan.
- Gombert, J.E., & Peereman, R. (2001). Training children with artificial alphabet. *Psychology*, 8(3), 338-357.
- Gordon, P. (1989). Levels of affixation in the acquisition of English morphology. *Journal of Memory and Language*, 28, 519-530.
- Goswami, U. (1986). Children's use of analogies in learning to read: A developmental study. *Journal of Experimental Child Psychology*, 42, 72-83.
- Goswami, U. (1988a). Children's use of analogies in learning to spell. *British Journal of Experimental Psychology*, 6, 21-33.
- Goswami, U. (1988b). Orthographic analogies and reading development. *The Quarterly Journal of Experimental Psychology*, 40A(2), 239-268.
- Goswami, U. (1990). Phonological priming and orthographic analogies in reading. *Journal of Experimental Child Psychology*, 49, 323-340.
- Goswami, U. (1995). Phonological development and reading by analogy: What is analogy, and what is it not? *Journal of Research in Reading*, 18(2), 139-145.
- Goswami, U. (1998). The role of analogies in the development of word recognition. In J.L. Metsala & L.C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 41-64). London: Lawrence Erlbaum Associates.
- Goswami, U. (1999). Orthographic analogies and phonological priming: A comment on Bowey, Vaughan and Hansen (1998). *Journal of Experimental Child Psychology*, 68, 108-133.
- Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Goswami, U., & East, M. (2000). Rhyme and analogy in beginning reading: Conceptual and methodological issues. *Applied Psycholinguistics*, 21(1), 63-93.
- Goswami, U., Gombert, J.E., & Fraca de Barrera, L.F. (1998). Children's orthographic representations and linguistic transparency: Nonsense word reading in English, French and Spanish. *Applied Psycholinguistics*, 19, 19-52.
- Goswami, U., Porpodas, C.D., & Weelwright, S. (1997). Children's orthographic representations in English and Greek. *European Journal of Psychology of Education*, 3, 273-292.
- Gough, P.B., & Hillinger, M.L. (1980). Learning to read: An unnatural act. *Bulletin of the Orton Society*, 30, 180-196.

- Grainger, J., Colé, P., & Segui, J. (1991). Masked morphological priming in visual word recognition. *Journal of Memory and Language*, *30*, 370-384.
- Hanson, V.L., Shankweiler, D., & Fisher, F.W. (1983). Determinants of spelling ability in deaf and hearing adults: Access to linguistic structure. *Cognition*, *14*, 323-344.
- Hargrave, A.C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, *15*(1), 75-90.
- Harm, M.W., & Seidenberg, M.S. (1999). Phonology, reading acquisition and dyslexia: Insights from connectionist models. *Psychological Review*, *106*, 491-528.
- Henderson, L. (1985). *Teaching spelling*. Boston: Houghton Mifflin.
- Henderson, A.J., & Shores, R.E. (1982). How learning disabled student's failure to attend to suffixes affects their oral reading performance. *Journal of Learning Disabilities*, *15*, 178-182.
- Henderson, L., Wallis, J., & Knight, D. (1984). Morphemic structure and lexical access. In H. Bouma, & D.G. Bouwhuis (Eds.), *Attention and performance 10: Control of language processes* (pp. 211-226). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Holmes, V.M., & O'Regan, J.K. (1992). Reading derivationally affixed French words. *Language & Cognitive Processes*, *7*, 163-192.
- Hulme, C., Hatcher, P.J., Nation, K., Brown, A., Adams, J., & Stuart, G. (2002). Phoneme awareness is a better predictor of early reading skill than onset-rime awareness. *Journal of Experimental Child Psychology*, *82*, 2-28.
- Hulme, C., Mutter, V.V., & Snowling, M. (1998). Segmentation does predict early progress in learning to read better than rhyme: A reply to Bryant. *Journal of Experimental Child Psychology*, *71*, 39-44.
- Hyona, J., Vainio, S., & Laine, M. (2002). A morphological effect obtains for isolated words but not for words in sentence context. *European Journal of Cognitive Psychology*, *14*, 417-433.
- Jacobs, A.M., Rey, A., Ziegler, J.C., & Grainger, J. (1998). MROM-p: An interactive activation multiple read-out model of phonological coding in visual word recognition. In J. Grainger, & A.M. Jacobs (Eds.), *Localist connectionist approaches to human recognition* (pp. 147-188). Mahawah, NJ: Lawrence Erlbaum Associates.
- Jansen, M.E. (1995). *Zur frage der allgemein Gültigkeit "englischer" Modelle des Lesenlernens [On the generality of "English" models of learning to read]*. Unpublished master's thesis, University of Salzburg, Salzburg, Austria.
- Jimenez, J.E. (1997). A reading-level match study of phonemic processes underlying reading disabilities in a transparent orthography. *Reading and Writing: An Interdisciplinary Journal*, *9*, 23-40.

- Jorm, A.F., Share, D.L., MacLean, R., & Mathews, R.G. (1984). Phonological recoding skills and learning to read – a longitudinal study –. *Applied Psycholinguistics*, 5(3), 201-207.
- Juel, C., Griffith, P.L., & Gough, P.B. (1986). Acquisition of literacy – a longitudinal study of children in 1st grade and 2nd grade. *Journal of Educational Psychology*, 78(4), 243-255.
- Khosmi, A. (1999). *LMC-R*. Paris : Les Editions du Centre de Psychologie Appliquée.
- Lamarche, R.M. (1986). La maîtrise de l'orthographe: essai de vérification de quelques hypothèses. *Revue Française de Pédagogie*, 74, 29-36.
- Landerl, K. (2000). Influences of orthographic consistency and reading instruction on the development of nonword reading skills. *European Journal of Psychology and Education*, 15(3), 239-257.
- Landerl, K., Wimmer, W., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: a German-English comparison. *Cognition*, 63, 315-334.
- Lanoë, C. (2000). La situation de lecture partagée : une routine pour l'acquisition du langage, de la lecture et de l'écriture. *Revue de Psychologie de l'Éducation*, 2, 70-93.
- Laudanna, A., Badecker, W., & Caramazza, A. (1992). Processing inflectional and derivational morphology. *Journal of Memory and Language*, 31, 333-348.
- Laudanna, A., & Burani, A. (1985). Address mechanisms to decomposed lexical entries. *Linguistics*, 23, 775-792.
- Laudanna, A., & Burani, C. (1995). Distributional properties of derivational affixes: Implications for processing. In L.B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 345-364). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Laudanna, A., Burani, C., & Cermele, A. (1994). Prefixes as processing units. *Language and Cognitive processes*, 9(3), 295-316.
- Laudanna, A., Cermele, A., & Caramazza, A. (1997). Morpho-lexical representation in naming. *Language and cognitive processes*, 12(1), 49-66.
- Laxon, V., Rickart, M., & Coltheart, V. (1992). Children read affixed words and non-words. *British Journal of Psychology*, 83, 407-423.
- Lecocq, P. (1992). *La lecture : processus, apprentissage, troubles*. Lille : Presses Universitaires de Lille.
- Lecocq, P. (1996). *E.C.O.S.S.E : une épreuve de compréhension syntaxico-sémantique*. Lille : Presses Universitaires de Lille.
- Lecocq, P., Casalis, S., Leuwers, C., & Watteau, N. (1996). *Apprentissage de la lecture et compréhension d'énoncés*. Lille : Presses Universitaires de Lille.
- Lefavrais, P. (1967). *Test de l'Allouette : Manuel*. Paris : Les Editions du Centre de Psychologie Appliquée.

- Lennox, C. & Siegel, L.S. (1994). The role of phonological and orthographic processes in learning to spell. In G.D.A. Brown, & N.C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 93-109). Chichester: John Wiley & Sons Ltd.
- Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX: A grade-level lexical database from French elementary school-readers. *Behavior Research, Methods, Instruments, & Computers*, 36, 156-166.
- Lewicki, P., Hill, T., & Czyzewska, M. (1992). Nonconscious acquisition of information. *American Psychologist*, 47, 796-801.
- Lieberman, Y. (1973). Segmentation of the spoken word and reading acquisition. *Bulletin of the Orton Society*, 23, 65-77.
- Lieberman, Y., Shankweiler, D., Fisher, W.F., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, 18, 201-212.
- Longtin, C., Segui, J., & Hallé, P.A. (2003). Morphological priming without morphological relationships. *Language and Cognitive Processes*, 18, 313-334.
- Mahony, D., Singson, M., & Mann, V. (2000). Reading ability and sensitivity to morphological relations. *Reading and Writing: An Interdisciplinary Journal*, 12(3-4), 192-218.
- Manelis, L., & Tharp, D.A. (1977). The processing of affixed words. *Memory and Cognition*, 5(6), 690-695.
- Manis, F.R., Custodio, R., & Szezulski, P.A. (1993). Development of phonologic and orthographic skills: A 2-year longitudinal study of dyslexic children. *Journal of experimental Child Psychology*, 56, 64-86.
- Mann, V.A. (1986). Phonological awareness: The role of reading experience. *Cognition*, 24, 65-92.
- Mann, V.A., & Wimmer, H. (2002). Phoneme awareness and pathways to literacy: a comparison of German and American children. *Reading and Writing: An Interdisciplinary Journal*, 15, 653-682.
- Marec-Breton, N. (2003). *Les traitements morphologiques dans l'apprentissage de la lecture*. Unpublished Doctorat de Psychologie, Université de Rennes 2, Rennes.
- Marec-Breton, N., Gombert, J.E., & Colé, P. (2005). Traitements morphologiques lors de la reconnaissance de mots écrits chez des apprentis lecteurs. *L'Année Psychologique*, 105, 9-45.
- Marouby-Terriou, G., & Denhière, G. (2002). Identifier l'écrit : influence des connaissances infra-lexicales. *Enfance*, 4, 381-407.
- 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.

- Marslen-Wilson, W.D., Komisarjevski-Tyler, L.K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, 101, 3-33.
- Martinet, C. (2001). *Le rôle de la phonologie dans l'acquisition des connaissances lexicales orthographiques*. Unpublished Doctorat de Psychologie, Université de Grenoble 2, Grenoble.
- 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, B11-B22.
- Mattingly, I.G. (1972). Reading, the linguistic process and linguistic awareness. In J.F. Kavanagh, & I.G. Mattingly (Eds.), *Language by ear and by eye: The relationships between speech and reading* (pp. 133-147). Cambridge, MA: MIT Press.
- Meulemans, T. (1998). Apprentissage implicite, mémoire implicite et développement. *Psychologie Française*, 43(1), 27-37.
- Meunier, F., & Segui, J. (1999). Frequency effects in auditory word recognition: The case of suffixed words. *Journal of Memory and Language*, 41, 327-344.
- Morais, J. (1991). Constraints on the development of phonemic awareness. In S.A. Brady, & D.P. Shankweiler (Eds.), *Phonological processes in literacy. A tribute to Isabelle Y. Liberman* (pp. 5-27). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Morais, J., Alegria, J., & Content, A. (1987). The relationship between segmental analysis and alphabetical literacy: An interactive view. *Cahiers de Psychologie Cognitive*, 7, 415-438.
- Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, 7, 323-331.
- Morton, J. (1980). The logogen model and orthographic structure. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 117-133). London: Academic Press.
- Mousty, P., & Alegria, J. (1996). L'acquisition de l'orthographe et ses troubles. In S. Carbonnel, P. Gillet, M.D. Martory, & S. Valdois (Eds.), *Approche cognitive des troubles de la lecture et de l'écriture chez l'enfant et l'adulte* (pp. 165-179). Marseille : Solal.
- Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1998). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology*, 71, 3-27.
- Nagy, W.E., & Anderson, R.C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, 19, 304-330.
- 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.

- Nation, K., & Hulme, C. (1998). The role of analogy in early spelling development. In C. Hulme, & R.M. Joshi (Eds.), *Reading and spelling: Development and disorders* (pp. 433-445). Mahwah, NJ: Lawrence Erlbaum Associates.
- Nunes, T., Bryant, P., & Bindman, M. (1998). Orthographe et grammaire: *the necsed move*. In L. Rieben, & M. Fayol (Eds.), *Des orthographes et leur acquisition* (pp. 101-124). Neuchâtel : Delachaux et Niestlé.
- Oney, B., & Goldman, S.R. (1984). Decoding and comprehension skills in Turkish and English – effects of the regularity of grapheme phoneme correspondences. *Journal of Educational Psychology*, 76(4), 556-568.
- Ouzoulias, A. (2001). L'émergence de la conscience phonémique : apprentissage sensoriel ou développement conceptuel? In G. Chauveau (Ed.), *Comprendre l'enfant apprenti lecteur. Recherches actuelles sur la psychologie de l'écrit* (pp. 101-127). Paris : Retz.
- Pacton, S. (2000). *L'apprentissage implicite en dehors du laboratoire : le cas des régularités orthographiques*. Unpublished Doctorat de Psychologie, Université de Bourgogne, Dijon.
- Pacton, S., & Fayol, M. (2000). The impact of phonological cues on children's judgements of nonwords: The case of double letters. *Current Psychology Letters*, 1, 39-54.
- Pacton, S., & Fayol, M. (2003). How do French children use morphosyntactic information when they spell adverbs and present participles? *Scientific Studies of Reading*, 7(3), 273-287.
- Pacton, S., Fayol, M., & Perruchet, P. (1999). L'apprentissage de l'orthographe lexicale : le cas des régularités. *Langue Française*, 124, 23-39.
- Pacton, S., Fayol, M., & Perruchet, P. (2002). The acquisition of untaught orthographic regularities in French. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 121-137). Dordrecht: Kluwer.
- Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of European Psychology: General*, 130(3), 1-25.
- Pastizzo, M.J., & Feldman, L.B. (2002). Discrepancies between orthographic and unrelated baselines in masked priming undermine a decompositional account of morphological facilitation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28(1), 244-249.
- Pavard, B. (1983). Traitement perceptif des mots affixés : mise en évidence d'un contrôle cognitif. *L'Année Psychologique*, 83, 443-464.
- Peereman, R., & Content, A. (1997). Orthographic and phonological neighborhoods in naming: not all neighbors are equally influential in orthographic space. *Journal of Memory and Language*, 37(3), 382-410.
- Peereman, R., & Content, A. (1999). LEXOP: A lexical database providing Orthography-Phonology statistics for French monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, 31, 376-379.

- Perruchet, P., & Nicolas, S. (1998). L'apprentissage implicite : un débat théorique. *Psychologie Française, 43(1)*, 13-25.
- Perruchet, P., & Pacton, P. (2004). Qu'apportent à la pédagogie les travaux de laboratoire sur l'apprentissage implicite ? *L'Année Psychologique, 104*, 121-146.
- Perruchet, P., Vinter, A., & Gallego, J. (1997). Implicit learning shapes new conscious percepts and representations. *Psychonomic Bulletin and Review, 4*, 43-48.
- Perruchet, P., & Vinter, A. (1998). Learning and development : The implicit knowledge assumption reconsidered. In M.I. Stadler, & P.A. Frensch (Eds.), *Handbook of implicit learning* (pp. 495-531). Thousand Oaks: Sage publications.
- Perry, C., Ziegler, J.C., & Coltheart, M. (2002). How predictable is spelling? An analysis of sound-spelling contingency in English. *The Quarterly Journal of Experimental Psychology, 55a*, 897-915.
- Pillon, A. (1998). The pseudoprefixation effect in visual word recognition: A true-neither nor orthographic-morphemic effect. *The Quarterly Journal of Experimental Psychology, 51A(1)*, 85-120.
- Plaut, D.C., & Gonnerman, L.M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to language processing. *Language and Cognitive Processes, 15*, 445-485.
- Plaut, D.C., McClelland, J.L., Seidenberg, M., & Patterson, K.E. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*, 56-115.
- Porpodas, C.D. (1999). Patterns of phonological and memory processing in beginning readers and spellers of Greek. *Journal of Learning Disabilities, 32(5)*, 406-416.
- Rack, J.P., Hulme, C., & Snowling, M. (1993). Learning to read: A theoretical synthesis. *Advances in Child Developmental Behavior, 24*, 99-132.
- Raven, J.C. (1976). *Coloured progressive matrices. Sets A, Ab, B*. Oxford: Oxford Psychologists' Press Ltd.
- Reber, A.S. (1976). Implicit learning of syntactic structure in synthetic languages. *Journal of Experimental Psychology Human Learning & Memory, 2*, 88-94.
- Reber, A.S. (1993). *Implicit learning and tacit knowledge*. Oxford: Oxford University Press.
- Reber, A.S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.
- Reber, A.S. & Allen, R. (1978). Analogic and abstraction strategies in synthetic grammar learning. *Cognition, 6*, 189-221.
- Reber, A.S. & Lewis, S. (1977). Implicit learning: An analysis of the form and structure of a body of tacit knowledge. *Cognition, 5*, 331-361.

- Reber, A.S., Walkanfeld, F.F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17, 888-896.
- Rey-Debove, J. (1984). *Le Robert Méthodique*. Paris : Dictionnaires Le Robert.
- Rittle-Johnson, B., & Siegler, R.S. (1999). Learning to spell: Variability, choice and change in children's strategy use. *Child development*, 70, 332-348.
- Robert, P. (1986). *Micro-Robert, dictionnaire du français primordial*. Paris : Dictionnaires Le Robert.
- Rubin, H. (1988). Morphological knowledge and early writing ability. *Language and Speech*, 31, 337-355.
- Rubin, G.S., Becker, C.A., & Freeman, R.H. (1979). Morphological structure and its effect on visual word recognition. *Journal of Verbal Learning and Verbal Behavior*, 18, 757-767.
- Sandra, D. (1994). The morphology of the mental lexicon: Internal word structure of viewed from a psycholinguistic perspective. *Language and Cognitive Processes*, 9, 227-269.
- Savigny, M. (1974). *Bat-Elem*. Issy-Les-Moulineaux : Editions de Psychologie Appliquée.
- Schriefers, H., Zwitserlood, P., & Roelofs, A. (1991). The identification of morphologically complex spoken words: Continuous processing or decomposition? *Journal of Memory and Language*, 30, 26-47.
- Seger, C.A. (1994). Implicit learning. *Psychological Bulletin*, 115, 163-196.
- Segui, J. (1984). The syllable: A basic perceptual unit in speech perception? In H. Bouma, & D.G. Bouwhuis (Eds.), *Attention and performance 10: Control of language processes* (pp. 165-181). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Segui, J., & Zubizarreta, M.L. (1985). Mental representation of morphologically complex words and lexical access. *Linguistics*, 23, 759-774
- Seidenberg, M.S., & Gonnerman, L.M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Science*, 4, 353-351.
- Seidenberg, M.S., & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523-568.
- Sénéchal, M. (2000a). Examen du lien entre la lecture de livres et le développement du vocabulaire chez l'enfant préscolaire. *Enfance*, 2, 169-186.
- Sénéchal, M. (2000b). Morphological effects in children's spelling of french words. *Canadian Journal of Experimental Psychology*, 54(2), 76-85.
- Seymour, P.H.K (1997). Foundation of orthographic development. In C.A. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell, research, theory and practice across languages* (pp. 319-337). Mahwah, NJ: Lawrence Erlbaum Associates.

- Seymour, P.H.K., Aro, M., & Erskine, J.M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143-174.
- Seymour, P.H.K., & Evans, H.M. (1993). The visual (orthographic) processor and developmental dyslexia. In D. Willows, R. Kruk, & E. Corcos (Eds.), *Visual processes in reading and reading disabilities* (pp. 347-376). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Shankweiler, D.P., Crain, S., Katz, L., Fowler, A., Liberman, A. Brady, S., Thornton, R., Lundquist, E., Dreyer, L., Fletcher, J.M., Stuebing, K.K., Shaywitz, S.E., & Shaywitz, B.A. (1995). Cognitive profiles of reading disabled children: Comparison of language skills in phonology, morphology and syntax. *Psychological Science*, *6*, 149-156.
- 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.
- Singson, M., Mahony, D., & Mann, V. (2000). The relation between reading ability and morphological skills: Evidence from derivational suffixes. *Reading and Writing: An Interdisciplinary Journal*, *12*, 219-252.
- Smith, P.T. (1995). Are morphemes really necessary? In L.B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 365-382). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Smith, P.T., & Sterling, C.M. (1982). Factors affecting the perceived morphemic structure of written words. *Journal of Verbal Learning and Verbal Behavior*, *21*, 704-721.
- 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* (pp. 111-128). Chichester: John Wiley & Sons Ltd.
- Sprenger-Charolles, L., & Casalis, S. (1996). Lire. Lecture/écriture : acquisition et troubles du développement. Paris : P.U.F.
- Spenger-Charolles, L., & Siegel, L.S. (1997). A longitudinal study of the effects of syllabic structure in the development of reading and spelling skills in French. *Applied Psycholinguistics*, *18*, 485-505.
- Sprenger-Charolles, L., Siegel, L., Béchenec, 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). Phonological mediation and orthographic factors in reading and spelling. *Journal of Experimental Child Psychology*, *68*, 134-155.

- Stolz, J.A., & Besner, D. (1998). Levels of representation in visual word recognition. A dissociation between morphological and semantic processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 24, 1642-1655.
- Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory and Cognition*, 7(4), 263-272.
- Taft, M. (1981). Prefix stripping revisited. *Journal of Verbal Learning and Verbal Behavior*, 20, 289-297.
- Taft, M. (1984). Evidence for an abstract lexical representation of word structure. *Memory and Cognition*, 12(3), 264-269.
- Taft, M. (1985). The decoding of words in lexical access: A review of morphographic approach. In D. Besner, T.G. Wallers, & G.E. Mackinnon (Eds.), *Reading research: Advances in theory and practice Vol. 5* (pp. 83-123). London: Academic Press.
- Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes*, 9(3), 271-294.
- Taft, M., & Forster, K.I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14, 638-647.
- Taft, M., & Forster, K.I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, 15, 607-620.
- Taft, M., Hambly, G., & Kinoshita, S. (1986). Visual and auditory recognition of prefixed words. *The Quarterly Journal of Experimental Psychology*, 38A, 351-366.
- Templeton, S., & Scarborough-Franks, L. (1985). The spelling's the thing: Knowledge of derivational morphology in orthography and phonology among older students. *Applied Psycholinguistics*, 6, 371-390.
- Thorstad, G. (1991). The effect of orthography on the acquisition of literacy skills. *British Journal of Psychology*, 82, 527-537.
- Totureau, C., Thévenin, M.G., & Fayol, M. (1997). The development of the understanding of number morphology in written French. In C.A. Perfetti, M. Fayol, & L. Rieben (Eds.), *Learning to spell* (pp. 97-114). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Treiman, R. (1989). Le rôle des unités intrasyllabiques dans l'apprentissage de la lecture. In L. Rieben, & C. Perfetti (Eds.), *L'Apprenti Lecteur* (pp. 241-259). Neuchâtel: Delachaux & Niestlé.
- Treiman, R. (1993). *Beginning to spell: A study of first-grade children*. New York: Oxford University Press.
- Treiman, R. (1994). Sources of information used by beginning spellers. In G.D.A. Brown & N.C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 72-91). Chichester: John Wiley & Sons Ltd.

- Treiman, R., & Cassar, M. (1997). L'acquisition de l'orthographe en anglais. In L. Rieben, M. Fayol, & C.A. Perfetti (Eds.), *Des orthographes et leur acquisition* (pp. 79-99). Neuchâtel: Delachaux et Niestlé.
- Treiman, R., Goswami, U., & Bruck, M. (1990). Not all nonwords are alike – implications for reading development and theory. *Memory and Cognition*, *18*(6), 559-567.
- Van Orden, G.C., Johnston, J.C., & Hale, B.L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *14*, 371-386.
- Wagner, R.K., & Torgesen, J.K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skill. *Psychological Bulletin*, *101*, 192-212.
- Wagner, R.K., Torgesen, J.K., & Rashotte, C.A. (1994). Development of reading-related phonological processing abilities: New evidence of bi-directional causality from a latent variable longitudinal study. *Developmental Psychology*, *30*, 73-87.
- Wimmer, H., & Goswami, U. (1994). The influence of orthographic consistency on reading development – word recognition in English and German children. *Cognition*, *51*(1), 91-103.
- Wimmer, H., Landerl, K., Linortner, R., & Hummer, P. (1991). The relationship of phonemic awareness to reading acquisition: More consequence than precondition, but still important. *Cognition*, *40*, 219-249.
- Ziegler, J.D., Jacobs, A.M., & Stone, G.O. (1996). Statistical analyses of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods Instruments and Computers*, *28*, 504-515.
- Ziegler, J.D., Perry, C., & Coltheart, M. (2000). The DRC model of visual word recognition and reading aloud: an extension to German. *European Journal of Cognitive Psychology*, *12*, 413-430.
- Ziegler, J.D., Perry, C., Jacobs, A.M., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science*, *12*, 379-384.
- Ziegler, J.C., Stone, G.O., & Jacobs, A.M. (1997). What's the pronunciation for –OUGH and the spelling for /u/? A database for computing feedforward and feedback inconsistency in English. *Behavior Research Methods, Instruments, and Computers*, *29*, 600-618.
- Zwitserslood, P. (1994). The role of semantic transparency in the processing and the representation of Dutch compounds. *Language and Cognitive Processes*, *9*(3), 341-368.

SOMMAIRE DES ANNEXES

Annexe 1 : Epreuves contrôles	241
1- Epreuves contrôles	241
1.1- LUM	241
1.2- EVIP	243
Annexe 2 : Expérience 1	245
2- Epreuves graphophonologiques	245
2.1- Epreuve graphophonologique implicite	245
2.2- Epreuve graphophonologique explicite	246
2.3- Epreuve graphophonologique de lecture	247
Annexe 3 : Expérience 2	249
3- Epreuves orthographiques	249
3.1- Epreuve orthographique implicite	249
3.2- Epreuve orthographique explicite	250
3.3- Epreuve orthographique de lecture	251
Annexe 4 : Expérience 3	254
4- Tâches d'évaluation des connaissances morphologiques	254
4.1- Epreuve morphologique implicite	254
4.2- Epreuve morphologique explicite	255
4.3- Epreuve morphologique de lecture	257

Annexe 1 : Epreuves contrôles

1- Epreuves contrôles

1.1- LUM

Cette épreuve fait partie de la batterie Lecture de Mots et Compréhension Révisée (LMC-R) de Khomsi (1999). Avec cette épreuve, il s'agit d'obtenir une première approximation concernant l'identification du mot écrit, tout en mesurant le niveau d'automatisation de cette compétence. C'est une lecture de mots, à voix haute et chronométrée, afin d'explorer la vitesse et la précision par une prise en compte des erreurs d'oralisation pour cette dernière. L'épreuve se présente sous la forme d'une liste de 105 mots à lire à voix haute en colonne (*cf.* ci-dessous). Les items sont écrits en minuscules (police Garamond, Taille 12) et de difficulté croissante (le premier mot de la liste est *il*, le dernier est *inquiétude*). Cette épreuve est chronométrée.

il	mars	naïf
un	jaune	curieux
le	parc	venir
lui	sept	départ
nu	bloc	écho
os	faim	devient
et	compter	mardi
fil	poisson	drôle
bol	herbe	histoire
tous	pied	glacé
mur	chemin	derrière
sac	joie	coloquinte
clé	phare	stylo
est	ciel	magasin
rue	coup	cueillir
petit	chorale	compagnon
camion	bruit	chronomètre
nom	monsieur	continuent
acide	agir	soleil
film	cousin	pharmacie
grand	unir	enveloppe

éléphant	travail	expliquer
trésor	yacht	magnifique
choc	animal	spécial
fier	surprise	fenouil
nouveau	doigt	gourmand
expert	sommeil	chirurgien
sportif	réunir	prévenir
merveilleux	briller	assiette
odeur	guêpe	dangereux
plongeant	feuillage	observer
parent	gymnastique	obscurité
idéal	fauteuil	descendre
obéir	obtenir	installeront
griller	jusque	inquiétude

La passation est individuelle pour l'ensemble des participants. Pour les enfants de GSM au niveau CM2, l'ensemble des items de *LUM* est proposé afin de déterminer le niveau d'expertise individuel en lecture. Il est précisé aux participants qu'ils vont devoir lire les mots par colonne, à voix haute, le plus rapidement possible et sans faire d'erreurs. Le chronomètre est déclenché dès le début de la lecture. Il est arrêté au bout d'une minute et le participant doit alors suspendre sa lecture. Le dernier item lu est souligné. Les items incorrectement oralisés sont retranscrits et les items sautés sont notés par un « S ». Deux notes sont relevées : le nombre d'items lus (IL) repérable par le numéro d'ordre du dernier item lu en une minute, ainsi que le nombre d'erreurs d'oralisation (EO). Une dernière note est calculée par soustraction des précédentes : il s'agit du nombre d'items lus correctement en une minute. Cette note sert de note globale pour l'épreuve (*LUM* sur 105). Les items sautés (« S ») ou non oralisés sont soustraits de la note IL avant de calculer la note *LUM*. La consigne est la suivante : « *Je vais te présenter une liste de mots. Tu vas devoir les lire à voix haute, le plus rapidement possible pendant une minute et sans faire de fautes. Les mots se lisent dans ce sens, de haut en bas en colonne (montrer). Tu as compris ? Tu n'as pas de questions ? Donc, dès que je donne le « top » du départ, tu dois lire le plus de mots possible, sans te tromper, jusqu'à ce que je donne le « top » de la fin* ».

1.2- EVIP

L'*EVIP* est une adaptation en langue française du *Peabody Picture Vocabulary Test - Revised (PPVT-R)*. Cette épreuve vise à évaluer l'étendue du vocabulaire acquis par un individu donné (versant réceptif). L'échelle est composée de cinq items d'entraînement suivis de 170 items rangés par ordre croissant de difficulté. Les items sont échelonnés en ordre de difficulté ascendante afin de permettre l'élimination des items trop faciles et trop difficiles pour un individu : proposer au participant des items beaucoup trop faciles ou beaucoup trop difficiles n'est pas pertinent au niveau de l'analyse des résultats et risque de surcroît de créer un effet de fatigue. Chaque item consiste en quatre dessins en noir et blanc numérotés de 1 à 4, présentés sur une feuille. Le participant doit choisir parmi les quatre images celle qui illustre le mieux la signification du mot stimulus qui est prononcé à haute voix par l'expérimentateur.

La passation est individuelle. Cinq items d'entraînement sont proposés puis débute la phase expérimentale. Les items expérimentaux sont au nombre de 170. Chaque participant ne passe que les items de l'échelle propre à son niveau selon sa classe d'âge. Il doit ainsi répondre aux 25 à 50 items de son niveau. Ces niveaux ont été établis à partir de l'étalonnage de l'échelle réalisé auprès d'un échantillon important d'enfants âgés de 2,6 ans à 18 ans et d'adultes âgés de 18 ans et plus. Cet étalonnage a permis d'effectuer une analyse des difficultés présentées par les items. La valeur des items aux différents niveaux d'âge a été comparée et ceci a alors conduit à sélectionner les items pertinents selon la classe d'âge. La durée de passation varie entre 8 et 15 minutes. La notation se fait au moment même de la passation. Cette épreuve est notée sur 170. Des consignes spécifiques s'adressent à trois catégories distinctes de participants : ceux de 7 ans ou moins, ceux de 8 ans et plus et les adultes.

La consigne pour les enfants de 7 ans ou moins est : « *Tu veux bien regarder des images avec moi ? Tu vois toutes les images sur la feuille (indiquer en pointant son doigt vers chaque image). Tu vas me montrer avec ton doigt chaque mot que je vais te dire. Par exemple, mets ton doigt sur... (premier item de la série d'entraînement) »*. Dès le début de la passation, si le participant choisit la mauvaise image, la bonne image est montrée du doigt, en disant : « *Non pas tout à fait, la bonne image, c'est celle-ci. Mais ce n'est pas grave »*. Il est expliqué en même temps mais brièvement pourquoi cette image est la bonne. L'item est alors répété jusqu'à ce qu'il soit réussi. Une fois que le participant a compris la règle qui consiste à montrer les réponses du doigt, l'évaluation commence. « *Maintenant on va regarder d'autres images. Tu montreras avec ton doigt l'image qui représente chaque mot que je te dis. Si je te dis des mots que tu ne connais pas, tu peux me le dire mais regarde quand même toutes les images et essaie. Tu as bien compris ? Tu n'as pas de questions ? »*.

La consigne pour les enfants de 8 ans et plus est la suivante : « *Tu connais beaucoup de mots ? Nous allons jouer à un jeu de vocabulaire. Attention, il y a quatre images sur cette feuille. Chaque image a un numéro (indiquer en montrant chaque image du doigt et en disant les numéros à tour de rôle : 1, 2, 3 et 4). Ecoute bien chaque mot que je te dis ; tu me diras alors le numéro de l'image qui représente ce mot, ou tu montreras cette image avec ton doigt. Voici le premier. Dis-moi le numéro de l'image, ou montre avec ton doigt l'image qui représente le mieux le mot (premier item de la série d'entraînement) »*. Si le participant indique la mauvaise image, la bonne réponse est donnée en expliquant pourquoi c'est la bonne. Une fois que le participant a bien compris, l'évaluation commence. « *Maintenant nous allons regarder d'autres images. Chaque fois que je dis un mot, toi, tu diras le numéro de son image, ou tu montreras avec ton doigt l'image qui représente le mieux ce mot. Si je te dis des mots que tu ne connais pas, tu peux me le dire mais regarde quand même toutes les images et essaie. Tu as bien compris ? Tu n'as pas de questions ?* ».

Pour les adultes, la consigne est : « *Chaque image a un numéro (indiquer en montrant chaque image du doigt et en disant les numéros à tour de rôle : 1, 2, 3 et 4). Ecoutez bien chaque mot ; vous me direz alors le numéro de l'image qui représente ce mot. Voici le premier. Dites-moi le numéro de l'image qui représente le mieux le mot (premier item de la série d'entraînement) »*. Une fois que le participant a bien compris, l'évaluation commence. « *Maintenant nous allons regarder d'autres images. Chaque fois que je vous dis un mot, vous m'indiquerez le numéro de son image. Vous avez bien compris ? Vous n'avez pas de questions ?* ».

Annexe 2 : Expérience 1

2- Epreuves graphophonologiques

La faisabilité des exercices, la compréhension des consignes ainsi que la familiarité aux items (mots) ont été testées au préalable en novembre et décembre 2003 dans des écoles maternelles et primaires différentes de celles sollicitées pour l'expérience proprement dite.

2.1- Epreuve graphophonologique implicite

Cette tâche comporte une liste de 12 paires de pseudo-mots : 8 constituent les items expérimentaux et 4 sont considérés comme jouant le rôle de distracteurs. Quatre paires de pseudo-mots sont utilisées pour l'entraînement.

Items	Paires
Entraînement	fita – gntv
	ratu – bkse
	nilo – mrzi
	tale – xlpq
Expérimentaux	thonou – chonou
	auvrou – aivrou
	piralue – piralum
	bamiran – bamirac
	givorin – givoris
	bracou – bracho
	patimme – patille
	narule – nagule
Distracteurs	votiche – vutiche
	nodape – nydape
	tapilon – zapilon
	bontra – bontro

Tableau présenté aux adultes :

1	fita	gntv
2	ratu	bkse
3	nilo	mrzi
4	tale	xlpg
5	thonou	chonou
6	auvron	aivron
7	piralue	piralum
8	bamiran	bamirac
9	givorin	givoris
10	bracou	bracho
11	patimme	patille
12	narule	nagule
13	votiche	vutiche
14	nodape	nydape
15	tapilon	zapilon
16	bontra	bontro

Les items consistants et inconsistants se différencient en termes de fréquence des bigrammes ($F(1,14) = 42,67 ; p < .0001$) et des trigrammes ($F(1,14) = 6,45 ; p < .05$).

2.2- Epreuve graphophonologique explicite

Cette tâche comporte 2 paires d'items pour l'entraînement et 10 paires pour l'expérience proprement dite. Chaque paire est constituée d'un item régulier et d'un item irrégulier dans la langue.

Items	Réguliers	Irréguliers
Entraînement	raconter	seconde
	nuit	huit
Expérimentaux	cil	fusil
	chose	chorale
	poète	poêle
	évier	hier
	grille	ville
	prix	dix
	bac	tabac
	miaou	août
	oiseau	oignon
	fil	gentil

Fréquence lexicale des items consistants et inconsistants utilisés dans l'épreuve graphophonologique explicite (Manulex)

Mots réguliers		Mots irréguliers	
cil	46	fusil	58
chose	66	chorale	44
poète	56	poêle	54
évier	47	hier	60
grille	58	ville	66
prix	61	dix	63
bac	51	tabac	50
miaou	46	août	55
oiseau	66	oignon	49
fil	61	gentil	60
Moyenne (écart-type)	55,80 (7,89)	Moyenne (écart-type)	55,90 (6,82)

Les deux catégories d'items ne se différencient pas de façon significative quant à la fréquence lexicale ($F < 1$).

2.3- Epreuve graphophonologique de lecture

Ce test de lecture comporte 16 items consistants (8 mots et 8 pseudo-mots) ainsi que 16 items inconsistants (8 mots et 8 pseudo-mots). Quatre items neutres constituent l'entraînement.

Items	Mots		Pseudo-mots	
Entraînement	faute	botte	tarola	fudin
	Consistants	Inconsistants	Consistants	Inconsistants
Expérimentaux	banane	monsieur	bertalie	baliden
	camarade	aquarium	falore	archérou
	drapeau	million	caronie	marlop
	fourmi	aiguille	vodomme	dafunix
	maladie	parfum	janitou	lautil
	poisson	respect	pourman	chotame
	salade	baptême	barulon	téjiroc
	lavabo	automne	rolane	silloude

	Nb lettres	Nb syllabes	Fréq bigrammes	Fréq trigrammes
Mots	6,94 (0,77)	2,25 (0,45)	4800,39 (3086,32)	685,12 (1479,52)
Pseudo-mots	6,94 (0,68)	2,5 (0,52)	4788,76 (3268,18)	204,27 (348,37)
Consistants	6,75 (0,68)	2,44 (0,51)	5430,56 (2771,15)	113,47 (120,36)
Inconsistants	7,13 (0,72)	2,31 (0,48)	4158,59 (3415,30)	775,91 (1478,21)

Aucune différence significative n'est obtenue entre les mots et les pseudo-mots en termes de nombre de lettres, de syllabes et de fréquence des trigrammes. Les mots et pseudo-mots diffèrent au niveau de la fréquence des bigrammes ($F(1,30) = 6,93 ; p < .05$).

Les items consistants ne se différencient pas des inconsistants pour le nombre de lettres et de syllabes. En revanche, pour les fréquences des bigrammes ($F(1,30) = 90,35 ; p < .0001$) et des trigrammes ($F(1,30) = 38,16 ; p < .0001$), les deux catégories se distinguent de façon significative.

Pour les mots, les consistants et inconsistants se différencient de façon significative pour les fréquences bigrammique et trigrammique (respectivement $F(1,14) = 170,67 ; p < .0001$ et $F(1,14) = 93,44 ; p < .0001$) mais pas pour les nombres de lettres et de syllabes. Pour les pseudo-mots, le même type de résultats est obtenu (bigrammes : $F(1,14) = 170,67 ; p < .0001$; trigrammes : $F(1,14) = 9,08 ; p < .01$).

Concernant le niveau de consistance, les mots consistants et les pseudo-mots consistants se distinguent aux fréquences des bigrammes et trigrammes (dans les deux cas : $F(1,14) = 42,67 ; p < .0001$). Pour le nombre de lettres et le nombre de syllabes, le F est inférieur à 1. Les mots inconsistants ne se différencient pas des pseudo-mots inconsistants sur l'ensemble des caractéristiques sauf pour la fréquence des bigrammes ($F(1,14) = 42,67 ; p < .0001$).

Enfin, les mots consistants et inconsistants ne diffèrent pas en termes de fréquence lexicale ($F < 1$).

Annexe 3 : Expérience 2

3- Epreuves orthographiques

3.1- Epreuve orthographique implicite

Vingt-quatre paires de pseudo-mots sont proposées : 16 paires d'items expérimentaux et 8 paires d'items distracteurs (en GSM, certaines paires ont été supprimées : elles figurent entre parenthèses). Quatre paires d'items sont proposées pour l'entraînement.

Items	Paires
Entraînement	balu – tlpz
	nufe – mfgc
	(ulit – jpgl)
	(dori – spkw)
Expérimentaux Bigramme position initiale	(camido – namido)
	cavure – navure
	caboli – naboli
	canite – nanite
	calufe – nalufe
	carubé – narubé
	catobe – natobe
	(caduli – naduli)
Expérimentaux Bigramme position finale	(romite – romive)
	latote – latove
	anoute – anouve
	corute – coruve
	pivate – pivave
	fulite – fulive
	bamate – bamave
	(donute – donuve)
Distracteurs Bigramme position initiale	(lavupé – favupé)
	latilo – fatilo
	(goliru – voliru)
	gorane – vorane
Distracteurs Bigramme position finale	(danise – danile)
	jotase – jotale
	(nilude – nilume)
	cofide – cofime

Tableau présenté aux adultes :

1	balu	tlpz
2	nufe	mfqc
3	ulit	jpgl
4	dori	spkw
5	camido	namido
6	cavure	navure
7	caboli	naboli
8	canite	nanite
9	calufe	nalufe
10	carubé	narubé
11	catobe	natobe
12	caduli	naduli
13	romite	romive
14	latote	latove
15	anoute	anouve
16	corute	coruve
17	pivate	pivave
18	fulite	fulive
19	bamate	bamate
20	donute	donuve
21	lavupé	favupé
22	latilo	fatilo
23	goliru	voliru
24	gorane	vorane
25	danise	danile
26	jotase	jotale
27	nilude	nilume
28	cofide	cofime

3.2- Epreuve orthographique explicite

Une liste de 10 paires de bigrammes est proposée : 2 paires pour les bigrammes expérimentaux et 8 paires pour les bigrammes saillants (distracteurs). Deux paires de bigrammes constituent l'entraînement à la tâche.

Bigrammes	Paires
Entraînement	ma – mu
	le – fe
Expérimental position Initiale	ca – na
Expérimental position Finale	te – ve
Saillants position Initiale	sa – za
	be – ie
	mi – my
	ac – aq
Saillants position Finale	ir – yr
	se – ze
	ce – ci
	al – am

3.3- Epreuve orthographique de lecture

Trente deux items constituent cette tâche de lecture à haute voix. Les items correspondent à 8 mots et 8 pseudo-mots comportant un bigramme (Big) fréquent ainsi que 8 mots et 8 pseudo-mots avec un bigramme rare en position initiale ou finale. Quatre items neutres font figure d'entraînement.

Items	Mots		Pseudo-mots	
Entraînement	cour	table	bilato	marode
	Fréquent	Rare	Fréquent	Rare
Expérimentaux Big Initial	camion	nageoire	caloude	narufe
	carotte	narine	cadute	nalompe
	cabane	nature	carombe	nalube
	canard	navire	catofe	navode
Expérimentaux Big Final	compote	consERVE	bofite	firanve
	cravate	sportive	possute	lutive
	tomate	lessive	falonte	toruve
	visite	salive	dufate	binouve

	Nb lettres	Nb phonèmes	Nb syllabes	Fréq bigrammes	Fréq trigrammes
Initial	6,43 (0,63)	5,06 (0,25)	2 (0)	6140,95 (2771,68)	540,96 (804,69)
Final	6,62 (0,72)	5,31 (0,70)	2,06 (0,25)	5404,83 (3426,49)	861,41 (1191,90)
Mots	6,68 (0,79)	5,37 (0,72)	2 (0)	7030,53 (3296)	1109,28 (1259,57)
Pseudo- mots	6,37 (0,50)	5 (0)	2 (0)	4515,25 (2339,71)	293,09 (425,59)
Fréquent	6,5 (0,52)	5,06 (0,25)	2 (0)	6467,47 (2198,68)	512,95 (570,20)
Rare	6,56 (0,81)	5,31 (0,70)	2,06 (0,25)	5078,31 (3721,38)	889,42 (1312,34)

Les mots ne se différencient pas de manière significative des pseudo-mots sur l'ensemble des caractéristiques du tableau ci-dessus. De même, les items à bigramme en position initiale ne se distinguent pas des items à bigramme en position finale (F inférieur à 1 sur toutes les caractéristiques hormis le nombre de syllabes où F est égal à 1 et le nombre de phonèmes où $F(1, 30) = 1,79$; *ns*). Les différences entre les items contenant un bigramme fréquent et ceux contenant un bigramme rare, quelle que soit la position, ne sont pas significatives. Les mots à bigramme en position initiale et les mots à bigramme en position finale ne se distinguent pas sur les différents facteurs de contrôle : F est inférieur à 1 pour toutes les caractéristiques sauf pour le nombre de phonèmes avec $F(1,14) = 2,07$; *ns* et pour le nombre de syllabes où F est égal à 1. Il en est de même pour les pseudo-mots (bigrammes en position initiale et en position finale) où F est inférieur à 1 pour l'ensemble des caractéristiques. Pour le niveau de fréquence du bigramme, quelle que soit la position du bigramme dans l'item, les mots à bigramme fréquent et ceux à bigramme rare ne diffèrent pas de façon significative. Les pseudo-mots à bigramme fréquent ne se distinguent pas des pseudo-mots à bigramme rare sauf pour la fréquence bigrammique ($F(1,14) = 14,05$; $p < .01$). Concernant la position du bigramme dans l'item, les mots ne diffèrent pas des pseudo-mots pour le bigramme en position initiale. Il en est de même lorsque le bigramme est en position finale. Les items à bigramme fréquent en position initiale ne se distinguent pas des items à bigramme rare. Lorsque le bigramme est en position finale, les items à bigramme fréquent ne se différencient pas non plus des items à bigramme rare.

Enfin, les mots à bigramme en position initiale et ceux à bigramme en position finale ne se différencient pas au niveau de la fréquence lexicale ($F(1,14) = 3,13$; *ns*). La différence entre les mots contenant un bigramme rare et les items incluant un bigramme fréquent n'est pas non

plus significative pour la fréquence lexicale, quelle que soit la position du bigramme ($F(1,14) = 4,17 ; ns$).

Annexe 4 : Expérience 3

4- Tâches d'évaluation des connaissances morphologiques

4.1- Epreuve morphologique implicite

Cette épreuve contient une liste de 24 paires de pseudo-mots : 16 paires pour l'expérience et 8 paires d'items distracteurs (les items entre parenthèses sont supprimés en GSM). Quatre paires d'items constituent l'entraînement.

Items	Paires
Entraînement	vosa – rfhg
	moti – sjbt
	(tipa – yvfc)
	(arlu – txnc)
Expérimentaux Préfixés	remobir – romebir
	(resibrer – risebrer)
	rerover – rorever
	relofer – rolefer
	(déloter – doléter)
	détiner – diténer
	dévoner – dovéner
	dénaler – danéler
Expérimentaux Suffixés	(birelage – lagebire)
	vaminage – nagivame
	dotilage – lagidote
	pivutage – tagipuve
	(biniteur – teuribin)
	donéveur – veurédon
	tonéreur – reuréton
	faléneur – neuréfal
Distracteurs Préfixés	(entonver – ontenver)
	enroufir – ourenfir
	(préjirer – prijéner)
	prédafer – pradéfer
Distracteurs Suffixés	(ronétoir – toiréron)
	talomoir – moirotal
	(pitesse – tessipe)
	argesse – gessare

Tableau présenté aux adultes :

1	vosa	rfhg
2	moti	sjbt
3	tipa	yvfc
4	arlu	txnc
5	remobir	romebir
6	resibrer	risebrer
7	rerover	rorever
8	relofer	rolefer
9	déloter	doléter
10	détiner	diténer
11	dévoner	dovéner
12	dénaler	danéler
13	birelage	lagebire
14	vaminage	nagivame
15	dotilage	lagidote
16	pivutage	tagipuve
17	biniteur	teuribin
18	donéveur	veurédon
19	tonéreur	reuréton
20	faléneur	neuréfal
21	entonver	ontenver
22	enroufir	ourenfir
23	préjirer	prijéner
24	prédafer	pradéfer
25	ronétoir	toiréron
26	talomoir	moirotal
27	pitesse	tessipe
28	argesse	gessare

4.2- Epreuve morphologique explicite

Cette tâche comprend une liste de 16 triplets de mots pour l'expérience. Deux triplets sont utilisés pour l'entraînement préalable.

Items	Triplets		
Entraînement Suffixé	invisible	incapable	insecte
Entraînement Préfixé	jupette	fillette	recette
Expérimentaux Préfixés	démêler	déloger	dévoré
	dévisser	décoller	déchirer
	démonter	déformer	déclarer
	délacer	dégeler	décider
	revoler	renouer	reculer
	recréer	relever	repérer
	remarier	reverser	renoncer
	regagner	redonner	refléter
Expérimentaux Suffixés	garage	triage	potage
	cirage	lavage	visage
	pliage	rasage	ménage
	bandage	séchage	fromage
	vendeur	danseur	facteur
	serveur	pêcheur	docteur
	voleur	joueur	auteur
	parleur	dormeur	pasteur

Tableau présenté aux adultes :

1	invisible	incapable	insecte
2	recette	jupette	fillette
3	démêler	déloger	dévoré
4	dévisser	décoller	déchirer
5	démonter	déformer	déclarer
6	délacer	dégeler	décider
7	revoler	renouer	reculer
8	recréer	relever	repérer
9	remarier	reverser	renoncer
10	regagner	redonner	refléter
11	garage	triage	potage
12	cirage	lavage	visage
13	pliage	rasage	ménage
14	bandage	séchage	fromage
15	vendeur	danseur	facteur
16	serveur	pêcheur	docteur
17	voleur	joueur	auteur
18	parleur	dormeur	pasteur

Les items affixés et pseudo-affixés ne se différencient pas au niveau de la fréquence lexicale ($F(1,43) = 3,55 ; ns$). Il en est de même pour les préfixés et suffixés ($F < 1$).

4.3- Epreuve morphologique de lecture

Cette épreuve est constituée d'une liste de 32 items à lire à haute voix dont 8 mots affixés, 8 mots pseudo-affixés, 8 pseudo-mots affixés et 8 pseudo-mots pseudo-affixés. L'entraînement consiste à lire 2 mots et 2 pseudo-mots neutres (sans construction morphologique spécifique).

Items	Mots		Pseudo-mots	
Entraînement	valise	fleur	pamoti	lipaber
	Affixés	Pseudo-affixés	Affixés	Pseudo-affixés
Expérimentaux Préfixés	déplier	désirer	débriser	démouder
	dépasser	déguiser	décasser	dénaller
	refermer	regretter	refaner	resaner
	reparler	renifler	reranger	rebaner
Expérimentaux Suffixés	gommage	message	mordage	boumage
	freinage	dommage	dirage	lorvage
	chasseur	douleur	gareur	tarmeur
	nageur	chaleur	cacheur	bouveur

	Nb lettres	Nb phonèmes	Nb syllabes	Fréq bigrammes	Fréq trigrammes
Préfixés	7,75 (0,58)	6,25 (0,44)	3 (0)	9568,56 (5111,40)	893,59 (887,95)
Suffixés	6,93 (0,57)	5,25 (0,45)	2,06 (0,25)	5092,07 (2445,09)	463,06 (618,11)
Mots	7,5 (0,73)	5,81 (0,83)	2,56 (0,51)	7641,80 (3855,95)	924,24 (918,99)
Pseudo-mots	7,19 (0,65)	5,69 (0,48)	2,5 (0,52)	7018,83 (5267,60)	432,41 (543,90)
Affixés	7,5 (0,75)	5,87 (0,83)	2,62 (0,52)	6937,24 (3619,29)	703,69 (958,03)
Pseudo-affixés	7,37 (0,62)	5,75 (0,68)	2,5 (0,52)	7396,37 (4420,68)	737,86 (792,02)

Les items préfixés se différencient de façon significative des items suffixés sur l'ensemble des caractéristiques du tableau ci-dessus, sauf pour la fréquence des trigrammes ($F(1,30) = 2,53$; *ns*).

Les différences entre les affixés et les pseudo-affixés ne sont pas significatives (tous les F sont inférieurs à 1).

Les mots ne se distinguent pas des pseudo-mots sur les différents contrôles effectués.

Pour les mots, les préfixés se distinguent des suffixés sur toutes les caractéristiques hormis les fréquences bigrammique et trigrammique (respectivement $F(1,14) = 3,66$; *ns* et $F <$

1). Dans le cas des pseudo-mots, les items préfixés diffèrent également des suffixés sauf pour la fréquence des trigrammes ($F(1,14) = 2,79$; *ns*) et le nombre de syllabes ($F < 1$).

Pour l'affixation, aucune différence significative n'est obtenue entre les mots affixés et les mots pseudo-affixés sur l'ensemble des facteurs contrôlés. Il en est de même pour les pseudo-mots affixés et les pseudo-mots pseudo-affixés.

Les items préfixés (mots et pseudo-mots), quel que soit leur niveau d'affixation, ne se différencient pas entre eux aux caractéristiques du tableau. Les items suffixés (mots et pseudo-mots), quel que soit leur niveau d'affixation, ne se différencient pas non plus entre pour l'ensemble des facteurs contrôlés.

Enfin, les mots préfixés et suffixés d'une part, et les mots affixés et pseudo-affixés d'autre part ne se différencient pas en termes de fréquence lexicale (respectivement $F(1,14) = 2,61$; *ns* et $F < 1$).

RESUME

Cette thèse étudie l'apparition des connaissances graphophonologiques, orthographiques, morphologiques et l'utilisation de ces connaissances en lecture. Les résultats indiquent que les connaissances explicites sur la consistance graphophonologique sont tardives, elles sont précédées par la manifestation d'une sensibilité graphophonologique. Dès le CP, cette consistance a une influence sur la vitesse et sur l'exactitude de la lecture des mots réguliers. Cet effet se retrouve chez le lecteur expert essentiellement pour les pseudo-mots. Dès le CP se manifeste également une sensibilité à la fréquence des bigrammes. En revanche, les enfants ne sont pas conscients de la fréquence relative des bigrammes isolés. La structure orthographique n'a pas d'effet significatif sur le traitement des mots et des pseudo-mots : les items contenant un bigramme fréquent en position initiale ou finale ne sont pas mieux lus que ceux avec un bigramme rare. Chez le lecteur expert, la présence d'un bigramme initial fréquent est facilitatrice. Enfin, une sensibilité à la structure morphologique écrite existe dès le CE1 et précède le développement de connaissances explicites. Dès le CP, la préfixation joue un rôle en lecture : les items préfixés sont mieux lus que les pseudo-préfixés. En revanche, la structure morphologique des suffixés ne facilite pas la lecture des items. Le caractère plurimorphémique des items n'influence pas non plus leur traitement chez l'expert. Ces résultats permettent de discuter l'impact précoce des connaissances orthographiques et morphologiques en lecture.

MOTS CLES : apprentissage de la lecture, développement, régularités graphophonologiques, orthographiques et morphologiques, connaissances implicites et explicites.

TITLE: Graphophonological, orthographic and morphological regularities: implicit learning and early impact on reading

ABSTRACT

This thesis studies the appearance of graphophonological, orthographic, morphological knowledge and the use of this knowledge in reading. The results show that explicit knowledge of graphophonological consistency comes later, and is preceded by a demonstration of graphophonological sensitivity. From the first grade, this consistency has an influence on the speed and on the precision of reading regular words. This effect is found in the expert reader mainly for pseudo-words. From the first grade there is also a sensitivity to the frequency of bigrams. On the other hand, children are not aware of the relative frequency of isolated bigrams. Spelling structure does not have a significant effect on the treatment of words and pseudo-words: items that contain a frequent bigram in an initial or final position are not read better than those with a rare bigram. In the expert reader, the presence of an initial frequent bigram is helpful. Finally, sensitivity to the written morphological structure is present from the second grade and precedes the development of explicit knowledge. From the first grade, prefixes play a role in reading: the prefixed items are read better than pseudo-prefixes. On the other hand, the morphological structure of suffixes does not help the reading of the items. When items are polymorphemic this does not influence their treatment by the expert. These results will lead to a discussion of the early impact of orthographic and morphological knowledge in reading.

KEY WORDS: learning to read, development, graphophonological, orthographic and morphological regularities, implicit and explicit knowledge.

DISCIPLINE : Psychologie

Université de Haute-Bretagne – Rennes 2

CRPCC – EA 1285

Laboratoire de Psychologie du Développement

Place du Recteur Henri Le Moal – CS 24 307

35043 Rennes Cedex